LATTICE POINT COUNTING FOR FINITE-ORDER MAPPING CLASSES

SPENCER DOWDALL AND HOWARD MASUR

We dedicate this paper to the memory of Maryam Mirzakhani.

Abstract. This paper concerns the lattice counting problem for the mapping class group of a surface S acting on Teichmüller space with the Teichmüller metric. In that problem the goal is to count the number of mapping classes that send a given point x into the ball of radius R centered about another point y. For the action of the entire group, Athreya, Bufetov, Eskin and Mirzakhani have shown this quantity is asymptotic to e^{hR}, where h is the dimension of the Teichmüller space. We instead consider only the action of finite-order elements of the group and show the associated count grows coarsely at the rate of $e^{hR/2}$, that is, with half the exponent. To obtain these quantitative estimates, we introduce a new notion in Teichmüller geometry, called complexity length, which reflects some aspects of the negative curvature of curve complexes and also has applications to counting problems.

Contents

1. Introduction

1.1. **Lattice point counting.** The goal of this paper is to count elements of the mapping class group via its action on Teichmüller space. When a group G acts on a metric space X by isometries, counting the number of orbit or "lattice" points in metric balls of increasing radius gives a measure of growth in the group as reflected

Date: May 5, 2023.

in the geometry of X. For example, the number of lattice points $\mathbb{Z}^n$ in a large metric ball in Euclidean space $\mathbb{R}^n$ is approximately the volume of the ball. Relatedly, in his Ph.D. dissertation, Margulis [Mar] considered the case of a compact negatively curved Riemannian manifold M and showed that for the isometric action of the fundamental group $\pi_1(M)$ on the universal cover $\widetilde{M}$, the number of lattice points in a ball of radius R is asymptotic to a constant times e^{hR}, where $h > 0$ is the topological entropy of the geodesic flow.

This paper concerns a refinement of this classical lattice point counting problem in the setting of Teichmüller geometry. Fix a connected, orientable surface S of genus g with p punctures such that the $\xi(S) = 3g - 3 + p$, termed its *complexity*, is positive. We consider the mapping class group $\mathrm{Mod}(S) = \mathrm{Homeo}^+(S)/\mathrm{Homeo}_0(S)$ of isotopy classes of orientation-preserving homeomorphisms of S, which acts isometrically and properly discontinuously on the Teichmüller space $\mathcal{T}(S)$ of marked hyperbolic metrics on S equipped with the Teichmüller metric $d_{\mathcal{T}(S)}$.

The geometric and dynamical theory of Teichmüller space bears striking parallels to that of negatively curved Riemannian manifolds. Motivated by this, Athreya, Bufetov, Eskin, and Mirzakhani [ABEM] drew on ideas from Margulis [Mar] to solve the analogous lattice point counting problem. For $x, y \in \mathcal{T}(S)$, let us write s_x for the finite cardinality of the stabilizer of x in $\mathrm{Mod}(S)$ and

$$\Lambda(x, y, R) = \left\{\phi \in \mathrm{Mod}(S) \mid d_{\mathcal{T}(S)}(\phi(x), y) \leqslant R\right\}$$

for the set of mapping classes that translate x to within distance R of y. The cardinality $|\Lambda(x, y, R)|$ then equals s_x times the number of orbit points $\mathrm{Mod}(S) \cdot x$ in the ball of radius R centered at y. Their result may then be stated as:

Theorem 1.1 (Athreya–Bufetov–Eskin–Mirzakhani [ABEM]). *There is a constant $\lambda > 0$ such that for all $x, y \in \mathcal{T}(S)$ one has $|\Lambda(x, y, R)| \sim \lambda s_x e^{h_S R}$, where $h_S = 2\xi(S) = 6g - 6 + 2p$ is the entropy of the Teichmüller geodesic flow, and the notation $f(R) \sim g(R)$ means that $f(R)/g(R) \to 1$ as $R \to \infty$.*

While this completely answers the lattice point counting problem for the full mapping class group, one might further refine it by considering the growth of certain naturally distinguished subgroups or subsets. This is related to the question of determining what a "typical" element of $\mathrm{Mod}(S)$ looks like.

The famous Nielsen–Thurston classification [Thu] states that every element of the mapping class group is either finite-order, reducible, or pseudo-Anosov (see Definition 3.6). Accordingly, we let

$$\Lambda_{\mathrm{fo}}(x, y, R), \qquad \Lambda_{\mathrm{red}}(x, y, R), \qquad \text{and} \qquad \Lambda_{\mathrm{pA}}(x, y, R)$$

denote the subsets of $\Lambda(x, y, R)$ consisting of finite-order, reducible, and pseudo-Anosov elements, respectively. Building on [ABEM], Maher [Mah1, Mah2] has used ideas from random walks to show that typical mapping classes are pseudo-Anosov in the sense that $|\Lambda_{\mathrm{pA}}(x, y, R)| \sim |\Lambda(x, y, R)|$. In particular, this shows that the proportion of finite-order and reducible lattice points tends to zero as $R \to \infty$, but it does not give any indication of the rate of convergence.

The purpose of this paper is to give quantitative estimates for the number of finite order mapping classes by counting their lattice points. We show:

Theorem 1.2. *For any $\delta > 0$ and pair of points $x, y \in \mathcal{T}(S)$, there are constants K_1, K_2, R_0 such that for all $R \geqslant R_0$ one has*

$$K_1 e^{\frac{h_S}{2}R} \leqslant |\Lambda_{\mathrm{fo}}(x, y, R)| \leqslant K_2 e^{(\frac{h_S}{2}+\delta)R}.$$

We remark that the exponents for the upper and lower bound do not quite coincide because of the δ in the exponent for the upper bound; we do not know if the δ can be removed. Nonetheless our main theorem shows that the growth rate of finite order elements is exponential with exponent essentially one half that of the entire group. In future work, we will use the techniques developed in this paper to additionally show the number $|\Lambda_{\mathrm{red}}(x, y, R)|$ of reducible elements grows coarsely at the rate of $e^{(h_S-1)R}$, with exponent one less than for the pseudo-Anosovs.

1.2. **Heuristics and hazards.** Let us describe a naive picture illustrating why one might expect the finite-order elements to grow at half the exponential rate of the whole group. The first observation is that there are only finitely many conjugacy classes of finite-order elements; thus it suffices to count each conjugacy class $[\phi_0]$ separately. Since the points x, y may be adjusted at the cost of increasing the constants K_1, K_2, we might as well assume $x = y$ is a fixed point x_0 for a given finite-order element ϕ_0. Now, the result of [ABEM] (Theorem 1.1) says there are approximately $e^{h_S R/2}$ mapping classes $f \in \mathrm{Mod}(S)$ so that $d_{\mathcal{T}(S)}(x_0, f(x_0)) \leqslant R/2$. Further, each of these produces a conjugate $\phi_f = f\phi_0 f^{-1}$ for which the translate $f(x_0)$ is a fixed point. The triangle inequality thus implies this finite-order element satisfies $d_{\mathcal{T}(S)}(x_0, \phi_f(x_0)) \leqslant R$.

This observation suffices for the lower bound in Theorem 1.2, provided the assignment $f \mapsto \phi_f$ is uniformly finite-to-one, as is the case when ϕ_0 has finite centralizer. This argument is carried out in detail in §13.1, where we prove the lower bound by constructing explicit examples.

For the upper bound, a hope might be that *all* (or at least most) elements $\phi \in [\phi_0]$ satisfying $d_{\mathcal{T}(S)}(x_0, \phi(x_0)) \leqslant R$ arise in this manner as $f\phi_0 f^{-1}$ for some element f with $d_{\mathcal{T}(S)}(x_0, f(x_0)) \leqslant R/2$. While this is, of course, too naive, the thrust of our argument is that the hope does hold in some moral sense, albeit in a rather complicated way involving an alternative understanding of distance.

The given element $\phi \in [\phi_0]$ can be expressed as a conjugate $\phi = f\phi_0 f^{-1}$ in possibly many ways, and choosing a conjugator f roughly corresponds to identifying a fixed point $f(x_0)$ of ϕ. The hope thus translates into finding a fixed point x_ϕ with $d_{\mathcal{T}(S)}(x_0, x_\phi) \leqslant R/2$. This, however, need not be possible: In §2 we provide an example of a finite-order ϕ_0 such that for every conjugate $\phi \in [\phi_0]$ with $d_{\mathcal{T}(S)}(x_0, \phi(x_0)) \leqslant R$, the closest fixed point x_ϕ satisfies the dual properties that: 1) up to additive error we have $d_{\mathcal{T}(S)}(x_0, x_\phi) \geqslant R$, and 2) the geodesic from x_0 to x_ϕ passes through the Teichmüller space $\mathcal{T}(V)$ of some subsurface V in such a way that the geodesic from x_ϕ on to $\phi(x_0)$ "backtracks" through the same Teichmüller space $\mathcal{T}(V)$, undoing the progress made in going from x_0 to x_ϕ.

The reasons why such backtracking is problematic are perhaps too technical to elaborate upon in this introduction. Suffice it to say that the theory of subsurface projections developed in [MM1, Raf1] shows that Teichmüller geodesics are governed by how they move through "thin regions" where the boundary curves ∂V of subsurfaces V become short. These regions behave like metric products, in which the Teichmüller space $\mathcal{T}(V)$ of the surface is one factor [Min], and are the main source of non-negative curvature and many headaches in $\mathcal{T}(S)$.

These two issues—that the closest fixed point may be too far away and that the piecewise geodesic path from x_0 to the fixed point and on to $\phi(x_0)$ may backtrack in subsurfaces—are the main obstacles. Our proof of the upper bound roughly divides into two separate parts overcoming these issues. The first part (§5–6) constructs good fixed points x_ϕ that minimize backtracking and branch points a_ϕ, b_ϕ that help mitigate it. The second, and much more elaborate, part (§7–12) ultimately shows, provided backtracking is controlled, that while the Teichmüller distance $d_{\mathcal{T}(S)}(x_0, x_\phi)$ may be much larger than $R/2$, there is a more apt measure of length that *is* on the order of $R/2$. Developing the theory of this length is major component of the paper, which introduces new ideas and techniques to Teichmüller theory that we hope may be of independent interest and lead to other applications.

1.3. **A new complexity length for Teichmüller space.** The impetus for our construction is the need to count points in a way that incorporates how geodesics move through Teichmüller spaces of subsurfaces.

For the purposes of counting, it is helpful to discretize $\mathcal{T}(S)$ by considering a *net* $\mathcal{N}(S)$; this is a c–separated subset whose $2\mathsf{c}$ balls cover $\mathcal{T}(S)$, for some constant c (see §3.10). Eskin and Mirzakhani introduced nets in [EM] and showed there is a uniform constant C_0 such that a ball or radius R about any thick point $x \in \mathcal{T}(S)$ contains at most $C_0 e^{h_S R}$ net points. When the thickness condition is removed and arbitrary centers are considered, they show that for any $\delta > 0$ there is some C_δ such that all balls of radius R contain at most $C_\delta e^{(h_S+\delta)R}$ net points. This is one explanation of where the δ comes from in our main theorem.

The key observation is as follows: If one fixes a thick center point x and moves distance R to a net point y by *only* moving in the Teichmüller space $\mathcal{T}(V)$ of a subsurface and not moving in the complement, then Minsky's product regions theorem [Min] says this behaves like a Teichmüller geodesic in $\mathcal{T}(V)$ and hence [EM] implies there should only be $C_0 e^{h_V R}$ such net points y. That is, imposing a restriction that the geodesic passes through Teichmüller spaces of subsurfaces cuts down on the number of net points that can be reached in distance R.

Our *complexity length* $\mathfrak{L}(x, y)$ (Definition 11.1) is designed to implement this observation in a rigorous way that accounts for the fact that geodesics can move through disjoint subsurfaces simultaneously. Although the construction is complicated, the rough idea is to take all the subsurfaces Z for which the curve complex projection (§3.11) $d_Z(x, y)$ is large, determined by a parameter C, and partition them by picking out a distinguished subfamily Ω, called a *witness family* (see §7), with the property that each such Z is minimally contained in a unique element V of Ω. This family comes with additional combinatorial structure (a *subordering*; Definition 7.14) that allows one to take the curve complex data from all these subsurfaces Z contributing to V and reassemble it, via the concept of *consistency* from [BKMM], into a pair of points $\widehat{x}_V^\Omega, \widehat{y}_V^\Omega \in \mathcal{T}(V)$ in the Teichmüller space of V with the property that $d_Z(\widehat{x}_V^\Omega, \widehat{y}_V^\Omega)$ and $d_Z(x, y)$ coarsely agree (up to only additive error) for each Z. The complexity length is then defined as

$$\mathfrak{L}(x, y) = \sum_{V \in \Omega} h_V d_{\mathcal{T}(V)}(\widehat{x}_V^\Omega, \widehat{y}_V^\Omega).$$

Since the distances are weighted by the exponents h_V used for counting, and since the pairs $\widehat{x}_V^\Omega, \widehat{y}_V^\Omega$ encode the data of the original points, we are able to prove the following analog of Eskin and Mirzakhani's [EM] net point counting result:

Theorem 1.3 (c.f. Theorem 12.1)**.** *For any sufficiently large parameter C, there exists $k \in \mathbb{N}$ such that each $x \in \mathcal{T}(S)$ has at most $kr^k e^r$ net points y within complexity length $r > 0$. That is: $\#\{y \in \mathcal{N}(\Sigma) \mid \mathfrak{L}(x,y) \leqslant r\} \leqslant kr^k e^r$.*

As explained in Remark 8.9, our complexity length is not dissimilar to Rafi's [Raf1] Distance Formula (Theorem 3.33), which roughly says the Teichmüller distance $d_{\mathcal{T}(S)}(x,y)$ is comparable, with multiplicative and additive error, to the sum of all large curve complex projections $d_Z(x,y)$. In fact, one finds that $\mathfrak{L}(x,y)$ and $d_{\mathcal{T}(S)}(x,y)$ also agree up to bounded multiplicative and additive error, since they both coarsely agree with the sum in the distance formula! However, there are two key differences between these perspectives:

The first is that the distance formula concerns *all* subsurfaces with large projection. The sum may therefore have arbitrarily many terms, and this ultimately contributes to a multiplicative error. But we cannot afford multiplicative error, since the distance R appears in the exponent in our main theorem and the whole point is to calculate the exponent. Throughout the construction we must therefore be careful to utilize witness families Ω of uniformly bounded cardinality, so that our sum has boundedly many terms and the various additive errors do not accumulate into a multiplicative error. By arranging things with great care, we are able to relate complexity length to Teichmüller distance with *only additive error*.

The second is that we sum over Teichmüller, rather than curve complex, distances, which facilitates the above application to counting. Nevertheless, we are are still able to tap into the hyperbolic geometry of curve complexes in the following sense. Let us say a triple (a,b,c) in $\mathcal{T}(S)$ is *θ–aligned* if for every subsurface Z the three pairwise curve complex projections satisfy the reverse triangle inequality:

$$d_Z(a,b) + d_Z(b,c) \leqslant d_Z(a,c) + \theta.$$

Since the curve complex $\mathcal{C}(Z)$ is hyperbolic, this is equivalent to saying the projection of b to $\mathcal{C}(Z)$ lies near the geodesic joining the projections of a and c. Because of the multiplicative error and arbitrary length of the sum, this inequality in each term of the distance formula does not translate into a reverse triangle inequality for Teichmüller distance: There are θ–aligned triples (a,b,c) for which $d_{\mathcal{T}(S)}(a,b) + d_{\mathcal{T}(S)}(b,c) - d_{\mathcal{T}(S)}(a,c)$ is arbitrarily large. Complexity length, however, does satisfy such a reverse triangle inequality. This requires the triple (a,b,c), or more generally tuple $(x_0,\dots,x_n)$, to be *strongly* θ–aligned (Definition 3.21), which adds a condition on the lengths of curves at b so that its projection to the Teichmüller space $\mathcal{T}(A)$ of each annulus A lies near the geodesic joining the projections of a and b. We then have the following key result:

Theorem 1.4 (c.f. Theorem 11.2)**.** *For any $n \geqslant 1$ and sufficiently large parameter C, there exists K such every strongly C–aligned n-tuple $(x_0,\dots,x_n)$ satisfies*

$$\mathfrak{L}(x_0,x_1) + \cdots + \mathfrak{L}(x_{n-1},x_n) \leqslant \left(h_S + \frac{n}{\mathsf{C}}\right) d_{\mathcal{T}(S)}(x_0,x_n) + K.$$

In particular $\mathfrak{L}(x,y) \leqslant K + (h_S + \frac{1}{\mathsf{C}})d_{\mathcal{T}(S)}(x,y)$ for any $x,y \in \mathcal{T}(S)$.

1.4. **Summary of proof.** Theorems 1.3 and 1.4 are the key features that enable complexity length to overcome the first main issue described in §1.2, namely of the closest fixed point being too far away. Roughly, the argument is as follows: If we were able to find a fixed point x_ϕ for $\phi \in [\phi_0]$ so that the triple $(x_0, x_\phi, \phi(x_0))$ were strongly aligned, then Theorem 1.4 would imply

$$\mathfrak{L}(x_0, x_\phi) + \mathfrak{L}(x_\phi, \phi(x_0)) \leqslant (h_S + \tfrac{2}{\mathsf{C}})d_{\mathcal{T}(S)}(x_0, \phi(x_0)) + K \leqslant (h_S + \delta)R + K,$$

provided C is chosen sufficiently large. By symmetry, the two terms on the left are equal and thus each at most $(h_S + \delta)R/2 + K$. Theorem 1.3 thus implies that for large R there are at most $e^{(h_S+2\delta)R/2}$ such fixed points x_ϕ. Since the multiplicity of the assignment $\phi \mapsto x_\phi$ is bounded by the uniform finiteness of the stabilizer of x_ϕ, this gives the desired upper bound on $|\Lambda_{\mathrm{fo}}(x_0, x_0, R)|$.

Complexity length thus enables the heuristic argument of §1.2 to work regardless of the distance to the closest fixed point, provided $(x_0, x_\phi, \phi(x_0))$ is strongly aligned. While this last condition need not hold, we circumvent it by utilizing a sort of barycenter for the triple $(x_0, x_\phi, \phi(x_0))$. In fact, for subtle technical considerations we construct a pair a_ϕ, b_ϕ of points so that $(x_0, a_\phi, b_\phi, \phi(x_0))$ is strongly aligned. We then count the number of such pairs a_ϕ, b_ϕ by the argument above, and carry out a reconstruction argument (Theorem 6.1) to show any pair arises as a_ϕ, b_ϕ for at most polynomially (in R) many elements $\phi \in [\phi_0]$ satisfying $d_{\mathcal{T}(S)}(x_0, \phi(x_0)) \leqslant R$. Together, these ingredients yield the upper bound on $|\Lambda_{\mathrm{fo}}(x_0, x_0, R)|$.

1.5. **Questions.** There are several natural questions prompted by this work. The most obvious is whether the δ in the upper bound of Theorem 1.2 can be removed. It arises from various technical considerations that manifest in the additive $\frac{n}{\mathsf{C}}$ term in Theorem 1.4. This is the result of a phenomenon that we term "badness" (§10) having to do with the fact that the witness family Ω may have pairs of nested subsurfaces $V_1 \sqsubset V_2$ with $h_{V_1} + h_{V_2} > h_S$ for which the distances $d_{\mathcal{T}(V_i)}(\widehat{x}^{\Omega}_{V_i}, \widehat{y}^{\Omega}_{V_i})$ appearing in complexity length correspond to a region during which the main Teichmüller geodesic $[x_0, x_n]$ is simultaneously moving through the Teichmüller spaces $\mathcal{T}(V_1)$ and $\mathcal{T}(V_2)$. Our construction endeavors to minimizes badness (§10.1), but it would be nice to find a solution eliminating it entirely. Even if the $\frac{n}{\mathsf{C}}$ term from Theorem 1.4 could be removed, there are still two polynomial factors, coming from Theorems 1.3 and 6.1, that enlarge the upper bound but are currently absorbed into the $e^{\delta R}$ factor in the statement of Theorem 1.2.

A related question is whether complexity length itself satisfies a reverse triangle inequality $\mathfrak{L}(a, b) + \mathfrak{L}(b, c) \leqslant \mathfrak{L}(a, c) + K$ for strongly aligned triples, rather than the hybrid formulation concerning both $\mathfrak{L}$ and $d_{\mathcal{T}(S)}$ in Theorem 1.4. While this may likely be the case, proving such a statement appears to be quite difficult.

A third question concerns the fact that our main theorem only provides coarse bounds with multiplicative error, rather than precise asymptotics as in Theorem 1.1). One reason Theorem 1.2 is so difficult is because it concerns intrinsically nongeneric phenomenon. In contrast to [ABEM, Mah1], where dynamics and ergodic theory are the main tools, our lattice points arise with vanishingly small probability that is undetectable by these tools. We must instead rely on coarse geometric arguments that lead to coarser bounds. Calculating the precise asymptotic growth of $|\Lambda_{\mathrm{fo}}(x, y, R)|$ will require completely different techniques.

Finally, as indicated in §1.2, we prove Theorem 1.2 by counting each conjugacy class $[\phi_0]$ of finite order elements separately, and one might wonder whether other conjugacy classes exhibit similar behavior. That is, for any element $\psi \in \mathrm{Mod}(S)$ and points $x, y \in \mathcal{T}(S)$ one may consider the growth of the set

$$\Lambda_\psi(x, y, R) = \left\{\psi' \in [\psi] \mid d_{\mathcal{T}(S)}(\psi'(x), y) \leqslant R\right\}.$$

Recent work of Han [Han1, Han2] shows that for every Dehn twist ψ, as well as most multitwists and most pseudo-Anosovs, the quantity $|\Lambda_\psi(x, y, R)|$ grows coarsely like $e^{h_S R/2}$. In light of this and Theorem 1.2, we propose the following:

Question 1.5. For any nontrivial element $\psi \in \mathrm{Mod}(S)$ and points $x, y \in \mathcal{T}(S)$, do there exist $K_1, K_2 > 0$ such that $K_1 e^{h_S R/2} \leqslant |\Lambda_\psi(x, y, R)| \leqslant K_2 e^{h_S R/2}$ for all large R? Furthermore, does $|\Lambda_\psi(x, y, R)| / e^{h_S R/2}$ converge as $R \to \infty$ and, if so, to what?

Just as Theorem 1.1 parallels Margulis's result [Mar] that the fundamental group $\pi_1(M)$ of a compact negatively curved manifold grows at the rate of e^{hR}, a positive answer to Question 1.5 would parallel work of Parkkonen–Paulin [PP] showing that each nontrivial conjugacy class in $\pi_1(M)$ essentially grows at the rate of $e^{hR/2}$.

1.6. **Outline.** The paper is organized as follows. In §2 we provide an example of a finite-order element ϕ_0 whose conjugates all have closest fixed points that are both far away from x_0 and exhibit backtracking. In §3 we collect the needed background material, surveying many of the ideas in the subject of the mapping class group, Teichmüller geometry, and the curve graph of a surface. Section 4 proves preliminary technical results that are needed in the sequel. This includes bounds on antichains that are used repeatedly in the construction of complexity length to control the size of witness families; an explanation of how alignment can be promoted to strong alignment, and an application of Gromov hyperbolicity of curve complexes to construct branch points with various properties.

The proof of the upper bound begins in §5, where we construct a good fixed point x_ϕ for each conjugate $\phi \in [\phi_0]$ along with a pair of branch points a_ϕ, b_ϕ for which the tuple $(x_0, a_\phi, b_\phi, \phi(x_0))$ is strongly aligned. The key properties enjoyed by these points are collected in Proposition 5.5. Next, in §6, for each possible pair a, b, we count the number of conjugates ϕ with $d_{\mathcal{T}(S)}(x_0, \phi(x_0)) \leqslant R$ whose branch points a_ϕ, b_ϕ are a, b. This count turns out to be polynomial in R; this is Theorem 6.1, the main result of the first part of the paper.

The second part of the paper, spanning §§7–12, is devoted to developing the theory of complexity length. In §7 we introduce the notion of *witness families* Ω with additional structures (wideness, insulation, subordering, and completeness) that will be needed in our constructions. Section 8 then defines the complexity $\mathfrak{L}(\Omega)$ of a witness family, associated to a strongly aligned tuple $(x_0, \ldots, x_n)$, by constructing *resolution points* $\widehat{x}_{iV}^{\Omega}$ in the Teichmüller spaces $\mathcal{T}(V)$ of each subsurface $V \in \Omega$. Section 9 then bounds each summand $h_V d_{\mathcal{T}(V)}(\widehat{x_{i-1}}_V^{\Omega}, \widehat{x}_{iV}^{\Omega})$ of complexity length in terms of the Lebesgue measure of a certain *contribution set* $\mathcal{A}_V^{\Omega}$ along the main Teichmüller geodesic $[x_0, x_n]$. This is perhaps the most intricate part of the argument and is accomplished by judicious use of Minsky's product regions (Theorem 3.11). Section 10 then introduces the notion of badness, which can lead to complexity length being strictly larger than $h_S d_{\mathcal{T}(S)}(x_0, x_n)$, and explains an iterative procedure for refining witness families and minimizing badness. It is here where we specify the various parameters of the definition and construct (in Proposition 10.13) witness families that simultaneously have controlled badness and uniformly bounded cardinality. In §11 we then define the complexity length $\mathfrak{L}(x_0, \ldots, x_n)$ of a strongly aligned tuple as an infimum of complexities $\mathfrak{L}(\Omega)$ of witness families satisfying certain properties. This culminates in Theorem 11.2 (c.f. 1.4 above) proving the key properties that complexity length satisfies a reverse triangle inequality and is bounded in terms if Teichmüller distance. Finally, in §12, we come back to the application of counting and bound, in Theorem 12.1 (c.f. 1.3 above) the number of net points within a given complexity length of x.

The proof of the main Theorem 1.2 is finally given in §13. The lower bound is handled by constructing an explicit example. At this point, with all the tools ready, the argument for the upper bound is not too difficult and roughly follows the sketch in §1.4 above. However, there are additional complications involving thin annuli and a resulting *savings* in complexity length that is needed to avoid overcounting the number of pairs a_ϕ, b_ϕ of branch points.

1.7. **Acknowledgments.** This paper owes an intellectual debt to Maryam Mirzakhani, who posed the question of counting finite-order elements back in 2006. We dedicate this paper to her memory in hopes that she would have found our answer interesting. We also thank Alex Eskin for numerous helpful conversations and for answering many questions related to the work in [ABEM, EM]. The authors would also like to thank Kasra Rafi for many helpful conversations. We gratefully acknowledge the support of the Simons Laufer Mathematical Sciences Institute, formerly Mathematical Sciences Research Institute, and its generous hospitality during the 2016 program on Geometric group theory, when the framework of complexity length was formalized, and during the 2022 Complementary Program, when the final components of the project were completed. The first named author was partially supported by NSF grants DMS-1711089 and DMS-2005368.

2. Example

We begin with the example promised in the introduction of a finite order mapping class ϕ_0 with fixed point at some x_0, a conjugate $\phi = w \circ \phi_0 \circ w^{-1}$ with closest fixed point $w(x_0)$ such that

- up to additive error $d_{\mathcal{T}(S)}(x_0, \phi(x_0)) = d_{\mathcal{T}(S)}(x_0, w(x_0))$.
- there is backtracking in the Teichmüller space of a subsurface

Let S be genus 7 surface cut into 7 disjoint 1-holed tori $\Sigma_1, \ldots, \Sigma_7$ and a 7 holed sphere Z. The map ϕ_0 is of order 7, rotating the tori sending Σ_i to Σ_{i+1} (Σ_7 to Σ_1), and fixes Z. Via the rotation we may identify points in $\mathcal{T}(\Sigma_i)$ and $\mathcal{T}(\Sigma_{i+1})$. These are copies of the hyperbolic plane $\mathbb{H}^2$ with the hyperbolic metric. Let f be a Anosov map on Σ_1, Σ_3, Σ_5 and Σ_7. For positive integers $j \leqslant n \leqslant m$ set $w = f^{(n)}$ on Σ_1, $w = f^{(m)}$ on Σ_3, $w = f^{(-m)}$ on Σ_5 and $w = f^{(j)}$ on Σ_7. Set w to be the identity on $\Sigma_2 \cup \Sigma_4 \cup \Sigma_6 \cup Z$. Then let $\phi = w \circ \phi_0 \circ w^{-1}$. With the identification of each $\mathcal{T}(\Sigma_i)$ with the hyperbolic plane $\mathbb{H}^2$ we assume f is of the form $z \to \lambda z$ on the copies of $\mathbb{H}^2$ corresponding to $\mathcal{T}(\Sigma_1), \mathcal{T}(\Sigma_3), \mathcal{T}(\Sigma_5), \mathcal{T}(\Sigma_7)$. Thus the imaginary axis is fixed by w. We may also choose the fixed point x_0 so that the lengths of the curves $\alpha_i = \partial\Sigma_i$ are moderate and i is the corresponding point in $\mathbb{H}^2$. Let μ_i the corresponding markings in Σ_i.

By the Minsky product formula [Min] it is easy to check that up to additive constants independent of j, n, m

$$d_{\mathcal{T}(S)}(x_0, w(x_0)) = d_{\mathcal{T}(S)}(x_0, \phi(x_0)) = m \log \lambda.$$

We now sketch the argument that any other fixed point u of ϕ, up to additive constant, satisfies $d_{\mathcal{T}(S)}(u, x_0) \geqslant m \log \lambda$. Let ν_i the Bers marking at u projected to Σ_i. Since $\phi(u) = u$ we have $\nu_4 = f^{(-m)}(\nu_3)$, $\nu_5 = f^{(-m)}(\nu_4)$, where again the rotation allows us to identify markings in different Σ_i. Thus $\nu_3 = f^{(2m)}(\nu_5)$.

This implies that the projection of u to the curve complex on Σ_3 and Σ_5 differ by $2m\log\lambda$ and so by the triangle inequality at least one of the two projections differs from the projection of x_0 by at least $\log\lambda$. Without loss of generality assume it is Σ_3. There must be interval an interval I (called an active interval in this paper) along $[u, x_0]$ where the boundary loop α_3 has length at most ϵ_0 and outside I the projections to the curve complex of Σ_3 only changes by an additive constant. This means that along I, we can consider all the points to lie in $\mathbb{H}^2 \times \mathcal{T}(S \backslash \Sigma_3)$. Again by [Min] we have up to an additive constant $|I| \geqslant m\log\lambda$. Therefore up to an additive constant the distance between u and x_0 must be at least $m\log\lambda$ and so $w(x_0)$ is the closest fixed point up to additive error.

Next we note that in the curve complex of Σ_1, that in going from x_0 to $w(x_0)$ we travel distance $n\log\lambda$ and then from $w(x_0)$ to $\phi(x_0)$ we backtrack distance $j\log\lambda$. We remark that the fixed point $w(x_0)$ will be called a good fixed point since there is not backtracking in *some* subsurface. In general we will find a good fixed point for any finite order ϕ. A second observation is that if we let y be a point with the same markings as $w(x_0)$ except in Σ_1 where we set the marking to agree with that of $\phi(x_0)$, then the three points $x_0, y, \phi(x_0)$ are aligned as moving from x_0 to y and then to $\phi(x_0)$ there is no backtracking in any domain. Another major goal will be to produce such points in general which we call good branch points. We will then do two major counts. In the first, given a good branch point, we will count the number of maps that determine the same branch point. The second count will be to determine the number of possible branch points. This is where we introduce complexity functions.

3. Background

Throughout, the term *surface* will indicate an oriented surface Σ homeomorphic to a closed surface minus a (possibly empty) finite set of points. The missing points are called *punctures* and are in bijective correspondence with the ends of Σ. We write $S_{g,p}$ for the connected genus $g \geqslant 0$ surface with $p \geqslant 0$ punctures, and define its *complexity* to be $\xi(S_{g,p}) := 3g - 3 + p$. In general, the complexity of a surface Σ is the sum $\xi(\Sigma) := \sum_i \xi(\Sigma_i)$ over the connected components Σ_i of Σ.

An *annulus* is a connected surface Σ of complexity -1 (i.e., $\Sigma = S_{0,2}$). Annuli are exceptional in several respects, and must be handled with care throughout our discussion.

The *entropy* of a connected surface Σ is defined to be $h_\Sigma := 2\,|\xi(\Sigma)|$ provided that $\xi(\Sigma) \geqslant -1$ and is defined to be zero otherwise, so that spheres, tori, once-punctured spheres, and thrice-punctured spheres have entropy equal to 0, and annuli have entropy 2. As for complexity, the entropy of a disconnected surface Σ is defined as the sum $h_\Sigma = \sum_i h_{\Sigma_i}$ of the entropies of its connected components Σ_i.

Convention 3.1. For the duration of this paper, we henceforth fix a connected surface S with $\xi(S) > 0$.

Notation 3.2. We use the notation $A \overset{+}{\prec} B$ or $B \overset{+}{\succ} A$ to mean that there is a universal constant c, depending only on the topology of S, such that $A \leqslant B + c$. The notation $A \overset{+}{=} B$ means that $A \overset{+}{\prec} B$ and $A \overset{+}{\succ} B$ both hold. We instead use the notation $\overset{+}{\prec}_x$, $\overset{+}{\succ}_x$ or $\overset{+}{=}_x$ to indicate that the implied constant c depends only on S and the quantity x.

3.1. **Curves.** The term "curve" will always refer to an isotopy class of essential simple closed curves. More precisely, a *curve* in a non-annular surface Σ is (the orientation-reversing-and-isotopy class of) an embedding $\mathbb{S}^1 \to \Sigma$ of the circle that is neither nullhomotopic nor homotopic into an end of Σ. A *curve in an annulus* is rather (the isotopy class of) an embedding of $\mathbb{S}^1$ whose complement consists of two annuli. Notice that each annulus Y has a unique curve; we call this the *core* of the annulus and denote it by ∂Y. The set of curves in Σ will be denoted by $\Gamma(\Sigma)$.

The *intersection number* of a pair of curves $\alpha, \beta \in \Gamma(\Sigma)$ is the minimum $i(\alpha, \beta)$ of the quantity $\left|a(\mathbb{S}^1) \cap b(\mathbb{S}^1)\right|$ over all representative embeddings $a, b\colon \mathbb{S}^1 \to \Sigma$. The curves α, β are *disjoint*, written $\alpha \perp \beta$, if they admit disjoint representatives, that is, if $i(\alpha, \beta) = 0$. Otherwise the curves *cut* each other and we write $\alpha \pitchfork \beta$. A *curve system* on Σ is a nonempty set of pairwise disjoint curves in Σ; being a set of curves, note that the elements of a curve system are necessarily distinct and thus nonisotopic. Curve systems α and β *cut* each other, denoted $\alpha \pitchfork \beta$, if $\alpha_0 \pitchfork \beta_0$ for some $\alpha_0 \in \alpha$ and $\beta_0 \in \beta$ (that is, if their union is not a curve system). Otherwise the curve systems are disjoint and we write $\alpha \perp \beta$.

3.2. **Subsurfaces.** Following [BKMM, §2.1], an *(essential) subsurface* of Σ is a subset $Y \subset \Sigma$ that is itself a surface and has the following structure:

- Y is a union of (not necessarily all) complementary components of a (possibly empty) embedding $C\colon (\sqcup_{i=1}^k \mathbb{S}^1) \to \Sigma$ whose components $C_1, \dots, C_k$ each define curves in Σ. The components of $C \cap \overline{Y}$ are then pairwise disjoint or isotopic curves in Σ and so determine a curve system ∂Y on Σ; these are the *boundary curves* of Y.
- No two components of Y are isotopic (equivalently, no two annuli components are isotopic).
- No component of Y is a thrice-punctured sphere.

Subsurfaces of Σ are identified when they are isotopic in Σ. We reserve the term *domain* for a connected subsurface of Σ. Note that Σ is itself a subsurface of Σ and has trivial boundary $\partial\Sigma = \varnothing$

Given a subsurface Y of Σ, the inclusion induces an injection $\Gamma(Y) \hookrightarrow \Gamma(\Sigma)$ that allows us to identify $\Gamma(Y)$ with the set of curves in Σ that are essential in Y. Note that a curve of ∂Y lies in $\Gamma(Y) \subset \Gamma(\Sigma)$ if and only if it is the core of an annulus component of Y. For any subsurface Y of S we use the notation $\overline{\Gamma}(Y) = \partial Y \cup \Gamma(Y)$ for the set of curves that are essential in Y or homotopic to a boundary component.

On the set of subsurfaces of Σ, define a relation $Z \sqsubset Y$ to mean $\Gamma(Z) \subset \Gamma(Y)$ as subsets of $\Gamma(\Sigma)$. In this case, one may adjust Z by an isotopy so that it is a bona fide subsurface of Y. The relation $\sqsubset$ may thus be safely read as "is an (essential) subsurface of," and we accordingly use the notation $Y \sqsubset \Sigma$ to mean that Y is a subsurface of Σ. We also use $Z \sqsubsetneq Y$ to mean $Z \sqsubset Y$ and $Z \neq Y$.

Lemma 3.3 (Behrstock–Kleiner–Minsky–Mosher [BKMM, Lemma 2.1])**.** *The set of subsurfaces of Σ (including $\varnothing$) is a lattice with partial order $\sqsubset$ and meet/join operations (termed "essential union/intersection") denoted by $\sqcup$ and $\sqcap$:*

- *Subsurfaces Z and Y are isotopic if and only if $\Gamma(Z) = \Gamma(Y)$.*
- *$Y \sqcup Z$ is the unique $\sqsubset$–minimal W so that $Y \sqsubset W$ and $Z \sqsubset W$,*
- *$Y \sqcap Z$ is the unique $\sqsubset$–maximal W so that $W \sqsubset Y$ and $W \sqsubset Z$.*

3.3. **Cutting.** We extend the binary relation $\pitchfork$ to the set of all curve systems and subsurfaces of Σ as follows: We have already defined $\alpha \pitchfork \beta$ when α, β are curves or curve systems on Σ. A curve system α is said to be *disjoint* from a subsurface $Y \sqsubset \Sigma$, denoted by $\alpha \perp Y$, if their isotopy classes have disjoint representatives; otherwise they are said to *cut* and we write $\alpha \pitchfork Y$.

Two subsurfaces Y and Z of Σ are said to be *disjoint*, denoted $Y \perp Z$, if they have disjoint representatives. They are *nested* if $Y \sqsubset Z$ or $Z \sqsubset Y$. Otherwise, Y and Z are said to *cut* (or *intersect transversely*), denoted $Y \pitchfork Z$. Observe that $Y \pitchfork Z$ is equivalent to $Y \pitchfork \partial Z$ and $Z \pitchfork \partial Y$. We say that Y and Z *intersect* if they are not disjoint. We will also need a relative form of cutting:

Definition 3.4 (Relative cutting)**.** Given a subsurface $V \sqsubset \Sigma$, we say that two subsurfaces Y_1, Y_2 of Σ *cut relative to* V if $Y_1' \pitchfork Y_2'$ for all subsurfaces $Y_i' \sqsubset Y_i$ that intersect V. In this case we write $Y_1 \pitchfork_V Y_2$.

Remark 3.5. Note that $Y_1 \pitchfork_V Y_2$ vacuously holds if either Y_1 or Y_2 is disjoint from V; thus $Y_1 \pitchfork_V Y_2$ does not necessarily imply $Y_1 \pitchfork Y_2$. However, if Y_1 and Y_2 both intersect V, then $Y_1 \pitchfork_V Y_2 \implies Y_1 \pitchfork Y_2$.

Relative cutting has the useful property, over cutting, in that it passes to subsurfaces. That is, $Y \pitchfork_V Z$ implies $Y' \pitchfork_V Z'$ for all subsurfaces $Y' \sqsubset Y$ and $Z' \sqsubset Z$.

3.4. **Mapping class group.** The *mapping class group* of a surface Σ is the quotient

$$\mathrm{Mod}(\Sigma) := \mathrm{Homeo}^+(\Sigma)/\mathrm{Homeo}_0(\Sigma)$$

of the group of orientation-preserving homeomorphisms of Σ by the normal subgroup $\mathrm{Homeo}_0(\Sigma)$ of homeomorphisms that are isotopic to the identity. Observe that $\mathrm{Mod}(\Sigma)$ acts on the sets of curves, curve systems, and subsurfaces of Σ preserving the relations $\pitchfork$, $\perp$, and $\sqsubset$. We write $D_\alpha \in \mathrm{Mod}(\Sigma)$ for the (left) *Dehn twist* about a curve α in Σ; see [FM, Chapter 3].

Definition/Theorem 3.6 (Nielsen–Thurston Classification [Thu])**.** Let Σ be a connected surface with $\chi(\Sigma) < 0$. An element $\phi \in \mathrm{Mod}(\Sigma)$ is said to be:

- *finite-order* if there is an integer $k \geqslant 1$ such that ϕ^k is trivial in $\mathrm{Mod}(\Sigma)$,
- *reducible* if it is not finite-order and there is a curve system C on Σ such that $\phi(C) = C$; any such C is called a *reducing system* for ϕ, A canonical reducing system has the property that for some integer k, for each complementary component of C, ϕ^k fixes the component and is either the identity or pseudo-Anosov.
- *pseudo-Anosov* if the set $\{\phi^k(\alpha) \mid k \in \mathbb{Z}\}$ is infinite for each curve α in Σ.

Moreover, each $\phi \in \mathrm{Mod}(\Sigma)$ falls into *exactly one* of these categories.

3.5. **Curve complexes.** Let Σ be a connected surface with $\xi(\Sigma) \geqslant 1$. The *curve graph* (or *curve complex*) of Σ is the simplicial graph $\mathcal{C}(\Sigma)$ with vertex set $\mathcal{C}_0(\Sigma) = \Gamma(\Sigma)$ and edges defined as follows: If $\xi(\Sigma) \geqslant 2$, then two curves α, β in Σ are joined by an edge in $\mathcal{C}(\Sigma)$ if they are disjoint. If $\xi(\Sigma) = 1$, then $\Sigma = S_{1,1}$ or $\Sigma = S_{0,4}$ and two curves α, β are joined by an edge if they intersect once in the case of $S_{1,1}$ or twice in the case of $S_{0,4}$; in either case $\mathcal{C}(\Sigma)$ is the well-known Farey graph.

The curve graph $\mathcal{C}(Y)$ of an annular subsurface $Y \sqsubset \Sigma$ is defined following [MM1]: The annular cover $A_Y \to \Sigma$ to which Y lifts homeomorphically admits a natural compactification $\bar{A}_Y$ coming from the usual compactification of $\tilde{\Sigma} \equiv \mathbb{H}^2$. An *(essential) arc* in $\bar{A}_Y$ is an isotopy class, rel endpoints, of properly embedded arcs whose endpoints lie on opposite components of $\partial \bar{A}_Y$. Two such arcs are said to be *disjoint* if they have representatives with disjoint interiors. We may then define the *curve graph* of Y to be the simplicial graph $\mathcal{C}(Y)$ whose vertices are arcs in $\bar{A}_Y$ with edges joining every pair of disjoint arcs.

Every curve graph is given the path metric in which each edge has length 1. A geodesic metric space is said to be *δ–hyperbolic* if each side of every geodesic triangle is contained in the δ–neighborhood of the union of the other two sides.

Theorem 3.7 (Masur–Minsky [MM2])**.** *There is an integer $\delta > 0$ depending only on S such that the curve graph $\mathcal{C}(\Sigma)$ of every domain $\Sigma \sqsubset S$ is δ–hyperbolic.*

The number δ may in fact be chosen independently of S [Aou, Bow, CRS, HPW]; e.g., $\delta = 17$ works. We emphasize that by our convention δ *is an integer*.

3.6. **Markings.** Following [MM1], a *(clean) marking* μ on a connected surface Σ with $\xi(\Sigma) = k \geqslant 1$ is a pair $\mu = (\mathrm{base}(\mu), t)$, where $\mathrm{base}(\mu) = (\beta_1, \dots, \beta_k)$ is a maximal curve system on Σ (a so-called *pants decomposition*) and $t = (t_1, \dots, t_k)$ is a tuple of curves on Σ such that β_i and t_i intersect minimally (either once or twice) subject to the condition that β_i and t_j are disjoint for all $i \neq j$. A *marking on an annulus* $Y \sqsubset \Sigma$ is a pair $\mu = (\mathrm{base}(\mu), t)$ where $\mathrm{base}(\mu) = \beta$ is the core of Y and $t \in \mathcal{C}_0(Y)$ is a vertex of the curve complex of Y. The curve system $\mathrm{base}(\mu)$ and tuple t are respectively called the *base* and *transversal* of the marking μ. (Our definition of marking is more restrictive than that used in [MM1] but shall suffice for our purposes).

A marking on a disconnected subsurface $\Sigma \sqsubset S$ is simply a choice of marking for each component of Σ. The set of markings on a surface Σ will be denoted $\mathcal{M}_0(\Sigma)$.

3.7. **Subsurface projections.** Let Σ be a connected surface with $\xi(\Sigma) \geqslant 1$ and $Y \sqsubset \Sigma$ a domain. We define a map $\pi_Y : \mathcal{C}_0(\Sigma) \to \mathcal{P}(\mathcal{C}_0(Y))$, with codomain the set of subsets of $\mathcal{C}_0(Y)$, as follows. For concreteness, fix a complete hyperbolic metric on Σ and realize Y such that ∂Y is a union of geodesics. First suppose $\xi(Y) \geqslant 1$ so that Y is not an annulus. If $\alpha \in \mathcal{C}_0(\Sigma) = \Gamma(\Sigma)$ is disjoint from Y we set $\pi_Y(\alpha) = \varnothing$ and if $\alpha \in \Gamma(Y)$ we set $\pi_Y(\alpha) = \alpha$. Otherwise $\alpha \pitchfork Y$ and the geodesic representative of α intersects Y in a collection of proper arcs. For each component a_i of $\alpha \cap Y$, the boundary of a regular neighborhood of $a_i \cup \partial Y$ in Y determines one or two curves in Y; the set of all curves obtained in this way is $\pi_Y(\alpha)$. For an annulus Y and a curve $\alpha \in \mathcal{C}_0(\Sigma) = \Gamma(\Sigma)$, we still set $\pi_Y(\alpha) = \varnothing$ when α is disjoint from Y. Otherwise $\alpha \pitchfork Y$ and we let $\pi_Y(\alpha)$ be the set of lifts of α that give essential arcs in the compactified annular cover $\bar{A}_Y$.

We extend the domain of π_Y to *sets* of curves by adopting the convention that $\pi_Y(A) = \bigcup_{\alpha \in A} \pi_Y(\alpha)$ for any set A of curves on Σ. We also define the notation

$$\mathrm{diam}_Y(A) := \mathrm{diam}_{\mathcal{C}(Y)}(\pi_Y(A)) \quad \text{and} \quad d_Y(A, B) := \mathrm{diam}_{\mathcal{C}(Y)}(\pi_Y(A) \cup \pi_Y(B))$$

for any sets or elements A and B in $\mathcal{C}_0(\Sigma)$. It is easy to see (e.g., [MM1, Lemma 2.3]) that $\mathrm{diam}_Y(\alpha) \leqslant 2$ for any curve system α on Σ. In particular, if $\alpha_0, \dots, \alpha_n$ is any path in $\mathcal{C}(\Sigma)$ such that $\alpha_i \pitchfork Y$ for all i, then $d_Y(\alpha_0, \alpha_n) \leqslant 2n$. There is a much stronger result for geodesics in $\mathcal{C}(\Sigma)$:

Theorem 3.8 (Bounded geodesic image [MM1])**.** *There is a constant $\mathsf{Q} > 0$ satisfying the following: Let Σ be a domain in S with $\xi(\Sigma) \geqslant 1$ and let $Y \subsetneq \Sigma$ be a proper subdomain. If g is a geodesic segment, ray, or biinfinite line in $\mathcal{C}(\Sigma)$ with $\pi_Y(\alpha) \neq \varnothing$ for all $\alpha \in g$, then $\mathrm{diam}_Y(g) \leqslant \mathsf{Q}$.*

If $\mu = (\mathrm{base}(\mu), t)$ is a marking in Σ, we define $\pi_Y(\mu) = \pi_Y(\mathrm{base}(\mu)) \cup \pi_Y(t)$; that is, $\pi_Y(\mu)$ is obtained by viewing μ as the set $\{\beta_1, \ldots, \beta_k, t_1, \ldots, t_k\}$ of base curves and transversals and projecting all these to Y. For any marking $\mu \in \mathcal{M}_0(\Sigma)$ and domain $Y \sqsubset \Sigma$, one easily finds that

$$\pi_Y(\mu) \neq \varnothing \quad \text{with} \quad \mathrm{diam}_Y(\mu) = \mathrm{diam}_{\mathcal{C}(Y)}(\pi_Y(\mu)) \leqslant 6.$$

We shall need the following elementary facts.

Lemma 3.9 ([BKMM, Lemma 2.12])**.** *There is a constant* k *so that the following holds for any nested domains* $V \sqsubset W \sqsubset \Sigma$ *in* S*. Let* α *denote any curve system or marking on* Σ*. Then* $\pi_V(\alpha)$ *is nonempty if and only if* $\pi_V(\pi_W(\alpha))$ *is nonempty and, moreover,* $\mathrm{diam}_{\mathcal{C}(V)}\left(\pi_V(\alpha) \cup \pi_V(\pi_W(\alpha))\right) \leqslant \mathsf{k}$.

Lemma 3.10. *For each* $C \geqslant 1$ *there exists* $C' \geqslant 1$ *with the following property. If* $\mu \in \mathcal{M}_0(\Sigma)$ *is a marking of a subsurface* Σ *in* S *and* α *is a curve system on* Σ *with* $d_V(\mu, \alpha) \leqslant C$ *for all domains* $V \sqsubset \Sigma$*, then there is a marking* $\mu' \in \mathcal{M}_0(\Sigma)$ *with* $\alpha \subset \mathrm{base}(\mu')$ *and* $d_V(\mu', \mu) \leqslant C'$ *for all domains* $V \sqsubset \Sigma$.

Proof. We follow the procedure, on page 798, of [BKMM, §2.2] to project the marking μ of Σ to a new marking μ' of Σ. Since $\pi_V(\mu)$ and $\pi_V(\alpha)$ coarsely agree for all domains $V \sqsubset \Sigma$, we may choose the components of α to be base curves in this inductive construction. This amounts to the following: For each curve $a \in \alpha$, let A denote the annulus with $\partial A = a$ and choose a transversal $t_a \in \pi_A(\mu) \subset \mathcal{C}(A)$. Then for each complementary component V of $\Sigma \backslash \alpha$, project μ to a marking μ_V of V as in [BKMM, §2.2]. These fit together to give a full marking (in the sense of [BKMM]) μ' of Σ that we may straighten to be a (clean) marking in our sense. The uniform bound on $d_V(\mu, \mu')$ is then a consequence of [BKMM, Lemma 2.10] $\square$

3.8. **Teichmüller space.** The *Teichmüller space* $\mathcal{T}(\Sigma)$ of a connected surface Σ with $\xi(\Sigma) \geqslant 1$ is the set of isotopy classes of marked hyperbolic structures on Σ. More precisely, $\mathcal{T}(\Sigma)$ is the space of pairs (X, f), where $f \colon \Sigma \to X$ is a homeomorphism between Σ and a complete, finite-area hyperbolic surface X, up to the equivalence relation $(X, f) \sim (Y, g)$ if there is an isometry $\Psi \colon X \to Y$ with $\Psi \circ f$ isotopic to g. Observe that the mapping class group $\mathrm{Mod}(\Sigma)$ naturally acts on $\mathcal{T}(\Sigma)$ by changing the marking: $\phi \cdot (X, f) = (X, f \circ \phi^{-1})$. This action is properly discontinuous and isometric with respect to the *Teichmüller metric* given by

$$d_{\mathcal{T}(\Sigma)}((X, f), (Y, g)) := \inf\left\{\tfrac{1}{2}\log \mathrm{QC}(\Phi) \mid \Phi \sim g \circ f^{-1}\right\},$$

where the infimum is over all quasiconformal maps homotopic to $g \circ f^{-1}$ and $\mathrm{QC}(\Phi)$ denotes the maximum dilatation of Φ. It is known that $d_{\mathcal{T}(\Sigma)}$ is a complete metric on $\mathcal{T}(\Sigma)$ and that the induced topology is homeomorphic to $\mathbb{R}^{h_\Sigma}$ (e.g., see [FM]). Moreover, $\mathcal{T}(\Sigma)$ is a *unique* geodesic metric space: every pair of points $x, y \in \mathcal{T}(\Sigma)$ are connected by a unique geodesic segment, which we denote by $[x, y]$.

For a point $x \in \mathcal{T}(\Sigma)$ and curve $\alpha \in \Gamma(\Sigma)$, we write $\ell_x(\alpha)$ for the length of the geodesic representative of α in the hyperbolic metric x. For any given $\epsilon > 0$, the *ϵ–thick part* of Teichmüller space is defined to be the subset

$$\mathcal{T}_\epsilon(\Sigma) = \{x \in \mathcal{T}(\Sigma) \mid \ell_x(\alpha) \geqslant \epsilon \text{ for all } \alpha \in \Gamma(\Sigma)\}$$

consisting of those metrics for which all curves have length at least ϵ. It is known that $\mathrm{Mod}(\Sigma)$ acts cocompactly on $\mathcal{T}_\epsilon(\Sigma)$ for every $\epsilon > 0$.

The conformal structure of an annulus is determined its modulus, and the usual notion of the Teichmüller space of the annulus is accordingly $\mathbb{R}_+$. In this paper we also include curve complex distance and so formally define the *Teichmüller space of an annulus* $A \sqsubset S$ with core $\alpha = \partial A$ to be the upper half plane

$$\mathcal{T}(A) = \mathcal{T}(\alpha) = \mathbb{H}^2 = \{x + iy \mid y > 0\}$$

equipped with the hyperbolic metric $ds^2 = \frac{dx^2+dy^2}{4y^2}$ of curvature -4. In this definition $1/y$ represents the hyperbolic length of the core ∂A and x measures twisting about A; so the thick part in this case is $\mathcal{T}_\epsilon(A) = \{x + iy \mid 0 < y \leqslant \frac{1}{\epsilon}\}$. However, due to their role in Minsky's product regions (Theorem 3.11 below), for annuli we are more concerned about the *thin part* $\mathcal{T}_\epsilon^{\leqslant}(\alpha) = \{x + iy \mid y \geqslant \frac{1}{\epsilon}\}$. Note that $\mathcal{T}(A)$ is homeomorphic to $\mathbb{R}^{h_A} = \mathbb{R}^2$.

3.9. **Product regions.** Given any $\epsilon > 0$ and curve system α on Σ, write $\mathcal{H}_{\epsilon,\alpha}(\Sigma) = \{x \in \mathcal{T}(\Sigma) \mid \ell_x(a) < \epsilon, \forall a \in \alpha\}$ for the set of hyperbolic metrics in which the curves in α are all shorter than ϵ. Let $\mathcal{P}(\Sigma|\alpha)$ denote the product space

$$\mathcal{P}(\Sigma|\alpha) := \prod_{V \subset \Sigma \backslash \alpha} \mathcal{T}(V) \times \prod_{a \in \alpha} \mathcal{T}(\alpha),$$

where the first product is over the connected components V of $\Sigma \backslash \alpha$, equipped with the *sup metric* $d_{\mathcal{P}(\Sigma|\alpha)} = \max_{V,a}\{d_{\mathcal{T}(V)}, d_{\mathcal{T}(\alpha)}\}$.

By using Fenchel–Nielsen coordinates adapted to the curve system α (see [Min]), one obtains a natural homeomorphism

$$\Phi_\alpha \colon \mathcal{T}(\Sigma) \to \mathcal{P}(\Sigma|\alpha)$$

under which the $\mathcal{T}(\alpha)$ component of $\Phi_\alpha(w)$ is $\tau_\alpha(w) + \frac{i}{\ell_\alpha(w)}$, where τ_α is the FN twist parameter of w for the curve α. The following foundational result of Minsky says that, for sufficiently small ϵ, the restriction of Φ_α to $\mathcal{H}_{\epsilon,\alpha}(\Sigma)$ distorts distances by a bounded *additive* amount:

Theorem 3.11 (Minsky, [Min])**.** *There exists $\mathsf{D}_0 \geqslant 0$, depending only on S, such that the following holds for all sufficiently small $\epsilon > 0$: For any domain $\Sigma \sqsubset S$ and any curve system α on Σ, all pairs of points $x, y \in \mathcal{H}_{\epsilon,\alpha}(\Sigma)$ satisfy*

$$\left|d_{\mathcal{T}(\Sigma)}(x, y) - d_{\mathcal{P}(\Sigma|\alpha)}(\Phi_\alpha(x), \Phi_\alpha(y))\right| \leqslant \mathsf{D}_0.$$

Moreover for every component $V \subset \Sigma \backslash \alpha$ and essential curve $\gamma \in \Gamma(V)$, the length of γ in x and in the $\mathcal{T}(V)$–component of $\Phi_\alpha(x)$ have ratio in $[\mathsf{D}_0^{-1}, \mathsf{D}_0]$.

3.10. **Volumes and nets.** The Teichmüller space $\mathcal{T}(\Sigma)$ admits a *holonomy measure* $\mathbf{m}$ defined as the push forward of the Masur–Veech measure on the space of unit area quadratic differentials; see [Mas, Vee]. Let us write $\mathrm{Ball}(x, r) \subset \mathcal{T}(\Sigma)$ for the metric ball of radius $r > 0$ centered at $x \in \mathcal{T}(\Sigma)$. We shall need two facts about $\mathbf{m}$. Firstly, Eskin and Mirzakhani [EM, Lemma 3.1] have shown that there exists a constant c such all balls of radius $\mathsf{c} \leqslant \mathsf{r} \leqslant 2\mathsf{c}$ have volume $\mathbf{m}(\mathrm{Ball}(x, r))$ uniformly bounded above and below. Secondly, Athreya, Bufetov, Eskin and Mirzakhani have calculated the volumes of large balls centered in the thick part:

Theorem 3.12 ([ABEM, Theorem 1.3])**.** *There exists a constant $C_1 \geqslant 1$ such that for every domain $\Sigma \sqsubset S$, thick point $x \in \mathcal{T}_{\epsilon_0}(\Sigma)$ and radius $r > 0$, one has* $\mathbf{m}(\mathrm{Ball}(x, r)) \leqslant C_1 e^{h_\Sigma r}$.

Note that these results in [ABEM] and [EM] are stated for closed surfaces of genus at least 2, but the proofs in fact hold for surfaces with $\xi(\Sigma) \geqslant 2$ (see, e.g., [Mah1]). When $\xi(\Sigma) = \pm 1$ and thus $h_\Sigma = 2$, for $\Sigma = S_{0,4}$ or $S_{1,1}$ or $S_{0,2}$, the Teichmüller space $\mathcal{T}(\Sigma)$ is isometric to the hyperbolic plane with constant curvature -4 and these calculations are elementary. For annuli $A \sqsubset S$ we will primarily be concerned with thin regions $\mathcal{T}_\epsilon^{\leqslant}(A)$ consisting of points x with $\ell_x(\partial A) \leqslant \epsilon$, whose volumes may be estimated as follows:

Lemma 3.13. *There exists C_2 so that if A is an annulus and $x \in \mathcal{T}_{\epsilon_0}(A)$ is thick, then the $\epsilon = e^{4\mathsf{c}}\epsilon_0$ thin region within $r > 0$ of x has* $\mathbf{m}\big(\mathrm{Ball}(x,r) \cap \mathcal{T}^{\leqslant}_{e^{4\mathsf{c}}\epsilon_0}(A)\big) \leqslant C_2 e^r$.

Proof. We use the disc model of hyperbolic space with the metric $\frac{|dz|^2}{4(1-|z|^2)^2}$.

We take the horocycle $|z - \frac{1}{2}| = \frac{1}{2}$ centered at $\frac{1}{2}$ of Euclidean radius $\frac{1}{2}$. Its interior is the horoball H. We want to bound the area of the intersection of H with the disc of Euclidean radius r centered at 0.

For sake of simplicity assume r is of the form $r = 1 - \frac{1}{2^n}$ for some large n. For a fixed j_0 and $j = j_0, \ldots, n-1$ consider the points in the horoball intersected with annulus $A_j = \{z : 1 - \frac{1}{2^j} \leqslant |z| \leqslant 1 - \frac{1}{2^{j+1}}$ We want to find the set of θ such that

$$|\frac{1}{2} + \frac{1}{2}e^{i\theta}| \geqslant 1 - \frac{1}{2^j}.$$

We find

$$\cos\theta \geqslant 2(1 - 1/2^j)^2 - 1 \geqslant 1 - 1/2^{j-1}.$$

For j large enough, $|\theta| \leqslant \frac{1}{2}$ and so

$$\theta^2 \leqslant 1/(2^{j-4})$$

or $\theta \leqslant \theta_0 = \frac{1}{2^{j/2-2}}$. Using the origin as base point we write the point $1/2 + 1/2e^{i\theta}$ as $se^{i\phi}$ and we see that

$$|\phi| \leqslant |\theta| \leqslant \theta_0.$$

The total area of the annulus A_j is of order 2^j and so the bound $|\phi| \leqslant \theta_0$ implies the area of $A_j \cap H$ is at most $C2^{j/2}$ for a constant C, and so is bounded by a multiple of the square root of the total area of A_j. We sum over $j \leqslant n$ to get the desired inequality. $\square$

Definition 3.14 (Fixed Nets)**.** For each domain $\Sigma \sqsubset S$, we henceforth fix a $(\mathsf{c}, 2\mathsf{c})$–net $\mathcal{N}(\Sigma)$ in $\mathcal{T}(\Sigma)$; this is a subset such that the c–balls centered on $\mathcal{N}(\Sigma)$ are all disjoint but the $2\mathsf{c}$–balls cover $\mathcal{T}(\Sigma)$. We additionally write $\mathcal{N}_{\epsilon_0}(\Sigma) = \mathcal{N}(\Sigma) \cap \mathcal{T}_{\epsilon_0}(\Sigma)$ for the set of thick net points and, for annuli $A \sqsubset S$, write $\mathcal{N}^{\leqslant}_\epsilon(A) = \mathcal{N}(A) \cap \mathcal{T}^{\leqslant}_\epsilon(A)$ for the set of thin net pints.

Since c and $2\mathsf{c}$ balls have uniformly controlled volumes, one finds, as in equation (17) of [EM], that the volume of any ball is comparable to the number of net points it contains. Thus Theorem 3.12 and Lemma 3.13 immediately give

Lemma 3.15. *There is a uniform constant $\mathsf{P} > 0$ such that for any domain $\Sigma \sqsubset S$, thick point $x \in \mathcal{T}_{\epsilon_0}(\Sigma)$, and radius $r > 0$, one has* $\#|\mathcal{N}(\Sigma) \cap \mathrm{Ball}(x,r)| \leqslant \mathsf{P}e^{h_\Sigma r}$. *Furthermore, if Σ is an annulus, then* $\#\left|\mathcal{N}^{\leqslant}_{e^{4\mathsf{c}}\epsilon_0}(A) \cap \mathrm{Ball}(x,r)\right| \leqslant \mathsf{P}e^r$.

3.11. **Projecting to curve complexes.** Every point $x \in \mathcal{T}(\Sigma)$ admits a *Bers marking* $\mu_x \in \mathcal{M}_0(\Sigma)$ constructed as follows: Greedily choose a shortest pants decomposition $\mathrm{base}(\mu_x)$ on the hyperbolic surface x, then choose a shortest-possible transversal t_i for each base curve $\beta_i \in \mathrm{base}(\mu_x)$. There is a universal *Bers constant* $L_0 > 0$, depending only on S, such that any Bers marking μ_x of any point $x \in \mathcal{T}(\Sigma)$ in the Teichmüller space of any non-annular domain $\Sigma \sqsubset S$ satisfies $\ell_x(\beta) \leqslant L_0$ for all $\beta \in \mathrm{base}(\mu_x)$. For an annulus, the Bers marking of a point $x \in \mathcal{T}(A)$ is just a pair $(\partial A, t)$ where the transversal $t \in \mathcal{C}_0(A)$ records the horizontal component of x.

Given any domain $V \sqsubset \Sigma$, the *projection* of $x \in \mathcal{T}(\Sigma)$ to $\mathcal{C}(V)$ is defined as

$$\pi_V(x) := \bigcup \big\{\pi_V(\mu_x) \mid \mu_x \text{ is a Bers marking on } x\big\}.$$

The *projection distance* in V of a pair of points $x, y \in \mathcal{T}(\Sigma)$ is then defined to be

$$d_V(x,y) := \mathrm{diam}_{\mathcal{C}(V)}\big(\pi_V(x) \cup \pi_V(y)\big).$$

This projection is coarsely Lipschitz for nonannuli: There is a constant $\mathsf{L} \geqslant 1$, depending only on S, such that for all domains $\Sigma \sqsubset S$ and $x, y \in \mathcal{T}(\Sigma)$ one has

$$d_V(x,y) \leqslant \mathsf{L} d_{\mathcal{T}(\Sigma)}(x,y) + \mathsf{L} \quad \text{for all nonannular subdomains } V \sqsubset \Sigma. \tag{3.16}$$

We caution that $\pi_A \colon \mathcal{T}(\Sigma) \to \mathcal{C}(A)$ is NOT coarsely Lipschitz for annuli $A \subsetneq \Sigma$, as is evident from the Distance Formula Theorem 3.33. However, we at least have the following coarse continuity for *every* domain $V \sqsubset \Sigma$ and point $x \in \mathcal{T}(\Sigma)$:

$$\mathrm{diam}_{\mathcal{C}(V)}\left(\bigcup_{y \in U} \pi_V(y)\right) \leqslant \mathsf{L} \quad \text{for some neighborhood } U \subset \mathcal{T}(\Sigma) \text{ of } x \tag{3.17}$$

In particular $\mathrm{diam}_{\mathcal{C}(V)}(\pi_V(x)) = d_V(x,x) \leqslant \mathsf{L}$ for any $x \in \mathcal{T}(\Sigma)$, meaning that all potential Bers markings on x have coarsely the same projection to $\mathcal{C}(V)$. For any set F of curves on Σ and any $x \in \mathcal{T}(\Sigma)$, we also adopt the notation

$$d_V(F,x) = \mathrm{diam}_{\mathcal{C}(V)}\big(\pi_V(F) \cup \pi_V(x)\big).$$

3.12. **Alignment.** We say that points in Teichmüller space are aligned if they do not backtrack when projected to curve complexes of subdomains. More precisely, for $\theta \geqslant 0$ an n–tuple $(x_0, \dots, x_n)$ in $\mathcal{T}(\Sigma)$ is said to be *θ–aligned in a domain $V \sqsubset \Sigma$* if for all indices $0 \leqslant i \leqslant j \leqslant k \leqslant n$ we have

$$d_V(x_i, x_j) + d_V(x_j, x_k) \leqslant d_V(x_i, x_k) + \theta.$$

The tuple is moreover *θ–aligned* if it is θ–aligned in every subdomain of Σ. This leads to the following easy consequence of hyperbolicity:

Lemma 3.18. *For any domains $V \sqsubset \Sigma \sqsubset S$, if a triple (x,z,y) in $\mathcal{T}(\Sigma)$ is θ–aligned in V, then $\pi_V(z)$ is contained in the $(\theta/2 + 4\delta + \mathsf{L})$–neighborhood of any $\mathcal{C}(V)$–geodesic connecting $\pi_V(x)$ to $\pi_V(y)$.*

Proof. Let $g = (\gamma_0, \dots, \gamma_k)$ be any geodesic with $\gamma_0 \in \pi_V(x)$ and $\gamma_k \in \pi_V(y)$. We will show that $\pi_V(z)$ lies within $\theta' = \theta/2 + 4\delta + \mathsf{L}$ of g. To this end, choose any $\beta \in \pi_V(z)$ and let α be a closest point to β along g. If $d_V(\beta, \alpha) \leqslant 2\delta$ we are done. Otherwise, we may choose α' on a geodesic from β to α with $d_V(\alpha', \alpha) = 2\delta$. Since α' is not within δ of g, applying δ–hyperbolicity to the geodesic triangles $\triangle(\gamma_0, \alpha, \beta)$

and $\triangle(\gamma_k, \alpha, \beta)$ ensures there are points β_x and β_y on geodesics $[\gamma_0, \beta]$ and $[\beta, \gamma_k]$, respectively, such that $d_V(\beta_x, \alpha'), d_V(\beta_y, \alpha') \leqslant \delta$. Whence

$$\begin{aligned} d_V(x,y) &\leqslant 2\mathsf{L} + d_V(\gamma_0, \gamma_k) \\ &\leqslant 2\mathsf{L} + d_V(\gamma_0, \beta_x) + d_V(\beta_x, \beta_y) + d_V(\beta_y, \gamma_k) \\ &\leqslant 2\mathsf{L} + d_V(\gamma_0, \beta) + d_V(\beta, \gamma_k) + 2\delta - (d_V(\beta_x, \beta) + d_V(\beta, \beta_y)) \\ &\leqslant 2\mathsf{L} + d_V(x,y) + \theta + 2\delta - (d_V(\beta_x, \beta) + d_V(\beta, \beta_y)) \end{aligned}$$

by (3.16) and alignment. Since $d_V(\beta, \alpha) \leqslant d_V(\beta, \beta_x) + 3\delta$ and similarly for β_y, the claimed inequality now follows:

$$2d_V(\beta, \alpha) \leqslant d_V(\beta_x, \beta) + d_V(\beta, \beta_y) + 6\delta \leqslant (2\mathsf{L} + \theta + 2\delta) + 6\delta = 2\theta'. \qquad \square$$

The following result of Rafi says that Teichmüller geodesics are uniformly aligned:

Theorem 3.19 (Rafi, [Raf2])**.** *There is a constant* B *such that for any domain* $\Sigma \sqsubset S$ *with* $\xi(\Sigma) \geqslant 1$*, any consecutive triple of points* a, b, c *along a geodesic in* $\mathcal{T}(\Sigma)$ *is* B*–aligned, that is:* $d_V(a,b) + d_V(b,c) \leqslant d_V(a,c) + \mathsf{B}$ *for all* $V \sqsubset \Sigma$.

This easily implies that Teichmüller geodesics project to within bounded Hausdorff distance of geodesics in curve graphs:

Lemma 3.20. *For any subdomains* $V \sqsubset \Sigma \sqsubset S$ *and any geodesic* $[x,y] \subset \mathcal{T}(\Sigma)$*, the projection* $\pi_V([x,y]) \subset \mathcal{C}(V)$ *has Hausdorff distance at most* $\mathsf{B} + 8\delta + 3\mathsf{L}$ *from any geodesic* g *connecting* $\pi_V(x)$ *to* $\pi_V(y)$.

Proof. Let $g = (\gamma_0, \dots, \gamma_k)$ be any geodesic in $\mathcal{C}(V)$ with $\gamma_0 \in \pi_V(x)$ and $\gamma_k \in \pi_V(y)$. Theorem 3.19 and Lemma 3.18 show that $\pi_V([x,y])$ lies within $B_0 = \mathsf{B}/2 + 4\delta + \mathsf{L}$ of g. Conversely, the fact that π_V is coarsely continuous (3.17) implies there is a subset $\{p_0, p_1, \dots, p_n\} \subset \pi_V([x,y])$ with $p_0 = \gamma_0$, $p_n = \gamma_k$ and $d_V(p_i, p_{i+1}) \leqslant 2\mathsf{L}$ for all $0 \leqslant i < n$. Let $q_i \in g$ be a closest point to p_i. Then $d_V(p_i, q_i) \leqslant B_0$ by the above, and hence $d_V(q_i, q_{i+1}) \leqslant 2B_0 + 2\mathsf{L}$. Since $q_0 = \gamma_0$ and $q_n = \gamma_k$, we see that g lies in the $B_0 + \mathsf{L}$ neighborhood of the set $\{q_0, \dots, q_n\}$. The claim now follows. $\square$

3.13. **Strong alignment.** While our above notion of alignment (§3.12) only concerns curve complex data, the construction of complexity length in §8.2 will require aligned tuples in which lengths of short curves vary roughly convexly. The precise condition is as follows:

Definition 3.21. We say that a tuple $(x_0, \dots, x_n)$ in $\mathcal{T}(\Sigma)$ is *strongly* θ*–aligned*, where $\theta \geqslant 1$, if it is θ–aligned and for every domain $V \sqsubset \Sigma$, there exist points $x_0^V, \dots, x_n^V$ appearing in order along $[x_0, x_n]$ such that for each $0 \leqslant i \leqslant n$ we have

- $d_V(x_i, x_i^V) \leqslant \theta$, and
- if V is an annulus then $\min\{\epsilon_0, \ell_{x_i}(\partial V)\}/\min\{\epsilon_0, \ell_{x_i^V}(\partial V)\}$ lies in $[\frac{1}{\theta}, \theta]$.

This is in fact only a minor strengthening of alignment, in that any aligned tuple may be superficially modified to achieve it:

Lemma 4.7. *For any* $\theta \geqslant 1$ *there exists* $\theta' \geqslant \theta$ *with the following property: For any* θ*–aligned tuple* $(x_0, \dots, x_n)$ *in* $\mathcal{T}(\Sigma)$*, there is a strongly* θ'*–aligned tuple* $(y_0, \dots, y_n)$ *such that* $x_0 = y_0$*,* $x_n = y_n$*, and such that for all* $0 \leqslant i \leqslant n$ *we have*

$$d_V(x_i, y_i) \overset{+}{\prec}_\theta 0 \quad \textit{for every domain } V \sqsubset \Sigma.$$

While the proof is not difficult, the formulation of strong alignment—involving separate tuples for all domain—leads to a somewhat technical and involved argument. As such, we defer the proof until §4.2 below. We note that, along the way, Corollary 4.4 will show the first bulleted condition of strong alignment is redundant, in that any aligned tuple automatically satisfies it for some larger constant.

3.14. **Active intervals.** The subsurface projection values $d_V(x, y)$ are closely related to the manner in which the Teichmüller geodesic $[x, y]$ interacts with product regions. This relationship is conveyed by the following results of Rafi.

In [Raf1, Theorem 3.1], Rafi proves that for all sufficiently small $\epsilon > 0$ there exists $\epsilon > \epsilon' > 0$ such that for any curve $\alpha \in \Gamma(\Sigma)$, every Teichmüller geodesic $[x, y]$ in $\mathcal{T}(\Sigma)$ has a possibly empty subinterval I such that $\ell_z(\alpha) \leqslant \epsilon$ for all $z \in I$ and $\ell_z(\alpha) \geqslant \epsilon'$ for all $z \in [x, y] \backslash I$. For a subsurface $Y \sqsubset \Sigma$, intersecting the intervals for the components of ∂Y (see [Raf1, Corollary 3.4]) produces a corresponding interval for Y that enjoys the following property:

Theorem 3.22 (Rafi [Raf1, Proposition 3.7])**.** *For each sufficiently small $\epsilon > 0$, there exist constants $0 < \epsilon' < \epsilon$ and $\mathsf{M}_\epsilon \geqslant 0$ such that for any domain $\Sigma \sqsubset S$, any Teichmüller geodesic $[x, y] \in \mathcal{T}(\Sigma)$, and any subdomain $V \sqsubset \Sigma$, there is a (possibly empty) connected interval $\tilde{\mathcal{I}}^\epsilon_V \subset [x, y]$ such that*

(1) $\ell_z(\alpha) < \epsilon$ for all $z \in \tilde{\mathcal{I}}^\epsilon_V$ and all $\alpha \in \partial V$.
(2) for all $z \in [x, y] \backslash \tilde{\mathcal{I}}^\epsilon_V$, some component β of ∂V has $\ell_z(\beta) \geqslant \epsilon'$.
(3) $d_V(w, z) \leqslant \mathsf{M}_\epsilon$ for every subinterval $[w, z] \subset [x, y]$ with $[w, z] \cap \tilde{\mathcal{I}}^\epsilon_V = \varnothing$.

Remark 3.23. Notice that item (1) implies $\tilde{\mathcal{I}}^\epsilon_V$ lies in the thin region $\mathcal{H}_{\epsilon,\partial V}(\Sigma)$ and hence, via Minsky's theorem, projects to a path in the $\mathcal{T}(V)$–factor of the product region $\mathcal{P}(\Sigma|\partial V)$. In a later result [Raf2, Theorem 5.3], Rafi produces for *non-annular* domains a related interval, defined in terms of expanding annuli, that satisfies (1) and (3) and additionally fellow travels a unit-speed Teichmüller geodesic in $\mathcal{T}(V)$. Since we will not need this latter property and must deal with annuli, we work the more rudimentary intervals from [Raf1] instead.

Recall that there is a universal *Margulis constant* such that on any complete hyperbolic surface, every pair of curves with hyperbolic length at most this value are disjoint. Hence, for small ϵ, if domains $U, V \sqsubset \Sigma$ have $\partial V \pitchfork \partial U$ then Theorem 3.22(1) implies $\tilde{\mathcal{I}}^\epsilon_V$ and $\tilde{\mathcal{I}}^\epsilon_U$ are disjoint. But it may be that $\tilde{\mathcal{I}}^\epsilon_V$ and $\tilde{\mathcal{I}}^\epsilon_U$ overlap when $V \pitchfork U$. To correct this, we will use a slight variation of $\tilde{\mathcal{I}}^\epsilon_V$.

Definition 3.24 (Uniform constants)**.** Fix once and for all a constant $\epsilon_0 > 0$ smaller than the Margulis constant and small enough for Theorems 3.11 and 3.22 to hold for ϵ_0. Let $\epsilon'_0 < \epsilon_0$ be the companion constant in Theorem 3.22. Define

$$\mathsf{M} = 100(\mathsf{M}_{\epsilon_0} + \delta + \mathsf{L} + \mathsf{B} + \mathsf{Q} + \mathsf{k} + \mathsf{K}),$$

where M_{ϵ_0} is from Theorem 3.22, δ is from Theorem 3.7, L is from (3.16), B is from Theorem 3.19, Q is from the Bounded Geodesic Image Theorem 3.8, k is from Lemma 3.9, and K is from the consistency Theorem 3.37 below.

Definition 3.25 (Active Intervals)**.** For any geodesic $[x, y] \sqsubset \mathcal{T}(\Sigma)$, where $\Sigma \sqsubset S$ is a domain, define the *active interval* of a subdomain $V \sqsubset \Sigma$ along $[x, y]$ as follows:

- For V is annular, $\mathcal{I}_V = \tilde{\mathcal{I}}^{\epsilon_0}_V$.

- For V nonannular, $\mathcal{I}_V$ is the intersection of all subintervals $[a,b] \subset [x,y]$ satisfying both $d_V(x,a) \leqslant 2\mathsf{M}_{\epsilon_0} + 5\mathsf{L}$ and $d_V(b,y) \leqslant 2\mathsf{M}_{\epsilon_0} + 5\mathsf{L}$.

Our active intervals have the following properties:

Lemma 3.26. *Let $[x,y]$ be a geodesic in $\mathcal{T}(\Sigma)$ and $V \sqsubset \Sigma$ a subdomain. Then*

(1) If $d_V(x,y) \geqslant \mathsf{M}$, then $\mathcal{I}_V$ is a nonempty, nondegenerate subinterval of $[x,y]$.
(2) $\mathcal{I}_V \subset \tilde{\mathcal{I}}_V^{\epsilon_0}$, and for all $z \in \mathcal{I}_V$ we have that $\ell_z(\alpha) < \epsilon_0$ for each component α of ∂V and that $\partial V \subset \mathrm{base}(\mu_z)$ for every Bers marking μ_z of z.
(3) If $[w,z]$ is a subinterval of $[x,y]\backslash\mathcal{I}_V$, then $d_V(w,z) < \mathsf{M}/3$.
(4) If $Y \sqsubset \Sigma$ is a domain with $Y \pitchfork V$, then $\mathcal{I}_Y \cap \mathcal{I}_V = \varnothing$. Moreover, if $\mathcal{I}_Y \neq \varnothing$, then $\mathcal{I}_U \cap \mathcal{I}_V = \varnothing$ for every subdomain $U \sqsubset Y$ with $\partial U \pitchfork V$.

Proof. For (1), if V is annular then $\mathcal{I}_V = \tilde{\mathcal{I}}_V^{\epsilon_0}$ by definition; hence $\mathcal{I}_V$ is an interval, and it being either empty or degenerate would imply $d_V(x,y) \leqslant 2\mathsf{M}_{\epsilon_0} + \mathsf{L} < \mathsf{M}$ by Theorem 3.22(3) and (3.17). Next suppose V is nonannular. Since $d_V(x,y) \geqslant \mathsf{M}$ and projections change coarsely continuously (3.17), we may find a nondegenerate subinterval $[w,z] \subset [x,y]$ so that the distances $d_V(x,w)$ and $d_V(z,y)$ are both within L of $2\mathsf{M}_{\epsilon_0}+7\mathsf{L}+\mathsf{B}$. Then every point $u \in [w,z]$ satisfies $d_V(x,u) \geqslant d_V(x,w) - \mathsf{B} > 2\mathsf{M}_{\epsilon_0} + 6\mathsf{L}$ by Theorem 3.19, and similarly $d_V(u,y) \geqslant 2\mathsf{M}_{\epsilon_0} + 6\mathsf{L}$. Thus each subinterval $[a,b]$ in Definition 3.25 contains $[w,z]$. As an intersection of intervals that contain $[w,z]$, $\mathcal{I}_V$ is thus indeed a nonempty, nondegenerate interval.

For (2), first observe that $\mathcal{I}_V \subset \tilde{\mathcal{I}}_V^{\epsilon_0}$; for annuli this is by definition, and for nonannuli it holds since Theorem 3.22(3) implies $\tilde{\mathcal{I}}_V^{\epsilon_0}$ qualifies as one of the intervals in the intersection defining $\mathcal{I}_V$. Now for $z \in \mathcal{I}_V$, Theorem 3.22(1) ensures that $\ell_z(\alpha) < \epsilon_0$ for each component α of ∂V. If μ_z is a Bers marking at z whose pants decomposition $\mathrm{base}(\mu_z)$ fails to contain some component $\alpha \in \partial V$, then we must have $\alpha \pitchfork \beta$ for some $\beta \in \mathrm{base}(\mu_z)$. Since $\mathrm{base}(\mu_z)$ is a *shortest* pants decomposition on z, it must be that $\ell_z(\beta) \leqslant \ell_z(\alpha) < \epsilon_0$. Since ϵ_0 is smaller than the Margulis constant, this forces α and β to be disjoint; a contradiction. Therefore (2) holds.

For annuli, (3) follows immediately from Theorem 3.22(3). If V is not an annulus, then $\mathcal{I}_V \cap [w,z] = \varnothing$ can only occur if $[w,z]$ is disjoint from some interval $[a,b]$, as in Definition 3.25, of the intersection yielding $\mathcal{I}_V$. Without loss of generality, we may assume $[w,z]$ is contained in $[x,a]$. Two applications of Theorem 3.19 then give $d_V(w,z) \leqslant d_V(x,a) + 2\mathsf{B} \leqslant \mathsf{M}_{\epsilon_0} + 5\mathsf{L} + 2\mathsf{B} < \mathsf{M}/3$.

For (4), we may assume $\mathcal{I}_Y$ is nonempty, for else the needed conclusions are immediate or vacuous. Thus it suffices to prove the 'moreover' conclusion for a subdomain $U \sqsubset Y$ satisfying $\partial U \pitchfork V$ (which could possibly be nested in V), since then we may apply it with $Y = U$. If V is an annulus, then the assumption $\partial U \pitchfork V$ reduces to $\partial U \pitchfork \partial V$. Similarly if Y is an annulus then necessarily $U = Y$ and now $Y \pitchfork V$ reduces to $\partial U \pitchfork \partial V$. In either case, the needed conclusion $\mathcal{I}_U \cap \mathcal{I}_V = \varnothing$ would follow from (2) and ϵ_0 being chosen smaller than the Margulis constant. Thus we may assume that neither V nor Y is an annulus. Finally, we also assume $\mathcal{I}_U$ is nonempty, for else there is nothing to prove, and write it as $\mathcal{I}_U = [a,b]$.

Since ∂V projects to Y and all points of $\tilde{\mathcal{I}}_V^{\epsilon_0}$ contain ∂V in their Bers marking by Theorem 3.22(1) (and the choice of ϵ_0), we observe that

$$d_Y(w,z) \leqslant 2\mathsf{L} \quad \text{for all } w,z \in \tilde{\mathcal{I}}_V^{\epsilon_0}.$$

If $\tilde{\mathcal{I}}_Y^{\epsilon_0}$ were contained in $\tilde{\mathcal{I}}_V^{\epsilon_0}$, then Theorem 3.22(3), together with (3.17), would evidently imply $d_Y(x,y) \leqslant 2\mathsf{M}_{\epsilon_0} + 4\mathsf{L}$. Thus $[x,x]$ and $[y,y]$ would both be valid intervals in the intersection from Definition 3.25 that, since Y is nonannular, defines $\mathcal{I}_Y$. As $\mathcal{I}_Y$ is nonempty, $\tilde{\mathcal{I}}_Y^{\epsilon_0}$ therefore cannot be contained in $\tilde{\mathcal{I}}_V^{\epsilon_0}$. Hence we may choose some point $w \in \tilde{\mathcal{I}}_Y^{\epsilon_0}$ outside of $\tilde{\mathcal{I}}_V^{\epsilon_0}$ and suppose, without loss of generality, that w lies in the same component of $[x,y] \backslash \tilde{\mathcal{I}}_V^{\epsilon_0}$ as x, so that $d_V(x,w) \leqslant \mathsf{M}_{\epsilon_0}$ by Theorem 3.22(3). The points $b \in \mathcal{I}_U$ and $w \in \tilde{\mathcal{I}}_Y^{\epsilon_0}$ now respectively contain ∂Y and ∂U in their Bers markings by (2) and Theorem 3.22(1). Since ∂U, ∂Y are disjoint and both project to V, it follows using the triangle inequality that

$$d_V(x,b) \leqslant d_V(x,w) + d_V(w,b) \leqslant \mathsf{M}_{\epsilon_0} + d_V(\partial Y, \partial U) \leqslant \mathsf{M}_{\epsilon_0} + 2.$$

By coarse continuity (3.17), this shows that $d_V(x,b') \leqslant \mathsf{M}_{\epsilon_0} + 3\mathsf{L}$ for some point b' immediately to the right of b. This means $[b', y]$ qualifies for the intersection defining $\mathcal{I}_V$ (recall that we have supposed V is nonannular). Therefore $\mathcal{I}_V$ is contained in $[b', y]$ and hence disjoint from $[a,b] = \mathcal{I}_Y$, as desired. □

This also leads to the following analog of the bounded geodesic image theorem, which roughly says that if Teichmüller points x, y have a large projection to a domain Z, then in any other domain V that cuts ∂Z, the $\mathcal{C}(V)$–geodesic from $\pi_V(x)$ to $\pi_V(y)$ must pass near ∂Z:

Corollary 3.27 (BGIT for Teichmüller space)**.** *For all domains $Z, V \sqsubset \Sigma \sqsubset S$, if Z has a nonempty active interval $\mathcal{I}_Z$ along a geodesic $[x,y]$ in $\mathcal{T}(\Sigma)$, so in particular if $d_Z(x,y) \geqslant \mathsf{M}$, then $d_V(x,\partial Z) + d_V(\partial Z, y) \leqslant d_V(x,y) + \mathsf{M}/3$.*

Proof. This is contentless when ∂Z is disjoint from V, since in that case $d_V(w,\partial Z) = \operatorname{diam}_{\mathcal{C}(V)}(\pi_V(w)) \leqslant \mathsf{L}$ for all w. By Lemma 3.26(2), the hypothesis provides a point $z \in \mathcal{I}_Z \subset [x,y]$ whose Bers markings all contains ∂Z. Hence Theorem 3.19 gives

$$d_V(x,\partial Z) + d_V(\partial Z, y) \leqslant d_V(x,z) + d_V(z,y) \leqslant d_V(x,y) + \mathsf{B} < d_V(x,y) + \mathsf{M}/3. \quad \square$$

3.15. **Time order.** The fact that cutting domains necessarily have disjoint active intervals (Lemma 3.26(4)) allows for the following definition:

Definition 3.28 (Time order)**.** Given a Teichmüller geodesic $[x,y]$ in $\mathcal{T}(\Sigma)$ and a pair $U, V \sqsubset \Sigma$ of domains, we write $U \lessdot V$ or $V \gtrdot U$, and say U *is time-ordered before* V *along* $[x,y]$, to mean that $U \pitchfork V$ and that $\mathcal{I}_U$ and $\mathcal{I}_V$ are nonempty along $[x,y]$ with $\mathcal{I}_U$ occurring before $\mathcal{I}_V$ when traveling from x to y.

While the geodesic in question and its orientation are both omitted from our notation, we will strive to make these clear from context so that the meaning of $U \lessdot V$ is unambiguous in our discussion of time-ordering. The following characterization of time-ordering follows immediately from Lemma 3.26:

Lemma 3.29 (Characterizing time-order)**.** *Let $[x,y]$ be a geodesic in $\mathcal{T}(\Sigma)$ and let $U, V \sqsubset \Sigma$ be subdomains with $U \pitchfork V$. Then $U \lessdot V$ implies $d_V(x,\partial U) < \mathsf{M}/3$ and $d_U(y,\partial V) < \mathsf{M}/3$. Accordingly, if $d_U(x,y)$ and $d_V(x,y)$ are at least M, then the following are equivalent:*
(1) $U \lessdot V$, (2) $d_V(x,\partial U) < \mathsf{M}/3$, (4) $d_U(x,\partial V) \geqslant 2\mathsf{M}/3$,
(3) $d_V(\partial U, y) \geqslant 2\mathsf{M}/3$, (5) $d_U(\partial V, y) < \mathsf{M}/3$.

Corollary 3.30 (Triple time-order and relative cutting)**.** *If $U, V, W \sqsubset \Sigma$ are subdomains with $d_V(x,y) \geqslant \mathsf{M}$ and $U \lessdot V \lessdot W$ along $[x,y]$, then $U \pitchfork_V W$.*

Proof. If the conclusion fails, we may find subdomains $U' \sqsubset U$ and $W' \sqsubset W$ that intersect V but have $\partial U' \perp \partial W'$. In particular, $\pi_V(\partial U') \neq \varnothing \neq \pi_V(\partial W')$. Since every curve system projects to a set of diameter at most 2 in $\mathcal{C}(V)$, this implies

$$d_V(\partial U, \partial W) \leqslant d_V(\partial U, \partial U') + d_V(\partial U', \partial W') + d_V(\partial W', \partial W) \leqslant 6 < \mathsf{M}/3.$$

With Lemma 3.29, this gives $d_V(x,y) \leqslant 2\mathsf{M}/3 + 6$, contradicting $d_V(x,y) \geqslant \mathsf{M}$. $\square$

Corollary 3.31 (Time-ordering subsurfaces)**.** *Suppose $U, V \sqsubset \Sigma$ satisfy $U \lessdot V$ along a geodesic $[x,y]$ in $\mathcal{T}(\Sigma)$. If $d_V(x,y) \geqslant \mathsf{M}$, then $U' \lessdot V$ for all $U' \sqsubset U$ with $\mathcal{I}_{U'} \neq \varnothing$ and $U' \pitchfork V$. Symmetrically, $d_U(x,y) \geqslant \mathsf{M}$ implies $U \lessdot V'$ for all $V' \sqsubset V$ with $\mathcal{I}_{V'} \neq \varnothing$ and $U \pitchfork V'$.*

Proof. Let $U' \sqsubset U$ be a subdomain with $\mathcal{I}_{U'} \neq \varnothing$ and $U' \pitchfork V$. If $U \lessdot V \lessdot U'$, then Corollary 3.30 would give $U \pitchfork_V U'$. As this is clearly false, we must have $U' \lessdot V$ as claimed. The proof for $V' \sqsubset V$ is similar. $\square$

We also have the following basic observation.

Lemma 3.32. *Let $U, V \sqsubsetneq Y \sqsubset \Sigma$ be domains with $U \lessdot V$ along a geodesic $[x,y] \in \mathcal{T}(\Sigma)$. Then $d_Y(x, \partial U) \leqslant d_Y(x, \partial V) + \mathsf{M}/3$.*

Proof. Choose points $u \in \mathcal{I}_U$ and $v \in \mathcal{I}_V$ and note these respectively contain ∂U, ∂V in their Bers markings. Since $U \lessdot V$ forces $u \in [x,v]$, Theorem 3.19 implies

$$d_Y(x, \partial U) \leqslant d_Y(x, u) \leqslant d_Y(x, v) + \mathsf{B} \leqslant d_Y(x, \partial V) + \mathsf{L} + \mathsf{B}. \qquad \square$$

3.16. **Distance estimates.** The following distance formula of Rafi [Raf1] says that Teichmüller distance can be estimated, with controlled multiplicative and additive error, in terms of projection distances:

Theorem 3.33 (Distance Formula; [Raf1])**.** *For each sufficiently large threshold T, there exists $K \geqslant 1$ such that for every domain $\Sigma \sqsubset S$ and all $x, y \in \mathcal{T}_{\epsilon_0}(\Sigma)$,*

$$\frac{1}{K} d_{\mathcal{T}(\Sigma)}(x,y) - K \leqslant \sum_V [d_V(x,y)]_T + \sum_A [\log d_A(x,y)]_T \leqslant K d_{\mathcal{T}(\Sigma)}(x,y) + K,$$

where the first summand is over all non-annular domains $V \sqsubset \Sigma$, the second over all annular domains $A \sqsubset S$, and where $[w]_T$ equals 0 when $w < T$ and otherwise equals w. Moreover, the rightmost inequality above holds for all $x, y \in \mathcal{T}(\Sigma)$.

While the full strength of this result is generally off limits to us, as we cannot afford multiplicative errors, we do make frequent use of the following consequence:

$$x, y \in \mathcal{T}_{\epsilon_0}(\Sigma), d_V(x,y) \leqslant T \text{ for all domains } V \sqsubset S \implies d_{\mathcal{T}(\Sigma)}(x,y) \overset{+}{\prec}_T 0. \tag{3.34}$$

In the case that x and y have short curves but are close in all curve complexes, we have the following variation based on Minsky's product regions.

Lemma 3.35. *Given $w, z \in \mathcal{T}(\Sigma)$, let σ_w (resp. σ_z) denote the multicurve consisting of all curves on w (resp. z) of length at most ϵ_0. Partition $\sigma_w = \delta_w \sqcup \gamma_w$ into those curves δ_w that are disjoint from σ_z and those γ_w that cut σ_z. Partition $\sigma_z = \delta_z \sqcup \gamma_z$ similarly. Set $R = \max\{R_1, R_2\}$ where*

$$R_1 = \max_{\alpha \in \delta_w \cup \delta_z} \frac{1}{2} \left| \log \frac{\min\{\ell_w(\alpha), \epsilon_0\}}{\min\{\ell_z(\alpha), \epsilon_0\}} \right|, \quad R_2 = \max_{\substack{\alpha \in \gamma_w, \beta \in \gamma_z \\ \alpha \pitchfork \beta}} \frac{1}{2} \left(\log \frac{\epsilon_0}{\ell_w(\alpha)} + \log \frac{\epsilon_0}{\ell_z(\beta)} \right).$$

(1) *If $k \geqslant 1$ is such that $\frac{1}{k} \leqslant \frac{\ell_w(\alpha)}{\ell_z(\alpha)} \leqslant k$ for all $\alpha \in \sigma_w \cup \sigma_z$, then $R \leqslant \log k$.*

(2) $d_{\mathcal{T}(\Sigma)}(w,z) \overset{+}{\succ} R$
(3) If $d_V(x,y) \leqslant T$ *for all domains* $V \sqsubset \Sigma$ *then* $d_{\mathcal{T}(\Sigma)}(w,z)) \overset{+}{\prec}_T R$.

Proof. Suppose the hypothesis of (1) holds. For every $\alpha \in \delta_w \cup \delta_z$ we then have $\min\{\ell_w(\alpha), \epsilon_0\} \leqslant \min\{k\ell_z(\alpha), \epsilon_0\} \leqslant k \min\{\ell_z(\alpha), \epsilon_0\}$, so that the α–term in the max defining R_1 is at most $\frac{1}{2}\log k$. Hence $R_1 \leqslant \log k$. Now consider $\alpha \in \gamma_w$ and $\beta \in \gamma_z$ with $\alpha \pitchfork \beta$. Since α and β cannot both be short on z, we have $\ell_z(\alpha) \geqslant \epsilon_0$ and the hypothesis gives $\ell_w(\alpha) \geqslant \epsilon_0/k$. Similarly $\ell_z(\beta) \geqslant \epsilon_0/k$. Hence $R_2 \leqslant \log k$.

For (2), if $x, y \in \mathcal{T}(\Sigma)$ are two points for which $\ell_x(\alpha), \ell_y(\alpha) \leqslant \epsilon_0$, then Minsky's Product Regions Theorem 3.11 implies

$$d_{\mathcal{T}(\Sigma)}(x,y) + \mathsf{D}_0 \geqslant \frac{1}{2}\left|\log\frac{\ell_x(\alpha)}{\ell_y(\alpha)}\right|$$

since $\Phi_\alpha(x), \Phi_\alpha(y) \in \mathcal{P}(\Sigma|\alpha)$ have at least this distance in the $\mathcal{T}(\alpha)$–factor of the product. Now let $\alpha \in \delta_w \cup \delta_z$ realize the max in R_1. If $\alpha \in \sigma_w \cap \sigma_z$ then the above inequality implies $d_{\mathcal{T}(\Sigma)}(w,z) \geqslant R_1 - \mathsf{D}_0$. If $\alpha \in \sigma_w$ but $\alpha \notin \sigma_z$, then we may choose a point $y \in [w,z]$ with $\ell_y(\alpha) = \epsilon_0$ and apply the above to get $d_{\mathcal{T}(\Sigma)}(w,z) \geqslant d_{\mathcal{T}(\Sigma)}(w,y) \geqslant R_1 - \mathsf{D}_0$. The symmetric reasoning applies if $\alpha \in \sigma_z \backslash \sigma_w$. Now consider any $\alpha \in \gamma_w$ and $\beta \in \gamma_z$ with $\alpha \pitchfork \beta$. The geodesic $[w,z]$ has to lengthen α to at least ϵ_0 before the intersecting curve β can become short; hence there are ordered points w', z' along $[w,z]$ with $\ell_{w'}(\alpha) = \ell_{z'}(\beta) = \epsilon_0$. Now the above observation bounds $d_{\mathcal{T}(\Sigma)}(w,w') + \mathsf{D}_0$ below by $\frac{1}{2}\log\frac{\epsilon_0}{\ell_w(\alpha)}$ and symmetrically for $d_{\mathcal{T}(\Sigma)}(z',z)$. Adding these together and taking a max over all such α, β proves $d_{\mathcal{T}(\Sigma)}(w,z) \geqslant R_2 - 2\mathsf{D}_0$.

For (3), we first claim there is a constant k so that $\ell_z(\alpha) \leqslant k$ for all $\alpha \in \delta_w \backslash \sigma_z$. To see this, let z' be the thick point obtained by lengthening every curve in σ_z to have length ϵ_0 (this can be done in Fenchel–Nielsen coordinates by, for example, adjusting the vertical component of $\Phi_{\sigma_z}(z)$ in each $\mathbb{H}^2$ factor of the product region $\mathcal{P}(\Sigma|\sigma_z)$). This adjustment changes neither subsurface projections nor the lengths of curves in $\delta_w \backslash \sigma_z$, since any such α is disjoint from and unequal to the curves in σ_z whose lengths are modified in the adjustment. Do the same to build w'. Then $d_V(w', z') \overset{+}{\prec} T$ for all V so (3.34) implies $d_{\mathcal{T}(\Sigma)}(w', z') \overset{+}{\prec}_T 0$. Now for $\alpha \in \delta_w \backslash \sigma_z$ we have $\ell_z(\alpha)$, $\ell_{z'}(\alpha)$, and $\ell_{w'}(\alpha) = \epsilon_0$ coarsely agree by construction of z' and the fact that w', z' have bounded distance. Hence $\ell_z(\alpha)$ is bounded, proving the claim.

By moving w and z a bounded distance, we may now assume all curves $\alpha \in \delta_w \backslash \sigma_z$ have $\ell_z(\alpha) = \epsilon_0$ and similarly that all $\beta \in \delta_z \backslash \sigma_w$ have $\ell_w(\beta) = \epsilon_0$. Therefore, letting $\mathcal{Y}$ be the multicurve $\delta_w \cup \delta_z$, we have $\ell_w(\alpha), \ell_z(\alpha) \leqslant \epsilon_0$ for all $\alpha \in \mathcal{Y}$. Hence by Minsky's theorem, $d_{\mathcal{T}(\Sigma)}(w,z)$ agrees up to additive error with the distance between $\Phi_{\mathcal{Y}}(w), \Phi_{\mathcal{Y}}(z)$ in $\mathcal{P}(\Sigma|\mathcal{Y}) = \mathcal{T}(\Sigma \backslash \mathcal{Y}) \times \prod_{\mathcal{Y}} \mathcal{T}(\alpha)$. Since the projections $d_\alpha(w,z)$ are bounded, the distance in the $\prod_{\mathcal{Y}} \mathcal{T}(\alpha)$ factor is by definition coarsely given by R_1.

We compute the distance in $\mathcal{T}(\Sigma \backslash \mathcal{Y})$. Using the identification of this factor with the product $\mathcal{P}(\Sigma \backslash \mathcal{Y} | \gamma_w)$, simultaneously lengthen all the curves in γ_w until they achieve length ϵ_0; for each $\alpha \in \gamma_w$ this takes distance $\frac{1}{2}\log\frac{\epsilon_0}{\ell_w(\alpha)}$. Parameterizing this path as $x(t) \in \mathcal{P}(\Sigma|\mathcal{Y})$ for $t > 0$, let us write $\gamma_w^t \subset \gamma_w$ for the curves that are still shorter than ϵ_0 at time t. The same argument as in the claim above shows that if $\beta \in \gamma_z$ is disjoint from γ_w^t, then β has uniformly bounded length at $x(t)$. Hence as soon as $\beta \in \gamma_z$ becomes disjoint from γ_w^t we may begin shortening β until it has length $\ell_z(\beta)$, which takes time $\frac{1}{2}\log\frac{\ell_{x(t)}(\beta)}{\ell_z(\beta)} \overset{+}{\prec} \frac{1}{2}\log\frac{\epsilon_0}{\ell_z(\beta)}$. In this way we build a path of length, up to additive error, R_2 from $\Phi_{\mathcal{Y}}(w)$ to a new point z' so that $\ell_{z'}(\beta) = \ell_z(\beta)$ for all $\beta \in \gamma_z$. As this procedure does not change subsurface projections, we see that $z', \Phi_{\mathcal{Y}}(z) \in \mathcal{P}(\Sigma|\mathcal{Y})$ have bounded distance in the $\mathcal{T}(\Sigma\backslash\mathcal{Y})$–factor. (To see this, look in the product $\mathcal{P}(\Sigma\backslash\mathcal{Y}|\gamma_z)$ and note that for each component V of $\Sigma\backslash(\mathcal{Y} \cup \gamma_z)$ these points are thick in $\mathcal{T}(V)$ and hence close by (3.34); further they are close in $\mathcal{T}(\beta)$ for $\beta \in \gamma_z$ by construction). Thus the distance from $\Phi_{\mathcal{Y}}(w)$ to $\Phi_{\mathcal{Y}}(z)$ in the $\mathcal{T}(\Sigma\backslash\mathcal{Y})$–factor is equal up to additive error to R_2. □

3.17. **Consistency.** By "undoing" the projection maps π_V, one may use curve complex data of subsurfaces to build points in Teichmüller space. This is accomplished by the work of Behrstock–Kleiner–Minsky–Mosher [BKMM] on consistency.

Definition 3.36 (Consistency)**.** Given a number $\theta \geqslant 1$ and a connected surface Σ, we say that a tuple $(z_V) \in \prod_{V \sqsubset \Sigma} \mathcal{C}(V)$ is *θ–consistent* if the following holds for all pairs of domains $U, V \sqsubset \Sigma$:

$$(1) \quad U \pitchfork V \quad \Longrightarrow \quad \min\left\{d_U(z_U, \partial V),\, d_V(z_V, \partial U)\right\} \leqslant \theta, \text{ and}$$

$$(2) \quad U \sqsubset V \quad \Longrightarrow \quad \min\left\{d_U(z_U, \pi_U(z_V)),\, d_V(z_V, \partial U)\right\} \leqslant \theta.$$

(Observe that if $\pi_U(z_V) = \varnothing$ in (2), then z_V is disjoint from U and ∂U so that $d_V(z_V, \partial U) \leqslant 1 \leqslant \theta$ is automatic).

The following result says that, up to bounded error, the consistent tuples in $\prod_{V \sqsubset \Sigma} \mathcal{C}(V)$ are exactly those obtained by projecting points in the Teichmüller space $\mathcal{T}(\Sigma)$. It was proven for the case of markings as Lemmas 4.1–4.2 and Theorem 4.3 of [BKMM]. However, since every marking $\mu \in \mathcal{M}_0(\Sigma)$ may be realized as the Bers marking of some thick point, the result holds for Teichmüller space as well:

Theorem 3.37 (Consistency and Realization [BKMM])**.** *There is a constant $\mathsf{K} \geqslant 1$ and function $\mathfrak{C}\colon \mathbb{R}_+ \to \mathbb{R}_+$ so that the following holds for every domain $\Sigma \sqsubset S$:*

- *For every $x \in \mathcal{T}(\Sigma)$, the projection tuple $(\pi_V(x))_{V \sqsubset \Sigma}$ is K–consistent.*
- *Conversely, every θ–consistent tuple $(z_V) \in \prod_{V \sqsubset \Sigma} \mathcal{C}(V)$ has a* realization *point $z \in \mathcal{T}(\Sigma)$ with $d_V(\pi_V(z), z_V) \leqslant \mathfrak{C}(\theta)$ for all domains $V \sqsubset \Sigma$. In fact we may assume $z \in \mathcal{N}_{\epsilon_0}(\Sigma)$ is a thick net point.*

Using consistency, one can easily see that the length of a curve $\alpha \in \Gamma(\Sigma)$ at a point $x \in \mathcal{T}(\Sigma)$ is related to the projection distances $d_V(x, \alpha)$ for domains $V \sqsubset \Sigma$:

Lemma 3.38. *For any non-annular domain $\Sigma \sqsubset S$, curve $\alpha \in \Gamma(\Sigma)$, and point $x \in \mathcal{T}_{\epsilon_0}(\Sigma)$, if $d_V(x, \alpha) \leqslant k$ for every domain $V \sqsubset \Sigma$, then $\ell_x(\alpha) \overset{+}{\prec}_k 0$.*

Proof. Let $Y \sqsubset \Sigma$ be a component of $\Sigma\backslash\alpha$. Define a tuple $(z_V) \in \prod_{V \sqsubset Y} \mathcal{C}(V)$ by $z_V = \pi_V(x)$ for each $V \sqsubset Y$. This tuple (z_V) is K–consistent by Theorem 3.37, and hence is realized by a point $x_Y \in \mathcal{T}_{\epsilon_0}(Y)$. Do this for each component of $\Sigma\backslash\alpha$, and choose a point $x_\alpha \in \mathcal{T}(\alpha)$ on the horocycle $y = 1/\epsilon_0$ and with twist parameter so that $\pi_\alpha(x_\alpha) = \pi_\alpha(x)$. These choices define a point in the product $z' \in \mathcal{P}(\Sigma|\alpha)$, and we let $z = \Phi_\alpha^{-1}(z') \in \mathcal{T}(\Sigma)$ be the corresponding point under Minsky's homeomorphism. Note that z is thick by construction and has $\ell_z(\alpha) = \epsilon_0$.

We claim $d_V(x, z) \overset{+}{\prec}_k 0$ for every domain $V \sqsubset \Sigma$. Indeed, if $V \perp \alpha$ then either V is the annulus with core α or $V \sqsubset Y$ for some component Y of $\Sigma\backslash\alpha$; in either case $\pi_V(z) = \pi_V(x_Y)$ coarsely agrees with $\pi_V(x)$ by construction. If instead $V \pitchfork \alpha$, then the fact that α is in every Bers marking at z implies $d_V(z, \alpha) \leqslant \mathsf{L}$. Thus $d_V(x, z) \leqslant d_V(x, \alpha) + d_V(\alpha, z) \leqslant k + \mathsf{L}$. This proves the claim and accordingly bounds $d_{\mathcal{T}(\Sigma)}(x, z)$ by (3.34) since x and z are thick. Since α has length ϵ_0 at z, it follows that $\ell_x(\alpha)$ is bounded. □

Corollary 3.39. *For any $k \geqslant 1$, domain $\Sigma \sqsubset S$, and point $x \in \mathcal{T}_{\epsilon_0}(\Sigma)$, we have*

$$\#\big\{Z \sqsubset \Sigma \mid d_V(x, \partial Z) \leqslant k \text{ for all domains } V \sqsubset \Sigma\big\} \overset{+}{\prec}_k 0$$

Proof. By Lemma 3.38, there is a number L such that if Z satisfies $d_V(x, \partial Z) \leqslant k$ for all $V \sqsubset \Sigma$, then each component α of ∂Z has $\ell_x(\alpha) \leqslant L$. On any hyperbolic surface y, there are only finitely many curves of length at most L. Varying y over a compact fundamental domain for the action of $\mathrm{Mod}(\Sigma)$ on $\mathcal{T}_{\epsilon_0}(\Sigma)$, we obtain a number $F = F(k, \epsilon_0)$ so that every point $x \in \mathcal{T}_{\epsilon_0}(\Sigma)$ has at most F curves of length at most L. Thus the number of subsurfaces $Z \sqsubset \Sigma$ whose boundary curves have length at most L on x is bounded in terms of F. □

4. Preliminaries – antichains, strong alignment, and branch points

4.1. **Antichains.** If Σ is a domain in S and Ω is a collection of subdomains of Σ, we typically write $\underline{\Omega}$ for the set of topologically maximal domains in Ω (that is, maximal with respect to the partial order $\sqsubset$ on Ω). Taking active intervals into account, for each geodesic $[x, y]$ in $\mathcal{T}(\Sigma)$ we may also consider the partial order $\prec_{[x,y]}$ on domains in Σ defined by

$$V \prec_{[x,y]} W \iff V \sqsubset W \text{ and } \mathcal{I}_V \subset \mathcal{I}_W \text{ along } [x, y].$$

We then write $\underline{\Omega}_x^y$ for the set of domains in Ω that are maximal with respect to $\prec_{[x,y]}$. Since the order $\prec_{[x,y]}$ is more restrictive than $\sqsubset$, we note that $\underline{\Omega} \subset \underline{\Omega}_x^y$.

For any collection Ω of domains and given integer i, we additionally write

$$|\Omega|_i = \#\{V \in \Omega \mid \xi(V) = i\}$$

for the number of domains in Ω with complexity i. The following is a variation of Rafi and Schleimer's bound on the cardinality of an antichain [RS, Lemma 5.1]. As the statement we need does not follow directly from the result in [RS], we include a proof in the same spirit as their argument.

Lemma 4.1. *Consider a domain $\Sigma \sqsubset S$ and a sequence of thresholds T_j, for $j \in \{-1, \ldots, \xi(\Sigma) + 1\}$, satisfying $\mathsf{M} \leqslant T_{\xi(\Sigma)+1} \leqslant T_{\xi(\Sigma)} \leqslant \cdots \leqslant T_{-1}$. For any geodesic $[x, y]$ in $\mathcal{T}(\Sigma)$ and any domain $W \sqsubset \Sigma$, the collections $\underline{\mathcal{P}}(W)$ and $\underline{\mathcal{P}}_x^y(W)$ of topologically maximal and $\prec_{[x,y]}$–maximal domains in the set*

$$\mathcal{P}(W) = \{V \sqsubset W \mid d_V(x, y) \geqslant T_{\xi(V)}\}$$

satisfy $|\underline{\mathcal{P}}(W)|_j \leqslant |\underline{\mathcal{P}}_x^y(W)|_j \leqslant (3T_{j+1})^{\xi(W)-j}$ for each $j \in \{-1, \ldots, \xi(\Sigma)\}$.

Proof. Fix some $j \in \{-1, \ldots, \xi(\Sigma)\}$. It suffices to prove the bound on $|\underline{\mathcal{P}}_x^y(W)|_j$. Since $\mathcal{P}(W)$ contains at most one domain (namely W) of complexity $\xi(W)$ or greater, it is clear that $|\underline{\mathcal{P}}_x^y(W)|_j = 0$ when $j > \xi(W)$ and that $|\underline{\mathcal{P}}_x^y(W)|_j \leqslant 1 = (3T_{j+1})^0$ when $j = \xi(W)$. We may therefore assume $j < \xi(\Sigma)$ and restrict to domains $W \sqsubset \Sigma$ with $\xi(W) > j$.

Given any domain $W \sqsubset \Sigma$, write $\underline{\Omega}_x^y(W)$ for the set of $\prec_{[x,y]}$–maximal domains in the collection

$$\Omega(W) = \{V \sqsubset W \mid d_V(x, y) \geqslant T_{j+1}\}.$$

We claim that $\underline{\Omega}_x^y(W)$ contains every domain of $\underline{\mathcal{P}}_x^y(W)$ of complexity j. Indeed, if $V \in \underline{\mathcal{P}}_x^y(W)$ has $\xi(V) = j$, then we must have $V \in \Omega(W)$ since $V \sqsubset W$ and $d_V(x, y) \geqslant T_j \geqslant T_{j+1}$ by definition of $\mathcal{P}(W)$. Thus if $V \notin \underline{\Omega}_x^y(W)$, it must be that $V \prec_{[x,y]} Z$ for some distinct $Z \in \Omega(W)$. This implies $V \subsetneq Z$, and hence $\xi(Z) \geqslant j+1$ and $T_{\xi(Z)} \leqslant T_{j+1}$ by monotonicity of the thresholds. Therefore, the fact

$Z \in \Omega(W)$ gives $Z \sqsubset W$ and $d_Z(x,y) \geqslant T_{j+1} \geqslant T_{\xi(Z)}$. But this implies $Z \in \mathcal{P}(W)$, contradicting the maximality of V in $\mathcal{P}(W)$. This proves $|\underline{\mathcal{P}}_x^y(W)|_j \leqslant |\underline{\Omega}_x^y(W)|_j$. It thus suffices to prove $|\underline{\Omega}_x^y(W)|_j \leqslant (3T_{j+1})^{\xi(W)-j}$ for every geodesic $[x,y]$ and domain $W \sqsubset \Sigma$ with $\xi(W) > j$.

The proof of this proceeds by induction on the complexity $k = \xi(W)$ of the domain W: For each each $k = j, \ldots, \xi(\Sigma)$, we will prove that $|\underline{\Omega}_x^y(W)|_j \leqslant (3T_{j+1})^{k-j}$ for every domain W with $\xi(W) \leqslant k$. We have already seen that this bound is immediate for $k = j$.

Let us therefore fix $k > j$ and assume that $|\underline{\Omega}_x^y(Z)|_j \leqslant (3T_{j+1})^{k-1-j}$ whenever $\xi(Z) < k$. Let $W \sqsubset \Sigma$ be any domain with $\xi(W) \leqslant k$ and choose curves $\alpha \in \pi_W(x)$ and $\beta \in \pi_W(y)$ realizing the distance $d_{\mathcal{C}(W)}(\alpha,\beta) = d_W(x,y)$. Fix also a geodesic $\alpha = \gamma_0, \ldots, \gamma_m = \beta$ joining α to β in $\mathcal{C}(W)$.

We claim that every domain $V \in \Omega(W)$ with $V \neq W$ is disjoint from one of the curves γ_i. This is immediate if α or β is disjoint from V. Otherwise, letting μ_x be a Bers marking on x with $\alpha \in \pi_W(\mu_x)$, Lemma 3.9 implies that $\pi_V(\alpha) \subset \pi_V(\pi_W(\mu_x))$ lies within k of $\pi_V(\mu_x)$. Hence $d_V(\alpha, x) \leqslant \mathsf{k} + 2\mathsf{L}$ and similarly for $d_V(\beta, y)$. Thus

$$d_V(\alpha,\beta) \geqslant d_V(x,y) - 2\mathsf{k} - 4\mathsf{L} \geqslant T_{j+1} \geqslant \mathsf{M} - 2\mathsf{k} - 4\mathsf{L} > \mathsf{Q}$$

by the specification of M in Definition 3.24. We may now invoke the Bounded Geodesic Image Theorem 3.8 to conclude $\pi_V(\gamma_i) = \varnothing$ for some i, as claimed.

We moreover claim that if $V \in \underline{\Omega}_x^y(W) \subset \Omega(W)$ is disjoint from γ_i, then

$$i < \tfrac{1}{2}T_{j+1} \qquad \text{or} \qquad m - i < \tfrac{1}{2}T_{j+1}.$$

Indeed, if this is not the case then necessarily $d_W(x,y) = m \geqslant T_{j+1}$. Therefore $W \in \Omega(W)$ by definition. For any point $v \in \mathcal{I}_V$, the Bers marking μ_v contains ∂V, which is disjoint from γ_i. Hence $d_W(v,\gamma_i) \leqslant \mathsf{L} + 2 \leqslant 3\mathsf{L}$. On the other hand

$$d_W(x,\gamma_i) \geqslant d_W(\alpha,\gamma_i) = i \qquad \text{and} \qquad d_W(y,\gamma_i) \geqslant d_W(\beta,\gamma_i) = m - i.$$

As these quantities are both at least $\frac{1}{2}T_{j+1}$, we therefore see that

$$d_W(x,v), d_W(y,v) \geqslant \tfrac{1}{2}T_{j+1} - 3\mathsf{L} \geqslant \tfrac{1}{2}\mathsf{M} - \tfrac{1}{7}\mathsf{M} > \tfrac{1}{3}\mathsf{M}.$$

But by Lemma 3.26, this is only possible if $v \in \mathcal{I}_W$. Thus we evidently have $\mathcal{I}_V \subset \mathcal{I}_W$, contradicting the $\prec_{[x,y]}$–maximality of V in $\Omega(W)$.

Therefore every $V \in \underline{\Omega}_x^y(W)$ satisfies $V = W$ or else $V \sqsubset Z$ for some Z in the set

$$\mathcal{Z} = \{Z \mid Z \text{ is a component of } W \backslash \gamma_i \text{ for some } i \text{ with } \max\{i, m-i\} < \tfrac{1}{2}T_{j+1}\}.$$

Further, since $\Omega(Z) \subset \Omega(W)$, every $V \sqsubset Z$ that is $\prec_{[x,y]}$–maximal in $\Omega(W)$ is also $\prec_{[x,y]}$–maximal in $\Omega(Z)$. Thus we have

$$\underline{\Omega}_x^y(W) \subset \{W\} \cup \bigcup_{Z \in \mathcal{Z}} \underline{\Omega}_x^y(Z).$$

In particular, $|\underline{\Omega}_x^y(W)|_j \leqslant \sum_{Z \in \mathcal{Z}} |\underline{\Omega}_x^y(Z)|_j \leqslant |\mathcal{Z}|\,(3T_{j+1})^{k-1-j} \leqslant (3T_{j+1})^{k-j}$ by our induction hypothesis and the fact that $|\mathcal{Z}| \leqslant 4(1 + \frac{1}{2}T_{j+1}) \leqslant 3T_{j+1}$. This concludes the induction and the proof of the lemma. □

4.2. **Promoting alignment to strong alignment.** Here we prove Lemma 4.7 and show that aligned tuples may be transformed into a strongly aligned ones by merely adjusting the lengths of certain curves while not affecting curve complex projections. We first show how to find, for each domain V, the ordered points x_i^V along the geodesic $[x_0, x_n]$ required in the Definition 3.21 of strong alignment.

Lemma 4.2. *If* $(x_0,\dots,x_n)$ *is* θ*–aligned in* $\mathcal{T}(\Sigma)$*, then for each domain* $V \sqsubset \Sigma$ *and* $0 \leqslant i \leqslant n$ *there is a point* $w_i \in [x_0, x_n]$ *so that* $d_V(x_i, w_i) \leqslant (\theta + \mathsf{M})/2$*. Clearly we may take* $w_0 = x_0$ *and* $w_n = x_n$*.*

Proof. Let g be any $\mathcal{C}(V)$ geodesic from $\pi_V(x_0)$ to $\pi_V(x_n)$. Lemmas 3.18 and 3.20 show that $\pi_V(x_i)$ lies within $\theta/2 + 4\delta + \mathsf{L}$ of g and that g lies within $\mathsf{B} + 8\delta + 3\mathsf{L}$ of $\pi_V([x_0, x_n])$. Whence there is a point $w_i \in [x_0, x_n]$ so that $d_V(x_i, w_i) \leqslant \theta/2 + \mathsf{B} + 12\delta + 4\mathsf{L} < (\theta + \mathsf{M})/2$. □

Lemma 4.3. *Let* $(x_0,\dots,x_n)$ *be* θ*–aligned in* $\mathcal{T}(\Sigma)$*, let* $V \sqsubset \Sigma$ *be a domain, and let* $w_0,\dots,w_n \in [x_0,x_n]$ *be points such that* $d_V(x_i,w_i) \leqslant T$ *for each* $0 \leqslant i \leqslant n$*. If* σ *is a permutation of* $\{0,\dots,n\}$ *such that the points* $w_{\sigma(0)},\dots,w_{\sigma(n)}$ *appear in order along* $[x_0,x_n]$*, then* $d_V(x_i, w_{\sigma(i)}) \leqslant 4T + \frac{3(\theta+\mathsf{B})}{2}$ *for each* $0 \leqslant i \leqslant n$*.*

Proof. Set $y_i = w_{\sigma(i)}$ for each i. We fix $0 \leqslant j \leqslant n$ and bound $d_V(x_j, y_j)$. By the pigeonhole principle, we may pick some $i \leqslant j$ so that $j \leqslant \sigma^{-1}(i)$. Similarly, there is some $k \geqslant j$ so that $\sigma^{-1}(k) \leqslant j$. Thus the points $w_k = y_{\sigma^{-1}(k)}$, y_j, and $y_{\sigma^{-1}(i)} = w_i$ appear in order along $[x_0, x_n]$. Since w_k appears before w_i when traveling from x_0 to x_n, two applications of Theorem 3.19 then give

$$\begin{aligned} &d_V(x_0,w_k) + 3d_V(w_k,w_i) + d_V(w_i,x_n) \\ &\qquad \leqslant d_V(x_0,w_i) + d_V(w_k,w_i) + d_V(w_k,x_n) + 2\mathsf{B} \\ &\qquad \leqslant d_V(x_0,x_i) + d_V(x_i,x_k) + d_V(x_k,x_n) + 4T + 2\mathsf{B}, \end{aligned}$$

where for the last inequality we have used the assumption that $d_V(x_i,w_i)$ and $d_V(x_k,w_k)$ are both at most T. By θ–alignment and the triangle inequality, the right hand side above at most

$$\begin{aligned} &d_V(x_0,x_n) + 2\theta + 4T + 2\mathsf{B} \\ &\qquad \leqslant d_V(x_0,w_k) + d_V(w_k,w_i) + d_V(w_i,x_n) + 2\theta + 4T + 2\mathsf{B}. \end{aligned}$$

Subtracting the beginning and end of this string of inequalities now gives

$$d_V(w_k,w_i) \leqslant \theta + \mathsf{B} + 2T.$$

On the other hand, using θ–alignment and Theorem 3.19, we have that

$$\begin{aligned} 2d_V(y_j,x_j) &\leqslant \big(d_V(y_j,w_i) + d_V(w_i,x_i) + d_V(x_i,x_j)\big) \\ &\quad + \big(d_V(y_j,w_k) + d_V(w_k,x_k) + d_V(x_k,x_j)\big) \\ &\leqslant \big(d_V(w_k,y_j) + d_V(y_j,w_i)\big) + 2T + \big(d_V(x_i,x_j) + d_V(x_j,x_k)\big) \\ &\leqslant d_V(w_k,w_i) + \mathsf{B} + 2T + d_V(x_i,x_k) + \theta \\ &\leqslant 2d_V(w_k,w_i) + \mathsf{B} + 4T + \theta \end{aligned}$$

Combining this with the previous inequality proves $d_V(y_j,x_j) \leqslant 4T + \frac{3(\theta+\mathsf{B})}{2}$. □

Corollary 4.4. *If* $(x_0,\dots,x_n)$ *is* θ*–aligned in* $\mathcal{T}(\Sigma)$ *and* $V \sqsubset \Sigma$ *is a domain, then there are ordered points* $y_0,\dots,y_n$ *along* $[x_0,x_n]$ *so that* $d_V(x_i,y_i) \leqslant 4(\theta + \mathsf{M})$*.*

We next determine which curves require adjusted lengths. To this end, given a tuple $(x_0,\dots,x_n)$ in $\mathcal{T}(\Sigma)$, let us say an annulus $A \sqsubset \Sigma$ *length-constrains* x_i if the following holds for *both* $y = x_0$ and $y = x_n$:

$$d_A(x_i,y) \geqslant 5(\theta + \mathsf{M}) \quad \text{or} \quad \ell_y(\partial A) < \epsilon_0'.$$

Observe that in this case A has a nonempty active interval $\mathcal{I}_A = \tilde{\mathcal{I}}_A^{\epsilon_0}$ along $[x_0, x_n]$ (c.f. Theorem 3.22 and Definition 3.25). Indeed, if $\ell_{x_0}(\partial A)$ or $\ell_{x_n}(\partial A)$ is less than ϵ_0' this is obvious. Otherwise $d_A(x_i, x_0), d_A(x_i, x_n) \geqslant 5(\theta + \mathsf{M})$ so that alignment implies $d_A(x_0, x_n) \geqslant 10\mathsf{M} + 9\theta$, showing that $\mathcal{I}_A$ is nonempty by Lemma 3.26.

Lemma 4.5. *Let $(x_0, \ldots, x_n)$ be θ–aligned. If annuli $A, B \sqsubset \Sigma$ both length constrain x_i, then the curves $\partial A, \partial B$ are disjoint.*

Proof. By contradiction, suppose $\partial A \pitchfork \partial B$. Since A and B both have active intervals along $[x_0, x_n]$, they are time-ordered and without loss of generality we may assume $A \lessdot B$. Thus $d_A(\partial B, x_n), d_B(x_0, \partial A) \leqslant \mathsf{M}/3$ by Lemma 3.29. Evidently $\ell_{x_n}(\partial A) \geqslant \epsilon_0'$, since $x_n \notin \tilde{\mathcal{I}}_A^{\epsilon_0} = \mathcal{I}_A$; hence the definition of length constraining gives $d_A(x_i, x_n) \geqslant 5(\theta + \mathsf{M})$. Similarly $d_B(x_0, x_i) \geqslant 5(\theta + \mathsf{M})$. It follows that

$$\left.\begin{aligned} d_A(x_i, \partial B) &\geqslant d_A(x_i, x_n) - d_A(\partial B, x_n) \\ d_B(\partial A, x_i) &\geqslant d_B(x_0, x_i) - d_B(\partial A, x_0) \end{aligned}\right\} \geqslant 5(\theta + \mathsf{M}) - \mathsf{M}/3 \geqslant 3\mathsf{M} > \mathsf{K}.$$

But this contradicts Theorem 3.37. □

Lemma 4.5 implies that the core curves of the annuli that length constrain x_i form a (possibly empty) multicurve. The next lemma says this multicurve is close to the Bers marking μ_{x_i} in all subsurfaces.

Lemma 4.6. *Let $(x_0, \ldots, x_n)$ be θ–aligned. If $A \sqsubset \Sigma$ length constrains x_i, then $d_V(x_i, \partial A) \stackrel{+}{\prec}_\theta 0$ for every domain $V \sqsubset \Sigma$.*

Proof. Let $V \sqsubset \Sigma$ be arbitrary. It suffices to suppose ∂A projects to V, for else $d_V(x_i, \partial A) = \mathrm{diam}_{\mathcal{C}(V)}(\pi_V(x_i)) \stackrel{+}{\prec} 0$. Since A has a nonempty thin interval along, we may choose a point $y \in [x_0, x_n]$ such that $\ell_y(\partial A) < \epsilon_0$. Thus ∂A is contained in every Bers marking at y. If $d_V(x_0, x_n) \leqslant \mathsf{M}$, then two applications of the triangle inequality followed by θ–alignment and Theorem 3.19 imply

$$\begin{aligned} 2d_V(\partial A, x_i) \leqslant 2d_V(y, x_i) &\leqslant d_V(x_0, y) + d_V(y, x_n) + d_V(x_0, x_i) + d_V(x_i, x_n) \\ &\leqslant 2d_V(x_0, x_n) + \mathsf{B} + \theta \leqslant 2\mathsf{M} + \mathsf{B} + \theta, \end{aligned}$$

as desired. Hence it remains to suppose $d_V(x_0, x_n) > \mathsf{M}$, which ensures V has a nonempty active interval $\mathcal{I}_V$ along $[x_0, x_n]$.

First suppose $A \pitchfork V$ and, by symmetry, that $A \lessdot V$ along $[x_0, x_n]$. Then evidently $x_n \notin \mathcal{I}_A$ and so (as in Lemma 4.5) the length constraint hypothesis implies $d_A(x_i, x_n) \geqslant 5(\theta + \mathsf{M})$. Time order also gives $d_A(\partial V, x_n) \leqslant \mathsf{M}/3$. Therefore we have $d_A(x_i, \partial V) \geqslant 5(\theta + \mathsf{M}) - \mathsf{M}/3 > \mathsf{K}$. Consistency of the point x_i (Theorem 3.37) now implies the desired bound $d_V(x_i, \partial A) \leqslant \mathsf{K}$.

Otherwise we necessarily have $A \sqsubset V$. Since $\partial A \in \pi_V(w)$ for all $w \in \mathcal{I}_A$, the entire interval projects to a set of diameter at most $2\mathsf{L}$ in $\mathcal{C}(V)$. In particular, we cannot have $\mathcal{I}_A = [x_0, x_n]$, as that would put us in the case $d_V(x_0, x_n) \leqslant 2\mathsf{L} < \mathsf{M}$ dispensed with above. By Lemma 4.2, we may choose a point $z' \in [x_0, x_n]$ with $d_V(x_i, z') \leqslant \theta + \mathsf{M}$. Since $\mathcal{I}_A \neq [x_0, x_n]$ and $\mathcal{I}_A$ projects to a set of diameter at most $2\mathsf{L}$, we may by (3.17) choose some $z \in [x_0, x_n] \setminus \mathcal{I}_A$ with $d_V(z, z') \leqslant 3\mathsf{L}$. The point z either lies before or after $\mathcal{I}_A$, let us suppose it is the former (the opposite possibility being symmetric). Then Lemma 3.26 gives $d_A(x_0, z) \leqslant \mathsf{M}/3$. This also implies $x_0 \notin \tilde{\mathcal{I}}_A^{\epsilon_0} = \mathcal{I}_A$ (since $\mathcal{I}_A$ connected) and hence $d_A(x_i, x_0) \geqslant 5(\theta + \mathsf{M})$ since A length constrains x_i. Combining these we find $d_A(x_i, z) \geqslant 4(\theta + \mathsf{M}) > \mathsf{M}$. By the bounded geodesic image theorem (specifically Corollary 3.27), we now conclude the desired bound

$$d_V(x_i, \partial A) \leqslant d_V(x_i, z) + \mathsf{M}/3 \leqslant d_V(x_i, z') + 3\mathsf{L} + \mathsf{M}/3 \leqslant \theta + 2\mathsf{M} + 3\mathsf{L}. \qquad \square$$

Using these results, we can now transform any aligned tuple into a strongly aligned one by modifying the lengths of core curves of length constraining annuli.

Lemma 4.7. *For any $\theta \geqslant 1$ there exists $\theta' \geqslant \theta$ with the following property: For any θ–aligned tuple $(x_0, \dots, x_n)$ in $\mathcal{T}(\Sigma)$, there is a strongly θ'–aligned tuple $(y_0, \dots, y_n)$ in $\mathcal{N}(\Sigma)$ such that $x_0 = y_0$, $x_n = y_n$, and such that for all $0 \leqslant i \leqslant n$ we have*

$$d_V(x_i, y_i) \stackrel{+}{\prec}_\theta 0 \quad \textit{for every domain } V \sqsubset \Sigma.$$

Moreover, for every annulus $A \sqsubset S$ and $0 < i < n$, we have $\ell_{y_i}(\partial A) \geqslant \epsilon_0$ unless each $y \in \{x_0, x_n\}$ satisfies $d_A(x_i, y) \geqslant 5(\theta + \mathsf{M})$ or $\ell_y(\partial A) < \epsilon_0{}'$.

Proof. We may assume that $\theta \geqslant \mathsf{M}$. Let $A \sqsubset \Sigma$ be any annulus. Let $u^A \in [x_0, x_n]$ be the rightmost point satisfying $d_A(x_0, u^A) \leqslant 18\theta$ and let $v^A \in [x_0, x_n]$ the leftmost point satisfying $d_A(v^A, x_n) \leqslant 18\theta$. Define a (possibly empty) subinterval $J_A \subset [x_0, x_n]$ as follows:

- If $\ell_{x_0}(\partial A), \ell_{x_n}(\partial A) < \epsilon_0'$, set $J_A = [x_0, x_n]$.
- If $\ell_{x_0}(\partial A) < \epsilon_0'$ and $\ell_{x_n}(\partial A) \geqslant \epsilon_0'$, set $J_A = [x_0, v^A]$
- If $\ell_{x_0}(\partial A) \geqslant \epsilon_0'$ and $\ell_{x_n}(\partial A) < \epsilon_0'$, set $J_A = [u^A, x_n]$.
- If $\ell_{x_0}(\partial A), \ell_{x_n}(\partial A) \geqslant \epsilon_0'$, then set $J_A = [u^A, v^A]$ provided that u^A occurs before v^A along $[x_0, x_n]$, and otherwise set $J_A = \varnothing$.

It is easy to see that $J_A \subset \mathcal{I}_A = \tilde{\mathcal{I}}_A^{\epsilon_0}$: This is immediate in the first case above. In the second case, if $v^A = x_0$ it is immediate, and if $v^A \neq x_0$ then necessarily $d_A(v^A, x_n) \geqslant 18\theta - \mathsf{L} > 17\mathsf{M}$ showing that $\mathcal{I}_A$ intersects $[v^A, x_n]$ so that indeed $[x_0, v^A] \subset \mathcal{I}_A$. The third case is similar. In the final case, if $J_A \neq \varnothing$ one similarly finds that $\mathcal{I}_A$ intersects both $[x_0, u^A]$ and $[v^A, x_n]$ so that again $J_A \subset \mathcal{I}_A$.

Notice that if $w \notin J_A$, then either $w \in [x_0, u^A]$ and we necessarily have both $\ell_{x_0}(\partial A) \geqslant \epsilon_0'$ and $d_A(x_0, w) \leqslant d_A(x_0, u^A) + \mathsf{B} \leqslant 19\theta$, or else $w \in [v^A, x_n]$ and we similarly have both $\ell_{x_n}(\partial A) \geqslant \epsilon_0'$ and $d_A(w, x_n) \leqslant 19\theta$. On the other hand, if $w \in J_A$ (which recall is contained in $\mathcal{I}_A$), then for both $z = x_0$ and $z = x_n$ we have

$$\ell_w(\partial A) < \epsilon_0 \quad \text{and} \quad \Big[d_A(z, w) \geqslant 18\theta \quad \text{or} \quad \ell_z(\partial A) < \epsilon_0'\Big]. \tag{4.8}$$

Let $x_0 = x_0^A, x_1^A, \dots, x_n^A = x_n$ be the ordered sequence of points along $[x_0, x_n]$ provided by Corollary 4.4 satisfying $d_A(x_i, x_i^A) \leqslant 8\theta$ for each i. Next define new points y_i^A as follows: If $x_i^A \in J_A$, then set $y_i^A = x_i^A$. If instead $x_i^A \notin J_A$, then define $y_i^A = x_0$ provided $x_i^A \in [x_0, u^A]$; if this fails, then necessarily $x_i^A \in [v^A, x_n]$ and we set $y_i^A = x_n$. With these definitions, we note that $x_0 = y_0^A, y_1^A, \dots, y_n^A = x_n$ appear in order along $[x_0, x_n]$. Furthermore, if $y_i^A \neq x_i^A$, then either $y_i^A = x_0$ with $x_i^A \in [x_0, u^A]$ or $y_i^A = x_n$ with $x_i^A \in [v^A, x_n]$. Hence the previous paragraph implies

$$d_A(x_i^A, y_i^A) \leqslant 19\theta \quad \text{for all } 0 \leqslant i \leqslant n.$$

Claim 4.9. *If $y_i^A \in J_A$ then A length constrains x_i.*

Proof. If $y_i^A \in J_A$, then by definition $y_i^A = x_i^A$ and hence $d_A(x_i, y_i^A) \leqslant 8\theta$. Since (4.8) holds for $w = y_i^A$, for both $z = x_0$ and $z = x_n$ we conclude that either $d_A(z, x_i) \geqslant 18\theta - 8\theta \geqslant 5(\theta + \mathsf{M})$ or $\ell_z(\partial A) < \epsilon_0'$. Thus A length constrains x_i. $\square$

For $0 < i < n$, fix a Bers marking μ_i at x_i, and let α_i be the set of core curves of annuli $A \sqsubset \Sigma$ that length constrain x_i. Lemma 4.5 implies that α_i is a multicurve in Σ, and Lemma 4.6 implies that $d_V(\mu_i, \alpha_i) \stackrel{+}{\prec}_\theta 0$ for every $V \sqsubset \Sigma$. Lemma 3.10 therefore provides a new marking ν_i of Σ such that $\alpha_i \subset \operatorname{base}(\nu_i)$ and $d_V(\mu_i, \nu_i) \stackrel{+}{\prec}_\theta 0$ for every domain $V \sqsubset \Sigma$.

We now construct points in $\mathcal{T}(\Sigma)$ by picking lengths for the curves of $\operatorname{base}(\nu_i)$. Specifically, use Fenchel–Nielsen coordinates to build a point $y_i \in \mathcal{T}(\Sigma)$ such that ν_i is a Bers marking at y_i and such that for each curve $\gamma \in \operatorname{base}(\nu_i)$: if $\gamma = \partial A$ for an annulus $A \sqsubset \Sigma$ with $y_i^A \in J_A$ then declare $\ell_{y_i}(\gamma) = \ell_{y_i^A}(\partial A)$, and if not then declare $\ell_{y_i}(\gamma) = \epsilon_0$. Notice that every curve $\gamma \in \operatorname{base}(\nu_i)$ satisfies $\ell_{y_i}(\gamma) \leqslant \epsilon_0$; thus by the Margulis lemma, every $\gamma \notin \operatorname{base}(\nu_i)$ satisfies $\ell_{y_i}(\gamma) \geqslant \epsilon_0$. Note that $y_1, \ldots, y_{n-1}$ immediately satisfy the final "moreover" conclusion of the lemma since the only potentially short curves on y_i are cores of annuli that length-constrain x_i. To complete the notation, also set $y_0 = x_0$ and $y_n = x_n$.

Notice that *every* annulus $A \sqsubset \Sigma$ now satisfies

$$\frac{\epsilon_0'}{\epsilon_0} \leqslant \frac{\min\{\epsilon_0, \ell_{y_i}(\partial A)\}}{\min\{\epsilon_0, \ell_{y_i^A}(\partial A)\}} \leqslant \frac{\epsilon_0}{\epsilon_0'}.$$

Indeed, the claim is immediate for $i \in \{0, n\}$, and for $0 < i < n$ we consider two cases: First suppose $y_i^A \notin J_A$. Then by construction $\ell_{y_i^A}(\partial A) \geqslant \epsilon_0'$ and regardless of whether or not $\partial A \in \operatorname{base}(\nu_i)$ we also have $\ell_{y_i}(\partial A) \geqslant \epsilon_0$. Next suppose $y_i^A \in J_A$. Then A constrains x_i by Claim 4.9; hence $\partial A \in \alpha_i \subset \operatorname{base}(\nu_i)$ by construction and consequently $\ell_{y_i}(\partial A) = \ell_{y_i^A}(\partial A)$ by fiat.

It remains to show the new tuple $(y_0, \ldots, y_n)$ satisfies the conclusion of the lemma. For every domain $V \sqsubset \Sigma$ we have

$$d_V(x_i, y_i) \leqslant d_V(x_i, \mu_i) + d_V(\mu_i, \nu_i) + d_V(\nu_i, y_i) \stackrel{+}{\prec}_\theta \mathsf{L} + 0 + \mathsf{L}$$

as required. Since $(x_0, \ldots, x_n)$ is θ–aligned, this also proves $(y_0, \ldots, y_n)$ is θ'–aligned for some θ' depending only on θ. Furthermore, for every annulus $A \sqsubset \Sigma$ we have

$$d_A(y_i, y_i^A) \leqslant d_A(y_i, x_i) + d_A(x_i, x_i^A) + d_A(x_i^A, y_i^A) \stackrel{+}{\prec}_\theta 0 + 8\theta + 19\theta.$$

This, together with the previous paragraph, shows that, after increasing θ' if necessary, both bullets of Definition 3.21 are satisfied for annuli. Finally, Corollary 4.4 ensures the first bullet also holds for nonannuli. $\square$

4.3. **Branch points.** Fix some domain $\Sigma \sqsubset S$. For any triple of points $y, x, z \in \mathcal{T}(\Sigma)$ and domain $V \sqsubset \Sigma$, hyperbolicity of $\mathcal{C}(V)$ implies that geodesics from $\pi_V(x)$ to $\pi_V(y)$ and $\pi_V(z)$ fellow travel for distance roughly equal to the Gromov product

$$(y|z)_x^V = \frac{1}{2}\left(d_V(x,y) + d_V(x,z) - d_V(y,z)\right) \leqslant \min\{d_V(x,y), d_V(x,z)\}. \tag{4.10}$$

More precisely, since $\pi_V(*)$ always has diameter at most L, it is a basic exercise in hyperbolic geometry (using, e.g. [ABC+, Proposition 2.1]) to prove the following:

Lemma 4.11. *If γ_y and γ_z are any $\mathcal{C}(V)$–geodesics from $\pi_V(x)$ to $\pi_V(y)$ and $\pi_V(z)$, then each $p \in \gamma_y$ with $d_V(\pi_V(x), p) \leqslant (y|z)_x^V + c$ lies within $8\delta + 4\mathsf{L} + c$ of γ_z.* $\square$

Hyperbolicity of $\mathcal{C}(V)$ (again, see [ABC+]) additionally implies that

$$(y|z)_x^V \geqslant \min\{(y|w)_x^V, (w|z)_x^V\} - 5\delta - 3\mathsf{L} \quad \text{for all } x, y, z, w \in \mathcal{T}(\Sigma), V \sqsubset \Sigma. \tag{4.12}$$

These observations allows us to perform the following construction to any collection of geodesics in $\mathcal{T}(\Sigma)$ with a common endpoint:

Lemma 4.13. *Consider any points $x, y_1, \ldots, y_n \in \mathcal{T}(\Sigma)$. For each domain $V \sqsubset S$, there is a "branch point" $\zeta_V \in \mathcal{C}(V)$ so that ζ_V lies within $8(\delta + \mathsf{L})$ of any geodesic from $\pi_V(x)$ to $\pi_V(y_i)$. Further, if $(y_j|y_k)^V_x \leqslant \min_{l,m}(y_l|y_m)^V_x + c$, then ζ_V lies within $24(\delta + \mathsf{L}) + c$ of any geodesic from $\pi_V(y_j)$ to $\pi_V(y_k)$.*

Proof. For $\beta, \gamma, \mu \in \mathcal{C}(V)$ let us write $2(\beta|\gamma)_\mu = d_V(\mu, \beta) + d_V(\mu, \gamma) - d_V(\beta, \gamma)$. Fix curves $\alpha \in \pi_V(x)$ and $\alpha_i \in \pi_V(y_i)$ for each $i = 1, \ldots n$. Now let $G = \min(y_l|y_m)^V_x$ and fix indices l, m achieving this minimum. Let $\zeta_V \in \mathcal{C}(V)$ be the point on a geodesic from α to α_l with

$$d(\alpha, \zeta_V) = (\alpha_l|\alpha_m)_\alpha \leqslant (y_l|y_m)^V_x + \mathsf{L} = G + \mathsf{L}.$$

Therefore $d_V(x, \zeta_V) - 2\mathsf{L} \leqslant G$ is smaller than any Gromov product $(y_l|y_i)^V_x$ and so Lemma 4.11 implies ζ_V lies within $8\delta + 6\mathsf{L}$ of any geodesic from $\pi_V(x)$ to $\pi_V(y_i)$.

Now suppose $(y_j|y_k)^V_x \leqslant G + c$ and, by the above, pick a point β on a geodesic from α to α_j with $d_V(\zeta_V, \beta) \leqslant 8\delta + 6\mathsf{L}$. Then

$$d_V(x, \beta) + 8\delta + 7\mathsf{L} \geqslant d_V(\alpha, \zeta_V) = (\alpha_l|\alpha_m)_\alpha \geqslant G - 2\mathsf{L} \geqslant (y_j|y_k)^V_x - 2\mathsf{L} - c.$$

Since $d_V(x, y_j) = (y_j|y_k)^V_x + (x|y_k)^V_{y_j}$ and β lies on a geodesic from $\pi_V(x)$ to $\pi_V(y_j)$, the above implies that

$$d_V(y_j, \beta) \leqslant d_V(x, y_j) - d_V(x, \beta) + 2\mathsf{L} \leqslant (x|y_k)^V_{y_j} + 8\delta + 11\mathsf{L} + c.$$

Thus Lemma 4.11 implies β lies within $16\delta + 15\mathsf{L} + c$ of any geodesic from $\pi_V(y_j)$ to $\pi_V(y_k)$. Since $d_V(\zeta_V, \beta) \leqslant 8\delta + 6\mathsf{L}$, the claim follows. ∎

The next step is to show that the tuple (ζ_V) of branch points is consistent:

Lemma 4.14. *The tuple $(\zeta_V) \in \prod_{V \sqsubset \Sigma} \mathcal{C}(V)$ is $7\mathsf{M}$–consistent.*

Proof. Let us call V "large" if $d_V(x, y_i) \geqslant \mathsf{M}$ for all $i = 1, \ldots, n$ and call V "small" otherwise. Note that $d_V(x, \zeta_V) \leqslant 2\mathsf{M}$ whenever V is small; this is because ζ_V lies within $8(\delta + \mathsf{L}) \leqslant \mathsf{M}$ of any geodesic from $\pi_V(x)$ to any $\pi_V(y_i)$ and therefore satisfies $d_V(x, \zeta_V) \leqslant d_V(x, y_i) + \mathsf{M}$ for some $i = 1, \ldots, n$.

Fix two domains $V, W \sqsubset \Sigma$. We must establish the inequalities in Definition 3.36 for the constant $7\mathsf{M}$. If V and W are both small, then $d_V(x, \zeta_V), d_W(x, \zeta_W) \leqslant 2\mathsf{M}$ and the claim follows from the fact that $(\pi_Z(x))_{Z \sqsubset \Sigma}$ is K–consistent (Theorem 3.37). Hence we may assume one of V, W is large.

First suppose $W \pitchfork V$ with V and W both large, then they are time-ordered along each geodesic $[x, y_i]$. The characterization (Lemma 3.29) implies that $W \lessdot V$ along $[x, y_i]$ iff $d_V(x, \partial W) < \mathsf{M}/3$. Hence we may suppose $W \lessdot V$ along each geodesic $[x, y_i]$ (the alternate possibility $V \lessdot W$ along each $[x, y_i]$ being symmetric). Now time-ordering implies $d_W(y_i, \partial V) \leqslant \mathsf{M}/3$ for each i. The fact that ζ_W lies within M of some geodesic from $\pi_W(y_j)$ to $\pi_W(y_k)$, which evidently has length at most $2\mathsf{M}/3$, thus implies $d_W(\zeta_W, y_j) \leqslant 2\mathsf{M}$ for some j. Therefore $d_W(\zeta_W, \partial V) \leqslant d_W(\zeta_W, y_j) + d_W(y_j, \partial V) \leqslant 3\mathsf{M}$ and we are done in this case.

Next suppose that V is small and W large with ∂W projecting to $\mathcal{C}(V)$. In this case $d_V(x, \zeta_V) \leqslant 2\mathsf{M}$ and we may pick i so that $d_V(x, y_i) < \mathsf{M}$. Since $d_W(x, y_i) \geqslant \mathsf{M}$ mean, Corollary 3.27 implies that

$$d_V(\zeta_V, \partial W) \leqslant 2\mathsf{M} + d_V(x, \partial W) + d_V(\partial W, y_i) \leqslant d_V(x, y_i) + 3\mathsf{M} < 4\mathsf{M}.$$

This establishes consistency when $W \sqsubset V$ with V small and W large, and (by symmetry) when $W \pitchfork V$ with at least one domain large.

Finally suppose $W \sqsubset V$ with V large. If $d_V(\partial W, \zeta_V) \leqslant 7\mathsf{M}$ we satisfy consistency, so it suffices to assume $d_V(\partial W, \zeta_V) > 7\mathsf{M}$. Fix curves $\alpha \in \pi_V(x)$ and $\alpha_i \in \pi_V(y_i)$ that each project to W (which we can do by Lemma 3.9). For each i, additionally fix a geodesic g_i in $\mathcal{C}(V)$ from α to α_i and a point $\beta_i \in g_i$ with $d_V(\zeta_V, \beta_i) \leqslant \mathsf{M}$. Fixing indices j, k realizing $\min(y_j|y_k)^V_x$, we also take a geodesic g from α_j to α_k that contains a curve β with $d_V(\zeta_V, \beta) \leqslant \mathsf{M}$. Note that every curve within $6\mathsf{M}$ of ζ_V cuts W (since ∂W is too far away); hence all curves within $5\mathsf{M}$ of β or any β_i also cut W. In particular, Theorem 3.8 gives $d_W(\zeta_V, \beta_i), d_W(\zeta_V, \beta) \leqslant \mathsf{Q}$.

We consider two subcases: Firstly suppose $d_W(\zeta_V, \alpha) > 2\mathsf{Q}$. Then for each i we have $d_W(\beta_i, \alpha) > \mathsf{Q}$, which by Theorem 3.8 implies that some curve along g_i between α and β_i is disjoint from W. This means every curve along g_i from β_i to α_i cuts W; indeed, the curves missing W have diameter 2 in in $\mathcal{C}(V)$ and all lie distance at least $5\mathsf{M}$ from β_i, thus such curves cannot occur along g_i both between α and β_i and between β_i and α_i. Therefore the Bounded Geodesic Image Theorem gives $d_W(\beta_i, \alpha_i) \leqslant \mathsf{Q}$. Since $d_W(\alpha_i, y_i) \leqslant \mathsf{k} + \mathsf{L}$ by Lemma 3.9, we conclude

$$d_W(y_i, \zeta_V) \leqslant d_W(y_i, \alpha_i) + d_W(\alpha_i, \beta_i) + d_W(\beta_i, \zeta_V) \leqslant 2\mathsf{Q} + \mathsf{k} + \mathsf{L} < \mathsf{M}$$

for each i. The fact that ζ_W lies within M of some geodesic from $\pi_W(y_m)$ to $\pi_W(y_l)$, and that this geodesic evidently has length at most $2\mathsf{M}$, implies $d_W(\zeta_W, y_m) \leqslant 3\mathsf{M}$. Therefore the triangle inequality gives $d_W(\zeta_W, \zeta_V) \leqslant 4\mathsf{M}$ as needed.

The final subcase is $d_W(\zeta_V, \alpha) \leqslant 2\mathsf{Q}$. First observe that either $d_W(\alpha_j, \beta) \leqslant \mathsf{Q}$ or $d_W(\alpha_k, \beta) \leqslant \mathsf{Q}$, since otherwise the Bounded Geodesic Image Theorem would imply that g has curves missing W both between α_j and β and between β and α_k. By symmetry let us suppose $d_W(\alpha_j, \beta) \leqslant \mathsf{Q}$ so that the triangle inequality gives $d_W(\alpha, \alpha_j) \leqslant 4\mathsf{Q}$. Since $d_W(\alpha, x), d_W(\alpha_j, y_j) \leqslant \mathsf{k} + \mathsf{L}$ by Lemma 3.9, this gives $d_W(x, y_j) \leqslant \mathsf{M}$ and implies $d_W(\zeta_W, x) \leqslant \min_i d_W(x, y_i) + \mathsf{M} \leqslant 2\mathsf{M}$. Therefore

$$d_W(\zeta_W, \zeta_V) \leqslant d_W(\zeta_W, x) + d_W(x, \alpha) + d_W(\alpha, \zeta_V) \leqslant 2\mathsf{M} + \mathsf{k} + \mathsf{L} + 2\mathsf{Q} \leqslant 3\mathsf{M},$$

which concludes the proof of the Lemma. □

Lemma 4.15 (Barycenters)**.** *There is a constant $\mathfrak{B} \geqslant 2\mathsf{M}$ such that for any domain $\Sigma \sqsubset S$, every ordered triple $y, x, z \in \mathcal{T}(\Sigma)$ has a* barycenter $b \in \mathcal{N}_{\epsilon_0}(\Sigma)$ *so that (x, b, y), (y, b, z), (z, b, x) are each $\mathfrak{B}$–aligned. Additionally, for any $\theta \geqslant 2\mathsf{M}$ there exist* annular-split barycenter *$y', z' \in \mathcal{N}(\Sigma)$ such that (y, y', x), (x, z', z), and (y, y', z', z) are each $\mathfrak{B}$–aligned, and so that for every domain $V \sqsubset \Sigma$:*

- *If V is an annulus and $(y|z)^V_x > \theta$, then*

$$d_V(y, y') \leqslant \mathfrak{B} \quad \text{and} \quad d_V(z', z) \leqslant \mathfrak{B}.$$

- *Otherwise* $\operatorname{diam}_{\mathcal{C}(V)} \pi_V(\{b, y', z'\}) \leqslant \mathfrak{B}$ *with both (y, z', x) and (x, y', z) $\mathfrak{B}$–aligned in V.*

Thus y' and z' coarsely agree with the branch point b in all domains, except for certain annuli for which y' and z' instead agree with y and z.

Proof. Let $(\beta_V) \in \prod_{V \sqsubset \Sigma} \mathcal{C}(V)$ be the $7\mathsf{M}$–consistent branch tuple from Lemmas 4.13–4.14. Theorem 3.37 then gives $b \in \mathcal{N}_{\epsilon_0}(\Sigma)$ so that $d_V(b, \beta_V) \leqslant \mathfrak{C}(7\mathsf{M})$ for every $V \sqsubset \Sigma$. Since $(y|z)^V_x \leqslant \min\{(y|y)^V_x, (z|z)^V_x\} + \mathsf{M}/2$ by (4.12), Lemma 4.13 ensures that β_V lies within M of any geodesic joining a pair of $\pi_V(x)$, $\pi_V(y)$ and

$\pi_V(z)$. These observations imply the triples (x,b,y), (x,b,z), and (y,b,z) are, for example, each $2(\mathfrak{C}(7\mathsf{M})+\mathsf{M})$–aligned.

Now define $(\xi_V),(\zeta_V) \in \prod_{V\sqsubset\Sigma}\mathcal{C}(V)$ so that $\xi_V=\zeta_V=\beta_V$ except for annuli A with $(y|z)_x^A > \theta$, in which case instead set $\xi_A=\pi_A(y)$ and $\zeta_A=\pi_A(z)$.

Claim 4.16. *(ξ_V) and (ζ_V) are R–consistent, for $R = 2\mathfrak{C}(7\mathsf{M})+11\mathsf{M}$.*

Proof. We prove the claim for (ξ_V), as the proof for (ζ_V) is symmetric. Since (β_V) is $7\mathsf{M}$–consistent, it suffices to check consistency for pairs A,V involving a domain A with $d_A(\xi_A,\beta_A) > 4\mathsf{M}$. In this case A must be an annulus with $(y|z)_x^A > \theta$. Since β_A lies within M of any geodesic from $\pi_A(y)$ to $\pi_A(z)$, we note that

$$d_A(\xi_A,\beta_A) = d_A(y,\beta_A) \leqslant d_A(y,\beta_A)+d_A(\beta_A,z) \leqslant d_A(y,z)+2\mathsf{M}.$$

Thus $d_A(y,z) > 2\mathsf{M}$, and additionally $\min\{d_A(y,x),d_A(x,z)\} > (y|z)_x^A > \theta \geqslant 2\mathsf{M}$ by (4.10). It follows that at least two of the quantities $d_A(y,b)$, $d_A(x,b)$, $d_A(z,b)$ must be larger than M, since otherwise the triangle inequality would bound the minimum of $d_A(y,x)$, $d_A(x,z)$, $d_A(y,z)$ by $2\mathsf{M}$. Without loss of generality, let suppose $d_A(y,b), d_A(z,b) > \mathsf{M}$, in which case Corollary 3.27 implies

$$\begin{aligned} d_V(y,z)+2d_V(b,\partial A) &\leqslant d_V(y,\partial A)+d_V(\partial A,b)+d_V(b,\partial A)+d_V(\partial A,z)\\ &\leqslant d_V(y,b)+d_V(b,z)+2\mathsf{M}/3\\ &\leqslant d_V(y,z)+2(\mathfrak{C}(7\mathsf{M})+\mathsf{M})+2\mathsf{M}/3. \end{aligned}$$

Thus $d_V(\beta_V,\partial A) \leqslant d_V(b,\partial A)+\mathfrak{C}(7\mathsf{M}) \leqslant 2(\mathfrak{C}(7\mathsf{M})+\mathsf{M})$. If $d_V(\xi_V,\beta_V)\leqslant 4\mathsf{M}$, this bounds $d_V(\xi_V,\partial A)$ and proves the claimed consistency. Otherwise $d_V(\xi_V,\beta_V) > 4\mathsf{M}$ which again means V is an annulus with $\xi_V=\pi_V(y)$. Therefore

$$\min\{d_A(\xi_A,\partial V),d_V(\xi_V,\partial A)\} = \min\{d_A(\pi_A(y),\partial V),d_V(\pi_V(y),\partial A)\} \leqslant \mathsf{K} \leqslant \mathsf{M}$$

by Theorem 3.37, and the required inequality is satisfied. ☐

Let $y',z' \in \mathcal{N}_{\epsilon_0}(\Sigma)$ be the thick net points provided by Theorem 3.37 realizing the R–consistent tuples $(\xi_V),(\zeta_V)$. Then for annuli A with $(y|z)_x^A > \theta$ the claim implies $d_A(y,y'),d_A(z,z') \leqslant \mathfrak{C}(R)$ which immediately gives $2\mathfrak{C}(R)$–alignment of (y,y',z',z), (y,y',x) and (x,z',z) in A. For all other domains V, $d_V(y',\beta_V)$ and $d_V(\beta_V,z')$ are at most $\mathfrak{C}(R)$, which gives $d_V(y',z') \leqslant 2\mathfrak{C}(R)$. Since β_V is a M–barycenter of $\{\pi_V(x),\pi_V(y),\pi_V(z)\}$, it also implies the three tuples above and (y,z',x), (x,y',z) are all $4(\mathfrak{C}(R)+\mathsf{M})$–aligned in V, as desired. ☐

Lemma 4.17. *The constant $\mathfrak{B}$ and net points $y',z' \in \mathcal{N}(\Sigma)$ from Lemma 4.15 can moreover be chosen so that (y,y',z',z) is strongly $\mathfrak{B}$–aligned. Further, if $y,z \in \mathcal{T}_{\epsilon_0}(\Sigma)$, then any annulus A with $d_A(y',z') > \mathfrak{B}$ satisfies $\ell_{y'}(\partial A),\ell_{z'}(\partial A) \geqslant \epsilon_0$.*

Proof. Let $\mathfrak{B}$ and $y_0,z_0 \in \mathcal{N}(S)$ be provided by Corollary 4.15 so that (y,y_0,z_0,z) is $\mathfrak{B}$–aligned. Next apply Lemma 4.7 to obtain a new constant $\mathfrak{B}' > \mathfrak{B}$ and net points $y_0',z_0' \in \mathcal{N}(S)$ such that (y,y_0',z_0',z) is strongly $\mathfrak{B}'$–aligned and so that $\max\{d_V(y_0,y_0'),d_V(z_0,z_0')\} \leqslant \mathfrak{B}'$ for all domains $V \sqsubset S$. This latter property implies these points y_0',z_0' additionally satisfy all the conclusions of Lemma 4.15 with the bound $\mathfrak{B}$ replaced by $(\mathfrak{B}+2\mathfrak{B}')$.

If a domain $V \sqsubset S$ now satisfies $d_V(y_0',z_0') > (\mathfrak{B}+2\mathfrak{B}')$, then it must be that V is an annulus with $(y|z)_x^V > \theta$ and, consequently,

$$\max\{d_V(y,y_0),d_V(z_0,z)\} \leqslant \mathfrak{B} < 5(\mathfrak{B}+\mathsf{M})$$

by construction in Lemma 4.15. Hence, if y, z are thick, the construction in Lemma 4.7 ensures we necessarily have $\ell_{y_0'}(\partial V) \geqslant \epsilon_0$ and $\ell_{z_0'}(\partial A) \geqslant \epsilon_0$. □

5. Finding good points

We now commence with the proof of the upper bound in Theorem 1.2. Since there are only finitely many conjugacy classes of finite-order elements in $\mathrm{Mod}(S)$, it suffices to perform the count for each conjugacy class separately:

Convention 5.1. We henceforth fix a finite-order element $\phi_0 \in \mathrm{Mod}(S)$ and a point $x_0 \in \mathcal{T}(S)$ such that $\phi_0(x_0) = x_0$; the existence of such a point was proven by Nielsen [Nie]. Let $m_0 \geqslant 2$ be the order of ϕ_0. Since there are only finitely many conjugacy classes of finite-order elements, we note that m_0 is universally bounded depending only on S and may furthermore suppose ϵ_0 is chosen so that $x_0 \in \mathcal{T}_{\epsilon_0}(S)$

Our first objective is to find fixed points for elements of the conjugacy class $[\phi_0]$ that enjoy certain nice properties. To begin, let x'_ϕ be any fixed point for $\phi \in [\phi_0]$; for example, if $\phi = f\phi_0 f^{-1}$ we may take $x'_\phi = f(x_0)$. Now apply Lemmas 4.13–4.14 to the points $x'_\phi, \phi(x_0), \dots, \phi^{m_0}(x_0)$ to get a $7\mathsf{M}$–consistent branch tuple (ζ_V) and corresponding thick point $w \in \mathcal{T}_{\epsilon_0}(S)$ provided by Theorem 3.37. Since ϕ fixes x'_ϕ and the set $\{\phi(x_0), \dots, \phi^{m_0}(x_0)\}$, it follows that $\phi(w)$ and w *both* coarsely satisfy the branch condition of Lemma 4.13 for the list $x'_\phi, \phi(x_0), \dots, \phi^{m_0}(x_0)$. In particular, for any domain $V \sqsubset S$, if indices j, k are chosen to achieve $\min_{j,k}(\phi^j(x_0)|\phi^k(x_0))^V_{x'_\phi}$, then for $y = w$ and $y = \phi(w)$ the triples

$$(x'_\phi, y, \phi^j(x_0)), \quad (\phi^j(x_0), y, \phi^k(x_0)), \quad \text{and} \quad (\phi^k(x_0), y, x'_\phi)$$

are each $(2\mathfrak{C}(7\mathsf{M}) + \mathsf{M})$–aligned in V. By Lemma 3.18 it follows that if $\triangle$ is a $\mathcal{C}(V)$ geodesic triangle with vertices in $\pi_V(x'_\phi)$, $\pi_V(\phi^j(x_0))$ and $\pi_V(\phi^k(x_0))$, then $\pi_V(w)$ and $\pi_V(\phi(w))$ both lie within $\mathfrak{C}(7\mathsf{M}) + \mathsf{M}$ of each side of $\triangle$. The set of such points has uniformly bounded diameter, hence we conclude $d_V(w, \phi(w)) \overset{+}{\prec} 0$. Since $w, \phi(w)$ are thick, the distance formula (3.34) now implies that $d_{\mathcal{T}(S)}(w, \phi(w)) \overset{+}{\prec} 0$.

Definition 5.2 (Good fixed point)**.** Apply Durham's result [Dur, Theorem 1.3] to the point w to obtain a fixed point x_ϕ for ϕ with $d_{\mathcal{T}(S)}(w, x_\phi) \overset{+}{\prec} 0$. Since w is uniformly thick, we may again adjust ϵ_0 if necessary so that $x_\phi \in \mathcal{T}_{\epsilon_0}(S)$.

By definition of the Gromov product, for each domain $V \sqsubset S$ we have

$$d_V(x_0, x_\phi) + d_V(x_\phi, \phi(x_0)) = d_V(x_0, \phi(x_0)) + 2(x_0|\phi(x_0))^V_{x_\phi}.$$

We thus view V as "backtracking" for ϕ if $(x_0|\phi(x_0))^V_{x_\phi}$ is large, since in this case $(x_0, x_\phi, \phi(x_0))$ is poorly aligned in V, and the $\mathcal{C}(V)$–geodesics from x_0 to x_ϕ and then to $\phi(x_0)$ fellow travel for a large distance. While in general it may be impossible to eliminate backtracking entirely, as in the example described in §2, our good fixed point x_ϕ minimizes it in the sense that there cannot be backtracking in the full orbit $\{\phi^i(V) \mid i \in \mathbb{Z}\}$ of any domain.

Lemma 5.3. *There exists $\Lambda \geqslant 2\mathsf{M}$, depending only on S, such that for every domain $V \sqsubset S$ we have that $(x_0|\phi(x_0))^{\phi^i(V)}_{x_\phi} \leqslant \Lambda$ for some $i \in \mathbb{Z}$.*

Proof. Let $G = \min_{j,k}(\phi^j(x_0)|\phi^k(x_0))^V_{x'_\phi}$ and choose $1 \leqslant j, k \leqslant m_0$ realizing this minimum. Observe that if $(\phi^i(x_0)|\phi^{i+1}(x_0))^V_{x'_\phi} > G'$ for all $i \in \mathbb{Z}$, then $|k - j| \leqslant m_0$

applications of (4.12) would imply $(\phi^j(x_0)|\phi^k(x_0))^V_{x'_\phi} > G' - m_0(5\delta + 3\mathsf{L})$; hence there exists i so that $(\phi^i(x_0)|\phi^{i+1}(x_0))^V_{x'_\phi} \leqslant G + m_0\mathsf{M}/2$. By its construction in Lemma 4.13, the branch point ζ_V thus lies within $24(\delta+\mathsf{L})+m_0\mathsf{M}/2 \leqslant m_0\mathsf{M}$ of any geodesic from $\pi_V(\phi^i(x_0))$ to $\pi_V(\phi^{i+1}(x_0))$. Since $d_V(w,\zeta_V) \leqslant \mathfrak{C}(7\mathsf{M})$, it follows that $(\phi^i(x_0), w, \phi^{i+1}(x_0))$ is $2(\mathfrak{C}(7\mathsf{M}) + m_0\mathsf{M})$–aligned in V, which is equivalent to saying $(\phi^i(x_0)|\phi^{i+1}(x_0))^V_w \leqslant \mathfrak{C}(7\mathsf{M}) + m_0\mathsf{M}$. Since x_ϕ is fixed, the uniform bound $d_V(w, x_\phi) \overset{+}{\prec} 0$ now implies $(x_0|\phi(x_0))^{\phi^{-i}(V)}_{x_\phi} = (\phi^i(x_0)|\phi^{i+1}(x_0))^V_{x_\phi} \overset{+}{\prec} 0$. □

Next apply Lemma 4.17 with constant $\theta = \Lambda$ to obtain the annular-split barycenters $a_\phi, b_\phi \in \mathcal{N}(S)$ for the ordered triple $(x_0, x_\phi, \phi(x_0))$. We summarize the key features of this construction as follows. To streamline notation, we define

$$R^\phi_V := d_V(x_\phi, b_\phi) \qquad \text{for each domain } V \sqsubset S \text{ and element } \phi \in [\phi_0]. \tag{5.4}$$

Proposition 5.5 (Good point properties)**.** *There is a constant $\Theta \geqslant 9\mathsf{M}$, depending only on S, such that for each $\phi \in [\phi_0]$ there exist points $x_\phi, a_\phi, b_\phi \in \mathcal{T}(S)$ such that*

(1) $x_\phi \in \mathcal{T}_{\epsilon_0}(S)$ is fixed by ϕ, and $a_\phi, b_\phi \in \mathcal{N}(S)$ are net points;
(2) the tuple $(x_0, a_\phi, b_\phi, \phi(x_0))$ is strongly Θ–aligned;
(3) each tuple (x_0, a_ϕ, x_ϕ) and $(x_\phi, b_\phi, \phi(x_0))$ is Θ–aligned;
(4) unless $V \sqsubset S$ is an annulus with $(x_0|\phi(x_0))^V_{x_\phi} > \Lambda$, we have $d_V(a_\phi, b_\phi) \leqslant \Theta$ and (x_0, b_ϕ, x_ϕ), $(x_\phi, a_\phi, \phi(x_0))$ are Θ–aligned in V;
(5) if $d_V(a_\phi, b_\phi) > \Theta$, then V is an annulus and $\ell_{a_\phi}(\partial V), \ell_{b_\phi}(\partial V) \geqslant \epsilon_0$;
(6) for each $V \sqsubset S$ there exists $j \in \mathbb{Z}$ so that $R^\phi_{\phi^j(V)} \leqslant \Theta$;
(7) if $R^\phi_{\phi(V)} \geqslant 7R^\phi_V + 7\Theta$, then $d_V(x_0, b_\phi) \geqslant 6\Theta > \mathsf{M}$;
(8) for any annulus $V \sqsubset S$ with $R^\phi_V > \Theta$, we have $d_V(b_\phi, \phi(x_0)) \leqslant \Theta$;
(9) for any nonannular $V \sqsubset S$, we have $R^\phi_V \overset{+}{\prec}_\Theta m_0\mathsf{L}d_{\mathcal{T}(S)}(x_0, \phi(x_0))$.

Proof. We take $\Theta = 4\Lambda + 4\mathfrak{B} + 9\mathsf{M}$. Then items (1)–(5) are immediate from Definition 5.2 and the construction of a_ϕ, b_ϕ in Lemmas 4.15–4.17.

For (6), Lemma 5.3 provides some $j \in \mathbb{Z}$ so that $(x_0|\phi(x_0))^{\phi^j(V)}_{x_\phi} \leqslant \Lambda$. Therefore the construction in Lemmas 4.15–4.17 implies that (x_ϕ, b_ϕ, x_0), $(x_\phi, b_\phi, \phi(x_0))$ and $(x_0, b_\phi, \phi(x_0))$ are all $\mathfrak{B}$–aligned in the domain $V' = \phi^j(V)$. Thus

$$\begin{aligned} 2R^\phi_{V'} &= \big(d_{V'}(x_\phi, b_\phi) + d_{V'}(b_\phi, x_0)\big) + \big(d_{V'}(x_\phi, b_\phi) + d_{V'}(b_\phi, \phi(x_0))\big) \\ &\qquad - \big(d_{V'}(x_0, b_\phi) + d_{V'}(b_\phi, \phi(x_0))\big) \\ &\leqslant d_{V'}(x_\phi, x_0) + \mathfrak{B} + d_{V'}(x_\phi, \phi(x_0)) + \mathfrak{B} - d_{V'}(x_0, \phi(x_0)) \\ &= 2(x_0|\phi(x_0))^{V'}_{x_\phi} + 2\mathfrak{B} \leqslant 2\Lambda + 2\mathfrak{B} \leqslant 2\Theta. \end{aligned}$$

For (7), since $(x_\phi, b_\phi, \phi(x_0))$ is Θ–aligned and x_ϕ is fixed, the hypothesis implies

$$d_V(x_0, x_\phi) = d_{\phi(V)}(\phi(x_0), x_\phi) \geqslant d_{\phi(V)}(b_\phi, x_\phi) - \Theta \geqslant 7d_V(x_\phi, b_\phi) + 6\Theta.$$

Thus by the triangle inequality we have

$$d_V(x_0, b_\phi) \geqslant d_V(x_0, x_\phi) - d_V(b_\phi, x_\phi) \geqslant 6\Theta > \mathsf{M}.$$

For (8), we consider an annulus V with $d_V(x_\phi, b_\phi) = R_V^\phi > \Theta$. If $d_V(a_\phi, b_\phi) \leqslant \mathfrak{B}$, then $\mathfrak{B}$–alignment of (x_ϕ, a_ϕ, x_0) and $(x_\phi, b_\phi, \phi(x_0))$ implies

$$\begin{aligned}
2(x_0|\phi(x_0))^V_{x_\phi} &= d_V(x_0, x_\phi) + d_V(x_\phi, \phi(x_0)) - d_V(x_0, \phi(x_0)) \\
&\geqslant d_V(x_0, a_\phi) + d_V(a_\phi, x_\phi) + d_V(x_\phi, b_\phi) + d_V(b_\phi, \phi(x_0)) - 2\mathfrak{B} \\
&\quad - (d_V(x_0, a_\phi) + d_V(a_\phi, b_\phi) + d_V(b_\phi, \phi(x_0))) \\
&= d_V(a_\phi, x_\phi) + d_V(x_\phi, b_\phi) - d_V(a_\phi, b_\phi) - 2\mathfrak{B} \\
&\geqslant 2d_V(b_\phi, x_\phi) - 2d_V(a_\phi, b_\phi) - 2\mathfrak{B} > 2(\Theta - 2\mathfrak{B}) > 2\Lambda.
\end{aligned}$$

Otherwise we evidently have $d_V(a_\phi, b_\phi) > \mathfrak{B}$. In either case, the construction in Lemmas 4.15–4.17 implies that $d_V(b_\phi, \phi(x_0)) \leqslant \mathfrak{B} \leqslant \Theta$.

Finally for (9), recall that by construction $(x_0, b_\phi, \phi(x_0))$, (x_ϕ, b_ϕ, x_0) and $(x_\phi, b_\phi, \phi(x_0))$ are each $\mathfrak{B}$–aligned in all nonannular domains. Hence for all nonannular Y we have

$$\begin{aligned}
R^\phi_{\phi(Y)} = d_{\phi(Y)}(x_\phi, b_\phi) &\leqslant d_{\phi(Y)}(x_\phi, \phi(x_0)) + \mathfrak{B} = d_Y(x_0, x_\phi) + \mathfrak{B} \\
&\leqslant d_Y(x_0, b_\phi) + d_Y(b_\phi, x_\phi) + \mathfrak{B} \\
&\leqslant d_Y(x_0, \phi(x_0)) + R^\phi_Y + 2\mathfrak{B}.
\end{aligned}$$

Now fix a nonannulus V and let $0 \leqslant j < m_0$ be the smallest integer so that $d_{\phi^{-j}(V)}(x_\phi, b_\phi) \leqslant \Theta$, which necessarily exists by (6). Applying the above estimate recursively with $Y = \phi^{-1}(V), \ldots, \phi^{-j}(V)$ we find

$$R^\phi_V \leqslant \left(\sum_{n=1}^{j} d_{\phi^{-n}(V)}(x_0, \phi(x_0)) + 2\mathfrak{B}\right) + \Theta \stackrel{+}{\prec}_\Theta \sum_{i=1}^{m_0} d_{\phi^i(V)}(x_0, \phi(x_0)).$$

Since V is nonannular, the Lipschitz bound (3.16) implies each term $d_{\phi^i(V)}(x_0, \phi(x_0))$ is at most $\mathsf{L} d_{\mathcal{T}(S)}(x_0, \phi(x_0)) + \mathsf{L}$. Thus this estimate gives the desired bound. $\square$

We remark that the failure of the Lipschitz estimate (3.16) for annuli in (9) is the entire reason we have utilized the adjusted barycenters a_ϕ, b_ϕ from Lemma 4.15 and the alternate conclusion (8).

6. Bounding the multiplicity of branch points

Recall that we have fixed an order $m_0 < \infty$ element $\phi_0 \in \mathrm{Mod}(S)$ and point $x_0 \in \mathcal{T}_{\epsilon_0}(S)$ with $\phi_0(x_0) = x_0$. For each $\phi \in [\phi_0]$, we have produced in Proposition 5.5 a fixed point x_ϕ for ϕ along with net points a_ϕ, b_ϕ satisfying various properties.

In this section we bound the multiplicity of any given pair $(a, b) \in \mathcal{N}(S)$:

Theorem 6.1. *There is a polynomial p such that that for any ordered pair $(a, b) \in \mathcal{N}(S)$ and $r \geqslant 0$, there are at most $p(r)$ finite order elements ϕ of $[\phi_0]$ for which $d_{\mathcal{T}(S)}(x_0, \phi(x_0)) \leqslant r$ and $(a_\phi, b_\phi) = (a, b)$.*

6.1. **Agreement.** We shall prove this by using curve complex data to effectively build-up a map ϕ on larger and larger subsurfaces. More precisely, we will bound the indicated subset of $[\phi_0]$ by partitioning into smaller and smaller subfamilies that agree on larger and larger subsurfaces.

We begin by establishing some general topological statements that will be useful.

Definition 6.2. We say $\phi, \psi \in \mathrm{Mod}(S)$ *agree* in a subsurface $A \sqsubset S$ if for each component Y of A we have $\phi(Y) = \psi(Y)$ and $\phi|_Y = \psi|_Y$ up to isotopy. So in particular $\phi(A) = \psi(A)$ as subsurfaces and, after adjusting say ψ by an isotopy,

$\psi^{-1}\phi$ pointwise fixes A and its boundary ∂A. In the case of an annular component Y, we additionally require

Remark 6.3. When Y is an annulus with $\phi|_Y = \psi_Y$ up to isotopy, we may always precompose with Dehn twists $T^n_{\partial Y}$ in Y so that ψ and $\phi \circ T^n_{\partial Y}$ agree in Y. The key point is that given any two curves α, β cutting Y we can choose $n \in \mathbb{N}$ so that $d_Y(\alpha, T^n_Y(\beta)) \leqslant 1$. Now, fix some curve α with $\alpha \pitchfork Y$ and choose n so that $d_Y(\alpha, T^n_Y \psi^{-1}\phi(\alpha)) \leqslant 1$. We claim that $d_Y(\beta, \psi^{-1}\phi T^n_Y(\beta)) \leqslant 3$ for every curve β. , so that ψ and ϕT^n_Y agree on Y. Indeed, choose $k \in \mathbb{N}$ so that $d_Y(T^k_A(\beta), \alpha_0) \leqslant 1$. Then, since T_Y and $\psi^{-1}\phi$ commute, we have

$$\begin{aligned} d_Y(\beta, \psi^{-1}\phi T^n_A(\beta)) &= d_Y(T^k_Y(\beta), T^k_Y\psi^{-1}\phi T^n_Y(\beta)) = d_Y(T^k_Y(\beta), T^n_Y\psi^{-1}\phi(T^k_Y(\beta))) \\ &\leqslant d_Y(T^k_Y(\beta), \alpha_0) + d_Y(\alpha_0, T^n_Y\psi^{-1}\phi(\alpha_0)) \\ &\quad + d_Y(T^n_Y\psi^{-1}\phi(\alpha_0), T^n_Y\psi^{-1}\phi(T^k_Y(\beta)) \leqslant 1 + 1 + 1 = 3. \end{aligned}$$

Annular components in A will be used to ensure mapping classes do not differ by Dehn twists about boundary components. For example, if $Y \sqsubset S$ is a torus with one boundary component, then agreement on Y conveys no information about twisting in Y; indeed the maps $\phi \circ T^n_{\partial Y}$ for $n \in \mathbb{Z}$ all agree on Y. If we let A be be the union of Y with a disjoint annulus parallel to ∂Y, then agreement on A additionally concerns Dehn twists about ∂Y so that the maps $\phi \circ T^n_{\partial Y}$, $n \in \mathbb{Z}$, no longer all agree in A.

Lemma 6.4. *For each $k > 0$ there exists a constant k' with the following property: For a given point $w \in \mathcal{T}(S)$ and pair of subsurfaces $A, B \sqsubset S$, let $\mathcal{F} \subset \mathrm{Mod}(S)$ be a set of mapping classes such that for all $\phi, \psi \in \mathcal{F}$ we have $\phi(A) = B$ and $d_W(\phi(w), \psi(w)) \leqslant k$ for each domain $W \sqsubset B$. Then, up to agreement on A, the collection $\{\phi|_A \colon A \to B \mid \phi \in \mathcal{F}\}$ has cardinality at most k'. That is, $\mathcal{F}$ may be partitioned into at most k' subfamilies in which all maps agree on A.*

Proof. There is a universal bound, depending on $\xi(S)$, on the number of components of A and the number of boundary components of each component. Hence, after partitioning $\mathcal{F}$ into boundedly many subfamilies, we may assume $\psi^{-1}\phi$ fixes up to isotopy each component and boundary component of A for all $\phi, \psi \in \mathcal{F}$. To prove the lemma, we must partition $\mathcal{F}$ to achieve agreement in each component Y of A.

First suppose Y is non-annular. Fix some $\phi \in \mathcal{F}$. Then for all $\psi \in \mathcal{F}$, the projections $\pi_V(\psi^{-1}\phi(w))$ and $\pi_V(w)$ coarsely agree in $\mathcal{C}(V)$ for each domain $V \sqsubset Y$. If $w' \in \mathcal{T}_{\epsilon_0}(Y)$ is a thick point realizing the consistent tuple $(\pi_V(w))_{V \sqsubset Y}$, it follows that $d_V(\psi^{-1}\phi(w'), w') \stackrel{+}{\prec}_k 0$ for all such V and hence that $d_{\mathcal{T}(Y)}(\psi^{-1}\phi(w'), w')$ is bounded by the distance formula. By proper discontinuity of the action on $\mathcal{T}(Y)$, there are boundedly many mapping classes $Y \to Y$ that coarsely fix w'.

Now suppose Y is an annulus. Fix $\phi \in \mathcal{F}$, set $V = \phi(Y)$, and let $\beta \in \mu_w$ be a curve in the short marking with $\beta \pitchfork Y$. The assumption implies that each $\psi \in \mathcal{F}$ satisfies $d_V(\phi(\beta), \psi(\beta)) \leqslant k$ and hence that there exists $n \in \mathbb{Z}$ with $|n| \leqslant k$ so that $d_V(T^n_V\psi(\beta), \phi(\beta)) \leqslant 1$. Letting $\mathcal{F}_n \subset \mathcal{F}$ denote those ψ that work with a given power n, we thus get a decomposition $\mathcal{F} = \mathcal{F}_{-k} \cup \cdots \cup \mathcal{F}_k$ into at most $2k + 1$ subsets. By the triangle inequality, all $\psi, \psi' \in \mathcal{F}_n$ satisfy $d_V(\psi(\beta), \psi'(\beta)) = d_V(T^n_V\psi(\beta), T^n_V\psi'(\beta)) \leqslant 2$. Now, for any other curve α, we may pick j so that

$d_Y(T_Y^j(\alpha),\beta) \leqslant 1$. Note that $\phi' T_Y^j = T_V^j \phi'$ for all $\phi' \in \mathcal{F}$. Then all $\psi, \psi' \in \mathcal{F}_n$ have

$$\begin{aligned} d_V(\psi(\alpha),\psi'(\alpha)) &= d_V(T_V^j\psi(\alpha), T_V^j\psi'(\alpha)) = d_V(\psi T_Y^j(\alpha), \psi' T_V^j(\alpha)) \\ &= d_V(\psi T_Y^j(\alpha), \psi(\beta)) + d_V(\psi(\beta),\psi'(\beta)) + d_V(\psi'(\beta), \psi' T_V^j(\alpha)) \\ &\leqslant 1+2+1 = 4. \end{aligned}$$

Hence we have partitioned into boundedly many subsets that each agree on Y. $\square$

6.2. **Backtracking domains.** Recall the streamlined notation $R_V^\phi = d_V(x_\phi, b_\phi)$ from (5.4). We view R_V^ϕ as a measure of backtracking in V since, aside from the exceptional annuli in Proposition 5.5(4), it coarsely agrees with the Gromov product $(x_0|\phi(x_0))_{x_\phi}^V$ and measures alignment of the triple $(x_0, x_\phi, \phi(x_0))$ in V.

Let $f\colon \mathbb{R}_+ \to \mathbb{R}_+$ be the function defined by

$$f(t) = 7t + 7\Theta$$

where $\Theta \geqslant 9\mathsf{M}$ is from Proposition 5.5. We say that a sequence $V, \phi(V), \ldots, \phi^j(V)$ *jumps* for $\phi \in [\phi_0]$ if for all $0 \leqslant i \leqslant j$

$$R^\phi_{\phi^i(V)} > f\left(R^\phi_{\phi^{-1}(V)}\right).$$

Remark 6.5. If $0 < i < j$ are such that $V, \ldots, \phi^i(V)$ and $\phi^{i+1}(V), \ldots, \phi^j(V)$ are both jump sequences, then the concatenation $V, \ldots, \phi^j(V)$ is also a jump sequence since for all $i+1 \leqslant k \leqslant j$ the monotonicity of f evidently implies

$$R^\phi_{\phi^k(V)} \geqslant f\big(R^\phi_{\phi^i(V)}\big) \geqslant f^2\big(R^\phi_{\phi^{-1}(V)}\big).$$

Definition 6.6. The set of *backtracking domains* $D(\phi)$ for $\phi \in [\phi_0]$ is the union of all jump sequences for ϕ, that is, the set of domains $\phi^i(V)$ for which there is a sequence $\{V, \phi(V), \ldots, \phi^i(V), \ldots, \phi^j(V)\}$ that jumps for ϕ. Remark 6.5 and Proposition 5.5(6) imply that for every $Z \in D(\phi)$ there exists $0 \leqslant i < m_0$ such that $\{\phi^{-i}(Z), \ldots, Z\}$ is a jump sequence contained in $D(\phi)$ and $\phi^{-i-1}(Z) \notin D(\phi)$. We call this i the backtracking *index* of Z in $D(\phi)$.

While each $Y \in D(\phi)$ satisfies $R_Y^\phi > 7\Theta$ by definition, the converse need not hold since there may exist domains Y with $R_Y^\phi > 7\Theta$ but whose orbit $\{\phi^i(Y) \mid i \in \mathbb{Z}\}$ does not contain a jump sequence. Nevertheless, we have:

Lemma 6.7. *Every domain $Z \in D(\phi)$ satisfies $R_Z^\phi = d_Z(x_\phi, b_\phi) > 7\Theta$. Dually, the collection $D(\phi)$ contains every domain $Z \sqsubset S$ with $R_Z^\phi > f^{m_0-1}(\Theta)$.*

Proof. The first claim follows immediately from the definition of jumping. We now suppose $Z \notin D(\phi)$ and show $R_Z^\phi \leqslant f^{m_0-1}(\Theta)$. If $R_Z^\phi \leqslant \Theta$ there is nothing to prove, so we assume $R_Z^\phi > \Theta$. Let us write $Z_k = \phi^k(Z)$ for $k \in \mathbb{Z}$. We may choose an integer $-m_0 < k < 0$ so that $R^\phi_{Z_k} \leqslant \Theta$ and $R^\phi_{Z_j} > \Theta$ for all $k < j \leqslant 0$.

For any $k \leqslant j < 0$, the sequence $\{Z_{j+1}, \ldots, Z_0\}$ evidently does not jump for ϕ; hence there is some $j < j' \leqslant 0$ so that $R^\phi_{Z_{j'}} \leqslant f(R^\phi_{Z_j})$. Starting with $k_0 = k$ and recursively using this observation to set $k_{j+1} = k_j'$ produces a sequence integers $k = k_0 < \cdots < k_n = 0$ such that $R^\phi_{Z_{k_{j+1}}} \leqslant f(R^\phi_{Z_{k_j}})$ for each j. Since $R^\phi_{Z_{k_0}} \leqslant \Theta$ and $n \leqslant k < m_0$, applying these inequalities inductively implies that

$$R_Z^\phi = R^\phi_{Z_{k_n}} \leqslant f^{\circ n}(R^\phi_{Z_{k_0}})) \leqslant f^{m_0-1}(\Theta). \qquad \square$$

For each $Y \in D(\phi)$, let $i \geqslant 0$ be the index of Y and define a constant C_Y by

$$C_Y := 6R^{\phi}_{\phi^{-i-1}(Y)} + 3\Theta + 2\mathsf{M}.$$

Also set $C_0 = f^{m_0}(\Theta)$. Since $\phi^{-i-1}(Y) \notin D(\phi)$ by definition of the index, we note that $9\mathsf{M} \leqslant \Theta \leqslant C_Y \leqslant C_0 - 2\mathsf{M}$ by Lemma 6.7 and the definition of f. In particular, C_Y is uniformly bounded.

Definition 6.8 (Orders on $D(\phi)$)**.** For $\phi \in [\phi_0]$ we define four asymmetric relations $\lhd_i$ on the set $D(\phi)$ of backtracking domains as follows. For $V, Y \in D(\phi)$,

- $V \lhd_0 Y$ if there exists $j \geqslant 1$ so that $\{V, \phi(V), \dots, \phi^j(V) = Y\} \subset D(\phi)$.
- $V \lhd_1 Y$ if $V \pitchfork Y$ and $V \lessdot Y$ along $[b_\phi, x_\phi]$.
- $V \lhd_2 Y$ if $V \subsetneq Y$ with $d_Y(b_\phi, \partial V) < C_Y$.
- $V \lhd_{\tilde{2}} Y$ if $V \subsetneq Y$ with $d_Y(b_\phi, \partial V) < C_Y + \mathsf{M}$.
- $V \lhd_3 Y$ if $V \supsetneq Y$ with $d_V(b_\phi, \partial Y) \geqslant C_0 + 2\mathsf{M}$.

Thus $\lhd_{\tilde{2}}$ is a weaker version of $\lhd_2$; it will serve a minor technical role. Notice that $\lhd_0$ and $\lhd_1$ are non-reflexive partial orders; in particular they are transitive.

In general, for any subcollection $\mathcal{W} \subset D(\phi)$ and $i \in \{0, 1, 2, \tilde{2}, 3\}$, we write $\mathcal{W}^i$ for the set of domains in $\mathcal{W}$ that are minimal with respect to the order $\lhd_i$, that is:

$$\mathcal{W}^i = \{Z \in \mathcal{W} \mid \nexists Y \in \mathcal{W} \text{ with } Y \lhd_i Z\}.$$

We also write $\mathcal{W}^\dagger = \mathcal{W}^1 \cap \mathcal{W}^2 \cap \mathcal{W}^3$ and $\mathcal{W}^\star = \mathcal{W}^0 \cap \mathcal{W}^\dagger$.

Lemma 6.9. *Let $\mathcal{W}$ be any subcollection of $D(\phi)$. If $U \in \mathcal{W}^1$ is such that the subcollection $\mathcal{U} = \{Z \in \mathcal{W} \mid Z \subsetneq U\}$ is nonempty, then $\mathcal{W}^1 \supset \mathcal{U}^1 \neq \varnothing$.*

Proof. Since $\mathcal{U}$ is nonempty and finite and $\lhd_1$ is transitive, $\mathcal{U}^1$ is nonempty. Now consider some $Y \in \mathcal{U}^1$ and take any $Z \in \mathcal{W}$ with $Z \pitchfork Y$. Then either $Z \subsetneq U$ and we have $Y \lessdot Z$ by virtue of $Y \in \mathcal{U}^1$, or else $Z \pitchfork U$ and hence $U \lessdot Z$, since $U \in \mathcal{W}^1$, and consequently $Y \lessdot Z$ by Corollary 3.31. Thus $Y \in \mathcal{W}^1$ as claimed. ☐

Lemma 6.10. *Let $\mathcal{W}$ be a subcollection of $D(\phi)$. If $V \in \mathcal{W}^1 \backslash \mathcal{W}^3$, then there exists some $Y \in \mathcal{W}^\dagger$ with $Y \supsetneq V$.*

Proof. The assumption $V \notin \mathcal{W}^3$ implies that

$$\Omega = \left\{Z \in \mathcal{W} \mid Z \supsetneq V \text{ and } d_Z(b_\phi, \partial V) \geqslant C_0 + \mathsf{M}\right\}$$

is nonempty. Hence we may choose a domain $Y \in \Omega$ maximizing the quantity $\xi(Y)$. We note that $d_Y(b_\phi, \partial V) \geqslant C_0 + \mathsf{M}$. We claim $Y \in \mathcal{W}^\dagger$. To see this, we consider any $Z \in \mathcal{W}$ and show that $Z \lhd_i Y$ fails for each of $i = 1, 2, 3$.

First consider the cases $Z \pitchfork Y$ and $Z \subsetneq Y$. Then the the multicurve ∂Z projects to $\mathcal{C}(Y)$. If ∂Z is disjoint from ∂V, then we have $d_Y(\partial Z, \partial V) \leqslant 2$ and hence $d_Y(b_\phi, \partial Z) \geqslant d_Y(b_\phi, \partial V) - 2 > C_0 + \mathsf{M} - 2 > C_Y$. When $Z \pitchfork Y$, this ensures $Y \lessdot Z$, and when $Z \subsetneq Y$ it precludes $Z \lhd_2 Y$. If, instead, ∂Z and ∂V are not disjoint, then $Z \pitchfork V$ and hence $V \lessdot Z$ by the assumption that $V \in \mathcal{W}^1$. If $Z \pitchfork Y$ this implies $Y \lessdot Z$ by Corollary 3.31, and if $Z \subsetneq Y$ it implies via Corollary 3.32 that

$$d_Y(b_\phi, \partial Z) \geqslant d_Y(b_\phi, \partial V) - \mathsf{M}/3 \geqslant C_0 + 2\mathsf{M}/3 > C_Y.$$

In either case, we may conclude that $Y \in \mathcal{W}^1 \cap \mathcal{W}^2$.

It remains to suppose $Z \supsetneq Y$. Since by construction Y is the largest complexity surface in Ω, we must have $Z \notin \Omega$. Hence $d_Z(b_\phi, \partial V) < C_0 + \mathsf{M}$. On the other hand, the containment $V \subsetneq Y$ gives $d_Z(b_\phi, \partial Y) \leqslant d_Z(b_\phi, \partial V) + 1 < C_0 + 2\mathsf{M}$ which precludes $Z \lhd_3 Y$ and consequently shows $Y \in \mathcal{W}^3$. ☐

Lemma 6.11. *Let $\mathcal{W} \subset D(\phi)$. If $V \in \mathcal{W}^1$, then there exists $Y \in \mathcal{W}^\dagger$ such that either $Y \sqsupset V$ or else $Y \subsetneq V$ with $Y \lhd_{\tilde{2}} V$.*

Proof. If $V \notin \mathcal{W}^3$, then Lemma 6.10 provides a domain satisfying the conclusion. We may henceforth assume $V \in \mathcal{W}^3$. If $V \in \mathcal{W}^2$, then $V \in \mathcal{W}^\dagger$ and the claim holds. Otherwise $V \notin \mathcal{W}^2$ and we have $V' \lhd_2 V$ for some $V' \in \mathcal{W}$. In particular,

$$\mathcal{V} = \{Z \in \mathcal{W} \mid Z \subsetneq V\}$$

is nonempty. The subcollection $\mathcal{V}^1$ is then nonempty, and we may choose a domain $Y' \in \mathcal{V}^1$ minimizing the quantity $\xi(Y')$. Firstly observe that $Y' \in \mathcal{W}^1$ by Lemma 6.9. Secondly, observe that $Y' \in \mathcal{W}^2$. Indeed, otherwise

$$\mathcal{Y} = \{Z \in \mathcal{W} \mid Z \subsetneq Y'\} = \{Z \in \mathcal{V} \mid Z \subsetneq Y'\}$$

is nonempty and hence contains some element $Z \in \mathcal{Y}^1$. But then $Z \in \mathcal{Y}^1 \subset \mathcal{V}^1$ by Lemma 6.9 with $\xi(Z) < \xi(Y')$, contradicting the choice of Y'.

Next observe that some $Y \sqsupset Y'$ satisfies $Y \in \mathcal{W}^\dagger$. Indeed, if $Y' \in \mathcal{W}^3$ we take $Y = Y'$ and if not then Lemma 6.10 provides such a Y. The domains Y and V cannot be disjoint, since they both contain Y', and nor can we have $Y \pitchfork V$, as then they could not both be minimal with respect to time order. If Y satisfies $Y \sqsupset V$, then the lemma is verified. The only remaining possibility is $Y \subsetneq V$, in which case we must show $Y \lhd_{\tilde{2}} V$. To see this, recall that we have $V' \lhd_2 V$. If the multicurves ∂Y and $\partial V'$ are disjoint, it follows that $d_V(b_\phi, \partial Y) < d_V(b_\phi, \partial V') + 1$. Otherwise $Y \pitchfork V'$ and we must have $Y \lessdot V'$ along $[b_\phi, x_\phi]$, by virtue of Y lying in $\mathcal{W}^1$; thus $d_V(b_\phi, \partial Y) \leqslant d_V(b_\phi, \partial V') + \mathsf{M}/3$ by Corollary 3.32. In either case, we conclude

$$d_V(b_\phi, \partial Y) \leqslant d_V(b_\phi, \partial V') + \mathsf{M}/3 < C_V + \mathsf{M},$$

which shows that $Y \lhd_{\tilde{2}} V$, as desired. □

We also have the following observation relating $\lhd_1$ and $\lhd_{\tilde{2}}$ to $\lhd_0$:

Lemma 6.12. *Suppose $Y, Z \in D(\phi)$ and that $Y \lhd_1 Z$ or $Y \lhd_{\tilde{2}} Z$. If a chain $\phi^{-j}(Y), \dots, Y$ is contained in $D(\phi)$, then the corresponding chain $\phi^{-j}(Z), \dots, Z$ is also contained in $D(\phi)$. That is, $\phi^{-j}(Y) \lhd_0 Y$ implies $\phi^{-j}(Z) \lhd_0 Z$.*

Proof. Suppose not. Let $i \geqslant 0$ be the index of Z, so that $\{\phi^{-i}(Z), \dots, Z\} \subset D(\phi)$ is a jump sequence but $\phi^{-i-1}(Z) \notin D(\phi)$. Set $Y' = \phi^{-i-1}(Y)$ and $Z' = \phi^{-i-1}(Z)$. The assumption implies $i < j$, so we have $Y' \in D(\phi)$ but $Z' \notin D(\phi)$.

We know that ∂Y projects to $\mathcal{C}(Z)$ and hence that

$$d_Z(b_\phi, x_\phi) \leqslant d_Z(b_\phi, \partial Y) + d_Z(\partial Y, x_\phi).$$

Now if $Y \lhd_1 Z$, then $Y \lessdot Z$ along $[b_\phi, x_\phi]$ and hence $d_Z(b_\phi, \partial Y) < \mathsf{M}/3$ by Lemma 3.29. If instead $Y \lhd_{\tilde{2}} Z$ then $d_Z(b_\phi, \partial Y) < C_Z + \mathsf{M}$ by assumption. In either case we have

$$R_Z^\phi = d_Z(b_\phi, x_\phi) < d_Z(\partial Y, x_\phi) + C_Z + \mathsf{M}. \tag{6.13}$$

We similarly know that $\partial Y'$ projects to $\mathcal{C}(Z')$. Since $Y' \in D(\phi)$, by Lemma 6.7, $d_{Y'}(x_\phi, b_\phi) \geqslant 7\Theta \geqslant \mathsf{M}$ and so Y' has an active interval along $[x_\phi, b_\phi]$. Thus we may choose a point $t \in [x_\phi, b_\phi]$ containing $\partial Y'$ in its Bers makings and apply Theorem 3.19 to conclude that

$$\begin{aligned} d_Z(\partial Y, x_\phi) = d_{Z'}(\partial Y', x_\phi) &\leqslant d_{Z'}(x_\phi, \partial Y') + d_{Z'}(\partial Y', b_\phi) \\ &\leqslant d_{Z'}(x_\phi, t) + d_{Z'}(t, b_\phi) \leqslant d_{Z'}(x_\phi, b_\phi) + \mathsf{B} = R_{Z'}^\phi + \mathsf{B}. \end{aligned}$$

Combining with (6.13) and using $Z' = \phi^{-i-1}(Z) \notin D(\phi)$, we now conclude

$$\begin{aligned} R^\phi_Z &< d_Z(\partial Y, x_\phi) + C_Z + \mathsf{M} \leqslant R^\phi_{Z'} + 2\mathsf{M} + C_Z \\ &= R^\phi_{Z'} + 2\mathsf{M} + 6R^\phi_{Z'} + 3\Theta + 2\mathsf{M} \\ &< f\left(R^\phi_{\phi^{-i-1}(Z)}\right). \end{aligned}$$

But this exactly means $\{\phi^{-i}(Z), \dots, Z\}$ does NOT jump for ϕ, a contradiction. □

6.3. **Initial domains for compatible subsurfaces.** In the spirit of "building up" our maps on larger and larger subsurfaces, for $\phi \in [\phi_0]$ and a possibly empty subsurface $A \sqsubset S$, we write

$$D_A(\phi) = \{Y \in D(\phi) \mid Y \not\sqsubset \phi(A)\}$$

for the backtracking domains whose preimages do not land in A. We emphasize that if $Y \in D(\phi)$ is an annulus, then $Y \sqsubset \phi(A)$ if and only if either Y is isotopic to an annular component of $\phi(A)$ or else $Y \subsetneq V$ for some component V of $\phi(A)$. Notice that for the empty subsurface we have $D_\varnothing(\phi) = D(\phi)$ and for the whole surface $A = S$ we have $D_S(\phi) = \varnothing$. We view $D_A(\phi)$ as the backtracking domains that we still need to account for once we "know" ϕ on A.

We use the notation $D^i_A(\phi) = (D_A(\phi))^i$ for minimal elements as in Definition 6.8.

Definition 6.14 (Initial domains)**.** Given $\phi \in [\phi_0]$, we say a domain $V \sqsubset S$ is *ϕ–initial* for a subsurface $A \sqsubset S$ if $V \in D^\star_A(\phi) = \bigcap_{i=0}^3 D^i_A(\phi)$. Note that we do not require minimality with respect to $\lhd_{\tilde{2}}$.

We will consider subsurfaces A that are constructed by successively adding initial domains, as follows:

Definition 6.15 (Compatible)**.** A subsurface $A \sqsubset S$ is *compatible* with $\phi \in [\phi_0]$ if either $A = S$ or else $A = \phi^{-1}(Z_0 \sqcup \dots \sqcup Z_m)$ for some sequence $\varnothing = Z_0, Z_1, \dots, Z_m$ in which each Z_{i+1}, with $0 \leqslant i < m$, is a ϕ–initial domain for $A_i = \phi^{-1}(Z_0 \sqcup \dots \sqcup Z_i)$ (Recall from Lemma 3.3 that $Z_0 \sqcup \dots \sqcup Z_i$ is the subsurface filled by $Z_0, \dots, Z_i$).

The next two lemmas explain how subsets $D_A(\phi)$ and relations $\lhd_i$ interact:

Lemma 6.16. *If $A \sqsubset S$ is compatible with $\phi \in [\phi_0]$ and domains $Y, Z \in D(\phi)$ satisfy $Y \lhd_1 Z$ and $Y \in D_A(\phi)$, then $Z \in D_A(\phi)$ as well.*

Proof. If $A = S$ then $D_A(\phi)$ is empty and the statement is vacuous. So assume $A \subsetneq S$ and let $\varnothing = Z_0, \dots, Z_m$ be the sequence of domains witnessing compatibility of A, so that $\phi(A) = Z_0 \sqcup \dots \sqcup Z_m$, and set $B = \phi(A)$. We must prove $Z \in D_A(\phi)$, which is equivalent to saying $Z \not\sqsubset B$. By means of contradiction, let us suppose $Z \sqsubset B$. Since $Y \pitchfork Z$ but $Y \not\sqsubset B$, we must have $Y \pitchfork B$. Since the Z_j fill B, it must be that $Y \pitchfork Z_j$ for some j (otherwise the subsurface B would be contained in $S \backslash \partial Y$).

By definition of compatibility, Z_j is ϕ–initial for $A_{j-1} = \phi^{-1}(Z_0 \sqcup \dots \sqcup Z_{j-1})$. Since $Y \in D_A(\phi) \subset D_{A_{j-1}}(\phi)$, the Definition 6.14 of initial ensures $Z_j \lessdot Y$ along $[b_\phi, x_\phi]$. Since $Y \lessdot Z$ along $[b_\phi, x_\phi]$ by assumption, it follows that $d_Y(\partial Z_j, \partial Z) \geqslant d_Y(b_\phi, x_\phi) - 2\mathsf{M}/3$. However, as ∂Z and ∂Z_j are both disjoint from ∂B, we also have $d_Y(\partial Z, \partial Z_j) \leqslant 4$. But this contradicts the estimate $d_Y(b_\phi, x_\phi) \geqslant 7\Theta$ from Lemma 6.7. □

Lemma 6.17. *Let $A \sqsubset S$ be compatible for $\phi \in [\phi_0]$ and let $Y, Z \in D_A(\phi)$. If $Z \in D^0_A(\phi)$ and $Y \lhd_i Z$ for $i = 1$ or $i = \tilde{2}$, then $Y \in D^0_A(\phi)$ as well.*

Proof. By means of contradiction, suppose $Y \notin D^0_A(\phi)$, meaning that there exists some $Y' \in D_A(\phi)$ with $Y' \lhd_0 Y$. Hence by definition $Y' = \phi^{-j}(Y)$ for some $j \geqslant 1$ with $\{\phi^{-j}(Y), \dots, Y\} \subset D(\phi)$. By Lemma 6.12 it follows that $\{\phi^{-j}(Z), \dots, Z\} \subset D(\phi)$ as well and therefore that $Z' = \phi^{-j}(Z) \lhd_0 Z$.

To obtain a contradiction, it suffices to show $Z' \in D_A(\phi)$. Since $Y', Z' \in D(\phi)$, we know that Y' and Z' both determine active intervals on $[b_\phi, x_\phi]$. First suppose $Y \lhd_1 Z$. Then $Y' \pitchfork Z'$ and $d_{Y'}(\partial Z', x_\phi) = d_Y(\partial Z, x_\phi) \leqslant \mathsf{M}/3$. By Lemma 3.29 this forces $Y' \lessdot Z'$ along $[b_\phi, x_\phi]$ and hence implies $Y' \lhd_1 Z'$. Since $Y' \in D_A(\phi)$, Lemma 6.16 therefore implies $Z' \in D_A(\phi)$ as well. If instead $Y \lhd_{\tilde{2}} Z$, then $Y' \subsetneq Z'$. Since $Y' \in D_A(\phi)$, we have $Y' \not\sqsubset \phi(A)$ and consequently $Z' \not\sqsubset \phi(A)$ as well. Thus $Z' \in D_A(\phi)$ as required. ☐

The next lemma ensures ϕ–initial domains exists whenever $D_A(\phi)$ is nonempty.

Lemma 6.18 (Initial domains exist)**.** *Let $A \sqsubset S$ be a compatible subsurface for $\phi \in [\phi_0]$. If $D_A(\phi)$ is nonempty, then $D^\star_A(\phi)$ is nonempty as well.*

Proof. Choose a domain $Z' \in D_A(\phi)$ maximizing the quantity $\xi(Z')$. Since $\lhd_0$ restricts to a partial order on the finite set $D_A(\phi)$, there exists $Z \in D^0_A(\phi)$ with $Z = Z'$ or $Z \lhd_0 Z'$. By definition of $\lhd_0$, we have $Z = \phi^{-j}(Z')$ for some $j \geqslant 0$.

Case 1: $Z \in D^1_A(\phi)$: Let $Y \in D^\dagger_A(\phi)$ be the domain provided by Lemma 6.11. If $Y \sqsupset Z$ then we must have $Y = Z \in D^0_A(\phi)$ by the maximality of $\xi(Z)$. Hence $Y \in D^\star_A(\phi)$ and we are done. Otherwise $Y \subsetneq Z$ with $Y \lhd_{\tilde{2}} Z$. Since $Z \in D^0_A(\phi)$, Lemma 6.17 now implies $Y \in D^0_A(\phi)$ and we again conclude $Y \in D^\star_A(\phi)$.

Case 2: $Z \notin D^1_A(\phi)$: Since $\lhd_1$ is a partial order on $D_A(\phi)$, there necessarily exists some $V \in D^1_A(\phi)$ with $V \lhd_1 Z$. Notice that $V \in D^0_A(\phi)$ by Lemma 6.17. Since $V \in D^1_A(\phi)$, we may invoke Lemma 6.11 to obtain a domain $Y \in D^\dagger_A(\phi)$. If $Y \subsetneq V$ with $Y \lhd_{\tilde{2}} V$, then the fact $V \in D^0_A(\phi)$ with Lemma 6.17 implies that $Y \in D^0_A(\phi)$ and hence $Y \in D^\star_A(\phi)$. If instead $Y \sqsupset V$, then the fact $V \pitchfork Z$ ensures we cannot have $Y \perp Z$ or $Y \sqsubset Z$. But $Z \sqsubset Y$ is also ruled out by the maximality of $\xi(Z)$. The only remaining possibility is $Y \pitchfork Z$. Since $V \lessdot Z$, Corollary 3.31 implies that $Y \lessdot Z$, which is to say $Y \lhd_1 Z$. Therefore $Y \in D^0_A(\phi)$ by Lemma 6.17 and we have found the desired domain in $D^\star_A(\phi)$. ☐

The next lemma says, in light of Corollary 3.39, that there are uniformly boundedly many options for the image $\phi(A)$ of a compatible subsurface.

Lemma 6.19 (Bounded compatibility)**.** *If $A \sqsubset S$ is a compatible subsurface for $\phi \in [\phi_0]$, then $B = \phi(A)$ satisfies $d_V(b_\phi, \partial B) \stackrel{+}{\prec}_\Theta 0$ for every domain $V \sqsubset S$.*

Proof. If $A = S$ then ∂S is empty and there is nothing to prove. So suppose $A \subsetneq S$ and let $\varnothing = Z_0, \dots, Z_m$ be the domains witnessing the compatibility of A. If $V \sqsubset B$ or $V \perp B$ there is nothing to prove, since then $d_V(b_\phi, \partial B)$ is just $\operatorname{diam}_{\mathcal{C}(V)} \pi_V(b_\phi) \leqslant \mathsf{L}$. Hence we assume ∂B projects to V. Since B is filled by the Z_i, we may choose $1 \leqslant j \leqslant m$ such that ∂Z_j projects to V. As the multicurves ∂Z_j and ∂B are evidently disjoint, it thus suffices to bound $d_V(b_\phi, \partial Z_j)$. Setting $B_j = Z_0 \sqcup \dots \sqcup Z_{j-1}$, by definition of compatibility we then know that $Z_j \in D(\phi)$ is initial for $A_j = \phi^{-1}(B_j)$.

Let us first suppose $V \in D(\phi)$. Then it must be that $V \in D_A(\phi)$ since $V \sqsubset \phi(A)$ was excluded above. Since $A_j \sqsubset A$, we also have $V \in D_A(\phi) \subset D_{A_j}(\phi)$. As Z_j is initial for A_j, if $V \pitchfork Z_j$, then the failure of $V \lhd_1 Z_j$ implies $Z_j \lessdot V$ along $[b_\phi, x_\phi]$

and hence $d_V(b_\phi, \partial Z_j) \leqslant \mathsf{M}$. If instead $Z_j \subsetneq V$, the failure of $V \lhd_3 Z_j$ implies $d_V(b_\phi, \partial Z_j) \leqslant C_0 + \mathsf{M}$. We are thus done in this case, as $V \perp Z_j$ and $V \sqsubset Z_j$ are precluded by ∂Z_j projecting to V.

Next suppose $V \notin D(\phi)$ so that $d_V(x_\phi, b_\phi) = R_V^\phi \leqslant f^{m_0}(\Theta)$ by Lemma 6.7. Since $Z_j \in D(\phi)$ ensures Z_j has an active interval along $[x_\phi, b_\phi]$, there is a point $t \in [b_\phi, x_\phi]$ with $\partial Z_j \subset \mathrm{base}(\mu_t)$ and consequently (by Theorem 3.19)

$$d_V(b_\phi, \partial Z_j) \leqslant d_V(b_\phi, t) \leqslant d_V(b_\phi, t) + d_V(t, x_\phi) \stackrel{+}{\asymp} d_V(b_\phi, x_\phi) \stackrel{+}{\prec} 0. \qquad \square$$

6.4. **Coherence.** Recall that our goal in Theorem 6.1 is to bound the number elements $\phi \in [\phi_0]$ producing a common pair of points (a_ϕ, b_ϕ).

Definition 6.20 (Coherence)**.** Given a subsurface $A \sqsubset S$, we say a family $\mathcal{F} \subset [\phi_0]$ is A–*coherent* if for all pairs $\phi, \psi \in \mathcal{F}$

- A is compatible with ϕ and ψ,
- ϕ and ψ agree on A, and
- $a_\phi = a_\psi$ and $b_\phi = b_\psi$.

The *displacement* of the family is $r(\mathcal{F}) = \max\{d_{\mathcal{T}(S)}(x_0, \phi(x_0)) \mid \phi \in \mathcal{F}\}$.

Definition 6.21 (Pre-initial)**.** Let $\mathcal{F} \subset [\phi_0]$ an A–coherent family. We say a domain $Y \sqsubset S$ is *pre-initial* for $\mathcal{F}$ if it is the preimage of some initial domain, that is, if $Y = \phi^{-1}(Z)$ some element $\phi \in \mathcal{F}$ and domain $Z \in D(\phi)$ that is ϕ–initial for A.

Lemma 6.22 (Boundedly many pre-initial domains)**.** *There is a degree* 1 *polynomial q_1 such that for any subsurface $A \sqsubset S$ and A–coherent family $\mathcal{F} \subset [\phi_0]$, the cardinality of the set of pre-initial domains for $\mathcal{F}$ is at most $q_1(r(\mathcal{F}))$.*

Proof. Let $\mathcal{P} = \{Y_1, Y_2, \dots\}$ be the set of pre-initial domains, and for each i choose $\phi_i \in \mathcal{F}$ such that $Y_i = \phi_i^{-1}(Z_i)$ for some ϕ_i–initial domain $Z_i \in D(\phi_i)$. Let $k_i < m_0$ be the index of Z_i, that is, the minimal $k_i \geqslant 0$ so that $\phi_i^{-k_i-1}(Z_i) \notin D(\phi_i)$. For each, k let $\mathcal{P}_k$ be the subset of pre-initial domains Y_i for which $k_i = k-1$. This gives a partition $\mathcal{P} = \mathcal{P}_1 \sqcup \dots \sqcup P_{m_0}$. Restricting to one subcollection $\mathcal{P}_k$ we henceforth assume $k = k_i + 1$ for all i.

Let us write $V_i = \phi_i^{-k}(Z_i)$ so that $V_i \notin D(\phi_i)$. Then for each i we have $\phi_i(V_i), \dots, \phi_i^k(V_i) \in D(\phi_i)$, with $Z_i = \phi_i^k(V_i)$ being ϕ_i–initial. The $\lhd_0$–minimality of Z_i in $D_A(\phi_i)$ implies that $\phi_i(V_i), \dots, \phi_i^{k-1}(V_i) \notin D_A(\phi_i)$. Thus for all $0 \leqslant n \leqslant k-2$ we have $\phi_i^{n+1}(V_i) \in D(\phi_i) \backslash D_A(\phi_i)$, meaning $\phi_i^{n+1}(V_i) \sqsubset \phi(A)$ or equivalently $\phi_i^n(V_i) \sqsubset A$. By assumption, all the maps $\phi \in \mathcal{F}$ agree on A; let us write $\psi \colon A \to B$ for this common restriction to A. With this notation we conclude that

$$Y_i = \phi_i^{-1}(Z_i) = \psi^{k-1}(V_i)$$

for all i. Since the domains Y_i are all distinct by assumption, it follows that the domains V_i and V_j are distinct whenever $i \neq j$.

By coherence, the points a_ϕ, b_ϕ for $\phi \in \mathcal{F}$ all agree; let us call these common points a, b. By Definition 6.6 of backtracking, having $V_i \notin D(\phi_i)$ and $\phi_i(V_i) \in D(\phi_i)$ means that

$$R^{\phi_i}_{\phi_i(V_i)} \geqslant f(R^{\phi_i}_{V_i}) = 7R^{\phi_i}_{V_i} + 7\Theta,$$

which by Proposition 5.5(7) implies $d_{V_i}(x_0, b) \geqslant 6\Theta$. Fixing $i = 1$ and considering the element ϕ_1, we know the triple $(x_0, b, \phi_1(x_0))$ is Θ–aligned in all domains. Therefore, for all $j \geqslant 1$ we have

$$d_{V_j}(x_0, \phi_1(x_0)) \geqslant d_{V_j}(x_0, b) - \Theta \geqslant 5\Theta.$$

We note that we may choose Θ large enough so that $\log(\Theta)$ is a valid threshold in the distance formula (Theorem 3.33). From this the number of domains $U \sqsubset S$ with $d_U(x_0, \phi_1(x_0)) \geqslant \Theta$ is bounded linearly in terms of $d_{\mathcal{T}(S)}(x_0, \phi_1(x_0)) \leqslant r(\mathcal{F})$, which finishes the proof of the lemma. $\square$

6.5. **Supercoherence.** We next consider coherent families with additional data:

Definition 6.23 (Supercoherence)**.** A family $\mathcal{F} \subset [\phi_0]$ will be called *supercoherent* for a subsurface $A \sqsubset S$ and domain $Y \sqsubset S$ if $\mathcal{F}$ is A–coherent and for all $\phi, \psi \in \mathcal{F}$:

- $Z_\phi = \phi(Y)$ is ϕ–initial for A,
- $\phi(A') = \psi(A')$ where A' is the subsurface filled by A and Y,
- if Y is nonannular, then $d_{Z_\phi}(b_\phi, x_\phi) = d_{Z_\psi}(b_\psi, x_\psi)$,
- if Y is annular, then $d_{Z_\phi}(b_\phi, \phi(x_0)) = d_{Z_\psi}(b_\psi, \psi(x_0))$.

Note that Z_ϕ being initial implies $Z_\phi \not\sqsubset \phi(A)$ and thus $Y \not\sqsubset A$. Also note that the subsurfaces $B = \phi(A)$ and $B' = \phi(A')$ (filled by B and Z_ϕ) are independent of ϕ. We also allow Y to denote the empty domain and say $\mathcal{F}$ is $(A, \varnothing)$–supercoherent to mean that it is A–coherent but that $D_A(\phi)$ is empty for each $\phi \in \mathcal{F}$.

Lemmas 6.19 and 6.22 allow us to easily partition coherent families into boundedly many supercoherent ones:

Lemma 6.24. *There is a degree* 2 *polynomial* q_2 *such that for any subsurface* $A \sqsubset S$*, any* A*–coherent family* $\mathcal{F} \subset [\phi_0]$ *may be partitioned into at most* $q_2(r(\mathcal{F}))$ *subfamilies* $\mathcal{F}'$ *that are each* (A, Y')*–supercoherent for some domain* Y'*.*

Proof. The elements $\phi \in \mathcal{F}$ for which $D_A(\phi)$ is empty comprise a subset of $\mathcal{F}$ that is $(A, \varnothing)$–supercoherent. Excising these, we henceforth suppose each $\phi \in \mathcal{F}$ has $D_A(\phi)$ nonempty and, in particular, that $A \neq S$. Accordingly, for each $\phi \in \mathcal{F}$ we may use Lemma 6.18 to choose some initial domain $Z_\phi \in D^*_A(\phi)$. We then set $Y_\phi = \phi^{-1}(Z_\phi)$ and let B'_ϕ denote the subsurface filled by Z_ϕ and $B = \phi(A)$. As $\phi^{-1}(B'_\phi)$ is clearly compatible for ϕ by construction, Lemma 6.19 and Corollary 3.39 provide a uniform bound k_0 (depending only on Θ) on the number of domains B'_ϕ produced in this way. Similarly Lemma 6.22 says there are at most $q_1(r(\mathcal{F}))$ possibilities for the domain Y_ϕ. Hence after partitioning into $k_0 q_1(r(\mathcal{F}))$ subfamilies we may assume $Y = Y_\phi$ and $B' = B'_\phi$ are independent of ϕ.

Now, if Y (and thus each Z_ϕ) is nonannular, then Proposition 5.5(9) provides a uniform constant k_1 so that each integer $d_{Z_\phi}(b_\phi, x_\phi) = R^\phi_{Z_\phi}$ is at most $k_1 d_{\mathcal{T}(S)}(x_0, \phi(x_0)) + k_1$. Thus we may further partition into at most $k_1 r(\mathcal{F}) + k_1$ subfamilies so that $d_{Z_\phi}(b_\phi, x_\phi)$ is independent of ϕ. If instead Y is annular, then for all $\phi \in \mathcal{F}$ we have $R^\phi_{Z_\phi} > 7\Theta$ by Lemma 6.7 since $Z_\phi \in D(\phi)$. Therefore Proposition 5.5(8) says $d_{Z_\phi}(b_\phi, \phi(x_0)) \leqslant \Theta$ and we may further partition into at most Θ subfamilies so that $d_{Z_\phi}(b_\phi, \phi(x_0))$ is independent of ϕ. Each subfamily $\mathcal{F}'$ produced in this way is then (A, Y)–supercoherent, where $Y = Y_\phi$ for any $\phi \in \mathcal{F}'$. $\square$

6.6. **Extending supercoherence.** The previous section shows that A–coherent families can be refined into (A, Y)–supercoherent ones. The remaining ingredient is to show that each (A, Y)–supercoherent family can be further refined into A'–coherent families for the enlarged subsurface A' filled by Y and A, or $A' = S$ in the case $Y = \varnothing$. This is the heart of our reconstructive argument.

Proposition 6.25. *There is a constant q_3 such that every (A, Y)–supercoherent family $\mathcal{F} \subset [\phi_0]$ can be partitioned into at most q_3 subfamilies $\mathcal{F}'$ that are each A'–coherent, where $A' = S$ when Y is the empty domain and otherwise A' is the subsurface filled by A and Y. In particular, in the latter case $\xi(A') > \xi(A)$.*

Proof. Since $\mathcal{F}$ is A–coherent, we know $a_\phi = a_\psi$ and $b_\phi = b_\psi$ for all $\phi, \psi \in \mathcal{F}$; let us write $a = a_\phi$ and $b = b_\phi$ for these common points. The definition of supercoherence ensures A' is compatible with each $\phi \in \mathcal{F}$; indeed, if $Y = \varnothing$ and $A' = S$ compatibility is automatic, and otherwise it follows from the compatibility of A (Definition 6.15) and fact that $Z_\phi = \phi(Y)$ is ϕ–initial for A. Supercoherence also gives $\phi(A') = \psi(A')$ for all $\phi, \psi \in \mathcal{F}$; let us call this common subsurface B'. Hence proving A'–coherence amounts to establishing agreement on A'. This is equivalent to showing the subset $\{(\phi\psi^{-1})|_{B'} \mid \phi, \psi \in \mathcal{F}\}$ of $\mathrm{Mod}(B')$ has uniformly bounded cardinality. By Lemma 6.4, for this it suffices to prove that

$$d_W(\phi\psi^{-1}(b), b) \stackrel{+}{\prec}_\Theta 0 \qquad \text{for all } \phi, \psi \in \mathcal{F} \text{ and all domains } W \sqsubset B'. \tag{6.26}$$

To set notation, for $W \sqsubset B'$ and $\phi, \psi \in \mathcal{F}$ we will write $V = \phi^{-1}(W) \sqsubset A'$ and $W' = \psi(V) \sqsubset B'$. Note that then

$$d_W(\phi\psi^{-1}(b), b) = d_V(\phi^{-1}(b), \psi^{-1}(b)) = d_{W'}(b, \psi\phi^{-1}(b)); \tag{6.27}$$

hence we are free to bound either of these three quantities. We first dispense with the case that Y is empty and, accordingly $B' = S = A'$:

Claim 6.28. *If $Y = \varnothing$, then $d_W(\phi\psi^{-1}(b), b) \stackrel{+}{\prec}_\Theta 0$ for all $W \sqsubset B'$ and $\phi, \psi \in \mathcal{F}$.*

Proof. First suppose $W \in D(\phi)$. Since $D_A(\phi) = \varnothing$ by definition of supercoherence, evidently $W \notin D_A(\phi)$ which means $W \sqsubset \phi(A)$ and hence $V \sqsubset A$. Since ϕ and ψ agree on A, we may apply the equal isometries $\phi = \psi \colon \mathcal{C}(V) \to \mathcal{C}(W)$ to conclude

$$d_V(\phi^{-1}(b), \psi^{-1}(b)) = d_{\phi(V)}(\phi\phi^{-1}(b), \psi\psi^{-1}(b)) = d_W(b, b) \leqslant \mathsf{L}.$$

If $W' \in D(\phi)$ we similarly conclude $d_V(\phi^{-1}(b), \psi^{-1}(b)) \leqslant \mathsf{L}$.

It remains to suppose $W \notin D(\phi)$ and $W' \notin D(\psi)$. By Lemma 6.7, this gives

$$d_V(\phi^{-1}(b), x_\phi) = d_W(b, x_\phi) \stackrel{+}{\prec}_\Theta 0 \quad \text{and} \quad d_V(\psi^{-1}(b), x_\psi) = d_{W'}(b, x_\psi) \stackrel{+}{\prec}_\Theta 0.$$

Hence by the triangle inequality it suffices to bound $d_V(x_\phi, x_\psi)$. Observe that there exists $n \geqslant 0$ so that $\phi^{-n}(V) \notin D(\phi)$ and $\psi^{-n}(V) \notin D(\psi)$; if $n+1 = m_0$ is the order of ϕ_0 (and thus of ϕ and ψ as well), then the condition is satisfied for $\phi^{-n}(V) = W$ and $\psi^{-n}(V) = W'$. Thus we may let $n \geqslant 0$ be the smallest integer so that both $\phi^{-n}(V) \notin D(\phi)$ and $\psi^{-n}(V) \notin D(\psi)$.

We claim that $\phi^{-i}(V) = \psi^{-i}(V)$ for each $0 \leqslant i \leqslant n$. Indeed, when $0 \leqslant i < n$ either $\phi^{-i}(V) \in D(\phi)$ and hence $\phi^{-i}(V) \sqsubset \phi(A) = B$, or else $\psi^{-i}(V) \in D(\psi)$ and hence $\psi^{-i}(V) \sqsubset \psi(A) = B$. In either case, inductively assuming $\phi^{-i}(V) = \psi^{-i}(V)$, the fact that the maps ϕ^{-1}, ψ^{-1} agree on B implies that $\phi^{-i-1}(V) = \psi^{-i-1}(V)$.

Let us write $V_i = \phi^{-i}(V) = \psi^{-i}(V)$ for $0 \leqslant i \leqslant n$. The facts that $V_n \notin D(\phi)$ and $V_n \notin D(\psi)$ now, by Lemma 6.7, give

$$d_{V_n}(x_\phi, x_\psi) \leqslant d_{V_n}(x_\phi, b) + d_{V_n}(b, x_\psi) = R^\phi_{V_n} + R^\psi_{V_n} \leqslant 2f^{m_0}(\Theta) \stackrel{+}{\prec}_\Theta 0.$$

As we have seen, the maps ϕ, ψ agree on each domain $V_n, \dots, V_1, V$. Successively applying the isometries $\phi = \psi \colon \mathcal{C}(V_{i+1}) \to \mathcal{C}(V_i)$ therefore gives the desired bound

$$d_V(x_\phi, x_\psi) = d_{V_1}(x_\phi, x_\psi) = \cdots = d_{V_n}(x_\phi, x_\psi) \stackrel{+}{\prec}_\Theta 0. \qquad \square$$

We henceforth assume that Y is nonempty and thus, by supercoherence, that $Z_\theta = \theta(Y) \sqsubset B'$ is initial in $D_A(\theta)$ for each θ. After partitioning into at most m_0 subfamilies, we may additionally assume each of these initial domains Z_θ has the same index k in $D(\theta)$. By definition, this means $\{\theta^{-k}(Z_\theta), \dots Z_\theta\} \subset D(\theta)$ but that $\theta^{-k-1}(Z_\theta) \notin D(\theta)$. Recalling that $\theta^{-1}(Z_\theta) = Y$, the fact that Z_θ is initial now forces $\theta^{-k+1}(Y), \dots, Y \notin D_A(\theta)$, which means $\theta^{-n}(Y) \sqsubset A$ for each $1 \leqslant n \leqslant k$. Since the elements $\theta \in \mathcal{F}$ all agree on A, it follows that for each $1 \leqslant n \leqslant k$ the common domain $Y_n = \theta^{-n}(Y)$ and map $\theta\colon Y_n \to Y_{n-1}$ are independent of $\theta \in \mathcal{F}$.

The fact that $Y_k = \theta^{-k}(Y) = \theta^{-k-1}(Z_\theta) \notin D(\theta)$ implies by Lemma 6.7 that

$$d_{Y_k}(x_\theta, b) \leqslant f^{m_0}(\Theta) \quad \text{for all } \theta \in \mathcal{F}. \tag{6.29}$$

Since the numbers $d_{Y_k}(x_\theta, b)$ are discrete and uniformly bounded, after further partitioning into at most $f^{m_0}(\Theta)$ subfamilies we may assume these distances all agree and hence that $R^\theta_{Y_k} = d_{Y_k}(x_\theta, b)$ is independent of θ. By definition, the constants $C_{Z_\theta} = 6R^\theta_{Y_k} + 3\Theta + 2\mathsf{M}$ are therefore also independent of $\theta \in \mathcal{F}$.

For any elements $\phi, \psi \in \mathcal{F}$, by (6.29) and the triangle inequality we have

$$d_{Y_k}(x_\phi, x_\psi) \leqslant d_{Y_k}(x_\phi, b) + d_{Y_k}(b, x_\psi) \leqslant 2f^{m_0}(\Theta).$$

Successively applying the isometries $\phi|_A = \psi|_A\colon \mathcal{C}(Y_k) \to \cdots \to \mathcal{C}(Y)$ thus gives

$$d_Y(x_\phi, x_\psi) = \cdots = d_{Y_k}(x_\phi, x_\psi) \leqslant 2f^{m_0}(\Theta) \quad \text{for all } \phi, \psi \in \mathcal{F}. \tag{6.30}$$

Using this, we next establish (6.26) for the domain $W = Z_\phi$:

Claim 6.31. $d_{Z_\phi}(\phi\psi^{-1}(b), b) < 3d_Y(x_\phi, x_\psi) + 3\Theta \stackrel{+}{\prec}_\Theta 0$ *for all* $\phi, \psi \in \mathcal{F}$.

Proof. We know from Proposition 5.5(3) that $(x_\psi, b, \psi(x_0))$ is Θ–aligned; hence:

$$d_{Z_\psi}(x_\psi, b) + d_{Z_\psi}(b, \psi(x_0)) \leqslant d_{Z_\psi}(x_\psi, \psi(x_0)) + \Theta$$

Applying the isometry $\psi^{-1}\colon \mathcal{C}(Z_\psi) \to \mathcal{C}(Y)$ thus gives

$$d_Y(x_\psi, \psi^{-1}(b)) + d_Y(\psi^{-1}(b), x_0) \leqslant d_Y(x_\psi, x_0) + \Theta.$$

We may swap x_ψ for x_ϕ at the cost of $d_Y(x_\psi, x_\phi)$ and then apply $\phi\colon \mathcal{C}(Y) \to \mathcal{C}(Z_\phi)$ to conclude $(x_\phi, \phi\psi^{-1}(b), \phi(x_0))$ is $(2d_Y(x_\phi, x_\psi) + \Theta)$–aligned in Z_ϕ:

$$d_{Z_\phi}(x_\phi, \phi\psi^{-1}(b)) + d_{Z_\phi}(\phi\psi^{-1}(b), \phi(x_0)) \leqslant d_{Z_\phi}(x_\phi, \phi(x_0)) + 2d_Y(x_\psi, x_\phi) + \Theta.$$

Since $(x_\phi, b, \phi(x_0))$ is Θ–aligned (Proposition 5.5(3)), Lemma 3.18 now implies that $\pi_{Z_\phi}(b)$ and $\pi_{Z_\phi}(\phi\psi^{-1}(b))$ respectively lie within $\frac{\Theta}{2} + 4\delta + \mathsf{L}$ and $d_Y(x_\phi, x_\psi) + \frac{\Theta}{2} + 4\delta + \mathsf{L}$ of any geodesic from $\pi_{Z_\phi}(\phi(x_0))$ to $\pi_{Z_\phi}(x_\phi)$.

If Y is nonannular, then supercoherence implies the distances

$$d_{Z_\phi}(\phi\psi^{-1}(b), x_\phi) = d_Y(\psi^{-1}(b), x_\phi), \quad \text{and}$$
$$d_{Z_\phi}(b, x_\phi) = d_{Z_\psi}(b, x_\psi) = d_Y(\psi^{-1}(b), x_\psi)$$

differ by at most $d_Y(x_\phi, x_\psi)$. Otherwise Y is annular and the distances

$$d_{Z_\phi}(b, \phi(x_0)) \quad \text{and} \quad d_{Z_\phi}(\phi\psi^{-1}(b), \phi(x_0)) = d_Y(\psi^{-1}(b), x_0) = d_{Z_\psi}(b, \psi(x_0))$$

agree by supercoherence. In either case, these estimates and the fact that $\pi_{Z\phi}(b)$ and $\pi_{Z_\phi}(\phi\psi^{-1}(b))$ both lie within controlled distance of a $\mathcal{C}(Z_\phi)$ geodesic from x_ϕ to $\phi(x_0)$ now imply the desired bound

$$d_{Z_\phi}(\phi\psi^{-1}(b), b) \leqslant 3d_Y(x_\psi, x_\phi) + 2\Theta + 16\delta + 5\mathsf{L} < 3d_Y(x_\phi, x_\psi) + 3\Theta. \qquad \square$$

This finishes the proof of Claim 6.31.

We continue with the proof of (6.26). Now let $A_0 = A'/A$ and $B_0 = B'/B$. Since all maps in $\mathcal{F}$ send A to B and A' to B', it follows that $\theta(A_0) = B_0$ for all $\theta \in \mathcal{F}$. After partitioning into boundedly many subfamilies, we additionally assume, for all $\theta, \theta' \in \mathcal{F}$, that $\theta'\theta^{-1}$ preserves each component and boundary component of B, B', and B_0. Note that if $A \perp Y$ then A' is just the union of A and Y so that $A_0 = Y$ and $B_0 = \theta(Y) = Z_\theta$ for each $\theta \in \mathcal{F}$.

Since we already know all ϕ, ψ agree on A, to prove agreement on A' it suffices to establish agreement on A_0 and on the boundary components of A that are essential in A'. The next claim essentially provides agreement on these boundary components

Claim 6.32. *$d_\beta(\partial Z_\phi, \partial Z_\psi) \overset{+}{\prec}_\Theta 0$ for each boundary component β of B.*

Proof. Let U be the annulus with core $\partial U = \beta$. By assumption $\phi\psi^{-1}(\beta) = \beta$ and $\phi\psi^{-1}(\partial Z_\psi) = \partial Z_\phi$. If ∂Z_ϕ is disjoint from β there is nothing to prove, so we assume $\partial Z_\phi \pitchfork \beta$, in which case $\partial Z_\psi \pitchfork \beta$ as well. Thus $Z_\phi \pitchfork U$ and $Z_\psi \pitchfork U$.

If $U \sqsubset B$ (that is, if B has an annular component isotopic to U) then agreement on A (coming from A–coherence) immediately implies the claim. So we suppose U is not a component of B. We claim that $d_U(b, \partial Z_\phi) \overset{+}{\prec}_\Theta 0$ and, symmetrically, $d_U(b, \partial Z_\psi) \overset{+}{\prec}_\Theta 0$. The claim will then follow from the triangle inequality.

By means of contradiction, suppose $d_U(b, \partial Z_\phi) > f^{m_0}(\Theta) + \mathsf{M}$. Since $Z_\phi \in D(\phi)$ satisfies $d_{Z_\phi}(b, x_\phi) \geqslant \mathsf{M}$, Corollary 3.27 implies that $d_U(b, \partial Z_\phi) \leqslant d_U(b, x_\phi) + \mathsf{M}/3$. Therefore $d_U(b, x_\phi) > f^{m_0}(\Theta)$ and consequently $U \in D(\phi)$ by Lemma 6.7. Since $U \not\pitchfork B = \phi(A)$, it follows that $U \in D_A(\phi)$. Now the fact that Z_ϕ is initial in $D_A(\phi)$ forces $Z_\phi \lhd_1 U$ in $D(\phi)$, meaning that Z_ϕ is time ordered before U along $[b, x_\phi]$. But this implies $d_U(b, \partial Z_\phi) \leqslant \mathsf{M}/3$ contradicting our above assumption. ☐

Now consider an arbitrary domain $W \sqsubset B_0$. Since B' is filled by B and Z_ϕ, and W is disjoint from B, it cannot be that Z_ϕ and W are disjoint. Further, $Z_\phi \sqsubset W$ occurs only in the case $W = Z_\phi$, which has been dealt with in Claim 6.31 above. Thus we may assume $W \sqsubsetneq Z_\phi$ or $W \pitchfork Z_\phi$. Let us deal with these two possibilities separately.

Claim 6.33. *If $W \sqsubsetneq Z_\phi$ (and hence $W' \sqsubsetneq Z_\psi$), then $d_W(\phi\psi^{-1}(b), b) \overset{+}{\prec}_\Theta 0$.*

Proof. First suppose $W \in D(\phi)$. Since $W \sqsubset B_0$, we have $W \not\pitchfork B = \phi(A)$ and consequently $W \in D_A(\phi)$ as well. Since Z_ϕ is initial in $D_A(\phi)$, it cannot be that $W \lhd_2 Z_\phi$. Thus by definition of $\lhd_2$ it must be that

$$d_{Z_\phi}(b, \partial W) \geqslant C_{Z_\phi} = 6 d_{Y_k}(b, x_\phi) + 3\Theta + 2\mathsf{M}.$$

Since $d_{Y_k}(x_\phi, b) = d_{Y_k}(x_\psi, b)$ by our assumption, the triangle inequality implies $d_{Y_k}(x_\phi, x_\psi) \leqslant 2 d_{Y_k}(x_\phi, b)$. From Claim 6.31 we also know that

$$d_{Z_\phi}(\phi\psi^{-1}(b), b) < 3 d_{Y_k}(x_\phi, x_\psi) + 3\Theta.$$

Combining these yields $d_{Z_\phi}(b, \partial W) > d_{Z_\phi}(\phi\psi^{-1}(b), b) + \mathsf{M}$. The BGIT (Corollary 3.27) therefore implies the bound $d_W(b, \phi\psi^{-1}(b)) < \mathsf{M} \overset{+}{\prec} 0$. If $W' \in D(\psi)$ the same reasoning bounds $d_{W'}(\psi\phi^{-1}(b), b)$.

It remains to suppose $W \notin D(\phi)$ and $W' \notin D(\psi)$ which, by Lemma 6.7, implies

$$d_V(\phi^{-1}(b), x_\phi) = d_W(b, x_\phi) \leqslant f^{m_0}(\Theta), \quad d_V(\psi^{-1}(b), x_\psi) = d_{W'}(b, x_\psi) \leqslant f^{m_0}(\Theta).$$

Hence in order to bound $d_V(\phi^{-1}(b), \psi^{-1}(b))$ it suffices to bound $d_V(x_\phi, x_\psi)$. For each $1 \leqslant j \leqslant k$, we have that $\phi^{-j}(W) \sqsubset Y_{j-1} \sqsupset \psi^{-j}(W')$. Since $\phi^{-1}(W) = V = \psi^{-1}(W')$ by construction and ϕ^{-1} agrees with ψ^{-1} on $Y_{j-1} \sqsubset B$, it follows by induction that $\phi^{-j-1}(W) = \psi^{-j-1}(W')$ for each $1 \leqslant j \leqslant k$ and that ϕ agrees with ψ on this domain, which for brevity we denote $V_j = \phi^{-j}(V) = \phi^{-j-1}(W)$. Applying the equal maps $\phi = \psi$ to V_j for $1 \leqslant j \leqslant k$ thus gives

$$d_{V_k}(x_\phi, x_\psi) = \cdots = d_V(x_\phi, x_\psi).$$

Hence it suffices to bound $d_{V_k}(x_\phi, x_\psi)$.

If $d_{V_k}(x_\phi, b), d_{V_k}(b, x_\psi) \leqslant \mathsf{M}$ we are done by the triangle inequality. Supposing instead $d_{V_k}(x_\phi, b) > \mathsf{M}$, then Corollary 3.27 implies $d_{Y_k}(x_\phi, \partial V_k) \leqslant d_{Y_k}(x_\phi, b) + \mathsf{M}$. As $d_{Y_k}(x_\phi, x_\psi) \leqslant 2d_{Y_k}(b, x_\phi)$, we thus also have $d_{Y_k}(x_\psi, \partial V_k) \leqslant 3d_{Y_k}(b, x_\phi) + \mathsf{M}$. Similar reasoning applies if $d_{V_k}(x_\psi, b) > \mathsf{M}$, so that in either case we may assume

$$d_{Y_k}(x_\phi, \partial V_k) \leqslant 3R^\phi_{Y_k} + \mathsf{M} \quad \text{and} \quad d_{Y_k}(x_\psi, \partial V_k) \leqslant 3R^\psi_{Y_k} + \mathsf{M}.$$

As the leftmost quantity above is invariant under applying ϕ, this gives

$$d_{\phi(Y_j)}(x_\phi, \partial\phi(V_j)) = d_{Y_k}(x_\phi, \partial V_k) \leqslant 3R^\phi_{Y_k} + \mathsf{M} \quad \text{for all } 0 \leqslant j \leqslant k, \tag{6.34}$$

On the other hand, the fact that $\{\phi(Y_k), \ldots, \phi(Y)\}$ is a jump sequence for ϕ ensures

$$d_{\phi(Y_j)}(b, x_\phi) \geqslant f(R^\phi_{Y_k}) = 7R^\phi_{Y_k} + 7\Theta \quad \text{for all } 0 \leqslant j \leqslant k.$$

Since (x_ϕ, b, x_0) and $(x_\phi, b, \phi(x_0))$ are Θ–aligned in Y_j by Proposition 5.5(4) (as $V_j \subsetneq Y_j$ ensures Y_j is nonannular), we claim this implies

$$d_{\phi(V_j)}(b, x_0), d_{\phi(V_j)}(b, \phi(x_0)) \leqslant \mathsf{M} \quad \text{for all } 0 \leqslant j \leqslant k.$$

Indeed if, say, $d_{\phi(V_j)}(b, x_0) > \mathsf{M}$ then we may choose $u \in [b, x_0]$ containing $\partial\phi(V_j)$ in is Bers marking and use alignment to conclude

$$\begin{aligned} d_{\phi(Y_j)}(x_\phi, \partial\phi(V_j)) + \mathsf{L} &\geqslant d_{\phi(Y_j)}(x_\phi, u) \geqslant d_{\phi(Y_j)}(x_\phi, x_0) - d_{\phi(Y_j)}(x_0, u) \\ &\geqslant d_{\phi(Y_j)}(x_\phi, b) + d_{\phi(Y_j)}(b, x_0) - d_{\phi(Y_j)}(x_0, u) - \Theta \\ &\geqslant d_{\phi(Y_j)}(x_\phi, b) + d_{\phi(Y_j)}(b, u) - \Theta - \mathsf{B} \geqslant 7R^\phi_{Y_k} + 5\Theta, \end{aligned}$$

contradicting (6.34). A similar contradiction arises if $d_{\phi(V_j)}(b, \phi(x_0)) > \mathsf{M}$. By the triangle inequality, for each $0 \leqslant j \leqslant k$ we now deduce

$$d_{V_j}(x_\phi, x_0) = d_{\phi(V_j)}(x_\phi, \phi(x_0)) \leqslant d_{\phi(V_j)}(x_\phi, x_0) + 2\mathsf{M}.$$

Applying this inductively for $j = k, \ldots, 1$ therefore gives

$$d_{V_k}(x_\phi, x_0) - 2k\mathsf{M} \leqslant d_V(x_\phi, x_0) = d_W(x_\phi, \phi(x_0)) \leqslant d_W(x_\phi, b) + \mathsf{M}.$$

Since $k < m_0$ and $d_W(x_\phi, b) \leqslant f^{m_0}(\Theta)$ by virtue of $W \notin D(\phi)$, we conclude that

$$d_{V_k}(x_\phi, x_0) \leqslant f^{m_0}(\Theta) + 2m_0\mathsf{M}.$$

A symmetric argument yields $d_{V_k}(x_\psi, x_0) \leqslant f^{m_0}(\Theta) + 2m_0\mathsf{M}$. The triangle inequality therefore gives $d_{V_k}(x_\phi, x_\psi) \stackrel{+}{\prec}_\Theta 0$ and completes the proof of the claim. □

Claim 6.35. *If $W \pitchfork Z_\phi$ (and hence $W' \pitchfork Z_\psi$) then $d_W(\phi\psi^{-1}(b), b) \stackrel{+}{\prec}_\Theta 0$.*

Proof. Observe first that $d_W(x_\phi, \partial Z_\phi) = d_V(x_\phi, \partial Y)$. If this quantity is larger than $\frac{3\mathsf{M}}{2}$, then, since $d_{Z_\phi}(b, x_\phi) \geqslant \mathsf{M}$, Corollary 3.27 implies that $d_W(x_\phi, b) \geqslant d_W(x_\phi, \partial Z_\phi) - \mathsf{M}/3 > \mathsf{M}$. Therefore W has an active interval and is time-ordered before Z_ϕ along $[x_\phi, b]$. In this case $d_{Z_\phi}(x_\phi, \partial W) \leqslant \mathsf{M}/3$ and so

$$d_{Z_\phi}(b, \partial W) \geqslant d_{Z_\phi}(b, x_\phi) - d_{Z_\phi}(\partial W, x_\phi) \geqslant f(R^\phi_{Y_k}) - \tfrac{\mathsf{M}}{3}$$

by the fact that Z_ϕ satisfies the jumping criterion to be in $D(\phi)$. Recalling Claim 6.31, the assumption $d_{Y_k}(b, x_\psi) = d_{Y_k}(b, x_\phi) = R^\phi_{Y_k}$, and the definition of f, we also have

$$d_{Z_\phi}(\phi\psi^{-1}(b), b) \leqslant 3d_{Y_k}(x_\phi, x_\psi) + 3\Theta \leqslant 6R^\phi_{Y_k} + 3\Theta < f(R^\phi_{Y_k}) - 2\mathsf{M}.$$

Therefore applying ϕ^{-1} to the left side of the previous two displayed inequalities gives

$$d_Y(\phi^{-1}(b), \partial V) \geqslant d_Y(\psi^{-1}(b), \phi^{-1}(b)) + \mathsf{M}.$$

By Corollary 3.27, it follows that we must have $d_V(\phi^{-1}(b), \psi^{-1}(b)) < \mathsf{M} \stackrel{+}{\prec}_\Theta 0$. This proves the claim in the case that $d_W(x_\phi, \partial Z_\phi) = d_V(x_\phi, \partial Y) \geqslant \frac{3\mathsf{M}}{2}$. The same reasoning applies when $d_{W'}(x_\psi, \partial Z_\psi) = d_V(x_\psi, \partial Y) \geqslant \frac{3\mathsf{M}}{2}$.

Now suppose $W \in D(\phi)$. Since W is disjoint from $B = \phi(A)$, we necessarily have $W \in D_A(\phi)$. The fact that Z_ϕ is initial in $D_A(\phi)$ thus implies $Z_\phi \lhd_1 W$ in $D(\phi)$. That is, W is time-ordered before Z_ϕ along $[x_\phi, b]$ and so $d_W(x_\phi, \partial Z) \geqslant d_W(x_\phi, b) - d_W(\partial Z, b) \geqslant 7\Theta - \mathsf{M}/3 > 6\mathsf{M}$ and we are done by the previous paragraph. The same conclusion holds if $W' \in D(\psi)$.

It remains to suppose $d_V(x_\phi, \partial Y), d_V(x_\psi, \partial Y) < \frac{3\mathsf{M}}{2}$ and $W \notin D(\phi)$, $W' \notin D(\psi)$. In particular $d_V(x_\phi, x_\psi) \leqslant d_V(x_\phi, \partial Y) + d_V(x_\psi, \partial Y) < 3\mathsf{M}$. Furthermore,

$$d_V(x_\phi, \phi^{-1}(b)) = d_W(x_\phi, b) \quad \text{and} \quad d_V(x_\psi, \psi^{-1}(b)) = d_{W'}(x_\psi, b)$$

are both bounded by $f^{m_0}(\Theta)$. Hence by the triangle inequality, $d_V(\phi^{-1}(b), \psi^{-1}(b))$ is bounded by $2f^{m_0}(\Theta)$ plus $d_V(x_\phi, x_\psi)$ and therefore by $2f^{m_0}(\Theta) + 3\mathsf{M} \stackrel{+}{\prec}_\Theta 0$. □

With these claims in hand, we may now complete the proof of the proposition. When Y is empty and $B' = S$, Claim 6.28 shows $d_W(\phi\psi^{-1}(b), b)) \stackrel{+}{\prec}_\Theta 0$ for all $\phi, \psi \in \mathcal{F}$. Thus by Lemma 6.4 we may partition into boundedly many subcollections to achieve agreement on $A' = S$. When Y is nonempty, we as above set $A_0 = A'/A$ and $B_0 = B' \backslash B$ and partition so that all maps send A to B and A_0 to B_0 inducing the same bijection of boundary components. By assumption we know all maps agree on A. Claims 6.31, 6.33 and 6.35 together with Lemma 6.4 imply that up to partitioning into boundedly many subfamilies we may assume all maps $A_0 \to B_0$ agree as well. If Y is disjoint from A, then A' is the disjoint union of A and A_0 and agreement on A' definitionally follows from agreement on A and A_0. Otherwise $Y \pitchfork A$ and B' is obtained by gluing B and B_0 along certain boundary components β of ∂B that are essential in B'. All of our maps $\phi\psi^{-1} \colon B' \to B'$ preserve these curves and are the identity on the complement $B' \backslash \partial B = B \sqcup B_0$. These compositions $\phi\psi^{-1}$ thus lie in the kernel of $\mathrm{Mod}(B') \to \mathrm{Mod}(B' \backslash \partial B) = \mathrm{Mod}(B) \times \mathrm{Mod}(B_0)$ and therefore consist of Dehn twists about these curves β. Finally, Claim 6.32 $d_\beta(\phi\psi^{-1}(\partial Z_\psi), \partial Z_\psi) = d_\beta(\partial Z_\phi, \partial Z_\psi) \stackrel{+}{\prec}_\Theta 0$ shows that only boundedly many Dehn twists about β arise among these compositions. Therefore pairs $\phi, \psi \in \mathcal{F}$ produce only boundedly many maps $\phi\psi^{-1} \colon B' \to B'$ and we may again partition so that all pairs agree as maps $A' \to B'$. □

We can now prove the main theorem of this section

Proof of Theorem 6.1. Given a pair points a, b, we consider the family $\mathcal{F}_0 \subset [\phi_0]$ consisting of those ϕ for which $a_\phi = a$ and $b_\phi = b$ with $d_{\mathcal{T}(S)}(x_0, \phi(x_0)) \leqslant r$. Let $A_0 = \varnothing$. The family $\mathcal{F}_0$ is trivially A_0–coherent with displacement $r(\mathcal{F}_0) \leqslant r$.

By induction, suppose we are given a subsurface A_k and subcollection $\mathcal{F}_k \subset \mathcal{F}_0$ that is A_k–coherent. If $A_k = S$, then all elements of $\mathcal{F}_k$ agree on S. Otherwise we apply Lemma 6.24 and subdivide $\mathcal{F}_k$ into at most $q_2(r(\mathcal{F}_k)) \leqslant q_2(r)$ subfamilies $\mathcal{F}'_k \subset \mathcal{F}_k$ that are each (A_k, Y'_k)–supercoherent for some domain Y'_k. Now apply Proposition 6.25 to further partition each $\mathcal{F}'_k$ into q_3 subfamilies $\mathcal{F}_{k+1}$ that are each A_{k+1}–coherent, where A_{k+1} is S when Y_k is empty and is otherwise the subsurface filled by A_k and Y_k.

Since each iteration $\mathcal{F}_k \rightsquigarrow \mathcal{F}_{k+1}$ yields coherence on a strictly larger subsurface $A_{k+1} \supsetneq A_k$, any chain $\mathcal{F}_0 \supset \mathcal{F}_1 \supset \dots$ produced in this way must terminate at $A_k = S$ within $k \leqslant \xi(S)$ steps. Since at each step each $\mathcal{F}_k$ is partitioned into at most $q_3 q_2(r)$ subfamilies, this iterative procedure ultimately produces at most $(q_3 q_2(r))^{\xi(S)}$ subfamilies $\mathcal{F}$ that are each S–coherent. By definition of coherence, all elements $\phi, \psi \in \mathcal{F}$ agree on S and thus in fact are equal. Hence in total there are at most $(q_3 q_2(r))^{\xi(S)}$ elements ϕ in the original collection $\mathcal{F}_0$ $\square$

7. Witness families

We now lay the foundation for our main technical construction of the "complexity length" between points of $\mathcal{T}(\Sigma)$. This notion of length is defined in terms of curve complex data of subsurfaces of Σ and relies on building and manipulating collections of subsurfaces with certain properties. In this section we introduce our terminology and establish basic operations and results about such families. The definition of complexity length will then be given in §8.2.

Given any parameter $\mathsf{C} \geqslant 2\mathsf{M}$, we once and for all fix a sequence of constants

$$\xi(S) + 30\mathsf{C}\frac{\epsilon_0}{{\epsilon_0}'} = \mathsf{N}_{\xi(S)+1} \leqslant \mathsf{N}_{\xi(S)} \leqslant \mathsf{N}_{\xi(S)-1} \leqslant \dots \leqslant \mathsf{N}_0 \leqslant \mathsf{N}_{-1} = \mathsf{N}, \tag{7.1}$$

where M and $\epsilon_0 > {\epsilon_0}'$ are from Definition 3.24. Note that $1 = \xi(S_{0,4})$ and $-1 = \xi(\text{annulus})$. The exact value of these constants (along with other related constants) will be specified later in a recursive manner (Proposition 10.13), but we stress that they depend only on C and S. By abuse of notation, for any domain V of S we set $\mathsf{N}_V := \mathsf{N}_{\xi(V)}$ and emphasize that these depend only on the integer $\xi(V)$ and not on the domain V.

Definition 7.2. For Σ a domain in S and $x, y \in \mathcal{T}(\Sigma)$ we consider the following collections of subdomains:

$$\Upsilon^c(x,y) = \{V \sqsubset \Sigma \mid d_V(x,y) \geqslant \mathsf{N}_V\}, \quad \text{and}$$
$$\Upsilon^\ell(x,y) = \{A \sqsubset \Sigma \mid A \text{ an annulus with } \min\{\ell_x(\partial A), \ell_x(\partial A)\} < \epsilon_0/\mathsf{N}_A < {\epsilon_0}'\}.$$

Thus Υ^c consists of those domains with big **c**urve complex distance, and $\Upsilon^\ell(x,y)$ consists of those curves (really annuli) with drastic **l**ength difference at x and y. Note that $|\Upsilon^\ell(x,y)| \leqslant 2\xi(\Sigma)$, since x and y each have at most $\xi(\Sigma)$ curves of length smaller than ϵ_0. We then define

$$\Upsilon(x,y) = \Upsilon^c(x,y) \cup \Upsilon^\ell(x,y)$$

and, when the points x, y are understood, abbreviate these simply as $\Upsilon, \Upsilon^c, \Upsilon^\ell$.

Remark 7.3. Observe that every $V \in \Upsilon$ has a nonempty active interval $\mathcal{I}_V$ along $[x,y]$. If $V \in \Upsilon^c$ this follows form Lemma 3.26(1). If instead $V \in \Upsilon^\ell$ then V is an annulus with at least one of $\ell_x(\partial V)$ or $\ell_y(\partial V)$ smaller than $\frac{1}{\mathsf{N}_V}\epsilon_0 \leqslant \frac{{\epsilon_0}'}{30\mathsf{C}\epsilon_0}\epsilon_0 < {\epsilon_0}'$. Hence $\mathcal{I}_V = \tilde{\mathcal{I}}_V^{\epsilon_0}$ is nonempty by Definition 3.25 and Theorem 3.22(2).

The distance formula (Theorem 3.33) basically says the domains V of $\Upsilon(x,y)$ (along with the projections $\pi_V(*)$ and lengths $\ell_*(\partial V)$ when V is an annulus) account for all of the data needed to estimate $d_{\mathcal{T}(\Sigma)}(x,y)$. However, the *multiplicative* error in the distance formula is unacceptable in our application because $d_{\mathcal{T}(\Sigma)}(x,y)$ goes into the exponent and the whole point is to calculate its coefficient.

Morally, this multiplicative error stems from the fact that the collection $\Upsilon(x,y)$ can be arbitrarily large. The point of witness families, defined next, is to partition $\Upsilon(x,y)$ into *uniformly boundedly many* subcollections (Definition 8.2). The data for each subcollection will then be recombined into a Teichmüller distance (Definition 8.7). If everything is done carefully, the weighted sum of these distances may be related to $d_{\mathcal{T}(\Sigma)}(x,y)$ with *only additive error* (Theorem 11.2).

Definition 7.4 (Witness family)**.** Let Σ be a domain in S. A collection Ω of domains of Σ is called a *witness family* for a geodesic segment $[x,y]$ in $\mathcal{T}(\Sigma)$ if:

(WF1) Every $V \in \Omega$ satisfies $V \in \Upsilon(x,y)$; that is $\Omega \subset \Upsilon(x,y)$.
(WF2) Every $Z \sqsubset \Sigma$ with $Z \in \Upsilon(x,y)$ satisfies $Z \sqsubset V$ for some $V \in \Omega$.
(WF3) If $Z \sqsubset W$ are such that $Z \in \Omega$ and $W \in \Upsilon(x,y)$, then either $W \in \Omega$ or else $W \pitchfork Z'$ for some $Z' \in \Omega$ with $Z \sqsubset Z'$.

7.1. **Supremums for witness families.** We will have need to discuss the minimal subsurfaces in a collection that contain a given subsurface:

Definition 7.5 (Minimal containment)**.** If Ω is a collection of subsurfaces of Σ and Z is an arbitrary subsurface of Σ, we use the notation $Z \preccurlyeq^{\Omega} W$ to mean that W is a minimal (with respect to inclusion) subsurface of Σ satisfying the conditions $Z \sqsubset W$ and $W \in \Omega$. If W is moreover the *unique* element of Ω such that $Z \preccurlyeq^{\Omega} W$, we write $W = \bar{Z}^{\Omega}$ and call W the Ω*–supremum* of Z.

Lemma 7.6. *Let Ω be a witness family for $[x,y]$ and suppose that $W \sqsubset \Sigma$ has an Ω–supremum $\bar{W}^{\Omega} \in \Omega$. Then $\bar{W}^{\Omega} \sqsubset V$ for every $V \in \Omega$ with $W \sqsubset V$.*

Proof. Let $V \in \Omega$ be such that $W \sqsubset V$. Then the family $\{V' \in \Omega \mid W \sqsubset V' \sqsubset V\}$ is nonempty and so contains a topologically minimal element V_0. By minimality, it must be that $W \preccurlyeq^{\Omega} V_0$. The assumed uniqueness of the Ω–supremum now implies that $\bar{W}^{\Omega} = V_0 \sqsubset V$, as claimed. ☐

7.2. **Complete witness families.** In general, a domain Z could satisfy $Z \preccurlyeq^{\Omega} W$ for multiple elements W of a witness family Ω. Our construction of Teichmüller resolutions below will utilize witness families for which every large-projection domain has a Ω–supremum:

Definition 7.7 (Completeness)**.** A witness family Ω for a geodesic segment $[x,y]$ in $\mathcal{T}(\Sigma)$ is said to be *complete* if every domain $Z \sqsubset \Sigma$ with $Z \in \Upsilon(x,y)$ has an Ω–supremum $\bar{Z}^{\Omega} \in \Omega$.

We next describe a criterion for completeness. Suppose that $A, B \sqsubset \Sigma$ are two domains with $A \pitchfork B$. Define $\mathcal{F}(A,B)$ to be the collection

$$\mathcal{F}(A,B) := \{Z \sqsubset \Sigma \mid Z \sqsubset A, Z \sqsubset B \text{ and } Z \in \Upsilon(x,y)\}$$

and let $\underline{\mathcal{F}}(A,B)$ be the subcollection of topologically maximal surfaces in $\mathcal{F}(A,B)$.

Lemma 7.8. *For every geodesic $[x,y]$ in $\mathcal{T}(\Sigma)$ and pair of domains $A, B \sqsubset \Sigma$ with $A \pitchfork B$, one has $|\underline{\mathcal{F}}(A,B)|_j \leqslant (2\mathsf{N}_{j+1})^{\xi(\Sigma)+2}$ for every $-1 \leqslant j \leqslant \xi(\Sigma)$.*

Proof. Consider $\mathcal{F}^c(A,B) = \{Z \sqsubset \Sigma \mid Z \sqsubset A, Z \sqsubset B \text{ and } Z \in \Upsilon^c(x,y)\}$. Applying Lemma 4.1 to the essential intersection $A \sqcap B$ (Lemma 3.3) immediately yields $|\underline{\mathcal{F}}^c(A,B)|_j \leqslant (2\mathsf{N}_{j+1})^{\xi(\Sigma)+1}$ for every $-1 \leqslant j \leqslant \xi(\Sigma)$. Since $\mathcal{F}(A,B) \subset \mathcal{F}^c(A,B) \cup \Upsilon^\ell(x,y)$ and $\Upsilon^\ell(x,y)$ consists of annuli, we see that $\underline{\mathcal{F}}(A,B) \subset \underline{\mathcal{F}}^c(A,B) \cup \Upsilon^\ell(x,y)$ and that the claimed bound follows for all $0 \leqslant j \leqslant \xi(\Sigma)$. For $j = -1$, the fact $\left|\Upsilon^\ell(x,y)\right| \leqslant 2\xi(\Sigma) \leqslant \mathsf{N}_0$ now gives $|\underline{\mathcal{F}}(A,B)|_{-1} \leqslant (2\mathsf{N}_0)^{\xi(\Sigma)+2}$. □

Let Ω be a witness family for $[x,y]$ in $\mathcal{T}(\Sigma)$. A *cutting pair in* Ω is a pair of subsurfaces $A, B \in \Omega$ such that $A \pitchfork B$. We say that the cutting pair (A,B) is *filled* in Ω if $\underline{\mathcal{F}}(A,B) \subset \Omega$ and that the pair is *unfilled* otherwise.

Lemma 7.9 (Filled to completeness)**.** *A witness family Ω for $[x,y]$ in $\mathcal{T}(\Sigma)$ is complete if and only if every cutting pair (A,B) in Ω is filled.*

Proof. First suppose every cutting pair in Ω is filled. Let $Z \sqsubset \Sigma$ be any domain with $Z \in \Upsilon(x,y)$. Condition (WF2) ensures there exists some $A \in \Omega$ with $Z \preccurlyeq^\Omega A$. Thus if Z fails to have an Ω–supremum, there is a second domain $B \in \Omega$ with $B \neq A$ and $Z \preccurlyeq^\Omega B$. By minimality, A and B cannot be nested and so we have $A \pitchfork B$. It follows that (A,B) is a cutting pair and that $Z \in \mathcal{F}(A,B)$. Therefore, by definition of $\underline{\mathcal{F}}(A,B)$, we have $Z \sqsubset Z_0$ for some $Z_0 \in \underline{\mathcal{F}}(A,B)$. The hypothesis that (A,B) is filled now implies that $Z_0 \in \Omega$. But since $Z \sqsubset Z_0 \sqsubset A, B$, this contradicts the minimality of $Z \preccurlyeq^\Omega A$ and $Z \preccurlyeq^\Omega B$. Therefore $\bar{Z}^\Omega$ exists.

Next suppose Ω is complete. Let (A,B) be a cutting pair in Ω and choose any $Z \in \underline{\mathcal{F}}(A,B)$. We must show $Z \in \Omega$. Since $Z \sqsubset A$, the Ω–supremum necessarily satisfies $\bar{Z}^\Omega \sqsubset A$ by Lemma 7.6. Similarly $\bar{Z}^\Omega \sqsubset B$. Therefore $\bar{Z}^\Omega \in \mathcal{F}(A,B)$ by definition. Since Z is a topologically maximal element of $\mathcal{F}(A,B)$ and $Z \sqsubset \bar{Z}^\Omega$ by definition, it follows that $Z = \bar{Z}^\Omega \in \Omega$. □

7.3. **Insulation.** Let $[x,y]$ be a Teichmüller geodesic in $\mathcal{T}(\Sigma)$. If $Z, V \sqsubset \Sigma$ are two domains with $Z \sqsubset V$ and $Z \neq V$, define

$$\mathcal{C}(V|_Z) = \{\alpha \in \mathcal{C}(V) \mid \alpha \text{ is essential or peripheral in } Z\} = \Gamma(Z) \cup \partial Z.$$

Observe that $\mathcal{C}(V|_Z)$ has diameter at most 2 in $\mathcal{C}(V)$. For any domain $E \sqsubset \Sigma$ and parameter $0 \leqslant t \leqslant d_E(x,y)$, we then define $\underline{\mathcal{L}}_t(E)$ and $\underline{\mathcal{R}}_t(E)$ to be the topologically maximal domains in the respective collections

$$\mathcal{L}_t(E) := \left\{Z \subsetneq E \mid Z \in \Upsilon(x,y) \text{ and } \exists \alpha \in \mathcal{C}(E|_Z) : d_E(\alpha, x) \in [t - 9\mathsf{C}, t + 9\mathsf{C}]\right\}$$

$$\mathcal{R}_t(E) := \left\{Z \subsetneq E \mid Z \in \Upsilon(x,y) \text{ and } \exists \alpha \in \mathcal{C}(E|_Z) : d_E(\alpha, y) \in [t - 9\mathsf{C}, t + 9\mathsf{C}]\right\},$$

where here $d_E(\alpha,x) = \operatorname{diam}_{\mathcal{C}(E)}(\pi_E(x) \cup \{\alpha\})$ and similarly for $d_E(\alpha,y)$. Note that by construction $\mathcal{L}_t(A) = \varnothing = \mathcal{R}_t(A)$ for any annulus A.

Lemma 7.10. *For every geodesic $[x,y]$ in $\mathcal{T}(\Sigma)$, domain $E \sqsubset \Sigma$, and parameter t with $0 \leqslant t \leqslant d_E(x,y) \geqslant \mathsf{N}_E$, one has $|\underline{\mathcal{L}}_t(E)|_j, |\underline{\mathcal{R}}_t(E)|_j \leqslant (2\mathsf{N}_{j+1})^{\xi(\Sigma)+3}$ for every $-1 \leqslant j \leqslant \xi(\Sigma)$.*

Proof. We give the proof for $\mathcal{L}_t(E)$: As in the proof of Lemma 7.8, let

$$\mathcal{L}_t^c(E) = \left\{Z \subsetneq E \mid Z \in \Upsilon^c(x,y) \text{ and } \exists \alpha \in \mathcal{C}(E|_Z) : d_E(\alpha, x) \in [t - 9\mathsf{C}, t + 9\mathsf{C}]\right\}$$

and observe that $\underline{\mathcal{L}}_t(E) \subset \underline{\mathcal{L}}_t^c(E) \cup \Upsilon^\ell(x,y)$. Since $\left|\Upsilon^\ell(x,y)\right| \leqslant 2\xi(\Sigma) \leqslant \mathsf{N}_{j+1}$, it therefore suffices to prove that $|\underline{\mathcal{L}}_t^c(E)|_j \leqslant (2\mathsf{N}_{j+1})^{\xi(\Sigma)+2}$ for all $-1 \leqslant j \leqslant \xi(\Sigma)$.

Choose $\alpha \in \pi_E(x)$ and $\beta \in \pi_E(y)$ realizing the distance $d_E(x,y) = d_{\mathcal{C}(E)}(\alpha,\beta)$ and fix a geodesic $\alpha = \gamma_0, \ldots, \gamma_m = \beta$ in $\mathcal{C}(E)$. As in the proof of Lemma 4.1, we claim that every $Z \subsetneq E$ with $d_Z(x, y) \geqslant \mathsf{N}_Z$ is disjoint from some curve γ_i. Indeed, this is immediate if α or β misses Z, and otherwise Lemma 3.9 ensures $d_{\mathcal{C}(Z)}(\pi_Z(\alpha), \pi_Z(\beta)) \geqslant d_Z(x, y) - 2\mathsf{k} \geqslant \mathsf{M} - 2\mathsf{k}$ so that the Bounded Geodesic Image Theorem implies $\pi_Z(\gamma_i) = \varnothing$ for some i.

Thus every $Z \in \mathcal{L}^c_t(E)$ is disjoint from some γ_i. Further, from the definition of $\mathcal{L}^c_t(E)$, we see that this curve γ_i must satisfy $d_E(\gamma_i, x) \in [t - 9\mathsf{C} - 1, t + 9\mathsf{C} + 1]$. Since $\operatorname{diam}_{\mathcal{C}(E)}(\pi_E(x)) \leqslant \mathsf{L}$, this implies $d_{\mathcal{C}(E)}(\alpha, \gamma_i) \in [t - 9\mathsf{C} - 2\mathsf{L}, t + 9\mathsf{C} + 2\mathsf{L}]$. Letting $\mathcal{W}$ denote the set of all components of $E \backslash \gamma_i$ obtained as i ranges between $\max\{0, t - 9\mathsf{C} - 2\mathsf{L}\}$ and $\min\{m, t + 9\mathsf{C} + 2\mathsf{L}\}$, it follows that

$$\mathcal{L}^c_t(E) \subset \bigcup_{W \in \mathcal{W}} \mathcal{P}(W),$$

where $\mathcal{P}(W)$ is as in Lemma 4.1. Now let $Z \in \underline{\mathcal{L}}^c_t(E)$ be a topologically maximal element of $\mathcal{L}^c_t(E)$ and choose $W \in \mathcal{W}$ such that $Z \in \mathcal{P}(W)$. If $Z \subsetneq V$ for some $V \in \mathcal{P}(W)$, then the facts $d_V(x, y) \geqslant \mathsf{N}_V$ and $Z \sqsubset V \subsetneq E$ with $Z \in \mathcal{L}^c_t(E)$ imply that $V \in \mathcal{L}^c_t(E)$ as well. But this contradicts the maximality of Z in $\mathcal{L}^c_t(E)$. Therefore Z is maximal in $\mathcal{P}(W)$ as well. This proves $\underline{\mathcal{L}}^c_t(E) \subset \bigcup_W \underline{\mathcal{P}}(W)$. We may now invoke Lemma 4.1 to conclude

$$|\underline{\mathcal{L}}^c_t(E)|_j \leqslant \sum_{W \in \mathcal{W}} |\underline{P}(W)|_j \leqslant |\mathcal{W}| \, (2\mathsf{N}_{j+1})^{\xi(\Sigma)+1} \leqslant 2(18\mathsf{C} + 5\mathsf{L})(2\mathsf{N}_{j+1})^{\xi(\Sigma)+1}$$

for every j. Since $18\mathsf{C} + 5\mathsf{L} < 23\mathsf{C} < \mathsf{N}_{j+1}$, the lemma follows. ∎

Definition 7.11 (Insulation)**.** If Ω is a witness family for $[x, y]$ in $\mathcal{T}(\Sigma)$, we say that $E \in \Omega$ is *insulated* in Ω if $\underline{\mathcal{L}}_0(E) \cup \underline{\mathcal{R}}_0(E) \subset \Omega$. The witness family Ω is said to be *insulated* if every $E \in \Omega$ is insulated in Ω.

The terminology stems from the observation that if E is insulated in Ω, then for every domain $Z \in \Upsilon(x, y)$ with $Z \lessdot^{\Omega} E$, ∂Z occurs towards the middle or "interior" of the geodesic $[\pi_E(x), \pi_E(y)]$ in $\mathcal{C}(E)$, rather than near the endpoints. This has the following useful consequence:

Lemma 7.12. *Let Ω be witness family for $[x, y]$ in $\mathcal{T}(\Sigma)$.*

(1) Suppose $V \in \Omega$ is insulated in Ω. If $Z \sqsubset \Sigma$ is such that $Z \lessdot^{\Omega} V$ with $Z \in \Upsilon(x, y)$, then the active intervals of Z and V satisfy $\mathcal{I}_Z \subset \mathcal{I}_V$.

(2) Suppose $V \in \Omega$ is insulated in Ω. If $Z, W \sqsubset \Sigma$ are such that $Z \lessdot^{\Omega} V$ with $Z \in \Upsilon(x, y)$ and W, V time-ordered along $[x, y]$, then $W \pitchfork Z$.

(3) Suppose Ω is insulated. If $V_1, V_2 \in \Omega$ are such that $V_1 \pitchfork V_2$ and $Z_1, Z_2 \sqsubset \Sigma$ are such that $Z_i \in \Upsilon(x, y)$ and $Z_i \lessdot^{\Omega} V_i$ for $i = 1, 2$, then $Z_1 \pitchfork Z_2$.

Proof. Suppose, contrary to (1), that $z \notin \mathcal{I}_V$ for some point $z \in \mathcal{I}_Z$. Then z lies in the same component of $[x, y] \backslash \mathcal{I}_V$ as either x or y. Without loss of generality, say x an z lie in the same component. Then $d_V(x, z) \leqslant \mathsf{M}/3$ by Lemma 3.26(3). Since the Bers marking at z contains ∂Z by Lemma 3.26(2), it follows that $d_V(x, \partial Z) \leqslant \mathsf{M}/3$ and thus that $Z \in \mathcal{L}_0(V)$ by definition. Hence $Z \sqsubset Z' \subsetneq V$ for some $Z' \in \underline{\mathcal{L}}_0(V)$. But then $Z' \in \Omega$ by the insulation of V, contradicting $Z \lessdot^{\Omega} V$.

We next prove (2): If W and Z do not cut, then ∂W and ∂Z are disjoint so that $d_V(\mathcal{C}(V|_Z), \partial W) \leqslant 2 + d_V(\partial W, \partial Z) \leqslant 3 < \mathsf{M}/2$. The fact that W and V are time-ordered implies $\min\{d_V(\partial W, x), d_V(\partial W, y)\} < \mathsf{M}/3$ by Lemma 3.29. Therefore $Z \in \mathcal{L}_0(V) \cup \mathcal{R}_0(V)$ and hence $Z \sqsubset Z'$ for some $Z' \in \underline{\mathcal{L}}_0(V) \cup \underline{\mathcal{R}}_0(V)$. Since V is insulated in Ω, it follows that $Z' \in \Omega$, contradicting $Z \lessdot^{\Omega} V$.

Conclusion (3) now follows easily. Since V_1 is insulated in Ω by hypothesis, (2) implies $Z_1 \pitchfork V_2$. Applying (2) again to the insulated $V_2 \in \Omega$, we conclude $Z_1 \pitchfork Z_2$. $\square$

We also note the following useful observation:

Lemma 7.13. *If Ω is an insulated witness family for $[x, y]$, then $\Upsilon^\ell(x, y) \subset \Omega$.*

Proof. Consider any annulus $A \in \Upsilon^\ell(x, y)$. Then either $\ell_x(\partial A) < \epsilon_0$ or $\ell_y(\partial A) < \epsilon_0$; by symmetry, let us suppose it is the former. By (WF2), there exists some $Z \in \Omega$ with $A \preccurlyeq^\Omega Z$. If $A = Z$ we are done. Otherwise ∂A is an essential curve in Z, and the fact $\ell_x(\partial A) < \epsilon_0$ implies that $d_Z(x, \partial A) \leqslant \mathsf{L} \leqslant \mathsf{M}$. Thus $A \in \mathcal{L}_0(Z)$ and hence $A \sqsubset Z'$ for some $Z' \in \underline{\mathcal{L}}_0(Z)$. Since Ω is insulated, we have $Z' \in \Omega$. But now the containments $A \sqsubset Z' \sqsubsetneq Z$ contradict $A \preccurlyeq^\Omega Z$. $\square$

7.4. **Subordered witness families.** To construct our complexity length, we will work with witness families that come equipped with the following structure:

Definition 7.14 (Subordering)**.** Let Ω be a witness family for the segment $[x, y]$ in $\mathcal{T}(\Sigma)$. A *subordering* on Ω is an ordering designation $Z \swarrow V$ exclusive or $V \searrow Z$ for every $Z, V \in \Omega$ with $Z \sqsubsetneq V$. This ordering data must satisfy:

(SO1) If $Z, W, V \in \Omega$ are such that $Z \sqsubsetneq W \sqsubsetneq V$ then
$$Z \swarrow V \iff W \swarrow V \quad \text{(and so } V \searrow Z \iff V \searrow W \text{ also).}$$
(SO2) If $Z, W, V \in \Omega$ are such that $Z \swarrow V \searrow W$, then $Z \pitchfork_V W$ and $Z \lessdot W$.
(SO3) If $Z, V \in \Omega$ and $W \sqsubset \Sigma$ with $W \in \Upsilon(x, y)$ are such that $Z \swarrow V \lessdot W$ or $W \lessdot V \searrow Z$, then $Z \pitchfork_V W$.
(SO4) If $Z, V \in \Omega$ are such that $Z \swarrow V$ (resp. $V \searrow Z$), then there does not exist any domain $W \in \Upsilon(x, y)$ with $\bar{W}^\Omega = V$, and $W \lessdot Z$ (resp. $Z \lessdot W$).

Remark 7.15. Condition (SO2) in fact implies condition (SO1), as can be seen by noting that if $Z \sqsubsetneq W \sqsubsetneq V$, then $Z \pitchfork_V W$ clearly fails so that $Z \swarrow V \searrow W$ and $W \swarrow V \searrow Z$ must both fail as well.

If Ω and Ω' are two subordered witness families with $\Omega \subset \Omega'$, we say that the subordering on Ω' *extends* the subordering on Ω if the ordering designations coming from Ω and Ω' agree on each pair $Z, V \in \Omega$ with $Z \sqsubsetneq V$. Subordered witness families enjoy the following property:

Lemma 7.16. *Let Ω be a subordered witness family for a geodesic $[x, y]$ in $\mathcal{T}(\Sigma)$. Fix a domain $V \in \Omega$ and let $W \sqsubsetneq V$ be a domain with $W \in \Upsilon(x, y)$ and $\bar{W}^\Omega = V$. Then for any $Z \in \Omega$ with $Z \swarrow V$ (resp. $V \searrow Z$) we have that either Z and W are disjoint, or $\mathcal{I}_Z$ occurs before (resp. after) $\mathcal{I}_W$.*

Proof. We only consider the case $Z \swarrow V$. If Z and W are disjoint, the lemma is satisfied. If $Z \pitchfork W$, then (SO4) ensures we have the time ordering $Z \lessdot W$ so that $\mathcal{I}_Z$ occurs before $\mathcal{I}_W$ as required. The possibility $W \sqsubset Z$ is ruled out by $\bar{W}^\Omega = V$, so it remains to consider the case $Z \sqsubsetneq W$.

Here, since $W \notin \Omega$, (WF3) provides $Z' \in \Omega$ so that $Z \sqsubset Z'$ and $Z' \pitchfork W$. Either $Z' \pitchfork V$, in which case (SO3) (applied to $Z \swarrow V$ and $Z \sqsubset Z'$) forces $Z' \lessdot V$ and hence $Z' \lessdot W$ by Corollary 3.31. Otherwise $Z' \sqsubset V$ and (SO1) gives $Z' \swarrow V$ so that we may invoke (SO4) (using $Z' \pitchfork W$) to again conclude $Z' \lessdot W$.

Since $\mathcal{I}_{Z'}$ is nonempty with $Z' \pitchfork W$, Lemma 3.26(4) now implies $\mathcal{I}_Z \cap \mathcal{I}_W = \varnothing$. In fact, since $Z' \lessdot W$ it must be that $\mathcal{I}_Z$ occurs before $\mathcal{I}_W$ along $[x, y]$. $\square$

Definition 7.17 (Encroachment)**.** Let Ω be a subordered witness family for $[x,y]$ in $\mathcal{T}(\Sigma)$. For $V \in \Omega$, the *left and right encroachments* of V in Ω are defined as

$$\mathcal{E}^{\ell}_{\Omega}(V) := \sup_{Z\in\Omega, Z\swarrow V} d_V(x, \mathcal{C}(V|_Z)) \quad \text{and} \quad \mathcal{E}^{r}_{\Omega}(V) := \sup_{Z\in\Omega, V\searrow Z} d_V(y, \mathcal{C}(V|_Z)),$$

respectively (where here $d_V(w, \mathcal{C}(V|_Z)) = \operatorname{diam}_{\mathcal{C}(V)}(\pi_V(w) \cup \mathcal{C}(V|Z)))$. The *encroachment* of V is then defined as $\mathcal{E}_{\Omega}(V) = \max\{\mathcal{E}^{\ell}_{\Omega}(V), \mathcal{E}^{r}_{\Omega}(V)\}$. To streamline notation, for each $W \notin \Omega$ these encroachments are set to zero:

$$\mathcal{E}^{\ell}_{\Omega}(W) = \mathcal{E}^{r}_{\Omega}(W) = \mathcal{E}_{\Omega}(W) = 0 \quad \text{when} \quad W \notin \Omega.$$

Definition 7.18 (Wide)**.** A subordered witness family Ω is *wide* if $\mathcal{E}_{\Omega}(V) \leqslant \mathsf{N}_V/3$ for all $V \in \Omega$.

We next describe two operations—refinement and augmentation—that may be used to enlarge a witness family and ultimately produce one that is both complete and insulated. For each operation, we must work to show that suborderings may be naturally extended to the new family.

7.5. **Refinement.** Lemma 7.9 suggests a means of making any witness family complete: simply add collections of the form $\underline{\mathcal{F}}(A,B)$ until every cutting pair is filled. This motivates the following operation:

Definition 7.19 (Refinement)**.** Let Ω be a witness family for $[x,y]$ and let (A,B) be a cutting pair in Ω. The *refinement of* Ω *along* (A,B) is the collection

$$\underline{\Omega}(A,B) := \Omega \cup \underline{\mathcal{F}}(A,B).$$

Thus $\underline{\Omega}(A,B) = \Omega$ if and only if (A,B) is a filled cutting pair in Ω. Fortunately refinement always produces a new witness family.

Lemma 7.20. *Let Ω be a witness family for $[x,y]$ in $\mathcal{T}(S)$ and let (A,B) be a cutting pair in Ω. Then the refinement $\underline{\Omega}(A,B)$ is a witness family for $[x,y]$ and the pair (A,B) is filled in $\underline{\Omega}(A,B)$.*

Proof. It is obvious that (A,B) is filled in $\underline{\Omega}(A,B)$, provided that $\underline{\Omega}(A,B)$ is a witness family. For this, conditions (WF1) and (WF2) are immediate; we verify (WF3). Let $Z \sqsubset W$ be such that $Z \in \underline{\Omega}(A,B)$ and $W \in \Upsilon(x,y)$. If $Z \in \Omega$, then W satisfies condition (WF3) because $\Omega \subset \underline{\Omega}(A,B)$ is a witness family. Therefore we may suppose $Z \notin \Omega$ so that $Z \in \underline{\mathcal{F}}(A,B)$. If $W \sqsubset A$ and $W \sqsubset B$, then the fact $W \in \Upsilon(x,y)$ implies $W \in \mathcal{F}(A,B)$ so that $W = Z \in \underline{\Omega}(A,B)$ by maximality of Z. If W cuts A or B, then we have verified (WF3) since A and B lie in $\underline{\Omega}(A,B)$ and contain Z. Therefore, let us suppose W cuts neither A nor B and that $[W \sqsubset A$ and $W \sqsubset B]$ fails. In this case we must have $A, B \sqsubset W$. But now (WF3), applied to $A \in \Omega$, implies that either $W \in \Omega \subset \underline{\Omega}(A,B)$ or else that $W \pitchfork Z'$ for some $Z' \in \Omega \subset \underline{\Omega}(A,B)$ with $Z' \sqsupset A \sqsupset Z$. This proves that $\underline{\Omega}(A,B)$ satisfies condition (WF3) and establishes the lemma. ∎

Furthermore, suborderings always extend to refinements:

Lemma 7.21 (Subordering refinements)**.** *If Ω is a subordered witness family for $[x,y]$ in $\mathcal{T}(\Sigma)$ and (A,B) is a cutting pair in Ω, then there is a unique subordering on the refinement $\underline{\Omega}(A,B)$ that extends the subordering on Ω.*

Proof. To simplify notation, write $\underline{\Omega} = \underline{\Omega}(A, B)$ and without loss of generality suppose $A \lessdot B$. We first show that conditions (SO1)–(SO4) uniquely determine a well-defined ordering designation $\swarrow\!\!\!\nearrow$ or $\nwarrow\!\!\!\searrow$ for each pair $Z, V \in \underline{\Omega}$ with $Z \sqsubsetneq V$:

(1) If $Z, V \in \Omega$ we must use the ordering designation from Ω: $Z \swarrow\!\!\!\nearrow V$ iff $Z \swarrow V$.
(2) If $V \in \Omega$ and $Z \notin \Omega$, then $Z \in \underline{\mathcal{F}}(A, B)$ so that (SO3) forces us to set $A \nwarrow\!\!\!\searrow Z \swarrow\!\!\!\nearrow B$ (since $Z \pitchfork_A B$ and $A \pitchfork_B Z$ both fail). We claim that exactly one of the following hold: (i) $A, B \sqsubset V$, (ii) $A \lessdot V$, or (iii) $V \lessdot B$. Firstly, it is impossible to have $V \sqsubset A, B$, as that would contradict the maximality of Z in $\mathcal{F}(A, B)$. Triple nesting $A \sqsubset V \sqsubset B$ or $B \sqsubset V \sqsubset A$ is also ruled out by $A \pitchfork B$. Thus if (i) fails, then V necessarily cuts A or B. If $A \pitchfork V$ then we must have $A \lessdot V$ along $[x, y]$, for otherwise we have $V \lessdot A \lessdot B$ and (by Corollary 3.30) $V \pitchfork_A B$, contradicting $Z \sqsubset V, A, B$. Similarly, if $V \pitchfork B$ then $V \lessdot B$ and we are in case (iii). To prove the claim, it remains to show (ii) and (iii) are mutually exclusive; but this is clear: since $A \pitchfork_V B$ fails (as $Z \sqsubset A, V, B$), Corollary 3.30 precludes $A \lessdot V \lessdot B$. We now suborder Z and V in each case:
   (i) If $A, B \sqsubset V$, then (SO2) implies $A \swarrow V \iff B \swarrow V$. In accordance with (SO1), we thus set $Z \swarrow\!\!\!\nearrow V$ in the case that $A \swarrow V$ and $B \swarrow V$ and set $V \nwarrow\!\!\!\searrow Z$ in the case that $V \searrow A$ and $V \searrow B$.
   (ii) If $A \lessdot V$ we declare $Z \swarrow\!\!\!\nearrow V$ in accordance with (SO3), since $\neg(A \pitchfork_V Z)$.
   (iii) If $V \lessdot B$ we similarly declare $V \nwarrow\!\!\!\searrow Z$.
(3) If $Z \in \Omega$ and $V \notin \Omega$, then $V \in \underline{\mathcal{F}}(A, B)$. Here (WF3) implies that $V \pitchfork Z'$ for some $Z' \in \Omega$ with $Z \sqsubset Z'$. Since $\neg(Z' \pitchfork_V Z)$, we thus declare $Z \swarrow\!\!\!\nearrow V$ if $Z' \lessdot V$ and $V \nwarrow\!\!\!\searrow Z$ if $V \lessdot Z'$ in accordance with (SO3). Note that this is well-defined: If $Z'_0 \in \Omega$ is any other such domain, $Z' \lessdot V \lessdot Z'_0$ is ruled out by Corollary 3.30 and the fact $\neg(Z' \pitchfork_V Z'_0)$.
(4) If $Z, V \notin \Omega$ then $Z, V \in \underline{\mathcal{F}}(A, B)$, contradicting the fact that no two surfaces in $\underline{\mathcal{F}}(A, B)$ can be properly nested. Therefore this case does not occur.

We have now established ordering designations on $\underline{\Omega}$ that extend those of Ω. We henceforth use $\swarrow$ and $\searrow$ for these orderings in both Ω and $\underline{\Omega}$, as the meaning is unambiguous. It remains to show these in fact give a subordering on $\underline{\Omega}$:

Condition (SO2): Let $Z, W, V \in \underline{\Omega}$ be such that $Z \swarrow V \searrow W$. If all three domains are in Ω the condition is clear. Suppose $V \notin \Omega$. Then $Z, W \in \Omega$ since no pair of domains in $\underline{\mathcal{F}}(A, B)$ are nested. By case (3) above, there must exist $Z', W' \in \Omega$ such that $Z \sqsubset Z'$, $W \sqsubset W'$, and $Z' \lessdot V \lessdot W'$. Corollaries 3.30–3.31 now imply $Z' \pitchfork_V W'$ and consequently $Z \pitchfork_V W$ and $Z \lessdot W$.

Next suppose $V \in \Omega$. Observe that if $Z, W \notin \Omega$, then case (2) above dictates that orderings for $Z \sqsubsetneq V$ and $W \sqsubsetneq V$ are both determined solely by the relationship of V to A and B so that in fact $Z \swarrow V \iff W \swarrow V$. As this is not the case, we see that at most one of Z, W can lie outside of Ω. By symmetry, let us suppose that $Z \notin \Omega$ (so that $Z \in \underline{\mathcal{F}}(A, B)$) and $W \in \Omega$. Since $Z \swarrow V$, the definition in (2) dictates that either $A, B \swarrow V$, or else $A \lessdot V$. If $A, B \swarrow V$, then we have $B \swarrow V \searrow W$ with $B, W, V \in \Omega$ so that (SO2) for Ω implies $B \lessdot W$ with $B \pitchfork_V W$. If $A, B \swarrow V$ fails, then we have $A \lessdot V \searrow W$ so that (SO3) for Ω implies $A \pitchfork_V W$. Since $Z \sqsubset A, B$, either of these outcomes implies $Z \pitchfork_V W$. Using Corollary 3.31, we may further conclude $Z \lessdot W$. Thus condition (SO2) is satisfied in this case.

Condition (SO1): This follows from the above and Remark 7.15.

Condition (SO3): Let $Z, V \in \underline{\Omega}$ and $W \in \Upsilon(x,y)$ be such that $Z \swarrow V \lessdot W$ (the case $W \lessdot V \searrow Z$ is similar). We must show $Z \pitchfork_V W$. As above, this is clear if $Z, V \in \Omega$, and at most one of Z or V can lie outside of Ω. Suppose first that $Z \notin \Omega$, so that $Z \in \underline{\mathcal{F}}(A,B)$, and $\Omega \in V$. Let us first consider the case (2i) above in which $A, B \sqsubset V$. Since $Z \swarrow V$, the definition dictates that $A \swarrow V$ as well. Therefore we may apply (SO3) for Ω to $A \swarrow V \lessdot W$ and conclude that $A \pitchfork_V W$. If we are not in case (2i), then (since $Z \swarrow V$) we must be in case (2ii) with $A \lessdot V$. Therefore we have $A \lessdot V \lessdot W$ and Corollary 3.30 gives $A \pitchfork_V W$. In either case, since relative cutting descends to the subsurface $Z \sqsubset A$, we may conclude $Z \pitchfork_V W$ as desired.

It remains to suppose that $Z \in \Omega$ and $V \notin \Omega$. Now case (3) dictates that there is a domain $Z' \in \Omega$ with $Z \sqsubset Z'$ and $Z' \lessdot V$. Therefore we have $Z \sqsubset Z' \lessdot V \lessdot W$ and may again conclude $Z' \pitchfork_V W$ and consequently $Z \pitchfork_V W$. This proves that $\underline{\Omega}$ satisfies condition (SO3).

Condition (SO4): Let $Z, V \in \underline{\Omega}$ be such that $Z \swarrow V$ and suppose that $W \in \Upsilon(x,y)$ satisfies $\bar{W}^{\underline{\Omega}} = V$. We show that $W \not\lessdot Z$. (The case $V \searrow Z$ is similar). This is clear if Z and W are disjoint or nested, so we may assume $Z \pitchfork W$. Note that this gives $W \neq V$ and, consequently $W \notin \underline{\Omega}$. As before, it suffices to suppose that exactly one of Z or V lies in Ω.

First suppose $V \in \Omega$ and $Z \notin \Omega$. We claim that the facts $V = \bar{W}^{\underline{\Omega}}$ and $\Omega \subset \underline{\Omega}$ imply $V = \bar{W}^{\Omega}$ as well. Indeed, if $V_0 \in \Omega$ is any domain with $W \sqsubset V_0$, then $V_0 \in \underline{\Omega}$ so that $V \sqsubset V_0$ by Lemma 7.6. Whence $V = \bar{W}^{\Omega}$ as claimed. Since $Z \notin \Omega$, we have $Z \in \underline{\mathcal{F}}(A,B)$ with $Z \swarrow V$ so that, by the definition in (2), either $A, B \sqsubset V$ with $B \swarrow V$, or else $A \lessdot V$. First consider the former case $A, B \sqsubset V$ with $B \swarrow V$. Since $W \lessdot^{\Omega} V$ and $B \in \Omega$ with $B \subsetneq V$, it cannot be that $W \sqsubset B$. Also we cannot have $B \sqsubset W$ or B disjoint from W because $W \pitchfork Z$. Therefore $W \pitchfork B$. Since $B, V \in \Omega$ with $B \swarrow V$ and $\bar{W}^{\Omega} = V$, we can now invoke (SO4) for Ω to conclude $B \lessdot W$. The desired time-ordering $Z \lessdot W$ now follows from Corollary 3.31. Next consider the latter case $A \lessdot V$. Now, A and W cannot be disjoint nor can $A \sqsubset W$ because $Z \pitchfork W$. If $W \pitchfork A$, then $A \lessdot V$ implies $A \lessdot W$ and $Z \lessdot W$ by Corollary 3.31. So it remains to suppose $W \sqsubset A \in \Omega$; but here Lemma 7.6 implies $V = \bar{W}^{\Omega} \sqsubset A$ contradicting $A \pitchfork V$. This proves that (SO4) holds when $V \in \Omega$ and $Z \notin \Omega$.

Next suppose $V \notin \Omega$ and $Z \in \Omega$. Since $Z \swarrow V$, the definition in (3) provides some $Z' \in \Omega$ such that $Z \sqsubset Z'$ and $Z' \lessdot V$. If $W \sqsubset Z'$, then since $Z' \in \Omega \subset \underline{\Omega}$, Lemma 7.6 implies that $V = \bar{W}^{\underline{\Omega}} \sqsubset Z'$, contradicting $Z' \pitchfork V$. Therefore $W \not\sqsubset Z'$. Neither can we have $Z' \sqsubset W$ or Z' and W disjoint (since $Z \pitchfork W$). Therefore $Z' \pitchfork W$ and we may invoke Corollary 3.31 to conclude $Z' \lessdot W$ and subsequently $Z \lessdot W$, as desired. This proves that (SO4) holds when $V \notin \Omega$ and $Z \in \Omega$ and completes the proof of Lemma 7.21. $\square$

One may now ask how encroachments in Ω and $\underline{\Omega}(A,B)$ are related:

Lemma 7.22 (Refined encroachments)**.** *Let Ω be a subordered witness family for $[x,y]$ in $\mathcal{T}(\Sigma)$ and let $\underline{\Omega} = \underline{\Omega}(A,B)$ be the subordered refinement along the the cutting pair (A,B). Then every domain $V \sqsubset \Sigma$ satisfies*

$$\mathcal{E}^{\ell}_{\underline{\Omega}}(V) \leqslant \max\{\mathcal{E}^{\ell}_{\Omega}(V), \mathsf{M}\} \quad \textit{and} \quad \mathcal{E}^{r}_{\underline{\Omega}}(V) \leqslant \max\{\mathcal{E}^{r}_{\Omega}(V), \mathsf{M}\}.$$

Proof. First suppose $V \notin \Omega$. As the claim is immediate for $V \notin \underline{\Omega}$, we assume $V \in \underline{\Omega} \backslash \Omega$ so that $\mathcal{E}_{\Omega}(V) = 0$. An examination of the proof of Lemma 7.21 shows that any $Z \in \underline{\Omega}$ with $Z \subsetneq V$ falls under case (3) and thus satisfies $Z \sqsubset Z'$ for some $Z' \in \Omega$ with $Z' \pitchfork V$. If $Z' \lessdot V$, so that $Z \swarrow V$, it follows that $d_V(x, \mathcal{C}(V|_Z)) \leqslant$

$1+d_V(x,\partial Z') < \mathsf{M}$ by Lemma 3.29. If instead $V \lessdot Z'$, so that $V \searrow Z$, we similarly have $d_V(y,\mathcal{C}(V|_Z)) < \mathsf{M}$. Thus $\mathcal{E}_{\underline{\Omega}}(V) \leqslant \mathsf{M}$ and the claim follows.

Next suppose $V \in \Omega$. Now the proof of Lemma 7.21 shows that every $Z \in \underline{\Omega}$ with $Z \swarrow V$ either satisfies $Z \in \Omega$, or else falls under case (2i) with $Z \sqsubset A \swarrow V$ or case (2ii) with $Z \sqsubset A \lessdot V$. In the former case we have $d_V(x,\mathcal{C}(V|_Z)) \leqslant \mathcal{E}^{\ell}_{\Omega}(V)$ by definition, and in the latter case we have $d_V(x,\mathcal{C}(V|_Z)) < \mathsf{M}$ as in the previous paragraph. For the middle case, we simply note that $\mathcal{C}(V|_Z) \subset \mathcal{C}(V|_A)$ and thus that $d_V(x,\mathcal{C}(V|_Z)) \leqslant d_V(x,\mathcal{C}(V|_A)) \leqslant \mathcal{E}^{\ell}_{\Omega}(V)$ by definition. This proves $\mathcal{E}^{\ell}_{\underline{\Omega}}(V) \leqslant \max\{\mathcal{E}^{\ell}_{\Omega}(V),\mathsf{M}\}$; the proof for $\mathcal{E}^{r}_{\underline{\Omega}}(V)$ is similar. □

7.6. **Augmentation.** We will repeatedly need to enlarge witness families Ω by adding sets of the form $\underline{\mathcal{L}}_t(E)$ or $\underline{\mathcal{R}}_t(E)$ (see §7.3) for $E \in \Omega$:

Definition 7.23 (Augmentation)**.** Let Ω be a witness family for $[x,y] \in \mathcal{T}(\Sigma)$. For any $E \in \Omega$ and $0 \leqslant t \leqslant d_E(x,y)$, the collections $\Omega \cup \underline{\mathcal{L}}_t(E)$ and $\Omega \cup \underline{\mathcal{R}}_t(E)$ are termed the *left* and *right augmentations* of Ω along E with parameter t.

Lemma 7.24. *If Ω is a witness family for $[x,y]$ in $\mathcal{T}(\Sigma)$ and $E \in \Omega$, then $\Omega\cup\underline{\mathcal{L}}_t(E)$ and $\Omega \cup \underline{\mathcal{R}}_t(E)$ are witness families for each $0 \leqslant t \leqslant d_E(x,y)$.*

Proof. We prove the claim for $\Omega' = \Omega \cup \underline{\mathcal{L}}_t(E)$; the proof for $\Omega \cup \underline{\mathcal{R}}_t(E)$ is identical. Conditions (WF1) and (WF2) are clear because $\Omega \subset \Omega'$ and each $Z \in \underline{\mathcal{L}}_t(E)$ satisfies $Z \in \Upsilon(x,y)$ by definition. For condition (WF3), suppose $Z \sqsubset W$ are such that $Z \in \Omega'$ and $W \in \Upsilon(x,y)$; we must show $W \in \Omega'$ or else $W \pitchfork Z'$ for some $Z' \in \Omega'$ with $Z \sqsubset Z'$. If $Z \in \Omega$, this follows from the fact that Ω is a witness family. Otherwise we have $Z \in \underline{\mathcal{L}}_t(E)$ so that $Z \sqsubset E$. If $E \pitchfork W$ we have satisfied (WF3). If $E \sqsubset W$, then we may apply (WF3) to $E \in \Omega$ to obtain our conclusion. It thus remains to suppose $Z \sqsubset W \sqsubsetneq E$. But now $W \in \Upsilon(x,y)$, $\mathcal{C}(E|_Z) \subset \mathcal{C}(E|_W)$ and $Z \in \mathcal{L}_t(E)$ together imply that $W \in \mathcal{L}_t(E)$. By maximality of Z, it follows that $W = Z \in \Omega'$. This establishes (WF3) for $\Omega \cup \underline{\mathcal{L}}_t(E)$ and proves the lemma. □

Extending suborderings to augmentations will require the following fact.

Lemma 7.25. *Let $[x,y]$ be a geodesic in $\mathcal{T}(\Sigma)$ and let $Z,W \sqsubset E \sqsubset \Sigma$ be domains such that $\{Z,W,E\} \subset \Upsilon(x,y)$.*

- *If $W \lessdot Z$ along $[x,y]$, then $d_E(x,\mathcal{C}(E|_W)) \leqslant d_E(x,\mathcal{C}(E|_Z))$.*
- *If $d_E(x,\mathcal{C}(E|_W)) \leqslant d_E(x,\mathcal{C}(E|_Z)) - 3$, then $W \lessdot Z$ along $[x,y]$.*

The same conclusions of course hold with the roles of x,y and W,Z swapped.

Proof. Consider the first claim. If $W \in \Upsilon^{\ell}(x,y)$, then W is an annulus and ∂W is short at either x or y. The time ordering $W \lessdot Z$ implies it must be that $\ell_x(\partial W) < \epsilon_0$. Hence $\mathcal{C}(E|_W)$ consists of the single curve ∂W, which is an element of any Bers marking μ_x at x. Thus

$$d_E(x,\mathcal{C}(E|_W)) = \operatorname{diam}_{\mathcal{C}(E)}(\pi_E(x)) \leqslant d_E(x,\mathcal{C}(E|_Z)$$

and the first claim holds when $W \in \Upsilon^{\ell}(x,y)$.

If $W \notin \Upsilon^{\ell}(x,y)$, then necessarily $d_W(x,y) \geqslant \mathsf{N}_W$. Let us set $k_W = d_E(x,\mathcal{C}(E|_W))$ and suppose on the contrary that $k_W > d_E(x,\mathcal{C}(E|_Z))$. Since E contains two subdomains that cut each other, E cannot be an annulus. Recalling that $\pi_E(x)$ is the set of all essential simple closed curves in $\mathcal{C}(E)$ achieved by projecting the curves of the Bers marking μ_x to E, it follows that $\pi_E(x)$ contains at least two distinct curves in $\mathcal{C}(E)$. In particular

$$k_W > d_E(x, \mathcal{C}(E|_Z)) = \operatorname{diam}_{\mathcal{C}(E)}(\pi_E(\mu_x) \cup \mathcal{C}(E|_Z)) \geqslant \operatorname{diam}_{\mathcal{C}(E)} \pi_E(\mu_x) \geqslant 1,$$

which gives $k_W \geqslant 2$.

Choose curves $\gamma \in \pi_E(x)$ and $\nu \in \mathcal{C}(E|_W)$ such that $d_E(\gamma, \nu) = k_W$. Choose also a curve $\zeta \in \partial Z$ that cuts W, and a geodesic $(\alpha_0, \dots, \alpha_m)$ in $\mathcal{C}(E)$ from $\alpha_0 = \gamma$ to $\alpha_m = \zeta$. The curve α_m cuts W by construction. Thus if $m = 0$, we trivially have $\pi_W(\alpha_i) \neq \varnothing$ for each $0 \leqslant i \leqslant m$. Otherwise $m \geqslant 1$ and for each $0 \leqslant i < m < k_W$ the curve α_i necessarily intersects ν (and consequently cuts W) by the fact that

$$d_E(\nu, \alpha_i) \geqslant d_E(\nu, \gamma) - d_E(\gamma, \alpha_i) = k_W - i \geqslant 2.$$

In any case we find that $\pi_W(\alpha_i) \neq \varnothing$ for all $0 \leqslant i \leqslant m$. It follows from the Bounded Geodesic Image Theorem (Theorem 3.8), that $d_W(\gamma, \zeta) \leqslant \mathsf{Q}$. Using $d_W(\partial Z, \zeta) \leqslant 2$ and Lemma 3.9 and recalling that $\mathsf{M} \geqslant 100(\mathsf{k} + \mathsf{Q} + 1)$ (Definition 3.24), this implies

$$d_W(x, \partial Z) \leqslant \mathsf{k} + 2 + d_W(\gamma, \zeta) \leqslant \mathsf{k} + 2 + \mathsf{Q} \leqslant \mathsf{M}/2.$$

However, the time ordering $W \lessdot Z$ implies $d_W(x, \partial Z) \geqslant 2\mathsf{M}/3$, a contradiction.

For the second claim, if W and Z were disjoint or nested, we would have $\operatorname{diam}_{\mathcal{C}(E)}(\mathcal{C}(E|_W) \cup \mathcal{C}(E|_Z)) \leqslant 2$; this can be seen by choosing a curve in $\partial W \cup \partial Z$ that is disjoint from every curve in $\mathcal{C}(E|_W) \cup \mathcal{C}(E|_Z)$. As this is incompatible with the hypothesis $d_E(x, \mathcal{C}(E|_W)) \leqslant d_E(x, \mathcal{C}(E|_Z)) - 3$, it must be that $W \pitchfork Z$. Thus either $Z \lessdot W$ or $W \lessdot Z$. But by the first part of the lemma, $W \lessdot Z$ is the only option compatible with the hypothesis. ◻

We may now extend suborderings to any augmentation in which the parameter and corresponding encroachment are controlled:

Lemma 7.26 (Subordering augmentations)**.** *Let Ω be a subordered witness family for $[x, y]$ in $\mathcal{T}(\Sigma)$ and suppose $E \in \Omega$ satisfies $\mathcal{E}_\Omega(E) \leqslant \mathsf{N}_E/3$. For each parameter $0 \leqslant t \leqslant \mathcal{E}^\ell_\Omega(E)$ (respectively, $0 \leqslant t \leqslant \mathcal{E}^r_\Omega(E)$) there is a natural subordering on $\Omega \cup \underline{\mathcal{L}}_t(E)$ (respectively, $\Omega \cup \underline{\mathcal{R}}_t(E)$) that extends the subordering on Ω.*

Proof. We prove the lemma for $\underline{\Omega} = \Omega \cup \underline{\mathcal{L}}_t(E)$; the proof for $\Omega \cup \underline{\mathcal{R}}_t(E)$ is symmetric. We first show that conditions (SO1)–(SO4) give rise to a natural ordering designation $\swarrow\!\!\!\nearrow$ or $\nwarrow\!\!\!\searrow$ for each pair $Z, V \in \underline{\Omega}$ with $Z \subsetneq V$:

1. If $Z, V \in \Omega$, we use the designation from Ω and set $Z \swarrow\!\!\!\nearrow V$ iff $Z \swarrow V$.
2. If $V \in \Omega$ and $Z \notin \Omega$, then $Z \in \underline{\mathcal{L}}_t(E)$ and we proceed as follows:
   - (i) If $V = E$, we set $Z \swarrow\!\!\!\nearrow V$. (For the case of $\Omega \cup \underline{\mathcal{R}}_t(E)$ with $Z \in \underline{\mathcal{R}}_t(E) \backslash \Omega$ we instead set $E \nwarrow\!\!\!\searrow Z$).
   - (ii) If $E \subsetneq V$, then we set $Z \swarrow\!\!\!\nearrow V \iff E \swarrow V$ and $V \nwarrow\!\!\!\searrow Z \iff V \searrow E$ in accordance with (SO1).
   - (iii) If $V \subsetneq E$, then the fact $V \sqsupset Z \in \mathcal{L}_t(E)$ implies $V \in \mathcal{L}_t(E)$, contradicting the maximality of Z in $\mathcal{L}_t(E)$. Hence $V \subsetneq E$ cannot occur.
   - (iv) If $V \pitchfork E$, we claim it must be that $V \lessdot E$ and therefore set $V \nwarrow\!\!\!\searrow Z$ in accordance with (SO3). (For the case of $\Omega \cup \underline{\mathcal{R}}_t(E)$ with $Z \in \underline{\mathcal{R}}_t(E)$, we instead have $E \lessdot V$ and accordingly set $Z \swarrow\!\!\!\nearrow V$.) To see this, note that $V \sqsupset Z \in \underline{\mathcal{L}}_t(E)$ implies there exists $\nu \in \mathcal{C}(E|_Z)$ with $d_E(x, \nu) \leqslant t + 9\mathsf{C}$. Since $t \leqslant \mathcal{E}^\ell_\Omega(E) \leqslant \mathsf{N}_E/3$, this gives $d_E(y, \nu) \geqslant \frac{2}{3}\mathsf{N}_E - 9\mathsf{C} \geqslant 5\mathsf{M}$. Now since ∂V and ν are disjoint, we have $d_E(y, \partial V) \geqslant 4\mathsf{M} - 2 > 2\mathsf{M}/3$, showing that $V \lessdot E$ by Lemma 3.29.

(3) If $V \notin \Omega$ and $Z \in \Omega$, then (WF3) provides some $Z' \in \Omega$ with $Z \sqsubset Z'$ and $Z' \pitchfork V$. Thus we set $Z \swarrow V$ if $Z' \lessdot V$ and $V \searrow Z$ if $V \lessdot Z'$ in accordance with (SO3). This is well-defined, as Corollary 3.30 ensures it is impossible to have the two such domains $Z'_1, Z'_2 \in \Omega$ with the time ordering $Z'_1 \lessdot V \lessdot Z'_2$.

(4) The case $Z, V \notin \Omega$ is ruled out by the fact domains in $\underline{\mathcal{L}}_t(E)$ are not nested.

The above establishes ordering designations on $\underline{\Omega}$ that extend those of Ω, and so we henceforth use $\swarrow$ and $\searrow$ for the designations in both Ω and $\underline{\Omega}$. To prove these give a subordering on $\underline{\Omega}$, it remains (by Remark 7.15) to verify (SO2)–(SO4):

Condition (SO2): Let $Z, W, V \in \underline{\Omega}$ be such that $Z \swarrow V \searrow W$. The condition is immediate if all three domains lie in Ω. If $V \notin \Omega$, then $Z, W \in \Omega$ because domains in $\underline{\mathcal{L}}_t(E)$ cannot be nested. As dictated by (3) above, we may choose $Z', W' \in \Omega$ such that $Z \sqsubset Z'$, $W \sqsubset W'$, and $Z' \lessdot V \lessdot W'$. By Corollaries 3.30–3.31, this implies $Z' \pitchfork_V W'$ and $Z \pitchfork_V W$ with $Z \lessdot W$, as desired.

Next suppose $V \in \Omega$. If $Z, W \notin \Omega$, then an examination of case (2) above shows that $Z \swarrow V \iff W \swarrow V$ since the ordering designations for $Z \subsetneq V$ and $W \subsetneq V$ are both determined by the relationship of V to E. As this is not the case, at most one of Z or W can lie outside of Ω. Let us first suppose $Z \in \Omega$ and $W \notin \Omega$, so that $W \in \underline{\mathcal{L}}_t(E)$. Now case (2) above dictates that the designation $V \searrow W$ must fall under (2ii) with $V \searrow E$ or else (2iv) in which $V \lessdot E$. In the first case $V \searrow E$ we have $Z \swarrow V \searrow E$ so that (SO2) for Ω ensures $Z \pitchfork_V E$ with $Z \lessdot E$, and in the second case $V \lessdot E$ we have $Z \swarrow V \lessdot E$ so that $Z \pitchfork_V E$ with $Z \lessdot E$ by (SO3) for Ω. In either case, we may conclude $Z \pitchfork_V W$ with $Z \lessdot W$, as desired.

It remains to suppose $V, W \in \Omega$ and $Z \notin \Omega$. By case (2) above, the designation $Z \swarrow V$ must fall under (2i) with $V = E$, or else (2ii) with $E \swarrow V$. In the latter case $E \swarrow V$ we have $E \swarrow V \searrow W$ so that we may use (SO2) to conclude $Z \pitchfork_V W$ and $Z \lessdot W$ as above. So let us restrict our attention to the case $V = E$. Since $E \searrow W$ and $Z \in \underline{\mathcal{L}}_t(E)$, we find that $d_E(y, \mathcal{C}(E|_W)) \leqslant \mathcal{E}^r_\Omega(E)$ and that

$$d_E(x, \mathcal{C}(E|_Z)) \leqslant t + 9\mathsf{C} \leqslant \mathcal{E}^\ell_\Omega(E) + 9\mathsf{C}.$$

Since $\mathcal{E}_\Omega(E) \leqslant \mathsf{N}_E/3$ and $d_E(x,y) \geqslant \mathsf{N}_E > 30\mathsf{C}$, the triangle inequality gives $d_E(x, \mathcal{C}(E|_Z)) \leqslant d_E(x, \mathcal{C}(E|_W)) - \mathsf{C}$ so that we may conclude $Z \lessdot W$ by Lemma 7.25. The fact $d_E(\partial Z, \partial W) \geqslant \mathsf{M} > 3$ further ensures that $Z \pitchfork_V W$, as desired.

Condition (SO3): Let $Z, V \in \underline{\Omega}$ and $W \in \Upsilon(x,y)$ be such that $Z \swarrow V \lessdot W$ or $W \lessdot V \searrow Z$. We may assume exactly one of Z or V lies outside of Ω. First suppose $V \in \Omega$ and $Z \notin \Omega$ so that $Z \in \underline{\mathcal{L}}_t(E)$ and the designation for $Z \subsetneq V$ is dictated by case (2) above. Let us examine these possibilities in turn: If $Z \subsetneq V$ falls under (2i), then $V = E$ and we must have $Z \swarrow V \lessdot W$. Here we find that $d_V(y, \partial W) \leqslant \mathsf{M}/3$ and that $d_V(x, \partial Z) \leqslant t + 9\mathsf{C} \leqslant \mathcal{E}^\ell_\Omega(V) + 9\mathsf{C}$. Since $d_V(x,y) \geqslant \mathsf{N}_V \geqslant 30\mathsf{C}$ and $\mathcal{E}_\Omega(V) \leqslant \mathsf{N}_V/3$, this together with the triangle inequality implies $d_V(\partial Z, \partial W) \geqslant \mathsf{C}$. We may thus conclude the desired $Z \pitchfork_V W$, for otherwise we would find that $d_V(\partial Z, \partial W) \leqslant \mathsf{M}/3$ exactly as in the proof of Corollary 3.30. If $Z \subsetneq V$ falls under (2ii), then (2ii) dictates $E \swarrow V \lessdot W$ in the case that $Z \swarrow V$ and instead dictates $W \lessdot V \searrow E$ in the case $V \searrow Z$. Either way, we may invoke (SO3) to conclude $E \pitchfork_V W$ and consequently $Z \pitchfork_V W$. Finally, if $Z \subsetneq V$ falls under (2iv), then it must be that $V \lessdot E$ and $V \searrow Z$. Hence we are in the case $W \lessdot V \searrow Z$ and in fact have $W \lessdot V \lessdot E$. Therefore $W \pitchfork_V E$ by Corollary 3.30 and we may conclude $W \pitchfork_V Z$ as desired.

Next suppose $V \notin \Omega$ and $Z \in \Omega$. Then (WF3) allows us to choose $Z' \in \Omega$ such that $Z \sqsubset Z'$. If $Z \swarrow V \lessdot W$ then (3) dictates that $Z' \lessdot V \lessdot W$, and if $W \lessdot V \searrow Z$ then (3) instead dictates $W \lessdot V \lessdot Z'$. Either way, we may invoke Corollary 3.30 to conclude $Z \pitchfork_V W$.

Condition (SO4): Let $Z, V \in \underline{\Omega}$ and $W \in \Upsilon(x,y)$ be such that $Z \sqsubsetneq V$, $Z \pitchfork W$, and $\bar{W}^{\underline{\Omega}} = V$. Observe that these facts imply $W \notin \underline{\Omega}$. We must show that $Z \swarrow V \implies Z \lessdot W$ and $V \searrow Z \implies W \lessdot Z$.

First suppose $V \notin \Omega$. Then we must have $Z \in \Omega$ because domains in $\underline{\mathcal{L}}_t(E)$ cannot be nested. Now (WF3) provides a domain $Z' \in \Omega$ such that $Z' \pitchfork V$ and $Z \sqsubset Z'$. Since $Z \pitchfork W$, it cannot be that Z' and W are disjoint. Nor can we have $Z' \sqsubset W$. If $W \sqsubset Z' \in \underline{\Omega}$, then the assumption $\bar{W}^{\underline{\Omega}} = V$ implies $V \sqsubset Z'$ by Lemma 7.6. As this contradicts $Z' \pitchfork V$, we must have $Z' \pitchfork W$. Now, if $Z \swarrow V$, then (3) dictates that $Z' \lessdot V$ so that we find $Z \lessdot W$ by Corollary 3.31. If instead $V \searrow Z$, then (3) dictates $V \lessdot Z'$ and we similarly deduce $W \lessdot Z$. This establishes (SO4) when $V \notin \Omega$.

Next suppose $V \in \Omega$. Then the facts $V \in \Omega \subset \underline{\Omega}$ and $\bar{W}^{\underline{\Omega}} = V$ imply $\bar{W}^{\Omega} = V$ as well (since Lemma 7.6 implies $V \sqsubset V_0$ for any $V_0 \in \Omega$ with $W \sqsubset V_0$). Thus if $Z \in \Omega$ as well we may invoke (SO4) to prove the claim. It therefore suffices to suppose $Z \notin \Omega$ so that $Z \in \underline{\mathcal{L}}_t(E)$. Let us examine the subcases of (2) in turn:

If $V = E$, then (2i) dictates $Z \swarrow V$ and we must show $Z \lessdot W$. If instead $W \lessdot Z$, then Lemma 7.25 implies that $d_E(x, \mathcal{C}(E|_W)) \leqslant d_E(x, \mathcal{C}(E|_Z)) \leqslant t + 9\mathsf{C}$. Since $\bar{W}^{\underline{\Omega}} = E$ it cannot be that $W \in \mathcal{L}_t(E)$, for then we would have $W \sqsubset W' \sqsubsetneq E$ for some $W' \in \underline{\mathcal{L}}_t(E) \subset \underline{\Omega}$. Thus $d_E(x, \mathcal{C}(E|_W)) \notin [t - 9\mathsf{C}, t + 9\mathsf{C}]$. Together, these inequalities give

$$d_E(x, \mathcal{C}(E|_W)) < t - 9\mathsf{C} \leqslant \mathcal{E}^{\ell}_{\Omega}(E) - 9\mathsf{C}.$$

By definition of encroachment, we may choose a domain $U \in \Omega$ such that $U \swarrow E$ and $d_E(x, \mathcal{C}(E|_U)) = \mathcal{E}^{\ell}_{\Omega}(V)$. By Lemma 7.25, the above inequality implies $W \lessdot U$ along $[x,y]$. We now have $U, V \in \Omega$ with $U \swarrow V$ and and $W \lessdot U$. Since we have seen $\bar{W}^{\Omega} = V$, this contradicts the fact that Ω satisfies (SO4). Therefore, $W \lessdot Z$ leads to a contradiction, and we may conclude the desired ordering $Z \lessdot W$.

If $E \sqsubsetneq V$, then (2ii) dictates $Z \swarrow V \iff E \swarrow V$. Observe that $W \lessdot^{\Omega} V$ precludes $W \sqsubset E$. We also cannot have $E \sqsubset W$ nor E and W disjoint (since $Z \pitchfork W$). Therefore $E \pitchfork W$, and we may apply (SO4) to $E \sqsubsetneq V$ in Ω to conclude $Z \swarrow V \implies E \swarrow V \implies E \lessdot W$, which in turn implies $Z \lessdot W$ as desired. If instead $V \searrow Z$, we similarly conclude $W \lessdot E$ and consequently $W \lessdot Z$.

If $V \pitchfork E$, then (2iv) dictates that $V \lessdot E$ and $V \searrow Z$. We cannot have $E \sqsubset W$ nor E and W disjoint (because $Z \pitchfork E$), and if $W \sqsubset E$, then Lemma 7.6 would imply $V = \bar{W}^{\Omega} \sqsubset E$, contradicting $V \pitchfork E$. Therefore it must be that $W \pitchfork E$. We may now apply Corollary 3.31 to $V \lessdot E$ to conclude $W \lessdot E$ and $W \lessdot Z$, as desired. This verifies (SO4) when $V \in \Omega$ and completes the proof of the lemma. ☐

As for refinements, we shall need to bound each augmentation's encroachments.

Lemma 7.27 (Augmented encroachments)**.** *Let Ω be a subordered witness family for $[x,y]$ in $\mathcal{T}(\Sigma)$, let $E \in \Omega$ satisfy $\mathcal{E}_{\Omega}(E) \leqslant \mathsf{N}_E/3$, and let $\underline{\Omega} = \Omega \cup \underline{\mathcal{L}}_t(E)$ with $0 \leqslant t \leqslant \mathcal{E}^{\ell}_{\Omega}(E)$ or $\underline{\Omega} = \Omega \cup \underline{\mathcal{R}}_t(E)$ with $0 \leqslant t \leqslant \mathcal{E}^{r}_{\Omega}(E)$ be the augmentation of Ω along E with parameter t. Then every domain $V \sqsubset \Sigma$ satisfies*

$$\mathcal{E}^{\ell}_{\underline{\Omega}}(V) \leqslant \begin{cases} \max\{\mathcal{E}^{\ell}_{\Omega}(E), t + 9\mathsf{C}\}, & \textit{if } V = E \textit{ and } \underline{\Omega} = \Omega \cup \underline{\mathcal{L}}_t(E) \\ \max\{\mathcal{E}^{\ell}_{\Omega}(V), \mathsf{M}\}, & \textit{else} \end{cases}$$

and

$$\mathcal{E}^r_{\underline{\Omega}}(V) \leqslant \begin{cases} \max\{\mathcal{E}^r_\Omega(E), t+9\mathsf{C}\}, & \textit{if } V = E \textit{ and } \underline{\Omega} = \Omega \cup \underline{\mathcal{R}}_t(E) \\ \max\{\mathcal{E}^r_\Omega(V), \mathsf{M}\}, & \textit{else} \end{cases}.$$

Proof. We prove the lemma for $\Omega \cup \underline{\mathcal{L}}_t(E)$; the proof for $\Omega \cup \underline{\mathcal{R}}_t(E)$ is symmetric. First suppose $V \notin \Omega$. The claim is immediate for $V \notin \underline{\Omega}$ (since then $\mathcal{E}_{\underline{\Omega}}(V) = 0$), so we assume $V \in \underline{\Omega} \backslash \Omega$. The proof of Lemma 7.26 shows that any $Z \in \underline{\Omega}$ with $Z \subsetneq V$ falls under case (3) and satisfies $Z \sqsubset Z'$ for some $Z' \in \Omega$ with $Z' \pitchfork V$. Therefore, as in the proof of Lemma 7.22, $Z \swarrow V$ implies $d_V(x, \mathcal{C}(V|_Z)) \leqslant \mathsf{M}$ and $V \searrow Z$ implies $d_V(y, \mathcal{C}(V|_Z)) \leqslant \mathsf{M}$. Thus the lemma holds for $V \notin \Omega$.

Next suppose $V \in \Omega$ and consider any $Z \in \underline{\Omega}$ with $Z \subsetneq V$. If $Z \swarrow V$, then the proof of Lemma 7.26 shows that either (1) $Z \in \Omega$, (2) $Z \in \underline{\mathcal{L}}_t(E)$ with $V = E$, or (3) $Z \sqsubset E \swarrow V$ so that $\mathcal{C}(V|_Z) \sqsubset \mathcal{C}(V|_E)$. In the former and latter cases we conclude the desired bound $d_V(x, \mathcal{C}(V|_Z)) \leqslant \mathcal{E}^\ell_\Omega(V)$. In the middle case, we have $V = E$ and instead find $d_V(x, \mathcal{C}(V|_Z)) \leqslant t + 9\mathsf{C}$ by the definition of $\underline{\mathcal{L}}_t(E)$. Thus we conclude the stated bound on $\mathcal{E}^\ell_{\underline{\Omega}}(V)$. If instead $V \searrow Z$, the proof of Lemma 7.26 now shows that either (1) $Z \in \Omega$, (2) $V \searrow E \sqsupset Z$ so that $\mathcal{C}(V|_Z) \subset \mathcal{C}(V|_E)$, or (3) $V \lessdot E \sqsupset Z$. In these three cases we may respectively bound $d_V(y, \mathcal{C}(V|_Z))$ by $\mathcal{E}^r_\Omega(V)$, $\mathcal{E}^r_\Omega(V)$, and M. Therefore $\mathcal{E}^r_{\underline{\Omega}}(V) \leqslant \max\{\mathcal{E}^r_\Omega(V), \mathsf{M}\}$ as claimed. ☐

7.7. **Completion.** We may now extend any witness family to a complete and insulated one:

Definition 7.28 (Insulated completion)**.** If Ω is a witness family for $[x, y]$ in $\mathcal{T}(\Sigma)$, define $\check{\Omega}$ to be

$$\check{\Omega} := \Omega \cup \left(\bigcup_{E \in \Omega} \underline{\mathcal{L}}_0(E) \cup \underline{\mathcal{R}}_0(E) \right) \cup \left(\bigcup_{(A,B) \text{ a cutting pair in } \Omega} \underline{\mathcal{F}}(A, B) \right).$$

The *insulated completion* of Ω is then defined to be $\overline{\Omega} = \cup_{i \in \mathbb{N}} \Omega_i$, where $\Omega_0, \Omega_1, \ldots$ is the sequence recursively defined by $\Omega_0 = \Omega$ and $\Omega_{i+1} = \check{\Omega}_i$.

Lemma 7.29. *Let Ω be a witness family for $[x, y]$ in $\mathcal{T}(\Sigma)$ and let $\Omega_0, \Omega_1, \ldots$ be the sequence recursively defined by $\Omega_0 = \Omega$ and $\Omega_{i+1} = \check{\Omega}_i$. Then*

(1) $\check{\Omega}$ is a witness family,
(2) $\overline{\Omega} = \Omega_{\xi(\Sigma)+1}$, and
(3) $\overline{\Omega}$ is a complete and insulated witness family.
(4) Any subordering on Ω extends to natural suborderings on $\check{\Omega}$ and $\overline{\Omega}$ whose encroachments, for each $V \sqsubset \Sigma$, satisfy

$$\mathcal{E}^\ell_{\check{\Omega}}(V), \mathcal{E}^\ell_{\overline{\Omega}}(V) \leqslant \max\{\mathcal{E}^\ell_\Omega(V), 9\mathsf{C}\} \qquad \textit{and} \qquad \mathcal{E}^r_{\check{\Omega}}(V), \mathcal{E}^r_{\overline{\Omega}}(V) \leqslant \max\{\mathcal{E}^r_\Omega(V), 9\mathsf{C}\}.$$

Proof. For (1), first observe that Ω is finite. This is because there are only finitely many subsurfaces $Z \sqsubset \Sigma$ with $d_Z(x, y) \geqslant \mathsf{M}$ and only finitely many annuli A with $\ell_x(\partial A) < \epsilon_0$ or $\ell_y(\partial A) < \epsilon_0$. The family $\check{\Omega}$ may be constructed by adding the finitely many families $\underline{\mathcal{L}}_0(E)$, $\underline{\mathcal{R}}_0(E)$, and $\underline{\mathcal{F}}(A, B)$ one at a time. Lemmas 7.20 and 7.24 show that each addition results in another witness family. Therefore the output $\check{\Omega}$ of those finitely many additions is a witness family.

For (2), let $k_0 = \max\{\xi(V) \mid V \in \Omega_0\}$, and for each $i > 0$ let $k_i = \max\{\xi(V) \mid V \in \Omega_i \backslash \Omega_{i-1}\}$ be the maximal complexity of any domain that was added during the ith iteration, with the convention that $k_i = -\infty$ if $\Omega_i = \Omega_{i-1}$ (in which case $\overline{\Omega} = \Omega_{i-1}$).

We claim that $k_{i-1} \geqslant 1+k_i$ for each $i > 0$. The result $\overline{\Omega} = \Omega_{\xi(\Sigma)+1}$ will then follow from the observation $k_0 \leqslant \xi(\Sigma)$. To see this, suppose $V \in \Omega_i \backslash \Omega_{i-1}$. Then either $V \in \underline{\mathcal{L}}_0(E) \cup \underline{\mathcal{R}}_0(E)$ for some $E \in \Omega_{i-1}$, or else $V \in \underline{\mathcal{F}}(A,B)$ for some cutting pair (A,B) in Ω_{i-1}. If $i = 1$, this shows that V is a proper subsurface of domain in Ω_0 and hence that $k_0 \geqslant 1+\xi(V)$. Next suppose $i > 1$. If $V \in \underline{\mathcal{L}}_0(E) \cup \underline{\mathcal{R}}_0(E)$ then we must have $E \notin \Omega_{i-2}$, for otherwise $V \in \Omega_{i-1} = \check{\Omega}_{i-2}$ by construction. Similarly, if $V \in \underline{\mathcal{F}}(A,B)$, then either $A \notin \Omega_{i-2}$ or $B \notin \Omega_{i-2}$ by the same reasoning. Therefore V is a proper subsurface of a domain in $\Omega_{i-1} \backslash \Omega_{i-2}$ and we may conclude the claimed inequality $k_i \leqslant k_{i-1} - 1$.

Combining (1) and (2), we see that $\overline{\Omega}$ is a witness family and that $\underline{\mathcal{F}}(A,B) \cup \underline{\mathcal{L}}_0(E) \cup \underline{\mathcal{R}}_0(E) \subset \overline{\Omega}$ for all $E \in \overline{\Omega}$ and all cutting pairs (A,B) in $\overline{\Omega}$. Therefore $\overline{\Omega}$ is complete by Lemma 7.9 and insulated by Definition 7.11, which proves (3). Finally, (4) follows from Lemmas 7.21, 7.22, 7.26, and 7.27. □

We may also use Lemmas 7.8 and 7.10 to control the cardinality of $\overline{\Omega}$.

Lemma 7.30. *For each $-1 \leqslant j \leqslant \xi(\Sigma)$, there exists a computable function $G_j \colon \mathbb{N}^{\xi(\Sigma)-j} \to \mathbb{N}$ depending only $\mathsf{N}_{\xi(\Sigma)}, \ldots, \mathsf{N}_{j+1}$ with the following property. If Ω is any witness family for any geodesic $[x,y]$ in $\mathcal{T}(\Sigma)$, then*

$$\left|\overline{\Omega}\right|_j - |\Omega|_j \leqslant G_j(K_{\xi(\Sigma)}, \ldots, K_{j+1})$$

for any tuple $(K_{\xi(\Sigma)}, \ldots, K_{j+1})$ satisfying $|\Omega|_i \leqslant K_i$ for each $\xi(\Sigma) \geqslant i > j$. In particular, there exists a computable function $G \colon \mathbb{N} \to \mathbb{N}$ such that $\left|\overline{\Omega}\right| \leqslant G(|\Omega|)$ for every witness family Ω.

Proof. For $j = \xi(\Sigma)$, we may take the constant function $G_{\xi(\Sigma)} \equiv 1$ since $\overline{\Omega}$ can contain at most one domain of this complexity. For the remaining j, we proceed inductively: Fix an integer $-1 \leqslant j < \xi(\Sigma)$ and suppose that the stipulated functions $G_{\xi(\Sigma)}, \ldots G_{j+1}$ have been constructed. Let us count the domains $V \in \overline{\Omega} \backslash \Omega$ of complexity $\xi(V) = j$. Any such V satisfies $V \in \underline{\mathcal{L}}_0(E) \cup \underline{\mathcal{R}}_0(E)$ for some $E \in \overline{\Omega}$ with $\xi(E) > j$, or else $V \in \underline{\mathcal{F}}(A,B)$ for some $A, B \in \overline{\Omega}$ with $\xi(A), \xi(B) > j$. Letting J denote the number of domains of $\overline{\Omega}$ of complexity at least $j+1$, it follows from Lemmas 7.8 and 7.10 that

$$\left|\overline{\Omega}\right|_j - |\Omega|_j = \left|\overline{\Omega} \backslash \Omega\right|_j \leqslant \binom{J}{2} (2\mathsf{N}_{j+1})^{\xi(\Sigma)+2} + 2J(2\mathsf{N}_{j+1})^{\xi(\Sigma)+3}.$$

However, our induction hypothesis gives

$$J \leqslant \sum_{i=j+1}^{\xi(\Sigma)} \left|\overline{\Omega}\right|_i \leqslant \left(G_{\xi(\Sigma)} + K_{\xi(\Sigma)}\right) + \cdots + \left(G_{j+1}(K_{\xi(\Sigma)}, \ldots K_{j+2}) + K_{j+1}\right).$$

Combining these inequalities shows that $\left|\overline{\Omega}\right|_j - |\Omega|_j$ is bounded above by function of $(K_{h_\Sigma}, \ldots, K_{j+1})$ that depends only on the thresholds $\mathsf{N}_{\xi(\Sigma)}, \ldots, \mathsf{N}_{j+1}$, as claimed. The final assertion of the lemma then holds for the function $G \colon \mathbb{N} \to \mathbb{N}$ defined as

$$G(x) = x + G_{\xi(\Sigma)} + \cdots + G_{-1}(x, \ldots, x). \qquad \square$$

8. Complexity of witness families

We next explain how the the structure of a witness family organizes curve complex projection data into a quantity that we call complexity. Let us first designate an acronym combining the many types of witness families that have been introduced in Definitions 7.7, 7.11, 7.14, and 7.18.

Terminology 8.1 (WISC)**.** A witness family is *WISC* if it is wide, insulated, subordered, and complete.

The starting point is the following notion that is suggested by completeness:

Definition 8.2 (Contribute)**.** If Ω is a complete witness family for a geodesic segment $[x,y] \in \mathcal{T}(\Sigma)$, we say that a domain $Z \sqsubset \Sigma$ *contributes to* $V \in \Omega$ if $Z \in \Upsilon(x,y)$ and $V = \bar{Z}^{\Omega}$.

Since every domain $Z \in \Upsilon(x,y)$ has a unique Ω–supremum, we may partition the domains of $\Upsilon(x,y)$ according to the elements of Ω they contribute to. We would like to somehow combine the data $\{(Z, d_Z(x,y)) \mid Z \text{ contributes to } V\}$ into a notion of "distance in V" that, when summed over all $V \in \Omega$, can be used for counting problems and is moreover related to the total Teichmüller distance $d_{\mathcal{T}(\Sigma)}(x,y)$. A subordering on Ω allows us to accomplish this by *resolving* x and y into points in the Teichmüller space $\mathcal{T}(V)$ for each $V \in \Omega$. In fact, we can resolve any point coarsely aligned between x and y.

8.1. **Teichmüller resolutions.** Recall the constant $\mathsf{C} \geqslant 0$ specified at the start of §7 (which determines the N_i).

Definition 8.3 (Projection tuple)**.** Let Ω be a WISC witness family for a geodesic $[x,y]$ in $\mathcal{T}(\Sigma)$. For each domain $V \in \Omega$ and point $w \in \mathcal{T}(\Sigma)$ satisfying

$$d_Z(x,w) + d_Z(w,y) \leqslant d_Z(x,y) + 9\mathsf{C} \quad \text{for all } Z \sqsubset V,$$

define its *projection tuple* to be the tuple $(\tilde{w}_Z) \in \prod_{Z \sqsubset V} \mathcal{C}(Z)$ given by:

$$\tilde{w}_Z := \begin{cases} \pi_Z(y), & \text{if } Z \in \Upsilon(x,y) \text{ and } \bar{Z}^{\Omega} \swarrow V \\ \pi_Z(x), & \text{if } Z \in \Upsilon(x,y) \text{ and } V \searrow \bar{Z}^{\Omega} \\ \pi_Z(w), & \text{else} \end{cases}.$$

In particular, $\tilde{w}_Z = \pi_Z(w)$ whenever $Z \sqsubset \Sigma$ contributes to V.

Proposition 8.4. *With the notation from Definition 8.3, the projection tuple $(\tilde{w}_Z) \in \prod_{Z \sqsubset V} \mathcal{C}(Z)$ is k–consistent for some constant k depending only on C.*

Proof. Let $U, Z \sqsubset V$ be arbitrary subdomains. We must show that:

$$\begin{aligned} U \pitchfork Z \quad &\Longrightarrow \quad \min\{d_U(\tilde{w}_U, \partial Z), d_Z(\tilde{w}_Z, \partial U)\} \leqslant k \\ U \sqsubset Z \quad &\Longrightarrow \quad \min\{d_U(\tilde{w}_U, \pi_U(\tilde{w}_Z)), d_Z(\tilde{w}_Z, \partial U)\} \leqslant k \end{aligned} \tag{8.5}$$

Note that for each $p \in \{x, w, y\}$ the pair $(\pi_U(p), \pi_Z(p))$ in $\mathcal{C}(U) \times \mathcal{C}(Z)$ satisfies these conditions with constant K by Theorem 3.37. Thus we may assume $\tilde{w}_U \neq \pi_U(w)$ or $\tilde{w}_Z \neq \pi_Z(w)$. We may additionally assume $d_U(x,y) \geqslant \mathsf{N}_U$ and $d_Z(x,y) \geqslant \mathsf{N}_Z$. Indeed, if say $d_Z(x,y) < \mathsf{N}_Z$, then by coarse alignment $\pi_Z(x) \cup \pi_Z(w) \cup \pi_Z(y)$ has diameter at most $9\mathsf{C} + \mathsf{N}_Z \leqslant 2\mathsf{N}$. Hence, regardless of whether $\tilde{w}_U$ is defined as $\pi_U(x)$, $\pi_U(w)$, or $\pi_U(y)$ we may move $\tilde{w}_Z$ by distance at most $2\mathsf{N}$ to arrive at a K–consistent pair $(\pi_U(p), \pi_Z(p))$ as above. In particular, $U, Z \in \Upsilon(x,y)$ and $\bar{U}^{\Omega}$ and $\bar{Z}^{\Omega}$ both exist by the completeness of Ω.

Suppose first $U \pitchfork Z$. By symmetry, we may suppose $\tilde{w}_Z \neq \pi_Z(w)$ so that $\bar{Z}^{\Omega}$ is a proper subsurface of V. We only consider the case $\bar{Z}^{\Omega} \swarrow V$ as the opposite case $V \searrow \bar{Z}^{\Omega}$ is symmetric. In this case $\tilde{w}_Z = \pi_Z(y)$ by definition. If we also have $\bar{U}^\Omega \swarrow V$, then $\tilde{w}_U = \pi_U(y)$ and the pair $(\tilde{w}_U, \tilde{w}_Z)$ is K–consistent. If instead $V \searrow \bar{U}^\Omega$, then (SO2) and Corollary 3.31 imply that $Z \lessdot U$ along $[x,y]$. Therefore $d_Z(\tilde{w}_Z, \partial U) = d_Z(y, \partial U) < \mathsf{M}/3$ by Lemma 3.29, and (8.5) is satisfied. The only remaining possibility is $\bar{U}^\Omega = V$. In this case we necessarily have $U \pitchfork (\bar{Z}^\Omega)$ (U is not disjoint from $\bar{Z}^\Omega$ since $U \pitchfork Z \sqsubset \bar{Z}^\Omega$, U is not contained in $\bar{Z}^\Omega$ as that would give $\bar{U}^\Omega \sqsubset \bar{Z}^\Omega$, and $\bar{Z}^\Omega$ is not contained in U since that would give $Z \sqsubset U$). Thus U and $\bar{Z}^\Omega$ are time-ordered. Since $\bar{Z}^\Omega \swarrow V$ and $\bar{U}^\Omega = V$, condition (SO4) forces $\bar{Z}^\Omega \lessdot U$ which in turn implies $Z \lessdot U$. Therefore $d_Z(\tilde{w}_Z, \partial U) = d_Z(y, \partial U) < \mathsf{M}/3$, as above, and we have verified (8.5) when $U \pitchfork Z$.

Next let us suppose that $U \sqsubset Z$. Then $\bar{U}^\Omega \sqsubset \bar{Z}^\Omega$. If $\bar{U}^\Omega = \bar{Z}^\Omega$, then consistency is automatically satisfied by definition of $\tilde{w}_U, \tilde{w}_Z$ and Theorem 3.37. So suppose $\bar{U}^\Omega \subsetneq \bar{Z}^\Omega$. If $\bar{Z}^\Omega \subsetneq V$, then condition (SO1) ensures that $\bar{U}^\Omega \swarrow V$ iff $\bar{Z}^\Omega \swarrow V$ and we again have consistency by Theorem 3.37. The only remaining possibility is $\bar{U}^\Omega \subsetneq \bar{Z}^\Omega = V$. Let us consider the case $\bar{U}^\Omega \swarrow V$ (the other case $V \searrow \bar{U}^\Omega$ being similar). In this case we have $\tilde{w}_Z = \pi_Z(w)$ and $\tilde{w}_U = \pi_U(y)$ by definition.

Claim 8.6. $d_Z(x, \partial U) < \mathcal{E}_\Omega(V) + \mathsf{M} \leqslant \mathsf{N}/2$.

Proof. We clearly cannot have $Z \sqsubset \bar{U}^\Omega$. If $Z \pitchfork \bar{U}^\Omega$ then (SO4) implies we must have $\bar{U}^\Omega \lessdot Z$ and consequently $d_Z(x, \partial(\bar{U}^\Omega)) \leqslant \mathsf{M}/3$. Since $\partial U \cup \partial(\bar{U}^\Omega)$ is a curve system on Σ we have $d_Z(\partial U, \partial(\bar{U}^\Omega)) \leqslant 2$. Thus $d_Z(x, \partial U) \leqslant 2 + \mathsf{M}/3 < \mathsf{M}$.

Since $\bar{U}^\Omega$ and Z cannot be disjoint, it remains to suppose $\bar{U}^\Omega \sqsubset Z$. There are two possibilities: Firstly, if $Z = V$, then

$$d_Z(x, \partial U) = d_V(x, \partial U) \leqslant d_V(x, \mathcal{C}(V|_{\bar{U}^\Omega})) \leqslant \mathcal{E}_\Omega(V) \leqslant \mathsf{N}/3$$

by the fact $\bar{U}^\Omega \swarrow V$. Secondly, if $Z \subsetneq V$, then $\bar{Z}^\Omega = V$ implies $Z \notin \Omega$ so that (WF3) provides some $Z' \in \Omega$ with $Z \pitchfork Z'$ and $\bar{U}^\Omega \sqsubset Z'$. Note that we must have $Z' \neq V$. If $Z' \subsetneq V$, then (SO1) implies $Z' \swarrow V$ so that (SO4) forces $Z' \lessdot Z$. Otherwise we have $Z' \pitchfork V$ so that (SO3) (using $\bar{U}^\Omega \swarrow V$ and $\bar{U}^\Omega \sqsubset Z'$) forces $Z' \lessdot V$ and we may again conclude $Z' \lessdot Z$ by Corollary 3.31. Therefore $d_Z(x, \partial Z') \leqslant \mathsf{M}/3$ so that we may use $U \sqsubset Z'$ to conclude $d_Z(x, \partial U) \leqslant \mathsf{M}/3 + 2 < \mathsf{M}$ as above. □

Since $\operatorname{diam}_{\mathcal{C}(U)}(\pi_U(w), \pi_U(\pi_Z(w)))$ is bounded by Lemma 3.9, verifying (8.5) amounts to bounding $\min\{d_U(y,w), d_Z(w, \partial U)\}$. Thus if $d_U(w,y) \leqslant \mathsf{N}$ we are done. Otherwise $d_U(w,y) > \mathsf{N}$, and applying Corollary 3.27 and Claim 8.6 gives

$$d_Z(w,x) + d_Z(x,y) \leqslant d_Z(w, \partial U) + d_Z(\partial U, y) + \mathsf{N} \leqslant d_Z(w,y) + 2\mathsf{N}.$$

On the other hand, the coarse alignment hypothesis on w gives

$$d_Z(x,w) + d_Z(w,y) \leqslant d_Z(x,y) + 9\mathsf{C}.$$

Combining these inequalities yields

$$2d_Z(x,w) + d_Z(w,y) - 9\mathsf{C} \leqslant d_Z(w,y) + 2\mathsf{N},$$

or equivalently $d_Z(w,x) \leqslant (2\mathsf{N} + 9\mathsf{C})/2 \leqslant 3\mathsf{N}/2$. Thus the triangle inequality and Claim 8.6 now give the desired bound $d_Z(w, \partial U) \leqslant 2\mathsf{N}$. □

Combining Proposition 8.4 with Theorem 3.37 and Lemma 3.10, we are now able to *resolve* w into the Teichmüller space of any $V \in \Omega$:

Definition 8.7 (Resolution point)**.** Let Ω be a WISC witness family for $[x,y]$ in $\mathcal{T}(\Sigma)$. For any $V \in \Omega$, there are coarsely well-defined *resolution points* $\widehat{x}_V^\Omega, \widehat{y}_V^\Omega \in \mathcal{T}(V)$ constructed as follows: Let $w \in \{x,y\}$. If V is nonannular, then $\widehat{w}_V^\Omega \in \mathcal{T}_{\epsilon_0}(V)$

is a thick point realizing the consistent tuple $(\tilde{w}_Z)_{Z\sqsubset V}$ from Definition 8.3. If V is an annulus, then the tuple$(\tilde{w}_Z)_{Z\sqsubset V}$ is a singleton $\tilde{w}_V \in \mathcal{C}(V)$, and we define $\widehat{w}_V^{\Omega}$ to be the point in $\mathcal{T}(V) = \mathbb{H}^2$ whose twist coordinate is given by $\tilde{w}_V = \pi_V(w)$, and whose length coordinate is $\frac{1}{\min\{\epsilon_0, \ell_w(\partial V)\}}$.

8.2. **Complexity via Teichmüller distance.** Given a WISC witness family Ω for a Teichmüller geodesic $[x,y]$ in $\mathcal{T}(\Sigma)$, Proposition 8.4 provides a pair $\widehat{x}_V^{\Omega}, \widehat{y}_V^{\Omega}$ of resolutions for each $V \in \Omega$. We now combine these into the following quantity:

Definition 8.8 (Complexity)**.** The *complexity* of a WISC witness family Ω for a geodesic $[x,y]$ in $\mathcal{T}(\Sigma)$ is the weighted sum

$$\mathfrak{L}(\Omega) = \sum_{V\in\Omega} h_V^* d_{\mathcal{T}(V)}(\widehat{x}_V^{\Omega}, \widehat{y}_V^{\Omega})$$

where $h_V^* = h_V$ for every nonannular domain, and for annuli A we set $h_A^* = 1$ in the case that $\widehat{x}_A^{\Omega}, \widehat{y}_A^{\Omega}$ are both ϵ_0–thick, and otherwise set $h_A^* = 2 = h_A$.

Remark 8.9. Let us highlight three features of this definition.

(1) The resolution points $\widehat{x}_V^{\Omega}, \widehat{y}_V^{\Omega}$ coarsely encode all the projection data of x, y, with the result that it is possible to reconstruct the original points from their resolutions. This allows one to relate complexity $\mathfrak{L}(\Omega)$ to counting problems, as we do in §12 below.

(2) It is helpful to compare this definition of $\mathfrak{L}(\Omega)$ to the distance formula Theorem 3.33. Indeed, if one applies the distance formula to each term $d_{\mathcal{T}(V)}(\widehat{x}_V^{\Omega}, \widehat{y}_V^{\Omega})$, the result is a weighted (by h_V^* and the multiplicative errors) sum of curve complex distances $d_Z(x,y)$ for all $Z \in \Upsilon(x,y)$. Thus $\mathfrak{L}(\Omega)$ is coarsely equivalent to $d_{\mathcal{T}(\Sigma)}(x,y)$ with some bounded but *unknown* multiplicative and additive error. The purpose of §§9–10 below to show that one can choose Ω carefully so that, up to only additive error, $\mathfrak{L}(\Omega)$ is bounded above by the *explicit* multiple $h_\Sigma d_{\mathcal{T}(\Sigma)}(x,y)$ (Theorem 11.2). This multiplicative control is crucial in our counting applications (Theorem 12.1) since the quantity $\mathfrak{L}(\Omega)$ appears in the exponent.

(3) The final and perhaps least apparent feature is that by decomposing the Teichmüller distance into separate subsurfaces, the quantity $\mathfrak{L}(\Omega)$ is able to tap into the hyperbolicity of curve complexes and promote alignment in curve complexes to a sort of alignment for complexity. That is, the definition is constructed with the heuristic that if (x,y,z) is an aligned triple in $\mathcal{T}(\Sigma)$ with associated WISC witness families $\Omega_x^y, \Omega_y^z, \Omega_x^z$, then one should morally expect $\mathfrak{L}(\Omega_x^y) + \mathfrak{L}(\Omega_y^z) \leqslant \mathfrak{L}(\Omega_x^z)$ up to additive error. To achieve this precise statement seems to be quite difficult. However, we will show in Theorem 11.2 that the witness families can be chosen so that, up to only additive error, $\mathfrak{L}(\Omega_x^y) + \mathfrak{L}(\Omega_y^z)$ is bounded above by $h_\Sigma d_{\mathcal{T}(\Sigma)}(x,z)$ provided (x,y,z) is strongly aligned. This feature together with the above-mentioned Theorem 12.1 make complexity a useful tool for counting orbit points of finite order and reducible mapping classes.

8.3. **Complexity of tuples.** To obtain the features indicated in Remark 8.9, we will work in a more general setting of tuples of witness families. Recall the parameter $\mathsf{C} \geqslant 2\mathsf{M}$ from §7 that determines the constants N_i and satisfies (7.1).

Definition 8.10. A *witness family for a strongly* C*–aligned tuple* $(x_0, \dots, x_n)$ in $\mathcal{T}(\Sigma)$ is a tuple $\Omega = (\Omega_1, \dots, \Omega_n)$ where each Ω_i is a witness family for $[x_{i-1}, x_i]$.

All of the notation and terminology from §7—such as subordering, refinement, augmentation, completion—are extended *componentwise* to the setting of witness families for tuples. Thus Ω has a given property provided it holds for each Ω_i. In particular, a subordering on Ω is a subordering on each Ω_i, and we will write $\swarrow_i$ and ${}_i\searrow$ for the subordering designations on Ω_i. We additionally define encroachments as $\mathcal{E}_\Omega(V) = \max_i \mathcal{E}_{\Omega_i}(V)$ and similarly for $\mathcal{E}^\ell_\Omega(V)$ and $\mathcal{E}^r_\Omega(V)$, and define the insulated completion of Ω to be $\overline{\Omega} = (\overline{\Omega_1}, \dots, \overline{\Omega_n})$.

Notation 8.11. When a strongly C–aligned tuple $(x_0, \dots, x_n)$ has been specified, we will use the shorthand $\Upsilon_i = \Upsilon(x_{i-1}, x_i)$ and similarly $\Upsilon^\ell_i = \Upsilon^\ell_i(x_{i-1}, x_i)$ and $\Upsilon^c_i = \Upsilon^c(x_{i-1}, x_i)$. Similarly, if $\Omega = (\Omega_1, \dots, \Omega_n)$ is a witness family for $(x_0, \dots, x_n)$, we will by abuse of notation write $V \in \Omega$ to mean that $V \in \cup_i \Omega_i$.

Definition 8.12. A witness family Ω for a strongly C–aligned tuple $(x_0, \dots, x_n)$ is WISC if each Ω_i is WISC, and in this case the *complexity* of Ω is defined as

$$\mathfrak{L}(\Omega) = \sum_{i=1}^{n} \mathfrak{L}(\Omega_i) = \sum_{i=1}^{n} \sum_{V \in \Omega_i} h^*_V d_{\mathcal{T}(V)}(\widehat{x_{i-1}}{}_V^{\Omega_i}, \widehat{x}_{iV}^{\Omega_i}).$$

In order to account for those annuli where we use $h^*_V = 1$ instead of $h^*_V = 2$ in the above formula, we also introduce the following:

Definition 8.13. The *savings* of a WISC witness family $\Omega = (\Omega_1, \dots, \Omega_n)$ is

$$\mathfrak{S}(\Omega) = \sum_{i=1}^{n} \sum_{V \in \Omega_i} (h_V - h^*_V) d_{\mathcal{T}(V)}(\widehat{x_{i-1}}{}_V^{\Omega_i}, \widehat{x}_{iV}^{\Omega_i}).$$

9. Bounding the contribution of a witness

We recall from the introduction that the reason witness families were introduced and the goal of the whole second half of the paper are Theorem 11.2 and Theorem 12.1 The first bounds the complexity of a collection of witness families defined by a strongly aligned set of points in terms of Teichmüller distance. The second counts net points in terms of complexity. Together they will give the desired count of net points in terms of Teichmüller distance. As a major first step towards proving Theorem 11.2 in this section we bound the distances $d_{\mathcal{T}(V)}(\widehat{x_{i-1}}{}_V^{\Omega}, \widehat{x}_{iV}^{\Omega})$ contributed by each individual witness; this is the content of Theorem 9.4. Throughout this section, we fix a WISC witness family $\Omega = (\Omega_1, \dots, \Omega_n)$ for a strongly C–aligned tuple $(x_0, \dots, x_n)$ in $\mathcal{T}(\Sigma)$. For each domain $V \sqsubset \Sigma$, we let $x^V_0, \dots, x^V_n \in [x_0, x_n]$ denote the points provided by Definition 3.21 (strong alignment) that appear in order along $[x_0, x_n]$ and satisfy $d_V(x_i, x^V_i) \leqslant \mathsf{C}$. In the case of an annulus, we furthermore assume the ratio of $\min\{\epsilon_0, \ell_{x_i}(\partial A)\}$ and $\min\{\epsilon_0, \ell_{x^V_i}(\partial A)\}$ is at most C. We also remind the reader that the collections $\Upsilon, \Upsilon^c, \Upsilon^\ell$ were introduced in Definition 7.2.

9.1. Contribution sets.

Estimating $\mathfrak{L}(\Omega)$ will involve a careful analysis of active intervals along the main Teichmüller geodesic $[x_0, x_n]$. To this end, we have the following basic observations.

Lemma 9.1. *If $V \in \Upsilon(x_{i-1}, x_i)$, for some $1 \leqslant i \leqslant n$, then V has a nonempty active interval $\mathcal{I}_V$ along $[x_0, x_n]$. In particular, this holds for each $V \in \cup_i \Omega_i$.*

Proof. Assume first $V \in \Upsilon^c(x_{i-1}, x_i)$. Then $d_V(x_{i-1}, x_i) \geqslant \mathsf{N}_V$. Hence by C–alignment of $(x_0, \ldots, x_n)$ we have

$$d_V(x_0, x_n) \geqslant d_V(x_0, x_{i-1}) + d_V(x_{i-1}, x_i) + d_V(x_i, x_n) - 2\mathsf{C} \geqslant \mathsf{N}_V - 2\mathsf{C} > \mathsf{M}.$$

Hence $\mathcal{I}_V \neq \varnothing$ by Lemma 3.26.

Otherwise $V \in \Upsilon^\ell(x_{i-1}, x_i)$ and V is an annulus with at least one of $\ell_{x_{i-1}}(\partial V)$ and $\ell_{x_i}(\partial V)$ smaller than ϵ_0/N_V Without loss of generality, we may therefore suppose $\ell_{x_i}(\partial V) \leqslant \epsilon_0/\mathsf{N}_V$. By strong C–alignment and the choice of N_V (7.1), this gives $\ell_{x_i^V}(\partial V) \leqslant \mathsf{C}\epsilon_0/\mathsf{N}_V < \epsilon_0{}'$. Therefore, V has a nonempty active interval $\mathcal{I}_V = \tilde{\mathcal{I}}_V^{\epsilon_0}$ along $[x_0, x_n]$ by Theorem 3.22(2) and Definition 3.25. □

Lemma 9.2. *If $Z \in \Upsilon^c(x_{i-1}, x_i)$ and $W \in \Upsilon(x_{i-1}, x_i)$ satisfy $Z \pitchfork W$, then Z and W are time-ordered compatibly along $[x_0, x_n]$ and $[x_{i-1}, x_i]$.*

Proof. We know from Lemma 9.1 and Remark 7.3 that Z and W have nonempty active intervals along both $[x_{i-1}, x_i]$ and $[x_0, x_n]$. Let us suppose that $Z \lessdot W$ along $[x_{i-1}, x_i]$ (the reverse possibility being handled similarly), and by contradiction that $W \lessdot Z$ along $[x_0, x_n]$. Then by time-ordering, $d_Z(\partial W, x_i)$ and $d_Z(x_0, \partial W)$ are at most $\mathsf{M}/3$. Hence

$$d_Z(x_0, x_i) \leqslant d_Z(x_0, \partial W) + d_Z(\partial W, x_i) \leqslant 2\mathsf{M}/3.$$

Since $Z \in \Upsilon_i^c$, alignment now gives the contradictory inequality

$$d_Z(x_0, x_i) \geqslant d_Z(x_0, x_{i-1}) + d_Z(x_{i-1}, x_i) - \mathsf{C} \geqslant \mathsf{N}_Z - \mathsf{C} \geqslant \mathsf{M}. \qquad \square$$

Recall (Definition 8.2) that $Z \sqsubset V$ contributes to $V \in \Omega_i$ if $Z \in \Upsilon_i$ and $V = \bar{Z}^{\Omega_i}$. For each $V \in \Omega$, we will now define a "contribution set" for V along $[x_0, x_n]$ by starting with the active interval $\mathcal{I}_V$, then removing the active interval $\mathcal{I}_W$ for any domain $W \in \Omega$ with $W \subsetneq V$, and finally adding the active intervals $\mathcal{I}_Z$ of any domain Z that contributes to V in some Ω_i. More precisely, for each $V \in \Omega$, we use Lemma 9.1 to define

$$\begin{aligned} M(V) &= \bigcup \{\mathcal{I}_W \mid W \in \Omega \text{ with } W \subsetneq V\} \subset [x_0, x_n], \quad \text{and} \\ C(V) &= \bigcup_{1 \leqslant i \leqslant n} C_i(V), \end{aligned}$$

where for each index $1 \leqslant i \leqslant n$ we define $C_i(V) = \varnothing$ if $V \notin \Omega_i$ and otherwise define

$$C_i(V) = \bigcup \{\mathcal{I}_Z \mid Z \subsetneq V \text{ contributes to } V \text{ in } \Omega_i\} \subset [x_0, x_n].$$

Definition 9.3 (Contribution set)**.** The *contribution set* of $V \in \cup_i \Omega_i$ is

$$\mathcal{A}_V^\Omega = \big(\mathcal{I}_V \backslash M(V)\big) \cup C(V) \subset [x_0, x_n].$$

We stress that all active intervals here are taken along the main geodesic $[x_0, x_n]$.

The following result is the heart of proving Theorem 11.2. It bounds Teichmüller distances in terms of size of active intervals of contribution sets.

Theorem 9.4. *If $V \in \Omega_i$, then $d_{\mathcal{T}(V)}(\widehat{x_{i-1}}_V^{\Omega_i}, \widehat{x_i}_V^{\Omega_i}) \stackrel{+}{\preceq}_{\mathsf{N}} \int_{x_{i-1}^V}^{x_i^V} \mathbb{1}_{\mathcal{A}_V^\Omega}$.*

We remark that the term on the left and the integrand both depend on the witness family, while the limits of integration just depend on the points x_{i-1} and x_i.

9.2. **Proving Theorem 9.4 for annuli.** We maintain the notation Ω, x_i, and x_i^V from the start of §9. Fix some index $1 \leqslant i \leqslant n$ and an annular domain $V \in \Omega_i$. So by strong alignment we have, in particular:

$$d_V(x_j, x_j^V) \leqslant \mathsf{C} \quad \text{and} \quad \frac{1}{\mathsf{C}} \leqslant \frac{\min\{\epsilon_0, \ell_{x_j}(\partial V)\}}{\min\{\epsilon_0, \ell_{x_j^V}(\partial V)\}} \leqslant \mathsf{C} \quad \text{for } j = i-1, i. \tag{9.5}$$

To ease notation, set $\hat{x}_j = \hat{x}_{jV}^{\Omega_i} \in \mathcal{T}(V) = \mathbb{H}^2_{\partial V}$ for $j \in \{i-1, i\}$, and recall that by definition these resolution points satisfy

$$d_V(\hat{x}_j, x_j) \stackrel{+}{\prec}_{\mathsf{C}} 0 \quad \text{and} \quad \ell_{\hat{x}_j}(\partial V) = \min\{\epsilon_0, \ell_{x_j}(\partial V)\} \quad \text{for } j = i-1, i. \tag{9.6}$$

The proof will follow easily from these facts:

Proof of Theorem 9.4–Annular case. Consider the active interval $\mathcal{I}_V$ of V along $[x_0, x_n]$. For each point $w \in \mathcal{I}_V$ we have $\ell_w(\partial V) < \epsilon_0$, and we write $w|_V$ for the $T(V)$–component of the point $\Phi_{\partial V}(w)$ in the product region $\mathcal{P}(\Sigma|\partial V)$. Since V is an annulus, there are no proper subdomains of V; hence by definition the contribution set is simply $\mathcal{A}_V^\Omega = \mathcal{I}_V$.

First suppose that $\mathcal{A}_V^\Omega \cap [x_{i-1}^V, x_i^V]$ is empty. Then $d_V(x_{i-1}^V, x_i^V) \leqslant \mathsf{M}$ and $\ell_{x_{i-1}^V}(\partial V), \ell_{x_i^V}(\partial V) \geqslant {\epsilon_0}'$. Therefore equations (9.5) and (9.6) above imply that

$$d_V(\hat{x}_{i-1}, \hat{x}_i) \stackrel{+}{\prec}_{\mathsf{C}} 0 \qquad \text{and} \qquad \frac{{\epsilon_0}'}{\mathsf{C}} \leqslant \ell_{\hat{x}_{i-1}}(\partial V), \ell_{\hat{x}_i}(\partial V) \leqslant \epsilon_0,$$

which together uniformly bound $d_{\mathcal{T}(V)}(\hat{x}_{i-1}, \hat{x}_i)$ in terms of C.

If $\mathcal{A}_V^\Omega \cap [x_{i-1}^V, x_i^V]$ is nonempty, then (being the intersection of intervals) it is necessarily an interval and we may write it as $[y, z] \subset [x_{i-1}^V, x_i^V]$. We claim that

$$d_{\mathcal{T}(V)}(\hat{x}_{i-1}, y|_V) \stackrel{+}{\prec}_{\mathsf{C}} 0 \qquad \text{and} \qquad d_{\mathcal{T}(V)}(\hat{x}_i, z|_V) \stackrel{+}{\prec}_{\mathsf{C}} 0. \tag{9.7}$$

By symmetry, let us just consider $d_{\mathcal{T}(V)}(\hat{x}_i, z|_V)$. To see this, note that if $z = x_i^V$ then obviously $\ell_z(\partial V) = \ell_{x_i^V}(\partial V)$, and otherwise we have both $\ell_{x_i^V}(\partial V) \geqslant {\epsilon_0}'$ and ${\epsilon_0}' \leqslant \ell_z(\partial V) \leqslant \epsilon_0$. Thus in either case equations (9.5)–(9.6) imply

$$\frac{{\epsilon_0}'}{\mathsf{C}\epsilon_0} \leqslant \frac{\min\{\epsilon_0, \ell_{x_i^V}(\partial V)\}}{\mathsf{C}\ell_z(\partial V)} \leqslant \frac{\ell_{\hat{x}_i}(\partial V)}{\ell_z(\partial V)} \leqslant \mathsf{C}\frac{\min\{\epsilon_0, \ell_{x_i^V}(\partial V)\}}{\min\{\epsilon_0, \ell_z(\partial V)\}} \leqslant \mathsf{C}\frac{\epsilon_0}{{\epsilon_0}'}.$$

Furthermore, since $[z, x_i^V]$ is disjoint from the interior of $\mathcal{I}_V$, Lemma 3.26 gives $d_V(z, x_i^V) \leqslant \mathsf{M}$. Combining with (9.5)–(9.6) we therefore have $d_V(\hat{x}_i, z) \stackrel{+}{\prec}_{\mathsf{C}} 0$. This proves $\hat{x}_i$ and $z|_V$ coarsely have the same horizontal coordinate in $\mathbb{H}^2_{\partial V}$, and the above bounds on $\ell_{\hat{x}_i}(\partial V)/\ell_z(\partial V)$ show they coarsely have the same vertical component. Therefore $d_{\mathcal{T}(V)}(\hat{x}_i, z|_V)$ is indeed bounded as claimed.

To conclude the argument, since y and z both lie in the thin region for ∂V, Minsky's product regions Theorem 3.11 implies that

$$d_{\mathcal{T}(V)}(y|_V, z|_V) \stackrel{+}{\asymp} d_{\mathcal{T}(\Sigma)}(y, z) = \int_y^z 1 \leqslant \int_{x_{i-1}^V}^{x_i^V} \mathbb{1}_{\mathcal{A}_V^\Omega}.$$

Combining this with (9.7) and the triangle inequality proves the proposition. $\square$

9.3. **Proving Theorem 9.4 for nonannuli.** We maintain the notation Ω, x_i, and x_i^V fixed at the start of §9. We also fix an index $1 \leqslant i \leqslant n$ and a nonannular domain $V \in \Omega_i$. Note that in this case $V \in \Upsilon^c(x_{i-1}, x_i)$ so that $d_V(x_{i-1}, x_i) \geqslant \mathsf{N}_V$.

9.3.1. *Setup.* We begin by identifying a subinterval of $[x_{i-1}^V, x_i^V]$ on which we have better control of resolution points.

Lemma 9.8. *There is a nonempty subinterval $J = [y, z] \subset [x_{i-1}^V, x_i^V]$ such that*

- *$d_V(x_{i-1}, y)$ and $d_V(x_i, z)$ are both at most $7\mathsf{C}$.*
- *For all $w \in J$ the distances $d_V(x_{i-1}, w)$ and $d_V(w, x_i)$ are both at least $3\mathsf{C}$.*

Furthermore J is contained in the active interval $\mathcal{I}_V$ of V along $[x_0, x_n]$.

Proof. Recall $\mathsf{C} \geqslant \mathsf{M} \geqslant \mathsf{L}$. We know $d_V(x_i, x_i^V), d_V(x_{i-1}, x_{i-1}^V) \leqslant \mathsf{C}$. Therefore

$$d_V(x_{i-1}^V, x_i^V) \geqslant d_V(x_{i-1}, x_i) - 2\mathsf{C} \geqslant \mathsf{N}_V - 2\mathsf{C} \geqslant 28\mathsf{C}.$$

Since $\pi_V : \mathcal{T}(\Sigma) \to \mathcal{C}(V)$ is coarsely L–Lipschitz and $\mathsf{L} \leqslant \mathsf{C}$, there must exists points $y, z \in [x_{i-1}^V, x_i^V]$ such that

$$5\mathsf{C} \leqslant d_V(x_{i-1}^V, y) \leqslant 6\mathsf{C} \qquad \text{and} \qquad 5\mathsf{C} \leqslant d_V(x_i^V, z) \leqslant 6\mathsf{C}.$$

Observe that necessarily y and z appear in order along $[x_{i-1}^V, x_i^V]$ for otherwise $y \in [z, x_i^V]$ and we may apply Theorem 3.19 (no backtracking) to conclude

$$\begin{aligned} d_V(x_{i-1}^V, x_i^V) &\leqslant d_V(x_{i-1}^V, y) + d_V(y, x_i^V) \leqslant 6\mathsf{C} + d_V(z, y) + d_V(y, x_i^V) \\ &\leqslant 6\mathsf{C} + d_V(z, x_i^V) + \mathsf{B} \leqslant 6\mathsf{C} + 6\mathsf{C} + \mathsf{B} < 13\mathsf{C}, \end{aligned}$$

which we have seen is false. By the triangle inequality, we also clearly have

$$d_V(x_{i-1}, y) \leqslant 7\mathsf{C} \quad \text{and} \quad d_V(x_i, z) \leqslant 7\mathsf{C}.$$

Finally, for any $w \in [x_0, z]$ Theorem 3.19 additionally gives

$$d_V(w, x_i^V) \geqslant d_V(w, z) + d_V(z, x_i^V) - \mathsf{B} \geqslant 5\mathsf{C} - \mathsf{B} \geqslant 4\mathsf{C} \tag{9.9}$$

so that $d_V(w, x_i) \geqslant 3\mathsf{C}$ by the triangle inequality. Similarly $d_V(x_{i-1}, w) \geqslant 3\mathsf{C}$ for all $w \in [y, x_n]$. This proves all $w \in J$ satisfy the second bullet point.

Finally, we know from Lemma 9.1 that V has a nonempty active interval along $[x_0, x_n]$. If $\mathcal{I}_V$ were disjoint from $[z, x_i^V]$, then we would have $d_V(z, x_i^V) \leqslant \mathsf{M}/3$ by Lemma 3.26(3). But this contradicts the implication $d_V(z, x_i^V) \geqslant 4\mathsf{C}$ of Equation (9.9). Thus $\mathcal{I}_V$ necessarily intersects $[z, x_i^V]$ and, similarly, $[x_{i-1}^V, y]$. Since $\mathcal{I}_V$ is an interval, the containment $J = [y, z] \subset \mathcal{I}_V$ follows. □

The interval J moreover contains the active interval of each domain contributing to V in Ω_i; this is a variation of Lemma 7.12(1) for this more general context of witness families for aligned tuples:

Lemma 9.10. *If $Z \subsetneq V$ contributes to V in Ω_i, then its active interval along $[x_0, x_n]$ lies in the interior of J. Further, $d_Z(x_{i-1}, x_{i-1}^V) \leqslant \mathsf{M}$ and $d_Z(x_i, x_i^V) \leqslant \mathsf{M}$.*

Proof. The fact that Z contributes to V implies $Z \in \Upsilon(x_{i-1}, x_i)$ but that $Z \notin \Omega_i$. Recall from Lemma 7.13 that $\Omega_i \supset \Upsilon^\ell(x_{i-1}, x_i)$; hence in fact $Z \in \Upsilon^c(x_{i-1}, x_i)$. If $d_V(x_i, \partial Z) \leqslant 9\mathsf{C}$, then by definition we would have $Z \in \mathcal{R}_0(V)$ for Ω_i and hence $Z \sqsubset Z'$ for some $Z' \in \underline{\mathcal{R}}_0(V)$. But since Ω_i is insulated, this would imply $Z' \in \Omega_i$ and contradict $\bar{Z}^{\Omega_i} = V$. Therefore

$$d_V(x_i, \partial Z) > 9\mathsf{C} > \mathsf{C} + \mathsf{M} \geqslant d_V(x_i, x_i^V) + \mathsf{M}.$$

Corollary 3.27 therefore implies $d_Z(x_i, x_i^V) < \mathsf{M}$. Similarly $d_Z(x_{i-1}, x_{i-1}^V) < \mathsf{M}$. Also observe that for all all $w \in [z, x_i^V]$ we have

$$d_V(x_i^V, w) \leqslant d_V(x_i^V, w) + d_V(w, z) \leqslant d_V(x_i^V, z) + \mathsf{B} \leqslant 6\mathsf{C} + \mathsf{B}$$

and therefore $d_V(x_i, w) \leqslant 8\mathsf{C}$. Similarly $d_V(x_{i-1}, w) \leqslant 8\mathsf{C}$ for all $w \in [x_{i-1}^V, y]$.

We know from Lemma 9.1 that Z has a nonempty active interval $\mathcal{I}_Z$ along $[x_0, x_n]$. We claim that $\mathcal{I}_Z$ is disjoint from $[z, x_i^V]$. Indeed, otherwise we would have $w \in [z, x_i^V] \cap \mathcal{I}_Z$ with $\partial Z \subset \mu_w$ and hence $d_V(x_i, \partial Z) \leqslant d_V(x_i, w) \leqslant 8\mathsf{C}$, contradicting the above lower bound

$$d_V(x_i, \partial Z) > 9\mathsf{C}.$$

Similarly $\mathcal{I}_Z$ must be disjoint from $[x_{i-1}^V, y]$.

Therefore, if $\mathcal{I}_Z$ is not contained in the interior of $J = [y, z]$, it is necessarily disjoint from $[x_{i-1}^V, x_i^V]$. This gives $d_Z(x_{i-1}^V, x_i^V) \leqslant \mathsf{M}/3$ and thus by the triangle inequality

$$\begin{aligned} d_Z(x_{i-1}, x_i) &\leqslant d_Z(x_{i-1}, x_{i-1}^V) + d_Z(x_{i-1}^V, x_i^V) + d_Z(x_i^V, x_i) \\ &\leqslant \mathsf{M} + \mathsf{M}/3 + \mathsf{M} < \mathsf{N}_Z. \end{aligned}$$

But this contradicts the fact, observed above, that $Z \in \Upsilon^c(x_{i-1}, x_i)$. □

The following observation will also be useful.

Lemma 9.11. *Suppose $W \subsetneq V$ has a nonempty active interval along $[x_0, x_n]$. If $\mathcal{I}_W$ intersects $[x_0, z]$ (resp. $[y, x_n]$) then $d_W(x_j, x_j^V) \leqslant \mathsf{M}$ for all $j \geqslant i$ (resp. all $j \leqslant i-1$). In particular, if $\mathcal{I}_W$ intersects J (as holds for every Z that contributes to V in Ω_i by Lemma 9.10), then $d_W(x_j, x_j^V) \leqslant \mathsf{M}$ for all $0 \leqslant j \leqslant n$.*

Proof. We suppose $\mathcal{I}_W \cap [x_0, z] \neq \varnothing$, the alternate hypothesis $\mathcal{I}_W \cap [y, x_n] \neq \varnothing$ being handled symmetrically. Fix any $j \geqslant i$. Pick some point $w \in \mathcal{I}_W \cap [x_0, z]$, so that $\partial W \subset \mu_w$. We then have $[z, x_i^V] \subset [w, x_j^V]$ and therefore (by Theorem 3.19)

$$d_V(w, x_j^V) \geqslant d_V(w, z) + d_V(z, x_i^V) + d_V(x_i^V, x_j^V) - 2\mathsf{B} \geqslant d_V(z, x_i^V) - 2\mathsf{B} \geqslant 4\mathsf{C}.$$

It follows that

$$d_V(\partial W, x_j^V) \geqslant d_V(w, x_j^V) - \mathsf{L} \geqslant 3\mathsf{C} > d_V(x_j, x_j^V) + \mathsf{M}.$$

Thus Corollary 3.27 gives the desired bound $d_W(x_j, x_j^V) \leqslant \mathsf{M}$. □

Corollary 9.12. *Suppose that $W \subsetneq V$ satisfies $W \in \Upsilon(x_{j-1}, x_j)$ for some $j \neq i$. Then $\mathcal{I}_W \cap J = \varnothing$.*

Proof. We assume $W \in \Upsilon(x_{j-1}, x_j)$ for $j > i$, the alternate case $j < i$ being symmetric. We know (Lemma 9.1) that W has a nonempty active interval $\mathcal{I}_W$ along $[x_0, x_n]$. To derive a contradiction, let us suppose there is a point $w \in \mathcal{I}_W \cap J$. It cannot be that $x_i^V \in \mathcal{I}_W$, since that would imply $[z, x_i^V] \subset \mathcal{I}_W$ and hence

$$d_V(z, x_i^V) \leqslant d_V(z, \partial W) + d_V(\partial W, x_i^V) \leqslant 2\mathsf{L} < 5\mathsf{C},$$

violating the choice of z in Lemma 9.8. Since $\mathcal{I}_W$ is an interval, we find that $[x_i^V, x_n] \supset [x_{j-1}^V, x_j^V]$ misses $\mathcal{I}_W$. Lemma 3.26(3) and Lemma 9.11 now give

$$d_W(x_{j-1}, x_j) \leqslant d_W(x_{j-1}, x_{j-1}^V) + d_W(x_{j-1}^V, x_j^V) + d_W(x_j^V, x_j) \leqslant 3\mathsf{M} < \mathsf{N}_W.$$

This shows $W \notin \Upsilon^c(x_{j-1}, x_j)$. Thus we must have $W \in \Upsilon^\ell(x_{j-1}, x_j)$. Choose $k \in \{j-1, j\}$ so that $\ell_{x_k}(\partial W) < \epsilon_0/\mathsf{N}_W$, and note that $k \geqslant i$. Since the curve ∂W is short at x_k, we evidently have $d_V(\partial W, x_k) \leqslant \mathsf{L}$. Since ∂W is also short at the chosen point $w \in \mathcal{I}_W \cap J$, this shows

$$d_V(w, x_k^V) \leqslant d_V(w, \partial W) + d_V(\partial W, x_k) + d_V(x_k, x_k^V) \leqslant 2\mathsf{L} + \mathsf{C}.$$

On the other hand the fact that $[z, x_i^V] \subset [w, x_k^V]$ gives (via Theorem 3.19)

$$d_V(w, x_k^V) \geqslant d_V(w, z) + d_V(z, x_i^V) + d_V(x_i^V, x_k^V) - 2\mathsf{B} \geqslant 5\mathsf{C} - 2\mathsf{B}.$$

As these inequalities are incompatible, we have derived our contradiction. $\square$

The following property of the interval J will play a key role in our argument.

Lemma 9.13. *If $w \in J$, then every domain $Z \sqsubset V$ satisfies*

$$d_Z(x_{i-1}, w) + d_Z(w, x_i) \leqslant d_Z(x_{i-1}, x_i) + 9\mathsf{C}.$$

Proof. Fix any domain $Z \sqsubset V$. First suppose that $d_Z(x_{i-1}, x_{i-1}^V)$ and $d_Z(x_i, x_i^V)$ are both at most $2\mathsf{C}$ (as is the case for $Z = V$). Then since $J \subset [x_{i-1}^V, x_i^V]$, Theorem 3.19 and the triangle inequality give

$$\begin{aligned} d_Z(x_{i-1}, w) + d_Z(w, x_i) &\leqslant d_Z(x_{i-1}^V, w) + d_Z(w, x_i^V) + 4\mathsf{C} \\ &\leqslant d_Z(x_{i-1}^V, x_i^V) + 4\mathsf{C} + \mathsf{B} \leqslant d_Z(x_{i-1}, x_i) + 9\mathsf{C}. \end{aligned}$$

So it suffices to assume at least one of the quantities is larger that $2\mathsf{C}$. Suppose then that $d_Z(x_i^V, x_i) > 2\mathsf{C} > \mathsf{M}$ (the other possibility is handled similarly). Then

$$d_V(x_i, \partial Z) \leqslant d_V(x_i, \partial Z) + d_V(\partial Z, x_i^V) \leqslant d_V(x_i, x_i^V) + \mathsf{M}/3 \leqslant 2\mathsf{C}$$

by Corollary 3.27. The triangle inequality therefore gives

$$d_V(x_{i-1}, \partial Z) \geqslant d_V(x_{i-1}, x_i) - d_V(x_i, \partial Z) \geqslant \mathsf{N}_V - 2\mathsf{C} \geqslant 28\mathsf{C}.$$

In particular, it must be the case that $d_Z(x_{i-1}, x_{i-1}^V) \leqslant 2\mathsf{C}$ (since otherwise the above argument would force $d_V(x_{i-1}, \partial Z) \leqslant 2\mathsf{C}$, which is false).

We next show that $d_Z(x_{i-1}^V, w) \leqslant \mathsf{M}$. Indeed, otherwise $d_Z(x_{i-1}^V, w) > \mathsf{M}$ and Z must have an active interval along $[x_{i-1}^V, w]$. Thus there is some point $u \in [x_{i-1}^V, w]$ that contains ∂Z in its Bers marking. Thus $d_V(x_i, u) \leqslant d_V(x_i, \partial Z) + 1 \leqslant 2\mathsf{C} + 1$. On the other hand equation (9.9) (in the proof of Lemma 9.8) gives $d_V(u, x_i^V) \geqslant 4\mathsf{C}$, which implies $d_V(u, x_i) \geqslant 3\mathsf{C}$; a contradiction.

We now know both $d_Z(x_{i-1}, x_{i-1}^V) \leqslant 2\mathsf{C}$ and $d_Z(x_{i-1}^V, w) \leqslant \mathsf{M}$. Combining these gives $d_Z(x_{i-1}, w) \leqslant 3\mathsf{C}$. It is now easy to conclude

$$\begin{aligned} d_Z(x_{i-1}, w) + d_Z(w, x_i) &\leqslant d_Z(x_{i-1}, w) + d_Z(w, x_{i-1}) + d_Z(x_{i-1}, x_i) \\ &\leqslant 3\mathsf{C} + 3\mathsf{C} + d_Z(x_{i-1}, x_i). \end{aligned}$$

$\square$

9.3.2. *Comparison points.* Lemma 9.13 and Proposition 8.4 imply that for each $w \in J$, the projection tuple $(\tilde{w}_Z) \in \prod_{Z \sqsubset V} \mathcal{C}(Z)$ from Definition 8.3 is k–consistent for some constant k depending only on C. We next use this fact together with the lengths of certain curves at w to define a point $\hat{w} \in \mathcal{T}(V)$ as follows:

Definition 9.14 (Comparison point)**.** For each point $w \in J$, consider the tuple $(\tilde{w}_Z)_{Z \sqsubset V}$ from Definition 8.3. Let α_w be the multicurve consisting of those curves $\gamma \in \Gamma(V)$ which are essential in V, have $\ell_w(\gamma) < \epsilon_0$, and satisfy

$$d_Z(\gamma, \tilde{w}_Z) = \mathrm{diam}_{\mathcal{C}(Z)}(\pi_Z(\gamma) \cup \tilde{w}_Z) \leqslant 2\mathsf{M} \text{ for every domain } Z \sqsubset V. \tag{9.15}$$

Using Proposition 8.4, Theorem 3.37, and Lemma 3.10, we may then build a marking μ of V that realizes the tuple $(\tilde{w}_Z)_Z$ and has $\alpha_w \subset \mathrm{base}(\mu)$. Working in

Fenchel–Nielsen coordinates for the pants decomposition base(μ), take $\hat{w} \in \mathcal{T}(V)$ to be the point whose Bers marking is μ and such that $\gamma \in \text{base}(\mu)$ has

$$\ell_{\hat{w}}(\gamma) = \begin{cases} \epsilon_0, & \text{if } \gamma \notin \alpha_w \\ \ell_w(\gamma), & \text{if } \gamma \in \alpha_w \end{cases}.$$

This comparison point satisfies (and is coarsely characterized by):

(1) $d_Z(\hat{w}, \tilde{w}_Z) \stackrel{+}{\prec}_{\mathsf{N}} 0$ for every domain $Z \sqsubset V$.
(2) If $\gamma \in \Gamma(V)$ is an essential curve in V, then $\ell_{\hat{w}}(\gamma) < \epsilon_0$ if and only if γ satisfies $\ell_w(\gamma) < \epsilon_0$ and (9.15). Further, in this case $\ell_{\hat{w}}(\gamma) = \ell_w(\gamma)$.

The next lemma shows that if $w \in \mathcal{I}_Z$ for some domain $Z \subsetneq V$ that contributes to V in Ω_i, then $\partial Z \subset \alpha_w$ and hence, by construction, $\ell_{\hat{w}}(\gamma) < \epsilon_0$ for each component γ of ∂Z that is essential in V. Thus the points $w \in \mathcal{T}(\Sigma)$ and $\hat{w} \in \mathcal{T}(V)$ both live in product regions for Z, and we may compare them as follows:

Lemma 9.16 (Comparisons in active intervals)**.** *Suppose $Z \subsetneq V$ contributes to V in Ω_i, For all $w \in \mathcal{I}_Z$ with corresponding comparison $\hat{w} \in \mathcal{T}(V)$, the following hold:*

(1) $\ell_w(\gamma) < \epsilon_0$ and $\ell_{\hat{w}}(\gamma) < \epsilon_0$ for each component $\gamma \in \partial Z \cap \Gamma(V)$.
(2) Writing $w|_Z$ and $\hat{w}|_Z$ for the $\mathcal{T}(Z)$–components of $\Phi_{\partial Z}(w) \in \mathcal{P}(\Sigma|\partial Z)$ and $\Phi_{\partial Z}(\hat{w}) \in \mathcal{P}(V|\partial Z)$, respectively, we have $d_{\mathcal{T}(Z)}(w|_Z, \hat{w}|_Z) \stackrel{+}{\prec}_{\mathsf{N}} 0$.

Proof. We will need the following observation.

Claim 9.17. *If $U \in \Upsilon(x_{i-1}, x_i)$ satisfies $\bar{U}^{\Omega_i} \swarrow_i V$ (resp. $V \searrow_i \bar{U}^{\Omega_i}$), then either Z and U are disjoint, or else $\mathcal{I}_U$ occurs before (resp. after) $\mathcal{I}_Z$ along $[x_0, x_n]$.*

Proof. Set $U' = \bar{U}^{\Omega_i}$ and, by symmetry, suppose $U' \swarrow_i V$. We may assume Z is not disjoint from U, and hence neither disjoint from U'. Note that we cannot have $Z \sqsubset U$ or $Z \sqsubset U'$, as that would imply $V = \bar{Z}^{\Omega_i} \sqsubset U' \subsetneq V$ by Lemma 7.6.

The fact that $Z \subsetneq V$ contributes to V implies $Z \notin \Omega_i$. We claim there is some $W \in \Omega_i$ such that $U' \sqsubset W$ and $W \pitchfork Z$. If $U' \pitchfork Z$ then we can simply take $W = U'$. Otherwise $U' \sqsubset Z$ and (WF3) (applied to $U' \in \Omega_i$ and $Z \notin \Omega_i$) provides such a W.

Since $W \pitchfork Z \sqsubset V$, we see that both $V \sqsubset W$ and $V \perp W$ are impossible. If $W \pitchfork V$, then (SO3) (applied to $U' \swarrow_i V$ and $V \sqsupset U' \sqsubset W$) forces $W \lessdot V$ and hence $W \lessdot Z$ along $[x_{i-1}, x_i]$ by Corollary 3.31. Otherwise $W \subsetneq V$ and (SO1) (using $U' \swarrow_i V$) implies $W \swarrow_i V$ so that we may invoke (SO4) (using $W \pitchfork Z$) to again conclude $W \lessdot Z$ along $[x_{i-1}, x_i]$. Note that the fact $Z \notin \Omega_i \supset \Upsilon^\ell(x_{i-1}, x_i)$ ensures that $Z \in \Upsilon^c(x_{i-1}, x_i)$. Hence Lemma 9.2 implies we have the same time-ordering $W \lessdot Z$ along $[x_0, x_n]$. Since $U \sqsubset W$, Lemma 3.26(4) now implies the intervals $\mathcal{I}_U$ and $\mathcal{I}_Z$ along $[x_0, x_n]$ are disjoint, and in fact it must be that $\mathcal{I}_U$ occurs before $\mathcal{I}_Z$. ☐

Returning to the lemma: Since $w \in \mathcal{I}_Z$, Lemma 3.26 implies $\ell_w(\alpha) < \epsilon_0$ for every component α of ∂Z. Hence, (1) will follow from the following fact:

Claim 9.18. *If $\gamma \in \mathcal{C}(V|_Z)$ satisfies $\ell_w(\gamma) < \epsilon_0$, then $\ell_{\hat{w}}(\gamma) = \ell_w(\gamma) < \epsilon_0$.*

Proof of claim. Given the hypotheses, by definition of $\hat{w}$ it suffices to show γ satisfies (9.15). Let $U \sqsubset V$ be any subdomain. If γ is disjoint from U, we trivially have $d_U(\gamma, \tilde{w}_U) = \text{diam}_{\mathcal{C}(U)}(\tilde{w}_U) \leqslant \mathsf{M}$. Observe also that

$$d_U(\gamma, w) = \text{diam}_{\mathcal{C}(U)}(\pi_U(\gamma) \cup \pi_U(w)) \leqslant \mathsf{L} \leqslant \mathsf{M}/2$$

owing to the fact that γ is short at w. Thus (9.15) is immediate when $\tilde{w}_U = \pi_U(w)$; this takes care of the case that U contributes to V. It remains to suppose, then, that $\gamma \pitchfork U$ and $U \in \Upsilon(x_{i-1}, x_i)$ with $\bar{U}^{\Omega_i} \neq V$. We write $U' = \bar{U}^{\Omega_i}$ and, by symmetry, assume $U' \swarrow_i V$. Then $\tilde{w}_U = \pi_U(x_i)$. As the curve $\gamma \in \mathcal{C}(V|_Z)$ cuts U, it cannot be that Z and U are disjoint. Claim 9.17 thus ensures $\mathcal{I}_U$ occurs before $\mathcal{I}_Z$ along $[x_0, x_n]$. Since $w \in \mathcal{I}_Z \subset J = [y, z] \subset [x_{i-1}^V, x_i^V]$ by Lemma 9.10, we now see that w and x_i^V lie in the same component of $[x_0, x_n] \backslash \mathcal{I}_U$. Whence $d_U(w, x_i^V) \leqslant \mathsf{M}/3$ by Lemma 3.26(3). We also see that $\mathcal{I}_U$ intersects $[x_0, z]$ and hence that $d_U(x_i, x_i^V) \leqslant \mathsf{M}$ by Lemma 9.11. Therefore $d_U(w, x_i) \leqslant 4\mathsf{M}/3$. Since $\tilde{w}_U = \pi_U(x_i)$ and we have already observed $d_U(\gamma, w) \leqslant \mathsf{M}/2$, we conclude that $d_U(\gamma, \tilde{w}_U) \leqslant 2\mathsf{M}$ and condition (9.15) is verified. □

It now follows from (1) that w and $\hat{w}$ lie in product regions for ∂Z, so we are justified in considering $w|_Z, \hat{w}|_Z \in \mathcal{T}(Z)$. By the distance formula [Raf1, Theorem 6.1], to bound $d_{\mathcal{T}(Z)}(w|_Z, \hat{w}|_Z)$ it suffices to show that $w|_Z$ and $\hat{w}|_Z$ have the same short curves and the same curve complex projections to all subsurfaces of Z.

First let $\beta \in \Gamma(Z)$ be an essential curve of Z. We claim that either $\ell_{w|_Z}(\beta)$ and $\ell_{\hat{w}|_Z}(\beta)$ are both at least ϵ_0, or else $\ell_{w|Z}(\beta)$ and $\ell_{\hat{w}|_Z}(\beta)$ coarsely agree. Indeed, by nature of the homeomorphism $\Phi_{\partial Z}$, the lengths $\ell_w(\beta)$ and $\ell_{w|_Z}(\beta)$ coarsely agree, as do $\ell_{\hat{w}}(\beta)$ and $\ell_{\hat{w}|_Z}(\beta)$. Thus it suffices to show either $\ell_w(\beta), \ell_{\hat{w}}(\beta) \geqslant \epsilon_0$ or else $\ell_w(\beta)$ and $\ell_{\hat{w}}(\beta)$ coarsely agree. But this follows from the construction of $\hat{w}$: if $\ell_w(\beta) < \epsilon_0$, then $\ell_{\hat{w}}(\beta) = \ell_w(\beta)$ by Claim 9.18. Conversely, if $\ell_{\hat{w}}(\beta) < \epsilon_0$, then we must have $\ell_{\hat{w}}(\beta) = \ell_w(\beta)$ by item (2) of Definition 9.14.

Next let $U \sqsubset Z$ be any domain in Z. Since the curves of ∂Z are all short at w, the Bers marking μ_w at w has $\partial Z \subset \operatorname{base}(\mu_w)$. Therefore, taking the curves of μ_w that are essential in Z defines a marking of μ' of Z, and in fact μ' is a Bers marking $\mu_{w|_Z}$ of $w|_Z$. Since $U \sqsubset Z$, we have $\pi_U(\mu_w) = \pi_U(\mu') = \pi_U(\mu_{w|_Z})$. Thus $d_U(w, w|_Z) \stackrel{+}{\prec} 0$. Similarly $d_U(\hat{w}, \hat{w}|_Z) \stackrel{+}{\prec} 0$. It therefore suffices to bound $d_U(w, \hat{w})$. By construction (Definition 9.14(1)) $d_U(\hat{w}, \tilde{w}_U) \stackrel{+}{\prec}_{\mathsf{N}} 0$ for $\tilde{w}_U$ as in Proposition 8.4. Thus we must bound $d_U(w, \tilde{w}_U)$. We consider the three possibilities of $\tilde{w}_U$: if $\tilde{w}_U = \pi_U(w)$ this is immediate. If not then $U \in \Upsilon(x_{i-1}, x_i)$ and $\bar{U}^{\Omega_i} \neq V$. Since U and Z are evidently not disjoint, if $\bar{U}^{\Omega_i} \swarrow_i V$ then Claim 9.17 implies that $\mathcal{I}_U$ occurs before $\mathcal{I}_Z$ along $[x_0, x_n]$. As above, (using Lemmas 3.26(3) and 9.11) it follows that $d_U(w, x_i^V) \leqslant \mathsf{M}/3$ and $d_U(x_i^V, x_i) \leqslant \mathsf{M}$ so that $d_U(w, \tilde{w}_U) = d_U(w, x_i) \leqslant 2\mathsf{M}$. If instead $V {}_i\!\searrow \bar{U}^{\Omega_i}$, we similarly obtain $d_U(w, \tilde{w}_U) = d_U(w, x_{i-1}) \leqslant 2\mathsf{M}$ and thereby establish (2). □

9.3.3. *The main argument.* With the requisite notation and setup established, we now work in earnest towards the proof of Theorem 9.4.

Definition 9.19 (The point $\bar{w}$)**.** Since $J \subset \mathcal{I}_V$ (Lemma 9.10), each point $w \in J$ lies in the thin region for the multicurve ∂V; accordingly we let $\bar{w}$ denote the $\mathcal{T}(V)$–component of product region point $\Phi_{\partial V}(w) \in \mathcal{P}(\Sigma|\partial V)$.

Our proof relies on comparing the points $\bar{w}, \hat{w} \in \mathcal{T}(V)$ for carefully chosen $w \in J$. To streamline notation, and mimic that used in Definition 9.14, we will set $\hat{x}_{i-1} = \widehat{x_{i-1}}_V^{\Omega_i}$ and $\hat{x}_i = \widehat{x}_{iV}^{\Omega_i}$; however we stress that $\hat{x}_{i-1}$ and $\hat{x}_i$ are defined by Definition 8.7 and are necessarily thick, whereas points $\hat{w}$ for $w \in J$ (from Definition 9.14) may be thin.

Remark 9.20. The fact that points $\hat{w}$, for $w \in J$, are allowed to be thin causes technical complications in the proof. However, allowing thinness is necessary in order for the crucial ingredient Lemma 9.16(2) to hold.

Strategy 9.21. The goal is to show that $d_{\mathcal{T}(V)}(\hat{x}_{i-1}, \hat{x}_i)$ is bounded, up to additive error, by $\int_{x_{i-1}^V}^{x_i^V} \mathbb{1}_{\mathcal{A}_V^\Omega}$. Since $J = [y, z] \subset [x_{i-1}^V, x_i^V]$, it suffices to instead work with $\int_y^z \mathbb{1}_{\mathcal{A}_V^\Omega}$. That is, we are concerned with the Lebesgue measure of $J \cap \mathcal{A}_V^\Omega$.

We will construct a piecewise geodesic path in $\mathcal{T}(\Sigma)$ from x_{i-1} to x_i with the property that each segment $[p, q]$ satisfies either $d_{\mathcal{T}(V)}(\hat{p}, \hat{q}) \stackrel{+}{\prec}_{\mathsf{C}} d_{\mathcal{T}(\Sigma)}(p, q)$, or else $d_Z(\hat{p}, \hat{q}) \stackrel{+}{\prec}_{\mathsf{C}} 0$ for every domain $Z \sqsubset V$; these two properties will be established in Lemmas 9.27 and 9.28 below. The piecewise path will consist of boundedly many segments—each of which is either $[x_{i-1}, y]$, $[z, x_i]$, or a subintervals of J—and will be constructed using breakpoints provided (essentially) by Lemma 9.29.

Furthermore, the segments $[p, q]$ with $d_{\mathcal{T}(V)}(\hat{p}, \hat{q}) \stackrel{+}{\prec}_{\mathsf{C}} d_{\mathcal{T}(\Sigma)}(p, q)$ will have total length at most $\int_y^z \mathbb{1}_{\mathcal{A}_V^\Omega}$. The triangle inequality thus implies $d_{\mathcal{T}(V)}(\hat{x}_{i-1}, \hat{x}_i)$ is at most $\int_y^z \mathbb{1}_{\mathcal{A}_V^\Omega}$ plus the sum of the lengths $d_{\mathcal{T}(V)}(\hat{p}, \hat{q})$ for the other segments $[p, q]$ with $d_Z(\hat{p}, \hat{q})$ bounded for all $Z \sqsubset V$. To complete the proof, we will use Minsky's product regions Theorem 3.11 to show these latter segments can be ignored.

To begin, let $\mathcal{D}$ denote the set of domains $Z \sqsubset \Sigma$ such that $Z \in \Upsilon(x_{i-1}, x_i)$ with $Z \sqsubsetneq V$ and $\bar{Z}^{\Omega_i} = V$. Thus $\mathcal{D}$ consists of all domains contributing to V in Ω_i except for V itself, and hence $C_i(V) = \cup_{Z \in \mathcal{D}} \mathcal{I}_Z$.

Definition 9.22. We say a subinterval $[p, q]$ of J is *squarely covered by* $\mathcal{D}$ if:

- the open interval (p, q) intersects $C_i(V)$, and
- whenever the open interval (p, q) intersects $\mathcal{I}_Z$ for some $Z \in \mathcal{D}$, then we have $[p, q] \subset \mathcal{I}_Y$ for some $Y \in \mathcal{D}$ with $\mathcal{I}_Z \subset \mathcal{I}_Y$ and $Z \sqsubset Y$.

Lemma 9.23. *If $[p, q] \subset J$ is squarely covered by $\mathcal{D}$, then $d_{\mathcal{T}(V)}(\hat{p}, \hat{q}) \stackrel{+}{\prec}_{\mathsf{C}} d_{\mathcal{T}(\Sigma)}(p, q)$.*

Proof. By hypothesis there exists $Z \in \mathcal{D}$ with $(p, q) \cap \mathcal{I}_Z \neq \varnothing$. If $Z \in \mathcal{D}$ is any such domain, then square covering further implies $[p, q] \subset \mathcal{I}_Y$ for some $Y \in \mathcal{D}$ with $\mathcal{I}_Z \subset \mathcal{I}_Y$ and $Z \sqsubset Y$. Let $\mathcal{Y}$ denote the set of topologically maximal domains in the collection

$$\{Y \in \mathcal{D} \mid [p, q] \subset \mathcal{I}_Y\}.$$

It follows from the above that $\mathcal{Y}$ is nonempty and moreover that if $Z \in \mathcal{D}$ satisfies $\mathcal{I}_Z \cap (p, q) \neq \varnothing$, then $Z \sqsubset Y$ for some $Y \in \mathcal{Y}$.

The domains in $\mathcal{Y}$ are evidently pairwise disjoint, since they cannot be nested and their active intervals overlap. Therefore $\partial\mathcal{Y} = \cup_{Y \in \mathcal{Y}} \partial Y$ defines a multicurve in V with the property that every element of $\mathcal{Y}$ is a component of $V \backslash \partial\mathcal{Y}$. By Lemma 9.16, each component γ of $\partial\mathcal{Y}$ satisfies $\ell_w(\gamma) < \epsilon_0$ and $\ell_{\hat{w}}(\gamma) < \epsilon_0$ for all $w \in [p, q]$. Consider the the product regions map $\Phi_{\partial\mathcal{Y}} \colon \mathcal{T}(V) \to \mathcal{P}(V|\partial\mathcal{Y})$. For each component Z of $V \backslash \partial\mathcal{Y}$ and each point $w \in [p, q]$, we may consider the projection $\hat{w}|_Z$ of $\Phi_{\partial\mathcal{Y}}(\hat{w})$ to $\mathcal{T}(Z)$.

Recall that $d_{\mathcal{P}(V|\partial\mathcal{Y})}(\Phi_{\partial\mathcal{Y}}(\hat{p}), \Phi_{\partial\mathcal{Y}}(\hat{q}))$ is the supremum of $d_{\mathcal{T}(Y)}(\hat{p}|_Y, \hat{q}|_Y)$ over all factors $\mathcal{T}(Y)$ of the product $\mathcal{P}(V|\partial\mathcal{Y})$, that is, over all components Y of $V \backslash \partial\mathcal{Y}$. Note that the components of the multicurve $\partial\mathcal{Y}$ count as annular components of $V \backslash \partial\mathcal{Y}$. Let Y be the component of $V \backslash \partial\mathcal{Y}$ maximizing this supremum. By Minsky's Theorem 3.11 we thus have

$$d_{\mathcal{T}(V)}(\hat{p}, \hat{q}) \stackrel{+}{\prec} d_{\mathcal{P}(V|\partial\mathcal{Y})}(\Phi_{\partial\mathcal{Y}}(\hat{p}), \Phi_{\partial\mathcal{Y}}(\hat{q})) = d_{\mathcal{T}(Y)}(\hat{p}|_Y, \hat{q}|_Y).$$

First suppose Y is not an element of $\mathcal{Y}$. We claim that $d_W(\hat{p}|_Y, \hat{q}|_Y)$ is uniformly bounded for all domains $W \sqsubset Y$. Note that by definition of product region factors, we have $d_W(\hat{p}|_Y, \hat{q}|_Y) \stackrel{+}{\asymp} d_W(\hat{p}, \hat{q})$. Clearly Y is the only component of $V \backslash \partial\mathcal{Y}$ containing W; since elements of $\mathcal{Y}$ *are* components of $V\backslash\partial\mathcal{Y}$ and yet $Y \notin \mathcal{Y}$, it follows that W cannot be contained in any element of $\mathcal{Y}$. If $W \in \mathcal{D}$, it follows that (p, q) is disjoint from $\mathcal{I}_W$, since otherwise W would be contained in an element of $\mathcal{Y}$ by construction. Hence in this case

$$d_W(\hat{p}, \hat{q}) \stackrel{+}{\asymp}_{\mathsf{C}} d_W(p, q) \leqslant \mathsf{M}/3.$$

If $W \notin \mathcal{D}$ but $W \in \Upsilon(x_{i-1}, x_i)$, then evidently $\bar{W}^{\Omega_i} \neq V$ and therefore the points $\tilde{p}_W$ and $\tilde{q}_W$ in the projection tuple (Definition 8.3) are equal (either $\pi_W(x_{i-1})$ or $\pi_W(x_i)$). Hence $d_W(\hat{p}, \hat{q}) \stackrel{+}{\asymp}_{\mathsf{C}} 0$ in this case as well. In the remaining case $W \notin \Upsilon(x_{i-1}, x_i)$ we have $d_W(x_{i-1}, x_i) < \mathsf{N}_W$ and therefore $d_W(p, x_i), d_W(q, x_i) \stackrel{+}{\prec}_{\mathsf{C}} \mathsf{N}_W$ by Lemma 9.13. Consequently $d_W(\hat{p}, \hat{q}) \stackrel{+}{\asymp}_{\mathsf{C}} d_W(p, q) \stackrel{+}{\asymp}_{\mathsf{C}} 0$ as before. Thus we have shown $d_W(\hat{p}|_Y, \hat{q}|_Y) \stackrel{+}{\asymp}_{\mathsf{C}} 0$ for every domain $W \sqsubset Y$.

Now let R denote the quantity from Lemma 3.35 for the pair p, q, and let $\hat{R}|_Y$ denote the corresponding quantity for the pair $\hat{p}|_Y, \hat{q}|_Y$. The lengths $\ell_{\hat{p}}(\gamma)$ and $\ell_{\hat{p}|_Y}(\gamma)$ are comparable for every essential curve γ in Y. Further, by construction, if $\ell_{\hat{p}}(\gamma) < \epsilon_0$, then $\ell_p(\gamma) = \ell_{\hat{p}}(\gamma) < \epsilon_0$. Thus every short curve at $\hat{p}|_Y$ is also short, with a comparable length, at p. The same holds for the points $\hat{q}|_Y$ and q. Therefore we evidently have $\hat{R}|_Y \stackrel{+}{\asymp} R$. Applying Lemma 3.35, and using our bound $d_W(\hat{p}|_Y, \hat{q}|_Y) \stackrel{+}{\asymp}_{\mathsf{C}} 0$ for all $W \sqsubset Y$, we now conclude

$$d_{\mathcal{T}(V)}(\hat{p}, \hat{q}) \stackrel{+}{\asymp} d_{\mathcal{T}(Y)}(\hat{p}|_Y, \hat{q}|_Y) \stackrel{+}{\asymp}_{\mathsf{C}} \hat{R}|_Y \stackrel{+}{\asymp} R \leqslant d_{\mathcal{T}(\Sigma)}(p, q).$$

It remains to suppose that Y is an element of $\mathcal{Y}$. Hence $[p, q] \subset \mathcal{I}_Y$. Using the product regions map $\Phi_{\partial\mathcal{Y}} \colon \mathcal{T}(\Sigma) \to \mathcal{P}(\Sigma|\partial\mathcal{Y})$ in the main Teichmüller space $\mathcal{T}(\Sigma)$, we may consider the $\mathcal{T}(Y)$–components $p|_Y$ and $q|_Y$ of $\Phi_{\partial\mathcal{Y}}(p)$ and $\Phi_{\partial Y}(q)$, respectively. We may now finally invoke Lemma 9.16(2) to obtain

$$d_{\mathcal{T}(Y)}(\hat{p}|_Y, \hat{q}|_Y) \stackrel{+}{\asymp}_{\mathsf{C}} d_{\mathcal{T}(Y)}(p|_Y, q|_Y).$$

Combining with the above estimate, and again using Theorem 3.11, we conclude

$$d_{\mathcal{T}(V)}(\hat{p}, \hat{q}) \stackrel{+}{\asymp} d_{\mathcal{T}(Y)}(\hat{p}|_Y, \hat{q}|_Y) \stackrel{+}{\asymp}_{\mathsf{C}} d_{\mathcal{T}(Y)}(p|_Y, q|_Y) \stackrel{+}{\asymp} d_{\mathcal{T}(\Sigma)}(p, q). \qquad \square$$

Recall from Definition 9.3 that $\mathcal{A}_V^\Omega = (\mathcal{I}_V \backslash M(V)) \cup C(V)$. Lemmas 9.8 and 9.10 together show that $C_i(V) \subset J \subset \mathcal{I}_V$. If we define $M_j(V) = \{\mathcal{I}_W \mid W \in \Omega_j \text{ with } W \subsetneq V\}$ then Lemma 9.12 furthermore shows that $M_j(V) \cap J = \varnothing$ and $C_j(V) \cap J = \varnothing$ for $j \neq i$; that is, we have $J \cap M(V) = J \cap M_i(V)$ and $J \cap C(V) = C_i(V)$. Combining these observations, we conclude that

$$\mathcal{A}_V^\Omega \cap J = \Big(\big(\mathcal{I}_V \backslash M(V)\big) \cup C(V)\Big) \cap J = \big(J \backslash M_i(V)\big) \cup C_i(V). \tag{9.24}$$

Since $M_i(V)$ is the union of the active intervals $\mathcal{I}_W$ of all domains $W \in \Omega_i$ with $W \swarrow_i V$ or $V \searrow_i W$, let us define

$$\mathcal{W}_- = \{W \in \Omega_i \mid W \swarrow_i V\} \qquad \text{and} \qquad \mathcal{W}_+ = \{W \in \Omega_i \mid V \searrow_i W\}.$$

Using these collections, we then define

$$y_1 = \sup\left(\{y\} \cup \bigcup_{W \in \mathcal{W}_-} \mathcal{I}_W\right) \qquad \text{and} \qquad z_1 = \inf\left(\{z\} \cup \bigcup_{W \in \mathcal{W}_+} \mathcal{I}_W\right),$$

where here each interval $\mathcal{I}_W$ is taken along $[x_0, x_n]$ and the supremum/infimum are taken with respect to the orientation of this interval from x_0 to x_n. Note that we have included $\{y\}$ and $\{z\}$ in the definition to both handle the case that $\mathcal{W}_\pm$ may be empty and to ensure $y_1, z_1 \in [y, z] = J$.

Lemma 9.25. *The points y_1, z_1 satisfy the following:*

(1) $d_V(x_{i-1}, y_1)$ and $d_V(z_1, x_i)$ are both at most $\mathsf{N}_V/3 + \mathsf{L}$.
(2) y_1 and z_1 lie in and occur in order along J.
(3) The interval $[y_1, z_1] \subset J$ is contained in $J \backslash M_i(V) \subset J \cap \mathcal{A}_V^\Omega$.
(4) Each point $w \in [z_1, z]$ satisfies $d_V(w, x_i) \stackrel{+}{\prec}_{\mathsf{C}} 0$.
(5) Each point $w \in [y, y_1]$ satisfies $d_V(x_{i-1}, w) \stackrel{+}{\prec}_{\mathsf{C}} 0$,

Proof. For (1), let us only consider $d_V(z_1, x_i)$. The construction of J (Lemma 9.8) ensures $d_V(x_i, z) \leqslant 7\mathsf{C}$. Hence the claim is immediate if $z_1 = z$. Otherwise, there is some $W \in \mathcal{W}_+$ so that $z_1 \in \mathcal{I}_W$. Thus ∂W is contained in the Bers marking at z_1 so that $d_V(z_1, x_i) \leqslant d_V(\partial W, x_i) + \mathsf{L}$. Since $V \overset{i}{\searrow} W$, the definition of encroachment and the fact that Ω_i is wide now give

$$d_V(x_i, z_1) \leqslant d_V(x_i, \partial W) + \mathsf{L} \leqslant \mathcal{E}_{\Omega_i}(V) + \mathsf{L} \leqslant \mathsf{N}_V/3 + \mathsf{L}.$$

For (2), since the pairs y, y_1 and z_1, z occur in order by construction, it suffices to show y_1, z_1 occur in order along $[x_0, x_n]$ as this will force $[y_1, z_1] \subset [y, z] = J$. By means of contradiction, let us instead suppose z_1, y_1 occur in order. First note that having $y_1 \in [x_i^V, x_n]$ would imply (by Theorem 3.19)

$$d_V(x_{i-1}^V, y_1) \geqslant d_V(x_{i-1}^V, x_i^V) - \mathsf{B} \geqslant d_V(x_{i-1}, x_i) - 2\mathsf{C} - \mathsf{B} \geqslant \mathsf{N}_V - 3\mathsf{C}$$

and hence $d_V(x_{i-1}, y_1) \geqslant \mathsf{N}_V - 4\mathsf{C}$. Since $\mathsf{N}_V \geqslant 30\mathsf{C}$, this is incompatible with (1). Hence we must in fact have $y_1 \in [z_1, x_i^V]$, in which case Theorem 3.19 now gives

$$d_V(y_1, x_i^V) \leqslant d_V(z_1, x_i^V) + \mathsf{B} \leqslant d_V(z_1, x_i) + \mathsf{C} + \mathsf{B} \leqslant \mathsf{N}_V/3 + 3\mathsf{C},$$

where we have again utilized (1). Using $V \in \Upsilon^c(x_{i-1}, x_i)$ together with one more application of (1), this now leads to the contradiction:

$$\begin{aligned}\mathsf{N}_V \leqslant d_V(x_{i-1}, x_i) &\leqslant d_V(x_{i-1}, y_1) + d_V(y_1, x_i^V) + d_V(x_i^V, x_i)\\ &\leqslant \mathsf{N}_V/3 + \mathsf{L} + \mathsf{N}_V/3 + 3\mathsf{C} + \mathsf{C} < \mathsf{N}_V.\end{aligned}$$

Since $[y_1, z_1] \subset J$, the assertion $[y_1, z_1] \subset J \backslash M_i(V)$ of (3) is clear: $M_i(V)$ is the union of intervals $\mathcal{I}_W$ for $W \subsetneq V$ with $W \in \mathcal{W}_- \cup \mathcal{W}_+$. By definition of the points y_1, z_1, if $W \in \mathcal{W}_-$ then $\mathcal{I}_W \subset [x_0, y_1]$ and if $W \in \mathcal{W}_+$ then $\mathcal{I}_W \subset [z_1, x_n]$. Hence $M_i(V)$ is disjoint from $[y_1, z_1]$, which proves the claim.

For (4), if $w \in [z_1, z] \subset [z_1, x_i^V]$, then as above Theorem 3.19 and (1) give

$$\begin{aligned}d_V(w, x_i) &\leqslant d_V(z_1, w) + d_V(w, x_i^V) + d_V(x_i^V, x_i)\\ &\leqslant d_V(z_1, x_i^V) + \mathsf{B} + \mathsf{C} \leqslant d_V(z_1, x_i) + 2\mathsf{C} + \mathsf{B} \stackrel{+}{\prec}_{\mathsf{C}} 0.\end{aligned}$$

The argument for (5) is symmetric. □

The significance of the subinterval $[y_1, z_1]$ is highlighted by the next lemma.

Lemma 9.26. *Every point $w \in [y_1, z_1]$ satisfies $d_{\mathcal{T}(V)}(\hat{w}, \bar{w}) \stackrel{+}{\prec}_{\mathsf{C}} 0$.*

Proof. We use Lemma 3.35 and show that $\hat{w}$ and $\bar{w}$ agree in all subsurfaces and have the same short curves with the same lengths. Consider any domain $Z \sqsubset V$ and let $\tilde{w}_Z \subset \mathcal{C}(Z)$ be as in Proposition 8.4. Then $d_Z(\hat{w}, \tilde{w}_Z) \stackrel{+}{\prec}_{\mathsf{C}} 0$ by construction.

Since $\bar{w}$ is simply the $\mathcal{T}(V)$–component of w, we also observe that $d_Z(w, \bar{w}) \stackrel{+}{\prec} 0$. Thus to bound $d_Z(\hat{w}, \bar{w})$ it suffices to bound $d_Z(w, \tilde{w}_Z)$.

If $Z \notin \Upsilon(x_{i-1}, x_i)$ or if $Z \in \Upsilon(x_{i-1}, x_i)$ with $\bar{Z}^{\Omega_i} = V$, then $\tilde{w}_Z = \pi_Z(w)$ by definition and hence $d_Z(w, \tilde{w}_Z) \stackrel{+}{\prec} 0$ is immediate. So suppose $Z \in \Upsilon(x_{i-1}, x_i)$ with $\bar{Z}^{\Omega_i} = W \subsetneq V$. We consider the case $W \swarrow_i V$, the opposite possibility $V {}_i\!\searrow W$ being similar. By definition we now have $\tilde{w}_Z = \pi_Z(x_i)$. On the other hand, the construction of y_1 implies $\mathcal{I}_W \subset [x_0, y_1]$. As $\mathcal{I}_Z \subset \mathcal{I}_W$ by Lemmas 9.8 and 9.10 (applied with W in place of V), it follows that $\mathcal{I}_Z$ is contained in $[x_0, y_1]$ and that w, x_i^V lie in the same component of $[x_0, x_n] \backslash \mathcal{I}_Z$. Therefore $d_Z(x_i^V, x_i) \leqslant \mathsf{M}$ by Lemma 9.11 and $d_Z(w, x_i^V) \leqslant \mathsf{M}/3$ by Lemma 3.26. The triangle inequality thus gives $d_Z(w, \tilde{w}_Z) = d_Z(w, x_i) \stackrel{+}{\prec} 0$ here as well.

By Lemma 3.35 it remains to bound the quantity R associated to the two points $\hat{w}, \bar{w} \in \mathcal{T}(V)$. For this, it suffices to bound the ratio $\ell_{\bar{w}}(\gamma)/\ell_{\hat{w}}(\gamma)$, from above and below, for every curve γ that is short on either $\bar{w}$ or $\hat{w}$. Note that $\ell_w(\gamma)$ and $\ell_{\bar{w}}(\gamma)$ agree up to bounded multiplicative error for all essential curves γ in V, thus we may instead bound the ratio $\ell_w(\gamma)/\ell_{\hat{w}}(\gamma)$. Suppose now that γ is an essential curve in V with $\ell_{\hat{w}}(\gamma) < \epsilon_0$. Then by Definition 9.14, $\ell_w(\gamma)/\ell_{\hat{w}}(\gamma) = 1$. Conversely, suppose γ is a curve in V with $\ell_w(\gamma) < \epsilon_0$. We show that γ satisfies condition (9.15), it will then follow from the definition of $\hat{w}$ that $\ell_{\hat{w}}(\gamma) = \ell_w(\gamma)$. Let $Z \sqsubset V$ be any domain. Since $\ell_w(\gamma) < \epsilon_0$, we have $\gamma \in \mu_w$. Thus $d_Z(\gamma, w) \leqslant \mathsf{L}$. Hence (9.15) is satisfied if $\tilde{w}_Z = \pi_Z(w)$. If $\tilde{w}_Z \neq \pi_Z(w)$, then $Z \in \Upsilon(x_{i-1}, x_i)$ with $W = \bar{Z}^{\Omega_i} \neq V$. Let us suppose $W \swarrow_i V$ so that $\tilde{w}_Z = \pi_Z(x_i)$, the reverse possibility $V {}_i\!\searrow W$ being similar. As above, we have that $d_Z(w, x_i) \leqslant 4\mathsf{M}/3$ and therefore conclude

$$d_Z(\gamma, \tilde{w}_Z) = d_Z(\gamma, x_i) \leqslant d_Z(\gamma, w) + d_Z(w, x_i) \leqslant \mathsf{L} + 4\mathsf{M}/3 \leqslant 2\mathsf{M}$$

as required. This establishes (9.15) for γ and proves the claim. □

We now establish the properties mentioned in Strategy 9.21 that will hold for the segments of the yet-to-be-constructed piecewise geodesic path from x_{i-1} to x_i.

Lemma 9.27. *Let $[p, q] \subset J$ be a subgeodesic satisfying either*

- *$[p, q]$ is squarely covered by $\mathcal{D}$, or*
- *$[p, q] \subset [y_1, z_1]$.*

Then $[p, q]$ is contained in $J \cap \mathcal{A}_V^\Omega$ and $d_{\mathcal{T}(V)}(\hat{p}, \hat{q}) \stackrel{+}{\prec}_{\mathsf{C}} d_{\mathcal{T}(\Sigma)}(p, q)$.

Proof. To see that $[p, q] \subset \mathcal{A}_V^\Omega$, we simply note that $[y_1, z_1]$ is contained in $J \backslash M_i(V)$ by Lemma 9.25(3), and that each squarely covered interval is contained in $\mathcal{I}_Y \subset C_i(V)$ for some $Y \in \mathcal{D}$. Thus clearly $[p, q] \subset (J \backslash M_i(V)) \cup C_i(V) = J \cap \mathcal{A}_V^\Omega$.

If $[p, q]$ is squarely covered by $\mathcal{D}$, the bound on $d_{\mathcal{T}(V)}(\hat{p}, \hat{q})$ is simply Lemma 9.23. If instead $[p, q] \subset [y_1, z_1]$, then Lemma 9.26 implies $d_{\mathcal{T}(V)}(\bar{p}, \hat{p})$ and $d_{\mathcal{T}(V)}(\bar{q}, \hat{q})$ are both bounded in terms of C. Therefore

$$d_{\mathcal{T}(V)}(\hat{p}, \hat{q}) \stackrel{+}{\prec}_{\mathsf{C}} d_{\mathcal{T}(V)}(\bar{p}, \bar{q})$$

Since the metric in $\mathcal{P}(\Sigma|\partial V)$ is a sup metric, by Minsky's Theorem 3.11 we have

$$d_{\mathcal{T}(V)}(\bar{p}, \bar{q}) \leqslant d_{\mathcal{P}(\Sigma|\partial V)}(\Phi_{\partial V}(p), \Phi_{\partial V}(q)) \leqslant d_{\mathcal{T}(\Sigma)}(p, q) + \mathsf{D}_0.$$

Combining with the previous inequality thus proves the lemma in this case. □

In contrast to Lemma 9.27, we have the following for certain subintervals of J:

Lemma 9.28. *Let $[p, q]$ be a geodesic segment in $\mathcal{T}(\Sigma)$ such that either*

- $[p,q] = [x_{i-1}, y]$, *or* $[p,q] = [z, x_i]$, *or*
- (p,q) *is contained in* J *and disjoint from* $C_i(V) \cup [y_1, z_1]$.

Then $d_Z(\hat{p}, \hat{q}) \stackrel{+}{\prec}_{\mathsf{C}} 0$ *for every domain* $Z \sqsubset V$.

Proof. To ease notation, set $J' = \{x_{i-1}\} \cup J \cup \{x_i\}$ and note that $p, q \in J'$. By Lemma 9.13, each point $w \in J'$ satisfies the condition of Definition 8.3

$$d_Z(x_{i-1}, w) + d_Z(w, x_i) \leqslant d_Z(x_{i-1}, x_i) + 9\mathsf{C} \quad \text{ for all } Z \sqsubset V$$

and so determines a consistent tuple $(\tilde{w}_Z) \in \prod_{Z \sqsubset V} \mathcal{C}(Z)$. Recall from Definitions 8.7 and 9.14 that $\hat{w} \in \mathcal{T}(V)$ then satisfies $d_Z(\tilde{w}_Z, \hat{w}) \stackrel{+}{\prec}_{\mathsf{C}}$ for any $Z \sqsubset V$.

Let us now fix a domain $Z \sqsubset V$ and bound $d_Z(\hat{p}, \hat{q})$ provided any of the conditions hold. First suppose $Z \notin \Upsilon(x_{i-1}, x_i)$, so that by Definition 8.3 $\tilde{p}_Z = \pi_Z(p)$ and $\tilde{q}_Z = \pi_Z(q)$. In this case we have $d_Z(x_{i-1}, x_i) < \mathsf{N}_V$, so the above condition implies

$$d_Z(x_{i-1}, w) + d_Z(w, x_i) \leqslant \mathsf{N}_V + 9\mathsf{C}$$

for every point $w \in J'$. Therefore we conclude that

$$d_Z(\hat{p}, \hat{q}) \stackrel{+}{\prec}_{\mathsf{C}} d_Z(\tilde{p}_Z, \tilde{q}_Z) = d_Z(p,q) \leqslant d_Z(p, x_i) + d_Z(x_i, q) \stackrel{+}{\prec} 2(\mathsf{N}_V + 9\mathsf{C}).$$

Next suppose $Z \in \Upsilon(x_{i-1}, x_i)$, and set $W = \bar{Z}^{\Omega_i} \in \Omega_i$. If $W \swarrow_i V$, then by definition $\tilde{p}_Z = \pi_Z(x_i) = \tilde{q}_Z$, and we have

$$d_Z(\hat{p}, \hat{q}) \stackrel{+}{\prec}_{\mathsf{C}} d_Z(\tilde{p}_Z, \tilde{q}_Z) = d_Z(x_i, x_i) \leqslant \mathsf{L}.$$

Similarly if $W {}_i\!\searrow V$, then $\tilde{p}_Z = \pi_Z(x_{i-1}) = \tilde{q}_Z$ and we again find $d_Z(\hat{p}, \hat{q}) \stackrel{+}{\prec}_{\mathsf{C}} 0$. The remaining possibility is $W = V$, in which Z contributes to V in Ω_i. In this case, $\tilde{p}_Z = \pi_Z(p)$ and $\tilde{q}_Z = \pi_Z(q)$, so that $d_Z(\hat{p}, \hat{q}) \stackrel{+}{\prec}_{\mathsf{C}} d_Z(p,q)$. Hence it suffices to bound this latter quantity $d_Z(p,q)$. We consider two cases:

First, suppose $Z = V$ itself. If $[p,q] = [x_{i-1}, y]$ or if $[p,q] = [z, x_i]$, then Lemma 9.8 provides the desired bound:

$$d_Z(p,q) \in \Big\{ d_V(x_{i-1}, y), d_V(z, x_i) \Big\} \leqslant 7\mathsf{C}.$$

Otherwise (p,q) is contained in J and disjoint from $C_i(V) \cup [y_1, z_1]$. It follows that $[p,q]$ is either contained in $[y, y_1]$ or $[z_1, z]$. In the latter case, Lemma 9.25(4) implies

$$d_V(p,q) \leqslant d_V(p, x_i) + d_V(x_i, q) \stackrel{+}{\prec}_{\mathsf{C}} 0,$$

and in the former case Lemma 9.25(5) similarly implies $d_V(p,q) \stackrel{+}{\prec}_{\mathsf{C}} 0$.

Second, suppose $Z \subsetneq V$. In that case we know that $\mathcal{I}_Z \subset J = [y, z]$ (Lemma 9.10) and that $d_Z(x_{i-1}, x_{i-1}^V)$ and $d_Z(x_i, x_i^V)$ are both at most M (Lemma 9.11). By Lemma 3.26 and the triangle inequality, it follows that

$$d_Z(z, x_i) \leqslant d_Z(z, x_i^V) + d_Z(x_i^V, x_i) \leqslant \mathsf{M}/3 + \mathsf{M} \leqslant 2\mathsf{M}$$

and similarly that $d_Z(x_{i-1}, y) \leqslant 2\mathsf{M}$. This handles the case that $[p,q]$ equals $[x_{i-1}, y]$ or $[z, x_i]$. If instead (p,q) is contained in J and disjoint from $C_i(V) \cup [y_1, z_1]$, then evidently $(p,q) \cap \mathcal{I}_Z = \varnothing$ due to the fact that $\mathcal{I}_Z \subset C_i(V)$ by definition. Therefore $d_Z(p,q) \leqslant \mathsf{M}/3$ by Lemma 3.26 and the lemma is proven. ☐

In order to decompose J into subsegments that satisfy either Lemma 9.27 or 9.28 above, we will use endpoints of active intervals $\mathcal{I}_Z$ for domains $Z \in \mathcal{D}$. To this end, let $\mathcal{D}_R$ denote the collection of all $Z \in \mathcal{D}$ such that $\mathcal{I}_Z$ intersects $[z_1, z]$. Define $\mathcal{D}_L$ symmetrically. Observe that $z_1 = z$ forces $\mathcal{D}_R = \varnothing$ (since each $Z \in \mathcal{D}$ has $\mathcal{I}_Z \subset (y, z)$ by Lemma 9.10) and similarly for $\mathcal{D}_L$. Using the notation of §4.1, we write $\underline{\mathcal{D}_L}_{x_0}^{x_n}$ and $\underline{\mathcal{D}_R}_{x_0}^{x_n}$ for the set of domains in $\mathcal{D}_L$ and $\mathcal{D}_R$, respectively, that are maximal with respect to the order

$$Z \prec_{[x_0,x_n]} Y \iff Z \sqsubset Y \text{ and } \mathcal{I}_Z \subset \mathcal{I}_Y \text{ along } [x_0, x_n].$$

Lemma 9.29. *The collections $\underline{\mathcal{D}_L}_{x_0}^{x_n}$ and $\underline{\mathcal{D}_R}_{x_0}^{x_n}$ have uniformly bounded cardinality. That is $\left|\underline{\mathcal{D}_L}_{x_0}^{x_n}\right|, \left|\underline{\mathcal{D}_R}_{x_0}^{x_n}\right| \stackrel{+}{\prec}_{\mathsf{N}} 0$.*

Proof. We only consider $\underline{\mathcal{D}_R}_{x_0}^{x_n}$. We may assume $z_1 \neq z$, for otherwise $\mathcal{D}_R = \varnothing$ and there is nothing to prove. Thus, by definition of z_1, we may choose $W \in \Omega_i$ such that $V \overset{i}{\searrow} W$ and so that z_1 is the left endpoint of $\mathcal{I}_W$. Since Ω_i is assumed to be wide, we have that $d_V(\partial W, x_i) \leqslant \mathcal{E}_{\Omega_i}(V) \leqslant \mathsf{N}_V/3$.

We claim that for every $Z \in \mathcal{D}_R$, the multicurves ∂Z and ∂W are disjoint. Indeed, the definition of $\mathcal{D}_R$ ensures that $\mathcal{I}_Z$ either intersects or occurs to the right of $\mathcal{I}_W$. If $\partial Z \pitchfork \partial W$, then $Z \pitchfork W$ and hence W is necessarily time-ordered before Z along $[x_0, x_n]$. Lemma 9.2 implies we also have the time ordering $W \lessdot Z$ along $[x_{i-1}, x_i]$. But, since Z contributes to V in Ω_i and $V \overset{i}{\searrow} W$, this contradicts (SO4).

Choose $\alpha \in \pi_V(x_{i-1})$ and $\beta \in \pi_V(x_i)$ realizing $d_V(x_{i-1}, x_i)$ and fix a geodesic $\beta = \gamma_0, \ldots, \gamma_m = \alpha$ in $\mathcal{C}(V)$. Since $\Upsilon^\ell(x_{i-1}, x_i) \subset \Omega_i$ by insulation, we have $\mathcal{D} \subset \Upsilon^c(x_{i-1}, x_i)$; that is $d_Z(x_{i-1}, x_i) \geqslant \mathsf{N}_Z$ for each $Z \in \mathcal{D}$. Exactly as in the proof of Lemma 4.1, the bounded geodesic image theorem implies that each $Z \in \mathcal{D}$ is disjoint from one of the curves γ_j. If we fix $Z \in \mathcal{D}_R$ and let $0 \leqslant j \leqslant m$ be such that Z is disjoint from γ_j, it follows that

$$j = d_V(\beta, \gamma_j) \leqslant d_V(x_i, \gamma_j) \leqslant d_V(x_i, \partial W) + d_V(\partial W, \partial Z) + d_V(\partial Z, \gamma_j) \leqslant \mathcal{E}_{\Omega_i}(V) + 2.$$

This proves that if we define

$$\mathcal{Y} = \{Y \mid Y \text{ is a connected component of } V \backslash \gamma_j \text{ for some } j \leqslant \mathcal{E}_{\Omega_i}(V) + 2\},$$

then each $Z \in \mathcal{D}_R$ satisfies $Z \sqsubset Y$ for some $Y \in \mathcal{Y}$. Notice that, since Ω_i is wide, $|\mathcal{Y}| \leqslant 2(\mathcal{E}_{\Omega_i}(V) + 3) \leqslant 2\mathsf{N}_V/3 + 6 \leqslant \mathsf{N}_V$.

For each $Y \in \mathcal{Y}$ we consider the collection

$$\mathcal{P}(Y) = \{U \sqsubset Y \mid d_U(x_{i-1}, x_i) \geqslant \mathsf{N}_U\}.$$

Now choose any $Z \in \underline{\mathcal{D}_R}_{x_0}^{x_n}$, that is a maximal element of $\mathcal{D}_R$ with respect to the partial order $\prec_{[x_0,x_n]}$. Choose some $Y \in \mathcal{Y}$ so that $Z \sqsubset Y$. Since $d_Z(x_{i-1}, x_i) \geqslant \mathsf{N}_Z$, we have $Z \in \mathcal{P}(Y)$ as well. We claim that furthermore $Z \in \underline{\mathcal{P}}_{x_0}^{x_n}(Y)$. To see this, consider any $U \in \mathcal{P}(Y)$ with $Z \sqsubset U$ and $\mathcal{I}_Z \subset \mathcal{I}_U$. Since $Z \sqsubset U \sqsubset Y \subsetneq V$, the fact $\bar{Z}^{\Omega_i} = V$ forces $\bar{U}^{\Omega_i} = V$ as well. As $U \subsetneq V$ and $U \in \Upsilon(x_{i-1}, x_i)$, we see that U contributes to V and in fact that $U \in \mathcal{D}$. Finally, since $\mathcal{I}_Z$ intersects $[z_1, z]$, the same holds for $\mathcal{I}_U \supset \mathcal{I}_Z$. Therefore $U \in \mathcal{D}_R$. Since Z is $\prec_{[x_0,x_n]}$–maximal in $\mathcal{D}_R$, it follows that $U = Z$. Hence $Z \in \underline{\mathcal{P}}_{x_0}^{x_n}(Y)$ as claimed. This proves that each element of $\underline{\mathcal{D}_R}_{x_0}^{x_n}$ is contained in $\underline{\mathcal{P}}_{x_0}^{x_n}(Y)$ for some $Y \in \mathcal{Y}$. Thus we have

$$\underline{\mathcal{D}_R}_{x_0}^{x_n} \subset \bigcup_{Y \in \mathcal{Y}} \underline{\mathcal{P}}_{x_0}^{x_n}(Y).$$

Applying Lemma 4.1 with the thresholds $\mathsf{N}_{\xi(\Sigma)} \leqslant \cdots \leqslant \mathsf{N}_{-1} = \mathsf{N}$ gives a bound $\left|\underline{\mathcal{P}}_{x_0}^{x_n}(Y)\right| \stackrel{+}{\prec}_{\mathsf{N}} 0$ for every Y. Since $|\mathcal{Y}| \leqslant \mathsf{N}$, we conclude $\left|\underline{\mathcal{D}_L}_x^y\right| \stackrel{+}{\prec}_{\mathsf{N}} 0$, as desired. $\square$

We are now finally ready to complete the proof of the Theorem:

Proof of Theorem 9.4–Nonannular case. Let E denote the union of $\{y, y_1, z_1, z\}$ with the set of all endpoints of active intervals $\mathcal{I}_Z$ for $Z \in \underline{\mathcal{D}_L}_{x_0}^{x_n}$ or $Z \in \underline{\mathcal{D}_R}_{x_0}^{x_n}$. This nonempty set is contained in J and has uniformly bounded cardinality by Lemma 9.29. Let us write $E = \{e_1, \dots, e_{k-1}\}$ ordered along J as

$$y = e_1 < e_2 < \cdots < e_{k-1} = z.$$

We also define $e_0 = x_{i-1}$ and $e_k = x_i$. The points $e_0, \dots, e_k$ therefore define a piecewise geodesic path in $\mathcal{T}(\Sigma)$ from x_{i-1} to x_i:

$$[e_0, e_1][e_1, e_2] \cdots [e_{k-1}, e_k]$$

We claim that each segment $[p, q]$ of this concatenation satisfies the hypotheses of either Lemma 9.28 or Lemma 9.27. Indeed, the first and last segments $[x_{i-1}, y]$ and $[z, x_i]$ satisfy Lemma 9.28 by fiat, and any subsegment of $[y_1, z_1]$ satisfies Lemma 9.27. If $[p, q]$ is not covered by the previous sentence, then $[p, q]$ is contained in $[y, y_1]$ or $[z_1, z]$. By symmetry, let us suppose it is the former. We may assume (p, q) intersects $C_i(V)$, for otherwise it satisfies Lemma 9.28. Now let $Z \in \mathcal{D}$ be any domain for which $\mathcal{I}_Z$ intersects (p, q). Since $[p, q] \subset [y, y_1]$ we evidently have $Z \in \mathcal{D}_L$ and may choose some $Y \in \underline{\mathcal{D}_L}_{x_0}^{x_n}$ with $Z \sqsubset Y$ and $\mathcal{I}_Z \subset \mathcal{I}_Y$. It follows that $[p, q]$ intersects $\mathcal{I}_Y$ as well. Since the points p, q are consecutive in the set E, which by definition contains both endpoints of $\mathcal{I}_Y$, it must be that $[p, q] \subset \mathcal{I}_Y$. Therefore $[p, q]$ is squarely covered by $\mathcal{D}$ and satisfies Lemma 9.27.

Taking resolutions produces a sequence of points $\hat{x}_{i-1} = \hat{e}_0, \dots, \hat{e}_k = \hat{x}_i$ in $\mathcal{T}(V)$. Let $P \subset \{1, \dots, k\}$ be the set of indices $1 \leqslant j \leqslant k$ such that the segment $[e_{j-1}, e_j]$ satisfies Lemma 9.27. Since the intervals $[e_{j-1}, e_j]$ with $j \in P$ have disjoint interiors and are each contained in $J \cap \mathcal{A}_V^\Omega$, applying Lemma 9.27 implies that

$$\sum_{j \in P} d_{\mathcal{T}(V)}(\hat{e}_{j-1}, \hat{e}_j) \overset{+}{\prec}_{\mathsf{C}} \sum_{j \in P} d_{T(\Sigma)}(e_{j-1}, e_j) \leqslant \int_y^z \mathbb{1}_{\mathcal{A}_V^\Omega}. \tag{9.30}$$

Note that in the first inequality above we have used the fact that k is uniformly bounded (Lemma 9.29) to combine the additive errors from each of the $|P| \leqslant k$ applications of Lemma 9.27 into a single additive error depending only on C.

Now let $Q = \{1, \dots, k\} \backslash P$ be the set of remaining indices. By the above, for each $j \in Q$ the segment $[e_{j-1}, e_j]$ satisfies Lemma 9.28; consequently we have

$$d_Z(\hat{e}_{j-1}, \hat{e}_j) \overset{+}{\prec}_{\mathsf{C}} 0 \quad \text{for each } j \in Q \text{ and every domain } Z \sqsubset V. \tag{9.31}$$

For each $j \in Q$ let Γ_j denote the set of essential curves α in V such that either $\ell_{\hat{e}_{j-1}}(\alpha) < \epsilon_0'/2$ or $\ell_{\hat{e}_j}(\alpha) < \epsilon_0'/2$. Since a point in $\mathcal{T}(V)$ can have at most $\xi(V)$ disjoint curves, we see that $|\Gamma_j| \leqslant 2\xi(S)$. Setting $\Gamma = \cup_{j \in Q} \Gamma_j$ now gives a set of uniformly bounded cardinality.

Note that if $\Gamma_j = \varnothing$, then the quantity $\hat{R}_j$ in Lemma 3.35 for the pair $\hat{e}_{j-1}, \hat{e}_j$ is uniformly bounded and hence that lemma implies $d_{\mathcal{T}(V)}(\hat{e}_{j-1}, \hat{e}_j) \overset{+}{\prec}_{\mathsf{C}} 0$. Thus if Γ were empty, combining the inequalities (9.30) and (9.31) above would prove the proposition. However, since the points $\hat{w}$ for $w \in J$ are allowed to be thin (c.f. Definition 9.14), Γ may be nonempty and we must work a bit harder.

Claim 9.32. *If $A \sqsubset V$ is an annulus with $\partial A = \alpha \in \Gamma$, then A does not contribute to V in Ω_i. Therefore* $\operatorname{diam}_{\mathcal{C}(A)}(\pi_A(\hat{e}_0) \cup \cdots \cup \pi_A(\hat{e}_k)) \overset{+}{\prec}_{\mathsf{C}}$.

Proof of claim. By contradiction, suppose that A contributes to V. The hypothesis implies there is some $0 \leqslant j \leqslant k$ such that $\ell_{\hat{e}_j}(\partial A) < \epsilon_0'/2$ and such that either

$j \in Q$ or $j+1 \in Q$. By construction, $\hat{e}_0 = \hat{x}_{i-1}$ and $\hat{e}_k = \hat{x}_i$ are both thick; therefore it must be that $0 < j < k$. The construction of $\hat{e}_j$ (Definition 9.14) implies in this case that $\ell_{e_j}(\partial A) = \ell_{\hat{e}_j}(\partial A) < \epsilon_0'/2$. Therefore the point $e_j \in J$ evidently lies in the interior of the active interval $\mathcal{I}_A$ of A. Since $\mathcal{I}_A \subset J$ by Lemma 9.10, this rules out both possibilities $e_j = y$ and $e_j = z$; hence in fact $1 < j < k-1$. Since e_j is in the *interior* of $\mathcal{I}_A$, we see that $\mathcal{I}_A \subset C_i(V)$ intersects the interiors of both $[e_{j-1}, e_j]$ and $[e_j, e_{j+1}]$. It follows neither (e_{j-1}, e_j) nor (e_j, e_{j+1}) is disjoint from $C_i(V)$, and thus that neither of these intervals satisfies Lemma 9.28. But this contradicts the assumption that either $j \in Q$ or $j+1 \in Q$. Hence A cannot contribute to V.

For the second conclusion, if $A \in \Upsilon(x_{i-1}, x_i)$ then the above implies that either $\bar{A}^{\Omega_i} \swarrow_i V$ or $V \searrow_i \bar{A}^{\Omega_i}$. In the former case we have $d_A(\hat{e_j}, x_i) \stackrel{+}{\prec}_{\mathsf{C}} 0$ for all $0 \leqslant j \leqslant k$, and in the latter case we have $d_A(\hat{e}_j, x_{i-1}) \stackrel{+}{\prec}_{\mathsf{C}} 0$. Otherwise $A \notin \Upsilon(x_{i-1}, x_i)$ so that $d_A(x_{i-1}, x_i) \leqslant \mathsf{N}_A$. In this case for each $0 \leqslant j \leqslant k$ we have $d_A(\hat{e}_j, e_j) \stackrel{+}{\prec}_{\mathsf{C}} 0$ by construction and, by Lemma 9.13, that

$$d_A(\hat{e}_j, x_i) \stackrel{+}{\prec}_{\mathsf{C}} d_A(e_j, x_i) \leqslant d_A(x_{i-1}, e_j) + d_A(e_j, x_i) \stackrel{+}{\prec}_{\mathsf{C}} d_A(x_{i-1}, x_i) \leqslant \mathsf{N}_A.$$

In any case, $\cup_j \pi_A(\hat{e}_j)$ lies within bounded distance of either $\pi_A(x_{i-1})$ or $\pi_A(x_i)$. $\square$

For any essential curve α on V, we now define a transformation f_α of $\mathcal{T}(V)$ to itself by utilizing the product region $\mathcal{P}(V|\alpha) = \mathcal{T}(V \backslash \alpha) \times \mathcal{T}(\alpha)$. Recalling that $\mathcal{T}(\alpha) = \mathbb{H}^2$, let $h_\alpha \colon \mathcal{T}(\alpha) \to \mathcal{T}(\alpha)$ be the map that pushes points vertically down to below the horizontal line $1/\epsilon_0'$; that is, $h_\alpha(x,y) = (x, \min\{y, \frac{1}{\epsilon_0'}\})$ for $(x,y) \in \mathbb{H}^2$. Conjugating with Φ_α then gives a transformation $f_\alpha = \Phi_\alpha^{-1} \circ (\mathrm{id} \times h_\alpha) \circ \Phi_\alpha$ from $\mathcal{T}(V)$ to itself. Observe that f_α is the identity on the complement of the thin region $\mathcal{H}_{\epsilon_0', \alpha}$, and therefore fixes every point of $w \in \mathcal{T}(V)$ with $\ell_w(\alpha) \geqslant \epsilon_0'$. The fact that f_α only makes α longer and does not affect twisting leads easily to the following:

Claim 9.33. *For every point $w \in \mathcal{T}(V)$ we have:*

- $d_Z(w, f_\alpha(w)) \stackrel{+}{\prec} 0$ *for every domain $Z \sqsubset V$.*
- $\log(\ell_{f_\alpha(w)}(\gamma)) \stackrel{+}{\succ} \log(\min\{\ell_w(\gamma), \epsilon_0'\})$ *for every essential curve γ on V.*

Proof of Claim. It is clear that any short marking at w is also a short marking for $f_\alpha(w)$; whence the first bullet. The second bullet is immediate for the curve $\alpha = \gamma$. If $\gamma \neq \alpha$ is disjoint from α, then γ is essential in $\mathcal{T}(V \backslash \alpha)$ and so the lengths $\ell_w(\gamma)$ and $\ell_{f_\alpha(w)}(\gamma)$ coarsely agree. Finally suppose $\gamma \pitchfork \alpha$. If $\ell_w(\alpha) \geqslant \epsilon_0'$ then $f_\alpha(w) = w$ and there is nothing to prove. Otherwise $\ell_w(\alpha) < \epsilon_0$ and so $\ell_{f_\alpha(w)}(\alpha) = \epsilon_0'$ by construction. Thus necessarily $\ell_{f_\alpha(w)}(\gamma) > \epsilon_0'$ since ϵ_0' is smaller than the Margulis constant. $\square$

Let us list the curves in Γ as $\Gamma = \{\alpha_1, \ldots, \alpha_m\}$ and write $f_t = f_{\alpha_t}$. For each $0 \leqslant j \leqslant k$, set $\hat{e}_j^0 = \hat{e}_j$ and then recursively set $\hat{e}_j^t = f_t(\hat{e}_j^{t-1})$ for $1 \leqslant t \leqslant m$. Since the points $\hat{e}_0$ and $\hat{e}_k$ are thick by construction, each map f_t fixes these two points and we have $\hat{e}_0^m = \hat{e}_0 = \hat{x}_{i-1}$ and $\hat{e}_k^m = \hat{e}_k = \hat{x}_i$. Hence to prove the proposition it suffices to bound $d_{\mathcal{T}(V)}(\hat{e}_0^m, \hat{e}_k^m)$.

Applying Claim 9.33 successively for the maps $f_1, \ldots, f_m$ gives $d_Z(\hat{e}_j, \hat{e}_j^m) \stackrel{+}{\prec}_{\mathsf{C}} 0$ for each $Z \sqsubset V$ (recall that $m \stackrel{+}{\prec}_{\mathsf{C}} 0$). Therefore (9.31) can now be restated as

$$d_Z(\hat{e}_{j-1}^m, \hat{e}_j^m) \stackrel{+}{\prec}_{\mathsf{C}} 0 \quad \text{for each } j \in Q \text{ and every domain } Z \sqsubset V.$$

Furthermore, when $j \in Q$, since Γ contains every short curve at $\hat{e}_{j-1}$ or $\hat{e}_j$ and each of these curves gets lengthened by one of the maps f_t, repeated applications of Claim 9.33 shows that the points $\hat{e}_{j-1}^m$ and $\hat{e}_j^m$ are uniformly thick. Therefore the quantity $\hat{R}^m_j$ from Lemma 3.35 associated to this pair is uniformly bounded above (in terms of C), and we may promote the above bound to yield the following:

$$d_{\mathcal{T}(V)}(\hat{e}^m_{j-1}, \hat{e}^m_j) \stackrel{+}{\prec}_{\mathsf{C}} 0 \quad \text{for each } j \in Q. \tag{9.34}$$

The only remaining step is to promote the bound in (9.30) to the new points $\hat{e}^m_{j-1}, \hat{e}^m_j$ for $j \in P$. The key point here is that the diameter bound from Claim 9.32 implies our transformations f_t are coarsely 1–Lipschitz for the points in question:

Claim 9.35. *For any $1 \leqslant j \leqslant k$ and $0 \leqslant t \leqslant m$ we have*

$$d_{\mathcal{T}(V)}(\hat{e}^t_{j-1}, \hat{e}^t_j) \stackrel{+}{\prec}_{\mathsf{C}} d_{\mathcal{T}(V)}(\hat{e}_{j-1}, \hat{e}_j).$$

Proof. We fix j and proceed by induction on t, with the claim being immediate for $t = 0$. Fix $t \geqslant 1$ and suppose the claim holds for $t-1$. To ease notation, set $p = \hat{e}^{t-1}_{j-1}$ and $q = \hat{e}^{t-1}_j$. Thus by the induction hypothesis it suffices to prove

$$d_{\mathcal{T}(V)}(f_t(p), f_t(q)) \stackrel{+}{\prec}_{\mathsf{C}} d_{\mathcal{T}(V)}(p, q). \tag{9.36}$$

Let $[a, b] = \mathcal{I}_{\alpha_t}$ be the (possibly empty) active interval for the curve α_t along the geodesic segment $[p, q]$. Since the length of α_t is at least $\epsilon_0{}'$ in the complement of $\mathcal{I}_{\alpha_t}$, the map f_t is the identity on this complement. Thus it suffices to suppose $[a, b]$ is nonempty, for otherwise f_t fixes both points p, q and (9.36) is immediate. As f_t is the identity on $[p, q]\backslash[a, b]$, we have $d_{\mathcal{T}(V)}(f_t(p), f_t(a)) = d_{\mathcal{T}(V)}(p, a)$ and similarly $d_{\mathcal{T}(V)}(f_t(b), f_t(q)) = d_{\mathcal{T}(V)}(b, q)$. Since

$$d_{\mathcal{T}(V)}(p, q) = d_{\mathcal{T}(V)}(p, a) + d_{\mathcal{T}(V)}(a, b) + d_{\mathcal{T}(V)}(b, q),$$

by the triangle inequality it therefore suffices to prove that

$$d_{\mathcal{T}(V)}(f_t(a), f_t(b)) \stackrel{+}{\prec}_{\mathsf{C}} d_{\mathcal{T}(V)}(a, b). \tag{9.37}$$

Now let $A \sqsubset V$ be the annulus with $\partial A = \alpha_t$. Combining Claim 9.32 with $t-1$ applications of Claim 9.33 implies that $d_A(p, q) \stackrel{+}{\prec}_{\mathsf{C}} 0$. By Theorem 3.19 it follows that $d_A(a, b) \stackrel{+}{\prec}_{\mathsf{C}}$. Let $a|_{\alpha_t}, a|^{-}_{\alpha_t}$ and $b|_{\alpha_t}, b|^{-}_{\alpha_t}$ respectively denote the $T(\alpha_t)$– and $T(V\backslash\alpha_t)$–components of the images $\Phi_{\alpha_t}(a), \Phi_{\alpha_t}(b) \in \mathcal{P}(V|\alpha_t)$. The previous sentence implies that the horizontal coordinates of $a|_{\alpha_t}$ and $b|_{\alpha_t}$ (viewed in $T(\alpha_t) = \mathbb{H}^2$) differ by an amount bounded in terms of C. On the other hand, since $a, b \in \mathcal{I}_{\alpha_t}$ we have $\ell_a(\alpha_t), \ell_b(\alpha_t) < \epsilon_0$. It follows that the vertical coordinates of $h_{\alpha_t}(a|_{\alpha_t})$ and $h_{\alpha_t}(b|_{\alpha_t})$ both lie between $\frac{1}{\epsilon_0}$ and $\frac{1}{\epsilon_0{}'}$. We conclude that $h_{\alpha_t}(a|_{\alpha_t})$ and $h_{\alpha_t}(b|_{\alpha_t})$ have uniformly bounded (in terms of C) distance in $\mathcal{T}(\alpha_t)$. Since the metric on $\mathcal{P}(V|\alpha_t)$ is a supremum, it follows that

$$\begin{aligned}
&d_{\mathcal{P}(V|\alpha_t)}\Big(\mathrm{id} \times h_{\alpha_t}(\Phi_{\alpha_t}(a)),\ \mathrm{id} \times h_{\alpha_t}(\Phi_{\alpha_t}(b))\Big) \\
&\qquad = \sup\Big\{d_{\mathcal{T}(V\backslash\alpha_t)}\big(a|^{-}_{\alpha_t}, b|^{-}_{\alpha_t}\big),\ d_{\mathcal{T}(\alpha_t)}\big(h_{\alpha_t}(a|_{\alpha_t}), h_{\alpha_t}(b|_{\alpha_t})\big)\Big\} \\
&\qquad \stackrel{+}{\prec}_{\mathsf{C}} d_{\mathcal{T}(V\backslash\alpha_t)}\big(a|^{-}_{\alpha_t}, b|^{-}_{\alpha_t}\big) \leqslant d_{\mathcal{P}(V|\alpha_t)}\big(\Phi_{\alpha_t}(a), \Phi_{\alpha_t}(b)\big).
\end{aligned}$$

Finally, the points $a, b, f_t(a) f_t(b)$ lie in the thin region $\mathcal{H}_{\epsilon_0,\alpha_t}(V)$ where Minsky's Theorem 3.11 ensures the maps $\Phi^{\pm1}_{\alpha_t}$ change distances by at most D_0. The last quantity above thus lies within D_0 of $d_{\mathcal{T}(V)}(a, b)$ and, recalling that $f_t = \Phi^{-1}_{\alpha_t} \circ (\mathrm{id} \times h_{\alpha_t}) \circ \Phi_{\alpha_t}$, the first quantity lies within D_0 of $d_{\mathcal{T}(V)}(f_t(a), f_t(b))$. Therefore the above estimate establishes (9.37) and completes the proof of the claim. $\square$

The proposition now follows easily by invoking the triangle inequality and successively applying equation (9.34), Claim 9.35, and equation (9.30):

$$d_{\mathcal{T}(V)}(\hat{x}_{i-1}, \hat{x}_i) = d_{\mathcal{T}(V)}(\hat{e}^m_0, \hat{e}^m_k) \leqslant \sum_{j\in P} d_{\mathcal{T}(V)}(\hat{e}^m_{j-1}, \hat{e}^m_j) + \sum_{j\in Q} d_{\mathcal{T}(V)}(\hat{e}^m_{j-1}, \hat{e}^m_j)$$
$$\stackrel{+}{\prec}_{\mathsf{C}} \sum_{j\in P}^{k} d_{\mathcal{T}(V)}(\hat{e}^m_{j-1}, \hat{e}^m_j) \stackrel{+}{\prec}_{\mathsf{C}} \sum_{j\in P}^{k} d_{\mathcal{T}(V)}(\hat{e}_{j-1}, \hat{e}_j) \stackrel{+}{\prec}_{\mathsf{C}} \int_y^z \mathbb{1}_{\mathcal{A}^\Omega_V}. \qquad \square$$

10. Dealing with badness

Continue to let $\Omega = (\Omega_1, \ldots, \Omega_n)$ be a WISC witness family for a strongly C–aligned tuple $(x_0, \ldots, x_n)$ in $\mathcal{T}(\Sigma)$. We want to use Theorem 9.4 to estimate both the complexity $\mathfrak{L}(\Omega)$ (Definition 8.12) and the savings $\mathfrak{S}(\Omega)$ of Ω (Definition 8.13). To this end, we will utilize weighted characteristic functions $\mathbb{1}_{\mathcal{A}^\Omega_V}$ of contribution sets: Again the ultimate goal is Theorem 11.2. The obstacle as has been suggested are the existence of nested sets. In this section we show how to modify a witness family, if necessary, to deal with this problem.

Definition 10.1 (Weight and savings)**.** For each $V \in \Omega$, use the points $x^V_0, \ldots, x^V_n$ to define an adjustment function $\xi_V \colon [x_0, x_n] \to \mathbb{R}$ whose value is 1 on those subintervals $[x^V_{i-1}, x^V_i]$ such that $V \in \Omega_i$ is an annulus with $\widehat{x_{i-1}}^{\Omega_i}_V, \widehat{x}^{\Omega_i}_{iV}$ both thick, and whose value is zero elsewhere. Thus $\xi_V \equiv 0$ for nonannular V. On $[x^V_{i-1}, x^V_i]$ the values of $h_V - \xi_V$ and ξ_V thus respectively agree with the coefficients h^*_V and $(h_V - h^*_V)$ of the $d_{\mathcal{T}(V)}(\widehat{x_{i-1}}^{\Omega_i}_V, \widehat{x}^{\Omega_i}_{iV})$ terms appearing in the complexity $\mathfrak{L}(\Omega_i)$ (Definition 8.8) and savings $\mathfrak{S}(\Omega)$ (Definition 8.13). Accordingly the *weight* and *savings functions* $[x_0, x_n] \to \mathbb{R}$ of $V \in \Omega$ are defined as the products $\omega_V = (h_V - \xi_V)\mathbb{1}_{\mathcal{A}^\Omega_V}$ and $\sigma_V = \xi_V \mathbb{1}_{\mathcal{A}^\Omega_V}$ with the characteristic function of the contribution set $\mathcal{A}^\Omega_V$ (from Definition 9.3). Summing now yields the *total weight* and *total savings* functions $\omega_\Omega, \sigma_\Omega \colon [x_0, x_n] \to \mathbb{R}$ of Ω:

$$\omega_\Omega = \sum_{V\in\Omega} \omega_V = \sum_{V\in\Omega} (h_V - \xi_V)\mathbb{1}_{\mathcal{A}^\Omega_V} \qquad \text{and} \qquad \sigma_\Omega = \sum_{V\in\Omega} \sigma_V = \sum_{V\in\Omega} \xi_V \mathbb{1}_{\mathcal{A}^\Omega_v}$$

Theorem 9.4 says that the individual terms $(h_V - h^*_V) d_{\mathcal{T}(V)}(\widehat{x_{i-1}}^{\Omega}_V, \widehat{x}^{\Omega}_{iV})$ and $h^*_V d_{\mathcal{T}(V)}(\widehat{x_{i-1}}^{\Omega}_V, \widehat{x}^{\Omega}_{iV})$ appearing in the savings $\mathfrak{S}(\Omega)$ and complexity $\mathfrak{L}(\Omega)$ are bounded by the respective integrals $\int_{x^V_{i-1}}^{x^V_i} \sigma_V$ and $\int_{x^V_{i-1}}^{x^V_i} \omega_V$. Since the points $x^V_0, \ldots, x^V_n$ appear in order along $[x_0, x_n]$, summing over all i and V shows that $\mathfrak{L}(\Omega)$ and $\mathfrak{S}(\Omega)$ are bounded by the integrals of the total weight ω_Ω and total savings σ_Ω functions over $[x_0, x_n]$ (up to an additive error depending on n, N, and the cardinality $|\Omega|$). If we knew $\omega_\Omega(p) + \sigma_\Omega(p) \leqslant h_\Sigma$ for all p, we would thus be able to bound $\mathfrak{L}(\Omega) + \mathfrak{S}(\Omega) \stackrel{+}{\prec} h_\Sigma d_{\mathcal{T}(\Sigma)}(x_0, x_n)$. While this inequality need not hold in general, it can only fail on the following sets:

Definition 10.2 (Bad set)**.** We say a point $p \in \mathcal{A}^\Omega_V$ is *bad for* V in Ω if there exists $Z \in \Omega$ such that $Z \subsetneq V$ and $p \in \mathcal{A}^\Omega_V \cap \mathcal{A}^\Omega_Z$. The *bad set for* V is then defined to be

$$\mathcal{B}^\Omega_V = \{p \in \mathcal{A}^\Omega_V \mid p \text{ is bad for } V\} \subset \mathcal{A}^\Omega_V.$$

with corresponding *badness function* $\beta_V = h_V \mathbb{1}_{\mathcal{B}^\Omega_V}$. As for the weight and savings, the total badness function is $\beta_\Omega = \sum_{V\in\Omega} \beta_V$

Remark 10.3. We note that $\mathcal{A}_V^\Omega \subset \mathcal{I}_V$ for every domain $V \in \Omega$. This is immediate from the definition when V annular, and when V is nonannular it follows from Lemmas 9.8–9.10. Therefore any pair of domains $V, W \in \Omega$ with $\mathcal{A}_V^\Omega \cap \mathcal{A}_W^\Omega \neq \varnothing$ must either be disjoint or nested (since $V \pitchfork W$ is precluded by Lemma 3.26(4)).

The relationship between badness and weight is made precise by the next lemma.

Lemma 10.4. *For any WISC witness Ω for a strongly C–aligned tuple $(x_0, \dots, x_n)$ in $\mathcal{T}(\Sigma)$, the weight, savings, and badness functions $\omega_\Omega, \sigma_\Omega, \beta_\Omega \colon [x_0, x_n] \to \mathbb{R}$ satisfy*

$$\omega_\Omega + \sigma_\Omega - \beta_\Omega \leqslant h_\Sigma.$$

Proof. Fix $p \in [x_0, x_n]$ and consider the subcollection

$$\mathcal{G}_p = \{V \in \Omega \mid p \in \mathcal{A}_V^\Omega \text{ and } p \notin \mathcal{B}_V^\Omega\}.$$

Notice that if $V \notin \mathcal{G}_p$, then necessarily $p \in (\mathcal{A}_\Omega^V \cap \mathcal{B}_\Omega^V) \cup ([x_0, x_n] \backslash \mathcal{A}_V^\Omega)$ and thus

$$(\omega_V + \sigma_V - \beta_V)(p) = h_V \mathbb{1}_{\mathcal{A}_V^\Omega}(p) - h_V \mathbb{1}_{\mathcal{B}_V^\Omega}(p) = 0.$$

On the other hand, if $V \in \mathcal{G}_p$ then $\beta_V(p) = 0$ so that

$$(\omega_V + \sigma_V - \beta_V)(p) = (\omega_V + \sigma_V)(p) = h_V \mathbb{1}_{\mathcal{A}_V^\Omega}(p) \leqslant h_V.$$

Summing over all V, we therefore have $(\omega_\Omega + \sigma_\Omega - \beta_\Omega)(p) \leqslant \sum_{V \in \mathcal{G}_p} h_V$. That this latter quantity is at most h_Σ follows from the observation that the domains in $\mathcal{G}_p$ are disjoint subdomains of Σ. Indeed, all $V, Z \in \mathcal{G}_p$ have $p \in \mathcal{A}_V^\Omega \cap \mathcal{A}_Z^\Omega$; hence cutting $V \pitchfork Z$ is impossible by Remark 10.3, and nesting $Z \subsetneq V$ is impossible by virtue of $V \in \mathcal{G}_p$ and the definition of $\mathcal{B}_V^\Omega$. $\square$

We remark that the lemma indicates that in trying to bound $\omega_\Omega + \sigma_\Omega$ in terms of h_Σ we need to bound β_Ω. The goal of this section is to construct witness families where that term is small.

We will also need the following feature of bad sets.

Lemma 10.5. *If $p \in \mathcal{B}_V^\Omega$, then there exists an index $1 \leqslant i \leqslant n$ and domains $Z, Y \subsetneq V$ such that $Z, V \in \Omega_i$. $Y \in \Upsilon(x_{i-1}, x_i)$ with $\bar{Y}^{\Omega_i} = V$, and $p \in \mathcal{A}_Z^\Omega \cap \mathcal{I}_Y$. Furthermore, i is the unique index for which p lies in the interior of $[x_{i-1}^V, x_i^V]$.*

Proof. By definition, there is some $Z \in \Omega$ such $Z \subsetneq V$ and $p \in \mathcal{A}_Z^\Omega$. Since $Z, V \in \Omega$ with $Z \subsetneq V$, it follows from the definition that $\mathcal{I}_Z \subset M(V)$. Since the contribution set is defined as $\mathcal{A}_V^\Omega = (\mathcal{I}_V \backslash M(V)) \cup C(V)$, it must be the case that $p \in C(V)$. In particular, $p \in C_i(V)$ for some $1 \leqslant i \leqslant n$, which means that $p \in \mathcal{I}_Y$ for some $Y \subsetneq V$ satisfying $Y \in \Upsilon(x_{i-1}, x_i)$ and $\bar{Y}^{\Omega_i} = V$. From Lemmas 9.8–9.10, we see that $p \in \mathcal{I}_Y \subset J \subset [x_{i-1}^V, x_i^V]$. Since $\mathcal{I}_Z$ evidently intersects J, by Corollary 9.12 we must have $Z \notin \Upsilon(x_{j-1}, x_j)$ for all $j \neq i$. Hence the fact $Z \in \Omega$ implies $Z \in \Omega_i$. $\square$

10.1. **Fixing badness.** If $\Omega = (\Omega_1, \dots, \Omega_n)$ and $\Omega' = (\Omega'_1, \dots, \Omega'_n)$ are both WISC witness families for a tuple $(x_0, \dots, x_n)$, we will write $\Omega \subset \Omega'$ to mean that $\Omega_i \subset \Omega'_i$ for each i and that the the subordering on Ω'_i extends the subordering on Ω_i. Note that in this case, for each $V \in \Omega$ we have $M^\Omega(V) \subset M^{\Omega'}(V)$ and $C^\Omega(V) \supset C^{\Omega'}(V)$ (since in Ω' there are more domains subordered below V and thus less domains contributing to V). Therefore we observe

$$V \in \Omega \subset \Omega' \implies \mathcal{A}_V^{\Omega'} \subset \mathcal{A}_V^\Omega.$$

Recall from §8.3 that we have defined $\mathcal{E}_\Omega(V) = \max_i \mathcal{E}_{\Omega_i}(V)$. If $V \in \Omega$ satisfies $\mathcal{E}_\Omega(V) \leqslant \frac{1}{3}\mathsf{N}_V - 9\mathsf{C}$, then for each $1 \leqslant i \leqslant n$ define

$$\Omega_i^+(V) := \begin{cases} \Omega_i \cup \underline{\mathcal{L}}_{\mathcal{E}^\ell_{\Omega_i}(V)}(V) \cup \underline{\mathcal{R}}_{\mathcal{E}^r_{\Omega_i}(V)}(V), & \text{if } V \in \Omega_i \\ \Omega_i, & \text{if } V \notin \Omega_i \end{cases}.$$

That is, $\Omega_i^+(V)$ is obtained by taking the left and then right augmentations of Ω_i along V with parameters $t = \mathcal{E}^\ell_{\Omega_i}(V)$ and $t = \mathcal{E}^r_{\Omega_i}(V)$, respectively. Observe that $\Omega_i^+(V)$ is a witness family by Lemma 7.24 and that it inherits a natural subordering by Lemma 7.26. Then let $\Omega^+(V) = (\Omega_1^+(V), \ldots, \Omega_n^+(V))$ be the associated witness family, and define

$$\widehat{\Omega}(V) = \overline{\Omega^+(V)} = \left(\overline{\Omega_1^+(V)}, \ldots, \overline{\Omega_n^+(V)}\right) = \left(\widehat{\Omega}_1(V), \ldots, \widehat{\Omega}_n(V)\right)$$

to be the insular completion of $\Omega^+(V)$, equipped with its natural subordering. Notice that, since $\widehat{\Omega}(V)$ is obtained from Ω by first performing left- and right-augmentations with parameters $\mathcal{E}^r_{\Omega_i}(V)$ and $\mathcal{E}^\ell_{\Omega_i}(V)$, and then a finite sequence of refinements and augmentations with parameter 0, Lemmas 7.22 and 7.27 imply that $\widehat{\Omega}(V)$ is again WISC (since Ω was wide and $\mathcal{E}_\Omega(V) \leqslant \frac{1}{3}\mathsf{N}_V - 9\mathsf{C}$). Observe that $\Omega \subset \widehat{\Omega}(V)$ and hence that $\mathcal{A}_V^{\widehat{\Omega}(V)} \subset \mathcal{A}_V^{\Omega}$. The point of this procedure is that it moves the bad set for V entirely off of itself:

Lemma 10.6. *Suppose that $\Omega^\dagger$, Ω, and $\Omega^\ddagger$ are insular, complete, subordered witness families for the strongly C–aligned tuple $(x_0, \ldots, x_n)$ in $\mathcal{T}(\Sigma)$, that $V \in \Omega^\dagger \subset \Omega$ satisfies $\mathcal{E}_\Omega(V) \leqslant \frac{1}{3}\mathsf{N}_V - 9\mathsf{C}$, and that $\Omega^\dagger \subset \Omega \subset \widehat{\Omega}(V) \subset \Omega^\ddagger$. Then*

$$\mathcal{B}_V^{\Omega^\dagger} \cap \mathcal{B}_V^{\Omega^\ddagger} = \varnothing.$$

Proof. Suppose on the contrary that there is some $p \in \mathcal{B}_V^{\Omega^\dagger} \cap \mathcal{B}_V^{\Omega^\ddagger}$. Let $1 \leqslant i \leqslant n$ be the unique index (cf Lemma 10.5) such that p is in the interior of $[x_{i-1}^V, x_i^V]$. By Lemma 10.5, for each $* \in \{\dagger, \ddagger\}$ we have $V \in \Omega_i^*$ and may choose subdomains $Z_*, Y_* \subsetneq V$ such that $Z_* \in \Omega_i^*$, that $Y_* \in \Upsilon_i$ contributes to V in Ω_i^*, and that $p \in \mathcal{I}_{Z_*} \cap \mathcal{I}_{Y_*}$. In particular, the domains $Z_\dagger, Y_\dagger, Z_\ddagger, Y_\ddagger$ must be pairwise disjoint or nested since we have

$$p \in \mathcal{I}_{Z_\dagger} \cap \mathcal{I}_{Y_\dagger} \cap \mathcal{I}_{Z_\ddagger} \cap \mathcal{I}_{Y_\dagger}.$$

Since $Z_\dagger, V \in \Omega_i^\dagger$ it must be that $Z_\dagger$ is subordered in $\Omega_i^\dagger$ with respect to V. Let us suppose $V \searrow_i Z_\dagger$ in $\Omega_i^\dagger$ (the reverse possibility $Z_\dagger \swarrow_i V$ being symmetric). Since $\partial Y_\ddagger$ and $\partial Z_\dagger$ are disjoint and $\Omega_i^\dagger \subset \Omega_i$, we see that

$$d_V(\partial Y_\ddagger, x_i) \leqslant d_V(\partial Z_\dagger, x_i) + 1 \leqslant \mathcal{E}^r_{\Omega_i^\dagger}(V) + 1 \leqslant \mathcal{E}^r_{\Omega_i}(V) + 1.$$

Since $\Omega_i^\ddagger \supset \Omega_i$ and $Y_\ddagger \in \Upsilon_i$ contributes to V in $\Omega_i^\ddagger$, it must be that $Y_\ddagger$ also contributes to V in Ω_i. By definition of encroachment we may choose a domain $U \in \Omega_i$ with $V \searrow_i U$ and $d_V(\mathcal{C}(V|_U), x_i) = \mathcal{E}^r_{\Omega_i}(V)$. We claim that $d_V(\partial Y_\ddagger, x_i) \geqslant \mathcal{E}^r_{\Omega_i}(V) - \mathsf{M}$. Indeed, if this were not the case then evidently $d_V(\partial Y_\ddagger, \partial U) \geqslant d_V(\partial Y_\ddagger, \mathcal{C}(V|_U)) - 1 \geqslant \mathsf{M} - 1$ which implies $\partial Y_\ddagger \pitchfork \partial U$. Since $Y_\ddagger$ contributes to V in Ω_i, (SO4) implies we must have the time ordering $Y_\ddagger < U$ along $[x_{i-1}, x_i]$. On the other hand, the facts that $Y_\ddagger$ and U both have active intervals along $[x_{i-1}, x_i]$ and that $d_V(\partial U, x_i) + 1 > d_V(\partial Y_\ddagger, x_i) + \mathsf{M}$ imply that we must have the time ordering $U < Y_\ddagger$; a contradiction.

The above two paragraphs show that

$$\mathcal{E}^r_{\Omega_i}(V) - \mathsf{M} \leqslant d_V(\partial Y_\ddagger, x_i) \leqslant \mathcal{E}^r_{\Omega_i}(V) + 1.$$

Therefore $Y_\ddagger$ is necessarily contained in the set $\mathcal{L}_{\mathcal{E}^r\Omega_i}(V)$ for the segment $[x_{i-1}, x_i]$. Hence we must have $Y_\ddagger \sqsubset W_\ddagger \subsetneq V$ for some domain

$$W_\ddagger \in \underline{\mathcal{L}}_{\mathcal{E}^r\Omega_i}(V) \subset \Omega_i^+(V) \subset \widehat{\Omega}_i(V) \subset \Omega_i^\ddagger.$$

But this contradicts the fact that $Y_\ddagger$ contributes to V in $\Omega_i^\ddagger$. □

10.2. **Limited Admissibility.** As conveyed in above, in the discussion before Lemma 10.4, bounding $\mathfrak{L}(\Omega)$ by Teichmüller distance requires controlling the the badness of the witness family tuple Ω. While it may not be possible to eliminate badness entirely, we will be content to minimize it by repeatedly applying the operation $\Omega \rightsquigarrow \widehat{\Omega}(V)$ along with Lemma 10.6 to move the badness somewhere else. Throughout this process we must carefully control the cardinality of the witness families so that sum of the additive errors from Theorem 9.4 does not blow up.

Recall that we have introduced (at the start of §7) an as-yet unspecified sequence of thresholds $\mathsf{N}_{\xi(S)+1}, \dots, \mathsf{N}_{-1}$. We will shortly explain how these are chosen recursively, together with accompanying bounds and fractions,

$$\Delta_j \geqslant 1 \quad \text{and} \quad 0 < \eta_j := \frac{1}{4(j+2)^3\mathsf{C}\Delta_j} < 1 \qquad \text{for } \xi(S) \geqslant j \geqslant -1,$$

in a manner that only depends on the parameter C and the global complexity $\xi(S)$. We continue to use the notation $\eta_V = \eta_{\xi(V)}$ and $\Delta_V = \Delta_{\xi(V)}$ for a domain $V \sqsubset S$. Before specifying these constants, let us mention the role they will play.

Definition 10.7 (Admissible and Limited)**.** A witness family $\Omega = (\Omega_1, \dots, \Omega_n)$ for a strongly C–aligned tuple $(x_0, \dots, x_n)$ in $\mathcal{T}(\Sigma)$ is called:

- *admissible* if $\left|\mathcal{B}^\Omega_V\right| \leqslant \eta_V d_{\mathcal{T}(\Sigma)}(x_0, x_n)$ for all $V \in \Omega$,
- *limited* if $|\Omega|_j \leqslant \Delta_j$ for every index $\xi(S) \geqslant j \geqslant -1$, where here

$$|\Omega|_j := \max_{1\leqslant i\leqslant n} |\Omega_i|_j = \max_{1\leqslant i\leqslant n} \#\{V \in \Omega_i \mid \xi(V) = j\}$$

Adding these conditions to our previous ones, we now say a witness family is *WISCAL* if it is wide, insulated, subordered, complete, admissible, and limited, or *WISCL* when we drop the admissibility condition.

The significance of such witness families is readily apparent:

Theorem 10.8. *Let $\Omega = (\Omega_1, \dots, \Omega_n)$ be a witness family for a strongly C–aligned tuple $(x_0, \dots, x_n)$ in $\mathcal{T}(\Sigma)$. If Ω is WISCAL, then its complexity and savings satisfy*

$$\mathfrak{L}(\Omega) + \mathfrak{S}(\Omega) \stackrel{+}{\prec}_{\mathsf{C},n} \left(h_\Sigma + \frac{n}{\mathsf{C}}\right) d_{\mathcal{T}(\Sigma)}(x_0, x_n).$$

Proof. Set $\Delta = \sum_j \Delta_j$. Since Ω is limited, each family Ω_i contains at most Δ_j domains of complexity j. Thence the full union $\cup_i \Omega_i$ contains at most $n\Delta \stackrel{+}{\prec}_{\mathsf{C},n} 0$ domains. For each domain $V \in \Omega$, since the points $x_0^V, \dots, x_n^V$ appear in order along $[x_0, x_n]$, Theorem 9.4 implies that

$$\sum_{1\leqslant i\leqslant n | V\in\Omega_i} h^*_V d_{\mathcal{T}(V)}(\widehat{x_{i-1}}_V^{\Omega_i}, \widehat{x}_{iV}^{\Omega_i}) + \sum_{1\leqslant i\leqslant n|V\in\Omega_i} (h_V - h^*_V) d_{\mathcal{T}(V)}(\widehat{x_{i-1}}_V^{\Omega_i}, \widehat{x}_{iV}^{\Omega_i})$$

$$\stackrel{+}{\prec}_{\mathsf{C},n} \sum_{i=1}^{n} \int_{x_{i-1}^V}^{x_i^V} \left((h_V - \xi_V) + \xi_V\right) \mathbb{1}_{\mathcal{A}_V^\Omega} = \int_{x_0}^{x_n} \left(\omega_V + \sigma_V\right).$$

Summing these $|\Omega| \leqslant n\Delta$ inequalities over all $V \in \Omega$, combining their additive errors, and applying Lemma 10.4 now yields

$$\mathfrak{L}(\Omega) + \mathfrak{S}(\Omega) \stackrel{+}{\prec}_{\mathsf{C},n} \int_{x_0}^{x_n} \left(\omega_\Omega + \sigma_\Omega\right) \leqslant \int_{x_0}^{x_n} (h_\Sigma + \beta_\Omega) = h_\Sigma d_{\mathcal{T}(\Sigma)}(x_0, x_n) + \sum_{V\in\Omega} h_V \left|\mathcal{B}_V^\Omega\right|.$$

Using the definition $\eta_j = (4(j+2)^3\mathsf{C}\Delta_j)^{-1}$, the fact that Ω is limited and admissible, and that $h_V \leqslant 2(\xi(V) + 2)$ for all $V \sqsubset \Sigma$, we thus conclude

$$\begin{aligned}
\mathfrak{L}(\Omega) + \mathfrak{S}(\Omega) &\stackrel{+}{\prec}_{\mathsf{C},n} h_\Sigma d_{\mathcal{T}(\Sigma)}(x_0, x_n) + \sum_{V\in\Omega} h_V \eta_V d_{\mathcal{T}(\Sigma)}(x_0, x_n) \\
&\leqslant d_{\mathcal{T}(\Sigma)}(x_0, x_n) \left(h_\Sigma + n \sum_{j=-1}^{\xi(\Sigma)} \left(\frac{2(j+2)}{4(j+2)^3\mathsf{C}\Delta_j} \right) \Delta_j \right) \\
&= d_{\mathcal{T}(\Sigma)}(x_0, x_n)) \left(h_\Sigma + \frac{n}{2\mathsf{C}} \sum_{j=-1}^{\xi(\Sigma)} \frac{1}{(j+2)^2} \right) \\
&= d_{\mathcal{T}(\Sigma)}(x_0, x_n) \left(h_\Sigma + \frac{n\pi^2}{12\mathsf{C}} \right) \leqslant d_{\mathcal{T}(\Sigma)}(x, y) \left(h_\Sigma + \frac{n}{\mathsf{C}} \right) \qquad \square
\end{aligned}$$

10.3. **Saturation.** It remains to prove that WISCAL witness families exist and, in the process, to specify all of the constants N_j, η_j, Δ_j. Our witness families will be constructed in the following iterative manner.

To begin with, suppose merely that our constants N_j and η_j have been specified arbitrarily subject to the conditions

$$(10.9) \quad \xi(S) + 30\mathsf{C}\frac{\epsilon_0}{\epsilon_0{}'} = \mathsf{N}_{\xi(S)+1} \leqslant \cdots \leqslant \mathsf{N}_{-1} = \mathsf{N} \quad \text{and} \quad 0 < \eta_j < 1 \quad \text{for all } j.$$

We continue to use the notation $\mathsf{N}_V = \mathsf{N}_{\xi(V)}$ and $\eta_V = \eta_{\xi(V)}$ for any domain $V \sqsubset S$.

Let $\Sigma \sqsubset S$ be any domain and let $(x_0, \ldots, x_n)$ be a strongly C–aligned tuple in $\mathcal{T}(\Sigma)$. For each $1 \leqslant i \leqslant n$, let Ω_i^0 be the set of topologically maximal domains in the collection

$$\Upsilon(x_{i-1}, x_i) = \Upsilon^c(x_{i-1}, x_i) \cup \Upsilon^\ell(x_{i-1}, x_i),$$

where we recall that $\Upsilon, \Upsilon^c, \Upsilon^\ell$ are the sets from Definition 7.2. Then Ω_i^0 is a witness family for $[x_{i-1}, x_i]$ by definition. Since $\Upsilon^\ell(x_{i-1}, x_i)$ consists of at most $2\xi(\Sigma)$ annuli, we see that Ω_i^0 consists of the topologically maximal domains in $\Upsilon^c(x_{i-1}, x_i)$ together with a subset of $\Upsilon^\ell(x_{i-1}, x_i)$. Thus the number of domains in Ω_i^0 of each complexity is uniformly bounded as described by Lemma 4.1. Since there are no nested domains in Ω_i^0, it is trivially subordered. Now let $\Omega^0 = (\Omega_1^0, \ldots, \Omega_n^0)$ be the associated subordered witness family for the tuple $(x_0, \ldots, x_n)$, and let $\Omega^1 = \overline{\Omega^0} = (\overline{\Omega_1^0}, \ldots, \overline{\Omega_n^0})$ be its insular completion (Definitions 7.28 and 8.10). Note that by Lemma 7.29 we have $\mathcal{E}_{\Omega^1}(V) \leqslant 9\mathsf{C}$ for all $V \in \Omega^1$.

For $k \in \mathbb{N}$, suppose that we have constructed an increasing chain $\Omega^1 \subset \cdots \subset \Omega^k$ of WISC witness families for $(x_0, \ldots, x_n)$. Among all domains $V \in \Omega^k$ satisfying

$$\mathcal{E}_{\Omega^k}(V) \leqslant \tfrac{1}{3}\mathsf{N}_V - 9\mathsf{C} \quad \text{and} \quad \left|\mathcal{B}_V^{\Omega^k}\right| > \eta_V d_{\mathcal{T}(\Sigma)}(x_0, x_n)$$

(that is, the Lebesgue measure of $\mathcal{B}_V^{\Omega^k} \subset [x_0, x_n]$ is more than η_V–percent of the total measure of $[x_0, x_n]$), choose one of maximal complexity and call it V_k. Using the operation $\Omega \rightsquigarrow \widehat{\Omega}(V) = \overline{\Omega^+(V)}$ from §10.1, we then define

$$\Omega^{k+1} = \widehat{\Omega^k}(V_k) = \left(\widehat{\Omega^k_1}(V_k), \ldots, \widehat{\Omega^k_n}(V_k)\right).$$

In this way, we obtain a list $V_1, V_2, \ldots$ of domains and a chain $\Omega^1 \subset \Omega^2 \subset \cdots$ of WISC witness families. In fact, this process must terminate in finitely many steps yielding a WISC witness family $\Omega := \cup_k \Omega^k$ with the property that every domain $V \in \Omega$ satisfies $\mathcal{E}_\Omega(V) > \frac{1}{3}\mathsf{N}_V - 9\mathsf{C}$ or $\left|\mathcal{B}_V^\Omega\right| < \eta_V d_{\mathcal{T}(\Sigma)}(x_0, x_n)$. Indeed, since there are only finitely many domains in each collection $\Upsilon(x_{i-1}, x_i)$, there are only finitely many possible witness families for $(x_0, \ldots, x_n)$. Since the bad sets $\mathcal{B}_{V_k}^{\Omega^k}$ and $\mathcal{B}_{V_k}^{\Omega^{k+1}}$ are disjoint by Lemma 10.6 (and $\mathcal{B}_{V_k}^{\Omega^k}$ is nonempty by choice of V_k), we see that $\Omega^k \subsetneq \Omega^{k+1}$ for each k. Therefore the families Ω^k are all distinct, showing that the process terminates in finitely many steps.

Definition 10.10. We refer to any family Ω obtained in this way as a *saturated* witness family for $(x_0, \ldots, x_n)$.

By choosing the constants N_j and η_j carefully, we will be able to bound the number of domains in Ω of each complexity and to moreover arrange that every domain $V \in \Omega$ satisfies $\mathcal{E}_\Omega(V) \leqslant \frac{1}{3}\mathsf{N}_V$ and $\left|\mathcal{B}_V^\Omega\right| \leqslant \eta_V d_{\mathcal{T}(\Sigma)}(x_0, x_n)$.

To this end, we first observe that a particular domain Z can appear in the list $V_1, V_2, \ldots$ at most $1/\eta_Z$ times. This is because if $k_1, \ldots, k_\ell$ are distinct indices with $Z = V_{k_1} = \cdots = V_{k_\ell}$, then Lemma 10.6 implies the bad sets $\mathcal{B}_Z^{\Omega^{k_1}}, \ldots, \mathcal{B}_Z^{\Omega^{k_\ell}}$ are all disjoint. Whence

$$d_{\mathcal{T}(\Sigma)}(x_0, x_n) \geqslant \left|\mathcal{B}_Z^{\Omega^{k_1}}\right| + \cdots + \left|\mathcal{B}_Z^{\Omega^{k_\ell}}\right| > \ell\eta_Z d_{\mathcal{T}(\Sigma)}(x_0, x_n)$$

and we have $\ell\eta_Z < 1$ as claimed. This has the following consequence:

Lemma 10.11. *If Ω is a saturated witness family, then each domain $Z \in \Omega$ satisfies $\mathcal{E}_\Omega(Z) \leqslant 9\mathsf{C}\left(1 + \frac{1}{\eta_V}\right)$.*

Proof. Suppose the saturated family is constructed as $\Omega = \cup_k \Omega^k$ for the increasing chain $\Omega^0 \subset \Omega^1 \cdots$ where each Ω^0_i is the set of topologically maximal domains in $\Upsilon(x_{i-1}, x_i)$, where $\Omega^1 = \overline{\Omega^0}$, and $\Omega^{k+1} = \widehat{\Omega}^k(V_k)$ for some domain $V_k \in \Omega_k$. Since there are no nested domains in any collection Ω^0_i for $1 \leqslant i \leqslant n$, we trivially have $\mathcal{E}_{\Omega^0}(W) = 0$ for every domain W. By Lemma 7.29, its insular completion Ω^1 satisfies $\mathcal{E}_{\Omega_1}(W) \leqslant 9\mathsf{C}$ for all $W \sqsubset \Sigma$.

For each augmentation $\Omega^k \rightsquigarrow \Omega^{k+1} = \overline{\Omega^{k+}(V_k)}$, Lemmas 7.27 and 7.29 together imply that the encroachment for any $Z \sqsubset \Sigma$ satisfies

$$\mathcal{E}_{\Omega^{k+1}}(Z) \leqslant \begin{cases} \mathcal{E}_{\Omega^k}(Z) + 9\mathsf{C}, & Z = V_k \\ \max\{\mathcal{E}_{\Omega^k}(Z), 9\mathsf{C}\}, & \text{else.} \end{cases}$$

Therefore, we conclude that $\mathcal{E}_\Omega(Z) \leqslant (m+1)9\mathsf{C}$, where $m = |\{k \in \mathbb{N} \mid Z = V_k\}|$ is the number of indices k for which $Z = V_k$. But we have observed that m is at most $1/\eta_V$. This proves the lemma. □

Next observe that if Ω is complete and insulated, then every domain added during an operation $\Omega \rightsquigarrow \widehat{\Omega}(V)$ is a proper subdomain of V. Indeed, fix some $1 \leqslant i \leqslant n$ and suppose $Z \in \widehat{\Omega}_i(V) \backslash \Omega_i$. If $Z \in \Omega_i^+(V) \backslash \Omega_i$, then by definition $Z \in \underline{\mathcal{L}}_{\mathcal{E}^\ell_{\Omega_i}(V)}(V) \cup \underline{\mathcal{R}}_{\mathcal{E}^r_{\Omega_i}(V)}(V)$ showing that Z is a proper subsurface of V by definition. Since Ω_i is already a complete and insular by assumption, it is clear from the construction (Definition 7.28) of the insular completion $\widehat{\Omega}_i(V) = \overline{\Omega_i^+(V)}$ that every $Z \in \overline{\Omega_i^+(V)} \backslash \Omega_i^+(V)$ satisfies $Z \subsetneq W$ for some $W \in \Omega_i^+(V) \backslash \Omega_i$. Therefore every $Z \in \Omega_{i+1} \backslash \Omega_i$ is a proper subsurface of V, as claimed.

Let us next analyze how the cardinalities of a family change under an operation $\Omega \rightsquigarrow \widehat{\Omega}(V)$. The previous paragraph shows that for each $\xi(\Sigma) \geqslant j \geqslant \xi(V)$ and $1 \leqslant i \leqslant n$, the number $|\Omega_i|_j$ of domains of complexity j stays constant. Hence:

$$\big|\widehat{\Omega}_i(V)\big|_j = |\Omega_i|_j \quad \text{for } \xi(\Sigma) \geqslant j \geqslant \xi(V), \quad 1 \leqslant i \leqslant n.$$

However, surfaces of lower complexity may be added during the augmentation step $\Omega \rightsquigarrow \Omega^+(V)$, and then during the completion step $\Omega^+(V) \rightsquigarrow \widehat{\Omega}(V)$: For each $-1 \leqslant j < \xi(V)$, Lemma 7.10 shows that

$$\big|\Omega_i^+(V)\big|_j \leqslant |\Omega_i|_j + 2(2\mathsf{N}_{j+1})^{\xi(\Sigma)+3} \qquad \text{for } -1 \leqslant j < \xi(V), \quad 1 \leqslant i \leqslant n.$$

Lemma 7.30 therefore implies that

$$\big|\widehat{\Omega}_i(V)\big|_j \leqslant |\Omega_i|_j + 2(2\mathsf{N}_{j+1})^{\xi(\Sigma)+3} + G_j(|\Omega_i|_{\xi(\Sigma)}, \ldots, |\Omega_i|_{j+1}),$$

where G_j is a function depending only on the thresholds $\mathsf{N}_{\xi(\Sigma)}, \ldots, \mathsf{N}_{j+1}$. To summarize, since $|\Omega|_j$ is defined as the maximum $\max_i |\Omega_i|_j$, for any operation $\Omega \rightsquigarrow \widehat{\Omega}(V)$ and complexity j, we have

$$\big|\widehat{\Omega}(V)\big|_j - |\Omega|_j \leqslant \begin{cases} C'_j + G_j(|\Omega|_{\xi(\Sigma)}, \ldots, |\Omega|_{j+1}), & -1 \leqslant j < \xi(V_i) \\ 0, & \xi(V_i) \leqslant j \leqslant \xi(\Sigma) \end{cases}, \tag{10.12}$$

where the number C'_j and function G_j depend only on the thresholds $\mathsf{N}_{\xi(\Sigma)}, \ldots, \mathsf{N}_{j+1}$.

With this, we can now specify our constants N_j, η_j, and Δ_j recursively:

Proposition 10.13 (Choosing the constants)**.** *For any $\mathsf{C} \geqslant \mathsf{M}$ and $n \geqslant 1$, there are constants N_j, Δ_j and η_j for $\xi(S) \geqslant j \geqslant -1$ satisfying (10.9), such that the following holds: For any domain $\Sigma \sqsubset S$ and strongly C–aligned tuple $(x_0, \ldots, x_n)$ in $\mathcal{T}(\Sigma)$, every saturated witness family Ω for $(x_0, \ldots, x_n)$ is WISCAL.*

Proof. By construction, every saturated family is WISC, but we must take care to ensure Ω is admissible and limited. Fix some complexity $j \leqslant \xi(S)$. Let us say a choice of constants $\Delta_{\xi(S)}, \ldots, \Delta_j$, fractions $\eta_{\xi(S)}, \ldots, \eta_j$, and thresholds $\mathsf{N}_{\xi(S)}, \ldots, \mathsf{N}_j$ is robust if, **irrespective of how the remaining fractions $\eta_{j-1}, \ldots, \eta_{-1}$ and thresholds $\mathsf{N}_{j-1}, \ldots, \mathsf{N}_{-1}$ are specified**, subject to equation (10.9), every saturated resolution family Ω as in the proposition statement satisfies

$$|\Omega|_m \leqslant \Delta_m \quad \text{and} \quad \mathcal{E}_\Omega(V) \leqslant \frac{1}{3}\mathsf{N}_V - 9\mathsf{C} \quad \text{ for all } \xi(V), m \geqslant j;$$

that is, if limited admissibility holds for all complexities at least j.

To begin the recursion, let us set

$$\Delta_{\xi(S)} = 1, \quad \eta_{\xi(S)} = \frac{1}{4(\xi(S)+2)^3\mathsf{C}\Delta_{\xi(S)}}, \quad \text{and} \quad \mathsf{N}_{\xi(S)} = 30\mathsf{C}\frac{\epsilon_0}{\epsilon_0{}'}\left(2 + \frac{1}{\eta_{\xi(S)}}\right).$$

Note this ensures $\mathsf{N}_{\xi(S)} \geqslant \xi(S) + 30\mathsf{C}\frac{\epsilon_0}{\epsilon_0{}'}$. We claim this choice is robust. Indeed, let Ω be any saturated witness family. Since there is only one domain of complexity $\xi(S)$, namely S itself, we trivially have $|\Omega|_{\xi(S)} \leqslant 1 = \Delta_{\xi(S)}$. Further, Lemma 10.11 and our choice of $\mathsf{N}_{\xi(S)}$ ensure that $\mathcal{E}_\Omega(S) \leqslant 9\mathsf{C}(1+\frac{1}{\eta_S}) < \frac{1}{3}\mathsf{N}_S - 9\mathsf{C}$.

By induction, fix some complexity $j < \xi(S)$ and suppose that we have already designated robust constants $\Delta_{\xi(S)}, \dots, \Delta_{j+1}$, fractions $\eta_{\xi(S)}, \dots, \eta_{j+1}$, and thresholds $\mathsf{N}_{\xi(S)}, \dots, \mathsf{N}_{j+1}$. Consider any saturated resolution family $\Omega = \cup_k \Omega^k$ for a tuple $(x_0, \dots, x_n)$ in $\mathcal{T}(\Sigma)$, where $\Omega^0 = (\Omega^0_1, \dots, \Omega^0_n)$ with each Ω^0_i equal to the set of maximal domains in $\Upsilon(x_{i-1}, x_i)$, where $\Omega^1 = \overline{\Omega^0}$, and $\Omega^{k+1} = \widehat{\Omega}^k(V_k)$ for each $k \geqslant 1$. Let us consider how many domains of complexity j can arise in Ω: By Lemma 4.1, there are constants $C_{\xi(\Sigma)}, \dots, C_j$ depending only on $\mathsf{N}_{\xi(\Sigma)}, \dots, \mathsf{N}_{j+1}$ such that the original family Ω^0 satisfies $\left|\Omega^0\right|_m \leqslant C_m$ for all $j \leqslant m \leqslant \xi(\Sigma)$. Therefore Lemma 7.30 implies that

$$\left|\Omega^1\right|_j = \left|\overline{\Omega^0}\right|_j \leqslant P_j,$$

where P_j is a constant depending only on $C_{\xi(\Sigma)}, \dots, C_j, \mathsf{N}_{\xi(\Sigma)}, \dots, \mathsf{N}_{j+1}$, and thus ultimately depending only on $\mathsf{N}_{\xi(S)}, \dots, \mathsf{N}_{j+1}$. Now, we have seen above that domains of complexity j are only added when we augment along a domain of complexity strictly larger than j. For each $m > j$ and k, our induction hypothesis ensures that $\left|\Omega^k\right|_m \leqslant |\Omega|_m \leqslant \Delta_m$. Thus there are at most $n\Delta_m$ domains of complexity m that are candidates for augmentation, and each such domain can occur on the list $V_1, V_2, \dots$ at most $1/\eta_m$ times.

Therefore, there are at most $\sum_{m=j+1}^{\xi(S)} n\Delta_m/\eta_m$ indices k such that

$$\left|\Omega^{k+1}\right|_j > \left|\Omega^k\right|_j.$$

For each such index k, our hypothesis $\left|\Omega^k\right|_m \leqslant |\Omega|_m \leqslant \Delta_m$ and equation (10.12) together imply the difference $\left|\Omega^{k+1}\right|_j - \left|\Omega^k\right|_j$ is bounded by a number Q_j depending only the thresholds $\mathsf{N}_{\xi(S)}, \dots, \mathsf{N}_{j+1}$ and constants $\Delta_{\xi(S)}, \dots, \Delta_{j+1}$. This proves $|\Omega|_j \leqslant \Delta_j$, where

$$\Delta_j := P_j + Q_j\left(n\frac{\Delta_{\xi(S)}}{\eta_{\xi(S)}} + \cdots + n\frac{\Delta_{j+1}}{\eta_{j+1}}\right)$$

is a constant depending only on n, our previously determined constants Δ_m, η_m, and N_m for $j < m \leqslant \xi(S)$. Now that we know $|\Omega|_j \leqslant \Delta_j$, we set

$$\eta_j := \frac{1}{4(j+2)^3\mathsf{C}\Delta_j} \qquad \text{and} \qquad \mathsf{N}_j := \max\left\{30\mathsf{C}\left(2+\frac{1}{\eta_j}\right), \mathsf{N}_{j+1}\right\}. \tag{10.14}$$

By Lemma 10.11, this d choice of N_j ensures that every domain $V \in \Omega$ with $\xi(V) = j$ satisfies $\mathcal{E}_\Omega(V) \leqslant 9\mathsf{C}(1+\frac{1}{\eta_j}) < \frac{1}{3}\mathsf{N}_V - 9\mathsf{C}$. Thus the constants $\Delta_{\xi(S)}, \dots, \Delta_j$, fractions $\eta_{\xi(S)}, \dots, \eta_j$ and thresholds $\mathsf{N}_{\xi(S)}, \dots, \mathsf{N}_j$ form a robust choice.

Proceeding recursively in this manner, we obtain a complete list of robust constants Δ_m, η_m, and N_m for $\xi(S) \geqslant m \geqslant -1$. Since these constants are robust, any saturated family Ω built using these constants is necessarily limited. Furthermore, since each $V \in \Omega$ satisfies $\mathcal{E}_\Omega(V) < \frac{1}{3}\mathsf{N}_V - 9\mathsf{C}$, the fact that Ω is saturated automatically implies that $\left|\mathcal{B}_V^\Omega\right| < \eta_V d_{\mathcal{T}(\Sigma)}(x_0, x_n)$. Hence Ω is also admissible. $\square$

11. Complexity Length

We are finally ready to define the key quantity coming from this lengthy construction, namely the complexity length of a tuple.

Definition 11.1. Let $\Sigma \sqsubset S$ be a domain. The *complexity length* and *savings* of a strongly C–aligned tuple $(x_0, \ldots, x_n)$ in $\mathcal{T}(\Sigma)$ are defined to be

$$\mathfrak{L}(x_0, \ldots, x_n) = \inf_\Omega \mathfrak{L}(\Omega) \qquad \text{and} \qquad \mathfrak{S}(x_0, \ldots, x_n) = \inf_\Omega \mathfrak{S}(\Omega),$$

where the infima are taken over all WISCL witness families Ω for the tuple, and where $\mathfrak{L}(\Omega)$ and $\mathfrak{S}(\Omega)$ are as given in Definitions 8.12–8.13. Note that the infima are achieved, since the set of such Ω is nonempty (e.g. by Proposition 10.13) and finite by virtue of the sets $\Upsilon(x_{i-1}, x_i)$ being finite.

We now have the following consequences of the construction:

Theorem 11.2. *Let $\Sigma \sqsubset S$ be a domain and $(x_0, \ldots, x_n)$ be a strongly C–aligned tuple in $\mathcal{T}(\Sigma)$. Then for any indices $0 = k_0 < k_1 < \cdots < k_m = n$ we have both*

$$\sum_{j=1}^m \mathfrak{L}(x_{k_{j-1}}, \ldots, x_{k_j}) \;\leqslant\; \mathfrak{L}(x_0, \ldots, x_n), \quad \textit{and}$$

$$\mathfrak{L}(x_0, \ldots, x_n) + \mathfrak{S}(x_0, \ldots, x_n) \;\stackrel{+}{\prec}_{C,n}\; \left(h_\Sigma + \frac{n}{\mathsf{C}}\right) d_{\mathcal{T}(\Sigma)}(x_0, x_n).$$

As a special case of the theorem, if (y, x, z) is a strongly C–aligned tuple in $\mathcal{T}(S)$, then $\mathfrak{L}(y, x) + \mathfrak{L}(x, z) \stackrel{+}{\prec}_{\mathsf{C}} (h_S + \frac{2}{\mathsf{C}}) d_{\mathcal{T}(S)}(y, z)$.

Proof. Let Ω' be a saturated witness family for $(x_0, \ldots, x_n)$. By Proposition 10.13, Ω' is WISCAL and hence satisfies $\mathfrak{L}(\Omega') + \mathfrak{S}(\Omega') \stackrel{+}{\prec}_{\mathsf{C},n} (h_\Sigma + \frac{n}{\mathsf{C}})\mathrm{d}_{\mathcal{T}(\Sigma)}(x_0, x_n)$ by Theorem 10.8. Next let Ω and Ω'' realize the infima from Definition 11.1, so that $\mathfrak{L}(x_0, \ldots, x_n) = \mathfrak{L}(\Omega)$ and $\mathfrak{S}(x_0, \ldots, x_n) = \mathfrak{S}(\Omega'')$. Since Ω' is a candidate for these infima, we trivially have $\mathfrak{L}(\Omega) + \mathfrak{S}(\Omega'') \leqslant \mathfrak{L}(\Omega') + \mathfrak{S}(\Omega')$. Combined with the previous observations, this proves the second assertion of theorem.

Next, note that by definition each subfamily $\Omega^j = (\Omega_{1+k_{j-1}}, \ldots, \Omega_{k_j})$ is a WISCL (but not necessarily admissible) witness family for the strongly C–aligned subtuple $(x_{k_{j-1}}, \ldots, x_{k_j})$. Since complexity length is an infimum, it follows that $\mathfrak{L}(x_{k_{j-1}}, \ldots, x_{k_j}) \leqslant \mathfrak{L}(\Omega^j)$. On the other hand, the complexity of a tuple witness family is exactly defined so that

$$\sum_{j=1}^m \mathfrak{L}(x_{k_{j-1}}, \ldots, x_{k_j}) \leqslant \sum_{j=1}^m \mathfrak{L}(\Omega^j) = \sum_{j=1}^m \left(\sum_{i=1+k_{j-1}}^{k_j} \mathfrak{L}(\Omega_i) \right) = \sum_{i=1}^n \mathfrak{L}(\Omega_i) = \mathfrak{L}(\Omega). \;\square$$

We note that the δ in the main theorem will come from this theorem. Namely given δ we will pick C large enough so that $\frac{n}{\mathsf{C}} < \delta$. This observation will be repeated in the last section.

12. Counting with complexity length

In §§7–11 we have gone to great lengths to define a quantity $\mathfrak{L}(x,y)$ that is essentially bounded above by $h_S d_{\mathcal{T}(S)}(x,y)$. We now count points with given complexity length. Recall from Definition 3.14 that for each domain Σ of S we have specified a $(\mathsf{c}, 2\mathsf{c})$ net $\mathcal{N}(\Sigma)$ in the Teichmüller space $\mathcal{T}(\Sigma)$. Our goal in this section is:

Theorem 12.1. *For any parameter $\mathsf{C} \geqslant \mathsf{M}$ there exists an integer $k \geqslant 1$ such that for any domain $\Sigma \sqsubset S$, point $x \in \mathcal{T}(\Sigma)$, and distance $r > 0$ we have*

$$\#\Big\{y \in \mathcal{N}(\Sigma) \mid \mathfrak{L}(x,y) \leqslant r\Big\} \leqslant kr^k e^r.$$

That is, there are at most kr^ke^r net points within complexity length r of x.

This should be compared with Lemma 3.15 (itself a consequence of Theorem 3.12 by [ABEM]), but with the Teichmüller distance replaced by complexity length.

Corollary 12.2. *There exists an integer $k \geqslant 1$ depending only on C, n such that for any domain $\Sigma \sqsubset S$, point $x \in \mathcal{T}(\Sigma)$, and distance $r > 0$, there are at most kr^ke^r tuples $(x_1, \dots, x_n)$ of net points such that $\mathfrak{L}(x, x_1, \dots, x_n) \leqslant r$.*

Proof. Since $\sum_{i=1}^n \mathfrak{L}(x_{i-1}, x_i) \leqslant \mathfrak{L}(x, x_1, \dots, x_n) \leqslant r$ by Theorem 11.2, for each integer partition $r \geqslant r_1 + \dots + r_n$ we count the number of strongly C–aligned tuples $(x_0, \dots, x_n)$ with $x_0 = x$ and $\mathfrak{L}(x_{i-1}, x_i) \leqslant r_i$. By Theorem 12.1, once x_{i-1} is determined there are at most $k(r_i)^k e^{r_i}$ options for the next net point x_i. Thus in total there are at most $k^n(r_1 \dots r_n)^k e^{r_1+\dots+r_n} \leqslant k^n r^{kn} e^r$ options for each of the at most r^n such partitions of r. $\square$

12.1. **Directed graphs.** The proof will require a bit of setup. Given a WISCL witness family Ω for a pair (x,y), we define a labeled directed graph $\mathcal{G} = \mathcal{G}(\Omega)$ as follows: The vertex set $\mathcal{V} = \mathcal{V}(\mathcal{G})$ is the set of domains in Ω with each vertex $Z \in \Omega$ labeled by the ordered pair $(h_Z^*, \lfloor d_{\mathcal{T}(Z)}(\widehat{x}_Z^\Omega, \widehat{y}_Z^\Omega)\rfloor)$, where the first entry h_Z^* is the weight used in calculating the complexity $\mathfrak{L}(\Omega)$ (Definition 8.8) and the second entry is the integer part of the Teichmüller distance between the resolution points $\widehat{x}_Z^\Omega, \widehat{y}_Z^\Omega$. Vertices $Y, Z \in \mathcal{V}$ are joined by directed labeled edge from Y to Z as follows:

- if $Y \pitchfork Z$ with $Y \lessdot Z$ along $[x,y]$, we have an edge $Y \overset{P}{\to} Z$ labeled "P;"
- if $Y \sqsubset Z$ with $Y \swarrow Z$, there is an edge $Y \overset{SW}{\to} Z$ labeled "SW;"
- if $Y \sqsupset Z$ with $Y \searrow Z$, there is an edge $Y \overset{SE}{\to} Z$ labeled "SE;"
- if Y and Z are disjoint (that is, $Y \perp Z$), there is no edge joining Y and Z.

Lemma 12.3. *These directed edges give a partial ordering on the vertices of $\mathcal{G}$.*

Proof. We need to prove transitivity. First consider a concatenation $W \overset{P}{\to} Y \overset{P}{\to} Z$. Then we have $W \pitchfork Z$ by Corollary 3.30 with $W \lessdot Z$ along $[x,y]$ by transitivity of time-order; hence $W \overset{P}{\to} Z$ as required. If the second edge is labeled $Y \overset{SE}{\to} Z$, then we must have $W \pitchfork Z$ by (SO3) and moreover $W \lessdot Z$ by Corollary 3.31. Finally, if the second edge is labeled $Y \overset{SW}{\to} Z$, then $Y \sqsubset Z$ and we cannot have $W \perp Z$ or $W \sqsupset Z$. If $W \pitchfork Z$, then necessarily $W \lessdot Z$ by Corollary 3.31 so that $W \overset{P}{\to} Z$ as needed. If instead $W \sqsubset Z$, then we must have $W \swarrow Z$ by (SO2), so that $W \overset{SW}{\to} Z$.

The cases $W \overset{SW}{\to} Y \overset{P}{\to} Z$ and $W \overset{SE}{\to} Y \overset{P}{\to} Z$ are handled by symmetric arguments as above. For the case $W \overset{SW}{\to} Y \overset{SE}{\to} Z$, axiom (SO2) exactly gives $W \overset{P}{\to} Z$. Similarly $W \overset{SW}{\to} Y \overset{SW}{\to} Z$, and the alternative with two SE edges, follows from (SO1).

For the last remaining case $W \overset{SE}{\to} Y \overset{SW}{\to} Z$, the domains W and Z cannot be disjoint because they both contain Y. If $W \pitchfork Z$ then we necessarily have $W \lessdot Z$ and thus $W \overset{P}{\to} Z$ by (SO3). Similarly if $W \sqsubset Z$, then (SO1) ensures $W \overset{SW}{\to} Z$ and, symmetrically, $W \sqsupset Z$ leads to $W \overset{SE}{\to} Z$. ☐

The labeled directed graph $\mathcal{G}$ is a combinatorial object that neither remembers the points x, y nor the family Ω giving rise to it. To emphasize this combinatorial structure, we will use lowercase letters $v \in \mathcal{V}$ to denote vertices of $\mathcal{G}$ and write $(h_v^*, d_v) \in \mathbb{N}^2$ for the label of the vertex. Our goal is, essentially, to count the number of witness families that give rise to a given labeled graph $\mathcal{G}$.

12.2. **Realizing initial subsets.** To this end, we say a subset $\mathcal{X} \subset \mathcal{V}$ *respects the partial order* if there is no directed edge from the complement $\mathcal{V} \backslash \mathcal{X}$ to $\mathcal{X}$, that is, if $Y \in \mathcal{X}$ implies $W \in \mathcal{X}$ for any directed edge $W \to Y$. For example, the subsets $\varnothing$ and $\mathcal{V}$ both respect the partial order.

Given our combinatorial graph $\mathcal{G}$, a subset $\mathcal{X}$ respecting the partial order, and an initial point $x \in \mathcal{T}(\Sigma)$, we say a pair of families Ω_1, Ω_2 are *equivalent over* $\mathcal{X}$ if

- each Ω_i is a WISCL witness family for a segment $[x, y_i]$ starting at x,
- the graph $\mathcal{G}(\Omega_i)$ associated to each family Ω_i is isomorphic to $\mathcal{G}$ via an isomorphism $f_i \colon \mathcal{G} \to \mathcal{G}(\Omega_i)$ of labeled directed graphs such that
- for each vertex $v \in \mathcal{X}$, the corresponding domains $f_i(v) \in \mathcal{V}(\mathcal{G}(\Omega_i)) = \Omega_i$ are equal, call it $f_1(v) = Z = f_2(v)$, and have the same resolution points, meaning that $\widehat{x}_Z^{\Omega_1} = \widehat{x}_Z^{\Omega_2}$ and $\widehat{y_1}_Z^{\Omega_1} = \widehat{y_2}_Z^{\Omega_2}$ in $\mathcal{T}(Z)$.

Definition 12.4. A *realization* of $\mathcal{X}$ relative to an initial point $x \in \mathcal{T}(\Sigma)$ is an equivalence class $\mathcal{R}$ of families over $\mathcal{X}$. We additionally say a segment $[x, y]$ *realizes* $\mathcal{R}$ if the equivalence class contains a witness family Ω for $[x, y]$. A realization of $\varnothing$ thus consists of no data, whereas a realization of $\mathcal{V}$ roughly consists of a witness family Ω giving rise to $\mathcal{G}$.

In general, a realization $\mathcal{R}$ of $\mathcal{X}$ determines for each vertex $v \in \mathcal{X}$ a domain $Z_v \sqsubset \Sigma$ and a pair of points $\hat{x}_v, \hat{y}_v \in \mathcal{T}(Z_v)$. These domains moreover satisfy the combinatorial conditions that if $v \overset{P}{\to} w$ then $Z_v \pitchfork Z_w$, if $v \overset{SW}{\to} w$ then $Z_v \sqsubset Z_w$, if $v \overset{SE}{\to} w$ then $Z_v \sqsupset Z_w$, and that Z_v, Z_w are disjoint if there is no edge joining v and w. Let us write $\Omega(\mathcal{R}) = \{Z_v \mid v \in \mathcal{X}\}$ for this set of domains.

Mimicking the notation from §7.1, let us say that $v \in \mathcal{X}$ *minimally contains* a domain U, denoted $U \lessdot^{\mathcal{X}} v$, if Z_v is a topologically minimal element of the set $\{Z_w \mid w \in \mathcal{X} \text{ and } U \sqsubset Z_w\}$.

Lemma 12.5. *If distinct vertices $v, w \in \mathcal{X}$ minimally contain $U \sqsubset \Sigma$, then for every segment $[x, y]$ realizing $\mathcal{R}$ we have $U \notin \Upsilon(x, y)$ and, in particular $d_U(x, y) \leqslant \mathsf{N}$.*

Proof. Let Ω be any witness family in $\mathcal{R}$ for the segment $[x, y]$. Then $Z_v, Z_w \in \Omega(\mathcal{R})$ both minimally contain U. By contradiction, let us suppose $U \in \Upsilon(x, y)$. The completeness of Ω then provides an Ω–supremum $U' = \bar{U}^{\Omega}$ which, by Lemma 7.6, satisfies $U' \sqsubset Z_v$ and $U' \sqsubset Z_w$. Observe that the domains Z_v, Z_w cannot be nested, as they both minimally contain U. Thus $Z_v \pitchfork Z_w$ and we may suppose, without loss of generality, that they are time-ordered $Z_v \lessdot Z_w$ along $[x, y]$. The supremum U' is an element of Ω nested inside Z_v and Z_w, hence it must be subordered with respect to them. In fact the subordering must be $Z_v \searrow U' \swarrow Z_w$, since the alternatives

$U' \swarrow Z_v \lessdot Z_w$ and $Z_v \lessdot Z_w \searrow U'$ are precluded by (SO2). This means that the directed graph $\mathcal{G}(\Omega)$ has edges $Z_v \overset{SE}{\to} U' \overset{SW}{\to} Z_w$. Since Z_w lies in $\Omega(\mathcal{R})$ and $\mathcal{X}$ respects the partial order, it must be the case that $U' \in \Omega(\mathcal{R})$ as well. But this contradicts the assumption that v and w both minimally containing U in $\mathcal{X}$. □

Corollary 12.6. *For any domain $U \sqsubset \Sigma$, the set $\{\pi_U(\hat{y}_v) \mid U \lessdot^{\mathcal{X}} v\}$ has uniformly bounded diameter in $\mathcal{C}(U)$.*

Proof. It suffices to bound the diameter $d_U(\hat{y}_v, \hat{y}_w)$ for any distinct pair of vertices $v, w \in \mathcal{X}$ satisfying both $U \lessdot^{\mathcal{X}} v$ and $U \lessdot^{\mathcal{X}} w$. Let Ω be any witness family in the equivalence class $\mathcal{R}$, say for a segment $[x, y]$. Then by definition $Z_v, Z_w \in \Omega$ with $\hat{y}_v = \widehat{y}^{\Omega}_{Z_v}$ and $\hat{y}_w = \widehat{y}^{\Omega}_{Z_w}$. Since $U \sqsubset Z_v$, the construction of resolution points (Definitions 8.3 & 8.7) implies that $\pi_U(\hat{y}_v) = \pi_U(\widehat{y}^{\Omega}_{Z_v})$ lies within a uniformly bounded distance of the set $\{\pi_U(x), \pi_U(y)\}$. The same holds for $\pi_U(\hat{y}_w)$. Since Lemma 12.5 ensures that $d_U(x, y) \leqslant \mathsf{N}$, the bound on $d_U(\hat{y}_v, \hat{y}_w)$ is therefore immediate. □

12.3. **Realization tuples.** We will show that the number of realizations of $\mathcal{X}$ is controlled by the labels on the vertices of $\mathcal{X}$. The first step is to show that a realization relative to $x \in \mathcal{T}(\Sigma)$ determines a companion point $p = p_{\mathcal{R}} \in \mathcal{T}(\Sigma)$. This point will be built via consistency from a tuple $(\tilde{p}_U) \in \Pi_{U \sqsubset \Sigma}\mathcal{C}(U)$ defined using only the data of the domains Z_v and the points $x \in \mathcal{T}(\Sigma)$ and $\hat{y}_v \in \mathcal{T}(Z_v)$ for $v \in \mathcal{X}$:

Definition 12.7 (Tuples from realizations)**.** Let $\mathcal{R}$ be a realization of $\mathcal{X}$ relative to $x \in \mathcal{T}(\Sigma)$, and define a tuple $(\tilde{p}_U) \in \Pi_{U \sqsubset \Sigma}\mathcal{C}(U)$ as follows: Given $U \sqsubset \Sigma$, if no vertices $v \in \mathcal{X}$ minimally contain U, then we set $\tilde{p}_U = \pi_U(x)$, and otherwise we choose some $v \in \mathcal{X}$ satisfying $U \lessdot^{\mathcal{X}} v$ and set $\tilde{p}_U = \pi_U(\hat{y}_v)$. Corollary 12.6 ensures this is coarsely well defined, independent of the chosen vertex v.

We observe the following:

Lemma 12.8. *If a segment $[x, y]$ realizes $\mathcal{R}$, then for each domain $U \sqsubset \Sigma$ we have:*

- *If $U \notin \Upsilon(x, y)$ then $\mathrm{diam}_{\mathcal{C}(U)}(\tilde{p}_U \cup \pi_U(x) \cup \pi_U(y)) \stackrel{+}{\prec}_{\mathsf{C}} 0$.*
- *If $U \in \Upsilon(x, y)$, then $d_U(\tilde{p}_U, y) \stackrel{+}{\prec}_{\mathsf{C}} 0$ provided its Ω–supremum $U' = \bar{U}^{\Omega}$ satisfies $U' \in \Omega(\mathcal{R})$, and $d_U(\tilde{p}_U, x) \stackrel{+}{\prec}_{\mathsf{C}} 0$ provided $U' \in \Omega(\mathcal{R})$.*

Proof. Let Ω be any witness family in $\mathcal{R}$ for $[x, y]$. Notice, as above, that for any $v \in \mathcal{X}$, the projection of $\hat{y}_v = \widehat{y}^{\Omega}_{Z_v}$ to the curve complex of any subsurface $Y \sqsubset Z_v$ is by construction coarsely either $\pi_Y(x)$ or $\pi_Y(y)$. In particular, by considering any $v \in \mathcal{X}$ minimally containing U, we see that $\tilde{p}_U = \pi_U(\hat{y}_v)$ is coarsely either $\pi_U(x)$ or $\pi_U(y)$. Therefore, if $U \notin \Upsilon(x, y)$ then $d_U(x, y) \leqslant \mathsf{N}_U$ and the first bullet follows.

Now suppose $U \in \Upsilon(x, y)$ and let $U' = \bar{U}^{\Omega}$. If $U' \in \Omega(\mathcal{R})$, then $U \lessdot^{\Omega(\mathcal{R})} U'$ so that by definition $\tilde{p}_U$ is coarsely given by $\pi_U(\widehat{y}^{\Omega}_{U'})$. However, since $U' = \bar{U}^{\Omega}$, the construction in Definitions 8.3–8.7 implies the projection of $\widehat{y}^{\Omega}_{U'}$ to $\mathcal{C}(U)$ coarsely agrees with $\pi_U(y)$. Hence $\tilde{p}_U = \pi_U(\widehat{y}^{\Omega}_{U'})$ is coarsely $\pi_U(y)$ as claimed.

Suppose, on the other hand, that $U' \notin \Omega(\mathcal{R})$. If there is no vertex of $\mathcal{X}$ that minimally contains U, then $\tilde{p}_U = \pi_U(x)$ by definition. However, if $U \lessdot^{\mathcal{X}} v$ for some $v \in \mathcal{X}$, then necessarily $U' \sqsubset Z_v$ by Lemma 7.6 and in fact we must have $U' \subsetneq Z_v$ since by assumption $Z_v \in \Omega(\mathcal{R})$ but $U' \notin \Omega(\mathcal{R})$. The domains U', Z_v are necessarily subordered in Ω, and the option $U' \swarrow Z_v$ is ruled out by the fact that $\mathcal{X} \cong \Omega(\mathcal{R})$ respects the partial order. Hence we have $Z_v \searrow U'$ so that by Definitions 8.3–8.7

the projection of the resolution point $\widehat{y}^{\Omega}_{Z_v}$ to U is coarsely $\pi_U(x)$. But since $U \lessdot^{\mathcal{X}} v$, this projection $\pi_U(\widehat{y}^{\Omega}_{Z_v})$ is by construction the U–coordinate $\tilde{p}_U$ of our tuple. Thus $\tilde{p}_U = \pi_U(\widehat{y}^{\Omega}_{Z_v})$ is coarsely $\pi_U(x)$, as claimed. ☐

Lemma 12.9. *The tuple $(\tilde{p}_U) \in \Pi_{U \sqsubset \Sigma}\mathcal{C}(U)$ is k–consistent, for some $k \stackrel{+}{\prec}_{\mathsf{C}} 0$.*

Proof. Let Ω be a witness family in the equivalence class $\mathcal{R}$, say for a segment $[x, y]$, so that we identify $\mathcal{X}$ with the subset $\Omega(\mathcal{R}) \subset \Omega$. Fix two domains $U, W \sqsubset \Sigma$ that either cut $U \pitchfork W$ or are nested $U \subsetneq W$ or $W \subsetneq U$. If either $U \notin \Upsilon(x, y)$ or $W \notin \Upsilon(x, y)$, then Lemma 12.8 implies $(\tilde{p}_U, \tilde{p}_W)$ is within bounded distance from either $(\pi_U(x), \pi_W(x))$ or $(\pi_U(y), \pi_W(y))$ and is hence consistent by Theorem 3.37.

We may therefore suppose $U, W \in \Upsilon(x, y)$. Let $U' = \bar{U}^{\Omega} \in \Omega$ and $W' = \bar{W}^{\Omega} \in \Omega$ be the Ω–suprema guaranteed by completeness. If neither U' nor W' is in $\Omega(\mathcal{R})$, Lemma 12.8 implies we coarsely have $\tilde{p}_U = \pi_U(x)$ and $\tilde{p}_W = \pi_W(x)$. If, on the other hand, U', W' both lie in $\Omega(\mathcal{R})$, then we coarsely have $\tilde{p}_U = \pi_U(y)$ and $\tilde{p}_W = \pi_W(y)$. In either case, the pair $(\tilde{p}_U, \tilde{p}_W)$ satisfies the consistency condition by Theorem 3.37. It therefore suffices to suppose exactly one of U' or W' lies in $\Omega(\mathcal{R})$ so that, without loss of generality, we suppose $U' \in \Omega(\mathcal{R})$ and $W' \notin \Omega(\mathcal{R})$. In particular, $U' \neq W'$ and, by Lemma 12.8, $d_U(\tilde{p}_U, y) \stackrel{+}{\prec}_{\mathsf{C}} 0$ and $d_W(\tilde{p}_W, x) \stackrel{+}{\prec}_{\mathsf{C}} 0$.

First suppose $U' \subsetneq W'$. The domains U', W' are then subordered in Ω, and the fact that $\mathcal{X}$ respects the partial order implies the subordering must be $U' \swarrow W'$. Note that in this case we must have $U \subsetneq W$ or $U \pitchfork W$, since the containment $W \sqsubset U$ would imply $W' \sqsubset U'$ by Lemma 7.6. Also, the domains U' and W cannot be disjoint, as ∂U projects to both of them, and nor can the be nested $W \sqsubset U'$, as that would again imply $W' \sqsubset U'$ by Lemma 7.6, contrary to our assumption. Thus either $U' \pitchfork W$ or $U' \subsetneq W$. If U' cuts W, then (SO4) with $U' \swarrow W' = \bar{W}^{\Omega}$ implies they must be time-ordered $U' \lessdot W$ along $[x, y]$. Since $d_W(\partial U, \partial U') \leqslant 2$, we thus obtain the desired bound

$$d_W(\tilde{p}_W, \partial U) \stackrel{+}{\prec}_{\mathsf{C}} d_W(x, \partial U') \leqslant \mathsf{M}$$

by Lemma 3.29. It instead $U' \subsetneq W$, there are two cases: Firstly, if $W = W'$, then the fact that Ω is wide with $U' \swarrow W'$ gives

$$d_W(\tilde{p}_W, \partial U) \stackrel{+}{\prec}_{\mathsf{N}} d_W(x, \partial U') = d_{W'}(x, \partial U') \leqslant \tfrac{1}{3}\mathsf{N}_W \leqslant \mathsf{N},$$

which is the desired condition for consistency. Secondly, if $W \neq W'$, then evidently $W \in \Upsilon(x, y)$ but $W \notin \Omega$. Since we have $U' \in \Omega$ with $U' \sqsubset W$, (WF3) provides some $Z \in \Omega$ with $U' \sqsubset Z$ and $Z \pitchfork W$. We claim that Z and W must be time-ordered $Z \lessdot W$ along $[x, y]$. Indeed, if $Z \subsetneq W'$ then (SO1) implies $Z \swarrow W'$ so that (SO4) forces $Z \lessdot W$; similarly if $Z \pitchfork W'$ then we must have $Z \lessdot W'$ since the alternative would give $U' \swarrow W' \lessdot Z$ and contradict (SO3) (since $U' \pitchfork_{W'} Z$ evidently fails). Therefore we conclude that $d_W(\tilde{p}_W, \partial U) \stackrel{+}{\prec}_{\mathsf{N}} d_W(x, \partial Z) \leqslant \mathsf{M}$, as desired.

A completely symmetric argument shows that the assumption $U' \supsetneq W'$ leads to the subordering $U' \searrow W'$. One finds that either $U \supsetneq W$ or $U \pitchfork W$, and $U \supsetneq W'$ or $U \pitchfork W'$, and that in any case $d_U(\tilde{p}_U, \partial W) \stackrel{+}{\prec}_{\mathsf{C}} d_U(y, \partial W')$ is uniformly bounded.

It remains to suppose $U' \pitchfork W'$. In this case the fact that $\mathcal{X}$ respects the partial order implies $U' \lessdot W'$ along $[x, y]$. We must also have $U \pitchfork W'$, since $U \sqsupset W'$ would give $U' \sqsupset U \sqsupset W'$ and $U \sqsubset W'$ would give $U' \sqsubset W'$ by Lemma 7.6. Similarly we necessarily have $U \pitchfork W$, since either alternative $U \sqsubset W$ or $U \sqsupset W$ would yield nesting $U' \sqsubset W'$ or $U' \sqsupset W'$ again by Lemma 7.6. Corollary 3.31

therefore implies that $U \lessdot W'$ and $U \lessdot W$ along $[x, y]$ which gives the desired bound $d_U(\tilde{p}_U, \partial W) \stackrel{+}{\prec}_{\mathsf{C}} d_U(y, \partial W) \leqslant \mathsf{M}$. $\square$

We also need to account for the potentiality of short curves:

Definition 12.10 (Short multicurve of a realization)**.** Let $\mathcal{R}$ be a realization of a subset $\mathcal{X} \subset \mathcal{G}$ that respects the partial order. The *short multicurve associated to* $\mathcal{R}$ is the multicurve $\alpha_{\mathcal{R}}$ consisting of all curves γ such that there is some $v \in \mathcal{X}$ so that Z_v is an annulus whose core ∂Z_v equals γ and is short at $\hat{y}_v \in \mathcal{T}(Z_v)$.

Lemma 12.11. *For any realization $\mathcal{R}$, we have that $\alpha_{\mathcal{R}}$ is indeed a multicurve and that $d_U(\tilde{p}_U, \alpha_{\mathcal{R}}) \stackrel{+}{\prec}_{\mathsf{C}} 0$ for every domain $U \sqsubset \Sigma$.*

Proof. Let Ω be an element of the equivalence class $\mathcal{R}$, say for a segment $[x, y]$. Given any component γ of $\alpha_{\mathcal{R}}$, we may choose some $v \in \mathcal{X}$ so that $A = Z_v$ is an annulus with $\gamma = \partial Z_v = \partial A$ short at the point $\hat{y}_v \in \mathcal{T}(A)$. By construction of this point $\hat{y}_v = \hat{y}^{\Omega}_A$ in Definition 8.7, we necessarily have $\ell_{\hat{y}_v}(\partial A) = \ell_y(\partial A)$; therefore $\gamma = \partial A$ is short at y. In particular, since all components of $\alpha_{\mathcal{R}}$ are short at the single point $y \in \mathcal{T}(\Sigma)$, $\alpha_{\mathcal{R}}$ is indeed a multicurve, as claimed.

For the second claim, it suffices to bound $d_U(\tilde{p}_U, \partial A)$ for all $U \sqsubset \Sigma$. If ∂A is disjoint from U, then $d_U(\tilde{p}_U, \partial A) = \operatorname{diam}_{\mathcal{C}(U)}(\tilde{p}_U)$ is bounded. So we may suppose $A \pitchfork U$ or $A \subsetneq U$. The fact that ∂A is in the Bers marking at y implies that $d_Y(\partial A, y) \leqslant \mathsf{L}$ for all domains $Y \sqsubset \Sigma$. In particular, since $d_U(y, \partial A) \leqslant \mathsf{L}$, it suffices to bound $d_U(\tilde{p}_U, y)$. If $d_U(x, y) \leqslant \mathsf{N}_U$, then $d_U(\tilde{p}_U, y) \stackrel{+}{\prec}_{\mathsf{C}} 0$ by Lemma 12.8, as needed. Hence we may suppose $d_U(x, y) \geqslant \mathsf{N}_U$, so that $U \in \Upsilon(x, y)$, and set $U' = \bar{U}^{\Omega}$. We claim that necessarily $U' \in \Omega(\mathcal{R})$ so that $d_U(\tilde{p}_U, y) \stackrel{+}{\prec}_{\mathsf{C}} 0$ by Lemma 12.8. Otherwise, the fact that $\mathcal{X}$ respects the partial order implies that the domains $A, U' \in \Omega$ must be related by $A \swarrow U'$ or $A \lessdot U'$. However, the latter option $A \lessdot U'$ would give $d_{U'}(x, \partial A) \leqslant \mathsf{M}$ and, since Ω is wide, the former option would give $d_{U'}(x, \partial A) \leqslant \mathsf{N}_{U'}/3$. Since $d_{U'}(\partial A, y) \leqslant \mathsf{L}$, either case implies $d_{U'}(x, y) < \mathsf{N}_{U'}$ and contradicts our assumption. $\square$

We can now use our realization data to reconstruct a point in Teichmüller space:

Definition 12.12 (Realization point)**.** Let $\mathcal{X} \subset \mathcal{G}$ be a subset respecting the partial order. To each realization $\mathcal{R}$ of $\mathcal{X}$ we associate a net point $p_{\mathcal{R}} \in \mathcal{N}(\Sigma)$ as follows: Let $\alpha_{\mathcal{R}}$ be the associated short multicurve and $(\tilde{p}_U) \in \Pi_{U \sqsubset \Sigma}\mathcal{C}(U)$ the associated tuple from Definition 12.7. By Theorem 3.37 and Lemma 3.10, we can build a marking μ on Σ so that $\alpha_{\mathcal{R}} \subset \operatorname{base}(\mu)$ and so that $d_U(\tilde{p}_U, \mu) \stackrel{+}{\prec}_{\mathsf{C}} 0$ for all $U \sqsubset \Sigma$. Now let $p_{\mathcal{R}} \in \mathcal{T}(\Sigma)$ be a net point that has μ as a Bers marking and so that for each component γ of $\alpha_{\mathcal{R}}$, the length $\ell_{p_{\mathcal{R}}}(\gamma)$ coarsely agrees with $\ell_{\hat{y}_v}(\gamma)$, where $v \in \mathcal{X}$ is the vertex so that $\gamma = \partial Z_v$, and so that all other components of $\operatorname{base}(\mu)$ coarsely have length ϵ_0.

Lemma 12.13. *Let $\mathcal{R}$ be any realization of the partially ordered set $\mathcal{X} = \mathcal{V}$ consisting of all vertices of the directed graph $\mathcal{G}$. Then for any segment $[x, y]$ realizing $\mathcal{R}$ we have $d_{\mathcal{T}(\Sigma)}(p_{\mathcal{R}}, y) \stackrel{+}{\prec}_{\mathsf{C}} 0$.*

Proof. Consider any domain $U \sqsubset \Sigma$. By Definition 12.12 we have $d_U(\tilde{p}_U, p_{\mathcal{R}}) \stackrel{+}{\prec}_{\mathsf{C}} 0$. Thus if $U \notin \Upsilon(x, y)$, then Lemma 12.8 implies $d_U(p_{\mathcal{R}}, y) \stackrel{+}{\prec}_{\mathsf{C}} 0$. If instead $U \in \Upsilon(x, y)$, then U has an Ω–supremum $U' = \bar{U}^{\Omega} \in \Omega$ by completeness. Clearly $U' \in \Omega(\mathcal{R})$, since our subset $\mathcal{V}$ consists of all vertices of $\mathcal{G} = \mathcal{G}(\Omega)$, and thus again we find $d_U(p_{\mathcal{R}}, y) \stackrel{+}{\prec}_{\mathsf{C}} 0$ by Lemma 12.8.

By the Distance formula, or rather Lemma 3.35, it remains to show that $p_{\mathcal{R}}$ and y have the same short curves with coarsely the same lengths. By construction of $p_{\mathcal{R}}$ in Definition 12.12 and $\hat{y}_v = \widehat{y}^{\Omega}_{Z_v}$ in Definition 8.7, each short curve γ on $p_{\mathcal{R}}$ is the core of an annulus $Z_v \in \Omega(\mathcal{R})$ on which $\ell_{\hat{y}_v}(\gamma)$ coarsely agrees with $\ell_\gamma(y)$. Conversely, for each short curve β on y, the annulus A with $\partial A = \beta$ satisfies $A \in \Upsilon^\ell(x, y)$. The fact that Ω is insulated implies (Lemma 7.13) that $\Upsilon^\ell(x, y) \subset \Omega$. Thus $A \in \Omega = \Omega(\mathcal{R})$ and, again by Definitions 12.12 and 8.7, the length of β at y coarsely agrees with the length of β on $\widehat{y}^{\Omega}_A$ and thus with the length of β at $p_{\mathcal{R}}$. $\square$

12.4. **Extending realizations.** Finally, we count how many ways there are to extend a realization to an enlarged subset that respects the partial order:

Proposition 12.14. *There is a constant C' depending only on C such that the following holds: Let $\mathcal{X} \subset \mathcal{V}$ be a subset that respects the partial order, and let $v \in \mathcal{V} \backslash \mathcal{X}$ be a vertex so that $\mathcal{X}' = \mathcal{X} \cup \{v\}$ also respects the partial order. Then for each realization $\mathcal{R}$ of $\mathcal{X}$ relative to a point $x \in \mathcal{T}(\Sigma)$, there are at most $\mathsf{C}' e^{h_v^* d_v}$ realizations $\mathcal{R}'$ of $\mathcal{X}'$ that extend $\mathcal{R}$.*

Proof. The equivalence class $\mathcal{R}$ is naturally partitioned into subsets $\mathcal{R}'$ that are each realizations of the larger set $\mathcal{X}'$. Picking such a $\mathcal{R}'$ that extends $\mathcal{R}$ amounts to specifying a domain $Z_v \sqsubset \Sigma$ along with a pair of net points $\hat{x}_v, \hat{y}_v \in \mathcal{N}(Z_v)$. We will show that there are at most boundedly many options for Z_v and $\hat{x}_v$ and that, once these are specified, at most $e^{h_v^* d_v}$ options for $\hat{y}_v$.

Let $(\tilde{p}_U)_{U \sqsubset \Sigma}$ be the tuple associated to $\mathcal{R}$ and $p_{\mathcal{R}}$ the realization point. Also let Ω be a witness family in $\mathcal{R}$, say for a segment $[x, y]$. Thus we have an identification of $\mathcal{G}(\Omega)$ with $\mathcal{G}$ under which $v \in \mathcal{V}$ corresponds to a domain $W \in \Omega$ and the pair $\hat{x}_v, \hat{y}_v$ are given by the resolution points $\widehat{x}^{\Omega}_W, \widehat{y}^{\Omega}_W$. These three pieces of data, W, $\widehat{x}^{\Omega}_W$, and $\widehat{y}^{\Omega}_W$, thus specify the realization $\mathcal{R}' \subset \mathcal{R}$ containing Ω.

Claim 12.15. *The domain W satisfies $d_U(p_{\mathcal{R}}, \partial W) \stackrel{+}{\prec}_{\mathsf{C}} 0$ for every domain $U \sqsubset \Sigma$ and, consequently, the number of such domains W is bounded by Corollary 3.39.*

Proof of Claim. Notice that $W \in \Upsilon(x, y)$ so that W has an active interval along $[x, y]$. It suffices to suppose ∂W projects to U, that is either $W \subsetneq U$ or $W \pitchfork U$, and to consider $d_U(\tilde{p}_U, \partial W)$. If $U \notin \Upsilon(x, y)$, then we have

$$d_U(\tilde{p}_U, \partial W) \stackrel{+}{\prec}_{\mathsf{C}} d_U(x, \partial W) \leqslant d_U(x, \partial W) + d_U(\partial W, y) \stackrel{+}{\prec} d_U(x, y) \leqslant \mathsf{N}_U.$$

by Lemma 12.8 and Corollary 3.27. If instead $U \in \Upsilon(x, y)$, we let $U' = \bar{U}^{\Omega}$. Note that the assumptions $W \subsetneq U$ or $W \pitchfork U$ imply that either $W \subsetneq U'$ or $W \pitchfork U'$. Therefore W and U' are related in the directed graph $\mathcal{G}(\Omega)$.

First suppose $U' \in \Omega(\mathcal{R})$, so that $d_U(\tilde{p}_U, y) \stackrel{+}{\prec}_{\mathsf{C}} 0$ by Lemma 12.8. Since $\mathcal{X}$ respects the partial order and $v \notin \mathcal{X}$ by assumption, we either have $U' \lessdot W$ along $[x, y]$ (if $U' \pitchfork W$) or $U' \searrow W$ in Ω (if $W \subsetneq U'$). In the the former case $U' \pitchfork W$ we necessarily have $U \pitchfork W$ as well (since we cannot have $W \subsetneq U \sqsubset U' \pitchfork W$). Therefore $U' \lessdot W$ implies $U \lessdot W$ by Corollary 3.31 so that $d_U(y, \partial W) \leqslant \mathsf{M}$ as desired. In the latter case $U' \searrow W$, we get $d_{U'}(y, \partial W) \leqslant \mathsf{N}$ by wideness of Ω. If $U' = U$ this is the desired bound. If instead $U \subsetneq U'$ then evidently $U \notin \Omega$. We claim there is some $W' \sqsupset W$ so that $W' \in \Omega$ and $W' \pitchfork U$. Indeed: if $W \pitchfork U$ we just take $W' = W$, and if $W \subsetneq U$ such a W' is provided by (WF3). Now observe that if $W' \pitchfork U'$, then we necessarily have $U' \lessdot W'$ along $[x, y]$ (since $W' \lessdot U' \searrow W$ is ruled out by (SO2)) and thus $U \lessdot W'$ by Corollary 3.31. Alternatively, if $W' \subsetneq U'$ then necessarily

$U' \searrow W'$ by (SO1) so that again we get $U \lessdot W'$ by (SO4). In any case $U \lessdot W'$ along $[x,y]$ and we obtain the desired bound:

$$d_U(\tilde{p}_U, \partial W) \stackrel{+}{\prec}_{\mathsf{C}} d_U(y, \partial W') \leqslant \mathsf{M}.$$

Next suppose $U' \notin \Omega(\mathcal{R})$ so that $d_U(\tilde{p}_U, x) \stackrel{+}{\prec}_{\mathsf{C}} 0$ by Lemma 12.8. Since $W \neq U'$, we see that U' cannot correspond to one of the vertices of $\mathcal{X}' = \mathcal{X} \cup \{v\}$ in $\mathcal{G}$. Since $\mathcal{X}'$ respects the partial order, this means that either $W \lessdot U'$ along $[x,y]$ (if $W \pitchfork U'$) or $W \swarrow U'$ in Ω (if $W \subsetneq U'$). A symmetric argument to the case $U' \in \Omega(\mathcal{R})$ above now shows that either $U' = U$ and $d_U(\partial W, x) \leqslant \mathsf{N}$ by wideness, or $W \sqsubset W' \lessdot U$ for some $W' \in \Omega$ so that $d_U(\tilde{p}_U, \partial W) \stackrel{+}{\prec}_{\mathsf{C}} d_U(x, \partial W') \leqslant \mathsf{M}$. □

Claim 12.16. *There are only boundedly many options for the net point $\hat{x}_v \in \mathcal{N}(W)$.*

Proof of Claim. We show that the net point $\hat{x}_v = \widehat{x}^{\Omega}_W$ is coarsely determined by the data captured by the original realization $\mathcal{R}$. Specifically, we will show that for any domain $U \sqsubset W$, if there is a unique vertex $u \in \mathcal{X}$ satisfying $U \preccurlyeq^{\mathcal{X}} u$, then $d_U(\widehat{x}^{\Omega}_W, \hat{y}_u) \stackrel{+}{\prec}_{\mathsf{C}} 0$, and otherwise $d_U(\widehat{x}^{\Omega}_W, x) \stackrel{+}{\prec}_{\mathsf{C}} 0$.

Let us first suppose there does not exist a unique vertex $u \in \mathcal{X}$ that minimally contains U. If $U \notin \Upsilon(x,y)$, then by construction in Definitions 8.3–8.7 we have $d_U(\widehat{x}^{\Omega}_W, x) \stackrel{+}{\prec}_{\mathsf{C}} 0$, which is the desired bound. If instead $U \in \Upsilon(x,y)$, then we consider $U' = \bar{U}^{\Omega}$ and note that $U' \sqsubset W$ by Lemma 7.6. By completeness, U' is the unique domain of Ω that minimally contains U. Hence the vertex $u \in \mathcal{V} = \mathcal{G}(\Omega)$ corresponding to U' cannot lie in $\mathcal{X}$, as that would contradict our assumption that $\mathcal{X}$ does not have a unique vertex minimally containing U. If $u = v$, that means $U' = W$ and hence that U contributes to W in Ω so that $d_U(\widehat{x}^{\Omega}_W, x) \stackrel{+}{\prec}_{\mathsf{C}} 0$ by construction. The remaining possibility is $u \notin \mathcal{X}' = \mathcal{X} \cup \{v\}$. In this case, the fact that $\mathcal{X}'$ respects the partial order means $U' \subsetneq W$ must be subordered $W \searrow U'$ in Ω. Therefore, by construction, we again have $d_U(\widehat{x}^{\Omega}_W, x) \stackrel{+}{\prec}_{\mathsf{C}} 0$ as desired.

Now suppose there does exist a unique vertex $u \in \mathcal{X}$ satisfying $U \preccurlyeq^{\mathcal{X}} u$. We must show $d_U(\widehat{x}^{\Omega}_W, \hat{y}_u) \stackrel{+}{\prec}_{\mathsf{C}} 0$. Let $Z = Z_u$ be the domain in Ω corresponding to u, so that $\hat{y}_u = \widehat{y}^{\Omega}_Z \in \mathcal{T}(Z)$. If $U \notin \Upsilon(x,y)$, then by definition $d_U(\widehat{y}^{\Omega}_Z, y) \stackrel{+}{\prec}_{\mathsf{C}} 0$ and $d_U(\widehat{x}^{\Omega}_W, x) \stackrel{+}{\prec}_{\mathsf{C}} 0$. Hence by the triangle inequality $d_U(\widehat{x}^{\Omega}_W, \hat{y}_u) \stackrel{+}{\prec}_{\mathsf{C}} d_U(x,y) \leqslant \mathsf{N}_U \stackrel{+}{\prec}_{\mathsf{C}} 0$, as needed. If instead $U \in \Upsilon(x,y)$, we again consider its supremum $U' = \bar{U}^{\Omega}$ and note that $Z \sqsupset U' \sqsubset W$ by Lemma 7.6. If $U' = Z$, that means U contributes to Z in Ω so that $d_U(\widehat{y}^{\Omega}_Z, y) \stackrel{+}{\prec}_{\mathsf{C}} 0$ by construction. Since $Z \in \Omega(\mathcal{R})$ and $W \notin \Omega(\mathcal{R})$, the fact that $\mathcal{X}$ respects the partial order implies the nested domains $Z = U' \subsetneq W$ are subordered $Z \swarrow W$. Therefore $d_U(\widehat{x}^{\Omega}_W, y) \stackrel{+}{\prec}_{\mathsf{C}} 0$ by construction and thus $d_U(\widehat{x}^{\Omega}_W, \hat{y}_u) \stackrel{+}{\prec}_{\mathsf{C}} 0$ by the triangle inequality. If $U' \subsetneq Z$, then the fact that Z minimally contains U in $\Omega(\mathcal{R})$ means we must have $U' \notin \Omega(\mathcal{R})$. Since $\mathcal{X}$ respects the partial order, the nested domains $U' \subsetneq Z$ must therefore be subordered $Z \searrow U'$ so that by construction $d_U(\widehat{y}^{\Omega}_Z, x) \stackrel{+}{\prec}_{\mathsf{C}} 0$. Now we either have $U' = W$, or else $U' \notin \Omega(\mathcal{R}')$ and therefore $W \searrow U'$ by the fact that $\mathcal{X}'$ respects the partial order. In either case we have $d_U(\widehat{x}^{\Omega}_W, x) \stackrel{+}{\prec}_{\mathsf{C}} 0$ by construction. Therefore we again obtain the desired bound $d_U(\widehat{x}^{\Omega}_W, \hat{y}_u) \stackrel{+}{\prec}_{\mathsf{C}} 0$ by the triangle inequality.

The above shows that that the curve complex projections $\pi_U(\hat{x}_v)$ of $\hat{x}_v = \widehat{x}^{\Omega}_W$ to domains $U \sqsubset W$ are coarsely determined by the point x and the data $\{(Z_u, \hat{y}_u) \mid u \in \mathcal{X}\}$ captured by the realization $\mathcal{R}$ of $\mathcal{X}$. When W is not an annulus, the point $\widehat{x}^{\Omega}_W$ is thick by construction and so coarsely determined by its curve complex projections. When W is an annulus, then by construction in Definition 8.7 the core ∂W has coarsely the same length at x and $\widehat{x}^{\Omega}_W$ so that again $\hat{x}_v$ is coarsely determined by the data of x and $\mathcal{R}$. In either case, we conclude there are only boundedly many options for the net point $\hat{x}_v$. $\square$

To conclude the proof of the proposition, the claims show that there are uniformly boundedly many possibilities for the next domain $W = Z_v$ and initial point $\hat{x}_v \in \mathcal{N}(W)$. To finish specifying a realization $\mathcal{R}'$ of $\mathcal{X}'$, it remains to choose a net point $\hat{y}_v \in \mathcal{N}(W)$. But according to the label (h_v^*, d_v) of the vertex $v \in V$, this net point must satisfy $d_{\mathcal{T}(W)}(\hat{x}_v, \hat{y}_v) \leqslant d_v$. Notice that by definition we have $h_v^* = h_W$, unless W is an annulus with both $\hat{x}_v = \widehat{x}_W^\Omega$ and $\hat{y}_v = \widehat{y}_W^\Omega$ ϵ_0–thick (Definition 8.8). In any case, Lemma 3.15, ensures that once $\hat{x}_v$ is specified there are at most $\mathsf{P}e^{h_v^* d_v}$ such net points $\hat{y}_v$. $\square$

12.5. **Finishing the count.** With these tools in hand, it is now a simple matter to complete the proof of Theorem 12.1

Proof of Theorem 12.1. We are given a point $x \in \mathcal{T}(\Sigma)$ and distance $r > 0$ and need to count the number of net points $y \in \mathcal{N}(\sigma)$ so that $\mathfrak{L}(x, y) \leqslant r$. By definition of complexity length, for each such point y there is a WISCL witness family Ω for the segment $[x, y]$ with $\mathfrak{L}(\Omega) = \mathfrak{L}(x, y) \leqslant r$. The corresponding directed graph $\mathcal{G}(\Omega)$ has exactly $|\Omega|$ vertices. Since Ω is limited, this number is bounded $|\Omega| \leqslant \Delta_{-1} + \cdots + \Delta_{\xi(\Sigma)} = \Delta$ in terms of C. Thus there are uniformly boundedly many options for the directed graph $\mathcal{G}(\Omega)$. Each edge has only three possible labels, and for each vertex v there are boundedly many options for the label h_v^*. The remaining vertex labels d_v satisfy $\sum_{v \in \mathcal{V}} h_v^* d_v = \mathfrak{L}(\Omega) \leqslant r$. Since there are at most $r^{|\Omega|}$ ways to partition the integer $\lfloor r \rfloor$ as a sum of $|\Omega|$ nonnegative integers, we conclude there is a constant C'' depending only on C such that there are at most $\mathsf{C}'' r^\Delta$ possibilities for the labeled directed graph $\mathcal{G}(\Omega)$.

Let us now fix such a labeled directed graph $\mathcal{G}$ and count the number of points y producing a witness family Ω with $\mathcal{G}(\Omega) = \mathcal{G}$. Using the partial order (Lemma 12.3), we can enumerate the finite vertex set $\mathcal{V} = \mathcal{V}(\mathcal{G}) = \{v_1, \ldots, v_{|\Omega|}\}$ so that each initial list $\mathcal{X}_i = \{v_1, \ldots, v_i\}$ respects the partial order. Let us count the number of possible realizations $\mathcal{R}_i$ of each of these sets. For the emptyset $\mathcal{X}_0 = \varnothing$, there is exactly one realization $\mathcal{R}_0$, and for each $1 \leqslant i \leqslant |\Omega|$, Proposition 12.14 implies there are at most $\mathsf{C}' e^{h_{v_i}^* d_{v_i}}$ realizations $\mathcal{R}_i$ of $\mathcal{X}_i$ extending each realization $\mathcal{R}_{i-1}$ of $\mathcal{X}_{i-1}$. Thus by induction there are at most $(\mathsf{C}')^i \prod_{j=1}^i e^{h_{v_i}^* d_{v_i}}$ realizations $\mathcal{R}_i$ of $\mathcal{X}_i$. In particular, we conclude that there are at most

$$(\mathsf{C}')^{|\Omega|} \exp\left(h_{v_1}^* d_{v_1} + \cdots + h_{v_{|\Omega|}}^* d_{v_{|\Omega|}}\right) \leqslant (\mathsf{C}')^\Delta \exp(r)$$

realizations of the full vertex set $\mathcal{V} = \mathcal{X}_{|\Omega|}$. Furthermore, by Lemma 12.13, each such realization $\mathcal{R}$ determines a point $p_{\mathcal{R}}$ that lies within bounded distance of the original point y; hence there are uniformly boundedly many net points y that admit a witness family Ω in the equivalence class $\mathcal{R}$. All together, there are at most $kr^\Delta e^r$ potential net points y for which $\mathfrak{L}(x, y) \leqslant r$, where k, Δ depend only on C. $\square$

13. Proving the main theorem

With all the setup in place, it is now fairly straightforward to prove Theorem 1.2. First observe that by the triangle inequality and fact that $\mathrm{Mod}(S)$ acts isometrically, for any points $x, y, x', y' \in \mathcal{T}(S)$ we have

$$\Lambda_{\mathrm{fo}}(x, y, R) \subset \Lambda_{\mathrm{fo}}\left(x', y', R + d_{\mathcal{T}(S)}(x, x') + d_{\mathcal{T}(S)}(y, y')\right).$$

Thus it suffices to prove the upper and lower bounds in Theorem 1.2 for one pair x, y, as it will then follow for any other pair x', y' with increased constants.

13.1. **Lower Bound.** To prove the lower bound in Theorem 1.2, we find a finite-order element ϕ_0 with centralizer the finite group generated by ϕ_0. We thank Dan Margalit for suggesting this example. Take a $4g+2$ regular polygon P with opposite sides identified and let ϕ_0 be the rotation of order $4g+2$ of the polygon. The quotient of P by ϕ_0 is a sphere with 3 marked points corresponding to the center of P, the identified vertices and the center of the edges. A sphere with 3 marked points has trivial mapping class group and therefore the centralizer of ϕ_0 is just the group generated by ϕ_0.

Now we can assume $x = y = x_0$ is fixed by ϕ_0. By Theorem 1.1 [ABEM] there exists $K > 0$ such that for all sufficiently large R there are at least $Ke^{h_S R/2}$ elements $w \in \mathrm{Mod}(S)$ so that the orbit point $w(x_0)$ lies in $\mathrm{Ball}(x_0, R/2)$. The point $w(x_0)$ is fixed by the finite order element $\phi_w = w\phi_0 w^{-1}$, and, since ϕ_w is an isometry,

$$d(x_0, \phi_w(x_0)) \leqslant d(x_0, w(x_0)) + d(w(x_0), \phi_w(x_0)) = 2d(x_0, w(x_0)) \leqslant R.$$

For the lower bound then, it is enough to show that the assignment $w \mapsto \phi_w$ is $4g+2$ to 1. Notice that if $\phi_{w_1} = \phi_{w_2}$, then

$$w_1\phi_0 w_1^{-1} = w_2\phi_0 w_2^{-1},$$

or $w_1^{-1}w_2$ is in the centralizer of ϕ_0. But this means $w_1\phi_0^j = w_2$ for some j.

13.2. **The upper bound.** Fix any $\delta > 0$ and choose the parameter C sufficiently large so that $\mathsf{C}\delta > 3$. Since there are only finitely many conjugacy classes of finite-order elements in $\mathrm{Mod}(S)$, it suffices to prove the upper bound for each conjugacy class separately. Let us therefore fix a finite order element ϕ_0 and take $x = y = x_0$ a fixed point. For each conjugate $\phi \in [\phi_0]$, we let a_ϕ, b_ϕ be the branch points from Proposition 5.5, so that $(x_0, a_\phi, b_\phi, \phi(x_0))$ is strongly Θ–aligned. We need:

Claim 13.1. *We have* $\mathfrak{L}(a_\phi, b_\phi) \overset{+}{\prec}_{\mathsf{C}} \mathfrak{S}(x_0, a_\phi, b_\phi, \phi(x_0))$ *and, consequently,*

$$\mathfrak{L}(x_0, a_\phi) + 2\mathfrak{L}(a_\phi, b_\phi) + \mathfrak{L}(b_\phi, \phi(x_0)) \overset{+}{\prec}_{\mathsf{C}} \mathfrak{L}(x_0, a_\phi, b_\phi, \phi(x_0)) + \mathfrak{S}(x_0, a_\phi, b_\phi, \phi(x_0))$$

Proof. Note that by Theorem 11.2, the first claim implies second. Recall from Proposition 5.5(4) that $d_V(a_\phi, b_\phi) \leqslant \Theta$ for all domains $V \sqsubset S$ except possibly some certain annuli A. By taking $\mathsf{C} \geqslant \Theta$, it follows from (7.1) that $\Upsilon(a_\phi, b_\phi)$ consists only of annuli. Since annuli are not nested, it follows from (WF2) that in fact $\Omega = \Upsilon(a_\phi, b_\phi)$ is the *only* allowed witness family for $[a_\phi, b_\phi]$. Let us call this set Υ. Notice that by Proposition 5.5(5) each such annulus A satisfies $\ell_{a_\phi}(\partial A), \ell_{b_\phi}(\partial B) \geqslant \epsilon_0$ and therefore, by construction (Definition 8.7), the resolution points $\widehat{a_\phi}_A^\Upsilon, \widehat{b_\phi}_A^\Upsilon$ are also ϵ_0–thick. In particular, for each such annulus A we use $h_A^* = 1$ when computing complexity length $\mathfrak{L}(\Upsilon)$ (Definition 8.8).

The savings $\mathfrak{S}(x_0, a_\phi, b_\phi, \phi(x_0))$ is defined an infimum over WISCL fitness families for the tuple, say realized by $\Omega = (\Omega_1, \Omega_2, \Omega_3)$. Since Ω_2 is a witness family for $[a_\phi, b_\phi]$, necessarily $\Omega_2 = \Upsilon$ as above. For each annulus $A \in \Omega_2$, we again use $h^*_A = 1$ for calculating complexity and hence $1 = h_A - h^*_A$ for savings. Therefore

$$\begin{aligned}\mathfrak{S}(x_0, a_\phi, b_\phi, \phi(x_0)) = \mathfrak{S}(\Omega) &= \sum_{i=1}^{3} \sum_{V \in \Omega_i} (h_V - h^*_V) d_{\mathcal{T}(V)}(\widehat{x_{i-1}}_V^{\Omega_i}, \widehat{x}_{iV}^{\Omega_i}) \\ &\geqslant \sum_{A \in \Omega_2} (h_A - h^*_A) d_{\mathcal{T}(A)}(\widehat{a_\phi}_A^{\Upsilon}, \widehat{b_\phi}_A^{\Upsilon}) \\ &= \sum_{A \in \Upsilon} (2-1) d_{\mathcal{T}(A)}(\widehat{a_\phi}_A^{\Upsilon}, \widehat{b_\phi}_A^{\Upsilon}) = \mathfrak{L}(\Upsilon) = \mathfrak{L}(a_\phi, b_\phi) \qquad \square\end{aligned}$$

Using this claim, we now apply Theorem 11.2, which says there is an additive constant C' depending only on C so that

$$\mathfrak{L}(x_0, a_\phi) + 2\mathfrak{L}(a_\phi, b_\phi) + \mathfrak{L}(b_\phi, \phi(x_0)) \leqslant \left(h_S + \frac{3}{\mathsf{C}}\right) d_{\mathcal{T}(S)}(x_0, \phi(x_0)) + 2\mathsf{C}'.$$

Let us suppose that $\mathfrak{L}(x_0, a_\phi) \leqslant \mathfrak{L}(b_\phi, \phi(x_0))$. If not, we may replace ϕ with ϕ^{-1} and, using the same fixed point x_ϕ, observe that $(x_0, a_{\phi^{-1}}, b_{\phi^{-1}}, \phi^{-1}(x_0))$ is Θ–strongly aligned with $b_{\phi^{-1}} = \phi^{-1}(a_\phi)$ and $a_{\phi^{-1}} = \phi^{-1}(b_\phi)$. In this case we have $\mathfrak{L}(x_0, a_{\phi^{-1}}) = \mathfrak{L}(\phi(x_0), b_\phi) < \mathfrak{L}(a_\phi, b_\phi)$ and so proceed in the same way counting ϕ^{-1}. Thus by symmetry we may indeed suppose $\mathfrak{L}(x_0, a_\phi) \leqslant \mathfrak{L}(b_\phi, \phi(x_0))$. It follows that $\mathfrak{L}(x_0, a_\phi) + \mathfrak{L}(a_\phi, b_\phi)$ is at most half the quantity above, and hence that

$$\mathfrak{L}(x_0, a_\phi) + \mathfrak{L}(a_\phi, b_\phi) \leqslant \left(h_S + \frac{3}{\mathsf{C}}\right) \frac{d_{\mathcal{T}(S)}(x_0, \phi(x_0))}{2} + \mathsf{C}' \leqslant (h_S + \delta) \frac{R}{2} + \mathsf{C}'.$$

Now, applying Corollary 12.2, we obtain a constant k such that the number of such pairs (a_ϕ, b_ϕ) is at most

$$k \left((h_S + \delta)\tfrac{R}{2} + \mathsf{C}'\right)^k e^{\frac{(h_S+\delta)}{2} R} e^{\mathsf{C}'} \leqslant k' R^k e^{\frac{(h_S+\delta)}{2} R},$$

for some larger constant $k' > k$ depending only on C, h_S, and δ. Lastly, Theorem 6.1 provides a polynomial p such that each such pair (a, b) arises as (a_ϕ, b_ϕ) for at most $p(R)$ elements $\phi \in [\phi_0]$ with $d_{\mathcal{T}(S)}(x_0, \phi(x_0)) \leqslant R$. Hence the total number of such ϕ is at most $P(R)$ times the the above, and we finally conclude that

$$|\Lambda_{\mathrm{fo}}(x_0, x_0, R)| \leqslant p(R) k' R^k e^{\left(\frac{h_S}{2} + \frac{\delta}{2}\right) R} \leqslant e^{\left(\frac{h_S}{2} + \delta\right) R}$$

holds for all sufficiently large R.

References

[ABC+] J. M. Alonso, T. Brady, D. Cooper, V. Ferlini, M. Lustig, M. Mihalik, M. Shapiro, and H. Short. Notes on word hyperbolic groups. In *Group theory from a geometrical viewpoint (Trieste, 1990)*, pages 3–63. World Sci. Publ., River Edge, NJ, 1991. Edited by Short.

[ABEM] Jayadev Athreya, Alexander Bufetov, Alex Eskin, and Maryam Mirzakhani. Lattice point asymptotics and volume growth on Teichmüller space. *Duke Math. J.*, 161(6):1055–1111, 2012.

[Aou] Tarik Aougab. Uniform hyperbolicity of the graphs of curves. *Geom. Topol.*, 17(5):2855–2875, 2013.

[BKMM] Jason Behrstock, Bruce Kleiner, Yair Minsky, and Lee Mosher. Geometry and rigidity of mapping class groups. *Geom. Topol.*, 16(2):781–888, 2012.

[Bow] Brian H. Bowditch. Uniform hyperbolicity of the curve graphs. *Pacific J. Math.*, 269(2):269–280, 2014.

[CRS] Matt Clay, Kasra Rafi, and Saul Schleimer. Uniform hyperbolicity of the curve graph via surgery sequences. *Algebr. Geom. Topol.*, 14(6):3325–3344, 2014.

[Dur] Matthew Gentry Durham. Elliptic actions on Teichmüller space. *Groups Geom. Dyn.*, 13(2):415–465, 2019.

[EM] Alex Eskin and Maryam Mirzakhani. Counting closed geodesics in moduli space. *J. Mod. Dyn.*, 5(1):71–105, 2011.

[FM] Benson Farb and Dan Margalit. *A primer on mapping class groups*, volume 49 of *Princeton Mathematical Series*. Princeton University Press, Princeton, NJ, 2012.

[Han1] Jiawei Han. Growth of pseudo-Anosov conjugacy classes in Teichmüller space, 2021. Preprint arXiv:2105.08640. To appear in *Groups Geom. Dyn.*

[Han2] Jiawei Han. Growth rate of dehn twist lattice points in Teichmüller space, 2021. Preprint arXiv:2105.08624. To appear in *J. Topol. Anal.*

[HPW] Sebastian Hensel, Piotr Przytycki, and Richard C. H. Webb. 1-slim triangles and uniform hyperbolicity for arc graphs and curve graphs. *J. Eur. Math. Soc. (JEMS)*, 17(4):755–762, 2015.

[Mah1] Joseph Maher. Asymptotics for pseudo-Anosov elements in Teichmüller lattices. *Geom. Funct. Anal.*, 20(2):527–544, 2010.

[Mah2] Joseph Maher. Random walks on the mapping class group. *Duke Math. J.*, 156(3):429–468, 2011.

[Mar] Grigoriy A. Margulis. *On some aspects of the theory of Anosov systems*. Springer Monographs in Mathematics. Springer-Verlag, Berlin, 2004. With a survey by Richard Sharp: Periodic orbits of hyperbolic flows, Translated from the Russian by Valentina Vladimirovna Szulikowska.

[Mas] Howard Masur. Interval exchange transformations and measured foliations. *Ann. of Math. (2)*, 115(1):169–200, 1982.

[Min] Yair N. Minsky. Extremal length estimates and product regions in Teichmüller space. *Duke Math. J.*, 83(2):249–286, 1996.

[MM1] H. A. Masur and Y. N. Minsky. Geometry of the complex of curves. II. Hierarchical structure. *Geom. Funct. Anal.*, 10(4):902–974, 2000.

[MM2] Howard A. Masur and Yair N. Minsky. Geometry of the complex of curves. I. Hyperbolicity. *Invent. Math.*, 138(1):103–149, 1999.

[Nie] Jakob Nielsen. Untersuchungen zur Topologie der geschlossenen zweiseitigen Flächen. III. *Acta Math.*, 58(1):87–167, 1932.

[PP] Jouni Parkkonen and Frédéric Paulin. On the hyperbolic orbital counting problem in conjugacy classes. *Math. Z.*, 279(3-4):1175–1196, 2015.

[Raf1] Kasra Rafi. A combinatorial model for the Teichmüller metric. *Geom. Funct. Anal.*, 17(3):936–959, 2007.

[Raf2] Kasra Rafi. Hyperbolicity in Teichmüller space. *Geom. Topol.*, 18(5):3025–3053, 2014.

[RS] Kasra Rafi and Saul Schleimer. Covers and the curve complex. *Geom. Topol.*, 13(4):2141–2162, 2009.

[Thu] William P. Thurston. On the geometry and dynamics of diffeomorphisms of surfaces. *Bull. Amer. Math. Soc. (N.S.)*, 19(2):417–431, 1988.

[Vee] William A. Veech. The Teichmüller geodesic flow. *Ann. of Math. (2)*, 124(3):441–530, 1986.

Department of Mathematics, Vanderbilt University, Nashville, TN 37240
Email address: `spencer.dowdall@vanderbilt.edu`

Department of Mathematics, University of Chicago, Chicago, IL 60637
Email address: `masurhoward@gmail.com`