La Matematica
https://doi.org/10.1007/s44007-024-00143-9

ORIGINAL RESEARCH ARTICLE

Intrinsic Interleaving Distance for Merge Trees

Ellen Gasparovic[1] · **Elizabeth Munch**[2] · **Steve Oudot**[3] · **Katharine Turner**[4] · **Bei Wang**[5] · **Yusu Wang**[6]

Received: 13 July 2023 / Revised: 13 August 2024 / Accepted: 25 November 2024

Abstract

A merge tree is a type of graph-based topological summary that tracks the evolution of connected components in the sublevel sets of scalar functions. Merge trees enjoy widespread applications in data analysis and scientific visualization. In this paper, we consider the problem of comparing two merge trees via the notion of interleaving distance in the metric space setting. We investigate several theoretical properties of such a metric. In particular, we show that the interleaving distance is intrinsic on the space of labeled merge trees and provide an algorithm to construct metric 1-centers for collections of labeled merge trees. We further prove that the intrinsic property of the interleaving distance also holds for the space of unlabeled merge trees. Our results provide practical recipes for performing statistics on merge trees.

Keywords Merge tree · Topological data analysis · Interleaving distance

✉ Bei Wang
beiwang@sci.utah.edu

Ellen Gasparovic
gasparoe@union.edu

Elizabeth Munch
muncheli@msu.edu

Steve Oudot
steve.oudot@inria.fr

Katharine Turner
katharine.turner@anu.edu.au

Yusu Wang
yusuwang@ucsd.edu

1 Union College, Schenectady, NY, USA

2 Michigan State University, East Lansing, MI, USA

3 Inria Saclay, Palaiseau, France

4 Australian National University, Canberra, ACT, Australia

5 University of Utah, Salt Lake City, UT, USA

6 University of California, San Diego, La Jolla, CA, USA

Published online: 16 December 2024

1 Introduction

Many applications in science and engineering use scalar functions to describe and model their data. For example, atmospheric scientists compare simulated data from the Weather Research and Forecasting (WRF) Model with daily surface observations in weather forecasts, where both simulated and observed parameters (such as surface temperature, pressure, precipitation, and wind speed) can be modeled as scalar functions. We are interested in comparing scalar functions by comparing their topological summaries. There are several types of summaries constructed from topological methods, including vector-based summaries such as persistence diagrams [1] and barcodes [2], as well as graph-based summaries such as merge trees, contour trees [3], and Reeb graphs [4].

The merge tree (sometimes referred to as a barrier tree [5]) for a given topological space $\mathbb{X}$ equipped with a continuous scalar function is a combinatorial construction that tracks the evolution of sublevel sets. For a given function $f : \mathbb{X} \to \mathbb{R}$, the merge tree encodes the connected components of the sublevel sets $f^{-1}(-\infty, a]$ for $a \in \mathbb{R}$. This construction is closely related to that of the Reeb graph [4], which analogously encodes connected components of the level sets $f^{-1}(a)$. The contour tree [3] is a special type of Reeb graph for a simply connected domain. Both merge trees and Reeb graphs are related to the level set topology through critical points of the scalar functions which give rise to them [6]. Furthermore, the mapper graph [7], which has found considerable success in applications, can be viewed as an approximation of a Reeb graph [8–10]. These constructions are referred to as graph-based summaries as the output object of study is always a graph G equipped with an induced real-valued function $f : G \to \mathbb{R}$. They have appeared in many contexts and applications over the last few decades [11–13]. Similar concepts also appeared within probability theory as trees created through excursion sets of random functions, and these trees are shown to be related to random branching processes (e.g. [14, 15]).

1.1 Related Work

Considerable recent effort has gone into understanding how to perform proper statistics on graph-based summaries. For instance, how does one define the mean of a collection of these objects? The first step toward answering this question is to determine a metric for the comparison of two summaries. This has been extensively studied recently with the creation of a veritable zoo of metric options for Reeb graphs and merge trees [16–25]; see two recent surveys [26, 27] and Sect. 2.3 for a discussion of some of these metrics. In particular, Carriére and Oudot [24] have investigated whether some of these metrics are *intrinsic* in the more general case of Reeb graphs; i.e., that the distance between two (close enough) graphs can be realized by a geodesic.

In this paper, we continue the investigation into the intrinsic-ness of these metrics with the more narrow view of merge trees. The main distance we study is the interleaving distance. This metric was originally given in the context of persistence modules [28, 29] as a generalization of the bottleneck distance, and has been ported to merge trees [16, 30] and Reeb graphs [17, 31] via a category-theoretic viewpoint [32, 33].

When restricting ourselves from Reeb graphs to merge trees, we can actually work in an even more restrictive setting that has desirable theoretical properties, namely, labeled merge trees. In this case, we study a data triple: a merge tree T with its function $f : T \to \mathbb{R}$, and a labeling $\pi : \{1, \ldots, n\} \to V(T)$ of its vertices, which at a minimum encompasses the leaves of T. The interleaving distance for labeled merge trees has been investigated in [34], where it is shown that the metric can be naturally realized as the L_∞-distance for a particular matrix construction. This matrix construction has already been discovered in the context of dendrograms [35] and phylogenetic trees [36], where the objects of interest are closely related to merge trees. The phylogenetic tree literature, in particular, provides a wealth of other options for metrics [37–48]. There has also been interest in that community for creating summaries of collections of phylogenetic trees [49–51]. A parallel line of work has also focused on constructing barycenters of merge trees [52, 53].

These ideas are also closely related to those of ultrametrics, a strengthening of the triangle inequality for a metric into a requirement that $d(x, y) \leq \max\{d(x, z), d(z, y)\}$ (for all z). Independent of the phylogenetic tree work, there has been extensive interest in what is known as Gelfand's Problem from the ultrametric literature, that is, to describe all finite ultrametric spaces up to isometry using graph theory. The answer to this question is exactly a restriction of the labeled merge tree, although their literature never calls it such [54–58].

Furthermore, our work has close ties with the literature on consensus of classification [59]. In the language used therein, labeled merge trees belong to the class of valued classification trees, with the path-length distance being actually induced by the function values. Geodesic midpoints between two merge trees are called *medians* and defined as Fréchet means in the considered metric between trees. Finding medians is a special instance of the so-called consensus problem, and it is known to have an easily computable solution when the metric between merge trees is chosen to be the ℓ^∞-distance between their corresponding ultra matrices [60], as is the case in our work. By contrast, the problem is known to be NP-hard when other ℓ^p-distances are put on the ultra matrices, typically when $p = 1$ or 2 [61], and in such situations one must resort to approximate solutions—e.g., mean-squares approximations in the case $p = 2$ [62, 63].

1.2 Our Contributions

In this paper, we prove that the interleaving distance is intrinsic on both the space of labeled merge trees as well as on the space of unlabeled merge trees. Furthermore, in the labeled merge tree setting, we provide explicit procedures for constructing geodesics and metric 1-centers. Our results mark important progress toward the goal of performing statistics on graph-based topological summaries to be used for topological data analysis and visualization. For instance, using results in this paper, Yan et al. [64] computed geodesics of merge trees and their structural averages for ensemble analysis and uncertainty visualization, and Curry et al. [65] utilized an estimation of the interleaving distance between unlabeled merge trees in order to classify and compare point cloud data.

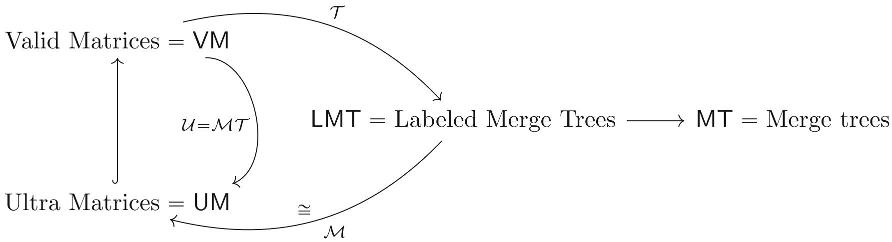


Fig. 1 A roadmap of key notations

In Sect. 2, we provide the necessary background on labeled merge trees and establish a correspondence between labeled merge trees and a particular class of matrices known as ultra matrices (Lemma 2.3). Then, in Sect. 3, we prove a stability result for the labeled interleaving distance d_I^L (Lemma 3.1), which we use both to show that d_I^L is *strictly intrinsic* on the space of labeled merge trees (Corollary 3.2) as well as to construct 1-centers for collections of labeled merge trees in Sect. 3.3.

Section 4 focuses on unlabeled merge trees and the interleaving distance. In particular, given two unlabeled merge trees, we show that the unlabeled interleaving distance between them is equal to the infimum over all finite labelings for the two trees of the labeled interleaving distance between them (Theorem 4.1). Section 4 concludes with the result that the interleaving distance is intrinsic on the space of unlabeled merge trees (Corollary 4.4). We end with a discussion of open problems and future work in Sect. 5.

2 Background

In this section, we give the basic definitions for our constructions of interest. We refer to Fig. 1 for an overview of notations. For the entirety of the section, we fix n, and denote $\{1, \ldots, n\}$ by $[n]$ and isomorphism by $\cong$. We note that some of the notions here appear in the literature under different names. For example, the concept of ultra matrix is the same as the one induced from the *ultra network* proposed in [66], both of which correspond to the distance matrix for a relaxed version of the ultrametric (see Definition 5). The concept of a merge tree is also the same as the *tree gram* in [66], both of which generalize the standard dendrogram. A dendrogram could be represented as an ultrametric, as shown by Jardine and Sibson [67], Hartigan [68], Carlsson and Mémoli [69].

2.1 Merge Trees and Labeled Merge Trees

First, we give the definition of a *merge tree* (which we shall also refer to as an *unlabeled merge tree* to contrast it with its labeled counterpart, defined subsequently) and related notions arising from the phylogenetic tree literature that we will make use of shortly.

In some cases—for instance, when a merge tree is constructed from sublevel sets of input data given by a topological space with a function $g : \mathbb{X} \to \mathbb{R}$—we prefer to think

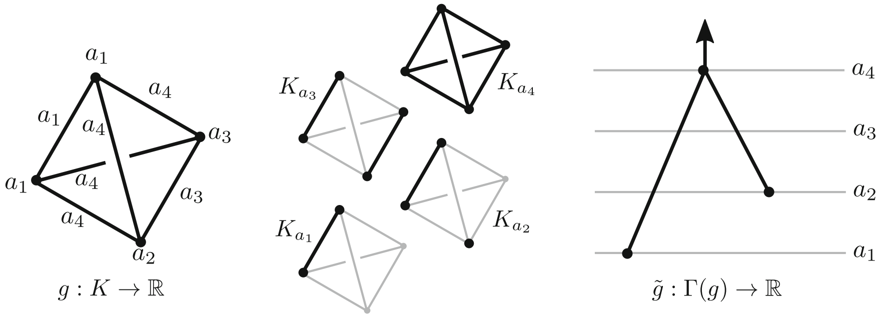


Fig. 2 A given function g defined on the complete graph K_4 is shown at left. In the middle are sublevelsets $K_{a_i} = g^{-1}(-\infty, a_i]$ for $a_1 < a_2 < a_3 < a_4$. At right is the resulting merge tree

of a merge tree as a continuous object which can be encoded with combinatorial data when necessary. Specifically, a merge tree is constructed from a function $g : \mathbb{X} \to \mathbb{R}$ by first building the epigraph, epi $f = \{(x, y) \in \mathbb{X} \times \mathbb{R} \mid y \geq f(x)\}$, with a function $\bar{f}$: epi $f \to \mathbb{R}$ given by $\bar{f}(x, y) = y$. Define an equivalence relation on epi f by setting $(x, y) \sim (x', y')$ iff the points are in the same connected component of the levelset $\bar{f}^{-1}(y) = \bar{f}^{-1}(y')$.

Definition 1 The merge tree of an input function $f : \mathbb{X} \to \mathbb{R}$ is the quotient space $\Gamma(f) = \text{epi } f / \sim$. The induced function $\tilde{f} : \Gamma(f) \to \mathbb{R}$ is given by $\tilde{f}([x, y]) = y$.

Note that the projection onto the first coordinate of the levelset $\bar{g}^{-1}(a)$ is exactly the sublevelset $g^{-1}[\infty, a)$, so we can equivalently compute this merge tree by studying the merging of connected components of the sublevelsets as a varies, hence the name.

We assume our input function g is nice enough to result in a a 1-dimensional stratified space equipped with a function, that is, there is a set of 0-cells inducing edges which are homeomorphic to intervals. This will happen, for instance, in the case where g is given by a Morse function on a manifold; or as we will use later, in the case where g is a function on a finite simplicial complex K where $g : K \to \mathbb{R}$ such that $\sigma \leq \tau$ implies $g(\sigma) \leq g(\tau)$. Consider the example of Fig. 2, where K is the complete graph on 4 vertices and we have $g : K \to \mathbb{R}$ given by the labels on the simplices, assuming $a_1 < a_2 < a_3 < a_4$. Then the resulting merge tree $\Gamma(g)$ is given on the right.

In this setting, we can then align the continuous notion of a merge tree with a more combinatorial version which will be given in Definition 2. A given collection of 0-cells for the stratification gives rise to a vertex set V, and then 1-cells are edges. We can then store the function in this setting by remembering $g(v)$ for the vertices. However, we next note a technicality of the two parallel viewpoints. In the stratified space setting, the set of 0-cells is not unique. replacing a merge tree edge $e = (u, v)$ with $f(u) < f(v)$ by a subdivision of that edge where the interior vertex w satisfies $f(u) < f(w) < f(v)$ does not change the inherent structure of the tree. We call such a modification a *monotone edge subdivision* and its inverse a *monotone edge simplification*. We consider two merge trees to be the same if one can be obtained from the other by a sequence of such subdivisions or simplifications.

Definition 2 A *merge tree* is a pair (T, f) of a finite rooted tree T with vertex set $V(T)$ and a function $f : V(T) \to \mathbb{R} \cup \{\infty\}$ such that adjacent vertices do not have equal function value, every non-root vertex has exactly one neighbor with higher function value, and the root (a degree one node) is the only point with the value ∞. Two merge trees are said to be equivalent if one can be obtained from the other of a series of monotone edge subdivisions or simplifications. The space of merge trees up to equivalence is denoted MT.

We choose to topologize MT using the topology induced by the interleaving distance, although we hold off on the full definition of this distance until Sect. 2.3. We commonly call the function f a *height function*, the non-root vertices with degree 1 are called *leaves*, and we let *depth*(u) denote the largest height difference between the vertex u in T and any node in the subtree rooted at u. All merge trees under consideration in this paper are assumed to be *finite*. When we wish to return to the stratified space setting, we denote this as $|T|$, the geometric realization of T where $f(x)$ is given by $f(v)$ when $x = v$ a vertex, and linear interpolation for points on the edges.

Furthermore, the merge tree structure induces a poset relation on the vertices of T. We say v is an *ancestor* of w and write $v \succ w$ if the unique path from v to w strictly decreases in f. This occurs if and only if w is in the subtree of v. We use $\mathrm{LCA}(v, w) \in T$ to mean the lowest common ancestor of v and w (or $\mathrm{LCA}_f(v, w)$ if the function needs to be emphasized), and $f(\mathrm{LCA}(v, w))$ for its function value. We have $\mathrm{LCA}(v, v) = v$. We abuse notation and write $\mathrm{LCA}(S)$ for the lowest common ancestor of any finite set $S \subset V(T)$.

Note that the merge tree as defined is closely related to the construction of a rooted, weighted tree. Indeed, there is a canonical weighting associated to any merge tree (T, f), namely, $\omega(u, v) = |f(u) - f(v)|$ for any two adjacent vertices u and v in the tree. However, because of the function setting, the merge tree requirements are stricter since, for instance, a merge tree (T, f) and its shift (i.e., translation) $(T, f + 100)$ are considered different as merge trees but induce the same weighting. The merge tree structure further provides a method for inducing a metric on the underlying tree vertices via the metric given by the length of the unique path between two points; i.e. $\delta_T(u, v) = \sum \omega(e)$ for the edges in the path from u to v.

We remind the reader that we use the terms *merge tree* and *unlabeled merge tree* interchangeably. In Sect. 3, we will be focusing on *labeled merge trees*, defined as follows.

Definition 3 A *labeled merge tree* is a triple (T, f, π) consisting of a merge tree (T, f) along with a map $\pi : [n] \to V(T)$ that is surjective on the set of leaves. When additional data are unnecessary or clear from context, we sometimes write T for (T, f, π). The space of labeled merge trees is denoted LMT.

Note that the topology on LMT comes from viewing it as a metric space with the labeled interleaving distance, as defined later in this section. Analogous to the unlabeled case, we consider two labeled merge trees to be the same if one can be obtained via edge contractions or insertions that respect the function values and existing labels.

When computing a merge tree from input data $g : \mathbb{X} \to \mathbb{R}$, note that labeled points on $\mathbb{X}$ give rise to labels on the resulting merge tree. For a point $x \in \mathbb{X}$, we

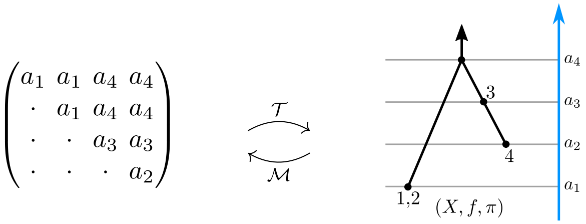


Fig. 3 An example of a labeled merge tree with two types of degenerate labels. As all matrices used are symmetric, we only show the upper triangular portion

have a point $(x, g(x)) \in \operatorname{epi} g$ and thus a point in the merge tree $[(x, g(x))] \in \Gamma(g)$. Given a labelling on $\mathbb{X}$, $\bar{\pi} : [n] \to \mathbb{X}$, we can push forward the labeling to define $\pi : [n] \to \Gamma(g)$ by $\pi[i] = [(\pi(i), g(\pi(i))]$.

Definition 3 is closely related to that of a *weighted, rooted X-tree* from the phylogenetic literature [70]. Specifically, given a set X, an X-tree is a pair (T, φ) where T is a tree and $\varphi : X \to V(T)$ is a map so that every vertex of degree at most 2 is in the image. The difference is that such weighted graphs do not keep track of function values, so that two different labeled merge trees that induce the same weighting might be considered to be the same X-tree. Thus, a labeled merge tree can be thought of as a weighted, labeled X-tree (where $X = [n]$) with $f(u)$ specified for a subset of vertices u that includes all leaves, and function values for the remaining vertices can be deduced from the weights on leaves.

As with X-trees, labels for our merge tree are allowed to go to vertices that are not leaves; we essentially think of these as degenerate labeled leaves. Furthermore, we do allow π to be non-injective, so a vertex can have multiple labels. See Fig. 3 for an example with labels on degenerate leaves and vertices with more than one label.

2.2 Relating Merge Trees and Matrices

In this section, we give the relationship between labeled merge trees and a particular class of matrices. Again, see Fig. 1 for an overview of notation.

We begin with the traditional notion of an *ultrametric* and our variant of it that relaxes one of the conditions, which will be closely related to our labeled merge trees.

Definition 4 An *ultrametric* is a function $d : X \times X \to \mathbb{R}$ such that for any $x, y, z \in X$,

- $d(x, y) \geq 0$ and is equal to 0 if and only if $x = y$,
- $d(x, y) = d(y, x)$, and
- $d(x, y) \leq \max\{d(x, z), d(z, y)\}$.

Definition 5 A *relaxed ultrametric* is a function $d : X \times X \to \mathbb{R}$ such that for any $x, y, z \in X$,

- $d(x, y) = d(y, x)$, and
- $d(x, y) \leq \max\{d(x, z), d(z, y)\}$.

It is well known that ultrametrics satisfy the isosceles triangle property. That is, for any triple x, y, z, at least two of $d(x, y)$, $d(y, z)$, and $d(x, z)$ must be equal.

Otherwise, assume without loss of generality that $d(x, y) < d(y, z) < d(x, z)$, and then $d(x, z) \not\leq \max\{d(x, y), d(y, z)\}$. Note that this further implies that the pair that are equal must be at least as big as the third value, since $d(x, y) = d(y, z) < d(x, z)$ still violates the ultrametric property. This means that relaxed ultrametrics still satisfy the isosceles triangle property.

When we have a set $X \cong [n]$, the information in a relaxed ultrametric can be stored as follows.

Definition 6 A symmetric matrix $M \in \mathbb{R}^{n\times n}$ is called *valid* if $M_{ii} \leq M_{ij}$ for all $1 \leq i, j \leq n$. A symmetric matrix M is called *ultra* if $M_{ij} \leq \max\{M_{ik}, M_{kj}\}$ for every $1 \leq k \leq n$. The spaces of valid and ultra matrices are denoted VM and UM, respectively, and note that $\mathsf{UM} \subseteq \mathsf{VM}$.

In particular, a relaxed ultrametric on $[n]$ is represented by an ultra matrix. As with merge trees, we will endow VM and UM with the topology induced by the relevant metric; in this case, the ℓ^∞ distance between matrices,

$$\|M - M'\|_\infty = \max_{i,j} |m_{i,j} - m'_{i,j}|.$$

Inspired by the cophenetic matrix construction of Cardona et al. [36] that is studied in relation to merge trees in [34], there is a natural way to associate a matrix to a labeled merge tree as follows.

Definition 7 The *induced matrix of a labeled merge tree* (T, f, π), denoted $\mathcal{M}(T, f, \pi) \in \mathbb{R}^{n\times n}$, is the matrix

$$\mathcal{M}(T, f, \pi)_{ij} = f(\mathrm{LCA}(\pi(i), \pi(j))).$$

See Fig. 3 for an example. We include the simple proof of the following result for completion.

Lemma 2.1 *The induced matrix of a labeled merge tree is an ultra matrix (Definition 6). That is,* $\mathcal{M}(T, f, \pi) \in \mathsf{UM}$ *for* $(T, f, \pi) \in \mathsf{LMT}$.

Proof Let $M = \mathcal{M}(T, f, \pi)$ for $(T, f, \pi) \in \mathsf{LMT}$. First, to check that it is a valid matrix, we see that M_{ii} is simply the function value $f(\pi(i))$. So, as $f(u) \leq f(LCA(u, v))$ by definition, we have

$$M_{ii} = f(\pi(i)) \leq f(\mathrm{LCA}(\pi(i), \pi(j))) = M_{ij}.$$

To check that M is an ultra matrix, let $u = \mathrm{LCA}(\pi(i), \pi(k))$, $v = \mathrm{LCA}(\pi(j), \pi(k))$, $w = \mathrm{LCA}(\pi(i), \pi(j), \pi(k))$. This means that $u \preceq w$ and $v \preceq w$. If u and v are not comparable, then there are two distinct paths from $\pi(k)$ to each of them, and thus we have a loop $\pi(j) \to u \to w \to v \to \pi(j)$, contradicting the tree property of T. If u and v are comparable, assume without loss of generality that $u \preceq v$; then v is a common ancestor for $\pi(i)$, $\pi(j)$, and $\pi(k)$, and thus $w \preceq v$. This implies $f(w) \leq f(v)$, and so for all k,

$$M_{ij} \leq f(w) \leq f(v) = \max\{f(u), f(v)\} = \max\{M_{ik}, M_{jk}\}.$$

□

We include an additional useful tool for working with the LCA which we will use throughout the paper.

Lemma 2.2 *Given a finite collection of points $S \subset |T|$, there exists a pair $\{x, y\} \subset S$ such that* $\mathrm{LCA}(S) = \mathrm{LCA}(x, y)$. *Moreover, for any $x \in S$, an element $y \in S$ exists such that* $\mathrm{LCA}(S) = \mathrm{LCA}(x, y)$.

Proof Because T is a tree, removing $w := \mathrm{LCA}(S)$ separates T into two disjoint subtrees (otherwise there exists a lower LCA). Thus, x and y can be chosen as elements in each of the respective tree. For the final statement of the lemma, if an x is given, then y can be chosen to be in the other subtree. □

A valid matrix may be viewed as representing a function f_M on a complete graph K of n vertices, with function value M_{ii} defined on vertex i and function value M_{ij} defined on edge (i, j). Note that because M is a valid matrix, any sublevel set of the resulting function $f : K \to \mathbb{R}$ satisfies the condition that every edge has equal or higher function value than either of its vertices. Given a valid matrix, one thus may obtain a labeled merge tree and subsequently an ultra matrix in the following way.

Definition 8 Let $M \in \mathbb{R}^{n\times n}$ be a valid matrix, K be a complete graph over n vertices, and $f_M : K \to \mathbb{R}$ be a function induced from M with $f_M(v_i) = M_{ii}$ and $f_M((v_i, v_j)) = M_{ij}$. The *labeled merge tree of a valid matrix M*, denoted as $\mathcal{T}(M)$, is the labeled merge tree of the complete graph K with the induced function f_M.

Specifically, given a valid matrix M, we can consider M to induce a function $f_M : K \to \mathbb{R}$ on the complete graph K on n vertices, as in the example of Fig. 2. Specifically, M is an $n \times n$ matrix and so we build the complete graph K on the vertex set $\{v_1, \ldots, v_n\}$. The function g is given by $g(v_i) = M_{i,i}$ for vertices, and $g(v_i, v_j) = M_{i,j}$ for edges. Because M is a valid matrix, this gives a well-defined map; in particular, $g(v_i) \leq g(v_i, v_j)$ for any $i \neq j$. The labeling is given by $\pi(i) = v_i$. We then compute the merge tree of g and push forward the labels; the result $\Gamma(g)$ is what we denote by $\mathcal{T}(M)$, the labeled merge tree of M.

This relationship gives a direct connection between labeled merge trees and ultra matrices; see the result below, which follows from Theorem 6 of [66] (although the proof there is written in the language of tree grams).

Lemma 2.3 (Theorem 6 in [66]) *$\mathcal{M}$ induces a bijection between labeled merge trees and ultra matrices.*

Putting this together, we can thus take an ultra matrix, turn it into a labeled merge tree and then back into an ultra matrix, i.e. $\mathcal{M}\mathcal{T}(M)$. This procedure corresponds to the maximal subdominant construction in [60]. Although $\mathcal{M}\mathcal{T}$ is the identity when restricted to ultra matrices, this is not the case when extending to only valid matrices. However, this construction does offer a method for turning a valid matrix into an ultra matrix.

Definition 9 The *ultra matrix of a valid matrix $M \in$ VM*, denoted $\mathcal{U}(M)$, is defined to be the induced matrix of $\mathcal{T}(M)$. That is, $\mathcal{U} = \mathcal{M}\mathcal{T}$.

2.3 Available Metrics

There are a number of metrics that may be defined on the space of (labeled) merge trees. Note that any metric defined on labeled merge trees can be extended to unlabeled merge trees by simply ignoring the labeling information, while likely turning the metric into a pseudometric. In this paper, we focus on interleaving distance d_I and labeled interleaving distance d_I^L. Other popular distances include the functional distortion distance d_{FD} [18] and the bottleneck distance d_B.
Interleaving Distance The interleaving distance is an idea arising from the generalization of the bottleneck distance for persistence diagrams to arbitrary persistence modules [28]. Generalizations abound [9, 32, 33], but the analog for merge trees was first given in [16]. We give a non-standard formulation here based on the concept of δ-good maps. In particular, in [30] a concept of δ-good map is given, and it is shown that one can then define interleaving distance based on this concept (see Theorem 7 of [30]). In what follows (Definition 10), we will use a slightly different formulation of the δ-good map from [30]. However, we show in Appendix A that these two concepts of δ-good maps are indeed equivalent.

Definition 10 Given two merge trees $(T, f), (T', f')$, a δ-*good* map $\alpha : (T, f) \to (T', f')$ is a continuous map on the metric trees such that the following properties hold:

(i) For any x in the geometric realization $|T|$, $f'(\alpha(x)) - f(x) = \delta$;
(ii) For any $w \in \mathrm{Im}(\alpha)$ with $x' := \mathrm{LCA}(\alpha^{-1}(w))$, $f(x') - f(u) \leq 2\delta$ for all $u \in \alpha^{-1}(w)$; and
(iii) For any $w \notin \mathrm{Im}(\alpha)$, $depth(w) \leq 2\delta$.

The *interleaving distance* is then defined to be

$$d_I((T, f), (T', f')) = \inf\{\delta \mid \exists\, \delta\text{-good } \alpha : (T, f) \to (T', f')\}.$$

One particularly useful property that we will use later is the following.

Lemma 2.4 *Let $\alpha : (T, f) \to (T', f')$ be a continuous map such that $f'(\alpha(x)) = f(x) + \delta$ for any $x \in |T|$. Assume $u \preceq v$. Then*

- *$\alpha(u) \preceq \alpha(v)$, and*
- *if w is the unique ancestor of $\alpha(u)$ with $f'(w) = f(v) + \delta$, then $w = \alpha(v)$.*

Proof Note that $u \preceq v$ implies that $f(u) \leq f(v)$ and further that the unique path γ from u to v in T is monotone increasing in f. Then the image of γ in T', $\alpha(\gamma)$, satisfies $f'(\alpha(\gamma(t))) = f(\gamma(t)) + \delta$ and thus is monotone increasing in f'. Thus, by definition, we have that $\alpha(u) \preceq \alpha(v)$. Further, the uniqueness of paths implies that if w is the unique ancestor with $f'(w) = f(v) + \delta$, then it must be the endpoint of γ, and so $w = \alpha(v)$. □

Labeled Interleaving Distance The following metric is closely related to one originally defined in [36] for comparing phylogenetic trees.

Definition 11 Given two labeled merge trees sharing the same set of n labels, the *labeled interleaving distance* is

$$d_I^L((T, f, \pi), (T', f', \pi')) = \|\mathcal{M}(T, f, \pi) - \mathcal{M}(T', f', \pi')\|_\infty.$$

The reason for calling such a distance an interleaving distance comes from [34] where it is shown that this metric arises as an interleaving distance on a particular category with a flow [33]. Note that because we need the labels in order to be able to have a well-defined matrix, this metric only works on labeled merge trees.

2.4 Intrinsic Metrics

Given a metric d on merge trees, we may define its intrinsic version as follows; see, e.g., [71].

Definition 12 Given two merge trees, let $\gamma : [0, 1] \to \mathrm{MT}$ be a continuous path in d such that $\gamma(0) = T$ and $\gamma(1) = T'$. The *length of* γ *induced by the distance* d is defined as

$$L_d(\gamma) = \sup_{n, \sum} \sum_{i=0}^{n-1} d(\gamma(t_i), \gamma(t_{i+1})),$$

where n ranges over $\mathbb{N}$ and $\sum$ ranges over all partitions $0 = t_0 \le t_1 \le \cdots \le t_n = 1$ of $[0, 1]$. The *intrinsic metric* $\hat{d}$ *induced by the distance* d is

$$\hat{d}(T, T') = \inf_{\gamma} L_d(\gamma).$$

Thus, the induced intrinsic metric on a metric space is the infimum of the lengths of all paths from one point to another. It is known that d is always less than or equal to $\hat{d}$.

A metric space is said to be a *length space* if the original metric d coincides with the intrinsic metric $\hat{d}$. Recall that a metric space is said to be a *geodesic space* if any two points in the space can be connected by a curve of length equal to the distance between the two points. In this case, the metric is said to be *strictly intrinsic*. Note that a geodesic space is necessarily a length space.

3 Geodesics and 1-Centers for Labeled Merge Trees

In this section, we prove an inequality involving the labeled interleaving distance and provide methods for constructing geodesics and 1-centers for collections of labeled merge trees.

3.1 More on the Labeled Interleaving Distance

The following result is not new. It follows from Theorem 2 of [66], which is slightly more general than the lemma below in the sense that the matrices M and M' are allowed to be non-valid as well. It is also a slight generalization of Lemma 15 of [69] (which is the following result restricted to the metric setting). It can also be deduced from Proposition 1 and Corollary 1 of [60] because the set of relaxed ultra matrices (which are the matrices corresponding to merge trees) is stable under translations along the diagonal. See Section 3 of [60] for the special case of ultra matrices (equivalently, labeled dendrograms), which adapts straightforwardly to our slightly more general setting. Nevertheless, here we provide simple and direct proofs, both for completeness and clarity.

Lemma 3.1 *For any pair of valid matrices $M, M' \in \mathsf{VM}$,*

$$d_I^L(\mathcal{T}(M), \mathcal{T}(M')) \leq \|M - M'\|_\infty.$$

Proof Let $\delta = \|M - M'\|_\infty$. Let $T = \mathcal{M}(M)$ and $T' = \mathcal{M}(M')$ be the associated merge trees, and $\widetilde{M} = \mathcal{U}(M)$ and $\widetilde{M}' = \mathcal{U}(M')$ the induced ultra matrices. Since, by definition, $d_I^L(\mathcal{T}(M), \mathcal{T}(M')) = \|\mathcal{U}(M) - \mathcal{U}(M')\|_\infty$, we will actually establish the inequality $\|\mathcal{U}(M) - \mathcal{U}(M')\|_\infty \leq \|M - M'\|_\infty$.

Consider any pair of (possibly equal) labels i and j with $1 \leq i \leq j \leq n$. We consider the vertices v_i and v_j in the complete graph K with $s, s' : K \to \mathbb{R}$ denoting the maps on K induced by M and M', respectively. Denote edge $e = v_i v_j$, so by definition $s(e) = M_{ij}$ and $s'(e) = M'_{ij}$. Because $\|M - M'\|_\infty \leq \delta$, we have that

$$s'(e) \leq s(e) + \delta \leq \widetilde{M_{ij}} + \delta.$$

Thus, v_i and v_j are in the same component of the $(\widetilde{M}_{ij} + \delta)$-sublevel set of s' and thus $\widetilde{M'}_{ij} \leq \widetilde{M}_{ij} + \delta$.

Symmetrically, for any $t < \widetilde{M}_{ij} - \delta$, v_i and v_j do not lie in the same connected component of the t-sublevel set of s'. Otherwise, by the same argument as above, v_i and v_j would belong to the same connected component of the $(t + \delta)$-sublevel set of T with $t + \delta < \widetilde{M}_{ij}$, a contradiction. Hence, $\widetilde{M}'_{ij} \geq \widetilde{M}_{ij} - \delta$. It follows that $|\widetilde{M}'_{ij} - \widetilde{M}_{ij}| \leq \delta$, and since this is true for all labels $1 \leq i \leq j \leq n$, the symmetric matrices $\widetilde{M}, \widetilde{M}'$ satisfy $\|\widetilde{M} - \widetilde{M}'\|_\infty \leq \delta$. Hence, $d_I^L(T, T') = \|\widetilde{M} - \widetilde{M}'\|_\infty \leq \delta$. ◻

3.2 Geodesics in LMT

The next corollary looks at the linear interpolation between the matrices associated to two labeled merge trees. Specifically, given any two labeled merge trees $T, T' \in \mathsf{LMT}$, we know that their associated matrices $M = \mathcal{M}(T)$, $M' = \mathcal{M}(T')$ are ultra matrices. We can define the line between them by setting $M^\lambda := (1 - \lambda)M + \lambda M'$ for $\lambda \in [0, 1]$. While not necessarily ultra matrices, it is easy to check that $M^\lambda \in \mathsf{VM}$ for all $\lambda \in [0, 1]$. We can then pull this back to a path of labeled merge trees by setting $T^\lambda = \mathcal{T}(M^\lambda)$.

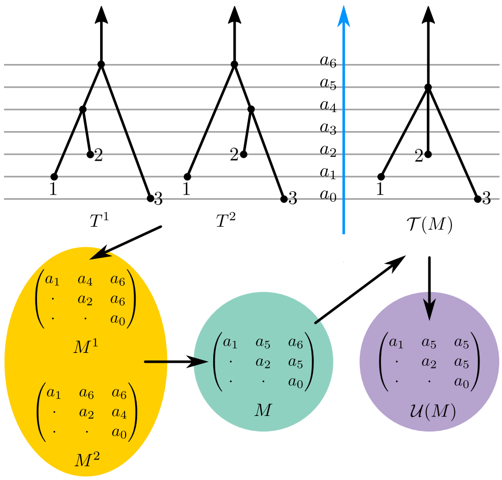


Fig. 4 An example of the averaging process for labeled merge trees. T^1 and T^2 are labeled merge trees with induced matrices M^1 and M^2. M is the pointwise average of M^1 and M^2, but is not an ultra matrix. The labeled merge tree $\mathcal{T}(M)$ is shown, whose induced matrix is the ultra matrix $\mathcal{U}(M)$

Corollary 3.2 (LMT Geodesics) *Given any two labeled merge trees $T, T' \in$ LMT, and their corresponding ultra matrices $M = \mathcal{M}(T)$, $M' = \mathcal{M}(T')$, the family of merge trees $\left\{T^\lambda := \mathcal{T}\left(M^\lambda\right)\right\}_{\lambda\in[0,1]}$ defines a geodesic between T and T' in the metric d_I^L. As a consequence, on the space of labeled merge trees, the metric d_I^L is strictly intrinsic.*

Proof Let δ denote the distance $d_I^L(T, T') = \|M - M'\|_\infty$. For any $0 \leq \lambda \leq \lambda' \leq 1$, the linearly interpolating matrices $M^\lambda, M^{\lambda'}$ satisfy $\|M^\lambda - M^{\lambda'}\|_\infty \leq (\lambda' - \lambda)\,\delta$. Hence, by Lemma 3.1, we have $d_I^L(T^\lambda, T^{\lambda'}) \leq (\lambda' - \lambda)\,\delta$. Since this is true for all $0 \leq \lambda \leq \lambda' \leq 1$, the triangle inequality implies that the family $\{T^\lambda\}_{\lambda\in[0,1]}$ defines a geodesic between T and T'. □

See the example of Fig. 4. Setting $\lambda = 1/2$, M^λ is the matrix (labeled M) shown in the middle green circle, and T^λ (labeled $\mathcal{T}(M)$) is the tree shown at the far right. Corollary 3.2 discusses the geodesics in the space of labeled merge trees. A metric space in general may have no geodesics; thus Corollary 3.2 provides an additional property for the space of interest. Furthermore, a geodesic can be used to perform shape morphing between a pair of merge trees (see [64]).

3.3 1-Centers in LMT

Our 1-center merge tree originates from the notion of a metric k-center in graph theory. Given m number of cities, one aims to build k facilities that minimize the maximum distance between a city to a facility.

Definition 13 Given a metric space (X, d), a 1-*center* $c \in X$ of a finite point set $P = \{p_1, \cdots, p_m\} \subset X$ is

$$c \in \arg\min_{x\in X} \max_{p\in P} d(x, p)$$

if it exists. That is, c is the center of a minimal enclosing ball of P.

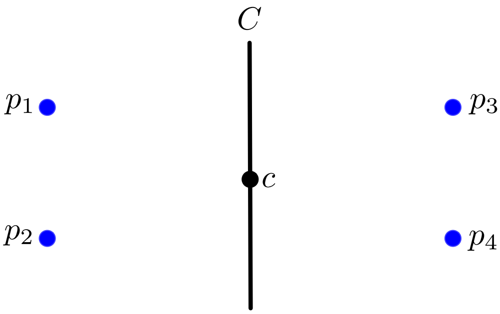


Fig. 5 An example of not unique 1-centers in a general metric space. $X = \mathbb{R}^2$ with metric $\|\cdot\|_\infty$ and $P = \{p_1, p_2, p_3, p_4\}$. The point c is the center of the range in each coordinate but every other point in C is also a 1-center

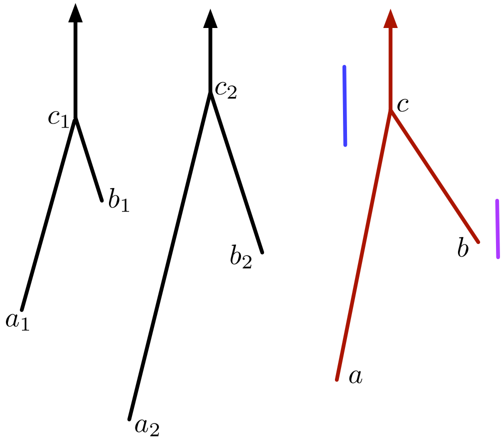


Fig. 6 An example of not unique 1-centers of a set of two labelled merge trees. All the 1-centers have the same tree shape, and the height of a must be the midpoint of a_1 and a_2 as the interval $|a_1 - a_2|$ is longer than both $|b_1 - b_2|$ and $|c_1 - c_2|$ but the heights of b and c can vary. It will be a 1-center whenever $b \in \left[\max\{b_1, b_2\} - \frac{|a_1-a_2|}{2}, \min\{b_1, b_2\} + \frac{|a_1-a_2|}{2}\right]$ (purple interval) and $c \in \left[\max\{c_1, c_2\} - \frac{|a_1-a_2|}{2}, \min\{c_1, c_2\} + \frac{|a_1-a_2|}{2}\right]$ (blue interval)

A metric 1-center will in general not be unique. For example, consider the metric space $\mathbb{R}^2$ with the l_∞ norm. If P is the set of points of the corners of rectangle (which is not a square), then only one of the coordinates of the 1-center is determined, see Fig. 5.

For $k = 1$, a metric 1-center of a finite set of labeled merge trees is one that minimizes the maximum distance to any other tree in the set. Again this is not in general unique. Analogously to the above example, we may have parameters that can be chosen within an interval as the range of these parameters in the set P is smaller. An example is shown in Fig. 6.

Here, we use the set notation $\in$ to indicate that c may not be unique. In the case of a finite collection of numbers χ in $\mathbb{R}$, the 1-center is simply the midpoint of the enclosing interval, $(\max(\chi) + \min(\chi))/2$. Now suppose we are given a collection of matrices $\{M^1, \ldots, M^N\}$. Let M_{mid} denote the matrix consisting of the entry-wise 1-center of the matrices, i.e., M_{ij} is the midpoint of the enclosing interval of numbers

$\{M_{ij}^1, M_{ij}^2, \ldots, M_{ij}^N\}$. It is easy to see that M_{mid} is a 1-center for these matrices in the space of all matrices equipped with the ℓ^∞ norm. A similar statement holds for a collection of valid matrices, and we include its simple proof for completeness.

Claim 3.3 *Let $M^1, \cdots, M^N$ be valid $n \times n$ matrices, and M_{mid} be the matrix consisting of the entry-wise 1-center of these matrices. Then M_{mid} must be valid as well and M_{mid} is a 1-center of $\{M^1, \ldots, M^N\}$ in the space of valid matrices equipped with the ℓ^∞ norm.*

Proof In what follows, all spaces of matrices are equipped with the ℓ^∞ norm. Since the space of valid matrices is a subspace of the space of all matrices, it follows that M_{mid} is a 1-center of $\{M^1, \ldots, M^N\}$ in the space of all matrices. Hence to prove the claim we only need to show that M_{mid} is a valid matrix. In other words, $(M_{mid})_{ii} \leq (M_{mid})_{ij}$ for any $i, j \in [n]$. To see why this holds, note that for any $i, j \in [n]$,

$$(M_{mid})_{ii} = \frac{\max_k(M_{ii}^k) + \min_k(M_{ii}^k)}{2} \leq \frac{\max_k(M_{ij}^k) + \min_k(M_{ij}^k)}{2} = (M_{mid})_{ij}.$$

The claim thus follows. □

M_{mid} as a 1-center of valid matrices is, by itself, a valid matrix, but may not be an ultra matrix, so we can replace it by its labeled merge tree (following the procedure described by Definition 8) and take its corresponding ultra matrix, thus turning it back to an ultra matrix.

The main result of this section is an algorithm to compute the 1-center of a collection of labeled merge trees under the labeled interleaving distance d_I^L. In particular, suppose we are given a set of labeled merge trees $\{T^1, \ldots, T^N\}$, whose corresponding induced matrices $\{M^1, \ldots, M^N\}$ are ultra (and thus also valid). We compute a 1-center valid matrix M_{mid} of $\{M^1, \ldots, M^N\}$ following Claim 3.3, and convert it to a labeled merge tree, denoted T^*. Then T^* is a 1-center of the labeled merge trees, see Fig. 4 for a simple example. The correctness of this procedure is established in the following Proposition 3.4.

Proposition 3.4 (LMT 1-Center) *Let $\{T^1, \ldots, T^N\}$ be a set of labeled merge trees, which gives rise to a set of ultra matrices $\{M^1, \ldots, M^N\}$. Let T^* be a merge tree constructed as above. Then T^* is a 1-center of $\{T^1, \ldots, T^N\}$. Furthermore, let $U^* = \mathcal{M}(T^*) = \mathcal{MT}(M_{mid})$ be the ultra matrix corresponding to T^*. Then U^* is a 1-center of the set of ultra matrices $\{M^1, \ldots, M^N\}$.*

Proof Recall that (the valid matrix) M_{mid} is a 1-center of ultra matrices $\{M^1, \ldots, M^N\}$ in the space of valid matrices following Claim 3.3. Set

$$\delta = \max_i \|M_{mid} - M^i\|_\infty.$$

Then $d_I^L(T, T^i) \leq \|M_{mid} - M^i\|_\infty$ by Lemma 3.1. It then follows that

$$\max_i d_I^L(T^*, T^i) \leq \max_i \|M_{mid} - M^i\|_\infty \leq \delta.$$

Thus $\{T_i\}_{i=1}^N$ is contained in a ball of radius δ centered at T^*.

We now show that this is in fact a *minimum enclosing ball* of $\{T^1, \ldots, T^N\}$ in the space of labeled merge trees, which would then imply that T^* is a 1-center for these merge trees. Specifically, assume there exists a $\widetilde{T}$ such that $\max_i d_I^L(\widetilde{T}, T^i) < \delta$. Set $\widetilde{U} = \mathcal{M}(\widetilde{T})$. Then for any i,

$$\|\widetilde{U} - M^i\|_\infty = d_I^L(\mathcal{T}(\widetilde{U}), \mathcal{T}(M^i)) = d_I^L(\widetilde{T}, T^i) < \delta.$$

Hence $\widetilde{U}$, as a valid matrix, gives rise to a smaller $\max_i \|\widetilde{U} - M^i\|$, which contradicts the assumption that M_{mid} is a 1-center within the space of valid matrices (i.e, $M_{mid} = \arg\min_M \max_i \|M - M^i\|$). Hence such a $\widetilde{T}$ cannot exist, and T^* is a 1-center for $\{T^1, \ldots, T^N\}$. By the relation between distance for ultra matrices and for their corresponding labeled merge trees, $U^* = \mathcal{M}(T^*)$ is a 1-center for $\{M^1, \ldots, M^N\}$, as well. □

Remark As a corollary of the above result, if we are given a collection of ultra matrices $\{M^1, \ldots, M^N\}$, then $U^* = \mathcal{M}\mathcal{T}(M_{mid})$ is a 1-center for them in the space of ultra matrices, where M_{mid} as defined earlier is the matrix consisting of the entry-wise 1-center of the input ultra matrices and M_{mid} is itself not necessarily a ultra matrix. Computing 1-centers for ultrametrics has been explored in the literature. While in general, this problem is NP-hard, for the case when we consider the ℓ^∞-norm on the space of ultrametrics (which is the same as our setting), it is known that there is a simple algorithm to compute it [72]. However, our approach above is completely different from the previous approach in [72], and has a different interpretation as well.

Fitting a distance matrix by ultrametrics and trees arises from the fields of mathematical psychology and evolutionary biology for such purposes as taxonomy and phylogenetic tree reconstruction. Farach et al. [73] studied the ultrametric ℓ^∞-fitting problem and proposed a rather involved algorithm that relies on computing cut-weights for edges in a minimum spanning tree. Agarwala et al. [74] studied the problem of fitting a distance matrix by a tree metric, and established that the tree fitting problem under the ℓ^∞ metric is NP-hard. They presented a polynomial 3-approximation algorithm using the Farris transform and modifying the approach of Farach et al. [73]. Chepoi and Fichet [72] studied the problem of finding the best ℓ^∞-fitting of distances by ultrametrics and tree metrics. Their problem formulation is as follows: given a vector u and a subset K of a real vector space, find the vector $\hat{u} \in K$ that is nearest to u in the ℓ^∞-norm [72]. They recovered and generalized the result of Agarwala et al. [74], but avoided solving a restricted ℓ^∞-approximation problem [72]. The provided general conditions on K under which a relationship between the subdominant of u and a best ℓ^∞-approximation holds [72]. Here, the subdominant u^* of u is the upper bound of the set $\{x \in K \mid x \prec u\}$, where $x \prec u$ means that all coordinates of x are smaller than or equal to those of u.

4 Interleaving Distances for Unlabeled Merge Trees

Moving to the unlabeled setting, we establish the existence of a certain labeling for a pair of merge trees that allows us to show that the interleaving distance for unlabeled merge trees is intrinsic.

Theorem 4.1 *Given two merge trees (T, f) and (T', f'), let L and L' be the respective leaf sets. Then*

$$d_I((T, f), (T', f')) = \inf_{\pi,\pi'} d_I^L((T, f, \pi), (T', f', \pi')) \tag{1}$$

where the infimum is taken over all finite labelings the two given merge trees, π and π', using at most $|L| + |L'|$ labels.

Note that the infimum in the statement implicitly is also taken over all vertex subdivisions of the tree since we can always add a vertex at a label point. Prior to proving the theorem, we will investigate the following construction of a labeling when given a δ-good map. First, note that given two labeled merge trees (T, f, π) and (T', f', π'), where $\pi : [n] \to V(T)$ and $\pi' : [n] \to V(T')$, the labeling information can be equivalently stored as an ordered collection of pairs $\Pi = \{(\pi(i), \pi'(i)) \mid i \in [n]\} \subseteq V(T) \times V(T')$. Since the order of the labels does not matter for this particular application, we will build Π iteratively. At the end, once we have constructed Π, we will assign the integers that index the labels to be compatible with Π.

Let L and L' denote the leaf sets for T and T', respectively. Assume we are given a δ-good map α as described in Definition 10. While this map is defined on the underlying metric trees, note that we can subdivide the trees so that $\alpha(v)$ is a vertex in T' for any vertex in T, and further that every point in the set $\alpha^{-1}(w)$ is a vertex in T if w is a vertex in T'.

Then, we construct the labeling Π as follows.

(S-1) Fix some $v \in L$, and let $w = \alpha(v)$. Then for every $u \in \alpha^{-1}(w)$, add (u, w) to Π and add the points to the vertex sets as necessary. Repeat this for every vertex in L.

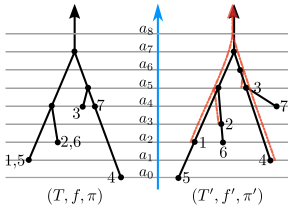
(S-2) For any leaf node $w \in L' \backslash \mathrm{Im}(\alpha)$, let x be its lowest ancestor contained in $\mathrm{Im}(\alpha)$. Let $u \in \alpha^{-1}(x)$ be a preimage of x from $|T|$. We assume that the same preimage u is always chosen for a given x; for example by assuming an ordering of the edges and vertices of T. Add (u, w) to Π and add each point to the respective vertex sets if necessary. Repeat for all leaves in L'.

(S-3) Fix an ordering on the pairs in $\Pi = \{(u_i, w_i) \mid i \in [n]\}$ and define $\pi(i) = u_i \in T$ and $\pi'(i) = w_i \in T'$.

Observe that since the preimage of any leaf node $w \in L' \cap \mathrm{Im}(\alpha)$ must be some vertex (or vertices) in L, any $w \in L' \cap \mathrm{Im}(\alpha)$ will be paired with some $u \in L$ by the process in (S-1), so this procedure does not miss any leaves in T'. See Fig. 7 for an example.

To use this construction to prove Theorem 4.1, we will use the following two lemmas.

$$\mathcal{M}(T,f,\pi) = \begin{pmatrix} a_1 & a_4 & a_7 & a_7 & a_1 & a_4 & a_7 \\ \cdot & a_2 & a_7 & a_7 & a_4 & a_2 & a_7 \\ \cdot & \cdot & a_4 & a_5 & a_7 & a_7 & \color{red}{a_5} \\ \cdot & \cdot & \cdot & a_0 & a_7 & a_7 & \color{red}{a_4} \\ \cdot & \cdot & \cdot & \cdot & a_1 & a_4 & a_7 \\ \cdot & \cdot & \cdot & \cdot & \cdot & a_2 & a_7 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & a_4 \end{pmatrix}$$



$$\mathcal{M}(T',f',\pi') = \begin{pmatrix} a_2 & a_5 & a_7 & a_7 & a_2 & a_5 & a_7 \\ \cdot & a_3 & a_7 & a_7 & a_5 & a_3 & a_7 \\ \cdot & \cdot & a_5 & a_5 & a_7 & a_7 & a_5 \\ \cdot & \cdot & \cdot & a_1 & a_7 & a_7 & a_5 \\ \cdot & \cdot & \cdot & \cdot & a_0 & a_5 & a_7 \\ \cdot & \cdot & \cdot & \cdot & \cdot & a_2 & a_7 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & a_4 \end{pmatrix}$$

Fig. 7 Given $\alpha : (T, f) \to (T', f')$, this is an example of the labeling induced by the procedure discussed after Theorem 4.1. The image of the map α is given by the red dashed lines, and α is δ-good for $\delta = a_{i+1} - a_i$. Labels 1–4 were generated in (S-1), the rest in (S-2). Note that there were two options for the location of label 7 in T. The other choice would be the same as the vertex labeled 3, and would only change the red entries in $\mathcal{M}(T, f, \pi)$

Lemma 4.2 *For any* $(u, w) \in \Pi$, $|f(u) - f'(w)| \le \delta$.

Proof If (u, w) is generated from (S-1) above, then the lemma holds by property (i) in the definition of the δ-good map α (see Definition 10). If (u, w) is generated from (S-2), then the lemma follows from property (iii) of the δ-good map α. Indeed, let x be the lowest ancestor of w contained in $\mathrm{Im}(\alpha)$, so that $\alpha(u) = x$. Then $0 \le f'(x) - f'(w) \le 2\delta$ and $f'(x) - f(u) = \delta$, implying that $|f'(w) - f(u)| \le \delta$. □

Lemma 4.3 *For any* $(u_1, w_1), (u_2, w_2) \in \Pi$,

$$|f(\mathrm{LCA}(u_1, u_2)) - f'(\mathrm{LCA}(w_1, w_2))| \le \delta.$$

Proof Assume we are given α, a δ-good map. If (u_i, w_i) is generated from (S-1), set $w_i' = w_i$. If (u_i, w_i) is generated via (S-2), then let w_i' be the lowest ancestor of w_i in $\mathrm{Im}(\alpha)$. In both cases, we have that $\alpha(u_i) = w_i'$ and $w_i \preceq w_i'$. Set $u_0 = \mathrm{LCA}(u_1, u_2)$, $w_0 = \mathrm{LCA}(w_1, w_2)$ and $w_0' = \mathrm{LCA}(w_1', w_2')$. This means, in particular, that $w_0 \prec w_0'$.

First, assume that both pairs are from (S-2), so that $w_1 \precneqq w_1'$ and $w_2 \precneqq w_2'$. Assume further that $w_0 \neq w_0'$. Since w_0 is the LCA of w_1 and w_2, the first assumption implies that $w_0 \preceq w_1'$ and $w_0 \preceq w_2'$. This means that at least one of $w_1' \preceq w_2'$ or $w_2' \preceq w_1'$ is true else we have a cycle. WLOG, assume $w_1' \preceq w_2'$. But then $w_2 \preceq w_0 \preceq w_1' \in \mathrm{Im}(\alpha)$ and since w_2' was defined as the lowest element above w_2 in the image of α, we have that $w_2' \preceq w_1'$. Hence $w_1' = w_2'$, and thus $w_1' = w_2' = w_0'$. Because we assumed in (S-2) that the same u is chosen for any given x, we can be assured that since $w_1' = w_2'$, $u_1 = u_2$. Hence $u_0 = u_1$. So

$$f(u_0) = f(u_1) = f'(w_1') - \delta = f'(w_0') - \delta.$$

Because α is δ-good, Definition 10(iii) implies

$$f(u_0) - \delta = f'(w_0') - 2\delta \le f'(w_0) \le f'(w_0') = f(u_0) + \delta$$

and so $|f(u_0) - f'(w_0)| \le \delta$, finishing the lemma for this case.

We show that in all remaining cases, $w_0 = w_0'$, and then use that to prove the lemma. We have already dealt with the case where both pairs come from (S-2). Assume that both pairs come from (S-1). It is immediate in this case that $w_0 = w_0'$. If exactly one comes from (S-1), assume that (u_1, w_1) comes from (S-1) and (u_2, w_2) comes from (S-2). This means $w_1 = w_1'$ and $w_2 \prec w_2'$. Since $w_1 \in \mathrm{Im}(\alpha)$ and $w_1 \prec w_0$, this means that $w_0 \in \mathrm{Im}(\alpha)$. If $w_0 \neq w_0'$, we would have a cycle of distinct elements $w_1 \prec w_0 \succ w_2 \prec w_2' \prec w_0' \succ w_1$ which is not possible in a tree. Thus $w_0 = w_0'$ in this case as well.

We will now prove the main claim, namely, that $|f(u_0) - f'(w_0)| \leq \delta$ in the case that $w_0 = w_0'$. To see that this is the case, assume that the claim does not hold; that is, either $f(u_0) - f'(w_0) > \delta$ or $f'(w_0) - f(u_0) > \delta$. Suppose first that $f'(w_0) - f(u_0) > \delta$, and consider $\alpha(u_0)$. Because $u_i \preceq u_0$ for $i = 1, 2$, by Lemma 2.4 we must have that $w_i' = \alpha(u_i) \preceq \alpha(u_0)$ for $i = 1, 2$. However, then $\alpha(u_0)$ is an ancestor of both w_1' and w_2' with

$$f'(\alpha(u_0)) = f(u_0) + \delta < f'(w_0),$$

contradicting the least common ancestor assumption of w_0.

Next, suppose $f(u_0) - f'(w_0) > \delta$ and consider $\alpha^{-1}(w_0)$. We claim that any point in $\alpha^{-1}(w_0)$ is a descendant of u_0; i.e., $v \preceq u_0$ for all $v \in \alpha^{-1}(w_0)$. Otherwise, we have that

$$f(\mathrm{LCA}(\alpha^{-1}(w_0))) > f(u_0) > f'(w_0) + \delta = f(v) + 2\delta$$

for any $v \in \alpha^{-1}(w_0)$, contradicting property (ii) of Definition 10. For $i = 1, 2$, let v_i be the unique ancestor of u_i with $f(v_i) = f'(w_0) - \delta$. By Lemma 2.4, since $\alpha(u_i) = w_i'$ and w_0 is the unique ancestor of w_i' with $f'(w_0) = f(v_i) + \delta$, this implies that $\alpha(v_i) = w_0$. That is, $v_i \in \alpha^{-1}(w_0)$. Further, $v_1 \neq v_2$. Otherwise if $v := v_1 = v_2$, then

$$f(v) = f'(w_0) - \delta < f(u_0) - 2\delta < f(u_0)$$

and thus v is a lower common ancestor of u_1 and u_2 than u_0, a contradiction. Hence, $\mathrm{LCA}(v_1, v_2) = u_0$. However,

$$f(u_0) - f(v_i) = f(u_0) - f'(w_0) + \delta > 2\delta.$$

This also contradicts property (ii) of Definition 10, finishing the proof of Lemma 4.3. □

Proof of Theorem 4.1 Say we have a δ-good map α for some $\delta \geq d_I((T, f), (T', f'))$. We construct the labelings π, π' as described above. Then Lemmas 4.2 and 4.3 imply that

$$d_I^L((T, f, \pi), (T', f', \pi')) \leq \delta.$$

As this is true for any δ,

$$\inf_{\Pi} d_I^L((T, f, \pi), (T', f', \pi')) \leq d_I((T, f), (T', f')).$$

To show the other inequality, assume we are given any pair of labelings π, π' and assume

$$d_I^L((T, f, \pi), (T', f', \pi')) = \delta.$$

We will construct the map α and show that it is δ-good.

For any $x \in |T|$, let $S_x \subseteq [n]$ be the labels in the subtree of x. Let y_i be the unique ancestor of $\pi'(i) \in |T'|$ for $i \in S_x$ with $f'(y_i) = f(x) + \delta$. First, we note that $y_i = y_j$ for all $i, j \in S_x$. Indeed, let $M = \mathcal{M}(T, f, \pi)$ and $M' = \mathcal{M}(T', f', \pi')$. By Lemma 2.2, $LCA(S_x) = \max_{i,j \in S_x} LCA(\pi(i), \pi(j))$. Then we know $M'_{ij} \leq \delta + M_{ij}$ for all pairs, and so for any $k \in S_x$,

$$\begin{aligned} f'(y_k) &= f(x) + \delta \geq f(\mathrm{LCA}(\pi(S_x))) + \delta \\ &= \max_{i,j \in S_x} M_{ij} + \delta \geq \max_{i,j \in S_x} M'_{ij} = f'(\mathrm{LCA}(\pi'(S_x))). \end{aligned}$$

Because every y_k has function value greater than the lowest common ancestor of $\pi'(S_x)$, the tree property implies that all y_k are equal. Thus, we can set $\alpha(x) = y_k$ for any $k \in S_x$ and it is well-defined and more-over continuous.

We need to ensure that the α constructed is δ-good as given in Definition 10. The map satisfies property (i) by construction, so we move on to (ii). Let $w \in \mathrm{Im}(\alpha)$ and set $x' = \mathrm{LCA}(\alpha^{-1}(w)) \in |T|$. Fix any $u \in \alpha^{-1}(w)$, and clearly $f(u) \leq f(x')$. By Lemma 2.2, x' must be $\mathrm{LCA}(u, u')$ for some other $u' \in \alpha^{-1}(w)$. Let i be a label in the subtree of u, and let j be a label in the subtree of u'. This further implies that $x' = \mathrm{LCA}(\pi(i), \pi(j))$. Set $w' = \mathrm{LCA}(\pi'(i), \pi'(j))$ and note that as $\pi'(i) \preceq w$ and $\pi'(j) \preceq w$, this implies that $w' \preceq w$. In particular, this means $f'(w') \leq f'(w)$. Further, by assumption $|f(x') - f'(w')| = |M_{ij} - M'_{ij}| \leq \delta$. Thus,

$$f(x') - f(u) \leq (f'(w) - f(u)) + (f(x') - f'(w')) + (f'(w') - f'(w)) \leq 2\delta$$

as the first part of the middle term is exactly δ, the second is $\leq \delta$, and the last is negative, showing that α satisfies property (ii).

Finally, we ensure property (iii). Let $w \in |T'| \setminus \mathrm{Im}(\alpha)$. Let i be the label of any leaf in the subtree of w, and set $y = \alpha(\pi(i))$ to be the image of the vertex labeled i in T. Then the tree property implies that $\pi'(i) \preceq w \preceq y$ and thus $f'(\pi'(i)) \leq f'(w) \leq f'(y)$. So,

$$\begin{aligned} |f'(w) - f'(\pi'(i))| \leq |f'(\pi'(i)) - f'(y)| &\leq |f'(\pi'(i)) - f(\pi(i))| + \delta \\ &= |M_{ii} - M'_{ii}| + \delta \leq 2\delta. \end{aligned}$$

As this is true for every leaf in the subtree of w, $\mathrm{depth}(w) \leq 2\delta$ and so α satisfies property (iii).

Thus, we have that $d_I((T, f), (T', f')) \leq d_I^L((T, f, \pi), (T', f', \pi'))$ for any given Π, completing the proof of the theorem. □

We conclude this section by showing that the interleaving distance is intrinsic on the space of finite (unlabeled) merge trees. Recall from Definition 12 that $\hat{d}$ denotes the intrinsic metric induced by a metric d.

Corollary 4.4 *For the space of finite (unlabeled) merge trees, $d_I = \hat{d}_I$.*

Proof Let (T, f) and (T', f') be two merge trees. Consider any $\varepsilon > 0$. According to Theorem 4.1, there are labelings π, π' of T, T' such that $d_I^L((T, f, \pi), (T', f', \pi')) \leq d_I((T, f), (T', f')) + \varepsilon$.

Now consider the space of labeled merge trees LMT. By Corollary 3.2, there exists a geodesic $\gamma : (T, f, \pi) \rightsquigarrow (T', f', \pi')$ in LMT such that the length $L_{d_I^L}(\gamma)$ is equal to $d_I^L((T, f, \pi), (T', f', \pi'))$.

Note that γ can be projected to a path γ' from T to T' in the space of (unlabeled) merge trees MT by simply ignoring the labeling. As $d_I((T, f), (T', f')) \leq d_I^L((T, f, \pi_1), (T', f', \pi_2))$ for any labelings π_1, π_2 between any two trees T and T', we have

$$\begin{aligned} \hat{d}_I((T, f), (T', f')) \leq L_{d_I}(\gamma') \leq L_{d_I^L}(\gamma) = d_I^L((T, f, \pi), (T', f', \pi')) \\ \leq d_I((T, f), (T', f')) + \varepsilon. \end{aligned} \tag{2}$$

On the other hand, by definition of the intrinsic metric $\hat{d}_I$ induced by d_I,

$$\hat{d}_I((T, f), (T', f')) \geq d_I((T, f), (T', f')). \tag{3}$$

Letting $\varepsilon \to 0$ in Eq. (2) and combining with Eq. (3), we obtain that $\hat{d}_I((T, f), (T', f')) = d_I((T, f), (T', f'))$. □

5 Concluding Remarks and Discussion

In this paper, we investigated whether interleaving-type distances for (finite) labeled or unlabeled merge trees are intrinsic or not, and presented positive answers in both cases. In the case of labeled trees, the geodesic between two labeled merge trees can be characterized and computed easily, and we also showed how to compute the 1-center of a set of labeled merge trees. In future work, it would be interesting to find a method to not only compute a 1-center but k-centers that could be incorporated into, say, a clustering method of an ensemble of merge trees. For unlabeled merge trees, however, computing the geodesic (even if just numerically estimating it) between two merge trees appears to be significantly harder, part of the reason being that it is NP-hard to approximate the interleaving distance between two merge trees, as pointed out in [30].

On the other hand, a simpler and easier to compute object is the bottleneck distance $d_B(T_1, T_2)$ between two (unlabeled) merge trees. We conjecture that the intrinsic distance $\hat{d}_B$ induced by d_B is in fact equivalent to $\hat{d}_I (= d_I)$.

Another natural question is whether (some of the) results for merge trees in this paper can be extended to contour trees. As a first question, can we characterize and compute the midpoint (i.e., the contour tree representing the 1-center) for two labeled contour trees under either $\hat{d}_I$, $\hat{d}_B$, or $\hat{d}_{FD}$ (where we remind the reader that d_{FD} denotes the functional distortion distance)? One idea is to compute the join and split trees of input contour trees, and compute the midpoint of the pair of join trees (resp., the pair of split trees). Note that each join or split tree can be viewed as a merge tree. Next we need to use the common ancestor information in both trees to construct a midpoint for the two contour trees. This step could be subtle: in particular, it is known [75] that in general, given a descending (join) tree T_J and an ascending (split) tree T_S with consistent functions associated to them, there may not exist a contour tree (or even a graph) whose join and split trees are equal to T_J and T_S, respectively. If such a contour tree exists, then it is unique, and the algorithm by Carr et al. [3] will compute this tree in near linear time.

Finally, understanding theoretical properties of distances between merge trees has many practical implications. For instance, in scientific visualization, such distances may be employed to study ensemble data sets that arise from scientific simulations (e.g., [52, 64]). Theorem 4.1 suggests the potential development of computing interleaving distances between unlabeled merge trees. Building on the work presented in this paper, Yan et al. [64] computed the structural average and geodesics of merge trees for uncertainty visualization. They explored various labeling strategies for computing interleaving distances between merge trees. Furthermore, Curry et al. [65] estimated the interleaving distance between unlabeled merge trees by searching for an optimal alignment between nodes in the trees with respect to a certain cost function; such estimation was used for classification and comparison of point cloud data. Moving beyond this paper, we envision a number of future applications in topological data analysis and visualization.

A Equivalence of δ-Good Map

We now prove Definition 10 is an equivalent version of the concept of a δ-good map as introduced in [30]. Recall that the original δ-good map is defined as follows.

Definition 14 [30] Given two merge trees $(T, f), (T', f')$, a δ-*good* map $\alpha : (T, f) \to (T', f')$ is a continuous map on the metric trees such that the following properties hold:

(P1) For any x in the geometric realization $|T|$, $f'(\alpha(x)) - f(x) = \delta$;
(P2) If $\alpha(u_1) \succeq \alpha(u_2)$, then we have that $u_1^{2\delta} \succeq u_2^{2\delta}$; and
(P3) If $w \in |T'|$ but $w \notin \mathrm{Im}(\alpha)$, then we have that $|f'(w^F) - f'(w)| \leq 2\delta$, where w^F is the lowest ancestor of w in $\mathrm{Im}(\alpha)$.
For any $w \notin \mathrm{Im}(\alpha)$, *depth*$(w) \leq 2\delta$.

Note that (P1) above is equivalent with condition (i) of our Definition 10. We will now show that properties (ii) and (iii) of Definition 10 are equivalent to properties (P2) and (P3) in the original definition above, respectively.
Equivalence of Definition 10(ii) with (P2) Above First, we show that Definition 10(ii) implies (P2). In particular, consider any $u_1, u_2 \in |T_1|$ such that $\alpha(u_1) \succeq \alpha(u_2)$. We aim to show that if Definition 10(ii) holds, then we must have $u_1^{2\delta} \succeq u_2^{2\delta}$.

To this end, let $u' \succeq u_2$ be the ancestor of u_2 such that $f(u') = f(u_1)$. (Note that as $\alpha(u_1) \succeq \alpha(u_2)$, we must have $f(u_1) \geq f(u_2)$.) By property (P1), we have that $f'(\alpha(u_1)) = f(u_1) + \delta$. It then follows from Lemma 2.4 that $\alpha(u') = \alpha(u_1)$ as $\alpha(u_1)$ is the unique ancestor of $\alpha(u_2)$ with f' value equal to $f(u') + \delta = f(u_1) + \delta$. This implies that both u_1 and u' have the same image under the map α. It then follows from property (ii) of Definition 10 that $f(x') \leq f(u_1) + 2\delta$ where $x' = \mathrm{LCA}(u_1, u')$, implying that $u_1^{2\delta} \succeq \mathrm{LCA}(u_1, u') \succeq \mathrm{LCA}(u_1, u_2)$. Since $f(u_2^{2\delta} \leq f(u_1^{2\delta})$, it then follows that $u_1^{2\delta} \succeq u_2^{2\delta}$.

Next, we show the opposite direction, namely (P2) above implies Definition 10(ii). In particular, consider any $w \in \mathrm{Im}(\alpha)$ with $x' = \mathrm{LCA}(\alpha^{-1}(w))$. Our goal is to show that $f(x') - f(u) \leq 2\delta$ for all $u \in \alpha^{-1}(w)$, assuming that (P2) holds. To see this, consider any two $u_1, u_2 \in \alpha^{-1}(w)$. Obviously, $\alpha(u_1) = \alpha(u_2) \Rightarrow \alpha(u_1) \succeq \alpha(u_2)$. It then follows from (P2) that $u_1^{2\delta} \succeq u_2^{2\delta}$. Since this holds for any two nodes in $\alpha^{-1}(w)$, it follows that $u_1^{2\delta} = u_2^{2\delta} \succeq x' = \mathrm{LCA}(\alpha^{-1}(w)$. Hence $f(x') \leq f(u) + 2\delta$, and property Definition 10(ii) thus holds.

Equivalence of Definition 10(iii) and (P3) First, assume (P3) holds, and we aim to show (iii) is true. Indeed, consider any $w \in |T'|$ but $w \notin \mathrm{Im}(\alpha)$. Now assume *depth*$(w) > 2\delta$. This means that there exists $w' \prec w$ (i.e, w' is a descendent of w) such that $f'(w) > f'(w') + 2\delta$. Note that by Lemma 2.4, if $w \notin \mathrm{Im}(\alpha)$, then no descendent of w could lie in $\mathrm{Im}(\alpha)$. It then follows that the lowest ancester of w' in $\mathrm{Im}(\alpha)$ is the same as that for w, which is w^F. Hence we have now found a point $w' \notin \mathrm{Im}(\alpha)$ but $f'(w^F) - f'(w') > f'(w) - f'(w') > 2\delta$, which contradicts property (P3). Hence the assumption is wrong and it must be that *depth*$(w) \leq 2\delta$ for any $w \in |T'| \backslash \mathrm{Im}(\alpha)$. That is, Definition 10(iii) holds.

Now consider the opposite direction. Assume Definition 10(iii) holds, and we aim to show (P3). Consider any $w \in |T'| \backslash \mathrm{Im}(\alpha)$ and let w^F be its lowest ancestor that is in $\mathrm{Im}(\alpha)$; note that as $w^F \succeq w$, $f'(w^F) \geq f'(w)$. Now suppose (P3) does not hold, meaning that $f'(w^F) - f'(w) > 2\delta$. This means that $w^{2\delta} \notin \mathrm{Im}(\alpha)$. Now take any node w' (not equal to $w^{2\delta}$ nor w^F) such that $w^{2\delta} \prec w' \prec w^F$. We have that $w' \notin \mathrm{Im}(\alpha)$ but $f'(w') > f'(w) + 2\delta$. This means that *depth*$(w') > 2\delta$, which contradicts the property stated in Definition 10(iii). Hence our assumption is wrong, and thus (P3) must hold.

Putting everything together, our concept of δ-good map as introduced in Definition 10 is equivalent to the one introduced in [30] (shown in Definition 14).

Acknowledgements Our initial research collaboration began during the Dagstuhl Seminar 17292: Topology, Computation and Data Analysis in July 2017. We thank all members of the breakout session on Reeb graphs for stimulating discussions. We are grateful to the Institute for Computational and Experimental Research in Mathematics (ICERM) for supporting us through the Collaborate@ICERM program in August 2018. EM was partially supported by National Science Foundation (NSF) through grants CMMI-1800466 and DMS-1800446. YW was partially supported by NSF through grants CCF-1740761 and DMS-1547357, as well as by National Institute of Health (NIH) under grant R01EB022899. BW was partially supported by NSF IIS-1513616, NSF DBI-1661375 and NIH R01EB022876.

Data availability Data sharing not applicable to this article as no datasets were generated or analyzed during the current study.

Declarations

Conflict of interest On behalf of all authors, the corresponding author states that there is no Conflict of interest.

References

1. Edelsbrunner, H., Letscher, D., Zomorodian, A.: Topological persistence and simplification. Discret. Comput. Geom. **28**, 511–533 (2002)
2. Ghrist, R.: Barcodes: the persistent topology of data. Bull. Am. Math. Soc. **45**(1), 61–75 (2008)
3. Carr, H., Snoeyink, J., Axen, U.: Computing contour trees in all dimensions. Comput. Geom. **24**(2), 75–94 (2003)
4. Reeb, G.: Sur les points singuliers d'une forme de Pfaff complèment intégrable ou d'une fonction numérique. Comptes Rendus de L'Académie ses Séances **222**, 847–849 (1946)
5. Flamm, C., Hofacker, I.L., Stadler, P.F., Wolfinger, M.T.: Barrier trees of degenerate landscapes. Zeitschrift für Physikalische Chemie **216**(2) (2002)
6. Milnor, J.: Morse Theory. Princeton University Press, New Jersey (1963)
7. Singh, G., Mémoli, F., Carlsson, G.: Topological methods for the analysis of high dimensional data sets and 3D object recognition. In: Eurographics Symposium on Point-Based Graphics, pp. 91–100 (2007)
8. Carrière, M., Oudot, S.: Structure and stability of the one-dimensional mapper. Found. Comput. Math. **18**(6), 1333–1396 (2018)
9. Munch, E., Wang, B.: Convergence between categorical representations of Reeb space and Mapper. In: Proceedings of the 32nd International Symposium on Computational Geometry, vol. 51, pp. 53–15316 (2016)
10. Carrière, M., Michel, B., Oudot, S.: Statistical analysis and parameter selection for mapper. J. Mach. Learn. Res. **19**, 1–39 (2018)
11. Weber, G., Bremer, P.-T., Pascucci, V.: Topological landscapes: a terrain metaphor for scientific data. IEEE Trans. Vis. Comput. Graph. **13**(6), 1416–1423 (2007)
12. Oesterling, P., Heine, C., Jaenicke, H., Scheuermann, G., Heyer, G.: Visualization of high-dimensional point clouds using their density distribution's topology. IEEE Trans. Vis. Comput. Graph. **17**(11), 1547–1559 (2011)
13. Widanagamaachchi, W., Jacques, A., Wang, B., Crosman, E., Bremer, P.-T., Pascucci, V., Horel, J.: Exploring the evolution of pressure-perturbations to understand atmospheric phenomena. In: IEEE Pacific Visualization Symposium (2017)
14. Le Gall, Jean-Francois.: Brownian excursions, trees and measure-valued branching processes. Ann. Probab. **19**(4), 1399–1439 (1991)
15. Evans, S.N.: Probability and Real Trees. Springer, Berlin (2006)
16. Morozov, D., Beketayev, K., Weber, G.: Interleaving distance between merge trees. In: Proceedings of Topology-Based Methods in Visualization (2013)
17. de Silva, V., Munch, E., Patel, A.: Categorified Reeb graphs. Discret. Comput. Geom. 1–53 (2016)
18. Bauer, U., Ge, X., Wang, Y.: Measuring distance between Reeb graphs. In: Proceedings of the 13th Annual Symposium on Computational Geometry, pp. 464–473 (2014)
19. Bauer, U., Munch, E., Wang, Y.: Strong equivalence of the interleaving and functional distortion metrics for Reeb graphs. In: Proceedings of the 31st International Symposium on Computational Geometry, vol. 34, pp. 461–475 (2015)
20. Bauer, U., Di Fabio, B., Landi, C.: An edit distance for Reeb graphs. In: Proceedings of the Eurographics Workshop on 3D Object Retrieval, pp. 27–34 (2016)
21. Di Fabio, B., Landi, C.: The edit distance for Reeb graphs of surfaces. Discret. Comput. Geom. **55**(2), 423–461 (2016)
22. Bauer, U., Landi, C., Memoli, F.: The Reeb graph edit distance is universal. Found. Comput. Math. **21**, 1441–1464 (2021)
23. Sridharamurthy, R., Masood, T.B., Kamakshidasan, A., Natarajan, V.: Edit distance between merge trees. IEEE Trans. Vis. Comput. Graph. **26**(3), 1518–1531 (2020)
24. Carrière, M., Oudot, S.: Local equivalence and intrinsic metrics between Reeb graphs. In: Proceedings of the 33rd International Symposium on Computational Geometry, vol. 77, pp. 25–12515 (2017)

25. Beketayev, K., Yeliussizov, D., Morozov, D., Weber, G.H., Hamann, B.: Measuring the distance between merge trees. In: Mathematics and Visualization, pp. 151–165. Springer, Berlin (2014)
26. Yan, L., Masood, T.B., Sridharamurthy, R., Rasheed, F., Natarajan, V., Hotz, I., Wang, B.: Scalar field comparison with topological descriptors: properties and applications for scientific visualization. Comput. Graph. Forum **40**(3), 599–633 (2021)
27. Bollen, B., Chambers, E., Levine, J.A., Munch, E.: Reeb graph metrics from the ground up. arXiv preprint arXiv:2110.05631 (2021)
28. Chazal, F., Cohen-Steiner, D., Glisse, M., Guibas, L.J., Oudot, S.Y.: Proximity of persistence modules and their diagrams. In: Proceedings of the 25th Annual Symposium on Computational Geometry, pp. 237–246 (2009)
29. Chazal, F., de Silva, V., Glisse, M., Oudot, S.: The Structure and Stability of Persistence Modules. Springer, Berlin (2016)
30. Touli, E.F., Wang, Y.: FPT-algorithms for computing Gromov–Hausdorff and interleaving distances between trees. In: Proceedings of the 27th Annual European Symposium on Algorithms, pp. 83–18314 (2019)
31. Curry, J.: Sheaves, cosheaves and applications. PhD thesis, University of Pennsylvania (2014)
32. Bubenik, P., de Silva, V., Scott, J.: Metrics for generalized persistence modules. Found. Comput. Math. **15**(6), 1501–1531 (2014)
33. de Silva, V., Munch, E., Stefanou, A.: Theory of interleavings on categories with a flow. Theory Appl. Categ. **33**(21), 583–607 (2018)
34. Munch, E., Stefanou, A.: The ℓ^∞-cophenetic metric for phylogenetic trees as an interleaving distance. In: Research in Data Science vol. 17, pp. 109–127. Springer, Cham (2019)
35. Sokal, R.R., Rohlf, F.J.: The comparison of dendrograms by objective methods. Taxon **11**(2), 33 (1962)
36. Cardona, G., Mir, A., Rosselló, F., Rotger, L., Sánchez, D.: Cophenetic metrics for phylogenetic trees, after Sokal and Rohlf. BMC Bioinform. **14**(1), 3 (2013)
37. Robinson, D.F., Foulds, L.R.: Comparison of weighted labelled trees. In: Combinatorial Mathematics VI, pp. 119–126. Springer, Berlin (1979)
38. Robinson, D.F., Foulds, L.R.: Comparison of phylogenetic trees. Math. Biosci. **53**(1), 131–147 (1981)
39. DasGupta, B., He, X., Jiang, T., Li, M., Tromp, J., Zhang, L.: On distances between phylogenetic trees. In: Proceedings of the 8th Annual ACM-SIAM Symposium on Discrete Algorithms, vol. 97, pp. 427–436 (1997)
40. DasGupta, B., He, X., Jiang, T., Li, M., Tromp, J.: On the linear-cost subtree-transfer distance between phylogenetic trees. Algorithmica **25**(2–3), 176–195 (1999)
41. Diaconis, P.W., Holmes, S.P.: Matchings and phylogenetic trees. Proc. Natl. Acad. Sci. **95**(25), 14600–14602 (1998)
42. Billera, L.J., Holmes, S.P., Vogtmann, K.: Geometry of the space of phylogenetic trees. Adv. Appl. Math. **27**(4), 733–767 (2001)
43. Bogdanowicz, D., Giaro, K.: On a matching distance between rooted phylogenetic trees. Int. J. Appl. Math. Comput. Sci. **23**(3), 669–684 (2013)
44. Bogdanowicz, D., Giaro, K.: Matching split distance for unrooted binary phylogenetic trees. IEEE/ACM Trans. Comput. Biol. Bioinf. **9**(1), 150–160 (2012)
45. Estabrook, G.F., McMorris, F.R., Meacham, C.A.: Comparison of undirected phylogenetic trees based on subtrees of four evolutionary units. Syst. Biol. **34**(2), 193–200 (1985)
46. Cardona, G., Llabrés, M., Rosselló, F., Valiente, G.: Nodal distances for rooted phylogenetic trees. J. Math. Biol. **61**(2), 253–276 (2009)
47. Choi, K., Gomez, S.M.: Comparison of phylogenetic trees through alignment of embedded evolutionary distances. BMC Bioinform. **10**(1), 423 (2009)
48. Lafond, M., El-Mabrouk, N., Huber, K.T., Moulton, V.: The complexity of comparing multiply-labelled trees by extending phylogenetic-tree metrics. Theoret. Comput. Sci. **760**, 15–34 (2019)
49. Miller, E., Owen, M., Provan, J.S.: Polyhedral computational geometry for averaging metric phylogenetic trees. Adv. Appl. Math. **68**, 51–91 (2015)
50. Markin, A., Eulenstein, O.: Cophenetic median trees under the Manhattan distance. In: Proceedings of the 8th ACM International Conference on Bioinformatics, Computational Biology, and Health Informatics, pp. 194–202 (2017)
51. Gavryushkin, A., Drummond, A.J.: The space of ultrametric phylogenetic trees. J. Theor. Biol. **403**, 197–208 (2016)

52. Pont, M., Vidal, J., Delon, J., Tierny, J.: Wasserstein distances, geodesics and barycenters of merge trees. IEEE Trans. Vis. Comput. Graph. **1**(28), 291–301 (2022)
53. Wetzels, F., Pont, M., Tierny, J., Garth, C.: Merge tree geodesics and barycenters with path mappings. In: IEEE Transactions on Visualization and Computer Graphics, pp. 1–11 (2023)
54. Gurvich, V., Vyalyi, M.: Characterizing (quasi-)ultrametric finite spaces in terms of (directed) graphs. Discret. Appl. Math. **160**(12), 1742–1756 (2012)
55. Hughes, B.: Trees and ultrametric spaces: a categorical equivalence. Adv. Math. **189**(1), 148–191 (2004)
56. Lemin, A.J.: The category of ultrametric spaces is isomorphic to the category of complete, atomic, tree-like, and real graduated lattices LAT^*. Algebra Universalis **50**(1), 35–49 (2003)
57. Dovgoshey, O., Petrov, E., Teichert, H.-M.: How rigid the finite ultrametric spaces can be? J. Fixed Point Theory Appl. **19**(2), 1083–1102 (2016)
58. Dovgoshey, O., Petrov, E.: From isomorphic rooted trees to isometric ultrametric spaces, p-Adic numbers. Ultrametric Anal. Appl. **10**(4), 287–298 (2018)
59. Leclerc, B.: Consensus of classifications: the case of trees. In: Rizzi, A., Vichi, M., Bock, H.-H. (eds.) Advances in Data Science and Classification, pp. 81–90. Springer, Berlin (1998)
60. Chepoi, V., Fichet, B.: ℓ^∞-approximation via subdominants. J. Math. Psychol. **44**(4), 600–616 (2000)
61. Barthélemy, J.-P., Leclerc, B.: The median procedure for partitions. Partitioning Data Sets **19**, 3–34 (1993)
62. De Soete, G.: A least squares algorithm for fitting additive trees to proximity data. Psychometrika **48**(4), 621–626 (1983)
63. Lapointe, F.-J., Cucumel, G.: The average consensus procedure: combination of weighted trees containing identical or overlapping sets of taxa. Syst. Biol. **46**(2), 306–312 (1997)
64. Yan, L., Wang, Y., Munch, E., Gasparovic, E., Wang, B.: A structural average of labeled merge trees for uncertainty visualization. IEEE Trans. Vis. Comput. Graph. **26**(1), 832–842 (2020)
65. Curry, J., Hang, H., Mio, W., Needham, T., Okutan, O.B.: Decorated merge trees for persistent topology. arXiv preprint arXiv:2103.15804 (2021)
66. Smith, Z., Chowdhury, S., Mémoli, F.: Hierarchical representations of network data with optimal distortion bounds. In: Proceedings of 50th Asilomar Conference on Signals, Systems and Computers, pp. 1834–1838 (2016)
67. Jardine, N., Sibson, R.: Mathematical Taxonomy. WileySeries in Probability and Mathematical Statistics, John Wiley & Sons Ltd., London (1971)
68. Hartigan, J.A.: Statistical theory in clustering. J. Classif. **2**, 63–76 (1985)
69. Carlsson, G., Mémoli, F.: Characterization, stability and convergence of hierarchical clustering methods. J. Mach. Learn. Res. **11**(47), 1425–1470 (2010)
70. Semple, C., Steel, M., Caplan, R.A.: Phylogenetics. Oxford University Press, Oxford (2003)
71. Burago, D., Burago, Y., Ivanov, S.: A Course in Metric Geometry, vol. 33. American Mathematical Society, Providence (2001)
72. Chepoi, V., Fichet, B.: l_∞-approximation via subdominants. J. Math. Psychol. **44**(4), 600–616 (2000)
73. Farach, M., Kannan, S., Warnow, T.: A robust model for finding optimal evolutionary trees. Algorithmica **13**, 155–179 (1995)
74. Agarwala, R., Bafna, V., Farach, M., Paterson, M., Thorup, M.: On the approximability of numerical taxonomy: fitting distances by tree metrics. In: Proceedings of the 7th Annual ACM-SIAM Symposium on Discrete Algorithms (1996)
75. Wang, S., Wang, Y., Wenger, R.: The JS-graph of join and split trees. In: Proceedings of the 30th Annual Symposium on Computational Geometry, pp. 539–548 (2014)