HOW MANY DIRECTIONS DETERMINE A SHAPE AND OTHER SUFFICIENCY RESULTS FOR TWO TOPOLOGICAL TRANSFORMS

JUSTIN CURRY, SAYAN MUKHERJEE, AND KATHARINE TURNER

ABSTRACT. In this paper we consider two topological transforms that are popular in applied topology: the Persistent Homology Transform and the Euler Characteristic Transform. Both of these transforms are of interest for their mathematical properties as well as their applications to science and engineering, because they provide a way of summarizing shapes in a topological, yet quantitative, way. Both transforms take a shape, viewed as a tame subset M of $\mathbb{R}^d$, and associates to each direction $v \in S^{d-1}$ a shape summary obtained by scanning M in the direction v. These shape summaries are either persistence diagrams or piecewise constant integer-valued functions called Euler curves. By using an inversion theorem of Schapira, we show that both transforms are injective on the space of shapes, i.e. each shape has a unique transform. Moreover, we prove that these transforms determine continuous maps from the sphere to the space of persistence diagrams, equipped with any Wasserstein p-distance, or the space of Euler curves, equipped with certain L^p norms. By making use of a stratified space structure on the sphere, induced by hyperplane divisions, we prove additional uniqueness results in terms of distributions on the space of Euler curves. Finally, our main result proves that any shape in a certain uncountable space of PL embedded shapes with plausible geometric bounds can be uniquely determined using only finitely many directions.

1. INTRODUCTION

In this paper we consider two topological transforms that are of theoretical and practical interest: the Euler Characteristic Transform (ECT) and the Persistent Homology Transform (PHT). At a high level both of these transforms take a shape, viewed as a subset M of $\mathbb{R}^d$, and associates to each direction $v \in S^{d-1}$ a shape summary obtained by scanning M in the direction v. This process of scanning has a Morse-theoretic and persistent-topological flavor—we study the topology of the sublevel sets of the height function $h_v = \langle v, \cdot \rangle|_M$ as the height varies. These evolving sublevel sets are summarized using either the *Euler Curve*, which records the Euler characteristic of each sublevel set, or the *persistence diagram*, which pairs

critical values of h_v in a computational way. In short, the ECT and the PHT are transforms that associate to any sufficiently tame subset $M \subset \mathbb{R}^d$ a map from the sphere to the space of Euler curves and persistence diagrams, respectively.

Before introducing the mathematical properties of these transforms, as well as the results proved in this paper, we point out some of the applications of interest. In both data science and computational geometry, quantifying differences in shape is a difficult problem. Part of the problem is structural: most statistical analyses operate on scalar-valued quantities, and it is very hard to summarize shape with a single number. To address this, Euler curves and persistence diagrams are slightly more complicated summary objects endowed with metrics amenable to detecting qualitative and quantitative differences in shape. To wit, in [20] the PHT was introduced and applied to the study of heel bones from primates. By quantifying the shapes of these heel bones and clustering these using certain metrics on persistence diagrams, the authors of that paper were able to automatically differentiate species of primates using quantified differences in phylogenetic expression of heel bone shape. A variant on the Euler Characteristic Transform was also used in [8] to predict clinical outcomes of brain tumors based on their shape. Here the applicability of topological methods is particularly compelling: aggressive tumors tend to grow "roots" that invade nearby tissues and these protrusions are often well-detected by changes in the Euler curves.

Of course, in each of these applications, one may wonder: To what extent are these shape summaries lossy? Topological invariants such as Euler characteristic and homology are obviously not faithful; there are non-homeomorphic spaces that appear to be identical when using these particular lenses for viewing shape. The surprising fact developed in [18, 20] and carried further in this paper and independently in [12] is that by considering the Euler curve for every possible direction one can completely recover a shape. Said differently, the Euler Characteristic Transform is a statistic that is sufficient to summarize compact, definable subsets of $\mathbb{R}^d$, see Theorem 3.5. Moreover, since homology determines Euler characteristic via an alternating sum of Betti numbers, we obtain injectivity results for the PHT as well, see Theorem 4.16.

The choice of our class of shapes is an important variable throughout the paper. We use "constructible sets" at first, these are compact subsets of $\mathbb{R}^d$ that can be constructed with finitely many geometric and logical operations. The theory for carving out this collection of sets comes from logic and the study of o-minimal/definable structures, as they provide a way of banishing shapes that are considered "too wild." One key property of working in this setting is that every constructible set can be triangulated. This is the natural setting for Schapira's inversion result [18], which is the engine that drives many of the results in our paper, specifically Theorems 3.5 and 4.16, and independently Theorem 5 and Corollary 6 in [12].

An additional reason to consider o-minimal sets is that they are naturally stratified into manifold pieces. This is important because it implies that the constructions that underly the ECT and PHT are also stratified, and that both of these two transforms can be described in terms of constructible sheaves. The upshot of this observation, which is developed in Section 5, is that the ECT of a particular shape M should be determined, in some sense, by finitely many directions, certainly if M is known in advance. Indeed, any o-minimal set induces a stratification of the sphere of directions, and whenever we restrict to a particular stratum the variation

of the transform can be described by homeomorphisms of the real line. To make precise how to relate the transform for two different directions in a fixed stratum, we specialize to the case where $M = K$ is a piecewise linearly (PL) embedded simplicial complex. This allows us to provide an explicit formula for the ECT and construct explicit linear relations for both transforms whenever two directions lie in the same stratum. This is the content of Lemma 5.19 and Proposition 5.21.

By considering a "generic" class of PL embedded simplicial complexes, we leverage the explicit formula for the Euler Characteristic Transform to produce a new measure-theoretic perspective on shapes. In Theorem 6.7 we prove that by considering the pushforward of the Lebesgue measure on S^{d-1} into the space of Euler curves (or persistence diagrams) one can uniquely determine a generic embedded simplicial complex up to rotation and reflection, i.e. an element of $O(d)$. The importance of this result for applications is that one can faithfully compare two un-aligned shapes simply by studying their associated distribution of Euler curves or persistence diagrams.

Finally, we extend the themes of stratification theory and injectivity results to answer our titular question, which is perhaps more accurately worded as "How many directions are needed to infer a shape?" Here the idea is that there is a shape hidden from view, perhaps cloaked by our sphere of directions, and we would like to learn as much about the shape as possible. Our mode of interrogation is that we can specify a direction and an oracle will tell us the Euler curve of the shape when viewed in that direction. Of course, by our earlier injectivity result, Theorem 3.5, we know that if we query all possible directions, then we can uniquely determine any constructible shape. However, a natural question of theoretical and practical importance is whether finitely many queries suffice. The main result of our paper, Theorem 7.14, shows that if we impose some a priori assumptions on our (uncountable) set of possible shapes, then finitely many queries indeed do suffice. However, unlike the popular parlor game *Twenty Questions,* the number of questions/directions we might need to infer our hidden shape is only bounded by

$$\Delta(d, \delta, k_\delta) = \left((d-1)\, k_\delta \left(\frac{2\delta}{\sin(\delta)} \right)^{d-1} + 1 \right) \left(1 + \frac{3}{\delta} \right)^d + O\left(\frac{dk_\delta}{\delta^{d-1}} \right)^{2d}.$$

The parameters above reflect a priori geometric assumptions that we make about our hidden shape. Specifically, we assume our shape is well modeled as a geometric simplicial complex embedded in $\mathbb{R}^d$, with a lower bound δ on the "curvature" at every vertex, and a uniform upper bound k_δ on the number of critical values when viewed in any given δ-ball of directions. We call this class of shapes $\mathcal{K}(d, \delta, k_\delta)$. Moreover, the two terms in the above sum are best understood as a two step interrogation procedure. The first term counts the number of directions that we'd sample the ECT at for any shape. The second term in the above sum counts the maximum number of questions we might ask, adapted to the answers given in the first round of questions.

2. Background on Euler calculus

Often in geometry and topology we require that our sets and maps have certain tameness properties. These tameness properties are exhibited by the following categories: piecewise linear, semialgebraic, and subanalytic sets and mappings.

Logicians have generalized and abstracted these categories into the notion of an o-minimal structure [22], which we now define.

Definition 2.1. An *o-minimal structure* $\mathcal{O} = \{\mathcal{O}_d\}$ specifies for each $d \geq 0$, a collection of subsets $\mathcal{O}_d$ of $\mathbb{R}^d$ closed under intersection and complement. These collections are related to each other by the following rules:

(1) If $A \in \mathcal{O}_d$, then $A \times \mathbb{R}$ and $\mathbb{R} \times A$ are both in $\mathcal{O}_{d+1}$; and
(2) If $A \in \mathcal{O}_{d+1}$, then $\pi(A) \in \mathcal{O}_d$ where $\pi : \mathbb{R}^{d+1} \to \mathbb{R}^d$ is axis-aligned projection.

We further require that $\mathcal{O}$ be closed with respect to all the operations of $\mathbb{R}$ that make it an ordered field, i.e. comparison $(<)$, addition $(+)$ and multiplication $(\cdot)$, and that $\mathcal{O}_1$ contain no more and no less than all finite unions of points and open intervals in $\mathbb{R}$. Elements of $\mathcal{O}$ are called *tame* or *definable* sets. Our assumptions guarantee that every semialgebraic set is definable [22, Ch. 2, 2.11]. A *definable map* $f : \mathbb{R}^d \to \mathbb{R}^n$ is one whose graph is definable. We define a *constructible set* as a compact definable subset of $\mathbb{R}^d$ and denote the set of all constructible subsets of $\mathbb{R}^d$ by $\mathrm{CS}(\mathbb{R}^d)$.

Remark 2.2. Our assumptions include more than is strictly required by van den Dries's notion of an o-minimal structure, which is defined in [22, Ch. 1, 2.1], in order to contain the sets we are most interested in, i.e. semialgebraic sets and their generalizations. What we have specified in Definition 2.1 is more accurately called the o-minimal structure generated by the ordered field $(\mathbb{R}, <, 0, 1, -, +, \cdot)$ with parameters/constants in $\mathbb{R}$, cf. [22, Ch. 1, §5]. As van den Dries describes in [22, Ch. 1, 5.1] and proves in [22, Ch. 2, 2.11] our notion of an o-minimal structure contains the collection of semialgebraic sets.

The advantage of o-minimal structures is that one can consider expansions of the collection of semialgebraic sets to include a restricted class of analytic functions, such as the square root function, without sacrificing desired tameness properties. Further expansions include a notion of exponentiation and the Pfaffian, which by Wilkie's theorem provides a model-complete theory [28].

Tame/definable sets play the role of measurable sets for an integration theory based on Euler characteristic called *The Euler Calculus*, see [10] for an expository review. The guarantee that definable sets can be measured by Euler characteristic is by virtue of Theorem 2.3.

Theorem 2.3 (Triangulation Theorem [22])**.** *Any tame set admits a definable bijection with a subcollection of open simplices in the geometric realization of a finite Euclidean simplicial complex. Moreover, this bijection can be made to respect a partition of a tame set into tame subsets.*

In view of the Triangulation Theorem, one can define the Euler characteristic of a tame set in terms of an alternating count of the number of simplices used in a definable triangulation.

Definition 2.4. If $X \in \mathcal{O}$ is tame and $h : X \to \cup \sigma_i$ is a definable bijection with a collection of open simplices, then the *definable Euler characteristic* of X is

$$\chi(X) := \sum_i (-1)^{\dim \sigma_i},$$

where $\dim \sigma_i$ denotes the dimension of the open simplex σ_i. We understand that $\chi(\varnothing) = 0$ since this corresponds to the empty sum.

As one might expect, the definable Euler characteristic does not depend on a particular choice of triangulation [22, pp.70-71], as it is a definable homeomorphism invariant. However, as the next example shows, the definable Euler characteristic is not a homotopy invariant.

Example 2.5. The definable Euler characteristic of the open unit interval $X = (0, 1)$ is -1. Note that X is contractible to a point, which has definable Euler characteristic 1. The reader might notice that these computations coincide with the compactly-supported Euler characteristic, which can be defined in terms of the ranks of compactly-supported cohomology or Borel-Moore homology groups. However, this equivalence only makes sense for locally compact sets and not all o-minimal sets are locally compact.

Remark 2.6. We will often drop the prefix "definable" and simply refer to "the Euler characteristic" of a definable set.

The definable Euler characteristic satisfies the *inclusion-exclusion rule*:

Proposition 2.7 ([22] p.71). *For tame subsets $A, B \in \mathcal{O}$ we have*
$$\chi(A \cup B) + \chi(A \cap B) = \chi(A) + \chi(B).$$

Consequently, the definable Euler characteristic specifies a valuation on any o-minimal structure, i.e. it serves the role of a measure without the requirement that sets be assigned a non-negative value. This allows us to develop an integration theory called Euler calculus that is well-defined for so-called constructible functions, which we now define.

Definition 2.8. A *constructible function* $\phi : X \to \mathbb{Z}$ is an integer-valued function on a tame set X with the property that every level set is tame and only finitely many level sets are non-empty. The set of constructible functions with domain X, denoted $\mathrm{CF}(X)$, is closed under pointwise addition and multiplication, thereby making $\mathrm{CF}(X)$ into a ring.

Definition 2.9. The *Euler integral* of a constructible function $\phi : X \to \mathbb{Z}$ is the sum of the Euler characteristics of each of its level sets, i.e.

$$\int \phi \, d\chi := \sum_{n=-\infty}^{\infty} n \cdot \chi(\phi^{-1}(n)).$$

Note that constructibility of ϕ implies that only finitely many of the $\phi^{-1}(n)$ are non-empty so this sum is well-defined.

Like any good calculus, there is an accompanying suite of canonical operations in this theory—pullback, pushforward, convolution, etc.

Definition 2.10. Let $f : X \to Y$ be a tame mapping between between definable sets. Let $\phi_Y : Y \to \mathbb{Z}$ be a constructible function on Y. The *pullback* of ϕ_Y along f is defined pointwise by

$$f^* \phi_Y(x) = \phi_Y(f(x)).$$

The pullback operation defines a ring homomorphism $f^* : \mathrm{CF}(Y) \to \mathrm{CF}(X)$.

The dual operation of pushing forward a constructible function along a tame map is given by integrating along the fibers.

Definition 2.11. The *pushforward* of a constructible function $\phi_X : X \to \mathbb{Z}$ along a tame map $f : X \to Y$ is given by

$$f_* \phi_X(y) = \int_{f^{-1}(y)} \phi_X d\chi.$$

This defines a group homomorphism $f_* : \mathrm{CF}(X) \to \mathrm{CF}(Y)$.

Putting these two operations together allows one to define our first topological transform: the Radon transform.

Definition 2.12. Suppose $S \subset X \times Y$ is a locally closed definable subset of the product of two definable sets. Let π_X and π_Y denote the projections from the product onto the indicated factors. The *Radon transform with respect to S* is the group homomorphism $\mathcal{R}_S : \mathrm{CF}(X) \to \mathrm{CF}(Y)$ that takes a constructible function on X, $\phi : X \to \mathbb{Z}$, pulls it back to the product space $X \times Y$, multiplies by the indicator function of S before pushing forward to Y. In equational form, the Radon transform is

$$\mathcal{R}_S(\phi) := \pi_{Y*}[(\pi_X^* \phi) 1_S].$$

The following inversion theorem of Schapira [18] gives a topological criterion for the invertibility of the transform $\mathcal{R}_S$ in terms of the subset $S \subset X \times Y$.

Theorem 2.13 ([18] Theorem 3.1). *If $S \subset X \times Y$ and $S' \subset Y \times X$ have fibers S_x and S'_x in Y satisfying*

  (1) $\chi(S_x \cap S'_x) = \chi_1$ *for all $x \in X$, and*
  (2) $\chi(S_x \cap S'_{x'}) = \chi_2$ *for all $x' \neq x \in X$,*

then for all $\phi \in CF(X)$,

$$(\mathcal{R}_{S'} \circ \mathcal{R}_S)\phi = (\chi_1 - \chi_2)\phi + \chi_2 \left(\int_X \phi d\chi \right) 1_X.$$

In the next section we show how to use Schapira's result to deduce the injectivity properties of the Euler Characteristic and Persistent Homology Transforms.

3. Injectivity of the Euler characteristic transform

In the previous section we introduced background on definable sets, constructible functions and the Radon transform. We now specialize this material to the study of persistent-type topological transforms on definable subsets of $\mathbb{R}^n$. What makes these transforms "persistent" is that they study the evolution of topological invariants with respect to a real parameter. In this section we begin with the simpler of these two invariants, the Euler characteristic.

Definition 3.1. The *Euler Characteristic Transform* takes a constructible function ϕ on $\mathbb{R}^d$ and returns a constructible function on $S^{d-1} \times \mathbb{R}$ whose value at a direction v and real parameter $t \in \mathbb{R}$ is the Euler integral of the restriction of ϕ to the half space $x \cdot v \leq t$. In equational form, we have

$$\mathrm{ECT} : \mathrm{CF}(\mathbb{R}^d) \to \mathrm{CF}(S^{d-1} \times \mathbb{R}) \qquad \text{where} \qquad \mathrm{ECT}(\phi)(v, t) := \int_{x \cdot v \leq t} \phi \, d\chi.$$

For our applications of interest, it suffices to restrict this transform to compact definable subsets of $\mathbb{R}^d$ (which we call constructible sets), where we identify a definable subset $M \in \mathcal{O}_d$ with its associated constructible indicator function $\phi =$

1_M. When we fix $v \in S^{d-1}$ and let $t \in \mathbb{R}$ vary, we refer to $\mathrm{ECT}(M)(v, -)$ as the *Euler curve* for direction v. This allows us to equivalently view the Euler Characteristic Transform for a fixed $M \in \mathrm{CS}(\mathbb{R}^d)$ as a map from the sphere to the space of Euler curves (constructible functions on $\mathbb{R}$).
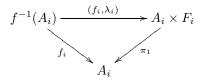
$$\mathrm{ECT}(M) : S^{d-1} \to \mathrm{CF}(\mathbb{R}) \qquad \text{where} \qquad \mathrm{ECT}(M)(v)(t) := \chi(M \cap \{x \mid x \cdot v \leq t\}).$$

The proof that the Euler curves are constructible is detailed in Lemma 3.4. We write $\mathrm{ECT}(M)(v)(\infty)$ as shorthand for the limit of $\mathrm{ECT}(M)(v)(t)$ as t goes to positive infinity. This limit exists by compactness and definability of M. Note that for every direction v this limit is the total Euler characteristic of M, i.e. $\mathrm{ECT}(M)(v)(\infty) = \chi(M)$.

Remark 3.2. The restriction to constructible sets CS permits us to ignore the difference between ordinary and compactly-supported Euler characteristic. This is because the intersection of a compact definable set M and the closed half-space $\{x \in \mathbb{R}^d \mid x \cdot v \leq t\}$ is compact and definable, and these two versions of Euler characteristic agree on compact sets. This restriction is not strictly necessary, but would require a reworking of certain aspects of persistent homology (reviewed in Section 4), which is beyond the scope of this paper.

The reader may wish to pause to consider why the ECT associates to a definable set M, viewed as the constructible function $\phi = 1_M$, a constructible function on $S^{d-1} \times \mathbb{R}$. This, along with other finiteness concerns, is addressed by the following lemma, which in turn depends on two versions of the Trivialization Theorem of o-minimal geometry.

Theorem 3.3 (The Trivialization Theorems of [22]). *If $S \subseteq \mathbb{R}^n$ and $A \subseteq \mathbb{R}^m$ are definable subsets and $f : S \to A$ is a continuous definable map, then the **Trivialization Theorem** [22, Ch. 9, 1.2] states that there is a finite partition of A into finitely many definable subsets $A_1, \ldots, A_M$ such that over each A_i there is a definable subset $F_i \subseteq \mathbb{R}^N$, for some N, a definable map $\lambda_i : f^{-1}(A_i) \to F_i$, such that the product map (f_i, λ_i), where $f_i := f|_{f^{-1}(A_i)}$, is a homeomorphism making the following diagram commute:*

$$
\begin{array}{ccc}
f^{-1}(A_i) & \xrightarrow{\ (f_i, \lambda_i)\ } & A_i \times F_i \\
& {\scriptstyle f_i} \searrow \quad \quad \swarrow {\scriptstyle \pi_1} & \\
& A_i &
\end{array}
$$

*Moreover, the **Trivialization Theorem with Distinguished Subsets** [22, Ch. 9, 1.7] states that if S has definable subsets $S_1, \ldots, S_k$, then we can refine the partition of A to guarantee that for each A_i there exists definable subsets $F_1, \ldots, F_k$ of F with $(f_i, \lambda_i)(f^{-1}(A_i) \cap S_j) = A_i \times F_j$. In other words, we can refine the fiber to respect the distinguished subsets of S.*

Lemma 3.4 (Finitely Many Topological Changes). *For any definable set $M \subseteq \mathbb{R}^d$ and any direction $v \in S^{d-1}$ the topological, i.e. homeomorphism, type of the sublevel sets $M_{v,t} = M \cap \{x \mid x \cdot v \leq t\}$ can only change finitely many times as a function of t. Moreover there is a uniform bound $\kappa_M \geq 0$, depending on M, on the number of changes in the topological type when viewed in any direction. As a consequence, for any direction $v \in S^{d-1}$ the Euler curve $\mathrm{ECT}(M)(v, -)$ is a constructible function over $\mathbb{R}$.*

Proof. One can associate to a definable $M \subseteq \mathbb{R}^d$ another definable set

$$X_M := \{(x, v, t) \in M \times S^{d-1} \times \mathbb{R} \mid x \cdot v \le t\}.$$

To see that X_M is definable we note that S^{d-1} is the set of points satisfying

$$x_1^2 + \cdots x_d^2 = 1$$

making it an algebraic set, which by [22, Ch. 2, 2.11] is definable. By taking products we can conclude that $\mathbb{R}^d \times S^{d-1} \times \mathbb{R}$ is definable as well. The collection of half-spaces $H := \{(x, v, t) \in \mathbb{R}^d \times S^{d-1} \times \mathbb{R} \mid x \cdot v \le t\}$ is also semialgebraic and hence definable. Consequently, the intersection of $M \times S^{d-1} \times \mathbb{R}$ and H, which is X_M, must be definable as well. Finally, by our Definition 2.1 we know the projection map onto the second and third factors $\pi_{2,3} : \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R} \to \mathbb{R}^d \times \mathbb{R}$ is definable and hence, by [22, Ch. 1, 2.3(ii)], the restriction of this projection map to X_M, i.e.

$$\pi : X_M \to S^{d-1} \times \mathbb{R} \qquad \text{where} \qquad \pi^{-1}(v, t) = M_{v,t}$$

is a continuous definable map as well.

By applying the Trivialization Theorem [22, Ch. 9, 1.2] to π we know that there is a finite partition of $S^{d-1} \times \mathbb{R}$ into definable sets $A_1, \ldots, A_n$ over which the map $\pi : X_M \to S^{d-1} \times \mathbb{R}$ is definably trivial, i.e. the map restricts to a trivial fiber bundle over each A_i. The trivialization over the sets A_i guarantees that the topological type of the sublevel sets $M_{v,t}$ changes only finitely many times when viewed as a function of v and t, but the statement of the lemma asks for finiteness when viewing the shape M in a fixed direction v. To obtain the statement we want, we view the A_i as distinguished subsets of $S^{d-1} \times \mathbb{R}$ and then apply the Trivialization Theorem with Distinguished Subsets [22, Ch. 9, 1.7] to the projection map $\pi_1 : S^{d-1} \times \mathbb{R} \to S^{d-1}$. This provides us with a partition of the sphere S^{d-1} into definable subsets $B_1, \ldots, B_m$ so that over each B_i there are n definable subsets $F_{i1}, \ldots, F_{in}$ of $\mathbb{R}$ so that the trivialization carries each $\pi_1^{-1}(B_i) \cap A_j$ to $B_i \times F_{ij}$. We note that each F_{ij} need not be connected, but each F_{ij} will be a finite union of singleton sets and connected intervals, i.e. definable 0 and 1-cells. The number of changes in the topology of $M_{v,t}$ for any $v \in B_i$ is then bounded by the sum of the number of 0 and 1-cells in the partition $F_{i1}, \ldots, F_{in}$ of $\mathbb{R}$. Taking the maximum of this number over $B_i, \ldots, B_m$ provides the uniform bounded κ_M on the number of topological changes when viewed in any direction. $\square$

This proof of Lemma 3.4 provides an important alternative characterization of the Euler Characteristic Transform: The ECT is simply the pushforward of the indicator function of X_M along the definable map π, which defines a constructible function on $S^{d-1} \times \mathbb{R}$.

We now use Schapira's inversion theorem, recalled as Theorem 2.13, to prove our first injectivity result.

Theorem 3.5. *Let $CS(\mathbb{R}^d)$ be the set of constructible sets, i.e. compact definable sets. The map $\mathrm{ECT} : CS(\mathbb{R}^d) \to CF(S^{d-1} \times \mathbb{R})$ is injective. Equivalently, if M and M' are two constructible sets that determined the same association of directions to Euler curves, then they are, in fact, the same set. Said symbolically:*

$$\mathrm{ECT}(M) = \mathrm{ECT}(M') : S^{d-1} \to CF(\mathbb{R}) \qquad \Rightarrow \qquad M = M'.$$

Proof. Let $M \in \mathrm{CS}(\mathbb{R}^d)$. Let W be the hyperplane defined by $\{x \cdot v = t\}$. By the inclusion-exclusion property of the definable Euler characteristic

$$
\begin{aligned}
\chi(M \cap W) &= \chi(\{x \in M : x \cdot v = t\}) \\
&= \chi(\{x \in M : x \cdot v \leq t\} \cap \{x \in M : x \cdot (-v) \leq -t\}) \\
&= \chi(\{x \in M : x \cdot v \leq t\}) + \chi(\{x \in M : x \cdot (-v) \leq -t\}) - \chi(M) \\
&= \mathrm{ECT}(M)(v, t) + \mathrm{ECT}(M)(-v, -t) - \mathrm{ECT}(M)(v)(\infty).
\end{aligned}
$$

This means that from $\mathrm{ECT}(M)$ we can deduce the $\chi(M \cap W)$ for all hyperplanes $W \in \mathrm{AffGr}_d$.

Let S be the subset of $\mathbb{R}^d \times \mathrm{AffGr}_d$ where $(x, W) \in S$ when x is in the hyperplane W. For simplicity, we denote the projection to $\mathbb{R}^d$ by π_1 and the projection to AffGr_d by π_2. For this choice of S the Radon transform of the indicator function 1_M for $M \in \mathcal{O}_d$ at $W \in \mathrm{AffGr}_d$ is

$$
\begin{aligned}
(\mathcal{R}_S 1_M)(W) &= (\pi_2)_*[(\pi_1^* 1_M) 1_S](W) \\
&= \int_{(x, W) \in S} (\pi_1^* 1_M) \, d\chi \\
&= \int_{x \in M \cap W} (\pi_1^* 1_M) \, d\chi \\
&= \chi(M \cap W).
\end{aligned}
$$

This implies that from $\mathrm{ECT}(M)$ we can derive $\mathcal{R}_S(1_M)$.

Similarly let S' be the subset of $\mathrm{AffGr}_d \times \mathbb{R}^d$ where $(W, x) \in S'$ when x is in the hyperplane W. For all $x \in \mathbb{R}^d$, $S_x \cap S'_x$ is the set of hyperplanes that go through x and hence is $S_x \cap S'_x = \mathbb{R}P^{d-1}$ and $\chi(S_x \cap S'_x) = \frac{1}{2}(1 + (-1)^{d-1})$. For all $x \neq x' \in \mathbb{R}^d$, $S_x \cap S'_{x'}$ is the set of hyperplanes that go through x and x' and hence is $S_x \cap S'_{x'} = \mathbb{R}P^{d-2}$ and $\chi(S_x \cap S'_{x'}) = \frac{1}{2}(1 + (-1)^{d-2})$. Applying Theorem 2.13 yields

$$
(\mathcal{R}_{S'} \circ \mathcal{R}_S)(1_M) = (-1)^{d-1} 1_M + \frac{1}{2}(1 + (-1)^{d-2})\chi(M) 1_{\mathbb{R}^d}.
$$

Note that if $\mathrm{ECT}(M) = \mathrm{ECT}(M')$ then $\mathcal{R}_S 1_M = \mathcal{R}_S 1_{M'}$, since ECT determines the Euler characteristic of every slice. Moreover, if $\mathrm{ECT}(M) = \mathrm{ECT}(M')$, then $\chi(M) = \chi(M')$, which by inspecting the inversion formula above further implies that $1_M = 1_{M'}$ and hence $M = M'$. $\qquad\square$

4. Injectivity of the persistent homology transform

The primary transform of interest for this paper is the Persistent Homology Transform, which was first introduced in [20] and was initially defined for a PL embedded simplicial complex in $\mathbb{R}^d$. The reader is encouraged to consult [20] (and [1] for a related precursor) for a more complete treatment of the PHT in that setting. Here, we illustrate how this transform can be defined for any constructible set $M \in \mathrm{CS}(\mathbb{R}^d)$.

As already noted, given a direction $v \in S^{d-1}$ and a value $t \in \mathbb{R}$, the sublevel set $M_{v,t} := \{x \in M \mid x \cdot v \leq t\}$ is the intersection of the constructible set M with a closed half-space. This intersection has various topological summaries, one of them being the (definable) Euler Characteristic χ. One can also consider the (cellular) homology with field coefficients H_k, which is defined in each degree $k \in \{0, 1, \ldots, n\}$. These are vector spaces that summarize topological content of any suitably nice

topological space X, which in this paper will always be spaces of the form $M_{v,t}$. In low degrees the interpretations of these homology vector spaces for a space X are as follows: $H_0(X)$ is a vector space with basis given by connected components, $H_1(X)$ is a vector space spanned by "holes" or closed loops that are not the boundaries of embedded disks, $H_2(X)$ is a vector space spanned by "voids" or closed two-dimensional (possibly self-intersecting) surfaces that are not the boundaries of an embedded three dimensional space. The higher homologies $H_k(X)$ are understood by analogy: these are vector spaces spanned by closed (i.e. without boundary) k-dimensional subspaces of X that are themselves not the boundaries of $k + 1$-dimensional spaces. The dimension of $H_k(X)$ is called the *Betti number* $\beta_k(X)$, which for $X \subseteq \mathbb{R}^d$ always satisfy $\beta_k(X) = 0$ for $k \geq d$.

The proof that ordinary Euler characteristic is a topological invariant is best understood via homology. Indeed, the Betti numbers determine the Euler characteristic via an alternating sum:

$$\chi(X) = \beta_0(X) - \beta_1(X) + \beta_2(X) - \beta_3(X) + \cdots$$

However, one feature that homology enjoys that the Euler characteristic does not is *functoriality*, which is the property that any continuous transformation of spaces $f : X \to Y$ induces a linear transformation of homology vector spaces $f_k : H_k(X) \to H_k(Y)$ for each degree k. This is the key feature that defines sublevel set persistent homology.

Definition 4.1. Let $M \in \mathrm{CS}(\mathbb{R}^d)$ be a compact definable set. For any vector $v \in \mathbb{R}^d$ define the *height function in direction* v, $h_v(x) = \langle v, x \rangle$ as the restriction of the inner product $\langle v, \cdot \rangle$ to points $x \in M$. The *sublevel set persistent homology group* in degree k between s and t is

$$PH_k(M, h_v)(s,t) = \mathrm{im}\, \iota_k^{s \to t} : H_k(M_{v,s}) \to H_k(M_{v,t}).$$

The remarkable feature of persistent homology is that one can encode the persistent homology groups for every pair of values $s \leq t$ using a finite number of points in the extended plane. This is done via the persistence diagram.

Definition 4.2. Let $\mathbb{R}^{2+}$ be the part of the extended plane that is above the diagonal, i.e. $\mathbb{R}^{2+} := \{(b,d) \in \{-\infty\} \cup \mathbb{R}) \times (\mathbb{R} \cup \{\infty\}) \mid b \leq d\}$. The *persistence diagram* in degree k associated to $M \in \mathrm{CS}(\mathbb{R}^d)$ when filtered by sublevel sets of $h_v(x) = \langle v, x \rangle$ is the *unique* finite multiset of points $\mathcal{B}_k = \{((b_i^k, d_i^k); n_i)\} \subset \mathbb{R}^{2+}$ with the property that for every pair of values $s \leq t$ the following equality holds

$$\dim PH_k(M, h_v)(s,t) := \mathrm{rank}\, \iota_k^{s \to t} = \mathrm{card}\{PH_k(M, h_v) \cap [-\infty, s] \times (t, \infty]\}.$$

Since the the persistence diagram $\mathcal{B}_k$ completely encodes the ranks of the sublevel set persistent homology groups $PH_k(M, h_v)$, we will pass between these two notations freely, depending on what needs to be emphasized. We further note that the persistence diagram is also called the *barcode*, which relies on the interpretation of each point $(b, d) \in \mathbb{R}^{2+}$ as an interval $I = [b, d) \subseteq \mathbb{R}$ where $b = \inf I$ and $d = \sup I$.

Remark 4.3 (Persistence and o-Minimality). The existence of the persistence diagram is a non-trivial result and depends on certain tameness properties of the maps $\iota_k^{s \to t}$. Finite dimensionality of all of the homology vector spaces $H_k(M_{v,t})$ suffices, but in the o-minimal setting things are even better behaved because the dimension of the persistent homology groups can only change finitely times. Recall that the Triangulation Theorem, Theorem 2.3, guarantees that for each v and t $M_{v,t}$ has a

finite decomposition into cells and so, the cellular homology is finite. The fact that the topology of $M_{v,t}$ can only change finitely many times as a function of t was proved in Lemma 3.4.

To consider persistence diagrams $PH_k(M, h_v)$ associated to different directions $v \in S^{d-1}$ we consider the set of all persistence diagrams, which can be topologized in several ways, as explained after Definition 4.4.

Definition 4.4. *Persistence Diagram Space*, written Dgm, is the set of all possible countable multisets of $\mathbb{R}^{2+} := \{(b, d) \in (\{-\infty\} \cup \mathbb{R}) \times (\mathbb{R} \cup \{\infty\}) \mid b \leq d\}$ where the number of points of the form (b, ∞) and $(-\infty, d)$ are finite and $\sum_{d-b<\infty} d - b < \infty$. Points of the form (b, ∞) or $(-\infty, d)$ are called *essential classes* and points of the form (b, d) where neither coordinate is ∞ are called *inessential classes*. For persistence diagrams that encode the sublevel set persistent homology $PH_k(M, h_v)$ of a constructible set M there are no points of the form $(-\infty, d)$.

We have used the term "space" with the implication that there is a topology on the set of all persistence diagrams. Indeed this topology comes from various choices of metrics on the set of persistence diagrams, which are phrased in terms of matchings of points between two persistence diagrams. Since a persistence diagram is technically a multiset, we append an additional coordinate to serve as a labelling index so that each $\mathcal{B}$ can be regarded as a genuine set.

Definition 4.5. Suppose $\mathcal{B} = \{(I; j) \mid (I; j) \in \mathbb{R}^{2+} \times \mathbb{N}\}$ and $\mathcal{B}' = \{(I'; j) \mid (I'; j) \in \mathbb{R}^{2+} \times \mathbb{N}\}$ are two persistence diagrams, viewed as sets rather than multisets; i.e. $(I; j)$ is to be interpreted as the jth copy of the interval I. A *matching* is a partial bijection $\sigma : \mathcal{B} \to \mathcal{B}'$, i.e. a choice of subset $\mathcal{M} \subseteq \mathcal{B}$, called the domain $\text{dom}(\sigma)$, and an injection $\sigma : \mathcal{M} \to \mathcal{B}'$. We write the complement of the domain of σ as $\text{dom}^c(\sigma) := \mathcal{B} - \text{dom}(\sigma)$ and the complement of the image of σ as $\text{im}^c(\sigma) := \mathcal{B}' - \text{im}(\sigma)$; collectively these are called the *unmatched points* of σ.

For this paper we will always promote a partial bijection σ to an actual bijection via the introduction of diagonal images. For a point $I = (b, d) \in \mathbb{R}^{2+}$ where neither coordinate is ∞, we define the *diagonal image of I* as $\Delta(I) = (\frac{b+d}{2}, \frac{b+d}{2})$. Associated to any partial bijection $\sigma : \mathcal{B} \to \mathcal{B}'$ is an actual bijection $\tilde{\sigma} : \mathcal{B}(\sigma) \to \mathcal{B}'(\sigma)$ where

$$\mathcal{B}(\sigma) := \text{dom}(\sigma) \cup \text{dom}^c(\sigma) \cup \bigcup_{(I', j') \in \text{im}^c(\sigma)} (\Delta(I'); j')$$

and

$$\mathcal{B}'(\sigma) := \text{im}(\sigma) \cup \text{im}^c(\sigma) \cup \bigcup_{(I, j) \in \text{dom}^c(\sigma)} (\Delta(I); j).$$

The map $\tilde{\sigma}$ now matches points that were previously unmatched by σ with their corresponding diagonal images. For ease of notation, we drop the superscript tilde on $\tilde{\sigma}$ and simply write σ for the extended map $\mathcal{B}(\sigma) \to \mathcal{B}'(\sigma)$.

The (p, q)-*cost* of a matching σ, where $\sigma_j(I)$ denotes the $\mathbb{R}^{2+}$ coordinates of $\sigma(I; j)$, is

$$W_{p,q}(\sigma) = \left(\sum_{(I,j) \in \mathcal{B}(\sigma)} \| I - \sigma_j(I) \|_q^p \right)^{1/p}.$$

As a reminder, the ℓ^q distance between two matched points $(I; j) = (b_j, d_j; j)$ and $\sigma(I, j) = (b_k, d_k; k)$ is

$$\|I - \sigma_j(I)\|_q = (|b_j - b_k|^q + |d_j - d_k|^q)^{1/q}$$

with the understanding that if $d_j = \infty$, then we must have $d_k = \infty$ and this distance collapses to $|b_j - b_k|$. The ℓ^∞ distance is $\|I - \sigma_j(I)\|_\infty = \max\{|b_j - b_k|, |d_j - d_k|\}$.

The *Wasserstein (p, q)-distance* between two diagrams $\mathcal{B}$ and $\mathcal{B}'$ is then the infimum of this matching cost over all matchings, i.e.

$$W_{p,q}(\mathcal{B}, \mathcal{B}') := \inf_{\sigma: \mathcal{B} \to \mathcal{B}'} W_{p,q}(\sigma).$$

As noted in [19], for fixed p, the Wasserstein (p, q)-distances are all bi-Lipschitz equivalent. The convention set by [7] was to refer to the Wasserstein (p, ∞)-distance as *the Wasserstein p-distance W_p*, but there are good reasons to adopt $q = p$ or $q = 1$ as conventions as well. Unless explicitly stated to the contrary, our default assumption is that $p = q$.

We note that the Wasserstein ∞-distance is also called the *bottleneck distance*, for which we reserve the special notation

$$d_B(\mathcal{B}, \mathcal{B}') := W_\infty(\mathcal{B}, \mathcal{B}') = \inf_{\sigma: \mathcal{B} \to \mathcal{B}'} \max_{(I,j) \in \mathcal{B}(\sigma)} \|I - \sigma_j(I)\|_\infty.$$

Although the bottleneck distance is the preferred distance for many theoretical purposes, in [20] the 1-Wasserstein distance was used. For more details about the geometry of the space of persistence diagrams under p-Wasserstein metrics see [21].

Definition 4.6. The *Persistent Homology Transform* PHT of a constructible set $M \in \mathrm{CS}(\mathbb{R}^d)$ is the map $\mathrm{PHT}(M) : S^{d-1} \to \mathrm{Dgm}^d$ that sends a direction $v \in S^{d-1}$ to the persistent diagrams gotten by filtering M in the direction of v, recording one diagram for each homological degree $0 \le k \le d - 1$, i.e.

$$\mathrm{PHT}(M) : v \mapsto (PH_0(M, h_v), PH_1(M, h_v), \ldots, PH_{d-1}(M, h_v)).$$

Letting the set M vary gives us the map

$$\mathrm{PHT} : \mathrm{CS}(\mathbb{R}^d) \to C(S^{d-1}, \mathrm{Dgm}^d),$$

where $C(S^{d-1}, \mathrm{Dgm}^d)$ is the set of continuous functions from S^{d-1} to Dgm^d, the latter being equipped with some Wasserstein p-distance.

Before moving on with the remainder of the paper, we offer a sheaf-theoretic interpretation of the Persistent Homology Transform, which is not necessary for the remainder of the paper. The reader that is uninterested in sheaves can safely ignore the following remark.

Remark 4.7 (Sheaf-Theoretic Definition of the PHT). Extending Lemma 3.4, we know that associated to any constructible set M is a space

$$X_M := \{(x, v, t) \in M \times S^{d-1} \times \mathbb{R} \mid x \cdot v \le t\}$$

and a map $\pi : X_M \to S^{d-1} \times \mathbb{R}$ whose fiber over (v, t) is the sublevel set $M_{v,t}$. The *derived Persistent Homology Transform* is the right derived pushforward of the constant sheaf on X_M along the map π, written $R\pi_* k_{X_M}$. The associated cohomology sheaves $R^i \pi_* k_{X_M}$ of this derived pushforward, called the Leray sheaves in [9], have stalk value at (v, t) the ith cohomology of the sublevel set $M_{v,t}$. If we restrict the sheaf $R^i \pi_* k_{X_M}$ to the subspace $\{v\} \times \mathbb{R}$, then one obtains a constructible sheaf that is equivalent to the persistent (co)homology of the filtration of M viewed

in the direction of v. The persistence diagram in degree i is simply the expression of this restricted sheaf in terms of a direct sum of indecomposable sheaves.

4.1. **Continuity of the PHT.** It should be noted that when M is an embedded simplicial complex, continuity of the resulting map $\mathrm{PHT}(M) : S^{d-1} \to \mathrm{Dgm}^d$ was proved as Lemma 2.1 of [20]. To ensure that the above definition generalizes to constructible sets, we must prove a generalization of that lemma.

Lemma 4.8 (cf. Lemma 2.1 [20]). *For a constructible set $M \in CS(\mathbb{R}^d)$, the map $\mathrm{PHT}(M) : S^{d-1} \to Dgm^d$ is continuous, where S^{d-1} is given the Euclidean distance and Dgm uses any Wasserstein p-distance with $1 \leq p \leq \infty$.*

Proof. First we note that since M is compact, there is a bound D_M on the distance from any point in M to the origin. This implies that for any two directions v_1 and v_2, the height functions h_{v_1} and h_{v_2} have a pointwise bound given by

$$|h_{v_1}(x) - h_{v_2}(x)| = |x \cdot v_1 - x \cdot v_2| \leq \|x\|_2 \cdot \|v_1 - v_2\|_2 \leq D_M \|v_1 - v_2\|_2$$

Consequently, we have that $\|h_{v_1} - h_{v_2}\|_\infty \leq D_M \|v_1 - v_2\|_2$.

The bottleneck stability theorem of [6] guarantees that for each homological degree $k \geq 0$ the bottleneck distance between the persistence diagrams is bounded by the L^∞ distance between the functions. It suffices to prove continuity of $\mathrm{PHT}(M)$ in each coordinate, so without loss of generality, we refer to the kth persistence diagram for h_{v_1} as $\mathcal{B}_1$ and the kth persistence diagram for h_{v_2} as $\mathcal{B}_2$. The bottleneck stability theorem guarantees that

$$d_B(\mathcal{B}_1, \mathcal{B}_2) \leq \|h_{v_1} - h_{v_2}\|_\infty \leq D_M \|v_1 - v_2\|_2.$$

This implies continuity of $\mathrm{PHT}(M)$ when using the bottleneck distance on persistence diagrams, which is the Wasserstein ∞-distance, because in order to make the left hand side less than ϵ, we need only bound the Euclidean distance $\|v_1 - v_2\|_2$ by $\delta = \epsilon/D_M$.

For general p we use Lemma 3.4, which showed that for a constructible set M there is a bound κ_M on the number of critical values when viewed in any direction v.

As a consequence of this lemma, the number of points in the persistence diagrams $\mathcal{B}_1$ and $\mathcal{B}_2$ are both bounded above by κ_M, since each point in the diagram corresponds to a topological change when filtering M in the direction v_i. Let $\sigma : \mathcal{B}_1 \to \mathcal{B}_2$ be a matching that realizes the bottleneck distance $\epsilon := d_B(\mathcal{B}_1, \mathcal{B}_2)$, so that in particular $\|I - \sigma_j(I)\|_\infty \leq \epsilon$ for all $(I, j) \in \mathcal{B}_1(\sigma)$. We note that it's possible that σ matches every point in $\mathcal{B}_1$ to the diagonal, which would lead to an augmented diagram $\mathcal{B}_1(\sigma)$ of cardinality at most $2\kappa_M$. Since the Wasserstein p-distance infimizes the (p, p)-cost over all matchings, we can now say that

$$W_p(\mathcal{B}_1, \mathcal{B}_2) \leq \left(\sum_{(I,j) \in \mathcal{B}_1(\sigma)} \|I - \sigma_j(I)\|_p^p \right)^{1/p} \leq (4\kappa_M \epsilon^p)^{1/p} = (4\kappa_M)^{1/p}\epsilon.$$

Since we can control ϵ in terms of $\|v_1 - v_2\|_2$, this proves continuity of the PHT for the Wasserstein p-distance for $p \in [1, \infty]$. $\square$

4.2. **Continuity of Betti curves and Euler curves.** The goal of this section will be the statement and proof of a persistent analog of the classical result that homology determines the Euler characteristic via Betti numbers. This persistent

analog will also feature the extension of the continuity result of Lemma 4.8 to the Betti Curve Transform (Definition 4.11), which will in turn imply continuity of the Euler Characteristic Transform. Each of these results will require the specification of a metric on constructible functions over the real line, i.e. $CF(\mathbb{R})$, which we now do.

Definition 4.9. For every $1 \leq p < \infty$ we define the L^p extended pseudo-metric (distance) on constructible functions as follows: For $f, g \in CF(\mathbb{R})$ we set

$$d_p(f, g) = \left(\int |f(t) - g(t)|^p dt \right)^{1/p}$$

when this integral exists and ∞ when it does not. Similarly, we define the L^∞ distance to be $d_\infty(f, g) := \sup_{t \in \mathbb{R}} |f(t) - g(t)|$ when the right hand side exists and ∞ when it does not.

We use this distance to prove the following preparatory lemma, which details the passage from persistence diagrams to constructible functions.

Lemma 4.10 (Continuity of the Diagram to Function Map). *Recall that every persistence diagram $\mathcal{B} = \{(I, j)\}$ can be viewed as a multiset of intervals in $\mathbb{R}$. Let $\Phi(\mathcal{B})$ be the constructible function associated to $\mathcal{B}$ that sums the indicator functions supported on each interval I appearing in $\mathcal{B}$, i.e. $\Phi(\mathcal{B}) = \sum_{(I,j) \in \mathcal{B}} \mathbb{1}_I$. Let Dgm_κ denote the subset of persistence diagram space with fewer than κ off-diagonal points.*

For every $p \in [1, \infty)$, the map $\Phi : Dgm_\kappa \to CF(\mathbb{R})$ is continuous when Dgm_κ is equipped with the Wasserstein p-distance and $CF(\mathbb{R})$ is equipped with the L^p distance of Definition 4.9. Additionally, this distance can be bound in terms of the bottleneck distance as follows

$$d_p(\Phi(\mathcal{B})), \Phi(\mathcal{B}'))^p \leq 2\kappa^p d_B(\mathcal{B}, \mathcal{B}').$$

Proof. Since $W_p(\mathcal{B}, \mathcal{B}') \geq d_B(\mathcal{B}, \mathcal{B}')$ for all $p \in [1, \infty)$ to show that that Φ is continuous when Dgm_κ is equipped with the Wasserstein q-distance it is sufficient to show it is continuous with respect to the bottleneck distance.

We note that we must exclude $p = \infty$ from the statement of the lemma, because a persistence diagram $\mathcal{B}$ is distance zero from the same diagram where a point is added on the diagonal, e.g. $\mathcal{B} \cup (b, b)$. In this case, the constructible functions $\Phi(\mathcal{B})$ and $\Phi(\mathcal{B} \cup (b, b)) = \Phi(\mathcal{B}) + \mathbb{1}_b$ are distance one away in the L^∞ distance, thus proving that continuity of this map for $p = \infty$ is not possible.

For $p \neq \infty$, we remark that the map Φ is intuitively continuous because a small variation in points in the diagram $\mathcal{B}$ results in a small variation in the endpoints of the simple functions making up $\Phi(\mathcal{B})$. By construction we know that

$$d_p(\Phi(\mathcal{B}), \Phi(\mathcal{B}'))^p = \int |\sum_{(I,j)} \mathbb{1}_I - \sum_{(I',j')} \mathbb{1}_{I'}|^p.$$

To quantify this small variation precisely, let $\sigma : \mathcal{B} \to \mathcal{B}'$ be a matching that realizes the 1-Wasserstein distance $W_1(\mathcal{B}, \mathcal{B}')$, where we work with the ℓ^1 distance between points in the persistence diagram. We note that since we assume the number of points in $\mathcal{B}$ and $\mathcal{B}'$ are both less than κ, then this infimum must be realized. Without loss of generality, we can assume that two off-diagonal points will be matched to each other whenever matching to their diagonal images has the same, or greater, cost.

Consider the function $f = \sum_{(I,j)} \mathbb{1}_I - \sum_{(I',j')} \mathbb{1}_{I'}$. To bound $\int |f|^p$ we will use bounds on $\int |f|$ and f. Let $I \Delta \sigma_j(I)$ denote the symmetric difference between intervals I and $\sigma_j(I)$. As already observed in [5, Prop. 1.2], we can bound the integral of $|f|$ by

$$\int |f| \leq \int |\sum_{(I,j)} \mathbb{1}_{I \Delta \sigma_j(I)}| = \sum_{(I,j)} \int |\mathbb{1}_{I \Delta \sigma_j(I)}| = W_1(\mathcal{B}, \mathcal{B}').$$

We also claim that $0 \leq |f(x)| \leq \kappa$ for all x. To see this, first observe that there are at most κ points in each of $\mathcal{B}$ and $\mathcal{B}'$. If $f(x) > \kappa$ then there must be intervals $I \in \mathcal{B}$ and $I' \in \mathcal{B}'$ such that $x \in I \cap I'$ and both I and I' are matched to the diagonal under σ. However, the cost of matching both I and I' to the diagonal is $\mu(I) + \mu(I')$, where μ denotes the length of each interval, whereas the cost of matching I to I' is $\mu(I \Delta I') \leq \mu(I) + \mu(I')$ as I and I intersect. This contradicts our assumption that σ was an optimal matching for $W_{1,1}$.

Using the bound $0 \leq |f(x)| \leq \kappa$, we know that $|f(x)|^p \leq \kappa^{p-1}|f(x)|$ for all x. Thus the integral $\int |f|^p$ is bounded by $\kappa^{p-1} \int |f|$ and hence

$$d_p(\Phi(\mathcal{B}), \Phi(\mathcal{B}'))^p \leq \kappa^{p-1} W_1(\mathcal{B}, \mathcal{B}').$$

To finish the proof we only need to observe that if $\mathcal{B}, \mathcal{B}' \in \mathrm{Dgm}_\kappa$ then both $W_1(\mathcal{B}, \mathcal{B}') \leq 2\kappa W_p(\mathcal{B}, \mathcal{B}')$ (for $p \in [1, \infty)$) and $W_1(\mathcal{B}, \mathcal{B}') \leq 2\kappa d_B(\mathcal{B}, \mathcal{B}')$. □

Recall that for fixed $M \in \mathrm{CS}(\mathbb{R}^d)$ and $v \in S^{d-1}$ the Betti number in degree k of the sublevel set $M_{v,t}$ is

$$\beta_k(M_{v,t}) = \dim H_k(M_{v,t}).$$

For each $t \in \mathbb{R}$, the right hand side of the above equation is determined by the cardinality of the intersection of the persistence diagram $PH_k(M, h_v)$ with the half-open quadrant $[-\infty, t] \times (t, \infty]$. However, by allowing t to vary, one obtains the *Betti curve* $\beta_{k,v}(t)$ for the direction v in degree k as the sum of indicator functions supported on the intervals appearing in the persistence diagram for $PH_k(M, h_v)$, i.e. $\Phi(PH_k(M, h_v))$ where Φ was defined in Lemma 4.10. These observations provide a third topological transform that stands between the PHT and the ECT.

Definition 4.11 (Betti Curve Transform). Let M be a constructible set and let $\mathrm{PHT}(M) : S^{d-1} \to \mathrm{Dgm}^d$ be the associated persistent homology transform. Applying the map $\Phi : \mathrm{Dgm} \to \mathrm{CF}(\mathbb{R})$ in each coordinate yields the *Betti Curve Transform* (BCT)

$$\mathrm{BCT}(M) : v \mapsto (\beta_{0,v}(t), \beta_{1,v}(t), \ldots, \beta_{d-1,v}(t)).$$

We now prove that the Betti Curve Transform is continuous as a function of the viewing direction $v \in S^{d-1}$. Although this is an immediate corollary of Lemma 4.10, we will provide a more concrete "for all ϵ, there exists a δ" proof using the Bottleneck distance bound proved earlier.

Corollary 4.12 (Continuity of Betti Curves). *Fix a constructible set M. For each $p \in [1, \infty)$, the Betti Curve Transform $\mathrm{BCT}(M) : S^{d-1} \to CF(\mathbb{R})^d$ is continuous, where each coordinate $\mathrm{BCT}(M)_k$ is viewed as a map from the sphere to $CF(\mathbb{R})$, the latter using the L^p distance on constructible functions.*
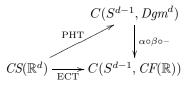
Proof. It suffices to prove continuity in each coordinate, which puts us in the setting of Lemma 4.10. As was proved in Lemma 3.4, we can bound the number of off-diagonal points in any persistence diagram $PH_k(M, h_v)$ by κ_M. This serves as the

value κ required by Lemma 4.10. Since Lemma 4.8 proved that the diagrams are continuous as a function of sphere direction, Lemma 4.10 proves that the resulting constructible functions also vary continuously. This completes the proof.

To see a more detailed special case of this argument, we note that bottleneck stability proves that for any pair of directions v_1 and v_2 the bottleneck distance between the associated diagrams is bounded as $d_B(\mathcal{B}_1, \mathcal{B}_2) \leq D_M \|v_1 - v_2\|_2$. Lemma 4.10 proves that the L^p distance between the associated Betti curves is bounded as $d_p(\Phi(\mathcal{B}_1), \Phi(\mathcal{B}_2)) \leq (2^{p+1} \kappa_M^p d_B(\mathcal{B}_1, \mathcal{B}_2))^{1/p}$. Consequently to make $d_p(\Phi(\mathcal{B}_1), \Phi(\mathcal{B}_2)) \leq \epsilon$ it suffices to make $\|v_1 - v_2\| \leq \delta$ where $\delta = \frac{\epsilon^p}{2\kappa_M^p D_M}$. $\qquad \square$

Proposition 4.13. *The Persistent Homology Transform (PHT) determines the Euler Characteristic Transform (ECT), i.e. we have the following commutative diagram of maps*

$$
\begin{array}{ccc}
 & & C(S^{d-1}, Dgm^d) \\
 & \nearrow^{\text{PHT}} & \downarrow^{\alpha \circ \beta \circ -} \\
CS(\mathbb{R}^d) & \xrightarrow[\text{ECT}]{} & C(S^{d-1}, CF(\mathbb{R}))
\end{array}
$$

where $\beta := \Phi^d : Dgm^d \to CF(\mathbb{R})^d$ is the map that takes the PHT to the Betti Curve Transform BCT. The map $\alpha : CF(\mathbb{R})^d \to CF(\mathbb{R})$ takes the alternating sum of d constructible functions pointwise. Finally, we recall that the set $C(S^{d-1}, CF(\mathbb{R}))$ refers to the set of continuous maps from S^{d-1}, equipped with the restriction of the Euclidean norm, to $CF(\mathbb{R})$, equipped with any L^p norm so long as $p \in [1, \infty)$; again the case $p = \infty$ is intentionally excluded, since continuity there is impossible.

Proof. Corollary 4.12 proves that each Betti curve is continuous as a function of $v \in S^{d-1}$. This proof used, Lemma 4.10, which proved that the map from $Dgm \to CF(\mathbb{R})$ that takes a persistence diagram to the sum of indicator functions supported on each interval is continuous when the domain is equipped with a Wasserstein p-norm and $CF(\mathbb{R})$ uses the L^p norm.

The alternating sum of d constructible functions, i.e. $\alpha : CF(\mathbb{R})^d \to CF(\mathbb{R})$, is continuous and so

$$
\alpha \circ \beta \circ \text{PHT}(M)(v, t) = \sum_{k=0}^{n} (-1)^k \beta_k(M_{v,t}) = \chi(M_{v,t}),
$$

which is $\text{ECT}(M)(v, t)$, is continuous as a function of v for any L^p norm so long as $p \neq \infty$. $\qquad \square$

When our constructible set M is a PL embedded geometric complex K, then we can say more about the continuity and image of the Euler Characteristic Transform, as the next remark indicates.

Remark 4.14. When M is a PL embedded simplicial complex K, we will show in Proposition 5.18 that, for any direction v, the Euler curve is right continuous. More precisely, the function space that ECT maps to consists of finite sums of indicator functions on half open intervals of the form $[a, b)$ where $b = \infty$ is allowed. If we endow the space of indicator functions with an L^p norm for finite p, then we get that if two Euler curves f and g differ at a filtration parameter t, then there is some $\epsilon > 0$ so that $(f - g)|_{[t, t+\epsilon)} \neq 0$, which in turn implies that $d_p(f, g) > \epsilon$. This

implies that in this setting, the Euler Characteristic Transform lands in a Hausdorff subspace of constructible functions $CF(\mathbb{R})$ equipped with some L^p norm.

Finally, we remind the reader of the sheaf-theoretic and K_0-theoretic perspective on the Euler Characteristc Transform.

Remark 4.15 (Grothendieck Group Interpretation). Continuing Remark 4.7, the reader familiar with the Grothendieck group of constructible sheaves (see [17] for a clear and concise treatment) will note that Proposition 4.13 coheres with the statement that the image of the (derived) Persistent Homology Transform in K_0 is the Euler Characteristic Transform.

4.3. **Injectivity of the PHT and BCT.** We can combine Proposition 4.13 with Theorem 3.5 to obtain a generalization of an injectivity result proved in [20] for simplicial complexes in $\mathbb{R}^2$ and $\mathbb{R}^3$.

Theorem 4.16. *Let $CS(\mathbb{R}^d)$ be the set of constructible sets, i.e. compact definable subsets of $\mathbb{R}^d$. The Persistent Homology Transform* PHT $: CS(\mathbb{R}^d) \to C(S^{d-1}, Dgm^d)$ *and Betti Curve Transform* BCT $: CS(\mathbb{R}^d) \to C(S^{d-1}, CF(\mathbb{R})^d)$ *are both injective.*

Proof. If two constructible sets M and M' have $\text{PHT}(M) = \text{PHT}(M')$ or $\text{BCT}(M) = \text{BCT}(M')$, then they have $\text{ECT}(M) = \text{ECT}(M')$, which by Theorem 3.5 implies that $M = M'$. $\qquad\qquad\square$

Remark 4.17 (Co-Discovery). In the final stages of preparing the first version of this article, a pre-print [12] by Rob Ghrist, Rachel Levanger, and Huy Mai appeared independently proving Theorem 4.16. The authors of that paper and this paper want to make clear that these results were independently discovered.

5. Stratified space structure of the ECT and PHT

Recall that the content of Lemma 3.4 was two-fold: (1) that for a constructible set M we could partition the space of directions and filtration values $S^{d-1} \times \mathbb{R}$ into finitely many regions over which the topological type of the sublevel sets $M_{v,t}$ do not change, and (2) that we could use this partition to obtain a partition of the sphere of directions so that when two directions v and w are in the same connected component of this partition, the persistent homology induced by h_v and h_w are equivalent in a certain sense. In this section we introduce the language of stratification theory to give a slightly tighter description of these decompositions in order to make issues of dimensionality and differentiability more transparent.

In particular, by using stratification theory we can refine the partitions of Lemma 3.4 into connected manifolds of varying dimension called *strata*. This will allow us to make certain arguments in terms of top-dimensional strata on the sphere S^{d-1}. When we eventually specialize our discussion to those M that are PL embedded geometric complexes in $\mathbb{R}^d$, we prove that the relationship between the persistent topology for h_v and h_w, when v and w are in the same stratum, is essentially linear. These observations will be critical in concluding that only finitely many directions are needed to infer a shape.

We now define what we mean by a stratified space structure. We note that there are many notions of a stratification, perhaps the most famous being the one due to Whitney [27]. Our notion of a stratification and of a stratified map will be a

combination of Whitney's condition with definability requirements. We follow the presentation in [16] in order to use the results proved there.

Definition 5.1. Suppose S_α and S_β are a pair of C^1 submanifolds of $\mathbb{R}^n$ with $S_\alpha \subseteq \overline{S_\beta}$. We let $T_x S_\alpha$ refer to the tangent space to S_α at x. The pair S_α and S_β satisfy *Whitney's condition* (*b*) *at* x if

> for every sequence $\{x_n\}$ in S_α converging to x and sequence $\{y_m\}$
> in S_β also converging to x where the tangent spaces $T_{y_k} S_\beta$ converge
> to τ and the secant lines $\ell_k := \langle x_k - y_k \rangle$ converge to ℓ then $\ell \subset \tau$.

This condition can be made to cohere with sets that are definable in some o-minimal structure.

Definition 5.2. A *definable C^p stratification of* $\mathbb{R}^n$ is a partition $\mathcal{S}$ of $\mathbb{R}^n$ into finitely many subsets, called strata, such that:

(1) Each stratum is a connected and definable C^p submanifold of $\mathbb{R}^n$.
(2) The Axiom of the Frontier holds, i.e. if $S_\alpha \cap \overline{S_\beta} \neq \varnothing$, then $S_\alpha \subseteq \overline{S_\beta}$.

We note that [15, Proposition 2.2.20] proves that the Axiom of the Frontier implies that the set of strata forms a partially ordered set (where $S_\alpha \leq S_\beta$ if and only if $S_\alpha \subseteq \overline{S_\beta}$) and that the frontier of any stratum, i.e. $\overline{S_\gamma} \setminus S_\gamma$, is a union of strata of strictly lower dimension.

A *definable C^p Whitney stratification* is a definable C^p stratification $\mathcal{S}$ such that for all $S_\alpha \leq S_\beta$ in $\mathcal{S}$ the pair satisfies Whitney's condition (*b*) at all points $x \in S_\alpha$.

We will be making use of two theorems, which say that stratifications and stratified maps can be made to respect pre-existing definable subsets. To be precise, we require Definition 5.3.

Definition 5.3. We say that $\mathcal{S}$ is *compatible* with a class $\mathcal{A}$ of subsets of $\mathbb{R}^n$ if each $A \in \mathcal{A}$ is a finite union of some strata in $\mathcal{S}$.

Theorem 5.4 (cf. Thm. 1.3 of [16])**.** *If $A_1, \cdots, A_k$ are definable sets in $\mathbb{R}^n$, then there exists a definable C^p Whitney stratification of $\mathbb{R}^n$ compatible with $\{A_1, \cdots, A_k\}$.*

Definition 5.5. Let $f : X \to Y$ be a definable map. A *C^p stratification of f* is a pair $(\mathcal{X}, \mathcal{Y})$, where $\mathcal{X}$ and $\mathcal{Y}$ are definable C^p Whitney stratifications of X and Y respectively. Moreover we require that for each $X_\alpha \in \mathcal{X}$ there be a $Y_\alpha \in \mathcal{Y}$, such that $f(X_\alpha) \subset Y_\alpha$ and $f|_{X_\alpha} : X_\alpha \to Y_\alpha$ is a C^p submersion.

Theorem 5.6 (cf. Thm. 2.2 of [16])**.** *Let $f : X \to Y$ be a continuous definable map. If $\mathcal{A}$ and $\mathcal{B}$ are finite collections of definable subsets of X and Y, respectively, then there exists a C^p stratification $(\mathcal{X}, \mathcal{Y})$ of f such that $\mathcal{X}$ is compatible with $\mathcal{A}$ and $\mathcal{Y}$ is compatible with $\mathcal{B}$.*

As observed in the proof of Lemma 3.4 and Remark 4.3, both the ECT and PHT can be viewed as auxiliary definable constructions associated to an o-minimal set M. Specifically, the ECT is gotten by the pushforward of the indicator function along the map

$$\pi : X_M \to S^{d-1} \times \mathbb{R}$$

and the PHT is associated to the Leray sheaves (the cohomology sheaves of the derived pushforward of the constant sheaf) of this map.

Since Theorem 5.6 assures us that this map is stratifiable, we get induced stratifications of the codomain of π as well as the sphere S^{d-1}. This is the content of the next lemma.

Lemma 5.7. *For a general o-minimal set $M \in \mathcal{O}_d$, the PHT or the ECT will induce a stratification of $S^{d-1} \times \mathbb{R}$ as well as a stratification of S^{d-1}. Moreover, in the induced stratification of the sphere S^{d-1}, a necessary condition for two directions v and w to be in the same stratum is that there is a stratum-preserving homeomorphism of the real line that induces an order preserving bijection between*

$$\bigcup_{k=0}^{d-1} \bigcup_{(I,j) \in PH_k(M,h_v)} \{b_j^k(I)\} \cup \{d_j^k(I)\} \rightarrow \bigcup_{k=0}^{d-1} \bigcup_{(I',j') \in PH_k(M,h_w)} \{b_{j'}^k(I')\} \cup \{d_{j'}^k(I')\}.$$

These two sets being the union of all the birth times and death times all the points (I, j) and (I', j') in each of the corresponding persistence diagrams in all degrees k, associated to filtering by h_v and h_w.

Proof. As already proved in Lemma 3.4

$$X_M := \{(x, v, t) \in M \times S^{d-1} \times \mathbb{R} \mid x \cdot v \leq t\}$$

is an element of $\mathcal{O}_{2d+1}$, and thus admits a definable C^p Whitney stratification compatible with X_M and X_M^c. The projection map $\pi : \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R} \rightarrow \mathbb{R}^d \times \mathbb{R}$ onto the last two factors is a continuous definable map. By Theorem 5.6 the map π admits a C^p stratification $(\mathcal{X}, \mathcal{Y})$ that is compatible with X_M and the trivializing partition $\{A_i\}$ of $S^{d-1} \times \mathbb{R}$ specified in the proof of Lemma 3.4.

As was done in Lemma 3.4, one can also project further from $S^{d-1} \times \mathbb{R}$ onto S^{d-1} and this map will be a continuous definable map, which in turn will admit a definable C^p stratification that is compatible with the trivializing partition $\{A_i\}$ of $S^{d-1} \times \mathbb{R}$ and $\{B_i\}$ of S^{d-1}.

Lemma 3.4 provides definable homeomorphisms $\pi^{-1}(B_i) \cap A_j \rightarrow B_i \times F_{ij}$, which can be pieced together to find a definable homeomorphism $\pi^{-1}(B_i) \rightarrow B_i \times \mathbb{R}$ where $\mathbb{R}$ is stratified in a way that is compatible with the the partition $F_{i1}, \ldots, F_{in}$ of $\mathbb{R}$. If v and w are in the same stratum of B_i, then restricting this homeomorphism to the pre-images of both will induce a zig-zag

$$\pi^{-1}(v) \hookrightarrow \pi^{-1}(B_i) \cong B_i \times \mathbb{R} \cong \pi^{-1}(B_i) \hookleftarrow \pi^{-1}(w)$$

that specializes to an order and stratum-preserving homeomorphism from $\pi^{-1}(v)$ to $\pi^{-1}(w)$. As a reminder, every topological change in the filtration induced by h_v is witnessed by a 0-cell in $\{F_{ij}\}$. Since each finite birth or death time in the persistence diagram $PH_k(M, h_v)$ corresponds to a homological change, and thus a topological change, the homeomorphism $\pi^{-1}(v) \cong \pi^{-1}(w)$ will match 0-cells of $\pi^{-1}(v)$, which include births/deaths of inessential features, with 0-cells in $\pi^{-1}(w)$. This completes the proof.   □

Lemma 5.7 simply guarantees the existence of a stratification of the sphere and gives criteria for determining when two directions are in the same stratum. To provide more explicit relationships between Euler curves or persistence diagrams associated to vectors in the same stratum, we specialize to sets $M \in \mathcal{O}_d$ that are (PL embedded) geometric simplicial complexes. We now recall some of the basic definitions.

Definition 5.8. A *geometric k-simplex* is the convex hull of $k + 1$ affinely independent points $v_0, v_1, \ldots v_k$ and is denoted $[v_0, v_1, \ldots, v_k]$. We call $[u_0, u_1, \ldots u_j]$ a *face* of $[v_0, v_1, \ldots v_k]$ if $\{u_0, u_1, \ldots u_j\} \subset \{v_0, v_1, \ldots v_k\}$.

Example 5.9. For example, consider three points $\{v_0, v_1, v_2\} \subset \mathbb{R}^d$ that determine a unique plane containing them. The 0-simplex $[v_0]$ is the vertex v_0. The 1-simplex $[v_0, v_1]$ is the edge between the vertices v_0 and v_1. Note that v_0 and v_1 are faces of $[v_0, v_1]$. The 2-simplex $[v_0, v_1, v_2]$ is the triangle bordered by the edges $[v_0, v_1]$, $[v_1, v_2]$ and $[v_0, v_2]$, which are also faces of $[v_0, v_1, v_2]$.

Remark 5.10 (Orientations). Traditionally, we view the order of vertices in a k-simplex as indicating an equivalence class of orientations, where if τ is a permutation then $[v_0, v_1, \ldots, v_k] = (-1)^{\mathrm{sgn}(\tau)}[v_{\tau(0)}, v_{\tau(1)}, \ldots, v_{\tau(k)}]$. However, for the purposes of this paper we can ignore orientation.

Definition 5.11. A *finite geometric simplicial complex* K is a finite set of geometric simplices such that

(1) Every face of a simplex in K is also in K;
(2) If two simplices σ_1, σ_2 are in K then their intersection is either empty or a face of both σ_1 and σ_2.

Remark 5.12 (Re-Triangulation). Strictly speaking we only care about the embedded image of the finite simplicial complex. We consider different triangulations as equivalent; our uniqueness results will always be up to re-triangulation.

For a pair of distinct points $x_i, x_j \in \mathbb{R}^d$ let $W(\{x_i, x_j\}) \subset \mathbb{R}^d$ be the hyperplane $\{v \in \mathbb{R}^d : x_i \cdot v = x_j \cdot v\}$ that is orthogonal to the vector $x_i - x_j$. This hyperplane divides the sphere S^{d-1} into two halves depending on whether $h_v(x_i) > h_v(x_j)$ or $h_v(x_j) > h_v(x_i)$. More generally, for a set of points $X = \{x_1, x_2, \ldots x_k\}$ we can define

$$W(X) := \bigcup_{i \neq j} W(\{x_i, x_j\})$$

as the union of the hyperplanes determined by each pair of distinct points.

Definition 5.13. Let $X = \{x_1, x_2, \ldots x_k\}$ be a finite set of points in $\mathbb{R}^d$ and let $W(X)$ be the union of hyperplanes determined by each pair of points. The hyperplane union $W(X)$ induces a stratification of S^{d-1}, which we call the *hyperplane division* of S^{d-1} by X. Define $\Sigma(X)$ as $S^{d-1} \setminus W(X)$.

Remark 5.14 (Stratified space structure). For a direction v, consider the simplicial complex $P(v)$ over the vertex set X where $\Delta \in P(v)$ if v is perpendicular to the affine plane spanned by the $x_i \in \Delta$. We can define a function $N : S^{d-1} \to \mathbb{Z}$ by

$$N(v) = \sum_{\text{locally maximal } \Delta \subset P(v)} \dim(\Delta).$$

If the points in X lie in generic position then the function N induces a stratified space structure on the sphere where the $d - k - 1$ dimensional strata correspond to the connected components of $N^{-1}(k)$. Notably the top dimensional strata (with dimension $d - 1$) are the connected components on $\Sigma(X)$. Notice that even when the hyperplanes are not in general position we can find, by virtue of Theorem 5.4, a definable C^p Whitney stratification of S^{d-1} that is compatible with the hyperplanes and all of their intersections.

Lemma 5.15. *Let $X = \{x_1, x_2, \ldots x_k\} \subset \mathbb{R}^d$. If $v_1, v_2 \in S^{d-1}$ are in the same stratum of $\Sigma(X)$, then the order of $\{v_1 \cdot x_i\}_{i=1}^k$ is the same as the order of the $\{v_2 \cdot x_i\}_{i=1}^k$.*

Proof. Observe that v_1 and v_2 lie in the same hemisphere of $W(\{x_i, x_j\})$ and so $h_{v_1}(x_i) > h_{v_1}(x_j)$ if and only if $h_{v_2}(x_i) > h_{v_2}(x_j)$.   □

Definition 5.16. For each vertex $x \in K$, the *star* of x, denote $\mathrm{St}(x)$ is the set of simplices containing x. Given a function $f : X \to \mathbb{R}$ we can define the *lower star* of x with respect to f, denoted $\mathrm{LwSt}(x, f)$, as the subset of simplices $\mathrm{St}(x)$ whose vertices have function values smaller than or equal to $f(x)$. Both stars and lower stars are generally not simplicial complexes as they are not closed under the face relation.

Remark 5.17 (Topological Interpretation of the Star). Although a geometric simplex is defined here in terms of convex hulls, which are closed when viewed as topological spaces, they are better viewed as topological spaces via their interiors. For example, if K is the simplicial complex consisting of the 1-simplex $[x, y]$ along with its two faces $[x]$ and $[y]$, then according to the above definition $\mathrm{St}(x) = \{[x], [x, y]\}$. If we identify each geometric simplex with its interior, then the star is the "open star," which in this case is the half-open interval $\mathrm{St}(x) = [x, y)$. Below we will give a combinatorial formula for computing the Euler characteristic of the star, which in this case is $\chi(\mathrm{St}(x)) = 1 - 1 = 0$. Note that if one is used to thinking in terms of Euler characteristic of the underlying space, then one must use compactly-supported Euler characteristic of the open star in order for these viewpoints to cohere.

For a finite geometric simplicial complex $K \subset \mathbb{R}^d$ with vertex set X, the height function $h_v : K \to \mathbb{R}$ is the piecewise linear extension of the restriction of h_v to the set of vertices X. When $v \notin W(X)$ then all the function values of h_v over X are unique. This implies that each simplex belongs to a unique lower star, namely to the vertex with the highest function value. We will make essential use of the following result.

Proposition 5.18 (Lem. 2.3 of [3] and §VI.3 of [13])**.** *For a piecewise linear function* $f : K \to \mathbb{R}$ *defined on a finite geometric simplicial complex* K *with vertex set* X*, we have that for every real number* $t \in \mathbb{R}$*, the sublevel set* $f^{-1}(-\infty, t]$ *is homotopic to* $\bigcup_{\{x \in X : v \cdot x \leq t\}} \mathrm{LwSt}(x, f)$.

Since both $\bigcup_{\{x \in X : v \cdot x \leq t\}} \mathrm{LwSt}(x, f)$ and $f^{-1}(-\infty, t]$ are also compact we conclude that they have the same definable Euler characteristic or Euler characteristic with compact support. We will use the notation

$$K^{(x,v)} = \left\{ \Delta \in \mathrm{St}(x) \subseteq K \mid x \cdot v = \max_{y \in \Delta} \{y \cdot v\} \right\}$$

to denote the lower star of x in the filtration by the height function in direction v.

Lemma 5.19. *Let* $K \subset \mathbb{R}^d$ *be a finite simplicial complex with vertex set* X*. Let* U *be a connected subset of* $\Sigma(X)$*. Then for fixed* $x \in X$*, the lower stars* $K^{(x,v)}$ *are all the same for all* $v \in U$*. We will sometimes denote the lower star by* $K^{(x,U)}$ *to highlight this consistency. Moreover, for all* $v \in U$ *we have the following formula for the Euler Characteristic Transform:*

$$\mathrm{ECT}(K, v) = \sum_{x \in X} 1_{\{[v \cdot x, \infty)\}} \chi(K^{(x,U)}).$$

Proof. We can see that the $K^{(x,v)}$ agree for all for all $v \in U$ by the observation that the vertices of X appear in the same order in each of the h_v. This $K^{(x,v)}$ is the subset of cells that are added at the height value $h_v(x)$.

Note that the change in the Euler characteristic of the sublevel sets of h_v as the height value passes $h_v(x)$ is

$$\chi(K^{(x,U)}) = \sum_{\Delta \in K^{(x,U)}} (-1)^{\dim \Delta}.$$

Here we use the notation $\chi(K^{(x,U)})$ for the (compactly-supported) Euler characteristic of this set of simplices. Consequently, we can write the Euler characteristic curve for K in direction $v \in U$ as

$$\mathrm{ECT}(K, v) = \sum_{x \in X} 1_{\{[v \cdot x, \infty)\}} \chi(K^{(x,U)}).$$

Note that many of the elements in this sum are zero. □

Example 5.20. As a simple example, let K be an embedded geometric 2-simplex $[x, y, z]$ along with all of its faces and suppose v is a direction in which, filtering by projection to v, the vertices appear in the order x, then y, then z. Note that $K^{(x,v)} = [x]$, which has Euler characteristic 1; $K^{(y,v)} = [x, y] \cup [y]$, which has Euler characteristic $(-1) + (-1)^0 = 0$; and $K^{(z,v)} = [x, y, z] \cup [y, z] \cup [x, z] \cup [z]$, which has Euler characteristic $(-1)^2 + (-1) + (-1) + (-1)^0 = 0$.

We now show that when two vectors v and w lie in the same connected component of $\Sigma(X)$, then the Euler curve or the persistence diagrams for h_v can be used to determine the Euler curve or persistence diagrams for h_w. We thank an anonymous referee for suggesting the following improved statement and proof.

Proposition 5.21. *Let $K \subset \mathbb{R}^d$ be a finite simplicial complex with known vertex set X. If $v, w \in S^{d-1}$ lie in the same connected component of $\Sigma(X)$, then given the Euler curve for v we can deduce the Euler curve for w. Similarly, given the kth persistence diagram associated to filtering by h_v we can deduce the k-th persistence diagram for filtering by h_w.*

Proof. Consider the height function $h_v : K \to \mathbb{R}$ on our PL embedded simplicial complex $K \subset \mathbb{R}^d$. Recall that the vertex set of K is X. Associated to the function h_v is another (typically discontinuous) function $k_v : K \to \mathbb{R}$ defined by setting $k_v(x) = h_v(\hat{x})$ where $\hat{x}$ is the unique vertex whose lower star contains the simplex in whose interior x belongs. Notice that this has the effect of increasing the value on the interior of a simplex to the maximum value obtained on any of its vertices. Consequently, for all $x \in K$ we have that $h_v(x) \leq k_v(x)$. This implies that for every $t \in \mathbb{R}$ we have the containment $k_v^{-1}(-\infty, t] \subseteq h_v^{-1}(-\infty, t]$. Notice that the sublevel set $k_v^{-1}(-\infty, t] = \bigcup_{\{x \in X : v \cdot x \leq t\}} \mathrm{LwSt}(x, h_v)$. By virtue of Proposition 5.18, we have that this inclusion of sublevel sets is a homotopy equivalence for every t, so h_v and k_v have identical persistence diagrams in all dimensions.

The virtue of defining the auxiliary function k_v is that it easy to compare with k_w when v and w belong to the same component of $\Sigma(X)$. This is because the level sets of k_v, k_w are empty except at the values of the inner product $x \cdot v$ and $x \cdot w$. At those values, the level sets are precisely the lower stars. As such, let $\phi : \mathbb{R} \to \mathbb{R}$ be any order-preserving homeomorphism that maps the values $x \cdot v$ to the values $x \cdot w$ for all $x \in X$. Such a ϕ exists because the order of the $\{x \cdot v\}$ and $\{x \cdot w\}$ for $x \in X$ are the same.

The mapping ϕ provides a bijection between these sets of values, with the property that the corresponding levelsets (i.e. lower stars) are equal, by virtue

of Lemma 5.19. This implies that $k_w = \phi \circ k_v$. It immediately follows that $\mathrm{ECT}(K, v) = \mathrm{ECT}(K, w) \circ \phi^{-1}$. Similarly, $\mathrm{PHT}(K, w)$ can be obtained from $\mathrm{PHT}(K, v)$ by transforming the persistence diagram plane using the function $(\phi, \phi) : \mathbb{R}^{2+} \to \mathbb{R}^{2+}$. This is because the birth time b_i and death time d_i of a feature in $\mathrm{PHT}(K, v)$ are transformed to a birth time $\phi(b_i)$ and death time $\phi(d_i)$ for a feature in $\mathrm{PHT}(K, w)$. $\square$

6. Uniqueness of the distributions of Euler curves up to $O(d)$ actions

In order to guarantee that two "close" shapes have close transforms—using either the ECT or PHT—one must first align or register the shapes' orientations in space. For example, if one wanted to compare simplicial versions of a lion and a tiger, we would need to first embed them in such a way that they are facing the same direction. To rephrase this, for centered shapes we would like to make them as close as possible by using actions of the special orthogonal group $SO(d)$. If we wished to optimize their alignment by also allowing reflections, we would like to make them as close as possible by using actions of the orthogonal group $O(d)$. In general, aligning or registering shapes is a challenging problem [4].

In this section we show how studying distributions on the space on Euler curves—or persistence diagrams, since homology determines Euler characteristic—can bypass this process. We can construct distributions of topological summaries through the pushforward under the topological transform of the uniform measure over the sphere. If μ is the Lesbesgue measure over S^{d-1} then the pushforward measure $\mathrm{ECT}_*(\mu)$ is a measure over the space of Euler curves defined by $\mathrm{ECT}_*(\mu)(A) = \mu(\mathrm{ECT}^{-1}(A))$. These pushforward measures are naturally invariant of $O(d)$ actions as the Lebesgue measure on the sphere is $O(d)$ invariant; acting on a shape and acting on the space of directions using the same element of $O(d)$ produces the same Euler curve. The key development of this section, Theorem 6.7, is the proof that for "generic" shapes the pushforward of the Lebesgue measure to the space of Euler curves uniquely determines that shape up to an $O(d)$ action. Additionally, our proof of this theorem is constructive. If we know the actual transforms of two generic shapes and we recognize they produce the same distribution of Euler curves, then we construct an element of $O(d)$ that relates them. However, the deeper implication is that knowing the distribution of Euler curves for a generic shape is a sufficient statistic for shape comparison. Continuing the aforementioned example, we can compare an arbitrarily embedded tiger and lion without ever aligning them. The theoretical ideas stated in this section were translated into an algorithm that "aligns" shapes without requiring correspondences in [26].

We now specify what we mean by a "generic shape."

Definition 6.1. A geometric simplicial complex K in $\mathbb{R}^d$ with vertex set X is *generic* if

(1) the Euler curves for the height functions h_v are distinct for all $v \in S^{d-1}$, and
(2) the vertex set X is in general position.

Our first task is to bound from below the number of "critical values" on a generic geometric complex when filtered in a generic direction. Since critical points and critical values are usually understood in a differentiable setting, and the shapes are

dealing with are only piecewise differentiable, we make precise what we mean by this intuitive term.

Definition 6.2. Let $f : K \to \mathbb{R}$ be a continuous function and $\chi_f : \mathbb{R} \to \mathbb{Z}$ the Euler curve of f. We say that t is an *Euler regular value* of $f : K \to \mathbb{R}$ if there is some open interval $I \subseteq \mathbb{R}$ containing t where χ_f is constant. An *Euler critical value* is a real number which is not an Euler regular value.

A real number t is called a *homological regular value* if there is some open interval $I \subseteq \mathbb{R}$ containing t such that for all $a < b$ in I the induced maps on homology by inclusion $(i_* : H_*(f^{-1}(-\infty, a]) \to H_*(f^{-1}(-\infty, b]))$ are all isomorphisms. A *homological critical value* is any real number that is not a homological regular value. For the purposes of this paper we will just use the term *critical value* when there is no chance for confusion.

Remark 6.3. There is some historical variation of the definition of a homological critical value and we refer the reader to [14] for a comparison and contrast of two candidate definitions along with several interesting examples.

We now state a lower bound on the number of critical values of a generic complex when filtered in a direction in $\Sigma(X)$.

Lemma 6.4. *If K is a generic finite geometric simplicial complex embedded in $\mathbb{R}^d$ with vertex set X then for all $v \in \Sigma(X)$ the height function h_v has at least $d - 1$ Euler critical values, i.e. values for which the Euler characteristic of the sublevel set changes.*

Furthermore, for each connected component $U \subset \Sigma(X)$ there exists $d-1$ linearly independent vertices $\{x_1^U, x_2^U, \ldots, x_{d-1}^U\}$ which are Euler critical for $ECT(K, v)$ for all $v \in U$.

Proof. Let U be a connected component of $\Sigma(X)$. The formula in Lemma 5.19 implies the number of Euler critical values for h_v ($v \in U$) is the number of vertices x such that $\chi(K^{(x,U)}) \neq 0$. Suppose that there are k critical vertices $x_1, x_2, \ldots x_k$, so that for all $v \in U$ $ECT(K, v) = \sum_{i=1}^{k} 1_{\{[v \cdot x_i, \infty)\}} \chi(K^{(x_i, U)})$.

Define a function $g : U \to \mathbb{R}^k$, with $g(v) = (v \cdot x_1, v \cdot x_2, \ldots, v \cdot x_k)$ which is clearly continuous and is injective by our genericity assumptions. As U is an open subset of a S^{d-1} there is a open subset $A \subset \mathbb{R}^{d-1}$ and a homeomorphism $f : A \to U$. Their composition, $g \circ f : A \to \mathbb{R}^k$, is continuous and injective. By Brouwer's Invariance of Domain Theorem we know that A is homeomorphic to $g(f(A))$ and thus $\dim(g(f(A))) = d - 1$. Since $f(U) \subset \mathbb{R}^k$ we conclude that $k \geq d - 1$.

By restricting to the first $d - 1$ critical vertices if needed, we have found vertices $\{x_1^U, x_2^U, \ldots, x_{d-1}^U\}$ which are Euler critical for $ECT(K, v)$ for all $v \in U$. We know that these x_i^U are linearly independent by our genericity assumptions. $\square$

Before proceeding to our main result, we need two technical lemmas.

Lemma 6.5. *Let $\ell : \mathbb{R}^d \to \mathbb{R}^d$ be a linear map. If there is a non-empty open subset U of S^{d-1} such that $\|\ell(v)\| = 1$ for all $v \in U$, then ℓ is an orthogonal transformation.*

Proof. Without loss of generality we can assume that U is an open ball in S^{d-1}. We want to show that for $v, w \in U$ that $\langle \ell(v), \ell(w) \rangle = \langle v, w \rangle$. Since $v, w \in U$, with

U an open ball in S^{d-1}, we know that $(v + w)/\|v + w\| \in U$ and hence

$$\|\ell(v + w)\| = \left\|\|v + w\|\ell\left(\frac{v + w}{\|v + w\|}\right)\right\| = \|v + w\|.$$

We thus can compute the inner product of $\ell(v)$ and $\ell(w)$ by

$$\begin{aligned}
\langle\ell(v), \ell(w)\rangle &= \frac{1}{2}(\langle\ell(v + w), \ell(v + w)\rangle - \langle\ell(v), \ell(v)\rangle - \langle\ell(w), \ell(w)\rangle) \\
&= \frac{1}{2}\left(\|\ell(v + w)\|^2 - \|\ell(v)\|^2 - \|\ell(w)\|^2\right) \\
&= \frac{1}{2}\left(\|v + w\|^2 - \|v\|^2 - \|w\|^2\right) \\
&= \langle v, w\rangle.
\end{aligned}$$

We can generate all of $\mathbb{R}^d$ from linear combinations of points in U. The inner product is bi-linear and so ℓ preserves the inner product over all of $\mathbb{R}^d$. As ℓ preserves the inner product over $\mathbb{R}^d$, ℓ must be an orthogonal transformation.   □

Lemma 6.6. *Let $\mathcal{S}$ be a definable stratification of S^{d-1}. Let $\mathcal{S}^k$ be the set that indexes connected components of k-dimensional strata within $\mathcal{S}$.*

Let G be the graph with vertex set that indexes top-dimensional strata, i.e. $V(G) := \mathcal{S}^{d-1}$, and with edge set that indexes codimension-1 strata, i.e. $E(G) := \mathcal{S}^{d-2}$. We say that i and j are connected by edge k if $S_k \leq S_i$ and $S_k \leq S_j$, i.e. $S_k \subset \overline{S_i}$ and $S_k \subseteq \overline{S_j}$. With this construction, the graph G is connected.

Proof. Let X_{d-3} be the union of strata of dimension $(d-3)$ and below in a definable stratification of S^{d-1}. Let $Y := S^{d-1} \backslash X_{d-3}$ be the complement of this union, which carries a definable stratification by restriction of the stratification $\mathcal{S}$ of S^{d-1}. By Alexander Duality we have that $\tilde{H}_0(Y) \cong \tilde{H}^{d-2}(X_{d-3}) = 0$, which implies that Y is connected. This proves that the union of top-dimensional strata and codimension-1 strata is a connected subspace of S^{d-1}.

Now consider any partition of the vertex set $V(G)$ into subsets V_0 and V_1. Let Y_0 and Y_1 denote the union of the corresponding top-dimensional strata represented by vertices in V_0 and V_1, respectively. Since the strata partition Y we have that $Y = \overline{Y_0} \cup \overline{Y_1}$ is expressible as a union of two closed (in Y) sets. Moreover, we know from the Axiom of the Frontier that $\overline{Y_0} \cap \overline{Y_1}$ is a union of $(d-2)$-strata. If this intersection is empty, then there are no $(d-2)$-strata and we have that Y is the union of a single top-dimensional stratum, thus proving that either V_0 or V_1 must be empty. If the intersection $\overline{Y_0} \cap \overline{Y_1}$ is non-empty, then there must be some $(d-2)$-dimensional stratum in this intersection and thus some vertex in V_0 must be connected by an edge with some vertex in V_1. This completes the proof.   □

The following is the main result of this section and is our formal statement about the uniqueness of a distribution over diagrams of curves up to an action by $O(d)$.

Theorem 6.7. *Let K and L be generic geometric simplicial complexes in $\mathbb{R}^d$. Let μ be Lesbesgue measure on S^{d-1}. If $\mathrm{ECT}(K)_*(\mu) = \mathrm{ECT}(L)_*(\mu)$ (that is the pushforward of the measures are the same), then there is some $\phi \in O(d)$ such that $L = \phi(K)$, that is to say that L is some combination of rotations and reflections of K.*

Proof. Recall that $\mathrm{ECT}(K)$ and $\mathrm{ECT}(L)$ are continuous and injective maps to the space of Euler curves. The hypothesis $\mathrm{ECT}(K)_*(\mu) = \mathrm{ECT}(L)_*(\mu)$ implies that the support of the pushforward measures is the same. We now argue that equality of supports implies equality of images. Firstly, we note that the support of $\mathrm{ECT}(K)_*(\mu)$ (respectively $\mathrm{ECT}(L)_*(\mu)$) contains the image of $\mathrm{ECT}(K)$ (respectively $\mathrm{ECT}(L)$). To see this, consider an Euler curve in the image and consider an open ball around that curve. Continuity implies that the pre-image is open and thus the Lebesgue measure is positive. Secondly, we show that any Euler curve not in the image of $\mathrm{ECT}(K)$ is not in the support of $\mathrm{ECT}(K)_*(\mu)$. To see this, we note that Remark 4.14 implies that $\mathrm{ECT}(K) : S^{d-1} \to \mathrm{CF}(\mathbb{R})$ is a continuous map from a compact space to a Hausdorff space, so the image is relatively closed in the subspace of possible Euler curves. If an Euler curve is not in the image of $\mathrm{ECT}(K)$, then there is an open set containing it that is disjoint from the image, which has measure zero according to the pushforward measure. Hence the support is a subset of the image. The above argument shows that $\mathrm{ECT}(K)_*(\mu) = \mathrm{ECT}(L)_*(\mu)$ implies the images of $\mathrm{ECT}(K)$ and $\mathrm{ECT}(L)$ are the same. Our genericity assumptions imply that $\mathrm{ECT}(L)$ is injective and thus we can define a bijection

$$\phi = \mathrm{ECT}(L)^{-1} \circ \mathrm{ECT}(K) : S^{d-1} \to S^{d-1}.$$

The map ϕ is both definable and a homeomorphism.

To see ϕ is a homeomorphism we need to show that

$$\mathrm{ECT}(L)^{-1} : \mathrm{im}(\mathrm{ECT}(L)) \to S^{d-1}$$

is continuous which we will do via a contrapositive version of an $\epsilon - \delta$ argument. For $v \in S^{d-1}$ define the function $f_v : S^{d-1} \to [0, \infty)$ by $f_v(w) = d_p(\mathrm{ECT}(L)(v), \mathrm{ECT}(L)(w))$ is continuous. The set $S^{d-1} \backslash B(v, \epsilon)$ is compact and hence $f_v(S^{d-1} \backslash B(v, \epsilon)) \subset [0, \infty)$ is a compact interval. This interval cannot contain 0 because $\mathrm{ECT}(L)$ is injective. This implies that there is some $\delta > 0$ with $f_v(w) > \delta$ for all $w \in S^{d-1} \backslash B(v, \epsilon)$. Since $v \in S^{d-1}$ was arbitrary, this means that for all $v \in S^{d-1}$ and all $\epsilon > 0$ there exists a $\delta > 0$ such that $d_{S^{d-1}}(v, w) > \epsilon$ implies that $d_p(\mathrm{ECT}(L)(v), \mathrm{ECT}(L)(w)) > \delta$.

We will now show that ϕ is definable. Let

$$A = \{(v, t, \mathrm{ECT}(K)(v, t)) | (v, t) \in S^{d-1} \times \mathbb{R}\}$$

and

$$B = \{(w, s, \mathrm{ECT}(L)(w, s)) : (w, s) \in S^{d-1} \times \mathbb{R}\}.$$

Both A and B are definable since the graph of definable functions is definable. Here we use that $\mathrm{ECT}(K)$ an $\mathrm{ECT}(L)$ are constructible (hence definable) functions on $S^{d-1} \times \mathbb{R}$.

Let P be the set of pairs of directions (v, w) such that the Euler curve of K in the direction of v, written $\mathrm{ECT}(K)(v)$, is the same as the Euler curve of L in the direction of w, written $\mathrm{ECT}(L)(w)$. Since the pullback of definable sets is definable [11, Lemma 11.1.15], we know that P is definable. P is then the graph of a definable map $\phi : S^{d-1} \to S^{d-1}$.

Let X and Y denote the vertex sets of K and L respectively. Let $W(X)$ and $W(Y)$ be the hyperplane partitions of X and Y, respectively, as defined in the previous section. By construction $W(X)$ is the union of $\binom{|X|}{2}$ hyperspheres, i.e. the intersections of the $d-1$ planes $W(\{x_i, x_j\})$ with S^{d-1}. We also know that $\phi^{-1}(W(Y))$ is homeomorphic to a union of $\binom{|Y|}{2}$ hyperspheres. Furthermore ϕ^{-1} is definable

and thus $W(X) \cap \phi^{-1}(W(Y))$ is definable. There exists a stratification of S^{d-1} into finitely many strata such that $W(X) \cap \phi^{-1}(W(Y))$ is contained in the lower dimensional strata. We will be looking at the restriction of ϕ to these top dimensional strata and showing that these restrictions are each the restriction of elements of $O(d)$, that is restrictions of an orthogonal map to their respective strata as the domain. We will then compare $d-1$ strata separated by a $d-2$ strata and show that their corresponding restrictions of ϕ agree as elements of $O(d)$.

Let $U \subset S^{d-1}$ be a connected open set with $U \cap W(X) = \emptyset$ and $\phi(U) \cap W(Y) = \emptyset$, equivalently $U \cap \phi^{-1}(W(Y)) = \emptyset$. This implies that U is a connected subset of $\Sigma(X)$ and $\phi(U)$ is a connected subset of $\Sigma(Y)$ which means we can apply Lemma 5.19 for both K and L separately.

Let X_U and $Y_{\phi(U)}$ respectively denote tffhe vertices of K or L such that $\chi(K^{(x,U)}) \neq 0$ or $\chi(L^{(x,\phi(U))}) \neq 0$. From Lemma 5.19 we know that

$$\mathrm{ECT}(K, v) = \sum_{x \in X_U} 1_{\{[v \cdot x, \infty)\}} \chi(K^{(x,U)})$$

and also that

$$\mathrm{ECT}(L, \phi(v)) = \sum_{y \in Y_{\phi(U)}} 1_{\{[\phi(v) \cdot y, \infty)\}} \chi(L^{(x,\phi(U))}).$$

Recall that the order of the values $\{v \cdot x\}$ over $x \in X_U$ is the same for all $v \in U$ and that this determines a total ordering over X_U. The same is true for the values $\{\phi(v) \cdot y\}$ over $y \in Y_{\phi(U)}$ again determining a total ordering over $Y_{\phi(U)}$. Since $\mathrm{ECT}(K_1, v) = \mathrm{ECT}(K_2, \phi(v))$ for all $v \in U$ we know that $|X_U| = |Y_{\phi(U)}|$ and furthermore we can consistently index the elements of X_U and $Y_{\phi(U)}$ such that

$$x_i \cdot v = y_i \cdot \phi(v)$$

for all $i = 1, \ldots, |X_U|$ and for all $v \in U$.

For clarity of exposition we shall first consider the case where $|Y_{\phi(U)}| \geq d$. From our genericity assumptions, $\{y_1, \ldots, y_d\}$ are linearly independent. This means that we can define the matrix

$$M_U = \begin{bmatrix} - & y_1 & - \\ - & y_2 & - \\ & \vdots & \\ - & y_d & - \end{bmatrix}^{-1} \begin{bmatrix} - & x_1 & - \\ - & x_2 & - \\ & \vdots & \\ - & x_d & - \end{bmatrix}$$

which has

$$M_U^T := \begin{bmatrix} | & | & & | \\ x_1 & x_2 & \ldots & x_d \\ | & | & & | \end{bmatrix} \begin{bmatrix} | & | & & | \\ y_1 & y_2 & \ldots & y_d \\ | & | & & | \end{bmatrix}^{-1}$$

and hence $M_U^T y_i = x_i$ for all i.

We will show that $\phi(v) = M_U v$ for all $v \in U$. For each $v \in U$ we have

$$y_i \cdot \phi(v) = x_i \cdot v = M_U^T y_i \cdot v = y_i \cdot M_U v.$$

Since this inner product holds for basis $\{y_i\}$ we know that $\phi(v) = M_U v$ for all $v \in U$.

Now suppose that we do not have $|Y_{\phi(U)}| < d$. We know that X_U and $Y_{\phi(U)}$ each must contain at least $d-1$ linearly independent points from Lemma 6.4. To

be able to use the arguments from the case where $|Y_{\phi(U)}| \geq d$ it will be sufficient to find vectors w_x and w_y such that w_x is linearly independent of X_U, w_y is linearly independent of $Y_{\phi(U)}$ and such that $\phi(v) \cdot w_y = v \cdot w_x$ for all $v \in U$. We can then apply the same reasoning merely substituting x_d (respectively y_d) with w_x (respectively w_y).

Let $\langle X_U \rangle$ and $\langle Y_{\phi(U)} \rangle$ denote the span of X_U and let w be perpendicular to $\langle X_U \rangle$. Note that $x_i \cdot w = 0$ for $i = 1, 2, \ldots, d-1$ by construction. We claim that $w \cdot v \neq 0$ for all $v \in U$. To prove this suppose that $v_0 \in U$ with $v_0 \cdot w = 0$. Since U is open there exists $a, b > 0$ such that both $av_0 + bw$ and $av_0 - bw$ are in U. However, $(av_0 + bw) \cdot x_i = av_0 \cdot x_i = (av_0 + bw) \cdot x_i$ for $i = 1, 2, \ldots, d-1$ and this would contradict the injectivity from domain U of the map $v \mapsto \{v \cdot x_1, v \cdot x_2, \ldots, v \cdot x_{d-1}\}$. As U is an open connected subset and $w \cdot v$ does not vanish on U we know either $v \cdot w > 0$ for all $v \in U$ or $v \cdot w < 0$ for all $v \in U$. By taking the negative if needed, let w_X denote the perpendicular to $\langle X_U \rangle$ such that $v \cdot w_X > 0$ for all $v \in U$.

Similarly we can define w_Y as the unit vector orthogonal to $\langle Y_{\phi(U)} \rangle$ (the span of $Y_{\phi(U)}$) such that $w_Y \cdot \phi(v) > 0$ for all $v \in U$.

Let $T : \langle X_U \rangle \to \langle Y_{\phi(U)} \rangle$ be the unique linear transformation which fixes the origin and such that $x_i \mapsto y_i$ for all i. Note that $v \cdot x = \phi(v) \cdot T(x)$ for each $x \in X$ and for all $v \in U$.

Let π_X and π_Y be the orthogonal projections of S^{d-1} onto $\langle X_U \rangle$ and $\langle Y_{\phi(U)} \rangle$ respectively. Projecting onto $\langle X_U \rangle$ and $\langle Y_{\phi(U)} \rangle$, we have

$$\pi_X(v) \cdot x = \pi_Y(\phi(v)) \cdot Tx.$$

The adjoint T^* of T (with respect to the Euclidean inner products) satisfies

$$\pi_X(v) \cdot x = T^*(\pi_Y(\phi(v))) \cdot x$$

for all $x \in X$ and all $v \in U$ which in turn implies

$$\pi_X(A) = T^* \pi_Y(\phi(A))$$

for all $A \subset U$.

We will use $\text{vol}_{S^{d-1}}$, vol_X and vol_Y to denote the Lebesgue measures over S^{d-1}, $\langle X_U \rangle$ and $\langle Y_{\phi(U)} \rangle$. Since T^* is linear there is a unique $\lambda_T > 0$ such that T^* scales volume by λ_T. For any $A \subset U$ we have

(6.1) $$\text{vol}_X(\pi_X(A)) = \lambda_T \text{vol}_Y(\pi_Y(\phi(A))).$$

Let $A_\epsilon \subset U$ be an open subset of U with diameter at most $\epsilon > 0$. We will be comparing the measures of A_ϵ, $\pi_X(A_\epsilon)$, $\phi(A_\epsilon)$ and $\pi_Y(\phi(A_\epsilon))$. If $v \in A_\epsilon$ then

$$\text{vol}_{S^{d-1}}(A_\epsilon) = \text{vol}_X \pi_X(A_\epsilon)(v \cdot w_x + O(\epsilon)).$$

We similarly have

$$\text{vol}_{S^{d-1}}(\phi(A_\epsilon)) = \text{vol}_Y \pi_Y(\phi(A_\epsilon))(\phi(v) \cdot w_y + O(\epsilon)).$$

Since ϕ is measure preserving we know that $\text{vol}_{S^{d-1}}(\phi(A_\epsilon)) = \text{vol}_{S^{d-1}}(A_\epsilon)$ and hence

(6.2) $$\text{vol}_X \pi_X(A_\epsilon)(v \cdot w_X + O(\epsilon)) = \text{vol}_Y \pi_Y(\phi(A_\epsilon))(\phi(v) \cdot w_Y + O(\epsilon)).$$

Together (6.1) and (6.2) imply that

$$\mathrm{vol}_X \, \pi_X(A_\epsilon)(v \cdot w_X + O(\epsilon)) = \mathrm{vol}_Y \, \pi_Y(\phi(A_\epsilon))(\phi(v) \cdot w_Y + O(\epsilon))$$
$$= \lambda_T \, \mathrm{vol}_X \, \pi_X(A_\epsilon)(\phi(v) \cdot w_Y + O(\epsilon))$$

and hence $(v \cdot w_X + O(\epsilon)) = \lambda_T(\phi(v) \cdot w_Y + O(\epsilon))$. By taking the limit as ϵ goes to zero we know $v \cdot w_X = \lambda_T \phi(v) \cdot w_Y$, with λ_T independent of $v \in U$. By setting $x_d = w_X$ and $y_d = \lambda_T w_Y$ we can state that $v \cdot x_d = \phi(v) \cdot y_d$ for all $v \in U$. This gives the dth linearly independent direction for the arguments above to imply that there is a linear map M_U such that ϕ and M_U agree on U.

Since M_U maps preserves distances for all $v \in U$ we can apply Lemma 6.5 to conclude that $M_U \in O(d)$.

Since ϕ is definable we have a stratification of S^{d-1} whose lower strata consist of $W(X_1) \cup \phi^{-1}(W(X_2))$. Set V be the set of the $(d-1)$ strata, E the $(d-2)$ strata, and f be the map which realises each element of V or E as its corresponding subset in S^{d-1}. Let G be the graph with vertex set V and edges E where the endpoints of edge e are the vertices v such that $f(e) \subset \overline{f(v)}$. We know from Lemma 6.6 that G is connected.

Consider adjacent $d-1$ dimensional strata U_1 and U_2 which are adjacent in G. This means that there is a $d-2$ strata within $\partial U_1 \cap \partial U_2$. As a $d-2$ dimensional subset of S^{d-1} we know that it must contains at least $d-1$ linearly independent points $\{v_1, v_2, \ldots v_{d-1}\} \subset \partial U_1 \cap \partial U_2$. Since ϕ is continuous we know that $M_{U_1}(v_i) = M_{U_2}(v_i)$ for all i. Choose $w \in S_{d-1}$ such that w is perpendicular to all the v_i (note that there are exactly two options). Since $\phi^{U_1}, \phi^{U_2} \in O(d)$ we can observe that $M_{U_1}(w)$ is perpendicular to $M_{U_1}(v_i)$ and $M_{U_2}(w)$ is perpendicular to $M_U U_2(v_i)$ for all i. As $M_U U_1(v_i) = \phi^{U_2}(v_i)$ for all i, this implies that $M_U U_2(w) = \pm M_U U_1(w)$.

Choose $v \in \partial U_1 \cap \partial U_2$ in the space of the $\{v_1, v_2, \ldots v_{d-1}\}$, and $a, b \neq 0$ such that $av - bw \in U_1$ and $av + bw \in U_2$. If $M_{U_2}(w) = -M_{U_1}(w)$ then we have

$$\phi(av - bw) = M_{U_1}(av - bw)$$
$$= aM_{U_1}(v) - bM_{U_1}(w)$$
$$= aM_{U_2}(v) + bM_{U_2}(w)$$
$$= M_{U_2}(av + bw)$$
$$= \phi(av + bw)$$

which contradicts the injectivity of ϕ. This implies that $M_{U_2}(w) = M_{U_1}(w)$ and thus $M_{U_2} = M_{U_1}$.

Since G is connected and $M_{U_2} = M_{U_1}$ for U_1 and U_2 adjacent vertices in G, we conclude that M_U must be the same for all U, which we denote M, and that $M \in O(d)$. We have two continuous maps $\phi, M|_{S^{d-1}} : S^{d-1} \to S^{d-1}$ that agree on an open dense subset (the union of the top strata) and hence must be the same over the entire domain. This implies that $\phi \in O(d)$ (by which we mean ϕ is the restriction of some $M \in O(d)$ to the unit sphere).

Now recall that by construction $\phi = \mathrm{ECT}(K_2)^{-1} \circ \mathrm{ECT}(K_1)$, so

$$\mathrm{ECT}(K_2)(\phi(v)) = \mathrm{ECT}(K_1)(v) \qquad \forall v \in S^{d-1}.$$

Since $\phi \in O(d)$ we can apply it to K_1 itself viewed as a subset of $\mathbb{R}^d$. Note that in this case we have the obvious formula

$$\mathrm{ECT}(K_1)(v) = \mathrm{ECT}(\phi(K_1))(\phi(v)) \qquad \forall v \in S^{d-1}.$$

Since both ECT and ϕ are injective we then have our desired implication.

$$\text{ECT}(K_2)(\phi(v)) = \text{ECT}(\phi(K_1))(\phi(v)) \quad \forall v \in S^{d-1} \qquad \Rightarrow \qquad K_2 = \phi(K_1).$$

<div style="text-align: right;">□</div>

7. The sufficiency of finitely many directions

In this section we reach the main result of our paper, which provides the first finite bound (to our knowledge) on the number of "inquiries" required to determine a shape belonging to a particular uncountable set of shapes, which we call $\mathcal{K}(d, \delta, k_\delta)$. These "inquiries" do not yield yes or no responses, but rather an Euler curve or persistent diagram.

The set of shapes belonging to $\mathcal{K}(d, \delta, k_\delta)$ can be described at a high level as follows:

(1) A "shape" is an embedded geometric simplicial complex $K \subset \mathbb{R}^d$. Two different embeddings of the same underlying combinatorial complex K are regarded as different shapes.

(2) Each embedded complex K is not "too flat" near any vertex. This guarantees that there is some set of directions with positive measure in which a non-trivial change of Euler characteristic (or homology) at that vertex can be observed. This measure is bounded below by a parameter δ, see Definition 7.3.

(3) No complex K has too many critical values when viewed in any given δ-ball of directions. Of course, for a fixed finite complex the number of critical values gotten by varying the direction is certainly bounded above, but we assume a uniform upper bound k_δ for our entire class of shapes $\mathcal{K}(d, \delta, k_\delta)$.

The proof of Theorem 7.14 proceeds in two steps. First, we show that the vertex locations of any member $K \in \mathcal{K}(d, \delta, k_\delta)$ can be determined simply by measuring changes in Euler characteristic when viewed along a fixed set of finitely many directions V. Once these vertices are located, the associated hyperplane division of the sphere described in Section 5 is then determined. Since we can provide a uniform bound on the number of vertices of any element of $\mathcal{K}(d, \delta, k_\delta)$, we then have a bound on the total possible number of top-dimensional strata of the sphere determined by hyperplane division. The second step is to then sample a direction from each individual top-dimensional stratum. Since Proposition 5.21 guarantees that we can interpolate the Euler Characteristic Transform over all of the top-dimensional strata, continuity guarantees that we can determine the entire transform using only these sampled directions. Finally, since Theorem 3.5 implies that $\text{ECT}(K)$ uniquely determines K, we then obtain the fact that any shape in $\mathcal{K}(d, \delta, k_\delta)$ is determined by finitely many draws from the sphere.

The mathematical ideas underlying Theorem 3.5 were central in the development of an algorithm that takes as input two classes of shapes and highlights the physical features that best describe the variation between them without requiring correspondences [26] and extended to characterize physical differences between classes of proteins in [25]. Again, we would like to make clear that the theory developed in this section is relevant to practical imaging applications.

As one might imagine, the above argument rests on many intermediary technical propositions and lemmas. The first lemma we introduce is an application of the

generalized pigeonhole principle, which will be used to pin down the locations of a vertex set of a fixed embedded simplicial complex.

Lemma 7.1. *Let $X = \{x_1, \ldots, x_l\} \subset \mathbb{R}^d$ be a finite set of points and let $D = \{v_1, \ldots, v_n\} \subset S^{d-1}$ be a set of directions in general position. We let $H(v_i, x_j)$ denote the hyperplane with normal vector v_i that passes through point x_j. Suppose n of the hyperplanes in $\mathcal{H} = \{H(v, x) \mid v \in D, x \in X\}$ intersect at $y \in \mathbb{R}^d$. If $n > (d-1)l$, then $y \in X$.*

Proof. The hypothesis implies that the set of hyperplanes $\mathcal{H}$ has at least n elements, because we've assumed that n of them intersect at some point $y \in \mathbb{R}^d$. Among these n hyperplanes consider the assignment of the plane $H(k)$ to some $x \in X$ with $x \in H(k)$. This defines a map from an n element set to the l element set X. Since $n \geq (d-1)l + 1$, the generalized pigeonhole principle implies at least one element of X is mapped to by d different hyperplanes, i.e. at least one element $\tilde{x} \in X$ is contained in d of the n hyperplanes containing y. Denote the normal vectors of these d hyperplanes by $v_{i_1}, \ldots, v_{i_d}$. Note that since $\tilde{x}$ and y are contained in these d hyperplanes we have the following system of equations

$$v_{i_k} \cdot \tilde{x} = v_{i_k} \cdot y \qquad \text{for} \qquad k = 1, \ldots, d.$$

Now we invoke the general position hypothesis, which says that the $\{v_{i_k}\}$ are linearly independent, and we deduce that $\tilde{x} = y$. $\qquad\square$

We now consider the two types of information that our two topological transforms can observe when looking in a direction $v \in S^{d-1}$. As has been the pattern for this paper, we first consider changes for the Euler Characteristic Transform and then deduce results about the Persistent Homology Transform.

Definition 7.2. Let $K \subset \mathbb{R}^d$ be a finite simplicial complex and $v \in S^{d-1}$. A vertex $x \in K$ is *Euler observable* in direction v if $\mathrm{ECT}(K)(v)$ changes value at $x \cdot v$. Given a subset of directions $A \subset S^{d-1}$, then a vertex $x \in X$ is Euler observable in A if x is observable for some $v \in A$. We will often say *observable* when the context is clear.

For the purposes of sampling the Euler Characteristic Transform, it is important to guarantee that each vertex is observable for some positive measure of directions. We make this precise by quantifying the above observability criterion via a real parameter δ.

Definition 7.3. Let $K \subset \mathbb{R}^d$ be a geometric simplicial complex and let $x \in K$ be a vertex of K. We say the vertex x is at least δ-*Euler observable* if there exists a ball of directions $B(v_x, \delta) \subset S^{d-1}$ such that the Euler characteristic of the sublevel sets of h_v changes at $v \cdot x$ for all $v \in B(v_x, \delta)$.

The idea of a δ-observable ball of directions at a vertex x is that the local geometry of the simplicial complex around x should not be "too flat" or similar to a hyperplane. This is indicated in the next example.

Example 7.4. Let K be a triangle in $\mathbb{R}^2$ with vertices $\{x, y, z\}$. The vertex y is observable in direction v exactly when $v \cdot y$ is smaller that both $v \cdot x$ and $v \cdot z$. If the angle at y is θ then y is δ-observable for all $\delta \leq \pi - \theta$.

This allows us to make precise our definition of $\mathcal{K}(d, \delta, k_\delta)$.

Definition 7.5. Let $\mathcal{K}(d, \delta, k_\delta)$ denote the set of all embedded simplicial complexes K in $\mathbb{R}^d$ with the following two properties:

(1) Every vertex $x \in K$ is at least δ-observable.
(2) For all $v \in S^{d-1}$, the number of $x \in X$ for which $h_w(x)$ is an Euler critical value of h_w for some $w \in B(v, \delta)$ is bounded by k_δ.

The second condition imposes a uniform upper bound on the number of critical points for all directions in δ neighborhood of direction v over all directions v on the sphere.

Before stating the true importance of the δ-observable condition, we remind the reader of a definition.

Definition 7.6. A δ-*net* in a metric space (M, d) is a subset N of points such that the union of δ-balls about these points $\cup_{p \in N} B(p, \delta)$ covers the entire space M.

Lemma 7.7. *Let $K \subset \mathbb{R}^d$ be a geometric simplicial complex. If a vertex $x \in K$ is δ-observable for some direction $v_x \in S^{d-1}$, then for any δ-net $V \subset S^{d-1}$ of directions there must be a direction $v \in V$ which can observe x.*

Proof. Since V is a δ-net there must be a $v \in V$ whose δ-ball includes v_x. Since x is δ-observable for v_x, it's observable for v as well. □

In order to extend the pigeonhole principle argument of Lemma 7.1 to find vertices that are δ-observable, we need a covering argument. This rests on the following technical lemma.

Lemma 7.8. *Let $r < \frac{\pi}{2}$ and fix a $w \in S^{d-1}$. We write $B_{S^{d-1}}(w, r) = \{v \in S^{d-1} : d(v, w) < r\}$ for the ball of directions about w of radius r. Recall that μ denotes the Lebesgue measure on S^{d-1} and ω_{d-1} is the Lesbesgue measure for the unit $(d-1)$-ball in $\mathbb{R}^{d-1}$. The following string of inequalities hold:*

$$\omega_{d-1} \sin(r)^{d-1} \leq \mu(B_{S^{d-1}}(w, r)) \leq \omega_{d-1} r^{d-1}.$$

Proof. The second inequality compares balls in the sphere to balls in Euclidean space. Since the sphere is positively curved the volume of a ball in the sphere is less than a ball of the same radius in Euclidean space of the same dimension.

The first inequality compares the volume of the $d-1$ dimensional ball in Euclidean space with radius r around w to the volume of the projection onto the tangent plane at w. The projection has radius $\sin(r)$. □

By knitting together the δ-observable condition with a uniform bound on the number of critical values in any ball of radius δ, we obtain an upper bound on the number of vertices.

Lemma 7.9. *The total number of vertices X for any $K \in \mathcal{K}(d, \delta, k_\delta)$ is bounded by*

$$|X| \leq \frac{dk_\delta}{\sin(\delta/2)^{d-1}}.$$

We know that $|X| = O(\frac{dk_\delta}{\delta^{d-1}})$.

Proof. The proof follows from an upper bound in the number of δ-balls needed to cover S^{d-1}. To do this we will use the relationship between packing and covering numbers. Let ω_{d-1} denote the volume of a ball in $\mathbb{R}^{d-1}$. From Lemma 7.8 we know that the area of a ball of radius $\delta/2$ within S^{d-1} is bounded below by $\omega_{d-1} \sin(\delta/2)^{d-1}$. We also know that the surface area of S^{d-1} is $d\omega_{d-1}$

This implies that the maximum number of balls of radius $\delta/2$ that can be packed into S^{d-1} is bounded above by

$$\frac{d\omega_{d-1}}{\omega_{d-1}\sin(\delta/2)^{d-1}} = \frac{d}{\sin(\delta/2)^{d-1}}.$$

Using the standard relationship that the covering number at radius δ is bounded above by the packing number at radius $\delta/2$ we know that there exists a covering of S^{d-1} by $\frac{d}{\sin(\delta/2)^{d-1}}$ balls of radius δ. In each of these balls in the cover at most k_δ different vertices can by observed. This implies that

$$|X| \le \frac{dk_\delta}{\sin(\delta/2)^{d-1}}.$$

<div align="right">□</div>

Definition 7.10. A (δ, C)-net is a subset of directions that meets every δ-ball in a set of cardinality C.

Proposition 7.11. *For $K \in \mathcal{K}(d, \delta, k_\delta)$ and the following constant which does not depend on K*

$$C = C(d, \delta, k_\delta) = (d-1)\,k_\delta + 1.$$

If V is a (δ, C)-net in S^{d-1} with the vectors in V in general position, then we can determine the location of the set of vertices of K using only the Euler curves generated from the directions in V.

Proof. The proof consists of first stating a constructive algorithm that will specify the location of all the vertices of K given Euler characteristic curves and then checking the correctness of the algorithm.

The algorithm declares a point x to be a vertex of K whenever there exists a set of different directions V_x, with the number of directions $|V_x| \ge C$ and $\mathrm{diam}(V_x) \le 2\delta$, such $v \cdot x$ is a critical value for height function h_v for all $v \in V_x$. To check the correctness of this algorithm we need to check that all the vertices are found and that no extra points are declared.

We first show that the algorithm will include every vertex in K. Consider a vertex x in K, since x is δ-Euler observable there is some ball $B(w, \delta)$ such that x is observed from every direction in $B(w, \delta)$. Let $V_x = V \cap B(w, \delta)$. By construction $\mathrm{diam}(V_x) \le \delta$ and $v \cdot x$ is a critical value for height function h_v for all $v \in V_x$. Since V is (δ, C)-net we know $|V_x| \ge C$. Thus our algorithm will declare x to be a vertex in K.

We now show that the algorithm will not declare any extra points. Suppose that $\tilde{x}$ is a point declared by our algorithm to be a vertex of K. This implies that there exists a vertex set $V_{\tilde{x}}$ and a ball of directions $B(w, \delta)$ such that both $V_{\tilde{x}} \subset B(w, \delta)$ and

$$|V_{\tilde{x}}| \ge (d-1)\,k_\delta + 1$$

and $\tilde{x} \cdot v$ is a critical value of the height function h_v for every $v \in V_{\tilde{x}}$. By assumption, we know that the number of observable vertices of K for any $v \in B(w, \delta)$ is at most k_δ. We then can apply our pigeonhole lemma, Lemma 7.1, with

$$n = |V_{\tilde{x}}| \qquad \text{and} \qquad l = k_\delta$$

to observe that $\tilde{x}$ must be one of these observable vertices of K and hence a vertex of K.   □

We now expand the above definitions and arguments to the homology setting.

Definition 7.12. A vertex x in a geometric simplicial complex $K \subset \mathbb{R}^d$ is *homology observable* if there is a direction $v \in S^{d-1}$ so that $\mathrm{PHT}(K)(v)$ contains an off diagonal point with either birth or death value $v \cdot x$. Moreover, we say that a vertex $x \in K$ is at least δ-*homology observable*, if it is homology observable for every direction in some ball $B(v_x, \delta) \subseteq S^{d-1}$ of directions.

The following statement is the analog to Proposition 7.11 for the Persistent Homology Transform, so we omit the proof as it is identical.

Corollary 7.13. *Suppose $K \subset \mathbb{R}^d$ is a finite simplicial complex such that every K is δ-homology observable with the same critical value conditions as in Proposition 7.11. Let V be a (δ, C)-net with the vectors in V all in general position. The location of the set of vertices of K can be determined using only the persistence diagrams generated from the directions in V.*

We now state the main theorem of the paper.

Theorem 7.14. *Any shape in $\mathcal{K}(d, \delta, k_\delta)$ can be determined using the* ECT *or the* PHT *using no more than*

$$\Delta(d, \delta, k_\delta) = ((d-1)\, k_\delta + 1)\left(1 + \frac{3}{\delta}\right)^d + O\left(\frac{d^{d+1} k_\delta^{2d}}{\delta^{2d(d-1)}}\right)$$

directions, where δ and k_δ are determined by the Euler or homological critical values, respectively.

Proof. We specify the proof for the ECT case, since the PHT case is identical.

The proof consists of showing that given a finite set of vectors $v_i \in V$ and the corresponding transform values $\mathrm{ECT}(K, v_i)$, we can recover $\mathrm{ECT}(K, v)$ for any direction $v \in S^{d-1}$. The proof proceeds as follows, after fixing some $K \in \mathcal{K}(d, \delta, k_\delta)$.

(1) Set V to be a general position (δ, C)-net in S^{d-1}. Here C is the constant from Proposition 7.11, $C = (d-1)\, k_\delta + 1$.

(2) Proposition 7.11 also states that the locations of the vertices X of the initial simplicial complex K can be recovered from the Euler characteristic curves generated from the directions in V.

(3) Now that we have the vertex set X of K, consider the hyperplane division $W(X)$ of S^{d-1} and its complement $\Sigma(X)$.

(4) For each connected component U_j of $\Sigma(X)$ pick a direction $u_j \in U_j$ and evaluate $\mathrm{ECT}(K, u_j)$. By Lemma 5.19 we can deduce the $\mathrm{ECT}(K, w)$ for any $w \in U_j$ from $\mathrm{ECT}(K, u_j)$.

(5) Given step (4), by continuity we know $\mathrm{ECT}(K, v)$ for $v \in W(X)$.

We now specify a bound on the number of directions in the set V, which rests on a bound for a (δ, C)-cover of the unit sphere S^{d-1}. We first state how to construct a (δ, C)-net. Consider the points V as the centers of a collection of C number δ'-nets with $\delta' = \frac{2}{3}\delta$ (although any $\delta' < \delta$ will suffice). Now perturb all the points in V slightly to put them into general position. The result of this procedure is a (δ, C)-net V with points in general position. By Lemma 5.2 in [24], the bound on the number of points needed for a δ'-net is $N = \left(1 + \frac{2}{\delta'}\right)^d$. This then proves the maximum number of directions required to specify the vertex set X for any element

$K \in \mathcal{K}(d, \delta, k_\delta)$ is

$$|V| \leq ((d-1)\,k_\delta + 1)\left(1 + \frac{3}{\delta}\right)^d.$$

Finally, we bound the number of directions required by Step (4) above. We note that $W(X)$ is the union of $n(X) = \binom{|X|}{2}$ hyperplanes. This divides the sphere up into at most

$$\sum_{j=0}^{d}\binom{n(X)}{j} \leq d\left(\frac{en(X)}{d}\right)^d,$$

note the above computation is standard computation used in both the uniform law of large numbers of sets [23] as well as the study of range spaces in discrete geometry [2]. Recalling Lemma 7.9 and using the fact that $n(X) = O(|X|^2)$, we see that the number of top dimensional strata is

$$O\left(d\left(\frac{n(X)}{d}\right)^d\right) = O\left(\frac{|X|^{2d}}{d^{d-1}}\right)$$

$$= O\left(\frac{d^{d+1}k_\delta^{2d}}{\delta^{2d(d-1)}}\right).$$

This completes the proof of the theorem. $\qquad\qquad\square$

8. Future directions

We conclude with a discussion of future directions that one might explore. We believe these questions would be intriguing to answer and important for geometry and statistics.

- In Sections 3 and 4 we used the class of compact, o-minimal sets to state our results. The primary purpose of this choice was to avoid inconsistencies between theory involving the two notions of Euler characteristic—the ordinary and compactly-supported/definable Euler characteristics. For general sets we cannot read off the definable Euler characteristic curve from the persistent homology. One direction of inquiry is considering a variant of persistent homology that naturally decategorifies to compactly-supported Euler characteristic. If we are restricted to filtrations only including locally compact sets then a natural choice would be a version of persistent Borel-Moore homology. It may be possible that we could use the full generality of Schapira's inversion theorem to prove the injectivity of the PHT for any locally compact o-minimal sets. One challenge with considering Borel-Moore persistent homology is that Borel-Moore homology is functorial with respect to proper continuous maps, but not to arbitrary continuous maps. When we work with non-compact definable sets, the inclusion maps of sublevel sets are continuous but are not proper continuous in general. This means we do not get the persistence module structure maps needed here. Another complication is that not all o-minimal sets are locally compact and little is defined, let alone understood, about Borel-Moore homology for sets that are not locally compact.
- In Section 5 we gave a complete description of the ECT and the PHT of finite simplicial complexes in terms of a stratified space decomposition of the sphere of directions, given by hyperplanes. We then showed that one can

interpolate the ECT by only knowing one curve from each stratum alongside knowing the vertex set. Both of these steps have natural analogs for shapes that aren't cut out by linear inequalities, but instead are determined by algebraic ones. How might one decompose the sphere and interpolate the ECT for shapes cut out using quadratic equations, for example?

- Continuing the above question to Section 6, one might ask whether more general constructible/semialgebraic sets have pushforward measures that are invariant to actions by $O(d)$.
- Finally, we note that Section 7 used piecewise linearity of the involved shapes in an essential way. There is no direct way to extend these results for arbitrary constructible sets, but there are a variety of related problems. One option is to explore whether it extends for some well-behaved family of shapes. For example, we could ask whether there is such a finiteness result for smooth algebraic manifolds with lower bounds on curvature. Another direction is a reconstruction up to some small error. One might be able to prove that given a set of directions, that we can construct a shape whose Euler curves agree with those for that sample direction such which must be close to the unknown shape with high probability.

Acknowledgments

References

[1] Pankaj K. Agarwal, Herbert Edelsbrunner, John Harer, and Yusu Wang, *Extreme elevation on a 2-manifold*, Discrete Comput. Geom. **36** (2006), no. 4, 553–572, DOI 10.1007/s00454-006-1265-8. MR2267546

[2] N. Alon, D. Haussler, and E. Welzl, *Partitioning and geometric embedding of range spaces of finite Vapnik-Chervonenkis dimension*, Proceedings of Symposium on Computational Geometry, pp. 331–340, 1987.

[3] Mladen Bestvina and Noel Brady, *Morse theory and finiteness properties of groups*, Invent. Math. **129** (1997), no. 3, 445–470, DOI 10.1007/s002220050168. MR1465330

[4] D. M. Boyer, J. Puente, J. T. Gladman, C. Glynn, S. Mukherjee, G. S. Yapuncich, and I. Daubechies, *A new fully automated approach for aligning and comparing Shapes*, The Ana. Rec. **258**, no. 1 `https://doi.org/10.1002/ar.23084`, 2015.

[5] P. Bubenik, J. Scott, and D. Stanley, "An algebraic Wasserstein distance for generalized persistence modules," arXiv preprint `https://arxiv.org/abs/1809.09654`.

[6] David Cohen-Steiner, Herbert Edelsbrunner, and John Harer, *Stability of persistence diagrams*, Discrete Comput. Geom. **37** (2007), no. 1, 103–120, DOI 10.1007/s00454-006-1276-5. MR2279866

[7] David Cohen-Steiner, Herbert Edelsbrunner, John Harer, and Yuriy Mileyko, *Lipschitz functions have L_p-stable persistence*, Found. Comput. Math. **10** (2010), no. 2, 127–139, DOI 10.1007/s10208-010-9060-6. MR2594441

[8] Lorin Crawford, Anthea Monod, Andrew X. Chen, Sayan Mukherjee, and Raúl Rabadán, *Predicting clinical outcomes in gioblastoma: An application of topological and functional data analysis*, J. Am. Stat. Assoc. **115** (2020), no. 531, 1139-1150, DOI 10.1080/01621459.2019.1671198.

[9] Justin Michael Curry, *Topological data analysis and cosheaves*, Jpn. J. Ind. Appl. Math. **32** (2015), no. 2, 333–371, DOI 10.1007/s13160-015-0173-9. MR3373695

[10] Justin Curry, Robert Ghrist, and Michael Robinson, *Euler calculus with applications to signals and sensing*, Advances in applied and computational topology, Proc. Sympos. Appl. Math., vol. 70, Amer. Math. Soc., Providence, RI, 2012, pp. 75–145, DOI 10.1090/psapm/070/589. MR2963602

[11] Justin Michael Curry, *Sheaves, cosheaves and applications*, ProQuest LLC, Ann Arbor, MI, 2014. Thesis (Ph.D.)–University of Pennsylvania. MR3259939

[12] Robert Ghrist, Rachel Levanger, and Huy Mai, *Persistent homology and Euler integral transforms*, J. Appl. Comput. Topol. **2** (2018), no. 1-2, 55–60, DOI 10.1007/s41468-018-0017-1. MR3873179

[13] Herbert Edelsbrunner and John L. Harer, *Computational topology*, American Mathematical Society, Providence, RI, 2010. An introduction, DOI 10.1090/mbk/069. MR2572029

[14] Dejan Govc, *On the definition of the homological critical value*, J. Homotopy Relat. Struct. **11** (2016), no. 1, 143–151, DOI 10.1007/s40062-015-0101-3. MR3462103

[15] Greg Friedman, *Singular intersection homology*, New Mathematical Monographs, vol. 33, Cambridge University Press, Cambridge, 2020, DOI 10.1017/9781316584446. MR4406774

[16] Ta Lê Loi, *Lecture 2: Stratifications in o-minimal structures*, The Japanese-Australian Workshop on Real and Complex Singularities—JARCS III, Proc. Centre Math. Appl. Austral. Nat. Univ., vol. 43, Austral. Nat. Univ., Canberra, 2010, pp. 31–39. MR2763234

[17] Pierre Schapira, *Operations on constructible functions*, J. Pure Appl. Algebra **72** (1991), no. 1, 83–93, DOI 10.1016/0022-4049(91)90131-K. MR1115569

[18] P. Schapira, *Tomography of constructible functions*, Applied algebra, algebraic algorithms and error-correcting codes (Paris, 1995), Lecture Notes in Comput. Sci., vol. 948, Springer, Berlin, 1995, pp. 427–435, DOI 10.1007/3-540-60114-7_33. MR1448182

[19] P. Skraba and K. Turner, *Wasserstein stability for persistence diagrams*, Preprint `arXiv:2006.16824`, 2020.

[20] Katharine Turner, Sayan Mukherjee, and Doug M. Boyer, *Persistent homology transform for modeling shapes and surfaces*, Inf. Inference **3** (2014), no. 4, 310–344, DOI 10.1093/imaiai/iau011. MR3311455

[21] Katharine Turner, *Medians of populations of persistence diagrams*, Homology Homotopy Appl. **22** (2020), no. 1, 255–282, DOI 10.4310/hha.2020.v22.n1.a15. MR4042792

[22] Lou van den Dries, *Tame topology and o-minimal structures*, London Mathematical Society Lecture Note Series, vol. 248, Cambridge University Press, Cambridge, 1998, DOI 10.1017/CBO9780511525919. MR1633348

[23] V. N. Vapnik and A. Ja. Červonenkis, *The uniform convergence of frequencies of the appearance of events to their probabilities* (Russian, with English summary), Teor. Verojatnost. i Primenen. **16** (1971), 264–279. MR0288823

[24] Roman Vershynin, *Introduction to the non-asymptotic analysis of random matrices*, Compressed sensing, Cambridge Univ. Press, Cambridge, 2012, pp. 210–268. MR2963170

[25] S. T. Wai, G. M. da Silva, H. Kirveslahti, E. Skeens, B. Feng, T. Sudijono, K. K. Yang, S. Mukherjee, B. Rubenstein, and L. Crawford, *A topological data analytic approach for discovering biophysical signatures in protein dynamics*, Preprint `arXiv:2021.07.28.454240`, 2021.

[26] Bruce Wang, Timothy Sudijono, Henry Kirveslahti, Tingran Gao, Douglas M. Boyer, Sayan Mukherjee, and Lorin Crawford, *A statistical pipeline for identifying physical features that differentiate classes of 3D shapes*, Ann. Appl. Stat. **15** (2021), no. 2, 638–661, DOI 10.1214/20-aoas1430. MR4298966

[27] Hassler Whitney, *Local properties of analytic varieties*, Differential and Combinatorial Topology (A Symposium in Honor of Marston Morse), Princeton Univ. Press, Princeton, N.J., 1965, pp. 205–244. MR0188486

[28] A. J. Wilkie, *Model completeness results for expansions of the ordered field of real numbers by restricted Pfaffian functions and the exponential function*, J. Amer. Math. Soc. **9** (1996), no. 4, 1051–1094, DOI 10.1090/S0894-0347-96-00216-0. MR1398816

Department of Mathematics and Statistics, University at Albany SUNY, Albany, New York 12222
  Email address: jmcurry@albany.edu

Departments of Statistical Science, Mathematics, Computer Science, Biostatistics & Bioinformatics, Duke University, Durham, North Carolina 27708; Center for Scalable Data Analytics and Artificial Intelligence at the Universität Leipzig, Germany; and the Max Planck Institute for Mathematics in the Natural Sciences, Leipzig, Germany
  Email address: sayan.mukherjee@mis.mpg.de

Mathematical Sciences Institute, Australian National University, Canberra, Australia
  Email address: katharine.turner@anu.edu.au