BICOMBING THE MAPPING CLASS GROUP AND TEICHMÜLLER SPACE VIA STABLE CUBICAL INTERVALS

MATTHEW GENTRY DURHAM

ABSTRACT. In this mostly expository article, we provide a new account of our proof with Minsky and Sisto that mapping class groups and Teichmüller spaces admit bicombings. More generally, we explain how the hierarchical hull of a pair of points in any colorable hierarchically hyperbolic space is quasi-isometric to a finite CAT(0) cube complex of bounded dimension, with the added property that perturbing the pair of points results in a uniformly bounded change to the cubical structure. Our approach is simplified and new in many aspects.

1. INTRODUCTION

The recent infusion of techniques from cube complexes into the study of mapping class groups of finite-type surfaces has had dramatic consequences, including toward our ability to control top-dimensional quasi-isometrically embedded flats [BHS21], build bicombings [DMS20, HHP23], and understand their geometry at infinity [DZ22, Ham21, Dur23], among others [ABD17, Pet21, DZ22, Zal25].

The central result facilitating most of the above work was the *cubical model theorem* of Behrstock–Hagen–Sisto [BHS21]. It says, roughly, that the coarse convex hull of a finite set of points in the mapping class group is quasi-isometrically approximated by a CAT(0) cube complex, which is a higher dimensional analogue of Gromov's tree approximation theorem for hyperbolic spaces. The cubical model theorem was later generalized by Bowditch [Bow19] (to certain coarse median spaces) and the author [Dur23] (to allow for points at infinity). All versions hold in the broader context of hierarchically hyperbolic spaces [BHS17b], a framework for studying a wide variety of spaces with blended negative and positive curvature, e.g. Teichmüller spaces [Bro02, Raf14, Dur16], that built off the seminal work of Masur–Minsky [MM99, MM00b] on the mapping class group; see Section 3.

The cubical approximation theorem allows one to build any number of cubically-defined (and hierarchically well-behaved) paths between a pair of points by importing them from the cubical model of their hull. This observation was at the core of two recent proofs that the mapping class group is *bicombable*[1]. The two proofs, due to Haettel–Hoda–Petyt [HHP23] and Minsky, Sisto, and the author [DMS20], are very different but both depend crucially on this cubical model machinery. See also a recent elegant proof of bicombability using wallspace techniques due to Petyt–Zalloum [PZS24].

The purpose of this paper is to explain a new proof of the two-point version of the main technical result of [DMS20] via the machinery in [Dur23], which we state here informally (see Theorem 9.9):

[1]Roughly, a bicombing on a metric space is a transitive family of quasi-geodesic paths that are stable under small perturbations of the endpoints; see Definition 2.16.

Theorem A. The hierarchical hull of a pair of points in a colorable HHS can be stably approximated by a CAT(0) cube complex.

With this tool in hand, bicombability is a straight-forward application of cubical techniques essentially via the arguments in [DMS20, Sections 5 and 6]:

Corollary B. Let S be a finite type surface. Then its mapping class group and Teichmüller space with the Teichmüller metric admit bicombings by hierarchy paths.

Remark 1.1. One can upgrade both of these statements to ensure coarse equivariance in the presence of an appropriate group action. In particular, in [DMS20] we showed that the bicombings in Corollary B are coarsely equivariant with respect to the mapping class group action, and that the mapping class group is *semi-hyperbolic* in the sense of [AB95]. This equivariance upgrade is relatively straight-forward once one has the bicombing, but we have omitted it since it adds yet another layer complexity to the statements and proofs in an already technical construction. We point the reader to [DMS20, Remark 2] for a discussion of how equivariance works for this construction.

We have several motivations for writing this paper. First, we wanted to give a simplified account of the cubical model results in [DMS20,DMS25] in less generality. Roughly speaking, the cubical model for a finite subset F of an HHS from [BHS21] is produced from a wallspace on a family of trees obtained by projecting F to each of the hyperbolic spaces in the hierarchy. The proof of bicombability in [DMS20] proceeds by using hyperbolic geometry techniques—culminating in the Stable Tree Theorem [DMS20, Theorem 3.2]—to produce a family of trees which is coarsely stable under bounded perturbations of F. In this paper, we simplify a number of aspects of this initial portion of the argument by limiting ourselves to modeling only a pair of points instead of arbitrary collections of points and rays.

The second motivation is to explain how to combine the stability techniques from [DMS20] with the new cubical model construction from [Dur23][2]. This involves a new, more elementary proof of the key Stable Trees Theorem for intervals; see Section 8 below. Notably, with Minsky and Sisto, we recently [DMS25] implemented a significant expansion of the techniques from [DMS20] utilizing the framework of [Dur23][3]. This allowed us to build an asymptotically CAT(0) metric [Kar11] on any colorable HHS, construct a $\mathcal{Z}$-boundary (in the sense of [Bes96, Dra06]) for any colorable HHG (see also [Ham25] for mapping class groups), and establish the Farrell–Jones Conjecture [FJ93] for many HHGs, including extra-large type Artin groups. This paper can be seen as explaining the simplest possible case of the main technical construction in [DMS25].

Unsurprisingly, combining the work in [DMS20] and [Dur23] involves an exposition of both, though we have mostly focused on where they meet. We summarize the main features of this (mostly) expository paper as follows:

- We give an exposition of the basic cubical model construction from [Dur23], which is significantly simplified from the much more general setting considered therein, (i.e., which allows for modeling points at infinity).

[2]This construction has many extra benefits as compared to the construction from [BHS21]. We point interested readers to the in-depth discussion in the introduction of [Dur23].

[3]The first version of this expository article appeared before [DMS25], but we have subsequently adapted it to account for that paper.

- We explain how to combine the stability techniques from [DMS20] with the cubical model construction of [Dur23]. In particular, this involves defining a refined notion of a *stable decomposition* (Definition 8.12), and giving a new proof of the Stable Trees Theorem 8.14 for intervals. We note that these stable decompositions also appear in [DMS25].
- Using the above, we prove bicombability of mapping class groups (and HHSes) via the stable cubulation techniques from [DMS20], while restricting ourselves to the simplified setting of cubical models for pairs of points. Notably, the bicombing construction from [HHP23] also depends on the cubical models (from [BHS21]), so one can also see this paper as supplying an exposition for some of the underlying technology of their construction.
- By focusing on the bicombing result (Corollary B), we are able significantly simplify the endgame of the path construction. In particular, we utilize Niblo–Reeves paths [NR98] in the cubical model instead of the more elaborate version of Niblo–Reeves contractions from our work with Minsky and Sisto [DMS20, Section 5].

1.1. **Background reading.** Our main focus in this article is explaining how to adapt the stable cubulation techniques from [DMS20] to the cubical model setup from [Dur23], which we note is done in much more general setting in [DMS25]. As such, the reader will benefit from a variety of contextual information and motivation in and around cube complexes and HHSes. We provide some suggested reading in the list below:

- See Sisto's notes [Sis17] and the background in Casals-Ruiz–Hagen–Kazachkov [CRHK21] for an introduction to HHSes. While we do give a proof of the harder bound of the Distance Formula in Corollary 6.11, we expect the reader to be familiar with the general framework.
- See Farb–Margalit [FM11] for an introduction to mapping class groups, and [BHS19, Section 11] and there references therein (particular, Masur–Minsky's work [MM99, MM00b]) for its hierarchically hyperbolic structure. Similarly, see Rafi [Raf07, Raf14] and the author [Dur16] for the hierarchical structure on Teichmüller space with the Teichmüller metric.
- For an introduction to bicombings and semihyperbolicity, see Alonso–Bridson [AB95].
- For background on cube complexes, see Sageev's notes [Sag14] and Schwer's book [Sch].
- For a discussion of cubical techniques in HHSes and related objects, see Zalloum's expository article [Zal23].

1.2. **Structure of the paper.** The paper begins with a high-level discussion of the endgame for the bicombing construction using cubical approximation techniques. In particular, in Section 2, we will extract the key facts that we prove about colorable hierarchically hyperbolic spaces, namely the Stable Cubes Theorem 9.9 and Stable Moves Theorem 2.13, to build a bicombing on any metric space possessing the properties described in the former. This includes the much simplified construction of the bicombing paths themselves via Niblo–Reeves paths in Subsection 2.3.

With this setup as motivation, we then proceed to describe the cubical model construction for HHSes. Sections 4 through 6 explain the interval version of the cubical model theorem from [Dur23]. The heart of the paper is in Sections 7 and

8, where we explain how to build the stabilized family of intervals which serve as input, as well as in Section 9, where we see how to plug this stabilized input into the cubulation machine.

1.3. **Acknowledgements.** This work was supported in part by NSF grant DMS-1906487 and NSF CAREER DMS-2441982. The main thanks go to Yair Minsky and Alessandro Sisto, as our papers [DMS20, DMS25] are at the center of this article. Thanks to Thomas Koberda for his efforts in coorganizing the first Riverside Workshop on Geometric Group Theory[4] in 2023 (and its subsequent versions), which was supported by NSF grant DMS-2234299. Thanks also to Harry Petyt for giving many detailed comments which greatly improved the paper. This paper was written for the conference proceedings and I appreciated Thomas' patience and encouragement while completing it. Thanks also to the other speakers—Anne Lonjou, Marissa Kawehi Loving, and Jean-Pierre Mutanguha—for their outstanding minicourses. Finally, thanks to the attendees of RivGGT23 where my minicourse covered some of the material in this paper, and the UCR graduate students who attended my topics course on the construction in [Dur23].

2. Beginning at the end

The purpose of this section is to prove the main consequence of the technical results we will spend the rest of the paper establishing. That is, we will begin by explaining how the Stable Cubical Intervals Theorem A (Theorem 9.9) is enough to prove bicombability of the mapping class groups and colorably hierarchically hyperbolic spaces, more generally.

We begin with a brief primer on CAT(0) cube complexes and a discussion of Niblo–Reeves paths and their stability under hyperplane deletions. In Subsection 2.2, we give a precise version of our stable cubulations result along with its associated path-building construction. In Subsection 2.4, we then define and discuss bicombings. In Subsection 2.5, we explain how to put these parts together to give a bicombing for any metric space satisfying some version of our stable cubulations setup.

2.1. **Wall spaces, cube complexes, and hyperplane deletions.** This subsection includes a quick set of definitions and references for readers unfamiliar with CAT(0) cube complexes.

Our approach here follows most closely Roller's approach [Rol16] to Sageev's cubulation machine [Sag95], which uses the notion of wall-space as introduced by Haglund-Paulin [HP98]. See Sageev's notes [Sag14], as well as Hruska-Wise [HW14], for excellent expositions on this subject. We note that we will later take an approach using medians in Section 6.

A *cube complex* is a CW complex obtained by gluing a collection of unit Euclidean cubes of various dimensions along their faces by a collection of (Euclidean) isometries. A cube complex Σ is *non-positively curved* when it satisfies Gromov's link condition, and every non-positively curved simply connected cube complex admits a unique CAT(0) metric.

[4] https://sites.google.com/view/riggthew-2023/home

A *hyperplane* in a CAT(0) cube complex Σ is a connected subspace whose intersection with each cube $\sigma = [0,1]^n$ is either $\varnothing$ or the subspace obtained by restricting exactly one coordinate to $\frac{1}{2}$. In particular, every hyperplane partitions Σ. This motivates the definition of a half-spaces and pocsets, following Roller [Rol16].

Definition 2.1 (Pocset)**.** Let $(\Sigma, <)$ be a partially-ordered set. A *pocset* structure on Σ has the additional structure of an order-reversing involution $A \mapsto A^*$ satisfying

(1) $A \neq A^*$ and A, A^* are $<$-incomparable;
(2) $A < B \implies B^* < A^*$.

- Given $A < B$, the *interval* between A and B is $[A, B] = \{C | A < C < B\}$.
- The pocset $(\Sigma, <)$ is *locally finite* if every interval is finite.
- Given $A, B \in \Sigma$, we say A is *transverse* to B if none of $A < B, A < B^*, A^* < B, A^* < B^*$ holds. The pocset $(\Sigma, <)$ has *finite width* if the maximal size of a pairwise transverse subset is finite.

In a CAT(0) cube complex, every vertex is contained in exactly one half-space associated to each hyperplane. The following notion abstracts the combinatorial structure of such a family of half-spaces:

Definition 2.2 (Ultrafilters and the DCC)**.** Given a locally finite pocset $(\Sigma, <)$, an *ultrafilter* α on Σ is a subset of Σ satisfying

(1) (Choice) Exactly one element of each pair $\{A, A^*\}$ is an element of α;
(2) (Consistency) $A \in \alpha$ and $A < B \implies B \in \alpha$.

- An ultrafilter α satisfies the *descending chain condition* (DCC) if every descending chain of elements terminates.

We can now define the dual cube complex to a locally finite, discrete, finite-width pocset:

Definition 2.3 (Dual cube complex)**.** Given a locally finite, discrete, finite-width pocset $(\Sigma, <)$, we define a CW complex $X(\Sigma)$ as follows:

(1) The 0-cells $X^{(0)}$ of X are the DCC ultrafilters on $(\Sigma, <)$.
(2) Two 0-cells α, β are connected by an edge in $X^{(1)}$ if and only if $|\alpha \triangle \beta| = 2$, i.e. $\alpha = (\beta - \{A\}) \cup \{A^*\}$ for some $A \in \Sigma$.
(3) Define $X^{(n)}$ inductively by adding an n-cube whenever the boundary of one appears in the $X^{(n-1)}$-skeleton.

The following is a theorem of Roller [Rol16], building on Sageev [Sag95]:

Theorem 2.4 (Sageev's machine)**.** *Given a locally finite, finite-width pocset $(\Sigma, <)$, the CW complex $X(\Sigma)$ is a finite dimensional CAT(0) cube complex, where the dimension is equal to the width of $(\Sigma, <)$.*

- *We call $X(\Sigma)$ the* dual cube complex *to $(\Sigma, <)$.*

We note that the dual cube complex to the pocset coming from the hyperplanes of a CAT(0) cube complex is (isomorphic to) that cube complex.

The following lemma is useful for obtaining isomorphisms between cube complexes whose wall-spaces we can directly compare. Its proof is an exercise in the basics of CAT(0) cube complexes which we leave to the reader.

Lemma 2.5. *Let $\mathcal{W}$, $\mathcal{W}'$ be wallspaces, and let $\iota : \mathcal{H}_{\mathcal{W}} \to \mathcal{H}_{\mathcal{W}'}$ be a bijection of their half-spaces, which preserves complements and disjointness. Then ι induces a*

cubical isomorphism $h : \mathcal{D}(\mathcal{W}) \to \mathcal{D}(\mathcal{W}')$ *between the corresponding dual CAT(0) cube complexes.*

Finally, in our stable cubulations theorem, we will want to realize maps between certain cube complexes as a sort of collapsing of a controlled piece of their combinatorial data. The right notion here is of a *hyperplane deletion*. The motivating example is where one has a wall-space $\mathcal{W}(\mathcal{H})$ on some set of walls $\mathcal{H}$, and one removes some collection of the walls $\mathcal{H} - \mathcal{G}$, to obtain a *restriction map* $\mathcal{D}(\mathcal{W}(\mathcal{H})) \to \mathcal{D}(\mathcal{W}(\mathcal{H} - \mathcal{G}))$ between their dual cube complexes in the sense of [CS11]. This map collapses the carriers of the hyperplanes in $\mathcal{G}$ (recall that a carrier of a hyperplane $\mathbf{h}$ is naturally isomorphic to $\mathbf{h} \times [0, 1]$).

2.2. **Stable cubes and stable moves.** We will prove Corollary B in an ostensibly more general framework than that of (colorable) HHSes[5]. That is, we only need a metric space which satisfies the following definitions. The first is effectively a weak two-point version of the cubical model property for HHSes from [BHS21], proven in Theorem 6.4 in this paper.

Definition 2.6 (Weak local quasi-cubicality)**.** Let X be a metric space. A B_0-*cubical model* for a pair $a, b \in X$ is a CAT(0) cube complex $\mathcal{Q}$ of dimension at most B_0 satisfying the following:

(1) There is a (B_0, B_0)-quasi-isometric embedding $\widehat{\Omega} : \mathcal{Q} \to X$ (the *cubical model map*),
(2) There are two *distinguished vertices* $\hat{a}, \hat{b} \in \mathcal{Q}$ so that any ℓ^∞-geodesic $\gamma \subset \mathcal{Q}$ between $\hat{a}, \hat{b}$ is sent to a (B_0, B_0)-quasi-geodesic $\widehat{\Omega}(\gamma) \subset X$ between a, b.

- We say X is B_0-*weakly locally quasi-cubical* (wLQC) if there exists $B_0 > 0$ so that any pair $a, b \in X$ admits a B_0-cubical model.

While CAT(0) cube complexes have many standard metrics (all of which are quasi-isometric), the convenient one for us is the ℓ^∞-metric, which, for vertices in a single unit cube, is just the sup-metric.

The next definition is extracted from Theorem 9.9 below. This property holds for colorable hierarchically hyperbolic spaces, e.g., the mapping class group and Teichmüller space; see [DMS20, Theorem 4.1] for the original, more general version, and [DMS25, Theorem 9.1] for the state-of-the-art. Part of our purpose here for extracting it from the HHS setting is to highlight the structure of the endgame.

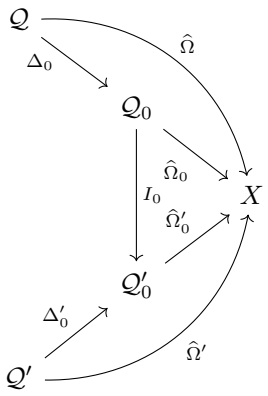
Definition 2.7 (Stable weak local quasicubicality)**.** We say that a metric space X is B_0-*stably wLQC* if there is a B_0 so that the following holds:

- For any $a, b, a', b' \in X$ with $d_X(a, a') \leqslant 1$ and $d_X(b, b') \leqslant 1$, there exist B_0-cubical models $\mathcal{Q}, \mathcal{Q}_0$ and $\mathcal{Q}', \mathcal{Q}'_0$ for a, b and a', b' so that the following

[5]We actually suspect that coarse median spaces satisfying some version of the property in Definition 2.7 will be HHSes, though it is unclear to us what exactly the right property is, e.g. perhaps a factor system in the sense of [BHS17b].

diagram commutes up to additive error B_0:

(1)



- In the above,
  (1) $\widehat{\Omega}, \widehat{\Omega}_0, \widehat{\Omega}', \widehat{\Omega}'_0$ are cubical model maps and hence (B_0, B_0)-quasi-isometric embeddings,
  (2) the map I_0 is a cubical isomorphism,
  (3) the maps Δ_0, Δ'_0 are induced by deleting at most B_0-many hyperplanes, and
  (4) Δ_0, Δ'_0 take the distinguished vertices associated to a, b and a', b' in $\mathcal{Q}, \mathcal{Q}'$ to those in $\mathcal{Q}_0, \mathcal{Q}'_0$, respectively.

Requirement (4) says that the way that the cube complexes $\mathcal{Q}, \mathcal{Q}_0$ model a, b align, and similarly for $\mathcal{Q}', \mathcal{Q}'_0$ and a', b'. Note that if X satisfies Definition 2.7 then it satisfies Definition 2.6.

Remark 2.8. Theorems 6.4 and 9.9 actually provide a lot more information about these cubical models. In particular, if $\mathcal{X}$ is an HHS and $a, b \in \mathcal{X}$, then the cube complexes $\mathcal{Q}$ and $\mathcal{Q}_0$ above model the *hierarchical hull* of a, b (Definition 3.12), which is a coarsely convex subspace which encodes the hierarchical geometry between a, b. Moreover, the combinatorial geometry of the cube complex encodes essentially all of this hierarchical data. See the introduction of [Dur23] for a detailed discussion.

2.3. **Stable moves.** In this subsection, we establish the key path-building portion of our bicombing construction, namely the Stable Moves Theorem 2.13. This involves proving that a well-studied family of paths in a cube complex due to Niblo–Reeves [NR98] behaves well under hyperplane deletion. These paths were at the center of their construction of a biautomatic structure on an CAT(0) cubical group, and so it is quite agreeable that they play a central role in our construction of a bicombing for our sort of quasi-cubical spaces. While we have maintained some of the structure and terminology from our work with Minsky and Sisto in [DMS20, Section 5], our limited setting allows for a much simpler path construction.

Let X be a finite CAT(0) cube complex with $X = \mathrm{hull}_X(a, b)$ the cubical hull[6] of $a, b \in X^0$. Set $\mathcal{H} = \mathcal{H}(a, b)$ to be the set of hyperplanes in X separating a, b, so that $X = \mathcal{X}(\mathcal{H})$ is the dual cube complex to $\mathcal{H}$. Recall that we can think of vertices of X as coherent orientations on hyperplanes in $\mathcal{H}$.

[6]The assumption that $X = \mathrm{hull}_X(a, b)$ is not necessary, but is convenient for our purposes.

Definition 2.9. Let $p \in X^0$. A hyperplane $\mathbf{h} \in \mathcal{H}$ is *transitional*, denoted $\mathbf{h} \in T(p, \mathcal{H})$, if

(1) $\mathbf{h}$ separates p from b, and
(2) p is *adjacent* to $\mathbf{h}$, i.e. $\mathbf{h}$ crosses some cube of which p is a vertex.

We note that there is a natural partial order $\prec$ on $\mathcal{H}$, where $\mathbf{h} \prec \mathbf{h}'$ when $\mathbf{h}$ separates $\mathbf{h}'$ from a, or equivalently $\mathbf{h}'$ separates $\mathbf{h}$ from b. In our special setting with $X = \mathrm{hull}_X(a, b)$, every vertex $p \in X^0$ occurs on some ℓ^1-geodesic from a to b. The property of transitionality for $\mathbf{h} \in T(p, \mathcal{H})$ is encoding that $\mathbf{h}$ is minimal in the partial order on $\mathcal{H}(p, b) \subset \mathcal{H}$.

The following lemma says that minimal elements in the partial order must cross:

Lemma 2.10. *If $\mathbf{h}_1, \mathbf{h}_2 \in T(p, \mathcal{H})$, then $\mathbf{h}_1$ crosses $\mathbf{h}_2$.*

Proof. If $\mathbf{h}_1$ and $\mathbf{h}_2$ do not cross, then, without loss of generality, there is a half-space of $\mathbf{h}_2$ containing $a, \mathbf{h}_1$. On the other hand, p is adjacent to $\mathbf{h}_2$. This implies that p, a are in different half-spaces of $\mathbf{h}_1$, which is a contradiction. □

With this lemma, we can define our *Niblo–Reeves* paths $p_0 = a, p_1, \ldots, p_n = b$ as follows [NR98]:

- Let $p_0 = a$.
- By Lemma 2.10, p_0 is the corner of a unique cube C_0 for which the set of hyperplanes that cross C_0 is precisely $T(p, \mathcal{H})$. Set p_1 to be the diagonally opposite corner of C_0. Equivalently, p_1 is obtained from p_0 by flipping orientations on the hyperplanes in $T(p, \mathcal{H})$.
- Define p_i for $i \geqslant 1$ recursively with p_i in place of p_0.

Thus, at each step $p_i \mapsto p_{i+1}$, the path crosses the set of $\prec$-minimal hyperplanes in $\mathcal{H}(p_i, b)$. The following proposition states the basic properties of the Niblo–Reeves paths:

Proposition 2.11. *Continuing with the above notation, the following hold:*

(1) *There exists some n so that $p_n = b$.*
(2) *For all $0 \leqslant i \leqslant n$, we have $d_X^\infty(p_i, p_{i+1}) = 1$.*
(3) *Any path obtained by concatenating ℓ^1-geodesics between p_i, p_{i+1} for $0 \leqslant i < n$ is an ℓ^1-geodesic.*

Proof. For (1), observe that the sequence p_i only crosses hyperplanes separating a from b. Each step $p_i \mapsto p_{i+1}$ of the sequence crosses the collection of $\prec$-minimal hyperplanes in $\mathcal{H}(p_i, b) \subset \mathcal{H}$. Now (1) is equivalent to the claim that $\mathcal{H} = \bigcup_i \mathcal{H}(p_i, b)$, which is to say that this process exhausts the set of hyperplanes separating a and b, which is clear because $\mathcal{H}$ is finite.

Item (2) is by construction, since p_i and p_{i+1} are opposite corners of a cube.

Finally, item (3) follows from the fact that $\mathcal{H}(p_i, b) \subset \mathcal{H}(p_{i+1}, b)$ with $\mathcal{H}(p_{i+1}, b) - \mathcal{H}(p_i, b) = \mathcal{H}(p_i, p_{i+1})$. Hence if $\mathbf{h} \in T(p_i, b)$, then $\mathbf{h}$ does not separate p_i from p_j for $j < i$. This says that the sequence $\{p_i\}$ does not cross any hyperplane twice, and it follows that the concatenation of the ℓ^1-geodesic segments will not either, making it an ℓ^1-geodesic. □

The following is a useful observation about this construction:

Proposition 2.12. *Let $\Phi : \mathcal{Q} \to \mathcal{Q}'$ be an isomorphism of cube complexes. If γ is the Niblo–Reeves path between $a, b \in \mathcal{Q}^0$, then $\Phi(\gamma)$ is the Niblo–Reeves path between $\Phi(a), \Phi(b) \in (\mathcal{Q}')^0$.*

We can now state the main result of this subsection.

Theorem 2.13 (Stable moves)**.** *Suppose that* $\mathbf{h} \in \mathcal{H}$ *is hyperplane, and let* $\Delta_{\mathbf{h}} : \mathcal{D}(\mathcal{H}) \to \mathcal{D}(\mathcal{H} - \{\mathbf{h}\})$ *denote the hyperplane collapse map. Set* $\Delta_{\mathbf{h}}(a) = a'$ *and* $\Delta_{\mathbf{h}}(b) = b'$*. Let* $p_0, \dots, p_n$ *denote the Niblo–Reeves path in* $X = \mathcal{D}(\mathcal{H})$ *between* a, b *and* $p'_0, \dots, p'_m$ *the path between* a', b' *in* $\mathcal{D}(\mathcal{H} - \{\mathbf{h}\})$*. Then the following hold:*

1. *We have* $n \leqslant m + 1$.
2. *For all* $1 \leqslant i \leqslant n$*, we have* $d^{\infty}_{\mathcal{D}(\mathcal{H}-\{\mathbf{h}\})}(\Delta_{\mathbf{h}}(p_i), p'_i) \leqslant 1$.

In other words, Niblo–Reeves paths are stable under hyperplane deletion maps. The following notation will be useful for us:

Definition 2.14. Let $\mathbf{h}' \in \mathcal{H}$. We set $t_{\mathbf{h}'}(a, \mathcal{H})$, called the *transition index* of $\mathbf{h}'$, to be the index i at which $\mathbf{h}' \in \mathcal{H}(p_i, p_{i+1})$. Similarly, if $\mathbf{h}' \in \mathcal{H} - \{\mathbf{h}\}$, we define $t_{\mathbf{h}'}(a', \mathcal{H} - \{\mathbf{h}\})$ to be the index i so that $\mathbf{h}' \in \mathcal{H}(p'_i, p'_{i+1})$.

Lemma 2.15. *If* $\mathbf{h}' \in \mathcal{H} - \{\mathbf{h}\}$*, then* $t_{\mathbf{h}'}(a, \mathcal{H}) = t_{\mathbf{h}'}(a', \mathcal{H} - \{\mathbf{h}\}) + \delta$ *for* $\delta \in \{0, 1\}$.

Proof. This is essentially [DMS20, Lemma 5.12], though the proof is basically immediate in our setting. Note that the transition index $t_{\mathbf{h}} = d_{\infty}(a, \mathbf{h}) + \frac{1}{2}$. Since the ℓ^{∞} distance between a' and $\mathbf{h}'$ measures the maximal length of a chain of hyperplanes separating them (plus $\frac{1}{2}$), that distance can be at most 1 more than $d_{\infty}(a, \mathbf{h})$. This completes the proof. □

Proof of the Stable Moves Theorem 2.13. Suppose $\mathbf{h}_1, \mathbf{h}_2$ separate $\Delta_{\mathbf{h}}(p_i)$ and p'_i in $X' = \mathcal{D}(\mathcal{H} - \{\mathbf{h}\})$ but do not cross, since we are done if all such hyperplanes cross.

Without loss of generality, we may assume that $\mathbf{h}_1$ separates $\Delta_{\mathbf{h}}(p_i)$ from $\mathbf{h}_2$. Let $j = t_{\mathbf{h}_1}(a', \mathcal{H} - \{\mathbf{h}\})$. Then $j < t_{\mathbf{h}_2}(a', \mathcal{H} - \{\mathbf{h}\}) < i$, with the latter inequality following because p'_{i+1} has already crossed $\mathbf{h}_2$ (and $\mathbf{h}_1$). Thus $j + 1 < i$.

But then $t_{\mathbf{h}_1}(a, \mathcal{H}) \leqslant t_{\mathbf{h}_1}(a', \mathcal{H} - \{\mathbf{h}\}) + 1 = j + 1 < i$, which says that $\mathbf{h}_1$ separates p_i from a, which is a contradiction. □

2.4. **Bicombings.** Our ultimate goal is to prove that mapping class groups and Teichmüller spaces admits bicombings (Theorem A). Roughly, a bicombing on a space X is a family of uniform quasi-geodesic paths that is transitive (i.e., connects any pair of points of X), and for which paths in the family between pairs of endpoints that are close to each other uniformly fellow-travel in a parameterized fashion. The notion of bicombing was first introduced by Thurston in the context of algorithmic problems in groups. It is a straightforward but informative exercise to prove that geodesics in hyperbolic and CAT(0) spaces satisfy this property, and this is precisely what the definition is intended to generalize.

The following definition is essentially from [AB95], as restated in [DMS20, Definition 6.4]:

Definition 2.16 (Bicombings)**.** A *discrete, bounded, quasi-geodesic bicombing* of a metric space X consists of a family of discrete paths $\{\Gamma_{x,y}\}_{x,y \in X}$ and a constant $\zeta > 0$ satisfying the following:

1. *Quasi-geodesic*: For any $x, y \in X$ with $d = d_X(x, y)$, there exists $n_{x,y} \leqslant d\zeta + \zeta$ so that the path $\Gamma_{x,y} : \{0, \dots, n_{x,y}\} \to X$ is a (ζ, ζ)-quasi-isometric embedding with $\Gamma_{x,y}(0) = x$ and $\Gamma_{x,y}(n_{x,y}) = y$; and

(2) *Fellow-traveling*: If $x', y' \in X$ and $d_X(x, x'), d_X(y, y') \leqslant 1$, then for all $t \in \{0, \dots, \max\{n_{x,y}, n_{x',y'}\}\}$, we have

$$d_X(\Gamma_{x,y}(t), \Gamma_{x',y'}(t)) \leqslant \zeta.$$

For our purposes, the fellow-traveling condition in item (2) is the key challenge. At this point, there are many constructions of uniform quasigeodesics in mapping class groups.

2.5. **Building the bicombing.** We are now ready to define the bicombing. Let X be B_0-stably wLQC (Definition 2.7). For any $a, b \in X$, let $\mathcal{Q}$ be a CAT(0) cube complex provided by Definition 2.7 and $\hat{a}, \hat{b} \in \mathcal{Q}$ the distinguished vertices associated to a, b.

We can define a family of paths $\{\Gamma_{a,b}\}_{a,b\in X}$ by taking $\Gamma_{a,b}$ to be the image under the quasi-isometry $\widehat{\Omega} : \mathcal{Q} \to X$ of the Niblo–Reeves path between $\hat{a}, \hat{b}$. By Definition 2.6, this is a uniform (B_0, B_0)-quasi-geodesic between a, b.

Theorem 2.17. *The family of paths $\{\Gamma_{a,b}\}_{a,b\in X}$ forms a discrete bounded quasi-geodesic bicombing on X.*

Proof. We need to confirm the properties in Definition 2.16. The paths are clearly discrete and are (B_0, B_0)-quasi-geodesics by assumption. So we need to show that they satisfy the fellow-traveling property in item (2) of that definition. This is a direct application of the properties in Definition 2.7 and Theorem 2.13.

Let $\hat{a}, \hat{b} \in \mathcal{Q}$ and $\hat{a}_0, \hat{b}_0 \in \mathcal{Q}_0$ be the distinguished vertices associated to a, b, and the same for $\hat{a}', \hat{b}' \in \mathcal{Q}'$ and $\hat{a}'_0, \hat{b}'_0 \in \mathcal{Q}'_0$ with a', b', respectively. Note that $\Delta_0(\hat{a}) = \hat{a}_0$, etc, by item (4) of Definition 2.7. Let $\Lambda, \Lambda_0, \Lambda', \Lambda'_0$ be the Niblo–Reeves paths in $\mathcal{Q}, \mathcal{Q}_0, \mathcal{Q}', \mathcal{Q}'_0$ connecting the respective pairs $(\hat{a}, \hat{b}), (\hat{a}_0, \hat{b}_0), (\hat{a}', \hat{b}'), (\hat{a}'_0, \hat{b}'_0)$ in their respective cube complexes $\mathcal{Q}, \mathcal{Q}_0, \mathcal{Q}', \mathcal{Q}'_0$. So in our notation, $\Gamma_{a,b} = \widehat{\Omega}(\Lambda)$ and $\Gamma_{a',b'} = \widehat{\Omega}'(\Lambda')$. We have the following estimate

$$\begin{aligned}
(2) \qquad d_X(\Gamma_{a,b}(i), \Gamma_{a',b'}(i)) &= d_X(\widehat{\Omega}(\Lambda(i)), \widehat{\Omega}'(\Lambda'(i))) \\
(3) &\leqslant d_X(\widehat{\Omega}_0 \circ \Delta_0(\Lambda(i)), \widehat{\Omega}'_0 \circ \Delta'_0(\Lambda'(i))) + 2B_0 \\
(4) &\leqslant d_X(\widehat{\Omega}_0(\Lambda_0(i)), \widehat{\Omega}'_0(\Lambda'_0(i))) + 4B_0 \\
(5) &\leqslant d_X(\widehat{\Omega}'_0 \circ I_0(\Lambda_0(i)), \widehat{\Omega}'_0(\Lambda'_0(i))) + 5B_0 \\
(6) &= d_X(\widehat{\Omega}'_0(\Lambda'_0(i)), \widehat{\Omega}'_0(\Lambda'_0(i))) \\
(7) &= 5B_0.
\end{aligned}$$

In the above, (3) follows from commutativity of the top and bottom triangles in diagram 1, (4) follows from Theorem 2.13, (5) follows from commutativity the middle triangle of diagram 1, and (6) follows from the fact that the cubical isomorphism $I_0 : \mathcal{Q}_0 \to \mathcal{Q}'_0$ identifies the paths Λ_0 and Λ'_0, i.e. $\Lambda_0(i) = \Lambda'_0(i)$ for each i (Proposition 2.12). This completes the proof. □

With this in hand, we can prove our bicombing result:

Proof of Corollary B. Mapping class groups and Teichmüller spaces of finite type surfaces are colorable HHSes (see Section 3 for references), so they are stably weakly locally quasi-cubical by Theorem 9.9. Hence Theorem 2.17 provides a bicombing by pushing forward the Niblo–Reeves path in an appropriate cubical model to

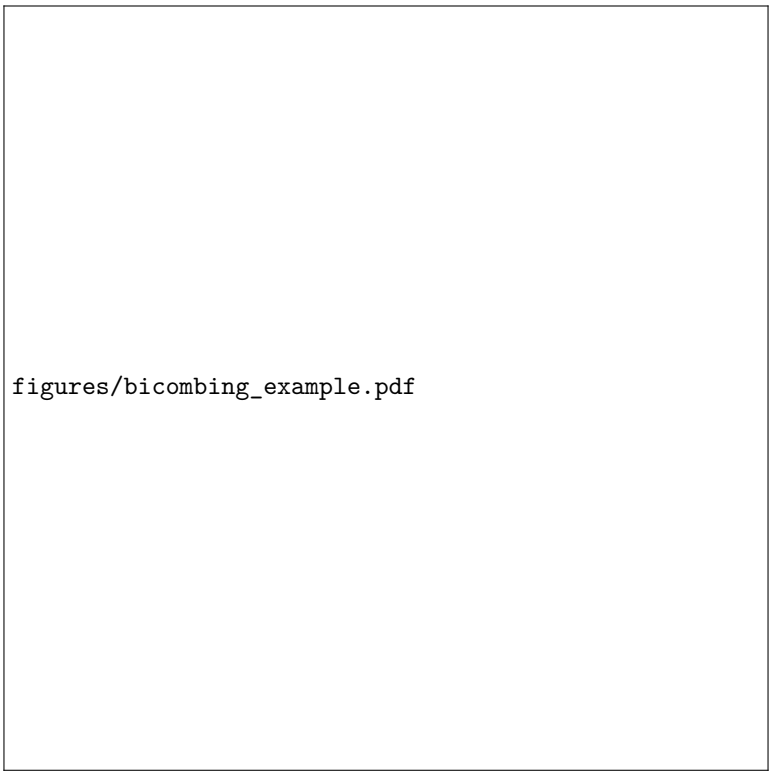


FIGURE 1. Stable Moves: A simple example of how Niblo–Reeves paths behave under hyperplane collapses. In each cube complex, the red vertices correspond to a, b. The orange hyperplanes are collapsed when passing to $\mathcal{Q}_0 = \mathcal{Q}'_0$. The Niblo–Reeves path in $\mathcal{Q}_0$ between a, b travels in a straight line between the red vertices. The Niblo–Reeves paths in $\mathcal{Q}$ (blue) and $\mathcal{Q}'$ (purple) descend to paths which are close to it the Niblo–Reeves path in $\mathcal{Q}_0$, and hence close to each other. This example is a modified version of [DMS20, Figure 25].

the respective space. Any such ℓ^∞-geodesic can be completed to an ℓ^1-geodesic by connecting vertices via ℓ^1-geodesic segments as in Proposition 2.11. But now Proposition 6.9 says that the images of these paths are hierarchy paths (Definition 3.9). This completes the proof. □

Remark 2.18. A reader familiar with both [BHS21] and [DMS20] will note that one usually uses the fact that the cubical model maps are *quasi-median* maps to get this last hierarchy path property. However, we have avoided discussions of median structures in this paper because they are unnecessary for our purposes. See [Dur23, Section 15] for how our cubical model construction works with medians.

3. A brief primer on HHSes

Hierarchically hyperbolic spaces were introduced by Behrstock–Hagen–Sisto in [BHS17b, BHS19], building off of work of Masur–Minsky [MM99, MM00b] on the mapping class group, among others [Bro02, Raf07, Raf14, Dur16, EMR17].

For our purposes, we will introduce the basic foundational concepts that we need. It turns out that the cubical model construction is a foundational-level construction, as one can use it to prove key facts like the Distance Formula 3.8; see Bowditch [Bow19] and [Dur23].

The following is a complete list (at the time of writing) of the spaces known to be hierarchically hyperbolic. With the exception of some cubical examples, all of them are colorable (Definition 7.1), and so our main results Theorem A and Corollary B apply:

- Hyperbolic spaces.
- Mapping class groups [MM99, MM00b, Aou13, BHS17b] and Teichmüller spaces, with both the Teichmüller [Raf07, Dur16] and Weil-Petersson metrics [Bro02].
- Many [HS20], but not all [She25], CAT(0) cubical groups, including all right-angled Artin and Coxeter groups [BHS17b].
- Fundamental groups of closed 3-manifolds without Nil or Sol summands [BHS19, HRSS22].
- Quotients of mapping class groups by large powers of Dehn twists [BHMS24a] and pseudo-Anosovs [BHS17a].
- Surface group extensions of Veech groups [DDLS23, DDLS24, Bon24] and multicurve stabilizers [Rus21] of mapping class groups.
- The genus-2 handlebody group [Che22].
- Artin groups of extra-large type [HMS24].
- Many free-by-cyclic groups [BGGH25].
- Admissible curve graphs [CR24].
- The class of HHSes is closed under products and appropriate combinations [BHS19], including graph products [BR22] and graph braid groups [Ber21].

We note that there are non-colorable cubical HHSes [Hag23] and non-HHS cubical groups [She25]. CAT(0) cube complexes are obviously locally modeled by cube complexes and, moreover, are bicombable [NR98], so they are not of interest in this paper in any event.

3.1. Basic setup of an HHS.

Roughly, a *hierarchically hyperbolic* structure is a pair $(\mathcal{X}, \mathfrak{S})$ where $\mathcal{X}$ is a uniformly quasigeodesic metric space and $\mathfrak{S}$ is an index set for a family of uniformly hyperbolic spaces $\{\mathcal{C}(U)\}_{U\in\mathfrak{S}}$. The ambient space $\mathcal{X}$ admits a family of uniformly coarsely Lipschitz projections $\pi_U : \mathcal{X} \to \mathcal{C}(U)$, and those projections, along with certain relations on the family $\mathfrak{S}$, largely control the coarse geometry of $\mathcal{X}$.

The following extracts some of the properties we need from the axiomatic setup [BHS19], see [Sis17] for a discussion.

(1) For each $U \in \mathfrak{S}$ there exist an $E_{\mathfrak{S}}$-hyperbolic space $\mathcal{C}(U)$ and a $E_{\mathfrak{S}}$-coarsely surjective $(E_{\mathfrak{S}}, E_{\mathfrak{S}})$-coarsely Lipschitz map $\pi_U : \mathcal{X} \to \mathcal{C}(U)$, with $E_{\mathfrak{S}}$ independent of U.
(2) The set $\mathfrak{S}$ has the following relations:

- It is equipped with a partial order $\sqsubset$ called *nesting*. If $\mathfrak{S} \neq \varnothing$, the set $\mathfrak{S}$ contains a unique $\sqsubset$-maximal element S.
- It has a symmetric, anti-reflexive relation $\perp$ called *orthogonality*. If $U \sqsubset V$, then U, V are not related via $\perp$. Furthermore, domains have *orthogonal containers*: if $U \in \mathfrak{S}$ and there is a domain orthogonal to U then there is a domain W such that $V \sqsubset W$ whenever $V \perp U$.
- For any distinct $U, V \in \mathfrak{S}$ unrelated via $\perp, \sqsubset$, then we say U, V are *transverse* and write $U \pitchfork V$.

(3) There exists an integer called the *complexity* of $\mathcal{X}$ and denoted $\xi(\mathfrak{S})$ such that whenever $U_1, U_2, \cdots U_n$ is a collection of pairwise non-transverse domains, then $n \leqslant \xi(\mathfrak{S})$.

(4) If $U \sqsubset V$ or $U \pitchfork V$, then there exists a set ρ_V^U of diameter at most $E_{\mathfrak{S}}$ in $\mathcal{C}(V)$.

(5) If $U \sqsubset V$, there is a map $\rho_U^V : \mathcal{C}(V) \to \mathcal{C}(U)$, called the *relative projection* of U to V which satisfies the following: For $x \in \mathcal{X}$ with $d_V(\pi_V(x), \rho_V^U) > E_{\mathfrak{S}}$, we have $\pi_U(x) \underset{E_{\mathfrak{S}}}{\asymp} \rho_U^V(\pi_V(x))$.

For two points $x, y \in \mathcal{X}$, it is standard to use $d_U(x, y)$ to denote $d_{\mathcal{C}(U)}(\pi_U(x), \pi_U(y))$, and similarly for subsets of $\mathcal{X}$. For a subset $A \subseteq \mathcal{X}$, we will also use $\operatorname{diam}_U(A)$ to denote the diameter of the set $\pi_U(A) \subset \mathcal{C}(U)$.

3.2. **Basic HHS facts.** We now collect some useful facts about HHSes from various places.

The following definition from [Beh06, BKMM12, BHS19], will be important for us. The idea here is that we want to combine all of the projections $\pi_U : \mathcal{X} \to \mathcal{C}(U)$ for each $U \in \mathfrak{S}$ into a global projection $\Pi : \mathcal{X} \to \prod_U \mathcal{C}(U)$. The following definition uses the relations on $\mathfrak{S}$ to constrain tuples in $\prod_U \mathcal{C}(U)$.

Definition 3.1 (Consistent tuple)**.** Let $\kappa \geqslant 0$ and let $(b_U) \in \prod_{U \in \mathfrak{S}} 2^{\mathcal{C}(U)}$ be a tuple such that for each $U \in \mathfrak{S}$, the U–coordinate b_U has diameter $\leqslant \kappa$. Then (b_U) is κ–***consistent*** if for all $V, W \in \mathfrak{S}$, we have

$$\min\{\mathbf{d}_V(b_V, \rho_V^W), \mathbf{d}_W(b_W, \rho_W^V)\} \leqslant \kappa$$

whenever $V \pitchfork W$ and

$$\min\{\mathbf{d}_W(x, \rho_W^V), \operatorname{diam}_V(b_V \cup \rho_V^W)\} \leqslant \kappa$$

whenever $V \sqsubset W$.

- We refer to the above inequalities as the *consistency inequalities*.

The following is a key fact for us. It says that the image of $\Pi : \mathcal{X} \to \prod_U \mathcal{C}(U)$ consists of uniformly consistent tuples:

Lemma 3.2. *There exists $\theta = \theta(\mathfrak{S}) > 0$ so that if $x \in \mathcal{X}$, then the tuple $(\pi_U(x)) \in \prod_{U \in \mathfrak{S}} \mathcal{C}(U)$ is θ-consistent.*

On the other hand, we would like to know whether all consistent tuples in $\prod_U \mathcal{C}(U)$ essentially come from $\mathcal{X}$. That is the content of the following Realization Theorem. In particular, says that any tuple satisfying Definition 3.1 can be coarsely realized via a point in $\mathcal{X}$. It is [BKMM12, Theorem 4.3] for mapping class groups, [EMR17, Theorem 3.2] for Teichmüller space with the Teichmüller metric, and [BHS19, Theorem 3.1] for the general HHS case:

Theorem 3.3 (Realization)**.** *For each $\kappa \geqslant 1$ there exists $\theta_e, \theta_u \geqslant 0$ such that the following holds. Let $\mathbf{b} = (b_U)_{U \in \mathfrak{S}} \in \prod_{U \in \mathfrak{S}} 2^{\mathcal{C}(U)}$ be κ-consistent.*

Then there exists $x \in X$ so that $d_U(b_U, \pi_W(x)) \leqslant \theta_e$ for all $U \in \mathfrak{S}$. Moreover, x is coarsely unique *in the sense that*

$$\operatorname{diam}_{\mathcal{X}}\{x \in \mathcal{X} | d_W(b_W, \pi_W(x)) \leqslant \theta_e \text{ for all } U \in \mathfrak{S}\} < \theta_u.$$

Remark 3.4. Notably, the proof of Theorem 3.3 in [BHS19] does not involve the hyperbolicity of the associated spaces $\{\mathcal{C}(U)\}_{U \in \mathfrak{S}}$, and actually follows from a weak form which is axiomatized into the HHS setup. In this sense, realization is a purely hierarchical fact which plays a similar role to the existence of a coarse median in Bowditch's analysis [Bow19].

Perhaps the main technique in HHSes is taking an object in the ambient space $\mathcal{X}$ and projecting it to all of the hyperbolic spaces, doing hyperbolic geometry, and then reassembling the pieces via hierarchical tools like the Realization Theorem 3.3. Toward that end, and recognizing that HHSes are fundamentally coarse spaces, we will usually want to restrict ourselves to the *relevant* collection of hyperbolic spaces to which an object has a sufficiently large projection.

Definition 3.5 (Relevant domains)**.** Given a subset $A \subset \mathcal{X}$ and $K > 0$, we say that $U \in \mathfrak{S}$ is *K-relevant* for A if $\operatorname{diam}_U(A) \geqslant K$. We denote the set of K-relevant domains for A by $\operatorname{Rel}_K(A)$.

The following Passing-Up lemma is a foundational fact about HHSes; in fact, it can be taken to be an axiom [Dur23, Proposition 4.15]. Roughly speaking, the passing-up lemma says that if $x, y \in \mathcal{X}$ have a large number of lower complexity relevant domains, then there must be some higher complexity relevant domain containing at least one of them, thereby allowing us to "pass up" the $\sqsubset$-lattice.

Lemma 3.6 (Passing-up)**.** *Let $\mathcal{X}$ be an HHS with constant $E_{\mathfrak{S}}$. For every $D > 0$ there is an integer $PU = PU(D) > 0$ such that if $V \in \mathfrak{S}$ and $x, y \in \mathcal{X}$ satisfy $d_{U_i}(x, y) > E_{\mathfrak{S}}$ for a collection of domains $\{U_i\}_{i=1}^{PU}$ with each $U_i \in \mathfrak{S}_V$, then there exists $W \in \mathfrak{S}_V$ and i with $U_i \sqsubsetneq W$ such that $d_W(x, y) > D$.*

As an example of how this lemma is foundational, we observe the following corollary. The original proof is embedded in the proof of Theorem 3.3 in [BHS19, Theorem 3.1], but not explicitly stated. We give a proof for completeness:

Corollary 3.7. *For any $x, y \in \mathcal{X}$, the set $\operatorname{Rel}_{E_{\mathfrak{S}}}(x, y)$ is finite.*

Proof. Let C be the coarse Lipschitz constant for all of the projections π_U for $U \in \mathfrak{S}$, and set $D > Cd_{\mathcal{X}}(x, y) + C$. If $\#\operatorname{Rel}_{E_{\mathfrak{S}}}(x, y) > P(D)$, then the Passing-Up Lemma 3.6 provides a domain $W \in \mathfrak{S}$ with $d_W(x, y) > D$, which is impossible. □

Another fundamental fact about HHSes is the following distance formula, which originated in work of Masur-Minsky [MM00b], and found various iterations elsewhere [Bro02,Raf07,Dur16,KK14] before being generalized to all HHSes in [BHS19]. See [BHS19, Theorem 4.5] for a proof from the now-standard HHS setup, and also [Bow19, Theorem 1.2] and [Dur23, Corollary C] from proofs using cubical model techniques, which we also recover here below in Corollary 6.11.

Theorem 3.8 (Distance formula)**.** *There exists $K_0 = K_0(\mathfrak{S}) > 0$ so that if $K > K_0$ and $x, y \in \mathcal{X}$, then*

$$d_{\mathcal{X}}(x, y) \asymp \sum_{U \in \mathfrak{S}} [d_U(x, y)]_K$$

where the coarse equality $\asymp$ depends only on $\mathfrak{S}$ and K.

Philosophically (for our purposes), the distance formula says that if we want to understand how to move between two points in an HHS, it should be enough to focus on the set of relevant domains between them.

The next definition is of a special kind of path in an HHS. The proof of their existence is always tied up in the proof of the more difficult bound of the distance formula, as is the case in this paper in Subsection 6.4. Otherwise, we will not need this notion.

Definition 3.9. A path $\gamma : I \to \mathcal{X}$ in an HHS $(\mathcal{X}, \mathfrak{S})$ is a *L-hierarchy path* if its projection $\pi_U(\gamma) \subset \mathcal{C}(U)$ is an unparameterized (L, L)-quasi-geodesic for each $U \in \mathfrak{S}$.

We end this subsection with a basic observation, which generalizes the observation that if two subsurfaces U, V are not transverse, then their boundary curves must be disjoint up to isotopy.

Lemma 3.10. *Suppose that $U, V, W \in \mathrm{Rel}_{K_0}(a, b)$ for $a, b \in \mathcal{X}$.*

1. *If $U, V \sqsubset W$ and U, V are not transverse, then $d_W(\rho^U_W, \rho^V_W) < E_{\mathfrak{S}}$.*
2. *If $U \pitchfork V$ and $W \sqsubset V$ with W, U not orthogonal, then $d_U(\rho^V_W, \rho^W_U) < E_{\mathfrak{S}}$.*

Finally, we end with a key HHS axiom, originally proven for mapping class groups in [MM00a, Theorem 3.1]:

Axiom 3.11 (Bounded geodesic image)**.** *If $U \sqsubset V$ and $\gamma \in \mathcal{C}(V)$ is a geodesic with $d_V(\rho^U_V, \gamma) > E_{\mathfrak{S}}$, then $diam(\rho^V_U(\gamma)) \leqslant E_{\mathfrak{S}}$. Furthermore, for any $x, y \in \mathcal{X}$ and any geodesic γ connecting $\pi_V(x), \pi_V(y)$, if $d_V(\rho^U_V, \gamma) > E_{\mathfrak{S}}$, then $d(\pi_U(x), \pi_U(y)) \leqslant E_{\mathfrak{S}}$.*

3.3. **Hierarchical hulls and related basic facts.** In our HHS setup, the basic object we will want to consider is the following hierarchical notion of a coarse convex hull. These *hierarchical hulls* are built out of consistent tuples of points whose coordinates are contained in the hyperbolic hulls in each domain. More specifically, given a pair of points $x, y \in \mathcal{X}$, for each, $U \in \mathfrak{S}$, we set $H_U = \mathrm{hull}_U(\pi_U(\{x, y\}))$, where hull_U denotes the hyperbolic hull in $\mathcal{C}(U)$.

Definition 3.12 (Hierarchical hull)**.** Given a constant $\theta > 0$, we define the *θ-hull* of x, y to be all $z \in \mathcal{X}$ so that $d_U(z, H_U) < \theta$ for all $U \in \mathfrak{S}$.

- We denote the θ-hull of x, y by $H_\theta = \mathrm{hull}_\theta(x, y)$.

One of the key facts about hierarchical hulls is that they are *hierarchically quasiconvex* [BHS19, Lemma 6.2], see [RST23] for more on convexity in HHSes. We do not need this definition for our purposes, but we do need the following consequence:

Lemma 3.13. *For any $\theta > 0$, there exists is a (J, J)-coarse Lipschitz retract $\mathfrak{g}_{H_\theta} : \mathcal{X} \to H_\theta$, where $J = J(\theta, \mathfrak{S}) > 0$.*

In the above lemma, the map $\mathfrak{g}_{H_\theta} : \mathcal{X} \to H_\theta$ is a hierarchical analogue of a cubical gate map. When $z \in \mathcal{X}$, the point $\mathfrak{g}_{H_\theta}(z) \in H_\theta$ is defined by coordinate-wise closest-point projection of $\pi_U(z)$ to H_U and then applying the Realization Theorem 3.3; see [BHS19, Proposition 6.3] and the surrounding text for details.

3.4. **Consequences of the Passing-up Lemma 3.6.** The following statements are useful for the sort of distance formula related hierarchical accounting that we need to do the proof of the cubical model theorem. The first says that any given relevant domain for a pair $x, y \in \mathcal{X}$ can only nest into boundedly-many other relevant domains for x, y. It first appeared in [DMS20, Lemma 2.15]; see [CRHK21, Lemma 11.13] for a different proof and [Dur23, Lemma 4.8] for a version that allows x, y to be at infinity.

Lemma 3.14 (Covering)**.** *For any HHS $(\mathcal{X}, \mathfrak{S})$, there exists $N = N(\mathfrak{S}) > 0$ so that the following holds. For any $x, y \in \mathcal{X}$ and all $U \in \mathrm{Rel}_{50E_{\mathfrak{S}}}(x, y)$, there are at most N domains $V \in \mathrm{Rel}_{50E_{\mathfrak{S}}}(x, y)$ so that $U \sqsubset V$.*

Another consequence of Lemma 3.6 is the following distance formula type statement, which is a less general version of [Dur23, Proposition 4.14]. Roughly, it says that the number of medium-sized relevant domains for a pair x, y can be accounted for (i.e., in terms of the domain projections) by large-sized domains higher up the hierarchy in a precise fashion.

Proposition 3.15 (Bounding large containers)**.** *Let $K_1 > 50E_{\mathfrak{S}}$, $K_2 \geqslant K_1 + 47E_{\mathfrak{S}}$, and $x, y \in \mathcal{X}$ with $\mathcal{U}_1 = \mathrm{Rel}_{K_1}(x, y), \mathcal{U}_2 = \mathrm{Rel}_{K_2}(x, y)$, and $\mathcal{V} = \mathcal{U}_1 - \mathcal{U}_2$. Then the following hold:*

1. *For all but $PU(K_2) - 1$ domains $V \in \mathcal{V}$, there exists $W \in \mathcal{U}_2$ with $V \sqsubset W$.*
2. *For any $W \in \mathcal{U}_2$, if $\mathcal{W} = \{V \in \mathcal{V} | V \sqsubset W\}$ then we have $\#\mathcal{W} \prec d_W(x, y)$ with constants depending only on $\mathfrak{S}$ and K_2.*

The above proposition is useful for proving the following slight generalization of the Distance Formula 3.8. It's useful for circumstances, such as in Section 5.1, where we want to make distance formula-type arguments where the summands are only coarsely equal to the standard distance formula terms. Probably the key example of this is when we want to replace $d_U(x, y)$ with $d_{T_U}(x, y)$, where each T_U is a uniform quasi-geodesic in $\mathcal{C}(U)$ between $\pi_U(x)$ and $\pi_U(y)$. This originally appeared in [Min10, Lemma 9.6] for mapping class groups, and see [BHMS24b, Theorem 2.9] (which depends on the distance formula) and [Dur23, Corollary 4.17] (which does not) for proofs in the HHS setting.

Corollary 3.16. *Let $x, y \in \mathcal{X}$. Suppose that $\sigma : \mathbb{R}_+ \to \mathbb{R}_+$ is a function satisfying*

$$\sigma(d_U(x, y)) \asymp d_U(x, y)$$

for each $U \in \mathfrak{S}$, where the constants in $\asymp$ are independent of U. Then for any $K_1 > 50E_{\mathfrak{S}}$, we have

$$\sum_{U \in \mathcal{U}} [d_U(x, y)]_{K_1} \asymp \sum_{U \in \mathcal{U}} [\sigma(d_U(x, y))]_{K_1}$$

where the constants in $\asymp$ depend only on $\mathfrak{S}$, K_1, and σ.

4. From hierarchical hulls to collapsed interval systems

In the next three sections, we will give an overview of the cubical model construction for the hierarchical hull of a pair of points in an HHS from [Dur23]. The original construction is due to Behrstock-Hagen-Sisto [BHS21], and see also Bowditch [Bow19] for a version in the more general setting of coarse median spaces satisfying HHS-like axioms.

The main goal of this section is to explain how to take a pair of points $a, b \in \mathcal{X}$ in an HHS $(\mathcal{X}, \mathfrak{S})$ and produce a family of intervals which encode the relevant hierarchical data in a fairly efficient manner. Roughly, the result of this process is a collection of intervals $\{\widehat{T}_U\}_{U\in\mathcal{U}}$ equipped with a family of relative projections satisfying certain HHS-like properties, where the index set $\mathcal{U} = \mathrm{Rel}_K(a, b)$ is the set of K-relevant domains for a, b (Definition 3.5). For any such collection of intervals, one can define a notion of 0-consistency (an exact form of Definition 3.1), which defines a special 0-consistent subset $\mathcal{Q}$ of the product $\prod_{U\in\mathcal{U}} \widehat{T}_U$. In Section 5 below we will explain why this set $\mathcal{Q}$ is quasi-isometric to the hierarchical hull of a, b in $\mathcal{X}$. Finally in Section 6, we will axiomatize the properties any such family of intervals satisfies, explain how it admits a natural wallspace structure, and show that the associated 0-consistent set $\mathcal{Q}$ is isometric to the cubical dual of this wallspace.

4.1. **Basic setup of the section.** For the rest of this section, we fix a constant $K \geqslant K_0$, where $K_0 = K_0(\mathfrak{S}) > 0$ is the constant from the Distance Formula 3.8. In this section, we will be able to allow K to be as large as necessary.

Fix also $a, b \in \mathcal{X}$. Let $\mathcal{U} = \mathrm{Rel}_K(a, b) = \{U \in \mathfrak{S} | d_U(a, b) > K\}$ denote their K-relevant set. To each $U \in \mathcal{U}$, let $\phi_U : T_U \to \mathcal{C}(U)$ denote a (C, C)-quasi-isometric embedding of an interval between a_U and b_U, so that $\phi_U(a_U) = \pi_U(a)$ and $\phi_U(b_U) = \pi_U(b)$. Here we can assume that $C = C(\mathfrak{S}) > 0$ depends only on the ambient HHS. Note that we are only taking a quasi-isometric embedding, instead of a geodesic, because that is the level of generality that we will require later.

We also set some terminology for a_U, b_U, since they will play a distinguished role below.

Definition 4.1 (Marked points on intervals)**.** For each $U \in \mathcal{U}$ the points $a_U, b_U \in T_U$ are called *marked points*.

We note that in the basic construction explained in this section, marked points will always been the endpoints of the intervals T_U. This will also be true of our collapsed stable trees defined in Subsection 9.2, though not necessarily of the stable intervals themselves.

The discussion in this section closely follows [Dur23, Sections 6 and 7]. Given the technical nature of this part, we will mostly state definitions and important properties, providing sketches where convenient but leaving the details for the reader to verify.

4.2. **Reduced interval systems.** Our first goal is to push the hierarchical data associated to a, b onto the family of intervals $\{T_U\}_{U\in\mathcal{U}}$. This involves producing a notion of relative projection between the intervals and verifying some HHS-like properties for them.

These *interval projections* are defined simply by closest-point projecting the relative projections of $(\mathcal{X}, \mathfrak{S})$ to the images of $\phi_U(T_U) \subset \mathcal{C}(U)$ then pulling back to T_U via ϕ_U^{-1}. We write δ_U^V, rather than ρ_U^V, for the relative interval projections among the T_U. These interval projections are interval versions of the "tree projections" described in [Dur23, Section 6].

The following lemma, which summarizes the basic properties that we need, says that the δ_U^V behave in essentially the same way as the ρ_U^V. Its proof is an exercise in the definitions and the basic properties of an HHS; see [Dur23, Lemma 6.13].

Lemma 4.2. *For any $C > 0$ there exists $R_C = R_C(C, \mathfrak{S}) > 0$ so that the following holds:*

1. *If $U \pitchfork V \in \mathcal{U}$, then both $\delta_U^V \subset T_U$ and $\delta_V^U \subset T_V$ are sets of diameter at most R_C which are R_C-close to a marked point of T_U, T_V, respectively. Moreover, $d_U(\phi_U(\delta_U^V), \rho_U^V) < R_C$, and similarly for δ_V^U.*
2. *If $U \sqsubset V \in \mathcal{U}$, then $\delta_V^U \subset T_V$ has diameter bounded by R_C. Moreover, $d_V(\phi_V(\delta_V^U), \rho_V^U) < R_C$.*
3. *If $U, V \sqsubset W \in \mathcal{U}$ with U, V not transverse, then $\operatorname{diam}_{T_W}(\delta_W^U \cup \delta_W^V) < R_C$.*
4. *If $U \pitchfork V$ and $W \sqsubset V$ and W, U not orthogonal, then $\operatorname{diam}_{T_U}(\delta_U^V \cup \delta_U^W) < R_C$.*
5. *(Bounded geodesic image) If $U \sqsubset V \in \mathcal{U}$, then any component $C \subset T_V - \mathcal{N}_{R_C}(\delta_V^U)$ satisfies $\operatorname{diam}_{T_U}(\delta_U^V(C)) < R_C$.*

It will be useful for us to abstract the properties obtained for the intervals T_U in the last subsection. We are doing this so that we can later plug in the intervals produced by the Stable Interval Theorem 8.14 below directly into this setup.

Definition 4.3 (Interval system)**.** Given $R > 0$, an *R-interval system* for $a, b \in \mathcal{X}$ and $\mathcal{U} = \operatorname{Rel}_K(a, b)$ is the following set of data:

1. For each $U \in \mathcal{U}$, we have an interval T_U and a (R, R)-quasi-isometric embedding $\phi_U : T_U \to \mathcal{C}(U)$ so that $d_U^{Haus}(\phi_U(T_U), \operatorname{hull}_U(a, b)) < R$.
2. For each $U \in \mathcal{U}$, there exists a marked point $a_U \in T_U$ with $d_{\mathcal{C}(U)}(\phi_U(a_U), \pi_U(a)) < R$, and similarly for b.
3. There are δ_U^V sets which satisfy the conclusions of Lemma 4.2.

Remark 4.4. When stabilizing the construction of the cubical models in Section 9, we will want to use the stable intervals provided by Theorem 8.14 below. For these, we will induce an interval system structure by a slightly more careful construction than just working directly with Gromov trees, e.g. item (2) in Definition 4.3 is immediate in that setting. But the upshot is the same: the rest of the material in this section also holds for that setup.

4.3. **Interval consistency.** The goal of this subsection is to define a map $\Psi : \operatorname{hull}_{\mathcal{X}}(a, b) \to \prod_{U \in \mathcal{U}} T_U$ and show that its image satisfies the appropriate analogue of Definition 3.1 in this setting. Toward that end, for any $\alpha > 0$, we set $\mathcal{Z}_\alpha \subset \prod_{U \in \mathcal{U}} T_U$ to be the set of α-consistent tuples.

We can now define our map $\Psi : H \to \prod_{U \in \mathcal{U}} T_U$ for $H = \operatorname{hull}_{\mathcal{X}}(a, b)$. Let $x \in H$ and for each $U \in \mathcal{U}$, let x_U denote a choice of a point in $\psi_U(x) = \phi_U^{-1}(p_U(\pi_U(x))) \subset T_U$, where $p_U : \mathcal{C}(U) \to \phi_U(T_U)$ is closest point projection. For each $x \in H$, this defines a tuple $(\psi_U(x)) = (x_U) \in \prod_{U \in \mathcal{U}} T_U$. That is, the coordinate-wise maps $\psi_U : H \to T_U$ combine to a global map $\Psi : H \to \prod_U T_U$.

Lemma 4.5. *For any $R > 0$, there exists $\alpha_0 = \alpha_0(\mathfrak{S}, R) > 0$ so that if $\{T_U\}_{U \in \mathcal{U}}$ is an R-interval system for $a, b \in \mathcal{X}$ and $x \in \operatorname{hull}_{\mathcal{X}}(a, b)$, then $\Psi(x) = (x_U)$ is α_0-consistent.*

Proof. Let $x \in \operatorname{hull}_{\mathcal{X}}(a, b)$. The corresponding tuple $(\pi_U(x_U))$ is $O(\theta, \mathfrak{S})$-consistent by Lemma 3.2. Moreover, by definition of $H = \operatorname{hull}_{\mathcal{X}}(a, b)$ (Definition 3.12), we have $d_U(\pi_U(x_U), H_U) < \theta$, where $H_U = \operatorname{hull}_{\mathcal{C}(U)}(a, b)$ and $\theta = \theta(\mathfrak{S}) > 0$. Hence $d_U(\pi_U(x_U), \phi_U(T_U)) < \theta'$ for some $\theta' = \theta'(\mathfrak{S}, \theta) > 0$. At this point, a straightforward argument using the properties in Lemma 4.2 gives the required consistency constant $\alpha_0 = \alpha_0(\mathfrak{S}, R)$. We leave the details to the reader. □

4.4. **Interval thickenings.** Our next goal is to define a way of collapsing the relative projection data associated to intervals T_U, i.e. the δ_U^V, in an interval system to remove coarseness from the setup. The idea is that Definition 4.3 roughly encodes the projection data up to a controlled error. In slightly more detail, observe that each interval T_U contains the relative projections δ_U^V for $V \sqsubset U$. We would like to collect close-by relative projections into single intervals called *cluster components*, so that we can later collapse them. One complication is that the δ_U^V need not be subintervals of T_U, so we need a couple of constants—one to determine what it means for relative projections to be "close", and the other to guarantee that cluster components which are not close satisfy a minimum distance bound.

We will use the following slightly more abstract setup, as it will be useful later during our stabilization process:

Definition 4.6 (Thickenings)**.** Let T be an interval with a decomposition $T = A \cup B$ into collections of segments. Given $r_1, r_2 > 0$, the (r_1, r_2)-*thickening* of T *along* B defines a decomposition $T = \mathbb{T}_e \cup \mathbb{T}_c$ into *edge* $\mathbb{T}_e$ and *cluster* $\mathbb{T}_c$ components as follows:

(1) First, take the r_1-neighborhoods in T of the components of B. Call the collection of these neighborhood $\mathcal{B}$, and their complement $T - \mathcal{B} = \mathcal{A}$.
(2) Second, connect any two subintervals in $\mathcal{B}$ which come within r_2 of each other in T by the geodesic between them. Call the resulting collection of subintervals $\mathbb{T}_c$ and set $\mathbb{T}_e = T - \mathbb{T}_c$.

In our current context, we can take B to be the collection of $\delta_U^V \subset T_U$ segments for $V \sqsubset U$ along with the marked points a_U, b_U on T_U (Definition 4.1). Thus a thickening of T_U along this collection B expands the δ_U^V segments and a_U, b_U and combines nearby expansions. For each $U \in \mathcal{U}$, we denote the resulting decomposition of the thickened interval by $T_U = \mathbb{T}_U^e \cup \mathbb{T}_U^c$. When we perform an (r_1, r_2)-thickening to each interval in an interval system, we refer to the resulting collection of thickened intervals as a (r_1, r_2)-*thickening* of the interval system.

Remark 4.7. In our context, we will be able to control the constants $r_1, r_2 > 0$ in terms of the ambient HHS, and, importantly, we will be able to choose our largeness constant K to be independent of them.

4.5. **Collapsed intervals and 0-consistency.** Let $\{T_U\}_{U \in \mathcal{U}}$ be an R-interval system for $a, b \in \mathcal{X}$. Fix $r_i = r_i(R, \mathfrak{S}) > 0$ for $i = 1, 2$ to be determined soon. For each $U \in \mathcal{U}$, let $T_U = \mathbb{T}_U^e \cup \mathbb{T}_U^c$ denote its (r_1, r_2)-thickening, as in Definition 4.6.

Let $q_U : T_U \to \widehat{T}_U$ denote the map which collapses each component of $\mathbb{T}_U^c$ to a point, each of which we call a *collapsed point*. The resulting object $\widehat{T}_U$ is an interval. Note that by construction $a_U, b_U \in \mathbb{T}_U^c$ for each $U \in \mathcal{U}$. Hence each of a_U, b_U determines a *marked point* $\hat{a}_U, \hat{b}_U \in \widehat{T}_U$, and these marked points $\hat{a}_U, \hat{b}_U$ will be the endpoints of the (collapsed) interval $\widehat{T}_U$. Moreover, the decomposition $T_U = \mathbb{T}_U^e \cup \mathbb{T}_U^c$ descends to a decomposition $\widehat{T}_U = \widehat{T}_U^e \cup \widehat{T}_U^c$, where the components of $\widehat{T}_U^c$ are the collapsed points corresponding to the components of $\mathbb{T}_U^c$, and the components of $\widehat{T}_U^e$ are isometric images of the components of $\mathbb{T}_U^e$.

Combining these q_U maps with our coordinate-wise projections $\psi_U : H \to T_U$, we get coordinate-wise maps $\hat{\psi}_U : H \to \widehat{T}_U$ to the collapsed intervals, and hence combining these gives a global map $\widehat{\Psi} : H \to \prod_U \widehat{T}_U$.

We can define *collapsed relative projections* $\hat{\delta}_U^V$ in the obvious way, by composing with the quotient maps q_U. Similarly, consistency in this setting is the appropriate analogue of Definition 3.1, except now all coarseness has been removed:

Definition 4.8. A tuple $(\hat{x}_U) \in \prod_{U \in \mathcal{U}} \widehat{T}_U$ is 0-*consistent* if the following hold:

(1) For any $U \pitchfork V \in \mathcal{U}$, either $\hat{x}_U = \hat{\delta}_U^V$ or $\hat{x}_V = \hat{\delta}_V^U$.
(2) For any $U \sqsubset V \in \mathcal{U}$, either $\hat{x}_V = \hat{\delta}_V^U$ or $\hat{x}_U \in \hat{\delta}_U^V(\hat{x}_V) \subset \widehat{T}_U$.

- We let $\mathcal{Q}$ denote set of 0-consistent tuples in $\prod_{U \in \mathcal{U}} \widehat{T}_U$.

We can prove that $\widehat{\Psi}$ has image in $\mathcal{Q}$:

Proposition 4.9. *For any $R > 0$, there exists $R_1 = R_1(R, \mathfrak{S}) > 0$ so that for any $r_1 > R_1$ and $r_2 > 0$, if $\{\widehat{T}_U\}_{U \in \mathcal{U}}$ are the collapsed intervals associated to an (r_1, r_2)-thickening of an R-interval system for $a, b \in \mathcal{X}$ and $x \in \mathrm{hull}_{\mathcal{X}}(a, b)$, then $\widehat{\Psi}(x) = (\hat{x}_U) \in \mathcal{Q}$.*

Proof. We need to prove 0-consistency of $\widehat{\Psi}(x)$ as in Definition 4.8.

First suppose that $U \pitchfork V \in \mathcal{U}$. Since $\Psi(x) = (x_V)$ is α_0-consistent by Lemma 4.5 where $\alpha_0 = \alpha_0(R, \theta, \mathfrak{S}) > 0$, we have, without loss of generality, that $d_{T_U}(x_U, \delta_U^V) < \alpha_0$. Hence $q_U(x_U) = \hat{\delta}_U^V$ as long as $r_1 > \alpha_0$.

Next suppose that $U \sqsubset V \in \mathcal{U}$. If $d_{T_V}(x_V, \delta_V^U) < \alpha_0 < r_1$, then $q_V(x_V) = \hat{\delta}_V^U$, and we are done. Otherwise, we have $\mathrm{diam}_{T_U}(x_U \cup \delta_U^V(x_V)) < \alpha_0$ while also $d_{T_V}(x_V, \delta_V^U) > \alpha_0$. By taking $\alpha_0 > R$, we have that $\mathrm{diam}_{T_U}(\delta_U^V(x_V) \cup f_U) < R$ for some $f \in \{a, b\}$ by the BGI property in item (4) of Definition 4.3. In particular, setting $r_1 > \alpha_0 + R$, then both x_U and $\delta_U^V(x_V)$ are within $\alpha_0 + R < r_1$ of f_U, and hence $q_U(x_U) = q_V(\delta_U^V(x_V))$. Finally, observe that $q_V^{-1}(q_V(x_V))$ is contained in the component of $T_V - \mathcal{N}_R(\delta_V^U)$ containing x_V, and hence $q_V(\delta_U^V(x_V)) = \hat{\delta}_U^V(q_V(x_V))$. This completes the proof. □

4.6. **Controlling the collapsed intervals.** In this subsection, we record the following technical lemma, which is in part a collapsed version of Lemma 4.2, but also contains the conclusion of Proposition 4.9. The statement and its proof are basically [Dur23, Lemma 7.13].

Lemma 4.10 (Collapsed interval control)**.** *Let $R > 0$ and $r_1 > R_1$ for $R_1 = R_1(R, \theta, \mathfrak{S}) > 0$ as in Proposition 4.9. There exists $R_2 = R_2(R, \mathfrak{S}) > 0$ so that for any $r_2 > R_2$, the following hold for any (r_1, r_2)-thickening of an R-interval system for $a, b \in \mathcal{X}$:*

(1) *For any $x \in H = \mathrm{hull}_{\mathcal{X}}(a, b)$, the tuple $\widehat{\Psi}(x) \in \mathcal{Q}$ is 0-consistent.*
(2) *If $U \pitchfork V \in \mathcal{U}$, then both $\hat{\delta}_U^V \in \widehat{T}_U$ and $\hat{\delta}_V^U \in \widehat{T}_V$ are in $\{\hat{a}_U, \hat{b}_U\}$.*
(3) *If $U \sqsubset V \in \mathcal{U}$, then $\hat{\delta}_V^U$ coincides with a collapsed point.*
(4) *For any $U \in \mathcal{U}$ and any $x, y \in H$, we have $d_U(x, y) \asymp d_{T_U}(x_u, y_U) \geqslant d_{\widehat{T}_U}(\hat{x}_U, \hat{y}_U)$, with the constants in $\asymp$ depending only on $\mathfrak{S}$ and R.*
(5) *If $x, y \in H$ and $U \in \mathcal{U}$ and $\hat{\psi}_U(x), \hat{\psi}_U(y)$ are distinct collapsed points, then $d_U(x, y) > 50E_{\mathfrak{S}}$.*
(6) *If $U \in \mathcal{U}$ is $\sqsubset_{\mathcal{U}}$-minimal, then $d_{\widehat{T}_U}(\hat{a}_U, \hat{b}_U) > K - 2r_1$, with the constants in $>$ depending only on $\mathfrak{S}$ and where $K = K(\mathfrak{S}) > 0$ is chosen independent of the setup.*

(7) *(Bounded geodesic image) Suppose* $V \sqsubset U$, $C \subset \widehat{T}_U - \hat{\delta}_U^V$ *is a component, then* $\hat{\delta}_V^U(C)$ *is a point. In particular, if* $f \in \{a, b\}$ *with* $\hat{f}_U$ *contained in* C*, then* $\hat{\delta}_V^U(C) = \hat{f}_V$.

(8) *If* $\hat{x} \in \mathcal{Q}$*, then* $\{U \in \mathcal{U} | \hat{x}_U \notin \{\hat{a}_U, \hat{b}_U\}\}$ *is pairwise non-transverse, and hence has cardinality bounded in terms of* $\mathfrak{S}$.

(9) *If* $\hat{x} \in \mathcal{Q}$*, then* $E(\hat{x}) = \{U \in \mathcal{U} | \hat{x}_U \in \widehat{T}_U^e\}$ *is pairwise orthogonal.*

Remark 4.11. Item (5) is notably the first time that the constant r_2 from Definition 4.6 plays a substantial role. This is important later in proving that the map $\widehat{\Psi} : H \to \mathcal{Q}$ between 0-consistent set $\mathcal{Q}$ and the hierarchical hull $H = \mathrm{hull}_{\mathcal{X}}(a, b)$ is a quasi-isometry. Another notable statement is item (6), which is about $\sqsubset_{\mathcal{U}}$-minimal domains, namely domains U for which no $V \in \mathcal{U}$ has $V \sqsubset U$. Here is where our largeness constant K plays a crucial role. In particular, since nothing nests into a $\sqsubset_{\mathcal{U}}$-minimal domain U, the map $q_U : T_U \to \widehat{T}_U$ only collapses the r_1-neighborhood of each endpoint, and hence the diameter of $\widehat{T}_U$ is coarsely bounded below by K. This is also important for confirming that $\widehat{\Psi} : H \to \mathcal{Q}$ is a quasi-isometry; see Section 5 below.

Proof. Item (1) is Proposition 4.9. Item (2)follows from item (2) of Definition 4.3 by choosing $r_1 > 2R$. Similarly, item (3) follows from Definition 4.3 (namely, item (1) of Lemma 4.2) by choosing $r_1 > 2R$.

Item (4) follows directly from the construction and item (3) of Definition 4.3.

Item (4) is an immediate consequence of the definition of the hull H (Definition 3.12) and the construction in this subsection. In particular, by assumption $\phi_U : T_U \to \mathcal{C}(U)$ is an (R, R)-quasi-geodesic between $\pi_U(a), \pi_U(b)$, giving the coarse equality, while the inequality follows because $q_U : T_U \to \widehat{T}_U$ collapses subintervals and therefore is distance non-increasing.

For item (5), if $\hat{\psi}_U(x), \hat{\psi}_U(y)$ are in distinct clusters, then $d_{T_U}(x, y) > r_2$, and hence $d_U(x, y) > r_2$ by item (4). Hence choosing r_2 large enough gives the desired lower bound.

For item (6): When $U \in \mathcal{U}$ is $\sqsubset_{\mathcal{U}}$-minimal with respect to $\mathcal{U}$, then $q_U : T_U \to \widehat{T}_U$ only collapses the r_1-neighborhoods of a_U, b_U in T_U. Since $d_U(a, b) > K$ by assumption, it follows that $d_{T_U}(a_U, b_U) > K$, while the map $q_U : T_U \to \widehat{T}_U$ only collapses the r_1-neighborhoods of a_U, b_U. Thus provided that $K - 2r_1 > r_2$ (which we can arrange), the claim follows.

For item (7): By item (1), if $\hat{f}_U$ is a marked point, then 0-consistency of $(\hat{f}_V) = \widehat{\Psi}(f)$ implies that $\hat{\delta}_V^U(\hat{f}_U) = \hat{f}_V$ because $d_{T_U}(\delta_U^V(f_U), f_V) < \alpha_0 < r_1$, where we note that $\alpha_0 = \alpha_0(R, \mathfrak{S}) > 0$ by Lemma 4.5. On the other hand, since $C \cap \hat{\delta}_U^V = \varnothing$, we have $d_{T_U}(q_U^{-1}(C), \delta_U^V) > R$. Hence item (4) of Definition 4.3 says that $\mathrm{diam}_{T_V}(\delta_V^U(C)) < R$, while also $\delta_V^U(f_U) \subset \delta_V^V(C)$. Hence $d_{T_V}(\delta_V^U(C), f_V) < \alpha_0 + R < r_1$, where we guarantee the last inequality by increasing r_1 as necessary. It follows then that $\hat{\delta}_V^U(C) = \hat{f}_V$, as claimed.

For item (8): Suppose U, V are transverse and $\hat{x}_U \notin \{\hat{a}_U, \hat{b}_U\}$. Since $\hat{\delta}_U^V \in \{\hat{a}_U, \hat{b}_U\}$ by item (2) and hence $\hat{x}_U \neq \hat{\delta}_U^V$, it follows from 0-consistency of $\hat{x}$ that $\hat{x}_V = \hat{\delta}_V^U$. Thus $\hat{x}_V \in \{\hat{a}_V, \hat{b}_V\}$ and so the set in the statement of item (8) is pairwise non-transverse, and hence has bounded complexity by item (2) in Subsection 3.1.

Finally, for item (9): Recall that $\widehat{T}^e_U = \widehat{T}_U - \widehat{T}^c_U$, where $\widehat{T}^c_U$ are the collapsed points. Hence $E(\hat{x})$ is the set of domains where $\hat{x}_U$ is not a collapsed point. Supposing that U, V are neither orthogonal nor (by item (8)) transverse. When $V \sqsubset U$ and $\hat{x}_U = \hat{\delta}^V_U$, we are done by items (1) and (3). When $V \sqsubset U$ and $\hat{x}_U \neq \hat{\delta}^V_U$, item (7) implies there is some $f \in \{a, b\}$ so that $\hat{\delta}^U_V(\hat{x}_U) = \hat{f}_V$, while 0-consistency says that $\hat{x}_V = \hat{\delta}^U_V(\hat{x}_U)$, giving $\hat{x}_V = \hat{f}_V$. Hence if U and V are not orthogonal, then one of $\hat{x}_U$ or $\hat{x}_V$ must be a collapsed point, proving this item. This completes the proof.

□

5. $\widehat{\Psi} : H \to \mathcal{Q}$ is a quasi-isometry

In this section, we explain why the map $\widehat{\Psi} : H \to \mathcal{Q}$ (Proposition 4.9) between the hierarchical hull $H = \mathrm{hull}_{\mathcal{X}}(a, b)$ of a pair of points $a, b \in \mathcal{X}$ and the 0-consistent set associated to any sufficiently wide (r_1, r_2)-thickening of an R-interval system for a, b is a quasi-isometry. There are two main parts to this:

(1) Proving that $\widehat{\Psi} : H \to \mathcal{Q}$ is a quasi-isometric embedding, and
(2) Proving that $\widehat{\Psi}$ is coarsely surjective. This is accomplished by defining a coarse inverse $\widehat{\Omega} : \mathcal{Q} \to H$, so that $\widehat{\Psi} \circ \widehat{\Omega}$ is coarsely the identity on $\mathcal{Q}$.

As we shall see, these two results have fairly different flavors—(1) is essentially purely hierarchical, while (2) has distinctly cubical elements. Item (1) essentially says that one can encode the distance formula in $\mathcal{X}$ as the ℓ^1-distance in a product of intervals. Its proof involves mostly distance formula-type arguments and we present it in full. On the other hand, the proof of item (2), which is contained in [Dur23, Sections 9–11], is rather involved and many of the details of the argument are not crucial to our expository purposes in this article. Thus, in this section, we will mostly explain the basics of the construction, including the points where our simpler setting simplifies things considerably.

The one exception here is the Tree Trimming Theorem 5.14, which is a distance-formula like tool that we need both here and for our stability results in Section 9. We will discuss it in detail and present its proof in Appendix A.

For the rest of this section, fix $a, b \in \mathcal{X}$ and let $H = \mathrm{hull}_{\mathcal{X}}(a, b)$ to be their hierarchical hull. Fix a largeness constant K, which we will be able to control relative to all other involved constants, and set $\mathcal{U} = \mathrm{Rel}_K(a, b)$. Finally, fix $R > 0$ and an R-interval system $\{T_U\}_{U \in \mathcal{U}}$ for a, b. Let $\{\widehat{T}_U\}_{U \in \mathcal{U}}$ denote the family of collapsed intervals corresponding to an (r_1, r_2)-thickening, where $r_i = r_i(R, \theta, \mathfrak{S}) > 0$ are provided by Proposition 4.9 and Lemma 4.10.

5.1. $\widehat{\Psi} : H \to \mathcal{Q}$ is a quasi-isometric embedding.

The content of this subsection is roughly contained in [Dur23, Section 8].

Before we state the result, we set some additional notation: For numbers A, B, the notation $[A]_B$ means 0 when $A < B$ and A when $A \geqslant B$. For $x \in H$ and $U \in \mathcal{U}$, we will write

- $x = \pi_U(x) \in \mathcal{C}(U)$;
- $x_U = \psi_U(x) \in T_U$;
- $\hat{x}_U = \hat{\psi}_U(x) \in \widehat{T}_U$.

Theorem 5.1. *There exists $K_1 = K_1(\mathfrak{S}) > 0$ so that for any $K_2 \geqslant K_1$ and any $x, y \in H$, we have that*

$$d_{\mathcal{X}}(a,b) \asymp \sum_{U \in \mathfrak{S}} [d_U(x,y)]_{K_2} \asymp \sum_{U \in \mathcal{U}} d_{\widehat{T}_U}(\hat{x}_U, \hat{y}_U)$$

where the constants in $\asymp$ only depend on $\mathfrak{S}$ and the choice of K_2.

We note that the left coarse equality of the above statement is simply the Distance Formula 3.8, hence the content is the right coarse equality. In fact, one can obtain the (more difficult) lower-bound in the distance formula by combining the above theorem with the construction of hierarchy paths via cubical models, as we do below in Corollary 6.11.

Before beginning the discussion of the proof properly, we make a couple of simplifying observations, which are all basically distance formula-type manipulations.

First, note that if $x, y \in H$ and $V \in \mathrm{Rel}_K(x,y)$ but $V \notin \mathcal{U}$, then $d_V(x,y) < K + \epsilon$ for some $\epsilon = \epsilon(\mathfrak{S}) > 0$. This is because $\pi_V(x), \pi_V(y) \in \mathcal{N}_\theta(\mathrm{hull}_V(a,b))$. Hence by taking $K_2 \geqslant K_1 > K + \epsilon$, we may increase the threshold to remove any extraneous domains. That is:

$$\sum_{U \in \mathfrak{S}} [d_U(x,y)]_{K_2} = \sum_{U \in \mathcal{U}} [d_U(x,y)]_{K_2}.$$

We next observe that for any $x, y \in H$, item (4) of Lemma 4.10 implies that $d_U(x,y) \asymp d_{T_U}(x_U, y_U)$, with constants depending only on $\mathfrak{S}$. Thus we can use Corollary 3.16 to replace the terms $d_U(x,y)$ with $d_{T_U}(x_U, y_U)$ in the sum, namely:

$$\sum_{U \in \mathcal{U}} [d_U(x,y)]_{K_2} \asymp \sum_{U \in \mathcal{U}} [d_{T_U}(x_U, y_U)]_{K_2}.$$

The two bounds in Theorem 5.1 are proven separately. The upper bound only requires the basic details from the construction of the $\widehat{T}_U$ and some distance formula tools. We have included the full proof of the proposition for clarity but also because it is used in proving the existence of hierarchy paths and the upper bound of the distance formula; moreover, the argument is similar to the proof of Corollary 3.16, which actually implies the proposition.

Proposition 5.2 (Upper bound)**.** *There exists $K_1 = K_1(\mathfrak{S}) > 0$ so that for any $K_2 > K_1$ and any $x, y \in H$ we have*

$$\sum_{U \in \mathcal{U}} [d_U(x,y)]_{K_2} \succ \sum_{U \in \mathcal{U}} d_{\widehat{T}_U}(\hat{x}_U, \hat{y}_U) \tag{8}$$

with the constants in $\succ$ depending only on $\mathfrak{S}$ and K_2.

Proof. As above, let $K_1 \geqslant K + \epsilon$ and set $K_2 \geqslant K_1 + 100E_{\mathfrak{S}}$. Let $\mathcal{U}_i = \mathrm{Rel}_{K_i}(x,y)$, so that $\mathcal{U}_2 \subset \mathcal{U}_1 \subset \mathcal{U}$. If $U \in \mathcal{U}$ satisfies $\hat{x}_U = \hat{y}_U$, then the corresponding term disappears on the right-hand side. Hence it suffices to account for the domains in $\mathcal{V} = \{V \in \mathcal{U} | d_V(x,y) < K_2 \text{ and } \hat{x}_U \neq \hat{y}_U\}$, which appear on the right-hand side (RHS) but not the left-hand side (LHS) of equation (8).

For this, item (9) of Lemma 4.10 says that $\hat{x}_U$ is not contained in some cluster point of $\widehat{T}_U$ for only boundedly-many U, and the same for $\hat{y}_U$. (In fact, we can take $\hat{x}_U = \hat{a}_U$ and $\hat{y}_U = \hat{b}_U$, be this will not simplify things.) This means that, up to ignoring a bounded collection of domains $\mathcal{V}' \subset \mathcal{V}$, if $V \in \mathcal{V} - \mathcal{V}'$ then $\hat{x}_V$ and $\hat{y}_V$ are at distinct cluster points, and so $d_V(x,y) > 50E_{\mathfrak{S}}$ by item (5) of Lemma 4.10, i.e., $\mathcal{V} - \mathcal{V}' \subset \mathrm{Rel}_{50E_{\mathfrak{S}}}(x,y)$, while also $d_{\widehat{T}_V}(\hat{x}_V, \hat{y}_V) \geqslant r_2 \geqslant 1$.

One finishes the proof with an application of Proposition 3.15 as follows. Set $\mathcal{U}_i = \mathrm{Rel}_{K_i}(x, y)$ and observe that $\mathcal{V} - \mathcal{V}' \subset \mathcal{U}_1 - \mathcal{U}_2$. Proposition 3.15 says that up to ignoring a subset $\mathcal{V}'' \subset \mathcal{V} - \mathcal{V}'$ with $\#\mathcal{V}'' \leqslant P(K_2)$, for each $V \in \mathcal{V}_0 = \mathcal{V} - (\mathcal{V}' \cup \mathcal{V}'')$ there exists some domain $W \in \mathcal{U}_2$ determining a term on the RHS with $V \sqsubset W$. Moreover, it says that the set of domains $\mathcal{W} = \{V \in \mathcal{V} | V \sqsubset W\}$ for which each such W accounts in this way satisfies $\#\mathcal{W} \leqslant (1 + A) \cdot d_W(x, y) + A$, for some $A = A(K_2, \mathfrak{S}) > 0$.

Combining these observations leads to the computation

$$\begin{aligned}
\sum_{U \in \mathcal{U}_2} d_U(x, y) &= \sum_{U \in \mathcal{U}_2} d_U(x, y) + \sum_{V \in \mathcal{V} - \mathcal{V}'} d_V(x, y) + \sum_{Z \in \mathcal{V}'} d_Z(x, y) \\
&\leqslant \sum_{U \in \mathcal{U}_2} ((1 + A) \cdot d_U(x, y) + A) + K_2 \cdot P(K_2) \\
&\leqslant (1 + 2A) \cdot \sum_{U \in \mathcal{U}_2} d_U(x, y) + K_2 \cdot P(K_2)
\end{aligned}$$

where the last inequality follows from the fact that $d_U(x, y) > K_2$ for each $U \in \mathcal{U}_2$. This completes the proof.

□

Proposition 5.3 (Lower bound)**.** *There exists $K_1 = K_1(\mathfrak{S}) > 0$ so that for any $K_2 > K_1$ and any $x, y \in H$ we have*

$$\sum_{U \in \mathcal{U}} [d_U(x, y)]_{K_2} \prec \sum_{U \in \mathcal{U}} d_{\hat{T}_U}(\hat{x}_U, \hat{y}_U) \tag{9}$$

with the constants in $\prec$ depending only on $\mathfrak{S}$ and K_2.

Proof. Observe that since $d_U(x, y) \prec d_{T_U}(x_U, y_U)$ for each U, it suffices by Corollary 3.16 to show that

$$\sum_{U \in \mathcal{U}} [d_{T_U}(x_U, y_U)]_{K_2} \prec \sum_{U \in \mathcal{U}} d_{\hat{T}_U}(\hat{x}_U, \hat{y}_U).$$

For this bound, there are possibly domains $U \in \mathcal{U}$ where $d_{T_U}(x_U, y_U) > K_2$ but there is a discrepancy $d_{T_U}(x_U, y_U) > d_{\hat{T}_U}(\hat{x}_U, \hat{y}_U)$. Let $\mathcal{W}$ denote the collection of such domains.

Since $\hat{T}_U$ is obtained from T_U by collapsing clusters, we need to analyze the clusters which separate x_U from y_U, i.e. those clusters $C_1, \ldots, C_n$ which intersect, in that order, the geodesic $[x_U, y_U]$ in T_U. The proof is mainly a BGIA 3.11 type argument along with an application of Proposition 3.15 and the Covering Lemma 3.14.

Continuing with the above notation, for each i, let $C_i' = C_i \cap [x_U, y_U]_{T_U} - (B_{\alpha_0 + R}(x_U) \cup B_{\alpha_0 + R}(y_U))$, where $\alpha_0 = \alpha(R, \mathfrak{S}) > 0$ is the interval consistency constant provided by Lemma 4.5; see Figure 2.

Observe that for each i, if we set $\mathcal{V}_i = \{V \in \mathcal{U} | V \sqsubset U \text{ and } \delta_U^V \subset C_i\}$ then $\{\delta_U^V | V \in \mathcal{V}_i\}$ is $(r_2 + 2r_1)$-dense in C_i. Hence we can find a subcollection $\mathcal{V}_i'$ so that

- Each $V \in \mathcal{V}_i'$ is $\sqsubset_{\mathcal{U}}$-minimal;
- For each $V \in \mathcal{V}'$, we have $\delta_U^V \subset C_i'$;
- $\{\delta_U^V | V \in \mathcal{V}_i\}$ is a $(2r_2 + 4r_1, r_2)$-net on C_i' (see Definition 5.4 below).

Note that the increase in the net constants in the third item is required to guarantee the second item. In particular, we have $\#\mathcal{V}_i > \mathrm{diam}_{T_U}(C_i') \geqslant \mathrm{diam}_{T_U}(C_i \cap [x_U, y_U]) - 2(\alpha_0 + R)$, where $\#\mathcal{V}_i$ denotes the cardinality of $\mathcal{V}_i$.

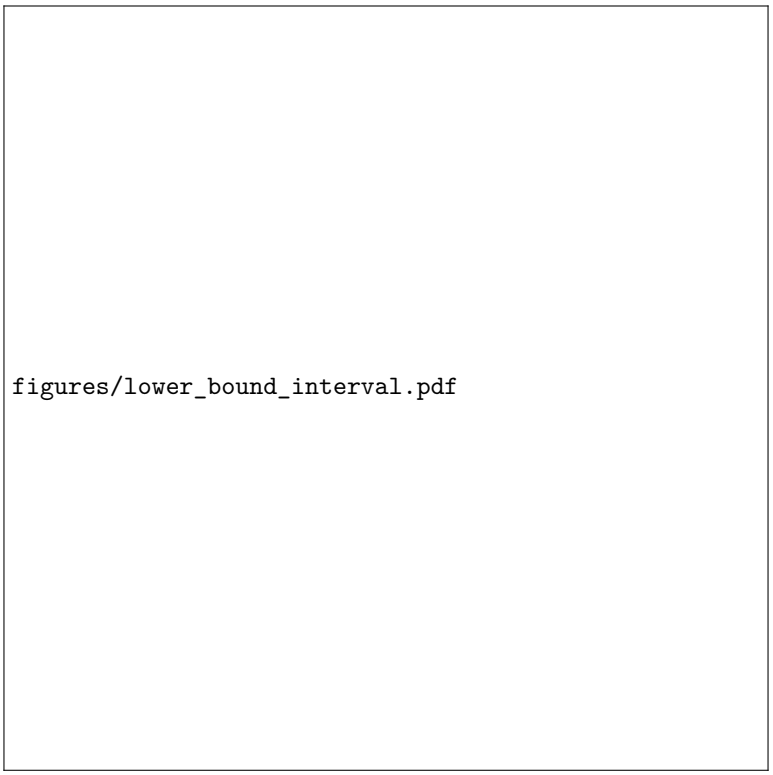


FIGURE 2. Proof of Proposition 5.3: The subcluster C' isolates the portion of C which intersects $[x_U, y_U]$ but avoids the $(\alpha_0 + R)$-neighborhoods of x_U and y_U. The (pink) net of δ-sets from $\sqsubset_{\mathcal{U}}$-minimal domains $V \sqsubset U$ with $\delta_U^V \subset C'$ coarsely encodes the length of C', with the BGIA property guaranteeing that $\hat{x}_V = \hat{a}_V$ and $\hat{y}_V = \hat{b}_V$ and hence $d_{\hat{T}_V}(\hat{x}_V, \hat{y}_V) \geqslant K - B^2 - 2r_1$ for each such V.

Let $V \in \mathcal{V}_i'$. Then, by construction, $d_{T_U}(\delta_U^V, x_U) > \alpha + R$ and $d_{T_U}(\delta_U^V, y_U) > \alpha_0 + R$. Hence, up to switching the roles of a and b, the BGI property (5) of Lemma 4.2 says that $d_{T_V}(\delta_V^U(a_V), \delta_V^U(x_V)) < R$ and $d_{T_V}(\delta_V^U(b_V), \delta_V^U(y_V)) < R$.

Lemma 4.5 says that the tuples $(a_U), (b_U), (x_U), (y_U)$ are all α_0-consistent. Thus since $d_{T_U}(a_U, \delta_U^V) > R + \alpha_0$ and similarly for b_U, x_U, y_U, α_0-consistency forces that $d_{T_V}(a_V, \delta_V^U(a_U)) < \alpha_0$, and similarly for b_V, x_V, y_V. As such, we get that $d_{T_V}(x_V, a_V) < \alpha_0 + R$ and $d_{T_V}(y_V, b_V) < \alpha_0 + R$, and so that $d_{T_V}(x_V, y_V) \geqslant d_{T_V}(a_V, b_V) - 2\alpha_0 - 2R$, from which we can conclude that $\hat{a}_V = \hat{x}_V$ and $\hat{b}_V = \hat{y}_V$ by requiring $r_1 > \alpha_0 + R$.

On the other hand, since each $V \in \mathcal{V}_i'$ is $\sqsubset_{\mathcal{U}}$-minimal and $d_V(a, b) > K$, item (4) of Lemma 4.10 says that $d_{\hat{T}_V}(\hat{a}_V, \hat{b}_V) > K$. Hence we obtain that $d_{\hat{T}_V}(\hat{x}_V, \hat{y}_V) = d_{\hat{T}_V}(\hat{a}_V, \hat{b}_V) > K$.

Thus for each $U \in \mathcal{W}$, we have

$$\begin{aligned} d_{T_U}(x_U, y_U) &= d_{\hat{T}_U}(\hat{x}_U, \hat{y}_U) + \sum_{i=1}^{n} \operatorname{diam}_{T_U}(C_i \cap [x_U, y_U]_{T_U}) \\ &\prec d_{\hat{T}_U}(\hat{x}_U, \hat{y}_U) + \sum_{i=1}^{n} \#\mathcal{V}_i' \\ &\prec d_{\hat{T}_U}(\hat{x}_U, \hat{y}_U) + \sum_{i=1}^{n} \left(\sum_{V \in \mathcal{V}_i'} d_{\hat{T}_V}(\hat{x}_V, \hat{y}_V) \right). \end{aligned}$$

Finally, the Covering Lemma 3.14 says that any domain $V \in \mathcal{U}$ can nest into only boundedly-many domains in $\mathcal{U}$, with bounds determined by $\mathfrak{S}$. Hence, each $\sqsubset_{\mathcal{U}}$-minimal domain $V \in \mathcal{U}$ must account for at most boundedly-many $U \in \mathcal{W}$ in the above way. Thus, using an argument similar to the end of the proof of Proposition 5.2, we can account for the distance lost by collapsing clusters in each domain in $\mathcal{W}$ with these large $\sqsubset_{\mathcal{U}}$-minimal domains by increasing the multiplicative constant for the estimate in equation (8) by a bounded amount. This completes the proof. □

We used the following definition in the above proof:

Definition 5.4 (Net)**.** Given constants $a, A \geqslant 0$, a (a, A)-*net* on a subspace $Y \subset X$ of a metric space is a collection $Z \subset Y$ of points so that:

(1) Z is a-coarsely dense in Y, i.e. for every $y \in Y$ there is some $z \in Z$ with $d(z, y) < a$, and
(2) For any $z, z' \in Z$, we have $d_X(z, z') \geqslant A$.

In particular, an (a, A)-net Z on Y satisfies $\#Z \asymp \operatorname{diam}(Y)$ with the constants depending only on a, A.

5.2. **The map $\widehat{\Omega} : \mathcal{Q} \to H$ and honing clusters.** With Theorem 5.1 in hand, in order to prove that $\widehat{\Psi} : H \to \mathcal{Q}$ is a quasi-isometry, we need to explain why $\widehat{\Psi}$ is coarsely surjective. The proof of this fact is quite involved and beyond the scope of an expository paper. As such, we will mostly present key ideas and not a full detailed proof because the proof is not greatly simplified versus the most general version, see [Dur23, Sections 9–11].

In this subsection, we will construct a map $\Omega : \mathcal{Q} \to \prod_{U \in \mathcal{U}} T_U$, and sketch why its image is β-consistent for some $\beta = \beta(\mathfrak{S}, R, r_1, r_2) > 0$. Upgrading this to a map $\widehat{\Omega} : \mathcal{Q} \to H$ involves standard HHS techniques. In Subsection 5.4, we will prove our Tree Trimming Theorem 5.14, which is a useful distance formula-like tool that will allow us modify our collapsed interval setup while coarsely preserving the geometry of $\mathcal{Q}$. This tool is a key ingredient in Subsection 5.6, where we sketch why $\widehat{\Omega}$ is a coarse inverse to $\widehat{\Psi}$, namely that $d_{\mathcal{Q}}(\widehat{\Psi} \circ \widehat{\Omega}(\hat{x}), \hat{x})$ is bounded independently of $\hat{x} \in \mathcal{Q}$.

We can now set about defining the map $\Omega : \mathcal{Q} \to \prod_{U \in \mathcal{U}} T_U$, which is defined coordinate-wise. To put things more precisely, we want to define Ω so that:

- For each $\hat{x} \in \mathcal{Q}$, the tuple $\Omega(\hat{x})$ is ω-consistent for some $\omega = \omega(\mathfrak{S}) > 0$, and
- when $\hat{x} = (\hat{x}_U) \in \mathcal{Q}$ and $(x_U) = \Omega(\hat{x})$, then $q_U(x_U) = \hat{x}_U$ for each $U \in \mathcal{U}$.

Unsurprisingly, the cluster subintervals, which are collapsed under each $q_U : T_U \to \widehat{T}_U$, play a key role. In particular, these cluster subintervals can have arbitrarily large diameter, and so if a coordinate $\hat{x}_U \in \widehat{T}_U$ lies in a cluster point of

$\widehat{T}_U$, then its preimage $q_U^{-1}(\hat{x}_U)$ is that whole cluster. Choosing a point within that cluster then requires utilizing the coordinates for $V \sqsubset U$ with $\hat{\delta}_U^V = \hat{x}_U$. We think of this process as *honing* the cluster.

Honing clusters: Let $\hat{x} = (\hat{x}_U) \in \mathcal{Q}$. When $U \in \mathcal{U}$ has $\hat{x}_U \in \widehat{T}_U^e$, then the definition is simple: $x_U = q_U^{-1}(\hat{x}_U)$, which is a unique point. For the other domains, we need some more setup.

Remark 5.5. The reader should be warned here that one has to deal with the case where all coordinates $\hat{x}_U$ are contained in a collapsed point. When this happens, there is not an obvious choice of collection of minimal domains to work with, and in order to determine coordinates back in the intervals T_U, one must use all of the hierarchical data at hand.

Suppose that $U \in \mathcal{U}$ and $\hat{x}_U \notin \widehat{T}_U^e$. Let C denote the cluster in T_U so that $q_U(C) = \hat{x}_U$. The first step is to associate a marked point $l_V \in \{a, b\}$ to each $\sqsubset_{\mathcal{U}}$-minimal domain $V \sqsubset U$ with $\delta_U^V \subset C$. We then use these l_V to hone the clusters of the larger $U \in \mathcal{U}$ in which such V are involved.

Suppose then that $U \in \mathcal{U}$ is $\sqsubset_{\mathcal{U}}$-minimal. Since $\widehat{T}_U$ has no internal cluster points, $\hat{x}_U$ is either a marked point—either $\hat{a}_U$ or $\hat{b}_U$—or a point in $\widehat{T}_U^e$. If the former, let l_U be the corresponding marked point—i.e. either a, b. If the latter, set $l_U = \varnothing$. We note that $\hat{a}_U \neq \hat{b}_U$ by construction when U is $\sqsubset_{\mathcal{U}}$-minimal.

Now suppose that $U \in \mathcal{U}$ is not $\sqsubset_{\mathcal{U}}$-minimal. We have that either $\hat{x}_U \in \widehat{T}_U^e$, in which case we set $l_U = \varnothing$ and $x_U = q_U^{-1}(\hat{x}_U)$. Otherwise, $\hat{x}_U$ is contained in a collapsed cluster $q_U(C)$ for the cluster $C \subset T_U$. Let $\mathcal{B}_{\min}(C)$ denote the set of $\sqsubset$-minimal domains $V \sqsubset U$ so that $\delta_U^V \subset C$. Observe that if $\mathcal{B}_{\min}(C) = \varnothing$, then $\operatorname{diam}_{T_U}(C)$ is bounded in terms of r_1, r_2 and hence in terms of $\mathfrak{S}$.

We now explain the honing process for domains in $\mathcal{B}_{\min}(C)$, which will associate to $\hat{x}_U$ a bounded diameter subset of the cluster C in T_U (Lemma 5.11).

Let $V \in \mathcal{B}_{\min}(C)$. Then we set

$$C_U^V = \operatorname{hull}_C(C \cap (\mathcal{N}_{r_2}(\delta_U^V) \cup l_V)),$$

which is a subinterval of C, see Figure 3. Finally, let $B_U(\hat{x}) = \bigcap_{V \in \mathcal{B}_{\min}(C)} C_U^V$ be the total intersection of the C_U^V.

In Lemma 5.9 below, we prove that if $V, W \in \mathcal{B}_{\min}(C)$, then $C_U^V \cap C_U^W \neq \varnothing$. Since each C_U^V is a subinterval, it follows from the Helly property for trees that $B_U(x)$ is nonempty. Finally, in Lemma 5.11 we prove that this intersection has diameter bounded in terms of $\mathfrak{S}$.

Thus we can define a tuple $\Omega(\hat{x}) = (x_U)_{U \in \mathcal{U}}$ as follows:

- For $U \in \mathcal{U}$ with $U \in E(\hat{x})$, we set $x_U = q_U^{-1}(\hat{x}_U)$.
- For $U \in \mathcal{U}$ with $U \notin E(\hat{x})$ and U $\sqsubset_{\mathcal{U}}$-minimal, we let x_U be any point in $q_U^{-1}(\hat{x}_U)$.
- Otherwise, we let $x_U \in B_U(\hat{x})$.

The following is immediate from the construction:

Lemma 5.6. *For every $\hat{x} \in \mathcal{Q}$ and $U \in \mathcal{U}$, we have $q_U(x_U) = \hat{x}_U$.*

In Proposition 5.12 below, we prove that (x_U) is a ω-consistent tuple for $\omega = \omega(\mathfrak{S}) > 0$.

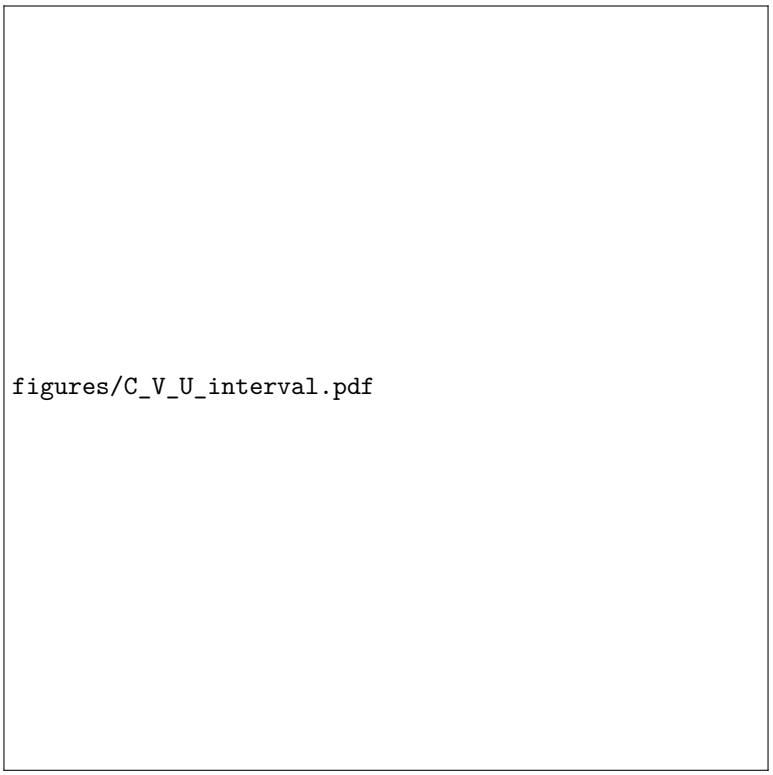


FIGURE 3. The cluster C (green) in T_U is identified with $\hat{x}_U$ in $\widehat{T}_U$. Since $V \in \mathcal{B}_{\min}(C)$, its shadow δ_U^V (orange) avoids all branches of T_U. The marked points and leaves l_V (pink) chosen by $\hat{x}_V$ determine the half-subtree C_U^V (blue) of C chosen by V.

Remark 5.7 (Cubical flavor). The process of cluster honing is one of the few places where one sees cubical-like features appearing in a non-explicitly cubical part of the construction. In particular, it is crucial that the two endpoints $\hat{a}_V, \hat{b}_V$ for every $\sqsubset_{\mathcal{U}}$-minimal V are distinct in $\widehat{T}_V$. Hence $\sqsubset$-minimal domains determine a choice of endpoint between a, b, like a wall choosing a half-space.

When U is not $\sqsubset_{\mathcal{U}}$-minimal and $\hat{\delta}_U^V = \hat{x}_U$, then the coordinate $\hat{x}_V$ either chooses neither endpoint of $\widehat{T}_V$ (if $\hat{x}_V \in \widehat{T}_V^e$), or chooses one of $\hat{a}_V, \hat{b}_V$. This choice then determines a sort of "half-interval" C_U^V of C, where we can think of δ_U^V as functioning like a hyperplane here. In Lemmas 5.9 and 5.11 below, we see that the intersection over all such "half-intervals" $B_U = \bigcap_{V \in \mathcal{B}_{\min}(C)} C_U^V$ is nonempty and has bounded diameter, and this ultra-filter-like behavior is forced by consistency of the tuple $\hat{x}$.

Remark 5.8 (A warning). The reader should be warned that one has to deal with the case of $\hat{x} \in \mathcal{Q}$ where $\hat{x}_U$ is a cluster or endpoint for every coordinate U. While, as in the previous remark, such points determine unambiguous coordinates in T_U for all $\sqsubset$-minimal domains $U \in \mathcal{U}$, there is no way to leverage this data alone into choosing coordinates in T_V for domains V higher up the $\sqsubset$-lattice. One must utilize the full set of relations between all of the domains.

FIGURE 4. The impossible situation in which $V \pitchfork W \in \mathcal{B}_{\min}(C)$ choose non-intersecting sides C_U^V (pink) and C_U^W (blue) of C (green). This forces inconsistency of $\hat{x}_V$ and $\hat{x}_W$.

5.3. **Bounding intersections.** In this subsection, we prove the supporting statements we need to define the map $\Omega : \mathcal{Q} \to \prod_{U \in \mathcal{U}} T_U$. Again fix $\hat{x} \in \mathcal{Q}$.

The first says that the C_U^V for $V \in \mathcal{B}_{\min}(C)$ have pairwise nonempty intersection:

Lemma 5.9. *If $V, W \in \mathcal{B}_{\min}(C)$, then $C_U^V \cap C_U^W \neq \varnothing$. Moreover, if $d_{T_U}(\delta_U^V, \delta_U^W) > r_2$, then at least one of $\delta_U^V \subset C_U^W$ or $\delta_U^W \subset C_U^V$ holds.*

Sketch of proof. We provide a sketch of the proof, see [Dur23, Lemma 9.9]. We shall observe item (1) during the course of the proof of the sketch.

By definition of $C_U^V = \mathrm{hull}_C((\mathcal{N}_{r_2}(\delta_U^V) \cup l_V) \cap C)$, we may assume that $\mathcal{N}_{r_2}(\delta_U^V) \cap \mathcal{N}_{r_2}(\delta_U^W) = \varnothing$, and hence that $V \pitchfork W$ by item (3) of Lemma 4.2. The idea now is that δ_U^V, δ_U^W must occur separately along the cluster C. Without loss of generality, we can assume that δ_U^V separates a_U from δ_U^W. One can then use the consistency of $(\hat{x}_U)$ and Lemma 4.2 and 4.10 to show that at least one of $l_V = b$ or $l_W = a$, which would force $C_U^V \cap C_U^W \neq \varnothing$, as required; see Figure 4.

□

As discussed above, we can define $B_U = \bigcap_{V \in \mathcal{B}_{\min}(C)} C_U^V$ and it must be nonempty by the Helly property for trees as long as $\mathcal{B}_{\min}(C) \neq \varnothing$.

Remark 5.10. The Helly property for trees says that if one has a pairwise-intersecting collection of subtrees then the collection has nonempty total intersection. In the general version of our current construction from [Dur23], the trees and subtrees in question need not be finite diameter. This is one place where the arguments in our current context are simpler than that considered in [Dur23].

The next lemma bounds the diameter of B_U:

Lemma 5.11. *There exists $B_1 = B_1(\mathfrak{S}, r_2) > 0$ such that for all $U \in \mathcal{U}$, we have* $\operatorname{diam}_{T_U}(B_U) < B_1$.

Proof. If B_U is too wide (relative to r_2), then we can find $V, W \in \mathcal{B}_{\min}(C)$ with $\delta_U^V, \delta_U^W \subset B_U$ and $r' < d_{T_U}(\delta_U^V, \delta_U^W) < 2r'$ for $r' = r'(r_2) > 0$, since the δ-sets for domains in $\mathcal{B}_{\min}(C)$ coarsely cover C with coarseness depending only on r_2. By Lemma 5.9, we must have at least one of $\delta_U^V \subset C_U^W$ or $\delta_U^W \subset C_U^V$, or possibly both. In fact, both must hold, for if $\delta_U^V \notin C_U^W$, then $\delta_U^V \notin B_U$, which is impossible. But then $C_U^V \cap C_U^W \subset \operatorname{hull}_C(\mathcal{N}_{r_2}(\delta_U^V) \cup \mathcal{N}_{r_2}(\delta_U^W))$, which has diameter bounded by $2r' + 2r_2$ since $d_{T_U}(\delta_U^V, \delta_U^W) < 2r'$. Since $B_U \subset C_U^V \cap C_U^W$, this completes the proof. □

We are now in place to state the consistency result:

Proposition 5.12. *There exist $\alpha_\omega = \alpha_\omega(\mathfrak{S}, r_1, r_2, R) > 0$ and a map $\Omega : \mathcal{Q} \to \mathcal{Y}$ so that for any $\hat{x} \in \mathcal{Q}$, the tuple $\Omega(\hat{x})$ is α_ω-consistent. Moreover, if $\Omega(\hat{x}) = (x_U)$, then $q_U(x_U) = \hat{x}_U$ for each $U \in \mathcal{U}$.*

The proof of this proposition is a somewhat elaborate exercise in the definition of Ω and the construction of the clusters, in which one utilizes many standard HHS-type arguments. Unlike other proofs in this paper, the proof of this fact in our current setting is not substantially different from proof in the general case in [Dur23, Subsection 9.5], beyond the need for working with $\sqsubset$-minimal *bipartite domains* as defined in [Dur23, Definition 9.2] instead of all $\sqsubset$-minimal domains. Thus instead of repeating it here, we leave it to the interested reader to check the details.

5.4. **Tree trimming.** In this subsection, we introduce a useful tool for modifying collapsed interval systems and their associated 0-consistent sets. These "tree-trimming" techniques will allow us to perform controlled collapses of subintervals across the various collapsed intervals $\widehat{T}_U$ while coarsely preserving the associated $\mathcal{Q}$-distance. We think of them as a more flexible version of the distance formula, as they allow us to modify distance terms instead of just throwing them out by increasing a threshold. For our purposes, they will be used in two different ways: We will need them in the proof of coarse surjectivity of $\widehat{\Psi} : H \to \mathcal{Q}$, and they will also be useful later in Subsection 6.3 to modify the cubical structure on $\mathcal{Q}$.

The main idea is simple: Given a family of subintervals $\mathcal{B}$ across of the family of intervals $\widehat{T}_U$, we want to collapse these subintervals $\Delta_U : \widehat{T}_U \to \widehat{T}_{U,0}$ coordinate-wise and then combine them into a global collapsing map $\Delta : \mathcal{Q} \to \mathcal{Q}'$, where $\mathcal{Q}'$ is the 0-consistent set of $\prod_{U \in \mathcal{U}} \widehat{T}_{U,0}$, defined the same way as in Definition 4.8. Formally, we use the following definition:

Definition 5.13 (Admissible trimming setup)**.** Given $B > 0$, a *B-admissible trimming setup* is a collection of *trimming intervals* $\mathcal{B}$ in the various $\widehat{T}_{U,e}$ satisfying the following:

- For each component $E \in \pi_0(\widehat{T}_{U,e})$, there are at most B-many components $D \in \mathcal{B}$ with $D \subset E$.
- For each component $D \in \mathcal{B}$ satisfies $|D| < B$.

Given a B-admissible trimming setup $\mathcal{B}$ on $\{\widehat{T}_U\}_{U \in \mathcal{U}}$, for each $U \in \mathcal{U}$ let $\Delta_U : \widehat{T}_U \to \widehat{T}_{U,0}$ denote the map which collapses each trimming interval in $\widehat{T}_U$ to a point. These allow us to induce projection data on the $\widehat{T}_{U,0}$ as follows: For $V \sqsubset U$ and $V \pitchfork U$, we set $(\hat{\delta}_U^V)_0 = \Delta_U(\hat{\delta}_U^V)$, and if $U \sqsubset V$, we set $(\hat{\delta}_U^V)_0 = \Delta_U \circ \hat{\delta}_U^V \circ \Delta_V^{-1}$. We let $\Delta : \prod_{\mathcal{U}} \widehat{T}_U \to \prod_{\mathcal{U}} \widehat{T}_{U,0}$ denote the product map.

Theorem 5.14 (Tree Trimming)**.** *There exists $B_{tt} = B_{tt}(\mathfrak{S}, K) > 0$ so that for any $B < B_{tt}$, there exists $L = L(\mathfrak{S}, B) > 0$ so that the following hold:*

(1) *We have $\Delta(\mathcal{Q}) \subset \mathcal{Q}'$.*
(2) *There is a map $\Xi : \mathcal{Q}' \to \mathcal{Q}$ so that $\Delta \circ \Xi = id_{\mathcal{Q}'}$. In particular, $\Delta(\mathcal{Q}) = \mathcal{Q}'$.*
(3) *The map $\Delta : \mathcal{Q} \to \mathcal{Q}'$ is an (L, L)-quasi-isometry.*

While the statement is quite believable in light of the Distance Formula 3.8, the proof of this theorem is quite involved. While these ideas were first developed in greater generality in [Dur23, Section 10]—where it also works when the underlying data involves more than two points and also hierarchy rays—the proof these is only slightly more involved than that required in our setting. Moreover, many of the technicalities from dealing with rays are dealt with in separate arguments, so the interested reader is encouraged to see a full proof there.

Instead, we will present a short sketch below, which will hopefully provide enough intuition to make the statement believable. We have included a proof of the stated version of Theorem 5.14 in Appendix A.

Sketch of proof. The fact that $\Delta(\mathcal{Q}) \subset \mathcal{Q}'$, namely that Δ preserves 0-consistency, is immediate from the fact that the Δ_U just collapse subintervals, so coordinates which were collapsed projections remain collapsed projections, showing item (1).

Item (2) is the most involved part. For each $U \in \mathcal{U}$, let $C_U = \Delta_U^{-1}(\hat{y}_U)$. The goal is to pick, for each $U \in \mathcal{U}$, a point $\hat{x}_U \in C_U$ and then prove 0-consistency of the tuple $\hat{x} = (\hat{x}_U)$. Note that $\Delta(\hat{x}) = \hat{y}$ for such a tuple, by construction.

The proof is similar in spirit to the cluster-honing procedure from Subsection 5.2. The main difference is that the collapsing subintervals are not automatically associated to a collection of domains, and this is where the work lies. The argument is inductive on level in the $\sqsubset$-lattice and requires a careful analysis of the construction and number of hierarchal techniques.

Finally, for item (3), since $\Delta|_{\mathcal{Q}}$ is surjective and distance non-increasing, it suffices to prove that $d_{\mathcal{Q}} \prec d_{\mathcal{Q}'}$, and for this it suffices to bound the distances in each $U \in \mathcal{U}$.

Given $\hat{x}, \hat{y} \in \mathcal{Q}$ with $\hat{x}' = \Delta(\hat{x})$ and $\hat{y}' = \Delta(\hat{y})$, the concern is that somehow one will have arbitrarily shrunk distances during the trimming process. However, each edge component of $\widehat{T}_U$ through which the geodesic $[\hat{x}_U, \hat{y}_U]$ passes can only contain boundedly-many subintervals collapsed. However, all but at most two of such edge components have endpoints which are cluster points labeled by $\hat{\delta}_U^V$ for V $\sqsubset$-minimal. Since for such V, the map $\Delta_V : \widehat{T}_V \to \widehat{T}'_V$ collapses subintervals of total diameter at most B^2, we can use these $\sqsubset$-minimal domains to account for all information lost up the hierarchy.

Making this precise takes care and utilizes Proposition 3.15. The interested reader can see [Dur23, Section 10]. □

5.5. **Converting to simplicial intervals.** One application of the Tree Trimming Theorem 5.14 is that we can replace the family of collapsed intervals $\widehat{T}_U$ with a family of simplicial intervals, where the cluster and marked points lie at vertices. The idea here is that one can trim each edge component $E \subset \widehat{T}_U^e$ to have a nonnegative integer length simply by collapsing some small segment of E of length less than 1. The family of such segments is an 1-admissible trimming setup, and this provides us a family of intervals $\{\widehat{T}'_U\}_{U\in\mathcal{U}}$ where $\Delta_U : \widehat{T}_U \to \widehat{T}'_U$ collapses these small segments, and $\Delta : \mathcal{Q} \to \mathcal{Q}'$ combines these to a map between their 0-consistent sets.

Corollary 5.15. *Suppose that $\{\widehat{T}_U\}_{U\in\mathcal{U}}$ are the collapsed intervals associated to an (r_1, r_2)-thickening of an R-interval system for $a, b \in \mathcal{X}$. There exists $L_{simp} = L_{simp}(R, \mathcal{X}) > 0$ so that $\Delta : \mathcal{Q} \to \mathcal{Q}'$ is an (L_{simp}, L_{simp})-quasi-isometry.*

- *In particular, up to a uniform quasi-isometry of $\mathcal{Q}$, we can assume that the collapsed intervals $\widehat{T}_U$ are all simplicial intervals with cluster and marked points at vertices.*

See [Dur23, Subsection 10.1] for more discussion. We note that the above corollary will not be used in the proof of the Main Theorem A, because we will carefully construct stable intervals to be simplicial. However, it is necessary for taking general intervals in each hyperbolic spaces and performing the above constructions to form a hierarchical family of intervals (see Definition 6.1, which is necessary for our proof of the distance formula below (Corollary 6.11).

5.6. **From $\Omega : \mathcal{Q} \to \prod_{U\in\mathcal{U}} T_U$ to $\widehat{\Omega} : \mathcal{Q} \to H$.** We return to our main task, namely defining a quasi-isometry $\widehat{\Omega} : \mathcal{Q} \to H$. So far, we have defined a map $\Omega : \mathcal{Q} \to \prod_{U\in\mathcal{U}} T_U$, which was the bulk of the work. Our next step is to upgrade this to a map $\widehat{\Omega} : \mathcal{Q} \to H$. Since this part of the argument mostly uses standard HHS techniques, we will just outline it, with appropriate citations to [Dur23].

For each $\hat{x} \in \mathcal{Q}$, we first take our tuple $\Omega(\hat{x}) = (x_U) \subset \prod_{\mathcal{U}} T_U$ and push it forward to a tuple $(y_U) \in \prod_{\mathcal{U}} \mathcal{C}(U)$ by $\phi_U(x_U) = y_U$, where $\phi_U : T_U \to \mathcal{C}(U)$ are the maps provided by the R-interval system structure (Definition 4.3). These combine to define a map $\Phi : \prod_{U\in\mathcal{U}} T_U \to \prod_{U\in\mathcal{U}} \mathcal{C}(U)$. By Proposition 5.12, there is some $\alpha_\omega = \alpha_\omega(\mathfrak{S}, R, r_1, r_2) > 0$ so that the tuple $\Omega(\hat{x}) \in \prod_{U\in\mathcal{U}} T_U$ is α_ω-consistent. Using just the R-system axioms and standard HHS techniques, [Dur23, Proposition 6.17] implies that $(y_U) \in \prod_{U\in\mathcal{U}} \mathcal{C}(U)$ is consistent in the sense of Definition 3.1 for some consistency constant depending only on $\mathfrak{S}, \alpha_\omega$.

Since $d_V(a, b) < K$ for all $V \in \mathfrak{S} - \mathcal{U}$, it follows that by choosing $y_V \in \pi_V(a)$ for all $V \in \mathfrak{S} - \mathcal{U}$, then the complete tuple $(y_U) \in \prod_{U\in\mathfrak{S}} \mathcal{C}(U)$ is β-consistent for $\beta = \beta(\mathfrak{S}, R, K) > 0$ (as α_ω only depended on $\mathfrak{S}$ and R).

Now with a complete consistent tuple in hand, we can apply the Realization Theorem 3.3 to define a map $\mathrm{Real}_\beta : \prod_{U\in\mathfrak{S}} \mathcal{C}(U) \to \mathcal{X}$ which takes each β-consistent tuple and produces a point $x \in \mathcal{X}$ whose projections coarsely coincide with the tuple, with error depending on $\beta, \mathfrak{S}$.

Finally, [Dur23, Lemma 6.6], which depends on the more general fact [BHS19, Proposition 6.3], provides a (J, J)-coarse retraction map $\mathrm{ret}_H : \mathcal{X} \to H$, where $H = \mathrm{hull}_{\mathcal{X}}(a, b)$ where $J = J(\mathfrak{S}) > 0$. This map is just closest-point projection to $\mathrm{hull}_U(a, b)$ in each $\mathcal{C}(U)$ plus Theorem 3.3.

Combining these, we get a map $\widehat{\Omega} : \mathcal{Q} \to H$ given by $\widehat{\Omega} = \mathrm{ret}_H \circ \mathrm{Real}_\beta \circ \Phi \circ \Omega$. Our main goal is to show that $\widehat{\Psi} \circ \widehat{\Omega} : \mathcal{Q} \to \mathcal{Q}$ is coarsely the identity, where $\widehat{\Psi} : H \to \mathcal{Q}$ is the map defined in Proposition 4.9.

Before analyzing this total composition, we want to look at an intermediate step. For convenience, if $\alpha > 0$, we let $\mathcal{Z}_\alpha$ denote the collection of α-consistent tuples in $\prod_{U \in \mathcal{U}} T_U$. Define

$$\Phi_\alpha : \mathcal{Z}_\alpha \to H$$

by $\mathrm{ret}_H \circ \mathrm{Real}_\beta \circ \Phi$.

The following proposition says that this map behaves coordinate-wise as one might expect upon composition with the coordinate maps $\psi_U : \mathcal{X} \to T_U$, which we recall are defined by $\psi_U = \phi_U^{-1} \circ p_U \circ \pi_U$, where $p_U : \mathcal{C}(U) \to \phi_U(T_U)$ is closest-point projection.

Proposition 5.16. *There exists $D_1 = D_1(K, \mathfrak{S}) > 0$ so that if $x = (x_V) \in \mathcal{Z}_{\alpha_\omega}$, then we have*

1. *$d_{T_U}(x_U, \psi_U \circ \Phi_{\alpha_\omega}(x)) < D_1$ for all $U \in \mathcal{U}$, and*
2. *$\#\{U \in \mathcal{U} | d_{T_U}(x_U, \psi_U \circ \Phi_{\alpha_\omega}(x)) > 2\alpha_\omega + 3R\} < D_1$.*

- *In particular, by choosing $L > 2\alpha_\omega + 3R$, we get*

$$\sum_{U \in \mathcal{U}} [d_{T_U}(x_U, \psi_U \circ \Phi_{\alpha_\omega}(x)]_L < D_1^2.$$

Sketch. This proposition is a special case of [Dur23, Proposition 11.4], and the proof is almost entirely the same. The proof of item (1) is a definition chase. It is basically saying that consistency in the interval setting is roughly consistency in the hierarchical setting, and that flowing from the former to the latter and back only moves coordinates a bounded amount.

On the other hand, the proof of (2)—which says that only boundedly-many coordinates move more than a bounded amount—is more involved, but also fairly standard. It uses a variation on the Passing-Up property from Lemma 3.6, and the BGI property from Definition 4.3. Set $\psi_U \circ \Phi_{\alpha_\omega}(x) = y = (y_U)$. The idea here is that if there are too many coordinates where $d_{T_U}(x_U, y_U) > 2\alpha_\omega + 3R$, then we can force the existence of some coordinate T_W where $d_{T_W}(x_W, y_W) > D_1$, which is impossible by (1). □

5.7. **The map $\widehat{\Omega} : \mathcal{Q} \to H$ and coarse surjectivity of $\widehat{\Psi}$.** We are ready to analyze the map $\widehat{\Omega} = \Phi_{\alpha_\omega} \circ \Omega : \mathcal{Q} \to H$. Since $\widehat{\Psi}(H) \subset \mathcal{Q}$ by Proposition 4.9, we see that $\widehat{\Psi} \circ \widehat{\Omega} : \mathcal{Q} \to \mathcal{Q}$. We prove that this map is coarsely the identity:

Proposition 5.17. *There exists $M_0 = M_0(R, \mathfrak{S}) > 0$ so that for any $\hat{x} \in \mathcal{Q}$, we have $d_{\mathcal{Q}}(\hat{x}, \widehat{\Psi} \circ \widehat{\Omega}(\hat{x})) < M_0$.*

Sketch. This proposition is a special case of the proof contained in [Dur23, Subsection 11.2]. Since the details of the proof are very similar to the original, we will again only provide a sketch.

Fix $\hat{x} \in \mathcal{Q}$ and let $\hat{y} \in \widehat{\Psi} \circ \widehat{\Omega}(\hat{x})$. We will argue coordinate-wise since that is how the ℓ^1 distance on $\mathcal{Q}$ is defined. The difference from the usual HHS style argument here is that $d_{\mathcal{Q}}$ is an ℓ^1 sum not involving a threshold. To get around that issue, we will make an argument using Proposition 5.16 and Tree Trimming 5.14 as follows. We show that all coordinate distances $d_{\widehat{T}_U}(\hat{x}_U, \hat{y}_U)$ are bounded in terms of $\mathfrak{S}$ and K, with all but boundedly-many (in terms of $\mathfrak{S}$) being bounded in terms of $\mathfrak{S}$.

We can then collapse the subintervals between $\hat{x}_U$ and $\hat{y}_U$ in the latter domains while only incurring a bounded error globally, resulting in images for $\hat{x}, \hat{y}$ which are bounded distance in the quotient, showing that $\hat{x}, \hat{y}$ were bounded distance in $\mathcal{Q}$.

Let $\Omega(\hat{x}) = (x_U)$ and $\Psi \circ \widehat{\Omega}(\hat{x}) = (y_U)$, and observe that since $q \circ \Omega(\hat{x}) = \hat{x}$, the tuples $(x_U), (y_U)$ are exactly the type in the statement of Proposition 5.16. Moreover, since $q \circ \Omega(\hat{y}) = \hat{y}$ by construction, while q is 1-Lipschitz by item (4) of Lemma 4.10, it will suffice to control the coordinate-wise distances of the $d_{T_U}(x_U, y_U)$ in order to utilize the above argument.

Let $\mathcal{V} = \{U \in \mathcal{U} | \hat{x}_U \neq \hat{y}_U\}$. One proceeds inductively on the $\sqsubset_{\mathcal{U}}$-level. Let $\mathcal{V}_{min} \subset \mathcal{V}$ denote the $\sqsubset_{\mathcal{U}}$-minimal domains in $\mathcal{V}$.

By item (8) of Lemma 4.10, for all but A-many $U \in \mathcal{V}_{min}$ with $A = A(\mathfrak{S}) > 0$, we have $\hat{x}_U, \hat{y}_U \in \{\hat{a}_U, \hat{b}_U\}$, and hence for all but at most A domains in $\mathcal{V}_{min}$, we have that $\hat{x}_U, \hat{y}_U$ are the (opposite) endpoints of $\widehat{T}_U$. By item (6) of that lemma, $d_{T_U}(a_U, b_U) > d_{\widehat{T}_U}(\hat{a}_U, \hat{b}_U) > K$, while K can be taken to be as much larger than $2\alpha_\omega + 3R$ as we need, as α_ω and R depend only on $\mathfrak{S}$. Since Proposition 5.16 bounds the number of domains $U \in \mathcal{U}$ where $d_{T_U}(x_U, y_U) > 2\alpha_\omega + 3R$, we can conclude that $\#\mathcal{V}_{min} < D_1 + A$.

Now let $\mathcal{V}' \subset \mathcal{V} - \mathcal{V}_{min}$ denote the set of non-minimal domains V where $\hat{x}_V \neq \hat{y}_V$ and $d_{T_V}(x_V, y_V) < 2\alpha_\omega + 3R$, which we note only excludes boundedly-many domains by item (2) of Proposition 5.16.

Since the maps $q_V : T_V \to \widehat{T}_V$ are 1-Lipschitz (item (4) of Lemma 4.10), we have $d_{\widehat{T}_V}(\hat{x}_V, \hat{y}_V) < 2\alpha_\omega + 3R$. The idea now is to use the Tree Trimming Theorem 5.14 to collapse the geodesics $[\hat{x}_V, \hat{y}_V]$ between $\hat{x}_V, \hat{y}_V$ in $\widehat{T}_V$ to points. Each interval $\widehat{T}_V$ contains at most one such geodesic, and each such geodesic has a length bounded independently of K. Hence the set of such intervals across both $\sqsubset$-minimal and non-minimal domains with this distance bound is an admissible trimming set in the sense required for Theorem 5.14.

If we take $\widehat{T}'_V$ to be the interval obtained from $\widehat{T}_V$ post-trimming and $\mathcal{Q}'$ the corresponding 0-consistent set, then the images $\hat{x}', \hat{y}'$ of $\hat{x}, \hat{y}$ now satisfy $\hat{x}'_V = \hat{y}'_V$ for all but boundedly-many V, and the remaining distances are all bounded by some function of D_1 by item (1) of Proposition 5.16. In particular, $d_{\mathcal{Q}'}(\hat{x}, \hat{y})$ is bounded. But the map $\Delta : \mathcal{Q} \to \mathcal{Q}'$ is a uniform quasi-isometry, so that means that $d_{\mathcal{Q}}(\hat{x}, \hat{y})$ is also bounded, which is what we wanted. We leave checking the details to the reader. □

6. Building a cube complex from a hierarchical family of intervals

The goal of this section is to explain how to take a special collection of intervals satisfying certain HHS-like properties, and to produce from them a CAT(0) cube complex. The details of this construction motivate the stabilization techniques we develop later in this article.

6.1. **Hierarchical families of intervals.**

In this subsection, we axiomatize the properties we need from a family of intervals in order to build a cube complex. The purpose of this axiomatization is to allow us to modify our construction in various ways while preserving the axioms, as any setup which satisfies the axioms produces a cube complex. The properties themselves are motivated by the HHS axioms and the properties in Lemma 4.10.

Definition 6.1 (Hierarchical family of intervals)**.** A *hierarchical family of intervals (HFI)* is the following collection of objects and properties:

(1) A finite index set $\mathcal{U}$ of *domains* with *relations* $\sqsubset, \pitchfork, \perp$ so that
   - $\sqsubset$ is anti-reflexive and anti-symmetric, and gives a partial order with a bound on the length of chains;
   - $\pitchfork$ is symmetric and anti-reflexive;
   - There is a unique $\sqsubset$-maximal element S;
   - Any pair of domains $U, V \in \mathcal{U}$ satisfies exactly one of the relations.

(2) There is a bound on the size of any pairwise $\perp$-comparable subset of $\mathcal{U}$.

(3) To each $U \in \mathcal{U}$ there is an associated finite simplicial interval $\widehat{T}_U$.

(4) A pair of labels $F = \{x, y\}$ which determined the endpoints, called *marked points* $\hat{x}_U, \hat{y}_U$, of the interval $\widehat{T}_U$ for each $U \in \mathcal{U}$.

(5) A family of *relative projections* determining:
   - For each $U, V \in \mathcal{U}$ with $V \sqsubset U$ or $V \pitchfork U$, there is a vertex $\hat{\delta}_U^V \in \widehat{T}_U^{(0)}$ called a *cluster point*.

(6) There exist projection maps $\hat{\delta}_V^U : \widehat{T}_U \to \widehat{T}_V$ when $V \sqsubset U$, satisfying the following properties:
   (a) (Consistency) For each $f \in F$, we have $\hat{\delta}_V^U(\hat{f}_U) = \hat{f}_V$.
   (b) (BGI) If $C \subset \widehat{T}_U - \hat{\delta}_U^V$ is a component, then $\hat{\delta}_V^U(C)$ coincides with $\hat{f}_V$ for $f \in F$ with $\hat{f}_U \in C$.

(7) If $U, V, W \in \mathcal{U}$ with $V, W \sqsubset U$, $V \pitchfork W$, and $\hat{\delta}_U^V \neq \hat{\delta}_U^W$, then $\hat{\delta}_V^U(\hat{\delta}_U^W) = \hat{\delta}_V^W \in \widehat{T}_V$.

Remark 6.2. In [Dur23, Definition 12.1], the definition of a *hierarchical family of trees* allows for the trees to have more than 2 leaves. This is only needed for modeling tuples of 3 or more points.

The following is an immediate consequence of our work in Section 4, especially Lemma 4.10:

Lemma 6.3. *Let $R > 0$, $r_1 > R_1 = R_1(R, \mathfrak{S}) > 0$ as in Proposition 4.9 and $r_2 > R_2 = R_2(R, \mathfrak{S}) > 0$ as in Lemma 4.10. For any (r_1, r_2)-thickening of an R-interval system for $a, b \in \mathcal{X}$, the associated family of collapsed intervals and their collapsed projections form a hierarchical family of intervals.*

As above in Definition 4.8, we can define a 0-consistent subset of the product $\prod_U \widehat{T}_U$. Our next goal is the following theorem:

Theorem 6.4. *The 0-consistent subset of an HFI is a CAT(0) cube complex. Moreover, there is a bijective correspondence between its hyperplanes and the edges of the factors.*

6.2. **Median graphs and dual cube complexes.** The proof of Theorem 6.4 that we give here will use the notion of median spaces. The connection between median graphs and CAT(0) cube complexes goes back to Roller's [Rol16] reframing of Sageev's construction [Sag95], see also [CN05, Nic13]. Notably, Chepoi [Che00, Theorem 6.1] showed that every median graph is the 1-skeleton of a unique CAT(0) cube complex. Bowditch explains these connections in detail in [Bow22, Section 11].

Let X be a space and $\mathsf{m} : X^3 \to X$ a ternary operator. We call (X, m) a *median algebra* if for all $a, b, c, d, e \in X$ we have

$$\mathsf{m}(a,b,c) = \mathsf{m}(b,a,c) = \mathsf{m}(c,a,b), \quad \mathsf{m}(a,a,b) = a, \quad \mathsf{m}(a,b,\mathsf{m}(c,d,e)) = \mathsf{m}(\mathsf{m}(a,b,c),\mathsf{m}(a,b,d),e).$$

A graph Γ is *median* if all choices of $v_1, v_2, v_3 \in \Gamma^{(0)}$, there exists a unique point lying on some geodesic from v_i, v_j for all i, j; then $\mathsf{m}(v_1, v_2, v_3)$ is that unique point.

Given $a, b \in X$, the (median) *interval* between a, b is $I(a,b) = \{x \in X | \mathsf{m}(a,x,b) = x\}$. We say a median algebra (X, m) is *discrete* when $I(a,b)$ is finite for every $a, b \in X$. A *subalgebra* is a subset $S \subset X$ which is closed under m. In particular, locally finite median graphs are discrete median algebras.

Every median algebra (X, m) determines a graph called its *adjacency graph* $\Gamma(X)$ as follows. We say that $a, b \in X$ are *adjacent* when $I(a,b) = \varnothing$. The vertices of $\Gamma(X)$ are the points of X, with $a, b \in X$ connected by an edge in $\Gamma(X)$ exactly when a, b are adjacent. This graph is median exactly when (X, m) is a discrete median algebra [Bow22, Lemma 16.1.2].

Every CAT(0) cube complex X admits a canonical median structure as follows. For any $a, b \in X^{(0)}$, let $J(a,b)$ denote the set of ℓ^1-geodesics between a, b. Then given $a, b, c \in X^{(0)}$, we have $\mathsf{m}(a,b,c) = J(a,b) \cap J(a,c) \cap J(b,c)$ is a single point, and (X, m) is a median structure with $X^{(1)}$ a median graph, with notably $J(a,b) = I(a,b)$ for all $a, b \in X^{(0)}$. Moreover, it is immediate that $X^{(1)} = \Gamma(X^{(0)})$, since median adjacent vertices in $X^{(0)}$ are exactly those spanning an edge in $X^{(1)}$.

Proof of Theorem 6.4. First, observe that $\mathcal{Y} = \prod_{U \in \mathcal{U}} \widehat{T}_U$ is a product of simplicial trees and hence a CAT(0) cube complex.

Next, observe that $\mathcal{Q} \subset \mathcal{Y}$ is a cubical subcomplex as follows. For any $\hat{x} \in \mathcal{Q}$, let $C \subset \mathcal{Y}$ be the minimal cube containing $\hat{x}$ in its interior. Presuming C is not a vertex, minimality says that $\hat{x}_U$ is in the interior of an edge in each of the defining coordinate intervals of C, whose domain labels are pairwise orthogonal by item (9) of Lemma 4.10. This means that we can freely alter those coordinates of $\hat{x}$ to be any choice of points in those edges, meaning $C \subset \mathcal{Q}$. Hence $\mathcal{Q}$ is a cubical subcomplex of $\mathcal{Y}$.

It is enough then to show that $\mathcal{Q}^{(0)} \subset \mathcal{Y}$ is a discrete median subalgebra. For once that is done, it will follow that $\Gamma(\mathcal{Q})$ is a median graph, but evidently $\Gamma(\mathcal{Q}) = \mathcal{Q}^{(1)}$, as $\mathcal{Q}$ is a cubical subcomplex of $\mathcal{Y}$. Since every median graph is the 1-skeleton of a unique CAT(0) cube complex [Che00, Theorem 6.1], it follows that $\mathcal{Q}$ must be that cube complex, completing the proof.

The claim boils down to three exercises that we leave for the reader:

- Any (finite) product of median spaces admits the structure of a median space, where the median on the product is the tuple of medians in the coordinates.
- For any $\hat{a}, \hat{b}, \hat{c} \in \mathcal{Q}^{(0)}$, we have $\left(\mathsf{m}_U(\hat{a}_U, \hat{b}_U, \hat{c}_U)\right)_U \in \mathcal{Q}$. This can be proven using a basic consistency argument (see [Dur23, Lemma 15.12]).
- With this median structure, $a, b \in \mathcal{Q}^{(0)}$ are median adjacent exactly when they span an edge in $\mathcal{Q}^{(1)}$.

Finally, observe that the hyperplanes of $\mathcal{Y}$ are of the form $\mathbf{h} \times \prod_{V \neq U} \widehat{T}_V$ for $\mathbf{h}$ some hyperplane in $\widehat{T}_U$. Since $\mathcal{Q}$ is a median subalgebra of $\mathcal{Y}$, the hyperplanes of $\mathcal{Q}$ are just hyperplanes of $\mathcal{Y}$ intersected with $\mathcal{Q}$, so the moreover statement follows. This completes the proof. □

6.3. **More tree trimming and hyperplane deletions.** In Subsection 5.4, we introduced our tree trimming techniques, which allowed us to take a family of collapsed intervals and collapse a further family of subintervals while maintaining control over their associated 0-consistent subsets. In this subsection, we need a further refinement applied to HFIs and their associated cube complexes.

To set up the statement, let $\{\widehat{T}_U\}_{U\in\mathcal{U}}$ be a hierarchical family of intervals, as in Definition 6.1. Fix a constant $B > 0$, which will control the size and number of subintervals we will collapse across all of the intervals. In our applications, we will be able to control the size of B, but notably not with respect to the largeness constant K, unlike in the statement of Theorem 5.14.

Definition 6.5 (Admissible simplicial trimming setup)**.** Given $B > 0$, a *B-admissible simplicial trimming setup* is a collection of *trimming intervals* $\mathcal{B}$ in the various $\widehat{T}_{U,e}$ satisfying the following:

(1) For each $D \in \mathcal{B}$ with $D \subset E \in \pi_0(\widehat{T}_{U,e})$, we have
   (a) D is a subinterval of E with endpoints at vertices, and
   (b) the length $l(D)$ of D is a positive integer whose length is bounded above by B.
(2) The number of subintervals in $\mathcal{B}$ is bounded by B, that is $\#\mathcal{B} < B$.

Remark 6.6. We note that the conditions in Definition 6.5 guarantee that each quotient interval $\widehat{T}_{U,0}$ is a simplicial interval of the usual type we consider. The requirement that each segment $D \in \mathcal{B}$ has endpoints at vertices guarantees that when we collapse such a segment, then it only affects the interval-hyperplanes it contains, preserving those in its complement.

We use the same associated notation as in Subsection 5.4. Furthermore, we note that if $\widehat{T}^e_U \subset \mathcal{B}$, then $\operatorname{diam}(\widehat{T}_U) < B^2$ and $\widehat{T}_{U,0}$ is a point. Let $\mathcal{U}_0 \subset \mathcal{U}$ denote the subset of domains for which $\widehat{T}_{U,0}$ is not a point. By ignoring the domains in $\mathcal{U} - \mathcal{U}_0$, we can consider the set of collapsed intervals $\{\widehat{T}_{U,0}\}_{U\in\mathcal{U}_0}$, where we remove any labels $\hat{\delta}^V_U$ from $\widehat{T}_U$ from any domain $V \sqsubset U$ with $V \in \mathcal{U} - \mathcal{U}_0$.

The following theorem expands on Theorem 5.14 in the HFI setting:

Theorem 6.7. *Let $B > 0$ and $\{\widehat{T}_U\}_{U\in\mathcal{U}}$ be an HFI. For any B-admissible simplicial trimming setup $\mathcal{B}$, the following hold:*

(1) *The collections $\{\widehat{T}_{U,0}\}_{U\in\mathcal{U}}$ and $\{\widehat{T}_{U,0}\}_{U\in\mathcal{U}_0}$ are HFIs. We let $\mathcal{Q}', \mathcal{Q}_0$ denote their corresponding 0-consistent subsets.*
(2) *We have $\Delta(\mathcal{Q}) = \mathcal{Q}'$, where $\Delta : \prod_U \widehat{T}_U \to \prod_U \widehat{T}'_U$ is the global collapsing map.*
(3) *There exists a map $\Xi : \mathcal{Q}' \to \mathcal{Q}$ with $\Delta \circ \Xi = id_{\mathcal{Q}'}$.*
(4) *The map $\Delta|_{\mathcal{Q}} : \mathcal{Q} \to \mathcal{Q}'$ is a $(1, B^2)$-quasi-isometry.*
(5) *At the level of cube complexes, the map $\Delta : \mathcal{Q} \to \mathcal{Q}'$ is a hyperplane deletion map deleting at most B^2-many hyperplanes.*

Proof. Item (1) is straightforward, since we are just collapsing simplicial subintervals. Items (2)–(3) follow from Theorem 5.14, where in the argument we just replace the HFI axioms (Definition 6.1) with the appropriate properties from Lemma 4.10. The fact that Δ deletes at most B^2-many hyperplanes from $\prod_U \widehat{T}_U$, so item (5) follows from the "moreover" part of Theorem 6.4. Finally, item (4) follows immediately from item (5). This completes the proof.

□

6.4. **Hierarchy paths and the distance formula.** With Theorem 6.4 in hand, we can make some observations about the cubical structure of $\mathcal{Q}$ and how it relates to its hierarchical nature. In particular, we get a proof of the (more difficult) lower bound of the Distance Formula 3.8 and the existence of hierarchy paths.

The following lemma is [Dur23, Lemma 14.7], and we have included its proof here for completeness.

Lemma 6.8. *Let $\hat{x}, \hat{y} \in \mathcal{Q}$. Let γ be any combinatorial geodesic in $\mathcal{Q}$ from $\hat{x}$ to $\hat{y}$. Then for all $U \in \mathcal{U}$, the projection $\hat{\pi}_U(\gamma) \subset \widehat{T}_U$ is an unparameterized geodesic between $\hat{x}_U$ and $\hat{y}_U$.*

Proof. Since combinatorial geodesics do not cross a hyperplane twice and the restriction quotient $\hat{\pi}_U : \mathcal{Q} \to \widehat{T}_U$ is 1-Lipschitz, it follows that $\hat{\pi}_U(\gamma)$ does not backtrack either. This completes the proof.

□

This lemma plays two roles: it's useful for proving that an isomorphism of HFIs induces an isomorphism of their cubical duals (Proposition 6.12), and, in the next proposition, we use it directly to show that combinatorial geodesics in a cubical model get sent to hierarchy paths (Definition 3.9) in the ambient HHS. That is the content of the next proposition, which is simplified version of [Dur23, Proposition 14.8]:

Proposition 6.9. *Let $a, b \in \mathcal{X}$, $K \geqslant K_0$ where $K_0 = K_0(\mathfrak{S}) > 0$ is from Theorem 3.8, $\mathcal{U} = \mathrm{Rel}_K(a, b)$, and let $H = \mathrm{hull}_{\mathcal{X}}(a, b)$. Let $\{T_U\}_{U \in \mathcal{U}}$ be any R-interval system for a, b and $\{\widehat{T}_U\}_{U \in \mathcal{U}}$ the HFI associated to any (r_1, r_2)-thickening for $r_1 \geqslant R_1, r_2 \geqslant R_2$ with r_1 a positive integer. Let $\mathcal{Q}$ be the corresponding 0-consistent set and $\widehat{\Omega} : \mathcal{Q} \to H$ the quasi-isometry constructed in Section 5.*

- *If $\hat{a}, \hat{b}$ are the points $\mathcal{Q}$ corresponding to a, b and γ is a combinatorial geodesic between them, then $\widehat{\Omega}(\gamma) \subset H$ is an (L_{hp}, L_{hp})-hierarchy path between a, b for $L_{hp} = L_{hp}(R, r_1, r_2, K, \mathfrak{S}) > 0$.*

Proof. The following is a fairly complete sketch of the proof, which is simpler than the more general version from [Dur23, Proposition 14.8].

Since the map $\widehat{\Omega} : \mathcal{Q} \to H$ is effectively defined coordinate-wise, its enough to consider the coordinate-wise maps $B_U : \widehat{T}_U \to T_U$ as described in Subsection 5.2, as the map $\phi_U : T_U \to \mathcal{C}(U)$ is an (R, R)-quasi-isometric embedding by assumption. For this, we can let γ_U denote the coordinate-wise projection of γ to each $\widehat{T}_U$, and we want to show that $B_U(\gamma_U)$ is an unparameterized quasi-geodesic in T_U.

By construction, $B_U(\gamma_U) \subset [a_U, b_U] = T_U$, so it suffices to show that this path never backtracks an unbounded amount. Proceeding by induction in the $\sqsubset_{\mathcal{U}}$-level in $\mathcal{U}$, if $U \in \mathcal{U}$ is $\sqsubset$-minimal, then the amount of collapsing in $q_U : T_U \to \widehat{T}_U$ is bounded by $2r_1$, and we are done by Lemma 6.8. For non-minimal domains $U \in \mathcal{U}$, note that by Lemma 6.8, backtracking in T_U is only possible if $\gamma_U(t), \gamma_U(s)$ are at the same collapsed point corresponding to some cluster C. If there are times $s < t$ so that $d_{T_U}(B_U(\gamma(s), b_U) - d_{T_U}(B_U(\gamma(s)), b_U) \gg 100r_2 + 100r_1$, then there is some $\sqsubset_{\mathcal{U}}$-minimal domain $V \sqsubset U$ with $\hat{\delta}_U^V = q_U(C)$ for which $B_U(\gamma_U(t))$ is on the side of a_U while $B_U(\gamma_U(s))$ is on the side of b_U, with both $(50r_2 + 50r_1)$-away from δ_U^V. But then consistency and the BGI properties from Lemma 4.2 implies that

$B_V(\gamma_V(t)), a_V$ and $B_V(\gamma_V(s)), b_V$ are R-close, and so $\gamma_V(t) = \hat{a}_V$ and $\gamma_V(s) = \hat{b}_V$, which contradicts the base case of this induction. This completes the proof. □

Remark 6.10. Since the map $\widehat{\Omega} : \mathcal{Q} \to \mathcal{X}$ is uniformly quasi-median [Dur23, Section 15], one could alternatively use the fact that median paths in HHSes are hierarchy paths [BHS21, Lemma 1.37].

As a consequence of Propositions 6.9 and 5.3, we obtain the lower bound of the Distance Formula 3.8:

Corollary 6.11. *There exists $K_0 = K_0(\mathfrak{S}) > 0$ such that for all $K > K_0$ and any $x, y \in \mathcal{X}$, we have*

$$\sum_{U \in \mathfrak{S}} [d_U(x, y)]_K \prec d_{\mathcal{X}}(x, y).$$

6.5. **Isomorphisms of HFIs and their associated cube complexes.** In this final subsection, we will discuss when two HFIs to determine isomorphic cube complexes. This statement, while unsurprising, motivates what happens in the subsequent sections, hence its placement here.

Let $\{\widehat{T}_U\}_{U \in \mathcal{U}}$ and $\{\widehat{T}'_U\}_{U \in \mathcal{U}}$ be two HFIs over a common index set $\mathcal{U}$ and common relations, i.e. $U, V \in \mathcal{U}$ have the same relation in both HFIs. Any family of maps $i_U : \widehat{T}_U \to \widehat{T}'_U$ induces a global map $I : \prod_{U \in \mathcal{U}} \widehat{T}_U \to \prod_{U \in \mathcal{U}} \widehat{T}'_U$, where $I((\hat{x}_U)) = (i_U(\hat{x}_U))$.

The following proposition, which is an interval version of [Dur23, Proposition 14.12], gives sufficient conditions for the induced map to be an isometry $I : \mathcal{Q} \to \mathcal{Q}'$ which induces a cubical isomorphism between their dual cube complexes.

Proposition 6.12. *Let $\{\widehat{T}_U\}_{U \in \mathcal{U}}$ and $\{\widehat{T}'_U\}_{U \in \mathcal{U}}$ be two HFIs over a common index set $\mathcal{U}$ with marked points $F = \{a, b\}$, and let $\mathcal{Q}, \mathcal{Q}'$ be their corresponding 0-consistent sets. Suppose that for each $U \in \mathcal{U}$, there is a simplicial isometry $i_U : \widehat{T}_U \to \widehat{T}'_U$ which also satisfies:*

1. *We have $i_U(\hat{a}_U) = \hat{a}'_U \in \widehat{T}'_U$ and $i_U(\hat{b}_U) = \hat{b}'_U \in \widehat{T}'_U$.*
2. *For all $V \in \mathcal{U}$ with $V \sqsubset U$ or $V \pitchfork U$, we have $i_U(\hat{\delta}^V_U) = \left(\hat{\delta}^V_U\right)'$.*
3. *For all $V \in \mathcal{U}$ with $U \sqsubset V$ and $\hat{x}_V \in \widehat{T}_V$, we have $\left(\hat{\delta}^V_U\right)'(i_V(\hat{x}_V)) = i_U(\hat{\delta}^V_U(\hat{x}_V))$.*

Then the corresponding map induces a cubical isomorphism $I : \mathcal{Q} \to \mathcal{Q}'$.

Sketch of proof. The proof of this version below is the same as the original, which is, in turn, an exercise in the construction from this section plus Lemma 6.8. The main idea is that items in the statement guarantee that the global map I induced from the coordinate-wise isometries preserves 0-consistency and distances, i.e. we have $I : \mathcal{Q} \to \mathcal{Q}'$ is an isometric embedding, with surjectivity following from a consistency argument. Since the interval-wise maps respect the simplicial structure, we have that I is also a cubical isomorphism via Lemma 2.5. We leave the details to the reader.

□

7. Controlling domains

In this section, we turn toward the proof of our main technical result, namely the Stable Cubical Interval Theorem 9.9, which is a stabilized version of the cubical

model construction for a pair of points a, b in any HHS $\mathcal{X}$. The goal is to show that if we perturb a, b a small amount to a pair of points a', b' with $d_{\mathcal{X}}(a, a') \leqslant 1$ and $d_{\mathcal{X}}(b, b') \leqslant 1$, then we can produce cubical models $\mathcal{Q}$ for a, b and $\mathcal{Q}'$ for a', b' which are isomorphic up to deleting boundedly-many hyperplanes. Doing so requires carefully controlling the main inputs into the cubulation machine—namely the families of intervals we construct from the projections of the points.

The first step in the process involves explaining how to gain control over the set of relevant domains (Definition 3.5) for a pair of points $F = \{a, b\} \subset \mathcal{X}$ and its perturbation $F' = \{a', b'\}$ as above. In the next sections, we explain how to take this projection data to build a family of stable intervals that we can plug into the cubulation machine.

The main bit of work for this first part has already been done for us by Bestvina–Bromberg–Fujiwara [BBF15] and Bestvina–Bromberg–Fujiwara–Sisto [BBFS20]. The work we did with Minsky and Sisto in [DMS20] mostly adapts their setup to the HHS setting, and makes some straight-forward but important observations.

In [DMS20, Subsections 2.3 and 2.4], we were dealing with the situation where F, F' were finite sets of points with cardinality possibly larger than 2. Nonetheless, all of the arguments there reduce to considering pairs of points, and so the proofs for our current setting are essentially the same. As such, we will mainly explain the setup and key statements, only giving minimal sketches of proofs.

7.1. **Colorability.** The hierarchical setting in which [BBFS20] works is for the following class, which was first defined in [DMS20, Definition 2.8] and later used in [HP23, Pet21]:

Definition 7.1 (Colorable HHS)**.** We say that an HHS $(\mathcal{X}, \mathfrak{S})$ is *colorable* if there exists a decomposition of $\mathfrak{S}$ into finitely many families $\mathfrak{S}_i$, so that each $\mathfrak{S}_i$ is pairwise-$\pitchfork$.

- We refer to the $\mathfrak{S}_i$ as *BBF families*.

Remark 7.2. The notion of colorability was inspired by work of Bestvina-Bromberg-Fujiwara [BBF15], who proved that the curve graph is finitely-colorable. As a consequence of this, the set of subsurfaces of a given finite-type surface S can be decomposed into finitely-many families of pairwise-interlocking subsurfaces, which the above definition directly generalizes. See [DMS20, Subsection 2.3] for a detailed discussion of this connection.

7.2. **Stable projections.** Our next goal is to state a theorem which equips any colorable HHS with a set of projections which satisfy certain key stability properties. We need some definitions first.

Given a pair of points $F = \{x, y\} \subset \mathcal{X}$ and $K > 0$, we let $\mathcal{U}(F) = \mathrm{Rel}_K(F)$. If $V \in \mathcal{U}(F)$, then we let $\mathcal{U}^V(F) = \{W \in \mathcal{U}(F) | W \subsetneq V\}$.

The following definition [DMS20, Definition 2.13] describes the domains we want to control when perturbing F:

Definition 7.3 (Involved domains)**.** Let $F = \{x, y\}, F' = \{x', y'\} \subset \mathcal{X}$ be so that $d_{\mathcal{X}}(x, x') \leqslant 1$ and $d_{\mathcal{X}}(y, y') \leqslant 1$. For $K > 0$, we say that $V \in \mathcal{U}(F) \cup \mathcal{U}(F')$ is *involved in the transition* between F and F' if one of the following holds:

(1) $\pi_V(F) \neq \pi_V(F')$.
(2) $\mathcal{U}^V(F) \neq \mathcal{U}^V(F')$.

The following, which combines the statements of [DMS20, Theorem 2.9 and Proposition 2.14], states that, for any colorable HHS $\mathcal{X}$, we can replace its given projections with coarsely equivalent projections with which we can control involved domains up to a bounded error:

Theorem 7.4. *Let $(\mathcal{X}, \mathfrak{S})$ be a colorable HHS with standard projections $\hat{\pi}_-, \hat{\rho}^-_-$. Then there exists $\theta > 0$ and new projections π_-, ρ^-_- with the same domains and ranges, respectively, such that the following hold:*

1. *The new projections π_-, ρ^-_- coarsely coincide with the old ones $\hat{\pi}_-, \hat{\rho}^-_-$. More precisely:*
   - (a) *If X, Y lie in different $\mathfrak{S}_j$ and $\hat{\rho}^X_Y$ is defined, then $\rho^X_Y = \hat{\rho}^X_Y$.*
   - (b) *If $X, Y \in \mathfrak{S}_j$, then $d^{Haus}_Y(\rho^X_Y, \hat{\rho}^X_Y) < \theta$.*
   - (c) *If $x \in \mathcal{X}$ and $Y \in \mathfrak{S}$, then $d^{Haus}_Y(\pi_Y(x), \hat{\pi}_Y(x)) < \theta$.*
2. *There exists $K_0 = K_0(\mathcal{X}) > 0$ and $N_0(\mathcal{X}) > 0$ so that if $K > K_0$, then the following holds. Let $x, y, x', y' \in \mathcal{X}$ with $d_{\mathcal{X}}(x, x') \leqslant 1$ and $d_{\mathcal{X}}(y, y') \leqslant 1$, and let $\mathcal{U} = \mathrm{Rel}_K(\{x, y\})$ and $\mathcal{U}' = \mathrm{Rel}_K(\{x', y'\})$ be the relevant sets for the new projections π_-. Then:*
   - (a) *$\#(\mathcal{U} \triangle \mathcal{U}') < N_0$, and if $V \in \mathcal{U} \triangle \mathcal{U}'$, then $d_V(x, y) \prec K$ and $d_V(x', y') \prec K$.*
   - (b) *There are at most N_0 domains $V \in \mathcal{U}(F) \cup \mathcal{U}(F')$ involved in the transition between F and F'.*

Sketch of proof. The existence of new projections satisfying the conditions of item (1) is essentially an immediate consequence of the construction in [BBFS20, Theorem 4.1]. Item (2), which regards stability of the relevant sets $\mathcal{U}$,$\mathcal{U}'$, requires a bit more work, but the key ideas are as follows:

Using [BHS19, Proposition 2.8] (which generalizes [CLM12, BKMM12, BBF15]), there is a partial order $\prec$ on $\mathrm{Rel}(x, y)$, which becomes a total order on $\mathrm{Rel}^i(x, y) = \mathrm{Rel}_K(x, y) \cap \mathfrak{S}_i$ for each i. In particular, we can think of each $\mathrm{Rel}^i(x, y)$ as an interval $X_1 \prec \cdots \prec X_n$. One of the additional properties that the new projections π_-, ρ^-_- satisfy (see [DMS20, Theorem 2.9]) is that if $X \prec Y$ along both $\mathrm{Rel}(x, y)$ and $\mathrm{Rel}(x', y')$, then $\pi_Y(x) = \pi_Y(x') = \rho^X_Y$ and $\pi_X(y) = \pi_X(y') = \rho^Y_X$. Using this, one can show that $\mathrm{Rel}(x, y)$ and $\mathrm{Rel}(x', y')$ can only differ at their first or last elements in the order $\prec$. Hence $\#\left(\mathrm{Rel}^i(x, y) \triangle \mathrm{Rel}^i(x', y')\right) \leqslant 4$; see [DMS20, Lemma 2.11 and Proposition 2.12]. Since there are only boundedly-many BBF families $\mathfrak{S}_i$, item (2a) follows easily.

A similar argument can now be used to control the number of involved domains of type (1) of Definition 7.3, see the proof of [DMS20, Proposition 2.14]. For involved domains of type (2), this is an immediate consequence of item (2a) and the Covering Lemma 3.14, which bounds the number of domains $W \in \mathcal{U}$ into which a given $V \in \mathcal{U}$ can nest. □

8. Stable Intervals

With the new HHS projections provided by Theorem 7.4, we now turn toward converting stabilized projection data into stable families of intervals, in partivuvlar stabilizing the input into the cubical model construction from Section 4.

Our ultimate goal is to produce cubical models for pairs of nearby points that differ by a controlled number of hyperplane deletions (Theorem 9.9). Since the hyperplanes in the cubical models come from edges of the simplicial intervals in

our HFIs (Theorem 6.4), it will suffice to build HFIs where the intervals only differ by boundedly-many bounded-size collections of edges. That is the content of the Stable Intervals Theorem 8.14.

We first explain the build-up to this theorem and defer the explicit statement. We will work in the setting of a hyperbolic space with a hierarchical-like configuration. The construction is a somewhat long and delicate exercise in hyperbolic geometry and we will mainly provide a detailed sketch.

8.1. **Basic setup and assumptions.** For the rest of this section, let us assume that $\mathcal{Z}$ is a δ-hyperbolic geodesic space. The first bit of notation we want to set is a choice of geodesic between any pair of points. That is, for any $x, y \in \mathcal{Z}$, we fix (once and for all) a geodesic $\lambda(x, y)$. More generally, if $C, C' \subset \mathcal{Z}$ are finite subsets, then we let $\lambda(C, C')$ denote a choice of geodesic between them whose length satisfies $|\lambda(C, C')| = d_{\mathcal{Z}}(C, C')$, and we require symmetry, namely $\lambda(C, C') = \lambda(C', C)$.

It will also be necessary to perform a slight modification on the geodesic function λ, so that its resulting length is an integer. To fix this, we let $\lambda_0(C, C')$ denote the restriction of $\lambda(C, C') : [0, d_{\mathcal{Z}}(C, C')] \to \mathcal{Z}$ to the interval $[0, \lfloor d_{\mathcal{Z}}(C, C') \rfloor]$. The effect of this is to trim $\lambda(A, B)$ so that it has positive integer length and their lengths satisfy $l(\lambda(C, C')) - l(\lambda_0(C, C')) < 1$. While not necessary, one could make this symmetric by trimming both the sides by half of $l(\lambda(C, C')) - l(\lambda_0(C, C'))$. In the course of the arguments, however, one can usually work with λ instead of λ_0.

Now, for the rest of this section, we fix a pair of points $\{a, b\} \subset \mathcal{Z}$. Furthermore, let $\mathcal{Y} \subset \mathcal{Z}$ be a finite (but possibly arbitrarily large) set of points of $\mathcal{Z}$ and let $\epsilon > 0$ so that $d_{\mathcal{Z}}(\lambda(a, b), y) < \epsilon/2$ for all $y \in \mathcal{Y}$. We call such an arrangement $(a, b; \mathcal{Y})$ an *ϵ-setup* in $\mathcal{Z}$.

In the hierarchical setting, $\mathcal{Z}$ will be one of the hyperbolic spaces $\mathcal{C}(U)$ for $U \in \mathcal{U}$, and the points $\mathcal{Y}$ will be the points in ρ_U^V for $V \sqsubset U \in \mathcal{U}$. The proximity condition of the points in $\mathcal{Y}$ to $\lambda(a, b)$ comes from BGIA 3.11.

Notation 8.1 ($O(\epsilon)$-notation)**.** In this section, we will often want to get controlled bounds over various distance quantities, where the bounds will depend on our choice of ϵ and the hyperbolicity constant δ for $\mathcal{Z}$. When a number A is bounded above as a function of ϵ, δ, we will simply write

$$A < O(\epsilon),$$

since δ is given to us, while we need some flexibility to choose ϵ.

The following lemma, which is essentially a consequence of the Morse Lemma for hyperbolic spaces, gives us control over how far $\lambda(C, C')$ can get from $\lambda(a, b)$ when $C, C' \subset \mathcal{Y}$, and are thus ϵ-close to $\lambda(a, b)$:

Lemma 8.2 (Fellow-traveling, δ-hyperbolic)**.** *Let $\mathcal{Z}$ be δ-hyperbolic and geodesic, and γ a geodesic between $a, b \in \mathcal{Z}$. For every $\epsilon > 0$ there exists $\epsilon' = O(\epsilon) > 0$ so that the following holds:*

- *Suppose γ' is a geodesic connecting points $x, y \in \mathcal{N}_\epsilon(\gamma)$, and that x', y' are closest points to x, y on γ, respectively. Then any geodesic $[x, y]$ satisfies $d_{\mathcal{Z}}^{Haus}([x, y], [x', y']) < \epsilon'$.*

As a consequence, if $C, C' \subset \mathcal{Y}$ are finite subsets, then $\lambda(C, C') \subset \mathcal{N}_{\epsilon'}(\lambda(a, b))$. Finally, it will be convenient for us to assume that $\epsilon' > \epsilon$.

8.2. **Clusters and the cluster graph.** Our first goal is to group the points of $\mathcal{Y}$ into *clusters* of nearby sets. To that end, let $E = E(\epsilon) > \epsilon$ be sufficiently large, to be determined below.

Let $\mathcal{C}_E$ denote the graph whose vertices are points of $\mathcal{Y} \cup \{a, b\}$, with two points $x, y \in \mathcal{Y} \cup \{a, b\}$ connected by an edge exactly when $d_{\mathcal{Z}}(x, y) < E$. The connected components of $\mathcal{C}_E$ are called *E-clusters*, or simply *clusters* when E is fixed.

For any E-cluster C, let $s(C) = \text{hull}_{\lambda(a,b)}(p_{\lambda(a,b)}(C))$ denote the hull (in $\lambda(a, b)$) of the closest point projection of C to $\lambda(a, b)$. The set $s(C)$ is called the *shadow* of C on $\lambda(a, b)$.

The following lemma controls how shadows of clusters are arranged along $\lambda(a, b)$. Its proof is straightforward (see [DMS20, Lemma 3.6]), so we leave it to the reader to check (:

Lemma 8.3. *If $E > 5\epsilon$ and $C \neq C'$ are distinct E-clusters, then $s(C) \cap s(C') = \varnothing$.*

In what follows, we will need to adjust the cluster constant E a number of times, but we will usually omit the constant and refer to them simply as clusters.

We now can define our notion of cluster separation using Lemma 8.2. The idea here is that clusters are ϵ-close to $\lambda(a, b)$, and hence geodesics connecting them stay ϵ'-close to $\lambda(a, b)$ by Lemma 8.2.

Definition 8.4. Let C_1, C_2, C_3 be clusters. Given $\epsilon' > 0$, we say C_2 *ϵ'-separates* C_1 from C_3 if there exists a minimal length geodesic connecting C_1 to C_3 which passes within $2\epsilon'$ of C_2.

- Given $E > 0$, the *E-cluster separation graph* for $(a, b; \mathcal{Y})$, denoted $\mathcal{G}_E = \mathcal{G}_{E,\epsilon'}(a, b; \mathcal{Y})$ is defined as follows: The vertices of $\mathcal{G}_E$ are clusters, and two vertices are connected by an edge if and only if they are not $2\epsilon'$-separated by another cluster.

The following is essentially a consequence of Lemma 8.3. It states that the cluster graph is an interval with the order of vertices corresponding to the order of the cluster shadows along $\lambda(a, b)$. We omit its proof, which uses only hyperbolic geometry and Lemma 8.3.

Proposition 8.5. *Let $\mathcal{Z}$ be δ-hyperbolic, $\lambda(a, b)$ a geodesic connecting $a, b \in \mathcal{Z}$, and $\mathcal{Y} \subset \mathcal{Z}$ finite with $\mathcal{Y} \subset \mathcal{N}_\epsilon(\lambda(a, b))$. Also let $\epsilon' = \epsilon'(\epsilon, \delta) > 0$ as in Lemma 8.2, with also $\epsilon' > \epsilon$.*

There exists $E = E(\epsilon, \delta) > 0$ so that the E-cluster separation graph $\mathcal{G}_{E,\epsilon'}(a, b; \mathcal{Y})$ is an interval whose order coincides with the order of cluster shadows along $\lambda(a, b)$, with the endpoint clusters containing a, b, respectively.

8.3. **Stable intervals defined.** We are now ready to describe the stable intervals.

By Proposition 8.5, the cluster separation graph is an interval. Our stable intervals admit a decomposition into two families of intervals, namely of *edge* components and *cluster* components. An edge component connects adjacent clusters at their nearest points, while a cluster component connects the endpoints of whatever edge components are connected to a given cluster.

Definition 8.6 (Stable interval)**.** Let $\mathcal{Z}$ be δ-hyperbolic, $a, b \in \mathcal{Z}$ with $\lambda(a, b)$ a geodesic connecting them, and $\mathcal{Y} \subset \mathcal{Z}$ finite with $\mathcal{Y} \subset \mathcal{N}_{\epsilon/2}(\lambda(a, b))$. The *stable interval* for this setup is the union of the following two families of geodesic segments:

FIGURE 5. Proof of Lemma 8.8: Clusters can have multiple pairs of closest points. In this example, the red and pink points on C_1 are equidistant from their matching point on C_2, while the red and pink cluster points in C_1 are both within 7ϵ of the endpoint of the shadow $s(C_1)$ nearest $s(C_2)$, and the same for the roles of C_1, C_2 switched. In subsequent arguments, this will allow us to conclude that connecting points between clusters are close even after we alter the clusters.

(1) (Edge components) For each adjacent pair of C, C' of clusters in $\mathcal{G} = \mathcal{G}_{E,\epsilon'}(a, b; \mathcal{Y})$, the *edge component* between them is $\lambda_0(C, C')$.
   - The edge components comprise the *edge forest*, which we denote by T_e.

(2) (Cluster components) For each cluster C, we let C_1, C_2 be the clusters adjacent to C in $\mathcal{G}$. Let $r(C) = C \cap (\lambda(C_1, C) \cup \lambda(C, C_2) \cup \{a, b\})$, where $\lambda(C_i, C) = \varnothing$ if C is an endpoint cluster, for the appropriate $i \in \{1, 2\}$. The cluster C then determines the *cluster component* $\mu(C) = \lambda(r(C))$.
   - The cluster components comprise the *cluster forest*, which we denote by T_c.

- The *stable interval* is $T(a, b; \mathcal{Y}) = T_e \sqcup T_c$, by which we mean the abstract union of these two forests, where we attach $\lambda_0(C_1, C_2)$ to the corresponding endpoints of $\mu(C_1), \mu(C_2)$ for adjacent C_1, C_2 as determined by $\mathcal{G}$.
- We let $\phi : T(a, b; \mathcal{Y}) \to \mathcal{Z}$ denote the map induced from the fact that its pieces are geodesics in $\mathcal{Z}$.

Remark 8.7. We note that stable intervals are abstractly intervals, but not necessarily as subsets of $\mathcal{Z}$. In particular, the points a, b need not be the "endpoints" of the stable interval T, either abstractly or in $\mathcal{Z}$. However, their images in the collapsed stable interval will be the endpoints, see Subsection 9.2.

The following lemma is the key to controlling the structure of $T(a, b; \mathcal{Y})$. We omit its proof since it again uses only hyperbolic geometry, though see Figure 5 for a schematic:

Lemma 8.8. *Let C_1, C_2 be distinct clusters along $\lambda(a, b)$. Suppose that $p_1 \in C_1$ and $p_2 \in C_2$ are closest points, while $x_1 \in s(C_1)$ and $x_2 \in s(C_2)$ are closest points. Then $d_{\mathcal{Z}}(x_i, p_i) \leqslant 7\epsilon$ for $i = 1, 2$.*

As a consequence, the following hold:

(1) *For distinct clusters C_1, C_2, we have $|d(C_1, C_2) - d(s(C_1), s(C_2))| < 14\epsilon$.*
(2) *For any $p \in T_e$, we have that $d_{\mathcal{Z}}(p, \mathcal{Y} \cup \{a, b\}) \leqslant d_{\mathcal{Z}}(p, \partial T_e) + O(\epsilon)$.*
(3) *For any components V, W of $T = T_e \cup T_c$, we have $\operatorname{diam}_{\mathcal{Z}}(s(V) \cap s(W)) < O(\epsilon)$.*

With the above lemma in hand, we can prove the following structural proposition:

Proposition 8.9. *There exists $L_0 = L_0(\delta, \epsilon) > 0$ so that the following hold:*

(1) *$T(a, b; \mathcal{Y})$ is an interval with endpoints contained in $\mathcal{Y} \cup \{a, b\}$ whose images in $\mathcal{Z}$ are $O(\epsilon)$ of $a, b \in \mathcal{Z}$.*
(2) *For every cluster C, we have that $d_{\mathcal{Z}}^{Haus}(\mu(C), C) < L_0$ and $\phi(T_C) \subset \mathcal{N}_{L_0}(\mathcal{Y} \cup \{a, b\})$.*

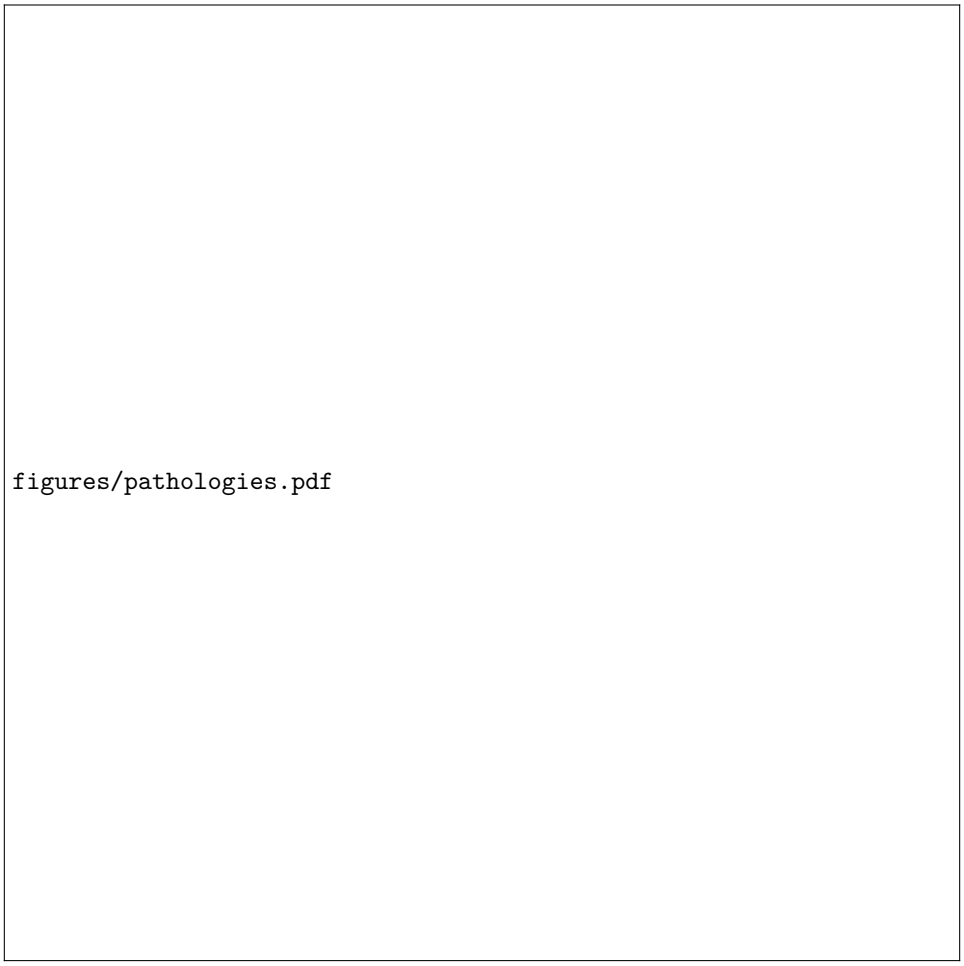


FIGURE 6. This figure, which is essentially [DMS20, Figure 11], is an example of two phenomena. First, the edge components of a stable interval can have bounded overlap in the space $\mathcal{Z}$. This can lead to a bounded multiplicative error in the quasi-isometric embedding $\phi : T(a, b; \mathcal{Y}) \to \mathcal{Z}$; see item (3) of Proposition 8.9. In the figure, the multiplicative constant is at least $\frac{102}{100}$. Second, the image of $T(a, b; \mathcal{Y})$ in $\mathcal{Z}$ need not itself be an interval, as the clusters can form leaves on the image.

(3) *The map $\phi : T(a, b; \mathcal{Y}) \to \mathcal{Z}$ is an (L_0, L_0)-quasi-isometric embedding with image L_0-Hausdorff close to $\lambda(a, b)$.*
(4) *Each edge component of T has positive integer length.*

Proof. For (1), observe that $T(a, b; \mathcal{Z})$ is an interval by construction and Proposition 8.5, and the end clusters contain a, b, respectively, also by that proposition. Since each cluster $C \subset \mathcal{N}_E(\lambda(a, b))$, if $a \in C$, then the outer-most endpoint of $\mu(C)$ must be within E of a, as required.

For (2), observe that each cluster C is within Hausdorff distance $O(\epsilon)$ of $s(C)$, while Lemma 8.8 implies that $\mu(C)$ is within Hausdorff distance $O(\epsilon)$ of $s(C)$. The conclusion follows.

Item (3) is essentially a consequence of Lemma 8.8 and the construction as follows. By Lemma 8.2, the image under the map $\phi : T \to \mathcal{Z}$ of each component of T_e and T_c is within $O(\epsilon)$ of its shadow on $\lambda(a, b)$. In other words, for any component V of T_e or T_c, the shadow map induces a $(1, O(\epsilon))$-quasi-isometry $s : \phi(V) \to s(\phi(V))$. On the other hand, Lemma 8.8 says every pair of components of T have shadows on $\lambda(a, b)$ with $O(\epsilon)$-bounded overlap. It follows that $\phi : T \to \mathcal{Z}$ is an $(O(\epsilon), O(\epsilon))$-quasi-isometric embedding, as required.

Finally, item (4) follows because we have used the function λ_0 as our connectors in Definition 8.6, which performs a small surgery at one of the ends of λ so that the result has integer length. This completes the proof. □

8.4. **Stable decompositions.** Having defined stable intervals above, we are almost ready to state the Stable Intervals Theorem 8.14. The motivation for the definitions in this section comes from our desire to plug things into our cubical model machinery. We describe how this works in Subsection 9.3 below, and we point the reader to it for motivation for what follows.

A stable decomposition involves decomposing a pair of stable intervals for a pair of admissible setups into subintervals, with some maps which identify and organize the various pieces. The purpose of the next definition is to give shorthand for requiring that the various maps involved respect how the various pieces are oriented towards the endpoints of the ambient intervals.

Definition 8.10 (Order-preserving maps)**.** Let T_1, T_2 be two intervals with the endpoints of T_i labeled by a_i, b_i. Suppose that $T_i = T_{i,1} \sqcup T_{i,2}$ are subdivisions of T_i into finite collections of subintervals with $\#T_{1,1} = \#T_{2,1}$ (and hence $\#T_{1,2} = \#T_{2,2}$).

- We say an isometry $i_{E,E'} : E \to E'$ between components $E \subset T_{1,1}$ and $E' \subset T_{2,1}$ is *order preserving* if $i_{E,E'}$ sends the endpoint of E closest to a_1 to the endpoint of E' closest to a_2.
- We say a bijection $\alpha : \pi_0(T_{1,1}) \to \pi_0(T_{2,1})$ is *order-preserving* if it coincides with the bijection obtained by recording the orders of appearance of the components of $T_{i,1}$ along T_i from a_i to b_i.

Remark 8.11. Using the above notation, note that an order-preserving bijection $\alpha : \pi_0(T_{1,1}) \to \pi_0(T_{2,1})$ automatically induces an order-preserving bijection $\beta : \pi_0(T_{1,2}) \to \pi_0(T_{2,2})$. Moreover, in this paper, we will only be working with intervals between a pair of points, which are usually denoted by (some form of) a, b. We will always take order-preserving to mean as one moves from the a-side to the b-side of the given intervals.

We also set one last bit of notation: Given a stable tree $T = T_e \cup T_c$ for an ϵ-setup $(\mathcal{Y}; \{a, b\})$, if $y \in \mathcal{Y} \cup \{a, b\}$, we let C_y denote the cluster containing y, and $\mu(C_y)$ the corresponding component of T^c.

The following definition contains the stability properties we want:

Definition 8.12 (Stable decomposition)**.** Let $\mathcal{Z}$ be δ-hyperbolic and geodesic.

(1) Given $N > 0$, two ϵ-setups $(a, b; \mathcal{Y})$ and $(a', b'; \mathcal{Y}')$ are (N, ϵ)-*admissible* if
   (a) $d_{\mathcal{Z}}(a, a'), d_{\mathcal{Z}}(b, b') \leqslant \epsilon$,

(b) $\mathcal{Y}, \mathcal{Y}' \subset \mathcal{N}_{\epsilon/2}(\lambda(a,b)) \cap \mathcal{N}_{\epsilon/2}(\lambda(a',b'))$,
(c) $|\mathcal{Y} \triangle \mathcal{Y}'| < N$.

(2) Given an ϵ-setup $(a, b; \mathcal{Y})$, an *edge decomposition* of its stable interval $T = T_e \cup T_c$ is a collection of subintervals $T_s \subset T_e$.

(3) Given $L > 0$, $\mathcal{Y}_0 \subset \mathcal{Y} \cap \mathcal{Y}'$, and two (N, ϵ)-admissible setups $(a, b; \mathcal{Y})$ and $(a', b'; \mathcal{Y}')$, we say that two edge decompositions $T_s \subset T_e$ and $T'_s \subset T'_e$ are $\mathcal{Y}_0$-*stably* L-*compatible* if
   (a) Each component of T_s and T'_s has positive integer length with endpoints at vertices of T_e and T'_e, respectively.
   (b) There is an order-preserving bijection $\alpha : \pi_0(T_s) \to \pi_0(T'_s)$ between the sets of stable components.
   (c) For each *stable pair* (E, E') identified by α, there exists an order-preserving isometry $i_{E,E'} : E \to E'$.
   (d) For all but at most L pairs of stable components (E, E'), we have $\phi(E) = \phi(E')$ are identical with $\phi(x) = \phi'(i_{E,E'}(x))$ for all $x \in E$.
   (e) For the (at most) L-many remaining stable pairs (E, E'), we have $d_{\mathcal{Z}}(\phi(x), \phi'(i_{E,E'}(x))) < L$ for all $x \in E$.
   (f) The complements $T_e - T_s$ and $T'_e - T'_s$ consist of at most L-many *unstable components* of diameter at most L.
   (g) The induced order-preserving bijection $\beta : \pi_0(T - T_s) \to \pi_0(T' - T'_s)$ satisfies:
      (i) (Endpoints) Let D_a denote the component of $T - T_s$ containing $\mu(C_a)$, and define $D_b, D_{a'}, D_{b'}$ similarly. Then $\beta(D_a) = D_{a'}$ and $\beta(D_b) = D_{b'}$.
      (ii) (Identifying clusters) For any $y \in \mathcal{Y}_0$, let D_y, D'_y denote the components of $T - T_s$ and $T' - T'_s$ containing $\mu(C_y), \mu(C'_y)$, respectively. Then $\beta(D_y) = D'_y$.

Remark 8.13. The above definition is a refinement of the properties contained in [DMS20, Stable Tree Theorem 3.2], where we have set up these stable decompositions to plug nicely into the cubical model machinery described earlier in the paper. How this process works is described below in more detail in Subsection 9.3, but for now we motivate some parts of the definition. Roughly speaking, the points $\mathcal{Y}$ and $\mathcal{Y}'$ stand in for the relative projection data on each interval, with each component of $T - T_s$ and $T' - T'_s$ being labeled by some such point. Eventually, we will collapse each of the components of $T - T_s$ and $T' - T'_s$ to points. Item (3g) says that the common labels $\mathcal{Y} \cap \mathcal{Y}'$ are identified, as well as the endpoint pairs a, a' and b, b'. The rest of the statement now says that the components of T_s and T'_s are in bijection with each other, where identified components are isometric, and the order-preserving properties explain how to glue these components to obtain an isometry between collapsed intervals. Finally, the proximity properties in items (3d) and (3e) are necessary for controlling the maps back into the ambient HHS; see the proof of Theorem 9.9. A more general version can be found in [DMS25, Definition 10.18].

8.5. **The stable interval theorem.** With Definition 8.12 in hand, we can now state the main result of this section, Theorem 8.14 below. It says that the stable intervals for a pair of admissible ϵ-setups admit compatible stable decompositions—that is, decompositions into subintervals which are in bijective correspondence with

idenitified pairs uniformly close, up to ignored a bounded collection of bounded-length subintervals. We first state this technical result, and then prove a corollary which gives the main upshot for our purposes of building stable cubical models.

For a bit of setup, recall the notion of a thickening of an interval from Subsection 4.4. This was a way of taking an interval with a decomposition $T = A \cup B$ into collections of segments, and expanding and combining one collection of the segments. Our stable intervals come with such a decomposition $T = T_e \cup T_c$, and we will always mean a thickening of a stable interval T to be a thickening of T along the components of the cluster forest T_c. Finally, we denote the components of the resulting thickening by $T = \mathbb{T}_e \cup \mathbb{T}_c$, and we note that $\mathbb{T}_e \subset T_e$ and $T_c \subset \mathbb{T}_c$.

Theorem 8.14 (Stable intervals)**.** *Let $\mathcal{Z}$ be δ-hyperbolic and geodesic. For any $\epsilon, N > 0$ and positive integers $r_1, r_2 > 0$, there exist $L_1 = L_1(\delta, \epsilon, N) > 0$ and $L_2 = L_2(\delta, \epsilon, N, r_1, r_2) > 0$ so that the following holds. Suppose $(a, b; \mathcal{Y})$ and $(a', b'; \mathcal{Y}')$ are (N, ϵ)-admissible ϵ-setups and let $T = T_e \cup T_c$ and $T' = T'_e \cup T'_c$ denote their stable intervals. Then*

1. *There exist $(\mathcal{Y} \cap \mathcal{Y}')$-stable L_1-compatible decompositions $T_s \subset T_e$ and $T'_s \subset T'_e$, and*
2. *There exist $(\mathcal{Y} \cap \mathcal{Y}')$-stable L_2-compatible edge decompositions $\mathbb{T}_s \subset \mathbb{T}_e$ and $\mathbb{T}'_s \subset \mathbb{T}'_e$ of the (r_1, r_2)-thickenings of T, T' along T_c, T'_c.*
   - *Moreover, we have $\mathbb{T}_s \subset T_s$ and $\mathbb{T}_s \subset T'_s$.*

Remark 8.15. In the HHS setting, we will be able to control each of the constants $\delta, \epsilon, N, r_1, r_2$ in terms of the ambient HHS structure. We have written the statement with two conclusions because the first conclusion is a mild reformulation of the original theorem [DMS20, Theorem 3.2], while the second is what we actually need for the cubulation machine discussed in this paper. Notably, the second statement will be an easy consequence of the first.

Before we move onto the proof, we observe the following corollary, whose statement and proof motivate the statement of Theorem 8.14. In particular, this corollary will allow us to verify the interval-wise condition in Proposition 6.12 and give us the cubical isomorphism we need for the Stable Cubulations Theorem 9.9; see Subsection 9.3.

Corollary 8.16. *Let $(a, b; \mathcal{Y})$ and $(a', b'; \mathcal{Y}')$ be (N, ϵ)-admissible ϵ-setups with $(\mathcal{Y} \cap \mathcal{Y}')$-stable L-compatible decompositions $T_s \subset T_e$ and $T'_s \subset T'_e$. Let $\Delta : T \to \widehat{T}$ and $\Delta' : T' \to \widehat{T}'$ denote the quotients obtained by collapsing each component of $T - T_s$ and $T' - T'_s$ to a point. The following hold:*

1. *$\widehat{T}$ and $\widehat{T}'$ are simplicial intervals where each collapsed component of $T - T_s$ and $T' - T'_s$ is a vertex, and*
2. *There exists an isometry $\Phi : \widehat{T} \to \widehat{T}'$ which is induced from the bijections $\alpha : \pi_0(T_s) \to \pi_0(T'_s)$ and $\beta : \pi_0(T - T_s) \to \pi_0(T' - T'_s)$, and the isometries of pairs of stable components $i_{E,\alpha(E)} : E \to \alpha(E)$ for $E \subset T_s$. Moreover, we have*
   (a) *$\Phi(\Delta(a)) = \Delta'(a')$ and $\Phi(\Delta(b)) = \Delta'(b')$,*
   (b) *If $y \in \mathcal{Y} \cap \mathcal{Y}'$ and $D_y \in \pi_0(T - T_s)$ and $D'_y \in \pi_0(T - T_s)$ contain $\mu(C_y)$ and $\mu(C'_y)$, respectively, then $\Phi(\Delta(D_y)) = \Delta'(D'_y)$.*

Proof. First, observe that $\widehat{T}$ and $\widehat{T}'$ are intervals because T, T' are and each component of $T - T_s$ and $T' - T'_s$ is a subinterval. They are simplicial intervals with the

given description because Theorem 8.14 (specifically item (4a) of Definition 8.12) provides that each component of T_s, T'_s has positive integer length, and so the distance between the points corresponding to pairs of collapsed components of $T - T_s$ and $T' - T'_s$ is a positive integer.

For the second conclusion, we can define $\Phi : \widehat{T} \to \widehat{T}'$ as follows: For each vertex $v \in \widehat{T}$ corresponding to a component $C \in \pi_0(T - T_s)$, we define $\Phi(v) = v'$, where v' is the vertex of $\widehat{T}'$ corresponding to $\beta(C) \in \pi_0(T' - T'_s)$. If now $x \in \widehat{T}$ is not such a vertex, then $x \in \Delta(E)$ for some component $E \in \pi_0(T_s)$. So we define $\Phi(x) = \Delta'(i_{E,\alpha(E)}(x))$. The fact that these piece-wise maps glue coherently is a direct consequence of the fact that α is order-preserving. The subitems of (2) are direct consequences of item (3g) of Definition 8.12. This completes the proof. □

Remark 8.17. Notably, the above corollary holds whenever we have a stable decomposition, so in particular when this is a stable decomposition arising from the thickened versions of our stable intervals.

8.6. **Detailed sketch of the proof.** In lieu of a full proof (see Appendix B), we provide a detailed sketch of Theorem 8.14 with figures illustrating the key technical arguments, which are easier to understand at a fairly close level than to write down carefully. Our proof strategy here follows more closely the proof of [DMS25, Theorem 10.23] than the original [DMS20, Theorem 3.1].

The goal is to show that the stable intervals (Definition 8.6) for pairs of admissible ϵ-setups admit stable decompositions in the sense of Definition 8.12. Roughly, these are decompositions of the intervals into collections of subintervals, which are either identical (as in item (d)), or the same length and uniformly close (as in item (e)), or are one of boundedly-many unstable components we can ignore (as in item (f)). We need to show how to constraints on the input data assumed in item (1) translates into control over the components of the stable intervals.

To begin the sketch, first observe that the two main assumptions of Theorem 8.14 on the setups $(a, b; \mathcal{Y})$ and $(a', b'; \mathcal{Y}')$ are that

- The sets of cluster points $\mathcal{Y}, \mathcal{Y}'$ are close to both $\lambda(a, b)$ and $\lambda(a', b')$ and have bounded symmetric difference, and
- The endpoints pairs (a, b) and (a', b') are within ϵ of each other.

This suggests the following iterated proof strategy:

(1) First, show that we can build a stable decomposition when we add a single nearby cluster point, while keeping the endpoints the same.
(2) Next, show that we can iteratively combine stable decompositions to obtain a stable decomposition.
(3) Finally, show that we can build a stable decomposition when we only wiggle a single endpoint and keep everything else (including the cluster sets) the same.

Steps (1) and (3) can be thought of as *moves*, where we only change one piece of data at a time, and we will see that type (3) moves are basically type (1) moves. For step (2), we will want to iteratively perform type (1) moves by adding the extra cluster points of $\mathcal{Y}' - \mathcal{Y}$ to the setup $(a, b; \mathcal{Y})$, thereby shifting from $(a, b; \mathcal{Y})$ to $(a, b; \mathcal{Y} \cup \mathcal{Y}')$. Then we make two type (3) moves to shift $(a, b; \mathcal{Y} \cup \mathcal{Y}') \to (a', b; \mathcal{Y} \cup \mathcal{Y}') \to (a', b'; \mathcal{Y} \cup \mathcal{Y}')$. Finally, we make another sequence of iterated type (1) moves to shift from $(a', b'; \mathcal{Y} \cup \mathcal{Y}')$ down to $(a', b'; \mathcal{Y}')$. Since this total sequence of moves is bounded by $N + 2$, we will be done.

For step (1), there are two main cases—the *split cluster* case, where we add a point z to $\mathcal{Y}$ far from all $\mathcal{Y}$-clusters; and the *affected cluster* case, where z is close to some existing $\mathcal{Y}$-cluster. See Figures 7 and 8, respectively, for relevant schematics. The philosophy here is simple: adding a cluster point only affects things locally, thus one only needs to make bounded local modifications.

Step (2) of the proof now involves iteratively applying step (1), and thus one must show that one can coherently combine two pairs of stable decompositions across a common setup. Figure 9 shows a schematic for the argument, in which one has three stable trees T_1, T_2, T_3 with associated stable decompositions, so that those for T_1, T_2 and T_2, T_3 are compatible. The idea is to intersect the two decompositions on T_2 and pull them back to T_1, T_3.

The associated stability constants increase a controlled amount with each iteration, but we are assuming that we only need to perform boundedly-many such iterations in Theorem 8.14.

Finally, step (3) is an application of steps (1) and (2). To get the idea, we can consider the following restricted setup, which we highlight to focus on what happens when we perturb an endpoint. Let $a, b, a' \in \mathcal{Z}$ and $\mathcal{Y} \subset \mathcal{Z}$ be finite with $\mathcal{Y} \subset \mathcal{N}_{\epsilon/2}(\lambda(a,b)) \cap \mathcal{N}_{\epsilon/2}(a',b)$. The following collection of setups, each of which differs by adding or deleting one cluster point, or by switching the role that a, a' or b, b' play between cluster or endpoint:

(1) $(a, b; \mathcal{Y})$,
(2) $(a, b; \mathcal{Y} \cup \{a\})$,
(3) $(a, b; \mathcal{Y} \cup \{a, a'\})$,
(4) $(a', b; \mathcal{Y} \cup \{a, a'\})$,
(5) $(a', b; \mathcal{Y} \cup \{a'\})$,
(6) $(a', b; \mathcal{Y})$.

Hence we can use steps (1),(2), (3) plus iteration to change the roles of endpoints once we have added all of the necessary cluster points. See Appendix B for the full argument.

FIGURE 7. Step (1) of the proof of Theorem 8.14, split cluster case: In the first figure, the two edge components of the two stable intervals are shown. The point z forms an individual cluster and therefore must be inserted as a vertex in the graph $\mathcal{G}$, between clusters C_1, C_2. Every cluster to the left of C_1 is unchanged, as are all of the connecting segments (in pink), and same for those to the right of C_2. The second figure shows the stable decomposition of the stable intervals. Here, we need to include the shadow of the cluster z on $\lambda(C_1, C_2)$ as an unstable component. This cuts $\lambda(C_1, C_2)$ into two segments, which are paired with $\lambda(C_1, z)$ and $\lambda(z, C_2)$, though these need to be slightly truncated so guarantee they have the same lengths; the unstable segments accomplishing this are red. The third figure shows the stable decomposition for their (r_1, r_2)-thickenings, which are subsets of the edge components $T_e \subset T$ and $T'_e \subset T$, are indicated in brown. Note that this can result in collapsing an entire edge component if its diameter is less than $r_2 - 2r_1$. These brown segments refine the stable decompositions $T_s \subset T_e$ and $T'_s \subset T'_e$.

FIGURE 8. Step (1) of the proof of Theorem 8.14, affected cluster case. Two phenomena are shown. In the top figure, there are three affected clusters, P, Q, R, for $(a, b; \mathcal{Y})$ which combine with z into a new cluster C' for $(a, b; \mathcal{Y} \cup \{z\})$. By definition, P, Q, R must be $O(\epsilon)$ apart, and actually are adjacent in $\mathcal{G}$. The clusters C_1, C_2 appear in both setups, and the subgraph of $\mathcal{G}$ to the left of C_1 coincides with the same subgraph for $\mathcal{G}'$, and similarly for C_2. So the corresponding (pink) components form (stable) components of $T_s \subset T_e$ and $T'_s \subset T'_e$. The components of T_e connecting P, Q, R are declared unstable and, in this example, the components $\lambda(C_1, P)$ and $\lambda(R, C_2)$ of T_e coincide with the components $\lambda(C_1, C')$ and $\lambda(C', C_2)$ of T'_e, and hence are stable. The second figure shows the stable decomposition for the (r_1, r_2)-thickenings, which are just the thickenings of the components of $T_s = T'_s$, and hence are also exactly the same. In the second example, the same discussion applies to the outside parts. For the middle of the graph, z combines with the cluster A, which results in pairs of segments $\lambda(C_1, A)$ and $\lambda(C_1, C')$, and $\lambda(A, C_2)$ and $\lambda(C', C_2)$, which have $O(\epsilon)$-close endpoints and are thus $O(\epsilon)$-Hausdorff close, with the same length up to error $O(\epsilon)$. Small surgeries provide the required stable components for the stable intervals T, T', shown in the third figure. The fourth figure shows the stable decomposition for the (R, r)-thickenings of this example.

FIGURE 9. The proof of Theorem 8.14, iterating stable decompositions: The idea for the iteration argument is to intersect the stable decompositions $T^1_{s,2} \cap T^3_{s,2} \subset T_2$ for T_1, T_3, respectively, and then use the isometries which identify the stable pairs for T_1, T_2 to give a refined stable decomposition for (T_1, T_2), which can be combined with a refinement for (T_2, T_3) via simple composition. In the schematic (which is *not* happening in $\mathcal{Z}$), we have lined up the stable intervals T_1, T_2, T_3 to indicate how to refine the pairwise stable decompositions for (T_1, T_2) and (T_2, T_3) for one for (T_1, T_3). The orange segments correspond to cluster components of the intervals, while the blue and green segments correspond to the unstable components for the pairs (T_1, T_2) and (T_2, T_3), respectively. The complement of the orange, blue, and green segments on T_2 form the refined stable decomposition on T_2, and the purple squares indicate how these segments induce refined stable decompositions on T_1 and T_3.

9. Stable cubical models

In this section, we explain how to combine the cubulation machinery described in Sections 4–6 with the Controlling Domains Theorem 7.4 established in Section 7 and the Stable Interval Theorem 8.14 from Section 8. Over the first two subsections, we prove that our stable intervals can be plugged into our interval system machinery to produce an HFI for any pair of points $a, b \in \mathcal{X}$ in an HHS.

The heart of the section is then explaining how to relate two such HFIs for a pair of points a, b and a perturbed pair a', b'. We explain this below in detail in Subsection 9.3.

9.1. Stable intervals form an interval system.

Our first goal is to explain how to take the family of stable intervals associated to a pair of points $a, b \in \mathcal{X}$ and produce a hierarchical family of intervals.

First, we need to induce a collection of relative projections on the stable intervals which turn them into an R-interval system (Definition 4.3) for a uniform $R = R(\mathfrak{S}) > 0$.

Let $U \in \mathcal{U}$. Recall that the BGIA 3.11 says that if $d_V(a, b)$ is sufficiently large for $V \in \mathfrak{S}$, then $\rho^V_U \subset \mathcal{N}_{E_{\mathfrak{S}}}(\lambda(a, b))$ when $V \sqsubset U$ and, even better, $\rho^V_U \subset \mathcal{N}_{E_{\mathfrak{S}}}(a \cup b)$ when $V \pitchfork U$. If we set $\mathcal{Y}_U = \{\rho^V_U | V \sqsubset U \in \mathcal{U}\}$, then $(\pi_U(a), \pi_U(b); \mathcal{Y}_U)$ is an ϵ-setup in $\mathcal{C}(U)$ for $\epsilon = 2E_{\mathfrak{S}}$ in the sense of Subsection 8.1. In particular, our stable interval construction (Proposition 8.9) applies.

For each $U \in \mathcal{U}$, let $T_U = T_{U,e} \cup T_{U,c}$ denote the stable interval for the setup $(a, b; \mathcal{Y}_U)$ in $\mathcal{C}(U)$. Recall that each component of $T_{U,c}$ coincides with $\mu(C)$ for some cluster C of points in $\mathcal{Y}_U \cup \{a, b\}$ (Subsection 8.2). Importantly, item (2) of Proposition 8.9 says that $\phi_U(\mu(C))$ is $O(\epsilon)$ (and hence $O(\mathfrak{S})$) Hausdorff close to C itself. This property is what makes the following definition work:

Definition 9.1 (Stable interval projections)**.** For each $U \in \mathcal{U}$, we define the following *stable interval projections*:

- For each $V \in \mathcal{U}$ with $V \sqsubset U$, let C be the cluster containing ρ_U^V. Define
$$\delta_U^V = \phi_U^{-1}(\mathrm{hull}_{\mu(C)}(p_{\phi_U(\mu(C))}(\rho_U^V))) \subset \mu(C),$$
where $p_U : \mathcal{C}(U) \to \phi_U(\mu(C))$ is the closest point projection to the geodesic segment $\phi_U(\mu(C)) = \mu(C)$.
- For each $V \in \mathcal{U}$ with $V \pitchfork U$, let $f \in \{a, b\}$ denote the point (provided by Definition 3.1) for which $d_U(\pi_U(f), \rho_U^V) < \epsilon$, and let C denote the cluster containing f. Again we define $\delta_U^V = \phi_U^{-1}(\mathrm{hull}_{\mu(C)}(p_{\phi_U(\mu(C))}(\rho_U^V))) \subset \mu(C)$.
- If $U \sqsubset V$, then we define the *interval projection map* $\delta_U^V : T_V \to T_U$ by $\delta_U^V = \phi_U^{-1} \circ p_{T_U} \circ \rho_U^V \circ \phi_V$.

The next lemma says that these projections turn the family of stable intervals into an interval system. The proof, which we leave to the reader, involves verifying the conditions of Definition 4.3 (and hence Lemma 4.2, and one does this via Lemma 3.10 and Proposition 8.9.

Lemma 9.2. *There exists $R_s = R_s(\mathfrak{S}) > 0$ so that the stable intervals $\{T_U\}_{U \in \mathcal{U}}$ endowed with the stable interval projections forms an R_s-interval system for a, b.*

9.2. **Thickenings of stable intervals, their collapses, and HFIs.** In this subsection, we explain why collapsed thickened interval systems of stable intevals form HFIs.

To begin, in Subsection 4.4, we took an interval system $\{T_U\}_{U \in \mathcal{U}}$ along with its relative projection data δ_U^V, and then thickened the intervals along the components B of a decomposition $T_U = A \cup B$, where B consists of these δ_U^V along with the marked points a_U, b_U. Then in Subsection 4.5, we collapsed along this thickening which induced collapsed relative projections on the collapsed intervals. On the other hand, the thickenings of stable intervals we considered in the Stable Intervals Theorem 8.14 were thickenings along the cluster components.

The following lemma, whose proof we leave to the reader, says that these thickenings coincide, and hence that the corresponding collapsed trees are identical:

Lemma 9.3. *For any $U \in \mathcal{U}$, let $T_U = A \cup B$ be the decomposition where B is the union of all δ_U^V for $V \sqsubset U$ along with a_U, b_U. Let $T_U = T_{U,e} \cup T_{U,c}$ denote the decomposition provided by Proposition 8.9. There exists $R_3 = R_3(\mathfrak{S}) > 0$ so that for any $r_1 > R_3$ and $r_2 > 0$, the (r_1, r_2)-thickenings of T_U along B and along $T_{U,c}$ are identical.*

The upshot of Lemmas 9.2 and 9.3 is that we can plug our stable intervals directly into our interval system machinery. The next step is to explain how to produce an HFI (Definition 6.1) from this system using the thickenings and collapsing procedures discussed in Subsections 4.4 and 4.5.

Continuing with our notation, Lemma 9.2 says that the stable intervals $\{T_U\}_{U \in \mathcal{U}}$ endowed with the interval projections via Definition 9.1 form an R_s-interval system for $a, b \in \mathcal{X}$, where $R_s = R_s(\mathfrak{S}) > 0$. Since $R_s = R_s(\mathfrak{S}) > 0$, Proposition 4.9 and Lemma 4.10 provide constants $R_1 = R_1(\mathfrak{S}) > 0$ and $R_2(\mathfrak{S}) > 0$ so that for any $r_1 \geqslant R_1$ and $r_2 \geqslant R_2$, the conclusions of Lemma 4.10 hold for the family of collapsed intervals $\{\widehat{T}_U\}_{U \in \mathcal{U}}$ obtained from the (r_1, r_2)-thickening of the system $\{T_U\}_{U \in \mathcal{U}}$. We also want to assume that $r_1 > R_3 = R_3(\mathfrak{S}) > 0$ so that Lemma 9.3 holds.

The next proposition states that as long as r_1 is a positive integer, the collapsed intervals $\widehat{T}_U$ are all simplicial intervals, where the collapsed points lie at vertices.

The proof is a straightforward application of the ideas in the preceding sections, so we leave it to the reader.

Proposition 9.4. *As long as $r_1 \geqslant \max\{R_1, R_3\}$ is a positive integer and $r_2 \geqslant R_2$, the family of collapsed stable intervals $\{\widehat{T}_U\}_{U\in\mathcal{U}}$ associated to an (r_1, r_2)-thickening of $\{T_U\}_{U\in\mathcal{U}}$ along $T_{U,c}$ and endowed with their collapsed projections forms a hierarchical family of intervals.*

We note the following immediate corollary of Proposition 9.4 and Theorem 6.4:

Corollary 9.5. *Let $a, b \in \mathcal{X}$, $K > K_0$ for K_0 as in Theorem 3.8 and $\mathcal{U} = \mathrm{Rel}_K(a, b)$. Let $\{\widehat{T}_U\}_{U\in\mathcal{U}}$ be the associated family of stable intervals. Then the hierarchical hull $H = \mathrm{hull}_{\mathcal{X}}(a, b)$ is quasi-isometric to the dual cube complex of the associated HFI, with constants depending only on $\mathfrak{S}$ and K.*

9.3. **From a perturbed pair to a family of stable decompositions.** In this subsection, we explain how the basic setup of this paper can be plugged into the machinery we have built so far.

For the rest of this section, assume that we have a colorable HHS $(\mathcal{X}, \mathfrak{S})$ endowed with the stable projections from Theorem 7.4. Let $K > K_0 = K_0(\mathfrak{S}) > 0$ be as in item (2) of that theorem.

Also for the rest of this section, fix $a, b, a', b' \in \mathcal{X}$ so that $d_{\mathcal{X}}(a, a'), d_{\mathcal{X}}(b, b') \leqslant 1$. We refer to the points a', b' as a *perturbed pair* from a, b. Let $\mathcal{U} = \mathrm{Rel}_K(a, b)$ and $\mathcal{U}' = \mathrm{Rel}_K(a', b')$. Let $\{T_U\}_{U\in\mathcal{U}}$ and $\{T'_U\}_{U\in\mathcal{U}'}$ denote the families of stable intervals produced as above.

The first thing to observe is that item (2)(a) of Controlling Domains Theorem 7.4 provides that

(1) $|\mathcal{U}\triangle\mathcal{U}'| < N_0 = N_0(\mathfrak{S})$, and
(2) Each domain $V \in \mathcal{U}\triangle\mathcal{U}'$ satisfies $d_V(a, b) \leqslant K'$ and $d_V(a', b') \leqslant K'$, where $|K - K'|$ is bounded in terms of $\mathfrak{S}$.

Item (2) plus Lemma 4.10 says that each of the associated collapsed intervals $\widehat{T}_V$ or $\widehat{T}'_V$ for $V \in \mathcal{U}\triangle\mathcal{U}'$ has bounded diameter. For the rest of the domains, namely the domains $U \in \mathcal{U}_0 = \mathcal{U} \cap \mathcal{U}'$, we want to obtain compatible stable decompositions between the intervals T_U and T'_U, which we do as follows.

First, observe that if $V \sqsubset U \in \mathcal{U}_0$, then $d_V(a, b) \geqslant K > E_{\mathfrak{S}}$ and $d_V(a', b') \geqslant K > E_{\mathfrak{S}}$, so the BGIA 3.11 implies that $\rho^V_U \subset \mathcal{N}_{E_{\mathfrak{S}}}(\lambda_U(a, b)) \cap \mathcal{N}_{E_{\mathfrak{S}}}(\lambda_U(a', b'))$, where λ_U is the function which assigns a geodesic in $\mathcal{C}(U)$ to any pair of points from Subsection 8.1. Hence taking $\epsilon = 2E_{\mathfrak{S}}$ in our stable tree setup suffices to make $(\pi_U(a), \pi_U(b); \mathcal{Y}_U)$ and $(\pi_U(a'), \pi_U(b'); \mathcal{Y}'_U)$ into ϵ-setups in $\mathcal{C}(U)$ (item (1) of Definition 8.12).

Moreover, for each $U \in \mathcal{U}_0$, Theorem 7.4 provides that

(a) The ρ-sets for the domains in $\mathcal{U}$ and $\mathcal{U}'$ which nest into U—namely $\mathcal{Y}_U$ and $\mathcal{Y}'_U$ respectively—satisfy $|\mathcal{Y}_U\triangle\mathcal{Y}'_U| < N_0 = N_0(\mathfrak{S})$, and
(b) If $\mathcal{U}_0^* \subset \mathcal{U}_0$ is the set of domains $U \in \mathcal{U}_0$ for which $\mathcal{Y}_U \neq \mathcal{Y}'_U$ or $\pi_U(a) \neq \pi_U(a')$ or $\pi_U(b) \neq \pi_U(b')$, then $\#\mathcal{U}_0^* < N_0$.

Hence, by item (a), these ϵ-setups are (N_0, ϵ)-admissible in the sense of item (2) of Definition 8.12 for all domains in $\mathcal{U}_0$, while item (b) says that for domains in $\mathcal{U}_0 - \mathcal{U}_0^*$, these ϵ-setups exactly coincide. Thus item (2) of the Stable Intervals Theorem 8.14 implies that the (r_1, r_2)-thickenings of their stable intervals T_U, T'_U admit L_s-compatible stable decompositions, where $L_s = L_s(\mathfrak{S}) > 0$ by choosing

$r_1 = \max\{R_1, R_3\}$ and $r_2 = R_2$, since the constants $R_i = R_i(\mathfrak{S}) > 0$ depend only on $\mathfrak{S}$. In particular, we have:

Proposition 9.6. *For each $U \in \mathcal{U}_0$, the (r_1, r_2)-thickenings of the stable intervals $T_U = T_{U,e} \cup T_{U,c}$ and $T'_U = T'_{U,e} \cup T'_{U,c}$ along $T_{U,c}, T'_{U,c}$ respectively admit L_s-compatible stable decompositions, where $L_s = L_s(\mathfrak{S}) > 0$.*

- *Moreover, if $U \in \mathcal{U}_0 - \mathcal{U}_0^*$, then $T_U = T'_U$ and these thickenings coincide exactly.*

For each $U \in \mathcal{U}_0$, let $\Delta_U : \widehat{T}_U \to \widehat{T}_{U,0}$ and $\Delta'_U : \widehat{T}'_U \to \widehat{T}'_{U,0}$ denote the maps which collapse the unstable components these decompositions to points; we note that these maps are the identity when $U \in \mathcal{U}_0 - \mathcal{U}_0^*$. By Definition 8.12 of a stable decomposition, the subintervals of the $\widehat{T}_U$ and $\widehat{T}'_U$ being collapsed for $U \in \mathcal{U}_0$ along with the whole intervals $\widehat{T}_V, \widehat{T}'_V$ (whichever is defined) for $V \in \mathcal{U} \triangle \mathcal{U}'$ collectively form $(L_s^2 + 2K'N_0)$-admissible simplicial trimming setups for $\{\widehat{T}_U\}_{U \in \mathcal{U}}$ and $\{\widehat{T}'_U\}_{U \in \mathcal{U}'}$, in the sense of Definition 6.5. This is because the set $(\mathcal{U} \cup \mathcal{U}') - \mathcal{U}_0$ has at most N_0-many domains by (1) above, each of which has diameter bounded by K', while the unstable pieces for the domains in $\mathcal{U}_0$ consist of at most L_2-many pieces of diameter at most L_2.

Set $L_2^2 + 2K'N_0 = S_0$, which we note depends only on $\mathfrak{S}$ and our choice of K. We now obtain the following from Theorem 6.7 and Proposition 9.6:

Proposition 9.7. *The following hold:*

1. *The collections $\{\widehat{T}_{U,0}\}_{U \in \mathcal{U}_0}$ and $\{\widehat{T}'_{U,0}\}_{U \in \mathcal{U}_0}$ are HFIs. Let $\mathcal{Q}_0, \mathcal{Q}'_0$ denote their corresponding 0-consistent subsets.*
2. *The combined maps $\Delta : \mathcal{Q} \to \mathcal{Q}_0$ and $\Delta' : \mathcal{Q}' \to \mathcal{Q}'_0$ satisfy $\Delta(\mathcal{Q}) = \mathcal{Q}_0$ and $\Delta'(\mathcal{Q}') = \mathcal{Q}'_0$ and are both $(1, S_0^2)$-quasi-isometries.*
3. *There exists maps $\Xi : \mathcal{Q}_0 \to \mathcal{Q}$ and $\Xi' : \mathcal{Q}'_0 \to \mathcal{Q}'$ with $\Delta \circ \Xi = id_{\mathcal{Q}_0}$ and $\Delta' \circ \Xi' = id_{\mathcal{Q}'_0}$.*
4. *At the level of cube complexes, the maps $\Delta\mathcal{Q} \to \mathcal{Q}_0$ and $\Delta' : \mathcal{Q}' \to \mathcal{Q}'_0$ delete at most S_0^2-many hyperplanes.*

Using this proposition, we can define maps $\widehat{\Omega}_0 : \mathcal{Q}_0 \to \mathcal{X}$ and $\widehat{\Omega}'_0 : \mathcal{Q}'_0 \to \mathcal{X}$ by $\widehat{\Omega}_0 = \widehat{\Omega} \circ \Xi$ and $\widehat{\Omega}'_0 = \widehat{\Omega}' \circ \Xi'$, where $\widehat{\Omega} : \mathcal{Q} \to \mathcal{X}$ and $\widehat{\Omega}' : \mathcal{Q} \to \mathcal{X}$ are the quasi-isometric embeddings defined Subsection 5.6.

For the final piece of setup, we want to observe that these HFIs $\{\widehat{T}_{U,0}\}_{U \in \mathcal{U}_0}$ and $\{\widehat{T}'_{U,0}\}_{U \in \mathcal{U}_0}$ admit an isomorphism in the sense of Proposition 6.12. For this, we require the following lemma, whose proof is a straightforward application of the definitions and Corollary 8.16.

Lemma 9.8. *For each $U \in \mathcal{U}_0^*$, the isometry $i_U : \widehat{T}_{U,0} \to \widehat{T}'_{U,0}$ provided by Corollary 8.16 satisfies the conditions of Proposition 6.12.*

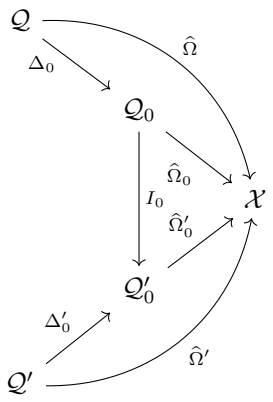
- *Thus these component-wise isometries combine to give an isometry $I_0 : \mathcal{Q}_0 \to \mathcal{Q}'_0$.*

9.4. **Statement of the stable cubulations theorem.** We are now ready to state the stable cubical intervals theorem:

Theorem 9.9 (Stable cubical intervals)**.** *Continuing with our standing notation, there exists $\Sigma = \Sigma(\mathfrak{S}, K) > 0$ so that the following diagram commutes up to error*

Σ:

(10)



- *In the above, $I_0 : \mathcal{Q}_0 \to \mathcal{Q}'_0$ is the isomorphism from Lemma 9.8.*
- *In particular, any colorable hierarchically hyperbolic space is Σ-stably weakly locally quasi-cubical (Definition 2.7).*

Proof. It remains to prove that the upper, lower, and middle triangles coarsely commute. The proof for the upper (and hence the lower) is essentially automatic by construction. In particular, by item (2) of Proposition 9.7, the map $\Delta_0 : \mathcal{Q} \to \mathcal{Q}_0$ is a $(1, S_0^2)$-quasi-isometry for $S_0 = S_0(\mathfrak{S}, K) > 0$. On the other hand, item (3) of that proposition implies that $\Delta_0 \circ \Xi = id_{\mathcal{Q}_0}$ while $\widehat{\Omega}_0 : \mathcal{Q}_0 \to \mathcal{X}$ is defined by $\widehat{\Omega}_0 = \widehat{\Omega} \circ \Xi$. Hence the top triangle commutes up to error S_0, and the same holds for the bottom triangle.

For the middle triangle, we want to show that $d_{\mathcal{X}}(\widehat{\Omega}_0(\hat{x}), \widehat{\Omega}'_0 \circ I_0(\hat{x}))$ is bounded in terms of $\mathfrak{S}, K$ for any $\hat{x} \in \mathcal{Q}_0$. However, we note that the maps $\widehat{\Omega}_0 : \mathcal{Q}_0 \to H$ and $\Delta_0 \circ \widehat{\Psi} : H \to \mathcal{Q}_0$ are quasi-inverses (with errors depending on $\mathfrak{S}, K$), and similarly for $\widehat{\Omega}'_0 : \mathcal{Q}'_0 \to H'$ and $\Delta_0 \circ \widehat{\Psi} : H' \to \mathcal{Q}'_0$. Here, $H = \mathrm{hull}_{\mathcal{X}}(a, b)$ and $H' = \mathrm{hull}_{\mathcal{X}}(a', b')$. But since $d_{\mathcal{X}}(a, a') \leqslant 1$ and $d_{\mathcal{X}}(b, b') \leqslant 1$, it follows from the Distance Formula 3.8 and Definition 3.12 that $d_{\mathcal{X}}^{Haus}(H, H')$ is bounded in terms of $\mathfrak{S}$.

Hence it suffices to show that for $x \in H \cap H'$ (so that $\widehat{\Psi}(x)$ and $\widehat{\Psi}'(x)$ are both defined), we have

$$d_{\mathcal{Q}'_0}(\Delta'_0 \circ \widehat{\Psi}'(x), I_0 \circ \Delta_0 \circ \widehat{\Psi}(x)) < \Sigma,$$

for $\Sigma = \Sigma(\mathfrak{S}, K)$. Set $\Delta'_0 \circ \widehat{\Psi}'(x) = \hat{x}'_0 = (\hat{x}'_{U,0})$ and $I_0 \circ \Delta_0 \circ \widehat{\Psi}(x) = \hat{y}'_0 = (\hat{y}'_{U,0})$, where we note that both $\hat{x}'_0, \hat{y}'_0 \in \mathcal{Q}'_0$.

We argue component-wise that $d_{\mathcal{Q}'_0}(\hat{x}'_0, \hat{y}'_0)$ is bounded. Recall that the map $\widehat{\Psi} : H \to \mathcal{Q}$ is defined component-wise (over the set $\mathcal{U}$) by $\hat{\psi}_U = q_U \circ \phi_U^{-1} \circ p_U \circ \pi_U$, where $\pi_U : \mathcal{X} \to \mathcal{C}(U)$ is the HHS projection, $p_U : \mathcal{C}(U) \to \phi_U(T_U)$ is closest point projection, $\phi_U : T_U \to \mathcal{C}(U)$ is the (L_0, L_0)-quasi-isometric embedding provided by Proposition 8.9 (here $L_0 = L_0(\mathfrak{S})$), and $q_U : T_U \to \widehat{T}_U$ is the map which collapses a uniform thickening of T_U along the cluster components, with constants depending only on $\mathfrak{S}$. The map $\widehat{\Psi}' : H' \to \mathcal{Q}'$ is defined similarly.

First, we can reduce to involved domains in $\mathcal{U}_0 = \mathcal{U} \cap \mathcal{U}'$. If $U \in \mathcal{U} \triangle \mathcal{U}'$, then a, b and a', b' have bounded diameter projections to U in terms of K and $\mathfrak{S}$, while also

each such domain is involved in the sense of Definition 7.3 and hence there are only boundedly-many by Theorem 7.4. Next observe that if $U \in \mathcal{U}_0$ is not involved, then the stable trees T_U, T'_U for the two setups are exactly the same, and hence have identical thickenings, and thus the stable decompositions the pair are the (identical) collection of edge components for each, and their corresponding collapsed trees $\widehat{T}_U$ and $\widehat{T}'_U$ are identical. Hence, for uninvolved $U \in \mathcal{U}_0$, we have $\hat{x}'_{U,0} = \hat{y}'_{U,0}$.

Let $U \in \mathcal{U}_0$ be involved. Since there are boundedly-many (in $\mathfrak{S}$) such domains by Theorem 7.4, it suffices to bound the diameter $d_{\widehat{T}'_{U,0}}(\hat{x}'_{U,0}, \hat{y}'_{U,0})$ in terms of $\mathfrak{S}$.

The distance $d_{\widehat{T}'_{U,0}}(\hat{x}'_{U,0}, \hat{y}'_{U,0})$ is measured in the edge components of $\widehat{T}'_{U,0}$ which separate them, and these edge components are stable components of T'_U. This means that if the distance $d_{\widehat{T}'_{U,0}}(\hat{x}'_{U,0}, \hat{y}'_{U,0})$ is very large, then there must be either

(1) some long stable component $E' \subset T'_{U,s}$ so that some long segment of $\Delta'_U(q'_U(E'))$ separates $\hat{x}'_{U,0}$ from $\hat{y}'_{U,0}$, or
(2) if no such long component exists, there must be some large number of such stable components separating them.

In case (1), let $E \subset T_{U,s}$ be the stable component paired with $E' \subset T'_{U,s}$. Note that $\phi_U(E) \cup \phi'_U(E')$ is uniformly quasi-convex (in $\mathfrak{S}$), and hence $\pi_U(x)$ has a uniformly bounded projection to it. However the fact that a long segment of $\Delta'_U(q_U(E')) \subset \widehat{T}'_{U,0}$ separates $\hat{x}'_{U,0}$ from $\hat{y}'_{U,0}$ means that some large diameter portion of $\phi_U(E) \cup \phi'_U(E')$ coarsely (in $\mathfrak{S}$) separates $p_U(\pi_U(x))$ from $p'_U(\pi_U(x))$, which is impossible.

In case (2), we can assume that no such long stable component exists, so that the distance $d_{\widehat{T}'_{U,0}}(\hat{x}'_{U,0}, \hat{y}'_{U,0})$ corresponds to some sequence $E'_1, \dots, E'_n \subset T'_{U,s}$ of short stable components of length bounded in terms of $\mathfrak{S}$. Let $E_i \subset T_{U,s}$ denote the stable component paired with E'_i. By Proposition 8.9 and the fact that $d_{\mathcal{X}}(a, a') \leqslant 1$ and $d_{\mathcal{X}}(b, b') \leqslant 1$, we have that $\phi_U(T_U)$ and $\phi'_U(T'_U)$ are within bounded (in $\mathfrak{S}$) Hausdorff distance in $\mathcal{C}(U)$. Since each of the E'_i are uniformly close (in $\mathfrak{S}$) to E_i and the bijection $\alpha : \pi_0(T_{U,s}) \to \pi_0(T'_{U,s})$ is order-preserving, it follows that the hulls $\mathrm{hull}_{\phi_U(T_U)}(E_1, E_n)$ and $\mathrm{hull}_{\phi'_U(T'_U)}(E'_1, E'_n)$ are within uniform (in $\mathfrak{S}$) Hausdorff distance of each other, and hence their union is uniformly quasi-convex (in $\mathfrak{S}$). Thus, if $\Delta'_U(q'_U(E'_1)), \dots, \Delta'_U(q'_U(E'_n))$ separate $\hat{x}'_{U,0}$ from $\hat{y}'_{U,0}$ in $\widehat{T}'_{U,0}$, then the union of these hulls coarsely (in $\mathfrak{S}$) separates $p_U(\pi_U(x))$ from $p'_U(\pi_U(x))$.

Since each of the segments E'_i is bounded length, the union of these hulls is long if the number of segments n is very large. Moreover, each pair of segments E'_i, E'_{i+1} is separated by some complementary component $C_i \subset T'_U - T'_{U,s}$. By construction, all but boundedly-many (in $\mathfrak{S}$) of these components corresponds to some cluster component of T'_U. Since there are only boundedly-many (in $\mathfrak{S}$) involved domains, by choosing n to be sufficiently large, there exists some C_j which contains a $\sqsubset_{\mathcal{U}_0}$-minimal domain $V \sqsubset U$ which is both $\sqsubset_{\mathcal{U}}$-minimal and $\sqsubset_{\mathcal{U}'}$-minimal. But this is impossible, for the BGIA 3.11 would imply that $d_V(\rho^U_V(p_U(\pi_U(x)), \rho^U_V(p'_U(\pi_U(x))))) > K$, and this cannot happen as we can also arrange $\pi_U(x)$ to be as far from ρ^V_U as necessary. Hence no such long sequence $E'_1, \dots E'_n$ can exist, completing the proof. □

Appendix A. Proof of the Tree Trimming Theorem 5.14

Below is a re-statement of Theorem 5.14. See Subsection 5.4 for the contextual definitions and notation.

Theorem A.1 (Tree Trimming)**.** *There exists $B_{tt} = B_{tt}(\mathfrak{S}, K) > 0$ so that for any $B < B_{tt}$, there exists $L = L(\mathfrak{S}, B) > 0$ so that the following hold:*

1. *We have $\Delta(\mathcal{Q}) \subset \mathcal{Q}'$.*
2. *There is a map $\Xi : \mathcal{Q}' \to \mathcal{Q}$ so that $\Delta \circ \Xi = id_{\mathcal{Q}'}$. In particular, $\Delta(\mathcal{Q}) = \mathcal{Q}'$.*
3. *The map $\Delta : \mathcal{Q} \to \mathcal{Q}'$ is an (L, L)-quasi-isometry.*

Proof. The fact that $\Delta(\mathcal{Q}) \subset \mathcal{Q}'$, namely that Δ preserves 0-consistency, is immediate from the fact that the Δ_U just collapse subintervals, so coordinates which were collapsed projections remain collapsed projections, showing item (1). Items (2) and (3) require more work and we provide a detailed sketch of both.

Item (2): Surjectivity of $\Delta : \mathcal{Q} \to \mathcal{Q}'$: Let $\hat{y} \in \mathcal{Q}'$. We will construct a point $\hat{x} \in \mathcal{Q}$ with $\Delta(\hat{x}) = \hat{y}$ by a coordinate-wise argument, thereby defining a map $\Xi : \mathcal{Q}' \to \mathcal{Q}$ with $\Delta \circ \Xi = id_{\mathcal{Q}'}$.

For each $U \in \mathcal{U}$, let $C_U = \Delta_U^{-1}(\hat{y}_U)$. Our goal is to pick, for each $U \in \mathcal{U}$, a point $\hat{x}_U \in C_U$ and then prove 0-consistency of the tuple $\hat{x} = (\hat{x}_U)$. Note that $\Delta(\hat{x}) = \hat{y}$ for such a tuple, by construction. The proof is by induction on the nesting level in $\mathcal{U}$, and is similar to the construction of the map $\widehat{\Omega} : \mathcal{Q} \to H$ as in Proposition 5.17. The main difference is that the collapsing subintervals are not automatically associated to a collection of domains, and this is where the work lies. The proof is basically the proof of the analogous fact in [Dur23, Theorem 10.2], except that our setup is much simpler. What follows is a very detailed sketch, which leaves some small bits to the reader.

For the base case, let $U \in \mathcal{U}$ be $\sqsubset_{\mathcal{U}}$-minimal. If $\hat{y}_U \in \widehat{T}_{U,0}$ is not contained in $\Delta_U(B)$ for some $B \in \mathcal{B}$, then $C_U \subset \widehat{T}_U$ is a point and we set $\hat{x}_U = C_U$. If $\hat{y}_U \in \Delta_U(B)$ for some $B \in \mathcal{B}$ and $B \subset \widehat{T}_U$ does not contain an endpoint, then let $\hat{x}_U \in C_U$ be any point. If instead C_U does contain an endpoint, then it is possible that it contains both endpoints, and we need to choose one.

The following claim, which is exactly [Dur23, Claim 10.3], does not require U to be $\sqsubset_{\mathcal{U}}$-minimal:

Claim A.2. If $W, Z \in \{V \in \mathcal{U} | V \pitchfork U, \text{ and } \hat{y}_V \neq (\hat{\delta}_V^U)_0\}$, then $\hat{\delta}_U^W = \hat{\delta}_U^Z$.

With Claim A.2 in hand, we can define $\hat{x}_U$ for $\sqsubset_{\mathcal{U}}$-minimal U as follows:

- If $\mathcal{W}_U^{\pitchfork} = \{V \in \mathcal{U} | V \pitchfork U \text{ and } \hat{y}_W \neq (\hat{\delta}_U^V)_0\} \neq \varnothing$, then we set $\hat{x}_U = \hat{\delta}_U^W$ for any (and hence all, by the claim) $W \in \mathcal{W}_U^{\pitchfork}$.
- If $\mathcal{W}_U^{\pitchfork} = \varnothing$, then we let $\hat{x}_U$ be any marked point in C_U.

At this point, we also want to set l_U to be equal to the endpoint $\hat{x}_U$ when U is $\sqsubset_{\mathcal{U}}$-minimal, where $l_U = \varnothing$ when C_U does not contain an endpoint of $\widehat{T}_U$. These labels will play a similar role as in the cluster honing process in Subsection 5.2.

For the inductive step, suppose that U is not $\sqsubset_{\mathcal{U}}$-minimal and that we have defined $\hat{x}_V$ for all $V \sqsubset U$ and, in particular, for all $\sqsubset_{\mathcal{U}}$-minimal $V \in \mathcal{U}$. As before, if C_U does not contain a marked or a projection $\hat{\delta}_U^V$ for some $V \sqsubset U$, then we let $\hat{x}_U \in C_U$ be any point. If C_U happens to contain one of the endpoints but no $\hat{\delta}_U^V$ for $V \sqsubset U$, then C_U contains only that one endpoint (since U is not $\sqsubset_{\mathcal{U}}$-minimal), and we let $\hat{x}_U \in C_U$ be any point

Otherwise, we set $C = C_U$ and let $\mathcal{V}_U$ denote the set of $\sqsubset_{\mathcal{U}}$-minimal domains $V \sqsubset U$ so that $\hat{\delta}_U^V \in C$ (or equivalently, $(\hat{\delta}_U^V)_0 = \hat{y}_U$). For each $V \in \mathcal{V}_U$, set $C_U^V = \mathrm{hull}_C(\hat{\delta}_U^V \cup \pi_C(\hat{x}_V))$, namely the convex hull in C of $\hat{\delta}_U^V$ and the projections of the marked point labeled by l_V down to the convex subset C. The following claim, which is exactly [Dur23, Claim 10.4], helps us choose a point in such a C_U:

Claim A.3. For every $V, W \in \mathcal{V}_U$, we have $C_U^V \cap C_U^W \neq \varnothing$. Moreover, at least one of $\hat{\delta}_U^V \in C_U^W$ or $\hat{\delta}_U^W \in C_U^V$ holds.

Proof. We may assume that $\hat{\delta}_U^V \neq \hat{\delta}_U^W$, and hence that $V \pitchfork W$. If $C_U^V \cap C_U^W = \varnothing$, then $\hat{\delta}_U^V \notin C_U^W$ and $\hat{\delta}_U^W \notin C_U^V$, and it follows that $l_V \neq l_W$. The BGI property in (7) of Lemma 4.10 plus item (4) of Lemma 4.2 imply that $\hat{\delta}_V^W \notin \Delta_V^{-1}(\hat{y}_V)$ and $\hat{\delta}_W^V \notin \Delta_W^{-1}(\hat{y}_W)$, and hence $\hat{y}_V \neq (\hat{\delta}_V^W)_0$ and $\hat{y}_W \neq (\hat{\delta}_W^V)_0$, which is a contradiction of 0-consistency of $\hat{y}$. This proves the claim. □

Hence the Helly property for trees implies that $\bigcap_{V \in \mathcal{V}_U} C_U^V \subset \widehat{T}_U$ is nonempty. This intersection, nonetheless, may contain many $\hat{\delta}_U^V$ for $V \in \mathcal{V}_U$. The following claim, which is exactly [Dur23, Claim 10.6] helps us make a consistent choice:

Claim A.4. Suppose $V, W \in \mathcal{W}_U^{\sqsubset}$, then $\hat{\delta}_U^V = \hat{\delta}_U^W$.

Proof of Claim A.4. For a contradiction, suppose that $\hat{\delta}_U^W \neq \hat{\delta}_U^V$, so that $V \pitchfork W$. Then again the BGI property in item (7) of Lemma 4.10 implies that $\hat{\delta}_W^V = \hat{\delta}_W^U(\hat{\delta}_U^V)$ and $\hat{\delta}_V^W = \hat{\delta}_V^U(\hat{\delta}_U^W)$. But then $\hat{y}_W \neq (\hat{\delta}_W^V)_0 = (\hat{\delta}_W^U)_0((\hat{\delta}_U^V)_0)$ and $\hat{y}_V \neq (\hat{\delta}_V^W)_0 = (\hat{\delta}_V^U)_0((\hat{\delta}_U^W)_0)$, which contradicts 0-consistency of $\hat{y}$, completing the proof. □

Finally, this last claim helps us make a consistent choice between $\hat{\delta}_U^V$ sets and marked points when C_U contains both:

Claim A.5. The following holds:

(1) If there exists $V \pitchfork U$ so that $\hat{y}_V \neq (\hat{\delta}_V^U)'$, then for all $W \in \mathcal{V}_U$ with $\hat{\delta}_U^W \neq \hat{\delta}_U^V$, we have $\hat{x}_W = \hat{\delta}_W^V$.
(2) If there exists $W \in \mathcal{V}_U$ with $\hat{y}_W \neq (\hat{\delta}_W^U)'(\hat{y}_U)$, then for all $V \pitchfork U$ with $\hat{\delta}_U^W \neq \hat{\delta}_U^V$, we have $\hat{y}_V = (\hat{\delta}_V^U)'$.

- In particular, if $V \in \mathcal{W}_U^{\pitchfork}$ and $W \in \mathcal{W}_U^{\sqsubset}$, then $\hat{\delta}_U^V = \hat{\delta}_U^W$.

Proof of Claim A.5. Suppose $V \pitchfork U$ as in item (1) of the claim, and let $W \in \mathcal{V}_U$ with $\hat{\delta}_U^W \neq \hat{\delta}_U^V$. Then $W \pitchfork V$, and since $W \sqsubset U$, we have $(\hat{\delta}_V^W)_0 = (\hat{\delta}_V^U)_0 \neq \hat{y}_V$, and hence $\hat{x}_W = \hat{\delta}_W^V$, by definition of $\hat{x}_W$.

Now suppose $W \in \mathcal{V}_U$, as in item (2), and that $V \pitchfork U$ with $\hat{\delta}_U^W \neq \hat{\delta}_U^V$. Then $W \pitchfork V$. By 0-consistency of $\hat{y}$, we have that $\hat{y}_U = (\hat{\delta}_U^W)_0$. But since $\hat{\delta}_U^V \neq \hat{\delta}_U^W$, we have $(\hat{\delta}_U^V)_0 \neq (\hat{\delta}_U^W)_0$, and so $\hat{y}_V = (\hat{\delta}_V^U)_0 = (\hat{\delta}_V^W)_0$ since $W \sqsubset U$. This completes the proof. □

We are finally ready to define our tuple $\hat{x}$, which we do domain-wise for each $U \in \mathcal{U}$:

(1) If $\hat{y}_U$ is not a marked or cluster point of $\widehat{T}_{U,0}$, then we let $\hat{x}_U \in C_U$ be any point in $\Delta_U^{-1}(\hat{y}_U)$.
(2) For any U of type (1): If $V \pitchfork U$ or $U \sqsubset V$, we set $\hat{x}_V = \hat{\delta}_V^U$, and if $U \sqsubset V$, we set $\hat{x}_V = \hat{\delta}_V^U(\hat{x}_U)$.

(3) If U is $\sqsubset_{\mathcal{U}}$-minimal and $\hat{y}_U$ is a marked point of $\widehat{T}_{U,0}$, then there are three subcases:
   (a) The set $\mathcal{W}_U^{\pitchfork} = \{W \in \mathcal{U} | W \pitchfork U \text{ and } (\hat{\delta}_W^U)' \neq \hat{y}_W\}$ is nonempty, and we set $\hat{x}_U = \hat{\delta}_U^W$ for any (and hence all by Claim A.2) $W \in \mathcal{W}_U^{\pitchfork}$;
   (b) $\mathcal{W}_U^{\sqsubset} = \{V \sqsubset U | (\hat{\delta}_V^U)'(\hat{y}_U) \neq \hat{y}_V \text{ and } (\hat{\delta}_W^U)'(\hat{y}_U) \neq \hat{y}_W\}$ is nonempty, and we set $\hat{x}_U = \hat{\delta}_U^V$ for any (and hence all by Claim A.4) $V \in \mathcal{W}_U^{\sqsubset}$; or
   (c) $\mathcal{W}_U^{\pitchfork}$ and $\mathcal{W}_U^{\sqsubset}$ are both empty, and we set $\hat{x}_U$ to be any point of $\widehat{T}_U^e$ contained in C_U, in particular a point which is not a cluster or marked point.

(4) Proceeding inductively, if U is not in cases (1)–(3) above, then $\bigcap_{V \in \mathcal{V}_U} C_U^V \neq \varnothing$ and has finite diameter. There are three subcases:
   (a) $\hat{x}_U = \hat{\delta}_U^W$ for any $W \in \mathcal{W}_U^{\pitchfork}$ if $\mathcal{W}_U^{\pitchfork} \neq \varnothing$,
   (b) $\hat{x}_U = \hat{\delta}_U^V$ for any $V \in \mathcal{W}_U^{\sqsubset}$ when $\mathcal{W}_U^{\sqsubset} \neq \varnothing$, or
   (c) Otherwise, both $\mathcal{W}_U^{\pitchfork}$ and $\mathcal{W}_U^{\sqsubset}$ are empty, and we set $\hat{x}_U$ to be any point in $\widehat{T}_U^e \cap C_U$.

Finally, we prove that $\hat{x} = (\hat{x}_U)$ is 0-consistent by a case-wise analysis, similar to the proof of Proposition 5.12. Showing this will complete the proof of surjectivity of $\Delta : \mathcal{Q} \to \mathcal{Q}'$. The proof closely follows the analogous part of the proof from [Dur23, Theorem 10.2].

$\underline{U \pitchfork V}$: We may assume that $\hat{x}_V \neq \hat{\delta}_V^U$, for we are done otherwise. Then $\hat{y}_V \neq (\hat{\delta}_V^U)_0$, and so $\hat{y}_U = (\hat{\delta}_U^V)_0$ by 0-consistency of $\hat{y}$. It follows that $\hat{\delta}_U^V \in C_U = \Delta_U^{-1}(\hat{y}_U)$, and necessarily $\hat{\delta}_U^V \in \{\hat{a}_V, \hat{b}_V\}$, so we may assume without loss of generality that $\hat{\delta}_U^V = \hat{a}_V$. If U is $\sqsubset_{\mathcal{U}}$-minimal, then our choice of $\hat{x}_U$ in (3) says that $\hat{x}_U = \hat{\delta}_U^V$, as required.

Now suppose that U is not $\sqsubset_{\mathcal{U}}$-minimal. Set $\mathcal{V}_U$ to be the set of $\sqsubset_{\mathcal{U}}$-minimal $W \sqsubset U$ with $\hat{\delta}_U^W \in C_U$. Let $W \in \mathcal{V}_U$ be so that $\hat{\delta}_U^W \neq \hat{\delta}_U^V$. Then $W \pitchfork V$ and $(\hat{\delta}_V^W)_0 = (\hat{\delta}_V^U)_0 \neq \hat{y}_V$, and so $\hat{x}_W = \hat{\delta}_W^V = \hat{a}_W$, with the first equality following from our choice in (3) and the last equality by 0-consistency of a. But this says that $a = l_W$, and hence $\hat{a}_U \in C_U^W$ for all $W \in \mathcal{V}_U$ (including those for which $\hat{\delta}_U^W = \hat{\delta}_U^V = \hat{a}_U$).

It follows then that $\hat{a}_U \in \bigcap_{W \in \mathcal{V}_U} C_U^W$ and so $\hat{x}_U = \hat{a}_U$ by definition of $\hat{x}$ in item (4a) above.

$\underline{V \sqsubset U}$: We may assume that $\hat{x}_V \neq \hat{\delta}_V^U(\hat{x}_U)$, for we are done otherwise. Then $\hat{y}_V \neq (\hat{\delta}_V^U)_0(\hat{y}_U)$, and 0-consistency of $\hat{y}$ implies that $\hat{y}_U = (\hat{\delta}_U^V)'$, and thus $\hat{\delta}_U^V \in C_U = \Delta_U^{-1}(\hat{y}_U)$. We claim that for all $W \in \mathcal{V}_U$, we have $\hat{\delta}_U^V \in C_U^W$.

To see this, is suffices to consider $W \in \mathcal{V}_U$ with $\hat{\delta}_U^W \neq \hat{\delta}_U^V$. Then $W \pitchfork V$. But $\hat{x}_V \neq \hat{\delta}_V^U(\hat{x}_U)$, so if also $\hat{\delta}_U^V \notin C_U^W$, then $\hat{x}_V \neq \hat{\delta}_V^U(\hat{\delta}_U^W) = \hat{\delta}_V^W$, while also $\hat{x}_W \neq \hat{\delta}_W^U(\hat{\delta}_U^V) = \hat{\delta}_W^V$, which both use the BGI property in item (7) of Lemma 4.10. But then $\hat{y}_V \neq (\hat{\delta}_V^W)_0$ and $\hat{y}_W \neq (\hat{\delta}_W^V)_0$, which contradicts 0-consistency of $\hat{y}$.

Hence $\hat{\delta}_U^V \in C_U^W$ for all $W \in \mathcal{V}_U$. It follows that $\hat{\delta}_U^V \in \bigcap_{W \in \mathcal{V}_U} C_U^W$, and then (4b) says that $\hat{x}_U = \hat{\delta}_U^V$.

Hence $(\hat{x})$ is 0-consistent. Since $\Delta(\hat{x}) = \hat{y}$ because $\hat{x}_U \in \Delta_U^{-1}(\hat{y}_U)$ for all $U \in \mathcal{U}$, we are done with the proof of surjectivity of $\Delta : \mathcal{Q} \to \mathcal{Q}'$, and this also defines the

map $\Xi : \mathcal{Q}' \to \mathcal{Q}$ by $\Xi(\hat{y}) = \hat{x}$, as defined above. This completes the proof of item (2).

Item (3): $\Delta : \mathcal{Q} \to \mathcal{Q}'$ is a quasi-isometry: The proof of this part proceeds similarly to its analogue in [Dur23, Theorem 10.2], though with some modest simplifications (see especially [Dur23, Subsection 9.1]).

Since $\Delta|_{\mathcal{Q}}$ is surjective and distance non-increasing, it suffices to prove that $d_{\mathcal{Q}} \prec d_{\mathcal{Q}'}$, and for this it suffices to bound the distances in each $U \in \mathcal{U}$.

Let $\hat{x}, \hat{y} \in \mathcal{Q}$ with $\hat{x}' = \Delta(\hat{x})$ and $\hat{y}' = \Delta(\hat{y})$. We may assume that $\hat{x}_U \neq \hat{y}_U$ for any $U \in \mathcal{U}$ that we will consider. We begin by making a useful reduction.

Observe that item (9) of Lemma 4.10 implies that there are at most boundedly-many domains $U \in \mathcal{U}$ for which $\hat{x}_U, \hat{y}_U$ are in the same component of $\widehat{T}^e_U$, where the bound is controlled by $\mathfrak{S}$. For such U, we have $d_{\widehat{T}_U}(\hat{x}_U, \hat{y}_U) \leqslant d_{\widehat{T}'_U}(\hat{x}'_U, \hat{y}'_U) + B^2$ because at most B-many subtrees of diameter B have been deleted from that edge. Hence we may deal with these boundedly-many domains with an additive constant in the distance estimate comparing $d_{\mathcal{Q}}(\hat{x}, \hat{y})$ and $d_{\mathcal{Q}'}(\hat{x}', \hat{y}')$. Hereafter, we assume that at least one of $\hat{x}_U, \hat{y}_U$ coincides with a cluster, or they are separated by at least one cluster.

There are two cases, (1) U is $\sqsubset_{\mathcal{U}}$-minimal and (2) otherwise.

Case (1): Then $\widehat{T}^e_U$ has one component. So there are at most B-many subtrees of diameter at most B in $\widehat{T}^e_U$ being collapsed. By item (9) of Lemma 4.10 there are at most boundedly-many (in $\mathfrak{S}$) such $\sqsubset_{\mathcal{U}}$-minimal U for which one of $\hat{x}_U, \hat{y}_U$ is not a cluster point. We may deal with such U by increasing the additive constant in the distance estimate.

The general case is where $\hat{x}_U, \hat{y}_U$ are the distinct endpoint clusters of the interval $\widehat{T}_U$. In this case,

$$d_{\widehat{T}'_U}(\hat{x}'_U, \hat{y}'_U) \geqslant d_{\widehat{T}_U}(\hat{x}_U, \hat{y}_U) - B^2 \geqslant K - B^2 - 2r_1,$$

by item (6) of Lemma 4.10. Hence by choosing $K = K(\mathfrak{S}) > 0$ sufficiently large and requiring that $B_{tt}^2 - 2r_1 < K$, we can guarantee that $d_{\widehat{T}'_U}(\hat{x}'_U, \hat{y}'_U) > 1$ allowing the estimate

$$d_{\widehat{T}_U}(\hat{x}_U, \hat{y}_U) \leqslant d_{\widehat{T}'_U}(\hat{x}'_U, \hat{y}'_U) + B^2 \leqslant (1 + B^2) d_{\widehat{T}'_U}(\hat{x}'_U, \hat{y}'_U).$$

Thus we can account for these domains via a bounded multiplicative error.

Case (2): Suppose that $U \in \mathcal{U}$ is not $\sqsubset_{\mathcal{U}}$-minimal. Since $\hat{x}_U \neq \hat{y}_U$, there is some collection of edge components $E_1, \dots, E_k \subset \widehat{T}^e_U$ separating $\hat{x}_U$ from $\hat{y}_U$, i.e. intersecting $[\hat{x}_U, \hat{y}_U]_{\widehat{T}_U}$. By assumption, each such component contains at most B-many subintervals in $\mathcal{B}$ of diameter at most B which are collapsed by $\Delta_U : \widehat{T}_U \to \widehat{T}'_U$. Moreover, adjacent intervals E_i, E_{i+1} are attached at some collapsed cluster C_i. For each such collapsed cluster C_i, choose a $\sqsubset_{\mathcal{U}}$-minimal domain $V_i \sqsubset U$ with $\hat{\delta}^{V_i}_U \in C_i$. Observe that

$$\begin{aligned} d_{\widehat{T}_U}(C_i, C_{i+1}) &\leqslant d_{\widehat{T}'_U}(\Delta_U(C_i), \Delta_U(C_{i+1})) + B^2 \\ &\leqslant d_{\widehat{T}'_U}(\Delta_U(C_i), \Delta_U(C_{i+1})) + B^2 d_{\widehat{T}'_{V_i}}(\hat{x}'_{V_i}, \hat{y}'_{V_i}). \end{aligned}$$

Hence we can account for the lost distance along $[\hat{x}_U, \hat{y}_U]$ between each pair of successive collapsed cluster points C_i, C_{i+1} (i.e., the distance lost in the edge E_i) by distances in the $\sqsubset_{\mathcal{U}}$-minimal domains V_i. Similarly, we can use V_1, V_k to account

for any lost distance between $\hat{x}_U, C_1$ and $C_n, \hat{y}_U$, respectively. Since each domain in $\mathcal{U}$ nests into at most boundedly-many domains in $\mathcal{U}$ by the Covering Lemma 3.14, the proof of case (2) is complete. This completes the proof of the theorem. □

Appendix B. Proof of the Stable Intervals Theorem 8.14

We will give a fairly complete accounting of the two main arguments involved in Theorem 8.14. See Subsection 8.6 for a detailed sketch and outline. The two key arguments involve showing that we can stably add one cluster point and then that we can iterate this process.

B.1. **Adding a cluster point.**

In this subsection, we explain how to build a stable decomposition when adding a cluster point to the setup. This is the main bit of work for Theorem 8.14.

Proposition B.1. *Let $\mathcal{Z}$ be a geodesic δ-hyperbolic space, $\epsilon > 0$, and $r_1, r_2 > 0$ be positive integers. Let $(a, b; \mathcal{Y})$ be an ϵ-setup and further fix a point $z \in \mathcal{N}_{\epsilon/2}(\lambda(a, b))$. Then*

1. *$(a, b; \mathcal{Y} \cup \{z\})$ is an ϵ-setup;*
2. *The two setups $(a, b; \mathcal{Y})$ and $(a, b; \mathcal{Y} \cup \{z\})$ are $(1, \epsilon)$-admissible;*
3. *There exists $L_1 = L_1(\delta, \epsilon) > 0$ so that the stable intervals T and T' for the setups $(a, b; \mathcal{Y})$ and $(a, b; \mathcal{Y} \cup \{z\})$ admit $\mathcal{Y}$-stable L_1-compatible decompositions;*
4. *There exists $L_2 = L_2(\delta, \epsilon, r_1, r_2) > 0$ so that the (r_1, r_2)-thickenings of T, T' admit $\mathcal{Y}$-stable L_2-compatible decompositions.*

Proof. Since the points a, b have not changed, the first two conclusions are automatic from the assumptions.

Choosing $E \gg \epsilon$, Proposition 8.5 implies that both $\mathcal{G} = \mathcal{G}_E(a, b; \mathcal{Y})$ and $\mathcal{G}' = \mathcal{G}_E(a, b; \mathcal{Y} \cup \{z\})$ are intervals. There are two main cases: (a) $d(z, C) > E$ for all clusters C, or (b) otherwise.

We will deal with these cases separately, though the arguments are similar. Moreover, for each case, we will want to deal both with the original decompositions of the stable tree $T = T_e \cup T_c$, as well as its (r_1, r_2)-thickening $T = \mathbb{T}_e \cup \mathbb{T}_c$. The compatible decompositions for the latter will be refined for the decomposition for the former, with possibly a slight increase in the error depending on r_1, r_2.

Case (a), z forms a split cluster: In this case, z forms its own cluster and, in particular, does not affect membership of any other cluster. Moreover, it cannot be an endpoint cluster of $\mathcal{G}'$, as those necessarily contain a and b. Hence there exist two clusters C_1, C_2 which are adjacent in $\mathcal{G}$ but are separated by z in $\mathcal{G}'$, say with C_1 on the side of a and C_2 on the side of b. In this sense, we think of z as splitting the clusters C_1, C_2.

Lemma 8.3 implies that z separates two clusters C, C' for $(a, b; \mathcal{Y})$ if and only if $s(z)$ separates $s(C), s(C')$ along $\lambda(a, b)$. It follows then that the subgraph of $\mathcal{G}$ connecting the cluster containing a to C_1 is identical to the subgraph of $\mathcal{G}_E(a, b; \mathcal{Y} \cup \{z\})$, and similarly for the subgraphs connecting C_2 to the cluster containing b. This says that for any pair C, C' on the a-side of z in $\mathcal{G}'$, $\lambda_0(C, C')$ is an edge component of both T and T', and similarly for such pairs on the b-side of z. Thus we can include each such component of $E \subset T_e$ and $E \subset T'_e$ in the respective stable decomposition $T_s \subset T_e$ and $T'_s \subset T'_e$, where the identity map is the map described in parts (3c) and

(3d) of Definition 8.12. By construction, each of these pieces has positive integer length, thus verifying property (3a).

To construct the rest of the stable decompositions, namely stable pieces of type (3e) in Definition 8.12, we need to compare the component $\lambda_0(C_1, C_2) \subset T_e$ to the two components $\lambda_0(C_1, z) \cup \lambda_0(z, C_2) \subset T'_e$. First, Lemma 8.8 says that the endpoints of $\lambda(C_1, C_2)$ and $\lambda(C_1, z)$ on C_1 are 14ϵ-close, and hence the endpoints of $\lambda_0(C_1, C_2)$ and $\lambda_0(C_1, z)$ are within $14\epsilon + 2$, and similarly for C_2. It follows from δ-hyperbolicity of $\mathcal{Z}$ that $\lambda_0(C_1, C_2)$ and $\lambda_0(C_1, z) \cup \lambda_0(z, C_2)$ are $O(\epsilon)$-Hausdorff close. Now, unless C_1, C_2 are $O(\epsilon)$-close in $\mathcal{Z}$—in which case we can declare all three of these geodesics to be unstable—then the shadow $s_T(z)$ of z on T will be properly contained in $\lambda_0(C_1, C_2)$ and have $O(\epsilon)$-bounded diameter. Hence we can decompose these geodesics into $O(\epsilon)$-close stable pieces and unstable pieces labeled by z, as described in property (3f), while satisfying the length requirement of (3a) and proximity requirements of (3d) and (3e); see Figure 7 for a heuristic picture. For the bijection $\alpha : \pi_0(T_s) \to \pi_0(T'_s)$, since T_s, T'_s have the same number of components, we can take let α denote the natural order-preserving bijection, i.e. the one obtained by moving along the geodesics $[a, b]_T$ and $[a, b]_{T'}$. By construction, this bijection identifies the components we want so that the related conditions (3d) and (3e) are satisfied.

Finally, we need to confirm the properties in (3g) for the order-preserving bijection $\beta : \pi_0(T - T_s) \to \pi_0(T' - T'_s)$ induced by α. Since β is order-preserving (with the order being "from a to b", as in Definition 8.10), (3(g)i), i.e. that the components of $T - T_s$ and $T' - T'_s$ associated to a and b are identified, is automatically satisfied. For (3(g)ii), note that $\mathcal{Y} \cap (\mathcal{Y} \cup \{z\}) = \mathcal{Y}$, and, by construction again, if $y \in \mathcal{Y}$, then β identifies the component D_y of $T - T_s$ corresponding to y with D'_y, the component of $T' - T'_s$ corresponding to y. In this case, this is because all such components are exactly the same. This proves part (1) of the statement.

For part (2) of the statement, we want to consider how the (r_1, r_2)-thickenings of T, T' along T_c, T'_c respectively affect the decompositions $T_s \subset T_e$ and $T'_s \subset T'_e$. First, observe that $\mathbb{T}_e \subset T_e$ and $\mathbb{T}'_e \subset T'_e$ by definition. In particular, consider a component $D \subset T_e$, where $D = \lambda_0(C, C')$ for adjacent clusters C, C'. Then the (r_1, r_2)-thickening of D either truncates D to $D_{thick} = D - (\mu_{r_1}(C) \cup \mu_{r_1}(C'))$, or possibly removes D entirely from $\mathbb{T}_e$ if $d_T(\mu(C), \mu(C')) < r_2 - 2r_1$. A similar statement holds for components of T'_e.

By the above, most components of T_s and T'_s are identical components of T_e, T'_e, and hence their thickenings are exactly the same by the above. For the remaining components, namely $\lambda_0(C_1, z), \lambda_0(z, C_2)$ of T'_e and $\lambda_0(C_1, C_2)$ of T_e, each is either truncated or removed as above. The remaining two components of the stable decomposition T_s correspond to $\lambda_0(C_1, z)$ and $\lambda_0(z, C_2)$, and so one can construct the corresponding components of $\mathbb{T}_s$ by applying the same procedure to the (r_1, r_2)-thickening of $\lambda_0(C_1, z)$ and $\lambda_0(z, C_2)$. Note that this can inflate the proximity constants, but only in terms of r_1, r_2. The rest of the proof goes through unchanged; see Figure 7.

Case (b), z affects other clusters: Suppose now instead that there exists some cluster C for $(a, b; \mathcal{Y})$ so that $d(z, C) < E$; we say C is an *affected cluster*.

Claim B.2. If $3E - 56\epsilon > 2E$, then there are at most three affected clusters, and they are adjacent in $\mathcal{G}$.

Proof. Suppose that $C_1, \dots, C_4$ are four clusters appearing in that order along $\lambda(a,b)$. By item (1) of Lemma 8.8, since $d(C_i, C_{i+1}) \geqslant E$, we have $d(s(C_i), s(C_{i+1})) > E - 14\epsilon$, for $i = 1, 2, 3$.

Since the shadows are disjoint segments of the geodesic $\lambda(a,b)$, we have that

$$d(s(C_1), s(C_4)) \geqslant d(s(C_1), s(C_2)) + d(s(C_2), s(C_3)) + d(s(C_3), s(C_4)) \geqslant 3E - 42\epsilon.$$

Hence if $3E - 56\epsilon \geqslant 2E$, then item (1) of Lemma 8.8 says that

$$d(C_1, C_4) \geqslant d(s(C_1), s(C_4)) - 14\epsilon \geqslant 3E - 56\epsilon > 2E.$$

This is impossible if z is E-close to both C_1, C_4. Hence there are at most three affected clusters and they are adjacent vertices in $\mathcal{G}$. □

Let C' denote the cluster formed by z and the affected clusters from $(a, b; \mathcal{Y})$. Note that the membership of all other $(a, b; \mathcal{Y})$-clusters are unchanged by the addition of z. Let C_1 and C_2 be the clusters adjacent to the affected clusters in $\mathcal{G}$ on the side of a and b, respectively. As in the split cluster case above, Lemma 8.3 implies that C_1 separates all clusters on the a-side of $\mathcal{G}$ from C', and similarly for C_2 and clusters on the b-side. Hence the subgraph of $\mathcal{G}$ connecting the cluster containing a to C_1 is the subgraph of $\mathcal{G}'$ connecting the (same) cluster containing a to C_1, and similarly for the b-side of $\mathcal{G}$ and $\mathcal{G}'$. As such, all of the corresponding components of T_e and T'_e are exactly the same, and have positive integer length by construction, verifying properties (3a), (3c), and (3d) of Definition 8.12 for these components.

The forests T_e and T'_e have some remaining components which we need to decompose. In T_e, there are the at-most 2 components of T_e which connect the affected clusters. By assumption, these segments are at most $O(\epsilon)$ long, so we can declare them to be unstable components.

We now explain how to handle the two remaining components. Let A be the affected cluster adjacent to C_1 in $\mathcal{G}$. It is enough to show that the geodesic segments $\lambda(C_1, A)$ and $\lambda(C_1, C')$ are $O(\epsilon)$-Hausdorff close, since then the lengths of $\lambda_0(C_1, A)$ and $\lambda_0(C_1, C')$ are the same up to error $O(\epsilon)$, and we can declare a small end segment of each unstable, so that their (stable) complements have the same (positive integer) length and are $O(\epsilon)$-close. But this endpoint proximity statement is again a consequence of Lemma 8.8, since the endpoints of $\lambda(C_1, A)$ and $\lambda(C_1, C')$ are within 7ϵ of the endpoint of $s(C_1)$ on $\lambda(a,b)$, while $s(C')$ contains $s(A)$ and its endpoint is at most $(E + \epsilon)$-away. Finally, the bijection $\beta : \pi_0(T - T_s) \to \pi_0(T' - T'_s)$ required for (3g) is defined similar to the split cluster case (a), with the proof of condition (3(g)i) being basically the same. For condition (3(g)ii), observe again the $\mathcal{Y} \cap (\mathcal{Y} \cup \{z\}) = \mathcal{Y}$. Assuming that $y \in \mathcal{Y}$ is contained in an affected cluster in the $(\mathcal{Y}; \{a, b\})$ setup—since otherwise, it is contained in an identical cluster in the $(\mathcal{Y} \cup \{z\}; \{a, b\})$ setup—then it corresponds to the single component which gets identified with the component of $T' - T'_s$ corresponding to the single cluster containing both y and z. This completes the proof of (1) in the affected cluster case. See Figure 8 for a heuristic picture.

The argument for (2) in the affected cluster case is essentially identical to the split cluster version: All but a few components of $T_s \subset T_e$ and $T'_s \subset T'_e$ are identical and thus have identical (r_1, r_2)-thickenings. And the remaining components of T_s, T'_s are constructed by pairing the corresponding components of T_e, T'_e adjacent to or connecting the affected clusters, and then either declaring them unstable or

modifying them pairwise to achieve the length bound. As above, this can be accomplished post-thickening; see Figure 8. This completes the proof of the theorem. □

B.2. **Iterating stable decompositions.** Next we move to proving our iteration statement (Proposition B.6). The main work is the base case of combining two pairs of stable decompositions across a common setup.

For this, suppose we have three setups $(a_i, b_i; \mathcal{Y}_i)$ for $i = 1, 2, 3$, satisfying that $\mathcal{Y}_i \subset \mathcal{N}_{\epsilon/2}(\lambda(a_j, b_k))$ for all i, j, k, and that $\mathcal{Y}_0 \subset \mathcal{Y}_i$ for each i. Suppose that the stable intervals associated to $(a_1, b_1; \mathcal{Y}_1)$ and $(a_2, b_2; \mathcal{Y}_2)$ admits $\mathcal{Y}_0$-stable L-compatible decompositions, and the same for $(a_2, b_2; \mathcal{Y}_2)$ and $(a_3, b_3; \mathcal{Y}_3)$. Our goal is to induce new stable decompositions on T_1, T_2 and T_2, T_3 which encode the stable decompositions from T_2, T_3 and T_1, T_2, respectively. The following describes the properties we want for these induced decompositions:

Definition B.3 (Refined stable decompositions)**.** Let $(a, b; \mathcal{Y})$ and $(a', b'; \mathcal{Y}')$ denote two admissible ϵ-setups and let $\mathcal{Y}_0 = \mathcal{Y} \cap \mathcal{Y}'$. Suppose we have two compatible $\mathcal{Y}_0$-stable decompositions $T_s \subset T$ and $T'_s \subset T'$. We say another pair of stable decompositions $T_{s,0} \subset T$ and $T'_{s,0} \subset T'$ *refine* $T_s \subset T$ and $T'_s \subset T'$ if the following hold:

(1) We have $T_{s,0} \subset T_s$ and $T'_{s,0} \subset T'_s$, and
(2) Let $\beta : \pi_0(T - T_s) \to \pi_0(T' - T'_s)$ and $\beta_0 : \pi_0(T - T_{s,0}) \to \pi_0(T' - T'_{s,0})$ denote the associated order-preserving bijections. Suppose $C_1 \in \pi_0(T - T_s)$ and let $C_{1,0} \in \pi_0(T - T_{s,0})$ denote the component containing C_1 (which exists because $T_{s,0} \subset T_s$). Then $\beta(C_1) \subset \beta_0(C_{1,0})$.

Remark B.4. Condition (2) in particular is used to verify the properties in (3g) of Definition 8.12 during our iteration procedure. For item (3(g)i), since β identifies the component of $T - T_s$ containing the cluster component associated to a to the component of $T' - T'_s$ containing the cluster component associated to a', the refinement condition (2) forces the same to hold for β_0. A similar observation works to verify item (3(g)ii).

Following our notation from before the above definition, let $\phi_i : T_i \to \mathcal{Z}$ denote the maps into $\mathcal{Z}$. Moreover, assume that we have

- a $\mathcal{Y}_0$-stable decomposition $T^1_{s,2} \subset T_{e,2}$ which is L-compatible with $T^2_{s,1} \subset T_{e,1}$ for the setups $(a_1, b_1; \mathcal{Y}_1)$ and $(a_2, b_2; \mathcal{Y}_2)$, and
- a $\mathcal{Y}_0$-stable decomposition $T^3_{s,2} \subset T_{e,2}$ which is L-compatible with $T^2_{s,3} \subset T_{e,3}$ for the setups $(a_2, b_2; \mathcal{Y}_2)$ and $(a_3, b_3; \mathcal{Y}_3)$.

Proposition B.5. *There exists $\mathcal{Y}_0$-stable $4L^2$-compatible decompositions $T^{2,3}_{s,1} \subset T_{e,1}$ and $T^{2,1}_{s,3} \subset T_{e,3}$ which refine $T^2_{s,1}$ and $T^2_{s,3}$.*

Proof. Note that the stable decompositions $T^1_{s,2}$ and $T^3_{s,2}$ both live on T_2. Set $T^{1,3}_{s,2} = T^1_{s,2} \cap T^3_{s,2}$. Observe that $T_{e,2} - T^{1,3}_{s,2} = (T_{e,2} - T^1_{s,2}) \cup (T_{e,2} - T^3_{s,2})$ has at most $2L$-many components each of diameter at most $4L^2$; note that the larger diameter bound accounts for unstable components to combine, but there are at most $2L$-many of them. Moreover, by item (3) of Definition 8.12, the endpoints of each component of $T^1_{s,2}$ and $T^3_{s,2}$ are at vertices of the components of $T_{e,2}$ which contain them, and hence this property holds for each component of $T^{1,3}_{s,2}$

We first produce refined decompositions of $T^{2,3}_{s,1} \subset T^2_{s,1}$ and $T^{1,3}_{s,2} \subset T^1_{s,2}$, and the same argument produces the refined decompositions $T^{1,3}_{s,2} \subset T^3_{s,2}$ and $T^{2,1}_{s,3} \subset T^2_{s,3}$. We will then see the properties of Definition B.3 allow us to combine these, using $T^{1,3}_{s,2}$ as a bridge to connect them.

We can induce a decomposition $T^{2,3}_{s,1}$ on T_1 by taking the collection of all $i^{-1}_{E,E'}(E' \cap T^{1,3}_{s,2})$ over all stable pairs (E, E') of T_1, T_2 as identified by the bijection $\alpha_{1,2} : \pi_0(T_{s,1}) \to \pi_0(T_{s,2})$. Observe that any component $V \subset T^{2,3}_{s,1}$ is contained in some component $V \subset E \subset T_{s,1} \subset T_{e,1}$, and the map $i_{E,E'}|_V : V \to i_{E,E'}(V)$ is an isometry, where $E' = \alpha_{1,2}(E)$, and vice versa. Thus the components of $T^{2,3}_{s,1}$ are in bijective correspondence with the components of $T^{1,3}_{s,2}$, with identified components being isometric. Moreover, the natural order-preserving bijection $\alpha^3_{1,2} : \pi_0(T^{2,3}_{s,1}) \to \pi_0(T^{1,3}_{s,2})$ (as in Definition 8.10) coincides with the bijective correspondence induced from $\alpha_{1,2}$ and the various order-preserving isometries $i_{E,E'}$. We also note that the various restrictions $i_{E,E'}|V$ are order-preserving since $i_{E,E'}$ is. Next observe that the images $\phi_1(V)$ and $\phi_2(i_{E,E'}(V))$ are either exactly the same or are L-Hausdorff close in $\mathcal{Z}$ as items (3d) and (3e) of Definition 8.12. Thus the decompositions $T^{2,3}_{s,1} \subset T_1$ and $T^{1,3}_{s,2} \subset T_2$ satisfy all the properties of Definition 8.12 except possibly (3g). We also want to show that they give a refinement as in Definition B.3.

To prove (3g), it will suffice to show the refinement property (Definition B.3), as explained in Remark B.4. Property (1), namely that $T^{2,3}_{s,1} \subset T^2_{s,1}$ and $T^{1,3}_{s,2} \subset T^1_{s,2}$, of Definition B.3 holds by construction. For property (2), we argument one component at a time. In particular, let $\beta^3_{1,2} : \pi_0(T_1 - T^{2,3}_{s,1}) \to \pi_0(T_2 - T^{1,3}_{s,2})$ be the natural order-preserving bijections (which exists via Remark 8.11 because $\alpha^3_{1,2}$ does).

First, since the (unstable) components of $T_1 - T^2_{s,1}$ only get expanded when passing to $T_1 - T^{2,3}_{s,1}$, it follows that the first component $C_0 \subset T_1 - T^2_{s,1}$ one encounters when moving from a_1 to b_1 along the geodesic $[a_1, b_1]_{T_1}$ is contained in the first such component of $C_{0,3} \subset T_1 - T^{2,3}_{s,1}$, and similarly for the setup on T_2. Hence these expanded clusters are identified by $\beta^3_{1,2}$, which we note also confirms item (3(g)i) of Definition 8.12.

We now proceed inductively along T_1 and T_2. Let C_1, C_1' denote the second pairs of components of $T_1 - T^2_{s,1}$ and $T_2 - T^1_{s,2}$ encountered after C_0, C_0'. Let E, E' denote the respective components of $T^2_{s,1}$ and $T^1_{s,2}$ connecting C_0, C_0' to C_1, C_1'. The components of $(T_1 - T^{2,3}_{s,1}) \cap E$ and $(T_2 - T^{1,3}_{s,2}) \cap E'$ are identified by the order-preserving isometry $i_{E,E'}$ in an order-preserving way. In particular, the two sets of components are in bijective correspondence. As a consequence, if $C_1^3 \subset T_1 - T^{2,3}_{s,1}$ is the component containing C_1, then $C_1' \subset \beta^3_{1,2}(C_1^3)$, as required. One can repeat this procedure on subsequent components of $T_1 - T^2_{s,1}$ and their paired components of $T_2 - T^1_{s,2}$ to see that $T^{2,3}_{s,1}$ and $T^{1,3}_{s,2}$ refine $T^2_{s,1}$ and $T^1_{s,2}$, respectively.

Combining the decompositions: We now want to see that $T^{2,3}_{s,1}$ and $T^{2,1}_{s,3}$ are $\mathcal{Y}_0$-stable $4L^2$-compatible decompositions for $(a_1, b_1; \mathcal{Y}_1)$ and $(a_3, b_3; \mathcal{Y}_3)$. For this, define $\alpha_{1,3} : \pi_0(T^{2,3}_{s,1}) \to \pi_0(T^{2,1}_{s,3})$ to be $\alpha_{1,3} = \alpha^1_{2,3} \circ \alpha^3_{1,2}$, and similarly set $\beta_{1,3} = \beta^1_{2,3} \circ \beta^3_{1,2}$. These compositions are order-preserving bijections because their constituent maps are, and moreover we can similarly induce order-preserving isometries on the components of $T^{2,3}_{s,1}$ and $T^{2,1}_{s,3}$ identified by $\alpha_{1,3}$, which satisfy the

proximity conditions of (3d) and (3e) of Definition 8.12 with a constant of $2L$. In particular, it remains to show the properties in (3g).

For this, we use the refinement property (2) of Definition B.3. This is basically automatic. For (1), the refinement property implies that $\beta^3_{1,2}$ identifies the component of $T_1 - T^{2,3}_{s,1}$ associated to a_1 to the component of $T_2 - T^{1,3}_{s,2}$ associated to a_2, and similarly for $\beta^1_{2,3}$. Hence $\beta_{1,3}$ identifies the components associated to a_1, a_3 and b_1, b_3. Similarly, if $y \in \mathcal{Y}_0 \subset \mathcal{Y}_1 \cap \mathcal{Y}_2 \cap \mathcal{Y}_3$, then the refinement property implies that $\beta^3_{1,2}$ identifies the component of $T_1 - T^{2,3}_{s,1}$ associated to y to the component of $T_2 - T^{1,3}_{s,2}$ associated to y, and similarly for $\beta^1_{2,3}$. Hence $\beta_{1,3}$ associates components associated to y in both of $T_1 - T^{2,3}_{s,1}$ and $T_3 - T^{2,1}_{s,3}$. This completes the proof of the proposition.

□

The following proposition is now a simple iterative application of Proposition B.5:

Proposition B.6. *For every $n \geqslant 2$, there exists $L_n = L_n(L, n) > 0$ so that the following holds. Suppose $(a_1, b_1; \mathcal{Y}_1), \dots, (a_n, b_n; \mathcal{Y}_n)$ is a sequence of pairwise admissible ϵ-setups and suppose that $\mathcal{Y}_0 \subset \bigcap_i \mathcal{Y}_i$. If each pair of consecutive setups admits $\mathcal{Y}_0$-stable L-compatible stable decompositions, then $(a_1, b_1; \mathcal{Y}_1)$ and $(a_n, b_n; \mathcal{Y}_n)$ admit $\mathcal{Y}_0$-stable L_n-compatible stable decompositions.*

B.2.1. *Completing the proof.* To prove Theorem 8.14, we have two remaining tasks:

(1) We need to explain how to build a stable decomposition when perturbing an endpoint but keeping the rest of the setup the same, and
(2) Show how to stick this into an iterative setup.

Perturbing an endpoint: This is an iterated application of Proposition B.1.

Consider the following restricted setup, which we highlight to focus on what happens when we perturb an endpoint. Let $a, b, a' \in \mathcal{Z}$ and $\mathcal{Y} \subset \mathcal{Z}$ be finite with $\mathcal{Y} \subset \mathcal{N}_{\epsilon/2}(\lambda(a, b)) \cap \mathcal{N}_{\epsilon/2}(a', b)$. Consider the following collection of setups, each of which differs by adding or deleting one cluster point, or by switching the role that a, a' or b, b' play between cluster or endpoint:

(1) $(a, b; \mathcal{Y})$,
(2) $(a, b; \mathcal{Y} \cup \{a\})$,
(3) $(a, b; \mathcal{Y} \cup \{a, a'\})$,
(4) $(a', b; \mathcal{Y} \cup \{a, a'\})$,
(5) $(a', b; \mathcal{Y} \cup \{a'\})$,
(6) $(a', b; \mathcal{Y})$.

First observe that $\mathcal{Y}$ is contained in each of the above sets of cluster points, so we want to see how to build $\mathcal{Y}$-stable decompositions for each, which implies the existence of $\mathcal{Y}_0$-stable decompositions for any $\mathcal{Y}_0 \subset \mathcal{Y}$. Next, observe that in the construction of the stable interval $T(a, b; \mathcal{Y})$, the roles of points of the endpoints a, b and the cluster points $\mathcal{Y}$ are equivalent in cluster formation process. The only special role that endpoints play in the Definition 8.6 of $T(a, b; \mathcal{Y})$ is in the segments defined *within* clusters, namely the $\mu(C)$ components. In particular, the edge components (i.e., the components of T_e) treat endpoints and cluster points equivalently. As the desired stable decompositions from Theorem 8.14 are subsets of the edge components, we are free to treat the points a, a', b, b' as cluster points in applying Proposition B.1 to obtain stable decompositions for each of the steps (1) → (2) → (3) and (4) → (5) → (6). The only exception is the step (3) → (4), where we are switching the roles of a and a'. Here Proposition B.1 does not directly apply, but, by the above discussion, these two setups have the exact

same clusters and thus the exact same edge components. Hence we get the desired bijections (with their proximity constraints) by identifying the edge components as stable components—which, again, are all exactly the same—and taking the bijection between the complementary components to be the obvious map.

In other words, provided an explanation for how to iteratively combine a sequence of stable decompositions, we can apply this process to move from setup (1) above to setup (6).

Putting it all together: In the more general setting of the theorem, the above sequence only appears in the middle of a longer sequence we need to consider. In particular, let $\mathcal{Y}' - \mathcal{Y} = \{x_1, \dots, x_n\}$ and $\mathcal{Y} - \mathcal{Y}' = \{z_1, \dots, z_m\}$, where $n, m < N$ by assumption. Set $\mathcal{Y}_i = \mathcal{Y} \cup \{x_1, \dots, x_i\}$ and $\mathcal{Y}'_i = \mathcal{Y}' \cup \{z_1, \dots, z_i\}$, and let $\mathcal{Y}_0 = \mathcal{Y} \cap \mathcal{Y}'$ and $\mathcal{Y}_{all} = \mathcal{Y} \cup \mathcal{Y}'$. Then, by the assumptions in our setup, we have two sequences of pairwise admissible ϵ-setups

$$(a, b; \mathcal{Y}) \to (a, b; \mathcal{Y}_1) \to \cdots \to (a, b; \mathcal{Y}_{all}) \to (a, b; \mathcal{Y}_{all} \cup \{a'\}) \to (a, b; \mathcal{Y}_{all} \cup \{a', b'\})$$

and

$$(a', b'; \mathcal{Y}') \to \cdots \to (a', b'; \mathcal{Y}_n) \to (a', b'; \mathcal{Y}_{all}) \to (a', b'; \mathcal{Y}_{all} \cup \{a\}) \to (a', b'; \mathcal{Y}_{all} \cup \{a, b\}),$$

which we note are connected in the middle by the above discussion, where we need to switch the roles of both a, a' and b, b'. In particular, the stable trees for the setups $(a, b; \mathcal{Y}_{all} \cup \{a', b'\})$ and $(a', b'; \mathcal{Y}_{all} \cup \{a, b\})$ have identical edge components.

By Proposition B.6, we get $\mathcal{Y}_0$-stable L_{n+2}-compatible stable decompositions for $(a, b; \mathcal{Y})$ and $(a, b; \mathcal{Y}_{all} \cup \{a', b'\})$, and $\mathcal{Y}_0$-stable L_{m+2}-compatible stable decompositions for $(a', b'; \mathcal{Y})$ and $(a', b'; \mathcal{Y}_{all} \cup \{a, b\})$. Since $(a, b; \mathcal{Y}_{all} \cup \{a', b'\})$ and $(a', b'; \mathcal{Y}_{all} \cup \{a, b\})$ determine identical edge components, we can use an identical argument to the proof of Proposition B.5 to induce refined $\mathcal{Y}_0$-stable $(L_n \cdot L_m)$-compatible stable decompositions for $(a, b; \mathcal{Y})$ and $(a', b'; \mathcal{Y})$. We leave the details to the reader. This completes the proof. □

References

[AB95] Juan M. Alonso and Martin R. Bridson. Semihyperbolic groups. *Proc. London Math. Soc. (3)*, 70(1):56–114, 1995. (Cited pp. 2, 3, and 9).

[ABD17] Carolyn Abbott, Jason Behrstock, and Matthew Gentry Durham. Largest acylindrical actions and stability in hierarchically hyperbolic groups. *arXiv preprint arXiv:1705.06219*, 2017. (Cited p. 1).

[Aou13] Tarik Aougab. Uniform hyperbolicity of the graphs of curves. *Geometry & Topology*, 17(5):2855–2875, 2013. (Cited p. 12).

[BBF15] Mladen Bestvina, Ken Bromberg, and Koji Fujiwara. Constructing group actions on quasi-trees and applications to mapping class groups. *Publ. Math. Inst. Hautes Études Sci.*, 122:1–64, 2015. (Cited pp. 40 and 41).

[BBFS20] Mladen Bestvina, Ken Bromberg, Koji Fujiwara, and Alessandro Sisto. Acylindrical actions on projection complexes. *Enseign. Math.*, 65(1-2):1–32, 2020. (Cited pp. 40 and 41).

[Beh06] Jason A. Behrstock. Asymptotic geometry of the mapping class group and Teichmüller space. *Geom. Topol.*, 10:1523–1578, 2006. (Cited p. 13).

[Ber21] Daniel James Solomon Berlyne. *Hierarchical hyperbolicity of graph products and graph braid groups*. PhD thesis, City University of New York, 2021. (Cited p. 12).

[Bes96] Mladen Bestvina. Local homology properties of boundaries of groups. *Michigan Mathematical Journal*, 43(1):123–139, 1996. (Cited p. 2).

[BGGH25] Eliot Bongiovanni, Pritam Ghosh, Funda Gültepe, and Mark Hagen. Characterizing hierarchically hyperbolic free by cyclic groups. *arXiv preprint arXiv:2508.15738*, 2025. (Cited p. 12).

[BHMS24a] Jason Behrstock, Mark Hagen, Alexandre Martin, and Alessandro Sisto. A combinatorial take on hierarchical hyperbolicity and applications to quotients of mapping class groups. *Journal of Topology*, 17(3):e12351, 2024. (Cited p. 12).

[BHMS24b] Jason Behrstock, Mark Hagen, Alexandre Martin, and Alessandro Sisto. A combinatorial take on hierarchical hyperbolicity and applications to quotients of mapping class groups. *J. Topol.*, 17(3):Paper No. e12351, 94, 2024. (Cited p. 16).

[BHS17a] Jason Behrstock, Mark F. Hagen, and Alessandro Sisto. Asymptotic dimension and small-cancellation for hierarchically hyperbolic spaces and groups. *Proc. Lond. Math. Soc. (3)*, 114(5):890–926, 2017. (Cited p. 12).

[BHS17b] Jason Behrstock, Mark F. Hagen, and Alessandro Sisto. Hierarchically hyperbolic spaces, I: Curve complexes for cubical groups. *Geom. Topol.*, 21(3):1731–1804, 2017. (Cited pp. 1, 6, 11, and 12).

[BHS19] Jason Behrstock, Mark Hagen, and Alessandro Sisto. Hierarchically hyperbolic spaces II: Combination theorems and the distance formula. *Pacific Journal of Mathematics*, 299(2):257–338, 2019. (Cited pp. 3, 11, 12, 13, 14, 15, 32, and 41).

[BHS21] Jason Behrstock, Mark F Hagen, and Alessandro Sisto. Quasiflats in hierarchically hyperbolic spaces. *Duke Mathematical Journal*, 1(1):1–88, 2021. (Cited pp. 1, 2, 3, 6, 11, 16, and 38).

[BKMM12] J. Behrstock, B. Kleiner, Y. Minsky, and L. Mosher. Geometry and rigidity of mapping class groups. *Geometry and Topology*, 16:781–888, 2012. (Cited pp. 13 and 41).

[Bon24] Eliot Bongiovanni. Extensions of finitely generated veech groups. *arXiv preprint arXiv:2406.11090*, 2024. (Cited p. 12).

[Bow19] Brian H. Bowditch. Convex hulls in coarse median spaces. *preprint*, 2019. (Cited pp. 1, 12, 14, and 16).

[Bow22] Brian H Bowditch. Median algebras. *preprint*, page 209, 2022. (Cited pp. 35 and 36).

[BR22] Daniel Berlyne and Jacob Russell. Hierarchical hyperbolicity of graph products. *Groups Geom. Dyn.*, 16(2):523–580, 2022. (Cited p. 12).

[Bro02] Jeffrey F. Brock. Pants decompositions and the Weil-Petersson metric. In *Complex manifolds and hyperbolic geometry (Guanajuato, 2001)*, volume 311 of *Contemp. Math.*, pages 27–40. Amer. Math. Soc., Providence, RI, 2002. (Cited pp. 1, 11, 12, and 14).

[Che00] Victor Chepoi. Graphs of some CAT(0) complexes. *Adv. in Appl. Math.*, 24(2):125–179, 2000. (Cited pp. 35 and 36).

[Che22] Marissa Chesser. Stable subgroups of the genus 2 handlebody group. *Algebraic & Geometric Topology*, 22(2):919–971, 2022. (Cited p. 12).

[CLM12] Matt T. Clay, Christopher J. Leininger, and Johanna Mangahas. The geometry of right-angled Artin subgroups of mapping class groups. *Groups Geom. Dyn.*, 6(2):249–278, 2012. (Cited p. 41).

[CN05] Indira Chatterji and Graham Niblo. From wall spaces to CAT(0) cube complexes. *Internat. J. Algebra Comput.*, 15(5-6):875–885, 2005. (Cited p. 35).

[CR24] Aaron Calderon and Jacob Russell. Hierarchical hyperbolicity of admissible curve graphs and the boundary of marked strata. *arXiv preprint arXiv:2409.06798*, 2024. (Cited p. 12).

[CRHK21] Montserrat Casals-Ruiz, Mark Hagen, and Ilya Kazachkov. Real cubings and asymptotic cones of hierarchically hyperbolic groups. *preprint*, 2021. (Cited pp. 3 and 15).

[CS11] Pierre-Emmanuel Caprace and Michah Sageev. Rank rigidity for cat (0) cube complexes. *Geometric and functional analysis*, 21(4):851–891, 2011. (Cited p. 6).

[DDLS23] Spencer Dowdall, Matthew G Durham, Christopher J Leininger, and Alessandro Sisto. Extensions of veech groups i: A hyperbolic action. *Journal of Topology*, 16(2):757–805, 2023. (Cited p. 12).

[DDLS24] Spencer Dowdall, Matthew G. Durham, Christopher J. Leininger, and Alessandro Sisto. Extensions of Veech groups II: Hierarchical hyperbolicity and quasi-isometric rigidity. *Comment. Math. Helv.*, 99(1):149–228, 2024. (Cited p. 12).

[DMS20] Matthew G Durham, Yair N Minsky, and Alessandro Sisto. Stable cubulations, bicombings, and barycenters. *arXiv preprint arXiv:2009.13647*, 2020. (Cited pp. 1, 2, 3, 4, 6, 7, 9, 10, 11, 15, 40, 41, 42, 45, 47, 48, and 49).

[DMS25] Matthew Gentry Durham, Yair Minsky, and Alessandro Sisto. Asymptotically cat (0) metrics, z-structures, and the farrell-jones conjecture. *arXiv preprint arXiv:2504.17048*, 2025. (Cited pp. 2, 3, 4, 6, 47, and 49).

[Dra06] A. N. Dranishnikov. On Bestvina-Mess formula. In *Topological and asymptotic aspects of group theory*, volume 394 of *Contemp. Math.*, pages 77–85. Amer. Math. Soc., Providence, RI, 2006. (Cited p. 2).

[Dur16] Matthew Gentry Durham. The augmented marking complex of a surface. *J. Lond. Math. Soc. (2)*, 94(3):933–969, 2016. (Cited pp. 1, 3, 11, 12, and 14).

[Dur23] Matthew Gentry Durham. Cubulating infinity in hierarchically hyperbolic spaces. *arXiv preprint arXiv:2308.13689*, 2023. (Cited pp. 1, 2, 3, 4, 7, 11, 12, 14, 15, 16, 17, 20, 22, 26, 28, 29, 30, 31, 32, 33, 35, 36, 37, 38, 39, 58, 59, 60, and 61).

[DZ22] Matthew Gentry Durham and Abdul Zalloum. The geometry of genericity in mapping class groups and teichmüller spaces via cat (0) cube complexes. *To appear in* Transactions of the AMS*, arXiv:2207.06516*, 2022. (Cited p. 1).

[EMR17] Alex Eskin, Howard Masur, and Kasra Rafi. Large-scale rank of Teichmüller space. *Duke Math. J.*, 166(8):1517–1572, 2017. (Cited pp. 11 and 13).

[FJ93] F Thomas Farrell and Lowell E Jones. Isomorphism conjectures in algebraic K-theory. *Journal of the American Mathematical Society*, 6(2):249–297, 1993. (Cited p. 2).

[FM11] Benson Farb and Dan Margalit. *A primer on mapping class groups (pms-49)*, volume 41. Princeton university press, 2011. (Cited p. 3).

[Hag23] Mark Hagen. Non-colorable hierarchically hyperbolic groups. *Internat. J. Algebra Comput.*, 33(2):337–350, 2023. (Cited p. 12).

[Ham21] Ursula Hamenstädt. Geometry of the mapping class groups iia: Cat (0) cube complexes and boundaries. *Preprint available at www. math. uni-bonn. de/people/ursula/noncube. pdf*, page 34, 2021. (Cited p. 1).

[Ham25] Ursula Hamenstädt. A Z-structure for the mapping class group. arXiv:2505.18808, 2025. (Cited p. 2).

[HHP23] Thomas Haettel, Nima Hoda, and Harry Petyt. Coarse injectivity, hierarchical hyperbolicity and semihyperbolicity. *Geom. Topol.*, 27(4):1587–1633, 2023. (Cited pp. 1 and 3).

[HMS24] Mark Hagen, Alexandre Martin, and Alessandro Sisto. Extra-large type Artin groups are hierarchically hyperbolic. *Math. Ann.*, 388(1):867–938, 2024. (Cited p. 12).

[HP98] Frédéric Haglund and Frédéric Paulin. *Simplicité de groupes d'automorphismes d'espaces à courbure négative*. Université de Paris-Sud. Département de Mathématique, 1998. (Cited p. 4).

[HP23] Mark F Hagen and Harry Petyt. Projection complexes and quasimedian maps. *Algebraic & Geometric Topology*, 22(7):3277–3304, 2023. (Cited p. 40).

[HRSS22] Mark Hagen, Jacob Russell, Alessandro Sisto, and Davide Spriano. Equivariant hierarchically hyperbolic structures for 3-manifold groups via quasimorphisms. *arXiv preprint arXiv:2206.12244*, 2022. (Cited p. 12).

[HS20] Mark F. Hagen and Tim Susse. On hierarchical hyperbolicity of cubical groups. *Israel J. Math.*, 236(1):45–89, 2020. (Cited p. 12).

[HW14] G Christopher Hruska and Daniel T Wise. Finiteness properties of cubulated groups. *Compositio Mathematica*, 150(3):453–506, 2014. (Cited p. 4).

[Kar11] Aditi Kar. Asymptotically CAT(0) groups. *Publicacions Matemàtiques*, pages 67–91, 2011. (Cited p. 2).

[KK14] Sang-Hyun Kim and Thomas Koberda. The geometry of the curve graph of a right-angled artin group. *International Journal of Algebra and Computation*, 24(02):121–169, 2014. (Cited p. 14).

[Min10] Yair Minsky. The classification of kleinian surface groups, i: Models and bounds. *Annals of Mathematics*, pages 1–107, 2010. (Cited p. 16).

[MM99] Howard A Masur and Yair N Minsky. Geometry of the complex of curves I: Hyperbolicity. *Inventiones mathematicae*, 138(1):103–149, 1999. (Cited pp. 1, 3, 11, and 12).

[MM00a] H. A. Masur and Y. N. Minsky. Geometry of the complex of curves. II. Hierarchical structure. *Geom. Funct. Anal.*, 10(4):902–974, 2000. (Cited p. 15).

[MM00b] Howard A Masur and Yair N Minsky. Geometry of the complex of curves II: Hierarchical structure. *Geometric and Functional Analysis*, 10(4):902–974, 2000. (Cited pp. 1, 3, 11, 12, and 14).

[Nic13] Bogdan Nica. Proper isometric actions of hyperbolic groups on L^p-spaces. *Compos. Math.*, 149(5):773–792, 2013. (Cited p. 35).

[NR98] G. A. Niblo and L. D. Reeves. The geometry of cube complexes and the complexity of their fundamental groups. *Topology*, 37(3):621–633, 1998. (Cited pp. 3, 7, 8, and 12).

[Pet21] Harry Petyt. Mapping class groups are quasicubical. *To appear in the* American Journal of Mathematics*, arXiv:2112.10681*, 2021. (Cited pp. 1 and 40).

[PZS24] Harry Petyt, Abdul Zalloum, and Davide Spriano. Constructing metric spaces from systems of walls. *arXiv preprint arXiv:2404.12057*, 2024. (Cited p. 1).

[Raf07] Kasra Rafi. A combinatorial model for the Teichmüller metric. *Geom. Funct. Anal.*, 17(3):936–959, 2007. (Cited pp. 3, 11, 12, and 14).

[Raf14] Kasra Rafi. Hyperbolicity in Teichmüller space. *Geometry & Topology*, 18(5):3025–3053, 2014. (Cited pp. 1, 3, and 11).

[Rol16] Martin Roller. Poc sets, median algebras and group actions. *arXiv preprint arXiv:1607.07747*, 2016. (Cited pp. 4, 5, and 35).

[RST23] Jacob Russell, Davide Spriano, and Hung Cong Tran. Convexity in hierarchically hyperbolic spaces. *Algebraic & Geometric Topology*, 23(3):1167–1248, 2023. (Cited p. 15).

[Rus21] Jacob Russell. Extensions of multicurve stabilizers are hierarchically hyperbolic. *To appear in* Geometry & Topology*, arXiv:2107.14116*, 2021. (Cited p. 12).

[Sag95] Michah Sageev. Ends of group pairs and non-positively curved cube complexes. *Proc. London Math. Soc. (3)*, 71(3):585–617, 1995. (Cited pp. 4, 5, and 35).

[Sag14] Michah Sageev. CAT(0) cube complexes and groups. In *Geometric group theory*, volume 21 of *IAS/Park City Math. Ser.*, pages 7–54. Amer. Math. Soc., Providence, RI, 2014. (Cited pp. 3 and 4).

[Sch] Petra Schwer. CAT(0) cube complexes. *Lecture Notes in Mathematics*, 2324. (Cited p. 3).

[She25] Sam Shepherd. A cubulation with no factor system. *Algebraic & Geometric Topology*, 25(1):255–266, 2025. (Cited p. 12).

[Sis17] Alessandro Sisto. What is a hierarchically hyperbolic space? *arXiv preprint arXiv:1707.00053*, 2017. (Cited pp. 3 and 12).

[Zal23] Abdul Zalloum. Injectivity, cubical approximations and equivariant wall structures beyond cat (0) cube complexes. *arXiv preprint arXiv:2305.02951*, 2023. (Cited p. 3).

[Zal25] Abdul Zalloum. Effective flipping, skewering and rank rigidity for cubulated groups with factor systems. *Transactions of the American Mathematical Society*, 2025. (Cited p. 1).

Department of Mathematics and Statistics, CUNY Hunter College, New York, NY
Email address: matthew.durham@hunter.edu