A Topological Similarity Measure Between Multi-Resolution Reeb Spaces

Yashwanth Ramamurthi, Tripti Agarwal, and Amit Chattopadhyay, *Member, IEEE*

Abstract—Searching similarity between a pair of shapes or data is an important problem in data analysis and visualization. The problem of computing similarity measures using scalar topology has been studied extensively and proven useful in the shape and data matching. Even though multi-field or multivariate (consists of multiple scalar fields) topology reveals richer topological features, research on building tools for computing similarity measures using multi-field topology is still in its infancy. In the current article, we propose a novel similarity measure between two piecewise-linear multi-fields based on their multi-resolution Reeb spaces - a newly developed data-structure that captures the topology of a multi-field. Overall, our method consists of two steps: (i) building a multi-resolution Reeb space corresponding to each of the multi-fields and (ii) proposing a similarity measure between two multi-resolution Reeb spaces by computing a list of topologically consistent matching pairs (of nodes) and the similarity between them. We demonstrate the effectiveness of the proposed similarity measure in detecting topological features from real time-varying multi-field data in two application domains - one from computational physics and one from computational chemistry.

Index Terms—Multi-field, reeb space, similarity measure, multi-resolution, joint contour net

1 Introduction

Topological similarity between shapes or data reveals important features useful to the domain scientists. Similarity measures using scalar topology have been successfully applied in shape matching [1], noise removal from point-cloud data [2], classification of bio-molecular or protein structures [3], symmetry detection [4], and periodicity analysis in time-dependant flows [5] to name a few. The design of these similarity or distance measures were mostly based on different scalar topology descriptors, such as, contour tree, Reeb graph, merge tree, Morse-Smale complex, extremum graph and persistence diagram.

Multi-field topology is usually richer than scalar topology, and therefore, more precise similarity measures are possible to design using multi-field topology, for better classification of shapes or data. Algorithms for the computation of multi-field topology have been designed using Jacobi sets [6], fiber topology [7], Reeb spaces [8] and multidimensional persistence [9]. Approximations of the Reeb space have been proposed using a mapper [10] and a joint contour net [11] to visualize topological features in multi-field data. However, to design similarity measures using multi-field topology one requires to generalize the existing scalar topology based techniques, both theoretically and computationally.

Towards this, we propose the first method for computing a topological similarity measure between two Reeb spaces. Our method generalizes the multi-resolution Reeb graph based similarity measure proposed by Hilaga *et al.* [1], which was designed for univariate or scalar data. However, a suitable extension of this method to measure topological similarity between multi-field data was an open problem. In the current paper, we contribute as follows:

- We introduce a novel multi-resolution Reeb space (MRS) data-structure, consisting of a finite sequence of Reeb spaces corresponding to an increasing sequence of resolutions, to capture the fiber-topology of a multi-field. The MRS data-structure generalizes the multi-resolution Reeb graph (MRG) structure [1] to multi-fields. Computing an MRS is non-trivial as compared to an MRG.
- We propose a novel technique for computing a list of topologically consistent matching pairs of nodes between two MRSs. Unlike finding the similarity between two MRGs [1], it is more challenging to find matching pairs between two Reeb spaces. Our method first constructs a multi-dimensional Reeb graph (MDRG) corresponding to each Reeb space and then applies a label propagation technique in the corresponding MDRGs to decide the topologically consistent matching pairs.
- Finally, we design a novel similarity measure between two MRS structures using the similarity values for the list of topologically consistent matching pairs. The similarity value between a matching pair is computed based on different geometric and topological properties of the corresponding fiber-components. We show, in particular, the proposed measure is a metric.
- We validate the proposed multi-field topology based similarity measure in detecting the nuclear scission

- *Yashwanth Ramamurthi and Amit Chattopadhyay are with the International Institute of Information Technology (IIIT), Bangalore, Karnataka 560100, India. E-mail: {yashwanth, a.chattopadhyay}@iiitb.ac.in.*
- *Tripti Agarwal is with the SCI, University of Utah, Salt Lake City, UT 84112 USA. E-mail: tripti.agarwal@iiitb.org.*

Manuscript received 11 Oct. 2020; revised 21 May 2021; accepted 1 June 2021. Date of publication 8 June 2021; date of current version 27 Oct. 2022.
(Corresponding author: Amit Chattopadhyay.)
Recommended for acceptance by J. Tierny.
Digital Object Identifier no. 10.1109/TVCG.2021.3087273

point in a time-varying multi-field data of Fermium-258 atom, described in [12]. In a second application, we show the effectiveness of our method in identifying the site of stable Pt-CO bond formation in a series of molecular orbital (for HOMO, LUMO, HOMO−1) data. It is observed that using a scalar field, the method fails to detect the correct site of Pt-CO stable bond formation. This case study emphasizes the necessity of developing multi-field topology based measures over scalar topology measures.

Outline. Section 2 discusses various topology-aware similarity and distance measures between scalar and multi-fields related to the current work. Section 3 provides the necessary background to understand the proposed multi-field similarity measure. Section 4 introduces the proposed multi-resolution Reeb space data-structure and an algorithm for computing the data-structure. Section 5 describes our algorithm for computing a similarity measure between two MRSs. In Section 6, we analyze the time-complexity of our algorithm. Section 7.1 provides the implementation details of the algorithm along with the experimental and comparison results using two real multi-field data. Finally, Section 8 provides a conclusion of the paper with a summary and future directions.

2 Related Work

Similarity and distance measures between scalar fields have been studied extensively using different topological descriptors, such as, contour trees, merge trees, Reeb graphs, extremum graphs and persistence diagrams. A survey paper by Biasotti *et al.* [13] gives an overview of different topological shape descriptors for scalar fields. Topological shape matching using Reeb graphs and their extensions have been studied extensively by different groups [14], [15], [16], [17], [18]. Hilaga *et al.* [1] proposed a multi-resolution Reeb graph based similarity measure and applied their method in the shape matching problem. A similar technique was proposed by Zhang *et al.* [3] to compute a similarity measure between two volumetric scalar data using their multi-resolution dual contour trees and showed an application in molecular shape matching [3]. However, generalizing these works for multi-fields is challenging.

Persistence diagram is one of the most efficient tools for the topological analysis of high-dimensional data [19]. A bottleneck distance was proposed between two persistence diagrams and proven stable [2]. Persistence and bottleneck distance have been widely applied in topological data analysis, noise removal, simplification and image analysis [2]. Dey *et al.* [20] proposed a persistence distortion metric for comparing metric graphs while Agarwal *et al.* [21] used the Gromov-Hausdorff distance [22] to compare metric trees. Carlsson *et al.* [23] proposed the theory of multi-dimensional persistence, although no generalization of the bottleneck distance is known between multi-fields.

Thomas *et al.* [4] introduced a contour clustering based technique for detecting multiscale symmetry in a volumetric scalar data. Narayanan *et al.* [24] proposed a topological distance between two extremum graphs for comparing two scalar fields. Various distance metrics have also been proposed based on merge trees. Beketayev *et al.* [25] proposed a metric between merge trees and showed the effectiveness of their metric by comparing it with the bottleneck distance. Saikia *et al.* [5] introduced a technique for computing the repeating topological pattern between two scalar data by comparing all subtrees between their merge trees using an extended branch decomposition graph (eBDG). In another work, Saikia *et al.* [26] proposed a histogram based distance measure by computing the histograms from the merge trees.

A survey paper by Gao *et al.* [27] discusses different inexact algorithms in the literature based on the graph edit distance. Fabio *et al.* [28] proposed an edit distance between Reeb graphs of surfaces and showed the stability property of the proposed metric. Recently, Sridharamurthy *et al.* [29] presented an edit distance-based metric between two merge trees for feature visualization in a time-varying scalar field data. Various metrics between Reeb graphs such as functional distortion metric [30] and interleaving distance [31] have been proposed.

Bauer *et al.* [33] showed the equivalence between the functional distortion metric and the interleaving distance proposed between two Reeb graphs. In a recent paper, Carrière *et al.* [34] showed that although bottleneck distance between two Reeb graphs is a pseudo-metric, it can discriminate a Reeb graph from the other Reeb graphs in a small enough neighbourhood, as efficiently as the Gromov-Hausdorff, functional distortion and interleaving distances.

However, research on computing topological similarity measures between multi-fields is still at a nascent stage. Singh *et al.* [10] proposed a mapper data-structure for capturing the topology of high-dimensional datasets and computed a dissimilarity matrix for the point-cloud data. Dey *et al.* [35] extended the work to a multiscale mapper (MSM) using a tower of covers and proved its stability by showing the bottleneck distance between the associated persistence diagrams is upper bounded by a distance between the maps. In the current paper, we propose a multi-resolution Reeb space (MRS) by considering a sequence of Reeb spaces at different resolutions. The sequence of resolutions of MRS can be considered as a special case of the tower of covers in MSM which makes MRS computation more geometric and MRS structure is simplified as a sequence of 1-dimensional simplicial complexes, whereas MSM is a tower of nerves. Moreover, the bottleneck distance between the persistence diagrams corresponding to two MSMs is a pseudo-metric between the MSMs (similar as the bottleneck distance between Reeb graphs in [30]), whereas we propose a similarity metric between two MRSs.

Carr *et al.* [11] developed a joint contour net (JCN) for a quantized approximation of the Reeb space which is a special case of the mapper structure. Duke *et al.* [12] showed an effective application of the JCN in visualizing nuclear scission in a multi-field density data. Recently, Agarwal *et al.* [36] proposed a distance metric between two multi-fields based on their fibre-component distributions and demonstrated its usefulness over scalar-topology. However, a fiber-component distribution is obtained by projecting a Reeb space onto the range, therefore, some features of the Reeb space might be lost. In the current paper, we propose an improved similarity measure by directly comparing two multi-resolution Reeb spaces.

3 Background

In this section, we describe the necessary background to understand the proposed similarity measure between two

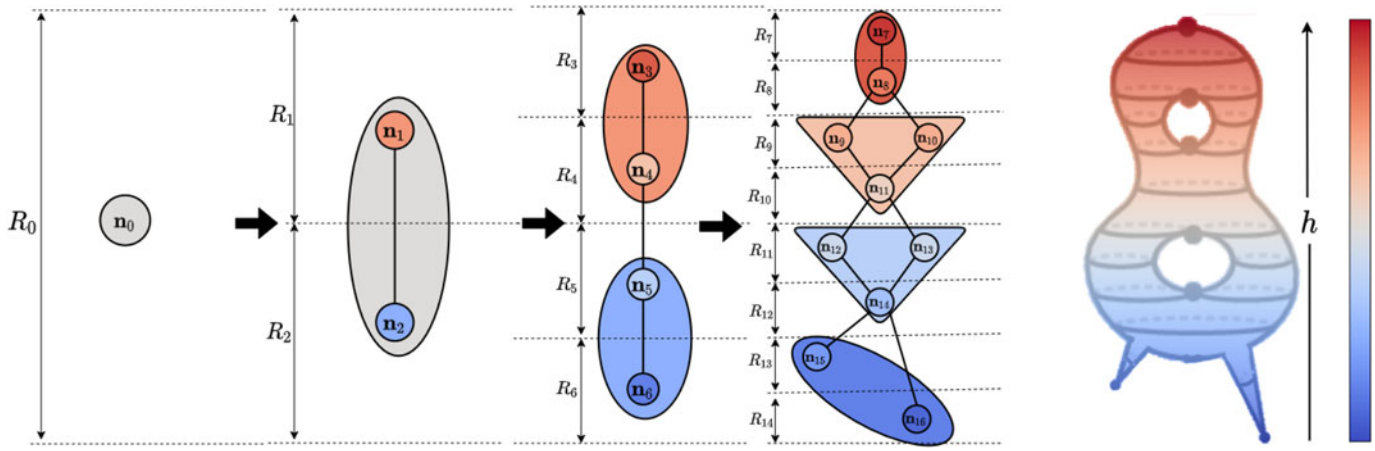


Fig. 1. An MRG of the height function h of a standing double torus with legs. The figure shows the MRG with four Reeb graphs at four different resolutions - coarser to finer Reeb graphs are shown from the left to right.

MRSs. More precisely, we briefly highlight the important tools for capturing the scalar and multi-field topology.

Piecewise-Linear Multi-Fields. Most of the data in scientific visualization comes as a discrete set of real values at the vertices (grid-points) of a mesh in a volumetric domain. Let $\mathcal{M}$ be a compact d-manifold data domain and $\mathbb{M}$ be a triangulation (mesh) of $\mathcal{M}$ whose vertices contain the data values. Let $\mathbf{V}(\mathbb{M}) = \{\mathbf{v}_0, \mathbf{v}_1, \ldots, \mathbf{v}_u\}$ be the set of vertices of $\mathbb{M}$. Mathematically, an r-dimensional multi-field data can be defined as a vertex map $\hat{\mathbf{f}} = (\hat{f}_1, \hat{f}_2, \ldots, \hat{f}_r) : \mathbf{V}(\mathbb{M}) \rightarrow \mathbb{R}^r$ which maps each vertex to an r-tuple vector of scalar values. From this discrete map we define a piecewise-linear (PL) multi-field $\mathbf{f} = (f_1, f_1, \ldots, f_r) : \mathbb{M} \rightarrow \mathbb{R}^r$ as $\mathbf{f}(\mathbf{x}) = \sum_{i=0}^{u} \beta_i \hat{\mathbf{f}}(\mathbf{v}_i)$ where $\mathbf{x} \in \sigma$ (a simplex of $\mathbb{M}$) has a unique convex combination of its vertices that can be expressed as $\mathbf{x} = \sum_{i=0}^{u} \beta_i \mathbf{v}_i$ with $\beta_i \geq 0$ and $\sum_{i=0}^{u} \beta_i = 1$. We note, $\mathbf{f}$ is continuous and the restriction of $\mathbf{f}$ over each simplex of $\mathbb{M}$ is linear. In the current paper, we consider any PL multi-field $\mathbf{f} : \mathbb{M} \rightarrow \mathbb{R}^r$ with the restriction $d \geq r \geq 1$ as our input multi-field. In particular, for $r = 1$, we obtain a PL scalar field.

Reeb Space and Reeb Graph. Given a PL multi-field $\mathbf{f} : \mathbb{M} \rightarrow \mathbb{R}^r$ and a range value $\mathbf{c} \in \mathbb{R}^r$, the inverse image $\mathbf{f}^{-1}(\mathbf{c}) = \{\mathbf{x} \in \mathbb{M} : \mathbf{f}(x) = \mathbf{c}\}$ is called a *fiber* and a connected component of the fiber is called a *fiber-component* [7], [37]. Mathematically, each fiber-component is an equivalence class obtained by an equivalence relation $\sim$ on $\mathbb{M}$, defined as $\mathbf{x} \sim \mathbf{y} \Leftrightarrow \mathbf{x}, \mathbf{y} \in \mathbb{M}$, $\mathbf{f}(\mathbf{x}) = \mathbf{f}(\mathbf{y}) = \mathbf{h}$ and $\mathbf{x}$, $\mathbf{y}$ belong to the same connected component of $\mathbf{f}^{-1}(\mathbf{h})$. Thus, the domain $\mathbb{M}$ is partitioned into a set of fiber-components by this equivalence relation. The Reeb space of $\mathbf{f}$, denoted by $\mathbb{W}_\mathbf{f}$, is a quotient space obtained by contracting each fiber-component to a unique point. The topology of $\mathbb{W}_\mathbf{f}$ is induced by the quotient map $q_\mathbf{f} : \mathbb{M} \rightarrow \mathbb{W}_\mathbf{f}$ which maps each point of $\mathbb{M}$, on a fiber-component, to the corresponding point of $\mathbb{W}_\mathbf{f}$ [8]. Geometrically, under some regularity conditions, $\mathbb{W}_\mathbf{f}$ is an r-dimensional polyhedron [37]. Fig. 3 shows an illustration of a bivariate field and its Reeb space.

In particular, for a PL scalar field $f : \mathbb{M} \rightarrow \mathbb{R}$, the fiber $f^{-1}(c) = \{x \in \mathbb{M} : f(x) = c\}$ is known as a *level set* of the isovalue $c \in \mathbb{R}$ and a connected component of the level set is called a *contour* instead of fiber-component. Under some regularity conditions, the Reeb space of the scalar field $f : \mathbb{M} \rightarrow \mathbb{R}$ is a 1-dimensional CW-complex or a graph structure, known as the Reeb graph and is denoted by $\mathcal{RG}_f$ [38]. Therefore, $\mathcal{RG}_f$ consists of a set of *nodes* and *arcs*, each arc connecting two of the nodes. Each point of the Reeb graph corresponds to a contour. In particular, the nodes of the Reeb graph correspond to the contours passing through the critical points [39] of f and the arcs connecting the nodes represent the contours which pass through the regular points (not critical) of f. If the domain $\mathbb{M}$ is simply-connected, then the Reeb graph does not contain any loop and is a tree which is known as the contour tree [40].

Multi-Resolution Reeb Graph. A multi-resolution Reeb graph (MRG) of a PL scalar field $f : \mathbb{M} \rightarrow \mathbb{R}$, proposed by Hilaga *et al.* [1], is a data-structure that computes a finite series of Reeb graphs at various levels of data resolutions. In practice, each Reeb graph is obtained by subdividing the data range into a set of $Q = 2^k$, $(k = 0, 1, \ldots, L-1)$ levels of resolution by a dyadic subdivision (as shown in Fig. 1). The domain $\mathbb{M}$ is partitioned into fat (or quantized) contours accordingly and the quantized Reeb graph at that resolution is the adjacency graph of the fat contours. The series of Reeb graphs in an MRG satisfy three properties: (i) a parent-child relationship is maintained between the nodes of the adjacent Reeb graphs at consecutive levels, (ii) by repeating the process of subdivision, when the levels of resolution goes to infinity the MRG converges to the actual Reeb graph of f and (iii) a Reeb graph at a particular resolution contains all the information about the coarser resolution Reeb graphs [1]. Fig. 1 shows the MRG of the height field of a standing double torus (with legs) at 4 different resolutions. In the current paper, we generalize the MRG structure to a MRS structure for capturing the multi-field topology.

Joint Contour Net. A joint contour net (JCN) of a PL multi-field $\mathbf{f} = (f_1, \ldots, f_r) : \mathbb{M} \rightarrow \mathbb{R}^r$, proposed by Carr *et al.* [11], is a quantized approximation of the Reeb space using a chosen number of quantization levels (or levels of resolution). The data range is first quantized or subdivided into $Q = q_1 \times q_2 \times \cdots \times q_r$ levels of quantization (here, range of f_i is quantized into q_i levels, $i = 1, 2, \ldots, r$). Thus the range of $\mathbf{f}$ is discretized into a finite set, such as a subset of $\mathbb{Z}^r$. In this case, for a quantized range value $\mathbf{h} \in \mathbb{Z}^r$, instead of a fiber we obtain a *quantized fiber* or *joint level set*, denoted as $\tilde{\mathbf{f}}^{-1}(\mathbf{h}) = \{\mathbf{x} \in \mathbb{M} : round(\mathbf{f}(\mathbf{x})) = \mathbf{h}\}$ where the *round* function is applied on each component field. Quantized fibers

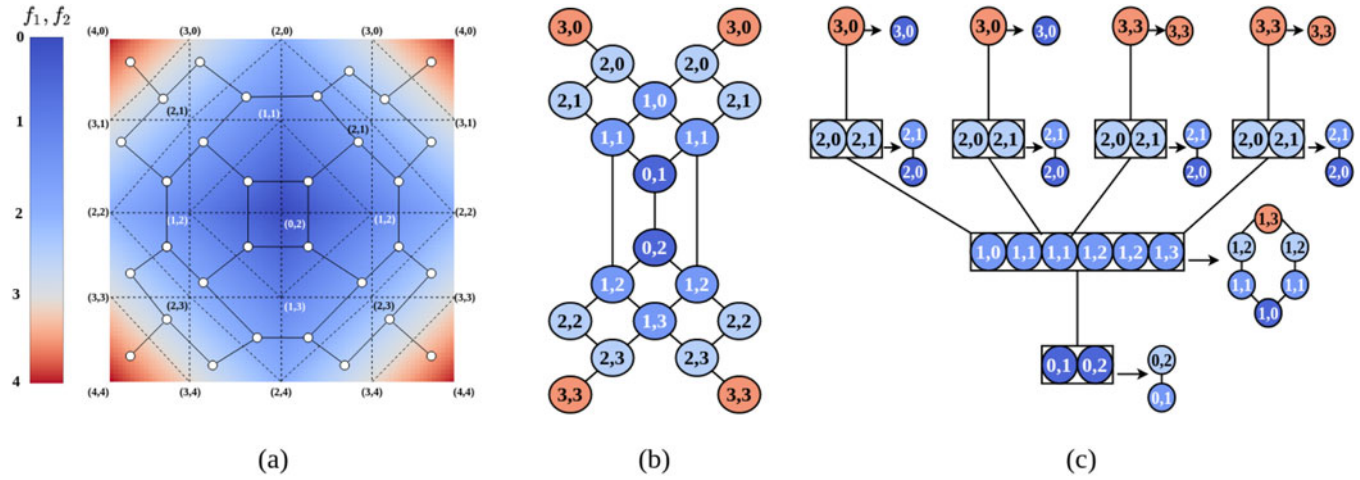


Fig. 2. (a) A PL bivariate data over a 2D mesh: The mesh coloring is based on the color map using the first field (f_1), (b) JCN at 4×4 resolution: the nodes are colored based on the values of f_1, (c) MDRG computed using the algorithm in [32]: the nodes in the first (second) dimension are colored based on the values of the first (second) field.

are not always connected. A connected component of a quantized fiber is called a *quantized fiber-component* or *joint contour*. The joint contour net is an adjacency graph of the joint contours or quantized fiber-components. Each node in the JCN corresponds to a joint contour and the adjacency between two joint contours is represented by an edge in the JCN. Fig. 2b shows an example of the JCN of a simulated bivariate dataset in Fig. 2a. In our method, we compute a multi-resolution Reeb space of a multi-field by computing a series of JCNs at different resolutions.

Multi Dimensional Reeb Graph. A Multi Dimensional Reeb Graph (MDRG) of a PL multi-field $f : \mathbb{M} \to \mathbb{R}^r$, proposed by Chattopadhyay *et al.* [32], [41], is a hierarchical decomposition of the Reeb space (or a joint contour net) into a set of Reeb graphs in different dimensions. In particular, to construct the MDRG for a PL bivariate field $\mathbf{f} = (f_1, f_2) : \mathbb{M} \to \mathbb{R}^2$, first we compute the Reeb graph $\mathcal{RG}_{f_1}$ of the field f_1 (in the first dimension). Now each point $p \in \mathcal{RG}_{f_1}$ corresponds to a contour, say C_p, of f_1. Then we restrict function f_2 (in the second dimension) on C_p and define the restricted function as $\widetilde{f}_2^p = f_2|_{C_p}$. Finally, we compute the Reeb graphs $\mathcal{RG}_{\widetilde{f}_2^p}$ for these restricted functions $\widetilde{f}_2^p$, in the second dimension. Thus MDRG of $\mathbf{f}$, denoted by $\mathrm{MDRG}_{\mathbf{f}}$, can be defined as: $\mathrm{MDRG}_{\mathbf{f}} = \{(p_1, p_2) : p_1 \in \mathcal{RG}_{f_1}, p_2 \in \mathcal{RG}_{\widetilde{f}_2^{p_1}}\}$. The definition can be extended for any PL multi-field $\mathbf{f} : \mathbb{M} \to \mathbb{R}^r$ with $d \geq r \geq 2$. Fig. 2b shows an example of a quantized Reeb space or JCN for a PL bivariate field (in Fig. 2a) and Fig. 2c shows its MDRG. In our algorithm of measuring the similarity between two JCNs, we use the corresponding MDRGs to decide a topologically consistent matching pair of nodes from the JCNs.

4 Multi-Resolution Extension of Reeb Space

In this section, we develop a new multi-resolution Reeb space (MRS) structure that captures the topology of a PL multi-field data for an increasing sequence of resolutions, similar as MRG of a PL scalar field [1]. Each Reeb space at a particular resolution is approximated by a JCN. The motivation behind using an MRS is to capture and compare the topological features which persist in its sequence JCNs, instead of just considering the topological features in a single JCN.

4.1 Overview

Let $\mathbf{f} = (f_1, f_2, \ldots, f_r) : \mathbb{M} \to \mathbb{R}^r$ be a PL multi-field and $\mathrm{JCN}_{\mathbf{f}}(\mathbf{q})$ denote the JCN of $\mathbf{f}$ with parameter-vector $\mathbf{q} = (q_1, q_2, \ldots, q_r)$, each parameter q_i being the levels of resolution of the component field f_i for $i \in \{1, 2, \ldots, r\}$. Thus, using $\mathbf{q}$, we obtain an approximated Reeb space $\mathrm{JCN}_{\mathbf{f}}(\mathbf{q})$ with $Q = q_1 \times q_2 \times \ldots \times q_r$ levels of resolution corresponding to $\mathbf{f}$. Now to construct a multi-resolution Reeb space, for simplicity, we consider a finite sequence $\{\mathbf{q}^{(k)} = (2^k, 2^k, \ldots, 2^k) : k = 0, 1, \ldots, L-1\}$ of $L\,(\geq 1)$ increasing levels of resolution of $\mathbf{f}$. Corresponding to this sequence of resolutions, we obtain a sequence $\{\mathrm{JCN}_{\mathbf{f},k} : k = 0, 1, \ldots, L-1\}$ where $\mathrm{JCN}_{\mathbf{f},k}$ represents the JCN of $\mathbf{f}$ corresponding to the resolution-vector $\mathbf{q}^{(k)}$. This sequence of JCNs defines a *Multi-resolution Reeb Space* (MRS) of $\mathbf{f}$ with L levels of resolution and is denoted by $\mathrm{MRS}_{\mathbf{f},L}$. The JCNs in $\mathrm{MRS}_{\mathbf{f},L}$ satisfy the following three properties:

(P1) There are parent-child relationships between the nodes of JCNs in adjacent resolutions, i.e between

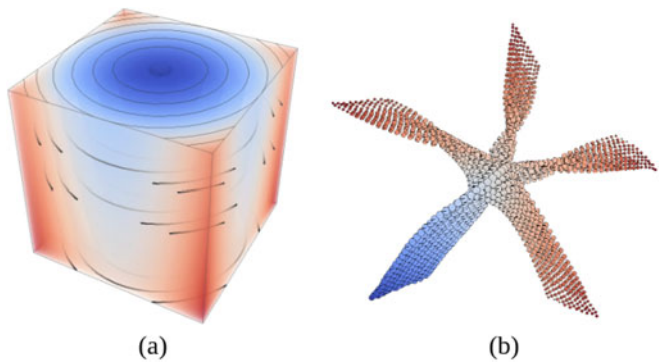


Fig. 3. (a) A PL bivariate data consisting of paraboloid and height fields in a 3D mesh. Each connected component of the black coloured curves is a fiber-component. (b) Reeb Space of the bivariate field. The color map of the domain and JCN is based on the value of the paraboloid field.

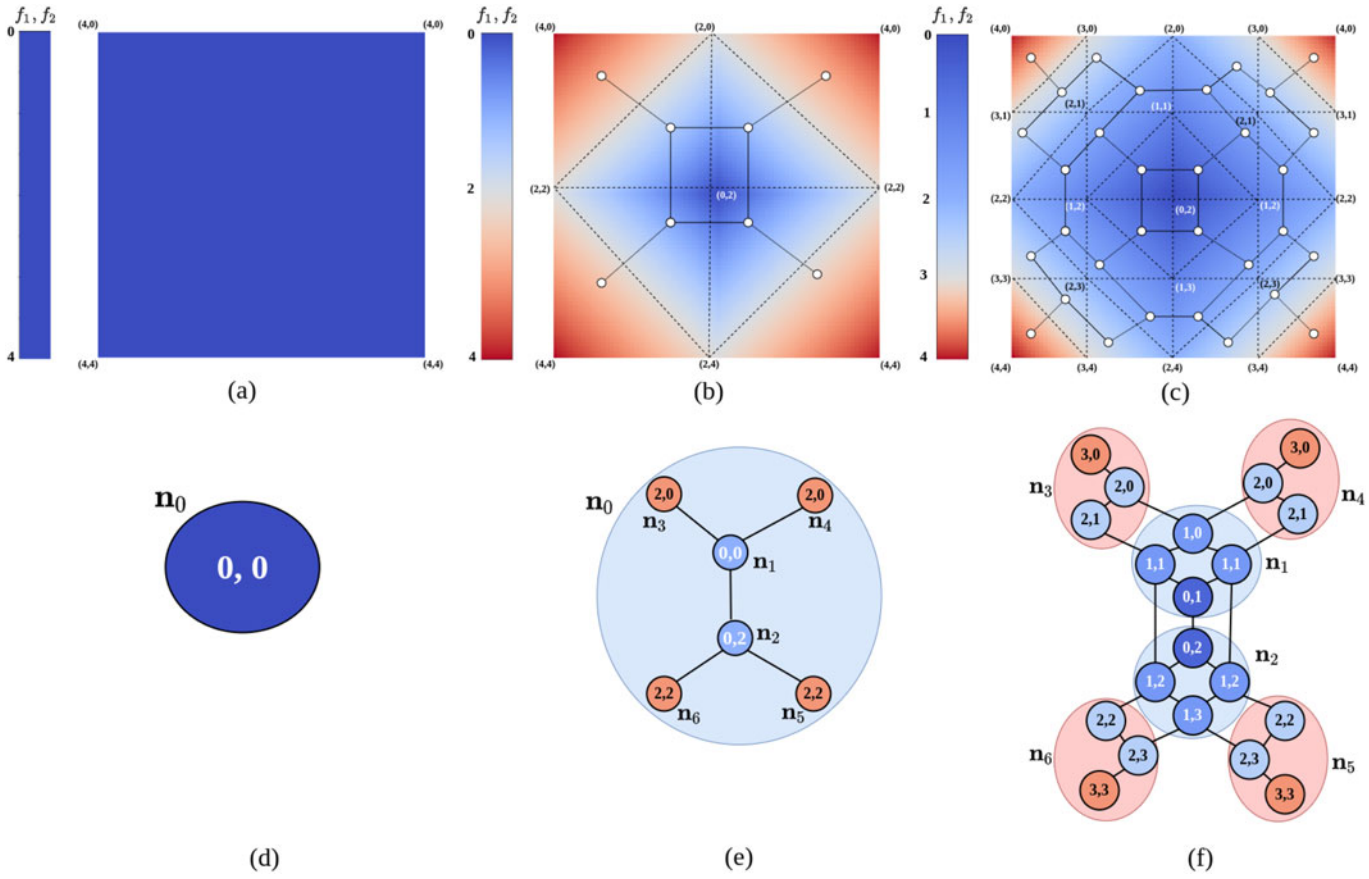


Fig. 4. Multi-resolution Reeb Space corresponding to a PL bivariate data: (ring, height). (a) Each component field is quantized into one level, and the corresponding JCN in (d) consists of only one node $\mathbf{n}_0$. (b) Each component field is quantized into two levels and the corresponding JCN is shown is (e). (c) Each component field is quantized into four levels, and the corresponding JCN is shown in (f). Parent-child relationships between the nodes of JCNs in consecutive resolutions are shown, e.g., $\mathbf{n}_0$ is parent of $\{\mathbf{n}_1, \mathbf{n}_2, \mathbf{n}_3, \mathbf{n}_4, \mathbf{n}_5, \mathbf{n}_6\}$. The meshes (a)-(c) and their corresponding JCNs (d)-(f) are colored by applying the respective color maps using the first scalar field.

the nodes of $\mathrm{JCN}_{\mathbf{f},k}$ and $\mathrm{JCN}_{\mathbf{f},k+1}$. In Fig. 4, node $\mathbf{n}_0$ of the JCN in (d) is the parent of the nodes $\{\mathbf{n}_1, \mathbf{n}_2, \mathbf{n}_3, \mathbf{n}_4, \mathbf{n}_5, \mathbf{n}_6\}$ of the JCN in (e).

(P2) Note that when the levels of resolution goes to infinity the JCN graph converges to the Reeb space of $\mathbf{f}$, i.e., $\mathrm{JCN}_{\mathbf{f},k}$ converges to the Reeb space $\mathbb{W}_{\mathbf{f}}$ as $k \to \infty$ (see [32] for details). In otherwords, we can say the multi-resolution Reeb space $\mathrm{MRS}_{\mathbf{f},L}$ converges to $\mathbb{W}_{\mathbf{f}}$ as the levels of resolution L tends to ∞.

(P3) A JCN of a certain resolution in $\mathrm{MRS}_{\mathbf{f},L}$ implicitly contains all the information of the JCNs of coarser resolutions, i.e., $\mathrm{JCN}_{\mathbf{f},k}$ contains all the information of $\mathrm{JCN}_{\mathbf{f},0}$, $\mathrm{JCN}_{\mathbf{f},1}$, $\ldots$, $\mathrm{JCN}_{\mathbf{f},k-1}$. Once a JCN of a certain resolution is constructed, the coarser resolution JCN can be constructed by grouping the adjacent nodes in the same coarser range interval, as shown in Fig. 4.

4.2 Construction of the Multi-Resolution Reeb Space

Let the components of the PL multi-field $\mathbf{f} = (f_1, f_2, \ldots, f_r)$ be quantized into $\mathbf{q}^{(L-1)} = (2^{L-1}, 2^{L-1}, \ldots, 2^{L-1})$ dyadic levels of resolutions where $L\,(\geq 1)$ is an integer. Then the joint contour net $\mathrm{JCN}_{\mathbf{f},L-1}$ of $\mathbf{f}$ with respect to the parameter vector $\mathbf{q}^{(L-1)}$ can be constructed using the algorithm described by Carr *et al.* [11].

To construct the multi-resolution Reeb space $\mathrm{MRS}_{\mathbf{f},L}$, we start with the finest resolution Reeb space $\mathrm{JCN}_{\mathbf{f},L-1}$ and from that compute the coarser resolution Reeb spaces by merging pairs of consecutive range intervals (for each of the component fields), sequentially. This is a generalization of the MRG construction by Hilaga *et al.* [1]. Algorithm 1 outlines the method for constructing a coarser resolution Reeb space $\mathrm{JCN}_{\mathbf{f},k-1}$ from the finer resolution Reeb space $\mathrm{JCN}_{\mathbf{f},k}$ ($k \geq 1$ is an integer). Let $\mathcal{R}_{\mathbf{f}} = [f_1^{\min}, f_1^{\max}] \times [f_2^{\min}, f_2^{\max}] \times \cdots \times [f_r^{\min}, f_r^{\max}]$ be the r-dimensional range interval of the multi-field $\mathbf{f}$ where $f_i^{\min} = \min_{\mathbf{x}\in\mathbb{M}} f_i(\mathbf{x})$ and $f_i^{\max} = \max_{\mathbf{x}\in\mathbb{M}} f_i(\mathbf{x})$ for $i = 1, 2, \ldots, r$. The range $[f_i^{\min}, f_i^{\max}]$ of the component field f_i is subdivided into $q = 2^k$ ($k \in \mathbb{N}$: set of natural numbers) dyadic levels of resolution or sub-intervals: $R_0^{(i)} = [x_0^{(i)}, x_1^{(i)})$, $R_1^{(i)} = [x_1^{(i)}, x_2^{(i)}), \ldots, R_{q-1}^{(i)} = [x_{q-1}^{(i)}, x_q^{(i)}]$, where $x_0^{(i)} = f_i^{\min}$ and $x_q^{(i)} = f_i^{\max}$ (for $i = 1, 2, \ldots, r$). Thus the range $\mathcal{R}_{\mathbf{f}}$ is subdivided into $Q = q \times q \times \cdots \times q$ (r times) dyadic levels of resolution or sub-intervals, denoted by $\mathcal{R}_{i_1 i_2 \ldots i_r} = R_{i_1}^{(1)} \times R_{i_2}^{(2)} \times \cdots \times R_{i_r}^{(r)}$ (where $i_1, i_2, \ldots, i_r = 0, 1, \ldots, q-1$).

For computing the coarser level Reeb space $\mathrm{JCN}_{\mathbf{f},k-1}$, we merge the adjacent levels of $\mathrm{JCN}_{\mathbf{f},k}$ (in pairs) and obtain the coarser sub-intervals as $\widetilde{\mathcal{R}}_{i_1 i_2 \ldots i_r} = \mathrm{merge}(R_{2i_1}^{(1)}, R_{2i_1+1}^{(1)}) \times \mathrm{merge}(R_{2i_2}^{(2)}, R_{2i_2+1}^{(2)}) \times \cdots \times \mathrm{merge}(R_{2i_r}^{(r)}, R_{2i_r+1}^{(r)})$ (for $i_1, i_2, \ldots, i_r = 0, 1, \ldots, \frac{q}{2} - 1$). So the levels of resolution of $\mathrm{JCN}_{\mathbf{f},k-1}$ reduces to $\widetilde{Q} = \frac{q}{2} \times \frac{q}{2} \times \cdots \times \frac{q}{2}$. We construct a Union-Find structure UF

[42] from the adjacency of the nodes of $\mathrm{JCN}_{\mathbf{f},k}$ with ranges in $\widetilde{R}_{i_1 i_2 \ldots i_r}$. Each connected component of the UF becomes a node of the coarser Reeb space $\mathrm{JCN}_{\mathbf{f},k-1}$ and the adjacencies of these new nodes are determined by their adjacencies in $\mathrm{JCN}_{\mathbf{f},k}$.

Algorithm 1. CREATECOARSERREEBSPACE

Input: $\mathrm{JCN}_{\mathbf{f},k}$
Output: $\mathrm{JCN}_{\mathbf{f},k-1}$
for $i_1 = 0$ to $\frac{q}{2} - 1$ **do**
  for $i_2 = 0$ to $\frac{q}{2} - 1$ **do**
    $\ddots$
    for $i_r = 0$ to $\frac{q}{2} - 1$ **do**
      % Range merging for coarser resolution
      $\widetilde{\mathcal{R}}_{i_1 i_2 \ldots i_r} = \mathrm{merge}(R^{(1)}_{2i_1}, R^{(1)}_{2i_1+1}) \times \mathrm{merge}(R^{(2)}_{2i_2}, R^{(2)}_{2i_2+1}) \times \cdots \times \mathrm{merge}(R^{(r)}_{2i_r}, R^{(r)}_{2i_r+1})$
      Create Union-Find Structure UF for the nodes of $\mathrm{JCN}_{\mathbf{f},k}$ in range $\widetilde{\mathcal{R}}_{i_1 i_2 \ldots i_r}$
      % Creating nodes of coarser JCN
      for each component C_j in UF **do**
        Create a node $\mathbf{n}_{C_j}$ in $\mathrm{JCN}_{\mathbf{f},k-1}$
        Map node-ids and field-values of the finer JCN to it
        Set node $\mathbf{n}_{C_j}$ as the parent for the nodes in C_j
      end for
    end for
  end for
  $\ddots$
end for
% Adding edges in coarser JCN
for each edge $e_1 e_2$ in $\mathrm{JCN}_{\mathbf{f},k}$ **do**
  if $e_1, e_2 \in$ components $C_j \neq C_l$ with $\mathbf{f}(e_1) \neq \mathbf{f}(e_2)$ **then**
    Add edge $e(\mathbf{n}_{C_j}, \mathbf{n}_{C_l})$ in $\mathrm{JCN}_{\mathbf{f},k-1}$ if not already present
  end if
end for
return $\mathrm{JCN}_{\mathbf{f},k-1}$

Algorithm 1 gives the pseudo-code of the method. Fig. 4 illustrates an MRS construction for a simple bivariate data in a 2D box domain with 3 levels of resolution. The construction of the MRS starts with the construction of the Reeb space with $Q = 2^2 \times 2^2$ levels at the finest resolution.

4.3 Node Attributes for Similarity Measure

To obtain a quantitative similarity measure between two multi-resolution Reeb spaces, we associate different attributes to the nodes of the MRS to quantify topological and geometrical properties of the fibers corresponding to a multi-field data by generalizing the scalar data attributes as in [3]. We consider the attribute set corresponding to a node $\mathbf{n}$, denoted by $\widetilde{\mathbf{n}}$, as $\widetilde{\mathbf{n}} = \{\mathcal{V}(\mathbf{n}), \mathcal{R}_0(\mathbf{n}), \mathcal{B}_0(\mathbf{n}), \mathcal{D}(\mathbf{n})\}$ where $\mathcal{V}(\mathbf{n}) = \frac{\mathrm{Volume}(\mathbf{n})}{\mathrm{Total\ Volume}(\mathbb{M})}$ is the normalized volume of the quantized fiber-component corresponding to the node $\mathbf{n}$, $\mathcal{R}_0(\mathbf{n}) = \frac{\mathrm{measure(range}(\mathbf{n}))}{\mathrm{measure(range}(\mathbf{f}))}$ is the normalized range, $\mathcal{B}_0(\mathbf{n})$ is the number of components of the joint level set corresponding to $\mathbf{n}$ and $\mathcal{D}(\mathbf{n})$ is the degree of $\mathbf{n}$ in the corresponding JCN. We denote any attribute corresponding to a node $\mathbf{n}$ by $\mathcal{A}(\mathbf{n})$.

5 A SIMILARITY MEASURE BETWEEN MRSS

Let $\mathrm{MRS}_{\mathbf{f},L}$ and $\mathrm{MRS}_{\mathbf{g},L}$ be two multi-resolution Reeb spaces with same levels of resolution (here, L) corresponding to two PL multi-fields $\mathbf{f}$ and $\mathbf{g}$, respectively. Our method of computing the similarity measure between two MRSs has two steps: 1. Creating a list of matching pairs of nodes from the respective MRSs and 2. Computing the similarity measure between the MRSs by defining a similarity measure between the pair of nodes in a matched pair. We describe these steps in the following subsections.

5.1 Creating Matching Pairs

To create the list of matching pairs between two multi-resolution Reeb spaces, denoted by MPAIR, we search from the coarser to the finer resolution Reeb spaces. Nodes $\mathbf{m} \in \mathrm{MRS}_{\mathbf{f},L}$ and $\mathbf{n} \in \mathrm{MRS}_{\mathbf{g},L}$ form a matching pair $(\mathbf{m}, \mathbf{n}) \in \mathrm{MPAIR}$ if they satisfy the following *matching rules* (generalizing the rules in [1], [3]):

(i) $\mathbf{m} \in \mathrm{MRS}_{\mathbf{f},L}$ and $\mathbf{n} \in \mathrm{MRS}_{\mathbf{g},L}$ do not belong to any other matched pair,
(ii) $\mathbf{m}$ and $\mathbf{n}$ belong to two JCNs, respectively, of same resolution and both nodes have the same range values,
(iii) Parent $\mathcal{P}(\mathbf{m})$ of $\mathbf{m}$ and parent $\mathcal{P}(\mathbf{n})$ of $\mathbf{n}$ must have been matched, i.e., $(\mathcal{P}(\mathbf{m}), \mathcal{P}(\mathbf{n}))$ is already a matched pair, except for the nodes in the coarsest resolution,
(iv) $\mathbf{m}$ and $\mathbf{n}$ must be topologically consistent. Unlike computing consistency using Reeb graphs in [1], here we consider the MDRGs of the corresponding Reeb spaces. That is, $\mathbf{m}$ and $\mathbf{n}$ should be in the same branches of the respective MDRGs as their matched siblings in all dimensions. In particular, for a bivariate field, to satisfy the topological consistency, a two-dimensional list of labels is maintained corresponding to each node of the JCN. Once two nodes are matched and labeled, their siblings in the same branch of the MDRG get the same labels (in each dimension), as shown in Fig. 7. The label propagation is described in detail in Section 5.2. Note that each coarsest resolution JCN consists of a single node, therefore, $\mathbf{m}$ and $\mathbf{n}$ belong to two such coarsest resolution JCNs are always topologically consistent.

Creating the list MPAIR of matching pairs between two MRSs is outlined in Algorithm 2.

Algorithm 2. CREATEMATCHINGPAIRS

Input: Multi-resolution Reeb spaces $\mathrm{MRS}_{\mathbf{f},L}$, $\mathrm{MRS}_{\mathbf{g},L}$
Output: MPAIR - list of matched pairs
for $k = 0, 1, \ldots, L-1$ **do**
  Create MDRGs corresponding to $\mathrm{JCN}_{\mathbf{f},k}$ and $\mathrm{JCN}_{\mathbf{g},k}$
  Add all the nodes of $\mathrm{JCN}_{\mathbf{f},k}$ to a priority queue $\mathbf{Q}$ where the priority of a node is set as its volume attribute.
  while $\mathbf{Q}$ is not empty **do**
    $\mathbf{m} \leftarrow \mathrm{Pop}(\mathbf{Q})$.
    Search for its best matching pair $\mathbf{n} \in \mathrm{JCN}_{\mathbf{g},k}$, satisfying the *matching rules* (i)-(iv).
    if $\mathbf{n}$ is found **then**
      $\mathrm{MPAIR} \leftarrow \mathrm{Add}(\{(\mathbf{m}, \mathbf{n}), k\})$
      Do label propagation in the MDRGs corresponding to the matched nodes $\mathbf{m}$ and $\mathbf{n}$.
    end if
  end while
end for
return MPAIR

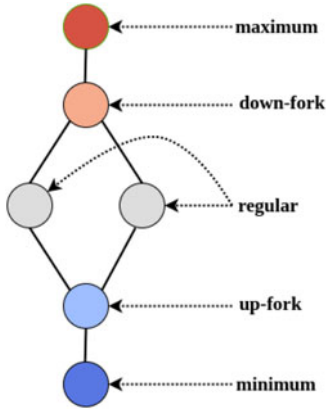


Fig. 5. Classification of nodes in a Reeb graph: maximum, minimum, up-fork, down-fork and regular.

5.2 Label Propagation

Generically, a node in a Reeb graph is classified as a minimum, maximum, up-fork, down-fork or regular, as illustrated in Fig. 5. A node is a minimum or maximum if its down-degree or up-degree is 0, respectively. A node is a down-fork if its down-degree is 2 and up-degree is 1. Similarly, a node is an up-fork if its up-degree is 2 and down-degree is 1. A regular node has both up-degree and down-degree as 1 (see [43]). We note, a branch corresponding to a node $\mathbf{m}$ in the Reeb graph consists of a monotonic increasing path from $\mathbf{m}$ to the upward direction until a maximum or an up-fork is encountered (upward branch), and a monotonic decreasing path from $\mathbf{m}$ to the downward direction until a minimum or a down-fork is reached (downward branch). It should be noted that if $\mathbf{m}$ itself is an up-fork or maximum, then it has no upward branch. Similarly, when $\mathbf{m}$ is a down-fork or minimum it has no downward branch.

Let the nodes $\mathbf{m} \in \mathrm{JCN}_{\mathbf{f},k}$ and $\mathbf{n} \in \mathrm{JCN}_{\mathbf{g},k}$ form a matched pair and let $(f_1(\mathbf{m}), f_2(\mathbf{m}), \ldots, f_r(\mathbf{m}))$ and $(g_1(\mathbf{n}), g_2(\mathbf{n}), \ldots, g_r(\mathbf{n}))$ be the corresponding field values. Now for the topological consistency rule, first we compute the MDRGs $\mathrm{MDRG}_{\mathbf{f},k}$ and $\mathrm{MDRG}_{\mathbf{g},k}$ corresponding to $\mathrm{JCN}_{\mathbf{f},k}$ and $\mathrm{JCN}_{\mathbf{g},k}$ and then do a label propagation in the branches of the MDRGs corresponding to $\mathbf{m}$ and $\mathbf{n}$. An MDRG is an r-dimensional hierarchy of Reeb graphs. In each dimension i of $\mathrm{MDRG}_{\mathbf{f},k}$, there is a Reeb Graph $\mathcal{RG}_i$ containing a unique node, say $\mathbf{m}_i$, corresponding to $\mathbf{m}$ and with field value $f_i(\mathbf{m})$ for $i = 1, 2, \ldots, r$.

In the first step, a label ℓ_1 is assigned to the node $\mathbf{m}_1$ and propagated to all the nodes in its branch in $\mathcal{RG}_1$ (and also to the corresponding JCN nodes). Next, we consider the Reeb graph $\mathcal{RG}_2$ in the second dimension which contains the node $\mathbf{m}_2$. We assign a label ℓ_2 to $\mathbf{m}_2$ and propagate this label to all the nodes in the branch of $\mathbf{m}_2$ in $\mathcal{RG}_2$ (and also to the corresponding JCN nodes). We continue this label propagation for all r dimensions and update the r-dimensional list of labels (maintained for each node of the JCN) for the nodes on the branches of $\mathbf{m}$ in the MDRG. These lists are used for matching the topological consistency of a pair of nodes. The label propagation for $\mathbf{n}$ in $\mathrm{MDRG}_{\mathbf{g},k}$ is done, similarly. Fig. 7 illustrates an example of label propagation for a matched pair in two MDRGs corresponding to two bivariate fields.

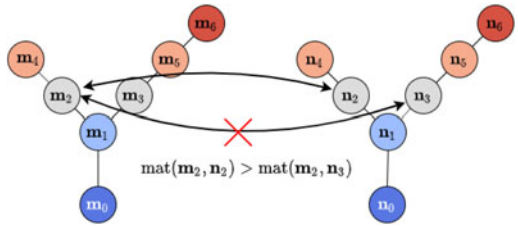


Fig. 6. Computing the unique matching pair using the matching function. We note, here the value of the matching function between $\mathbf{m}_2$ and $\mathbf{n}_2$ is greater then its value between $\mathbf{m}_2$ and $\mathbf{n}_3$.

5.3 Computing the Unique Matching Pair

If there is more than one node in $\mathrm{MRS}_{\mathbf{g},L}$ that can match with $\mathbf{m} \in \mathrm{MRS}_{\mathbf{f},L}$, then a unique node is selected based on their matching value. The node which has the highest matching value is selected. Suppose, a node $\mathbf{m} \in \mathrm{MRS}_{\mathbf{f},L}$ can be matched with nodes $\mathbf{n}_1, \mathbf{n}_2 \in \mathrm{MRS}_{\mathbf{g},L}$. Then the ambiguity is broken by defining a loss function between the nodes and between their adjacent nodes. The higher the loss, the lesser the similarity between the nodes. The loss function between two nodes is defined as

$$\epsilon(\mathcal{A}(\mathbf{m}), \mathcal{A}(\mathbf{n})) = 1 - \widetilde{\varphi}(\mathcal{A}(\mathbf{m}), \mathcal{A}(\mathbf{n})), \tag{1}$$

where $\widetilde{\varphi}$ will be defined in Equation (4). The similarity decreases as the loss increases.

Next, we define the matching function for a pair of nodes using the attributes of their adjacent nodes in the respective MDRGs. Let $\mathrm{adj}(\mathbf{m}, \mathcal{RG}_{f_i})$ denote the collection of adjacent nodes of $\mathbf{m}$ in $\mathcal{RG}_{f_i}$ whose f_i values are in $[s_i, t_i)$ for $i = 1, 2, \ldots, r$. Let $\overline{\mathrm{adj}}(\mathbf{m}, \mathcal{RG}_{f_i})$ denote the sum of the attributes of the nodes in $\mathrm{adj}(\mathbf{m}, \mathcal{RG}_{f_i})$. Then the matching function between the pair of nodes $\mathbf{m}$ and $\mathbf{n}$ is defined as

$$\begin{aligned}\mathrm{mat}(\mathbf{m}, \mathbf{n}) = &- \epsilon(\mathcal{A}(\mathbf{m}), \mathcal{A}(\mathbf{n})) \\ &- \sum_{f_r \in [s_r, t_r)} \epsilon(\overline{\mathrm{adj}}(\mathbf{m}, \mathcal{RG}_{f_r}), \overline{\mathrm{adj}}(\mathbf{n}, \mathcal{RG}_{g_r})) \\ &- \sum_{f_{r-1} \in [s_{r-1}, t_{r-1})} \epsilon(\overline{\mathrm{adj}}(\mathbf{m}, \mathcal{RG}_{f_{r-1}}), \overline{\mathrm{adj}}(\mathbf{n}, \mathcal{RG}_{g_{r-1}})) \\ &- \quad \cdots \\ &- \sum_{f_1 \in [s_1, t_1)} \epsilon(\overline{\mathrm{adj}}(\mathbf{m}, \mathcal{RG}_{f_1}), \overline{\mathrm{adj}}(\mathbf{n}, \mathcal{RG}_{g_1})).\end{aligned} \tag{2}$$

Our method decides a unique matching pair by computing the matching function in Equation (2) for all possible candidates. First, it compares by computing the part $-\epsilon(\mathcal{A}(\mathbf{m}), \mathcal{A}(\mathbf{n}))$ in Equation (2) for all pairs. If that can decide a unique matching pair, it stops. Otherwise, the algorithm incrementally adds the contributions $-\epsilon(\overline{\mathrm{adj}}(\mathbf{m}, \mathcal{RG}_{f_r}), \overline{\mathrm{adj}}(\mathbf{n}, \mathcal{RG}_{g_r}))$ of the adjacent nodes in the Reeb graph $\mathcal{RG}_{f_r}$ by varying the range $[s_r, t_r)$, till all adjacent nodes are added in the corresponding branches. If it obtains a unique matching pair at any intermediate stage, it stops. Otherwise, the algorithm goes to the next level Reeb graphs $\mathcal{RG}_{f_{r-1}}$ and $\mathcal{RG}_{g_{r-1}}$. to add the contribution of the adjacent nodes as before. It continues till it can decide a unique matching pair (see Fig. 6). Generically, this method converges and decides a unique matching pair.

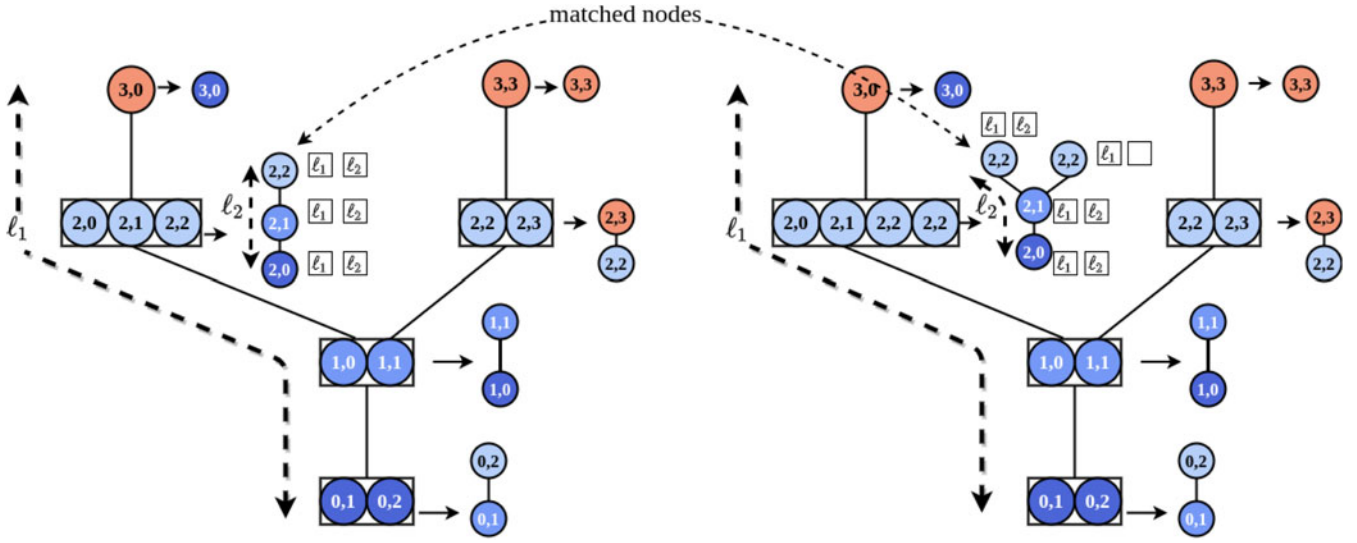


Fig. 7. Label propagation is demonstrated for a matching pair in two MDRGs. For a bivariate field, a two-dimensional list of labels needs to be maintained. The coloring of the nodes is similar to the MDRG in Fig. 2.

5.4 Similarity Calculation

Following Zhang *et al.* [3], first we define a real-valued similarity function φ for each matched pair $(\mathbf{m}, \mathbf{n}) \in \mathcal{M}$ as

$$\varphi(\mathbf{m}, \mathbf{n}) = \omega_1 \widetilde{\varphi}(\mathcal{V}(\mathbf{m}), \mathcal{V}(\mathbf{n})) + \omega_2 \widetilde{\varphi}(\mathcal{R}_0(\mathbf{m}), \mathcal{R}_0(\mathbf{n})) + \omega_3 \widetilde{\varphi}(\mathcal{B}_0(\mathbf{m}), \mathcal{B}_0(\mathbf{n})) + \omega_4 \widetilde{\varphi}(\mathcal{D}(\mathbf{m}), \mathcal{D}(\mathbf{n})), \tag{3}$$

where the weights ω_i satisfy $0 \leq \omega_i \leq 1$ for $i = 1, 2, 3, 4$ and $\sum_{i=1}^{4} \omega_i = 1$. The function $\widetilde{\varphi} : \mathbb{R}^+ \times \mathbb{R}^+ \rightarrow \mathbb{R}^+$ is defined as

$$\widetilde{\varphi}(x_1, x_2) = \frac{\min(x_1, x_2)}{\max(x_1, x_2)}. \tag{4}$$

Thus, we have $0 \leq \varphi(\mathbf{m}, \mathbf{n}) \leq \varphi(\mathbf{m}, \mathbf{m}) = \varphi(\mathbf{n}, \mathbf{n}) = 1$.

Next, we define the similarity function Φ between two kth resolution JCNs $\mathrm{JCN}_{\mathbf{f},k}$ and $\mathrm{JCN}_{\mathbf{g},k}$, by the weighed sum of the similarities for all pairs $(\mathbf{m}_i, \mathbf{n}_i) \in \mathcal{M}$ with $\mathbf{m}_i \in \mathrm{JCN}_{\mathbf{f},k}$ and $\mathbf{n}_i \in \mathrm{JCN}_{\mathbf{g},k}$, as

$$\Phi(\mathrm{JCN}_{\mathbf{f},k}, \mathrm{JCN}_{\mathbf{g},k}) = \sum_{i=1}^{s} \frac{\mathcal{V}(\mathbf{m}_i) + \mathcal{V}(\mathbf{n}_i)}{2} \varphi(\mathbf{m}_i, \mathbf{n}_i), \tag{5}$$

where s is the total number of such pairs. Finally, we define the similarity $\overline{\Phi}$ between two MRSs $\mathrm{MRS}_{\mathbf{f},L}$ and $\mathrm{MRS}_{\mathbf{g},L}$ as

$$\overline{\Phi}(\mathrm{MRS}_{\mathbf{f},L}, \mathrm{MRS}_{\mathbf{g},L}) = \frac{1}{L} \sum_{k=0}^{L-1} \Phi(\mathrm{JCN}_{\mathbf{f},k}, \mathrm{JCN}_{\mathbf{g},k}). \tag{6}$$

We note, $0 \leq \overline{\Phi}(\mathrm{MRS}_{\mathbf{f},L}, \mathrm{MRS}_{\mathbf{g},L}) \leq \overline{\Phi}(\mathrm{MRS}_{\mathbf{f},L}, \mathrm{MRS}_{\mathbf{f},L}) = 1$. Thus we have the following theorem.

Theorem 5.1. *Let $\mathcal{F}_{\mathbb{M}}$ be the space of all multi-resolution Reeb spaces corresponding to different r-dimensional PL multi-fields over a triangulated domain $\mathbb{M}$ and let*

$$\delta(\mathrm{MRS}_{\mathbf{f},L}, \mathrm{MRS}_{\mathbf{g},L}) = 1 - \overline{\Phi}(\mathrm{MRS}_{\mathbf{f},L}, \mathrm{MRS}_{\mathbf{g},L}). \tag{7}$$

If $\omega_1 = 1$ and $\omega_2 = \omega_3 = \omega_4 = 0$ in Equation (3) (i.e., if we use only the volume attribute in the similarity function), then δ is a metric. In other words, in this case $(\mathcal{F}_{\mathbb{M}}, \delta)$ is a metric space.

Proof. See Appendix A, which can be found on the Computer Society Digital Library at http://doi.ieeecomputersociety.org/10.1109/TVCG.2021.3087273. □

Note that the MRS structure, being a special case of MSM, enjoys the stability property under map perturbation, as described by Dey *et al.* [35]. However, to show the stability of MRS structures with respect to the proposed distance measure in Equation (7), i.e., to show small perturbation of a multi-field gives a small variation in the corresponding MRS, requires further theoretical analysis. Some experimental results on the stability of δ is shown in Appendix B, available in the online supplemental material.

6 Computational Complexity

In this section, we give the computational complexity of Algorithms 1 and 2.

Part I: Time Complexity for Computing an MRS. The computational cost of a JCN as dicussed in [11] depends on the number of simplices $|F|$ in the input mesh, the levels of quantization q_i for each component field f_i, the number of functions r and the dimension d of the domain. Using the algorithm in [11], computing a quantized Reeb space at the finest resolution costs $\mathcal{O}(r|E| + |E|\alpha(|E|))$, where $|E|$ is the number of edges in the JCN and α is the inverse Ackermann function [42]. Here, $|E| = \mathcal{O}((2r + d)N_{\mathbf{f}})$ where $N_{\mathbf{f}}$ is the total number of fragments (a fragment is a part of a quantized contour in a simplex). Again, $N_{\mathbf{f}} = \mathcal{O}(Q|F|)$ where $Q = \Pi_{i=1}^{r} q_i$.

Algorithm 1, for computing a coarser JCN from a JCN, depends on the number of fields r, the levels of resolutions $\mathbf{q} = (q_1, \ldots, q_r)$, the number of nodes $|V|$ and the number of edges $|E|$ of the input JCN. The time-complexity for grouping the nodes into $\widetilde{Q} = \frac{Q}{2^r}$ range intervals (denoted by $\widetilde{R}_{i_1 i_2 \ldots i_r}$ in Algorithm 1) is $\mathcal{O}(\widetilde{Q}r|V|)$, since r comparisons of the range values are required for each node to decide in which interval it belongs to.

Next, in line 7 of Algorithm 1, the computational cost for creating the Union-Find structure UF for the set of nodes corresponding to each range interval $\widetilde{R}_{i_1 i_2 \ldots i_r}$ with $|V|$ nodes (in the worst case) is given by $\mathcal{O}(|V| + |E|\alpha(|V|))$ where α is the inverse Ackermann function. The number of components in UF is at most $|V|$. Thus the total time for creating the nodes in the coarser JCN and the child node-ids from the finer JCN costs $\mathcal{O}(|V|)$ and is followed directly from the UF structure (line 9-13, Algorithm 1).

Finally, the time complexity for adding edges to the coarser JCN is $\mathcal{O}(|E|\alpha(|V|))$ since it requires two FIND operations per edge and each find operation takes $\mathcal{O}(\alpha(|V|))$ time (line 18-23, Algorithm 1). Therefore, the total time complexity for creating the coarser resolution Reeb space is given as $\mathcal{O}(\widetilde{Q}r|V| + \widetilde{Q}(|V| + |E|\alpha(|V|)) + |E|\alpha(|V|))$ or $\mathcal{O}(\widetilde{Q}(r|V| + |E|\alpha(|V|)))$.

For constructing a MRS with L levels of resolutions we need to call CreateCoarserReebSpace function $L-1$ times. Therefore, roughly, the time complexity for computing a MRS with L levels of resolutions is $\mathcal{O}(LQ\frac{2^{Lr}-1}{2^{(L-1)r}(2^r-1)}(r|V| + |E|\alpha(|V|)))$.

Part II: Complexity for Computing the Similarity Metric. Algorithm 2, for creating matching pairs, takes two Multi-Resolution Reeb Spaces $\mathrm{MRS}_{\mathbf{f},L}$ and $\mathrm{MRS}_{\mathbf{g},L}$ as input. Let $|V|$ and $|V'|$ be the number of nodes in the finest resolution JCN of $\mathrm{MRS}_{\mathbf{f},L}$ and $\mathrm{MRS}_{\mathbf{g},L}$, respectively. The worst case run time for constructing an MDRG (line 2) from a JCN with $|V|$ nodes, $|E|$ edges and r number of fields, using the algorithm in [32], is $\mathcal{O}(r|V|(|V| + |E|\alpha(|V|) + |V|\log(|V|)))$. The time complexity for creating the priority queue $\mathbf{Q}$ (in line 3) for the JCN with $|V|$ nodes is $\mathcal{O}(|V|\log|V|)$.

In line 6 of Algorithm 2, a matching pair is created based on the four matching rules in Section 5.1 and then added to create the MPAIR list. The first matching rule for creating a pair can be checked in $\mathcal{O}(1)$ time by maintaining a boolean variable corresponding to each node in the JCN which denotes whether the node is already matched or not. The second rule of checking whether the pair of nodes have the same range values takes $\mathcal{O}(r)$ time. The third rule to check if their parent nodes also have matched, requires examining MPAIR in the parent resolution which takes $\mathcal{O}(\min(|V|,|V'|))$ time. To check the fourth rule of topological consistency for a pair of nodes $\mathbf{m} \in \mathrm{JCN}_{\mathbf{f},k}$ and $\mathbf{n} \in \mathrm{JCN}_{\mathbf{g},k}$, we need to check whether $\mathbf{m}$ and $\mathbf{n}$ have the identical r-dimensional lists of labels obtained by label propagation in the respective MDRGs (as discussed in Section 5.2). The worst case time-complexity for checking whether a two such r-dimensional lists are identical is $\mathcal{O}(r\min(|V|,|V'|))$.

Therefore, the time complexity for finding a matching pair for $\mathbf{m}$ is $\mathcal{O}(|V'|(r + \min(|V|,|V'|) + r\min(|V|,|V'|))$ or $\mathcal{O}(r|V'|\min(|V|,|V'|))$. Then the time complexity for label propagation corresponding to a matched pair $(\mathbf{m},\mathbf{n})$ (line 9, Algorithm 2) is $\mathcal{O}(r(|V| + |V'|))$. Finally, the time complexity of the algorithm CreateMatchingPair can be given as $\mathcal{O}(Lr|V||V'|\min(|V|,|V'|) + r|V|(|V| + |E|\alpha(|V|) + |V|\log(|V|)) + r|V'|(|V'| + |E'|\alpha(|V'|) + |V'|\log(|V'|)))$.

7 Implementation and Experimental Results

In this section, we briefly discuss the implementation of our algorithm. Then we show the experimental results of our method on two real datasets.

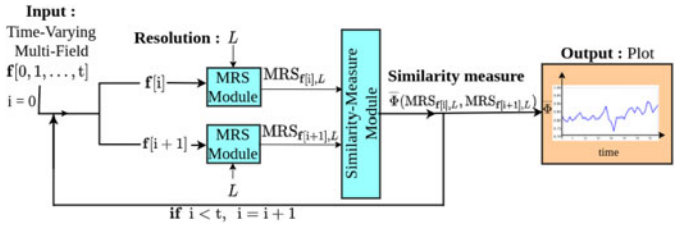


Fig. 8. Implementation pipeline.

7.1 Implementation

Sections 4 and 5 describe the algorithm for constructing the MRS data-structure and the method for computing the similarity measure between two MRSs, respectively. We implement our algorithms in C++, using the Visualization Toolkit (VTK) [44] of version 6.1. The implementation is designed to work for a generic pair of multi-field data. However, for our current experiments it takes a time-varying multi-field data as input and computes the similarity measure between two data at consecutive time stamps. Thus we obtain a plot of the similarity measure over time as output. In our implementation, we use the joint contour net [11] and multi-dimensional Reeb graph [32] as library classes. Moreover, we build the following classes as part of our main implementation.

I. *MRS Module:* This module implements Algorithm 1 to obtain the MRS structure as a series of Reeb spaces for different resolutions (JCNs). The inputs for constructing an MRS are the multi-field and the required levels of resolution. The first part of the implementation is for computing the finest resolution JCN from the input multi-field. In the second part, it computes a series of coarser JCNs from the finest resolution JCN.

II. *Similarity-Measure Module:* This module implements Algorithm 2 for computing the similarity measure between two MRSs as defined in Equation (6). We store the similarity measure between two multi-fields at consecutive time-stamps for creating the final similarity plot given a time-varying multi-field data.

The pipeline of our implementation is outlined in Fig. 8.

Next, we discuss the application of our implementation in two computational problems to validate the method for detecting the topological features in a series of (or time-varying) multi-field data.

7.2 Nuclear Fission: Fermium Atom Dataset

The Fermium-258 (or Fe-258) atom dataset of nuclear fission is used for experimentation (see [12] for a detailed description of the dataset). The aim is to detect the correct scission point using the proposed algorithm. The Fe-258 atom dataset comprises of the symmetric Compact Fission (sCF) data, where the nucleus of the atom splits into almost two equal parts. The sCF data is equivalent to the trajectory in an n-dimensional collective space (defined by the degrees of freedom in the fission) defined in DFT [12]. This trajectory consists of 40 sites and each site consists of 3 scalar fields, i.e., spatial densities of proton $\boldsymbol{p}$, neutron $\boldsymbol{n}$ and nucleon $\boldsymbol{t}$ (proton + neutron) inside the nucleus, defined on a regular $19 \times 19 \times 19$ grid. As discussed in the paper [45], the dataset

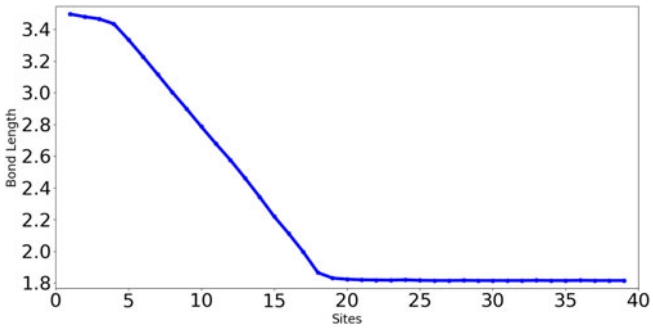


Fig. 9. The plot of the bond length between C and Pt atoms at different sites. It can be observed from the plot that the bond length between the C and Pt atoms stabilizes at site 21.

shows a slope for the energy plot and not a sudden discontinuity, and hence the plot is not able to detect the correct scission point. Duke *et al*. [12] describe a JCN-based visualization method to detect the scission point where they visually identify the topological split by observing each of the JCN structures for all 40 sites, which is a tedious process. We propose to detect the topological feature using our similarity measure plot.

Observation and Results. We compute the similarity plots for the Fe-258 fission dataset using different combinations of component scalar fields (proton: ***p***, neutron: ***n*** and nucleon: ***t*** densities) and with four MRS resolutions ($L = 2, 3, 4, 5$), as shown in Fig. 10. From our experiments, we observe that equal weights of the attributes in, Equation (3), give the best results for most of the cases. Therefore, the similarity values in all our experiments are computed for equal weights ($\omega_1 = \omega_2 = \omega_3 = \omega_4 = 0.25$). In Fig. 10, we observe the most significant dip in each of the plots is at site 260. This is the site where the nucleus splits into two fragments, as observed by Duke *et al.* [12]. For the resolution $L = 2$ the plots show only one sharp dip, at the site of the nuclear scission. As the MRS resolution increases, we observe other dips in the similarity plots. These dips correspond to the other topological features in the data. However, the most significant dip persists in all the resolutions. This illustrates the effectiveness of a multi-resolution method. In this particular dataset, we observe that the scission point can be captured using only the proton density ***p*** (similarly for neutron density ***n*** or nucleon density ***t***) as shown in the first row of Fig. 10. However, in our next experiment we see that the individual scalar fields are not sufficient to capture a topological feature and the necessity of a multi-field based technique.

7.3 Adsorption: Pt-CO Stable Bond Formation

Adsorption is a process in which gas molecules (adsorbate) gets adsorbed on a metal surface (adsorbent). This surface phenomenon is of interest to scientists because of its application in various fields such as molecular electronics, corrosion,

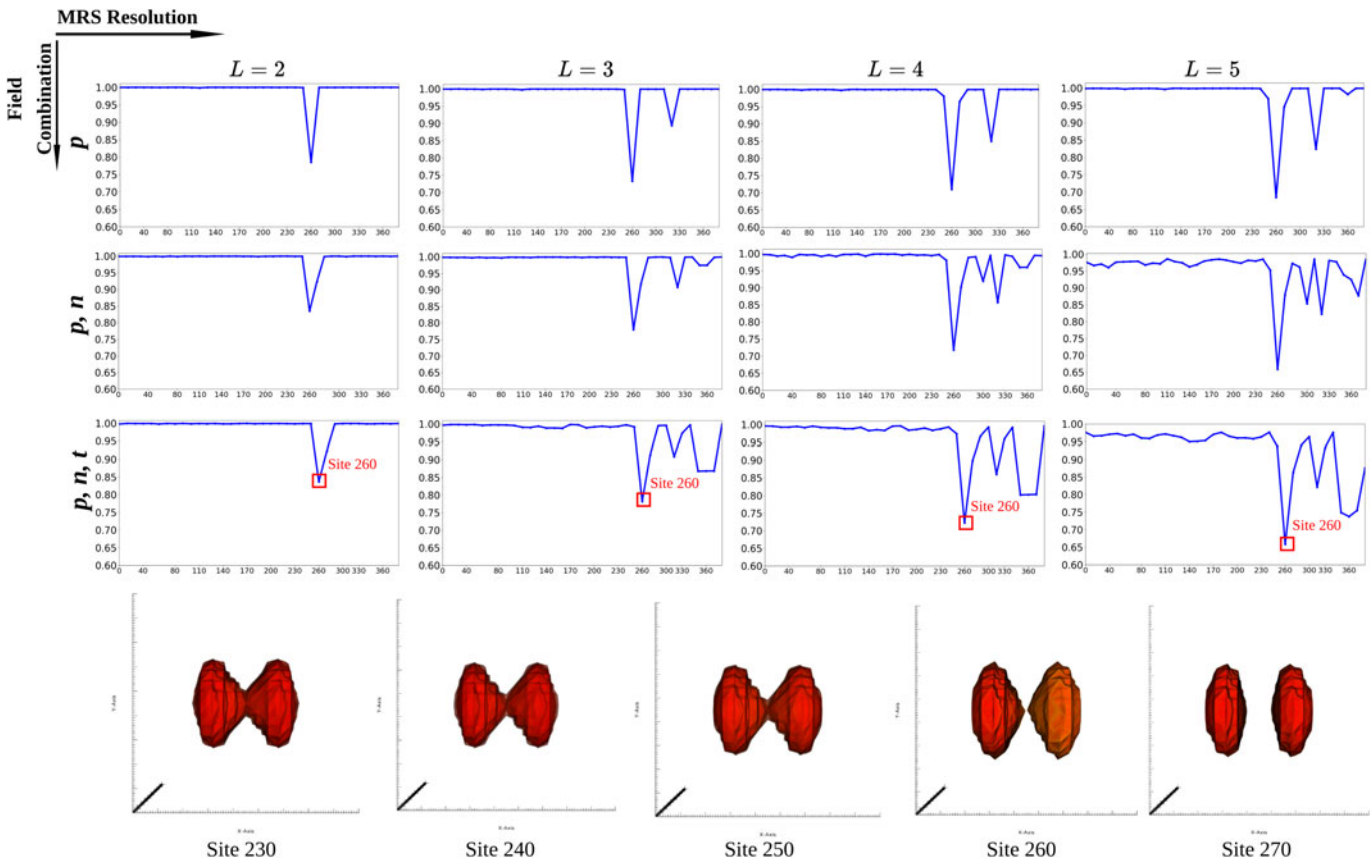


Fig. 10. Plots for Fe-258 data. *Top-three rows:* Similarity plots for time-varying Fermium atom data. Each row shows the plots using 4 different MRS resolutions ($L = 2, 3, 4, 5$). Each column shows the plots for 3 different field combinations. First row shows the plots for the scalar field (***p***), second row shows the plots for the bivariate field (***p, n***), and the third row shows the plots for the tri-variate field (***p, n, t***). It can be observed from the plots that the most significant dip is captured at site 260. *Bottom-row:* The nuclear scission is visualized using fibers at Sites: $230 - 270$ in which the split happens at Site 260. This site corresponds to the 'lowest' dip in the similarity plots.

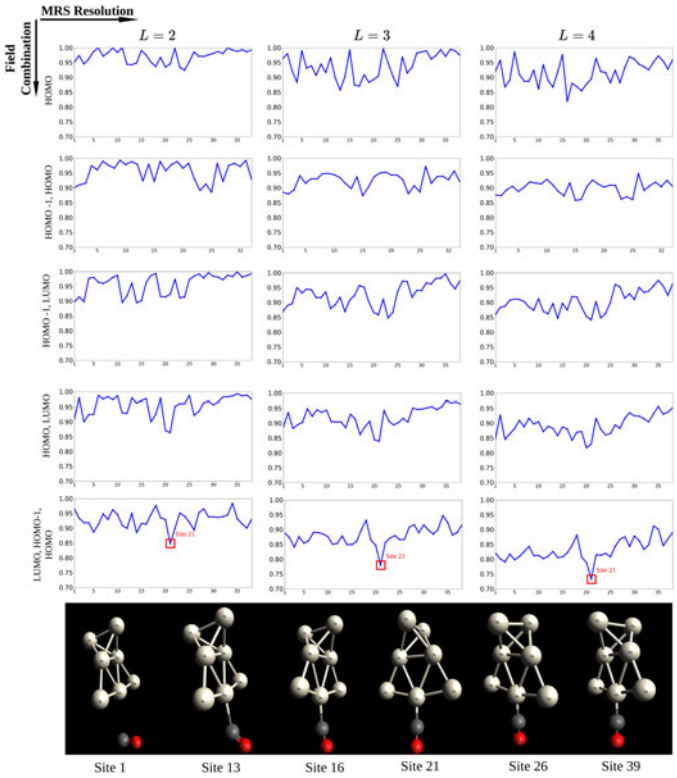


Fig. 11. Plots for Pt-CO data. Columns correspond to the plots for 3 different MRS resolutions ($L = 2, 3, 4$). First five rows correspond to the plots for 5 different field combinations. The first row shows the plots for the field (HOMO); the second and third rows show the plots for the fields (HOMO-1, HOMO) and (HOMO-1, LUMO) respectively; the fourth row shows the plots for the fields (HOMO, LUMO) and the fifth row shows the plots for the fields (LUMO, HOMO-1, HOMO). The last row shows the geometry of the Pt-CO bond formation at site 13. However, the plots of the fifth row show the most significant dip at site 21, which corresponds to the stable bond formation between Pt and C atoms.

electrochemistry, and heterogeneous catalysis [46], [47]. The adsorption of Carbon Monoxide (CO) molecule on the Platinum (Pt) surface is particularly of interest because of its industrial applications in automobile emission, fuel cells and other catalytic processes [48], [49]. Therefore, an atomic level understanding of the interaction of the CO molecule with the platinum surface is of utmost importance.

To validate our method, we study the interaction of Pt_7 cluster with CO molecule (Fig. 11, last row). As discussed in [48], a Pt-CO bond is formed when the CO molecule gets adsorbed on the Pt_7 cluster and the internal bond of CO molecule becomes weaker. A *stable bond* is formed between the C atom of CO molecule and one of the Pt atoms of Pt_7 when the adsorption process is complete.The data shows that the bond between the C and Pt atoms becomes stable when the bond length between the atoms is 1.84Å (see the plot in Fig. 9). In the current experiment, we aim to detect the correct site of stable bond formation between Pt_7 and CO molecules.

Our present Pt-CO dataset consists of electron density distributions of HOMO (Highest Occupied Molecular Orbital), LUMO (Lowest Unoccupied Molecular Orbital) and HOMO-1. Orbital numbers 69, 70, 71 correspond to HOMO-1, HOMO, LUMO respectively. We consider a series of 40 datasets, generated by quantum mechanical computations for varying distances between the carbon atom of the CO molecule and the platinum surface [36].

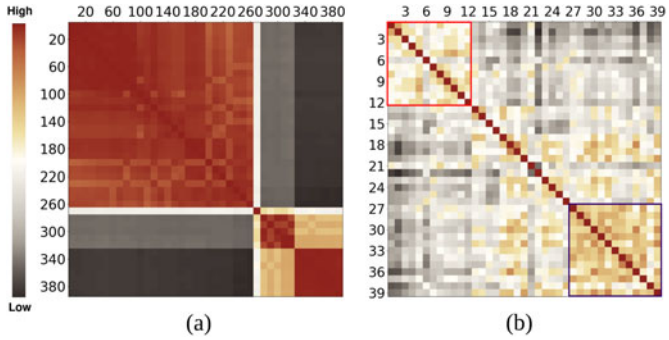


Fig. 12. Similarity matrices. (a) Similarity matrix between all sites of the Fe-258 atom data (using fields p, n, t) with MRS resolution $L = 3$. Two clusters can be observed by the sites before and after the nuclear scission point (site 260), respectively. (b) Similarity matrix of the Pt-CO bond detection data (using orbital densities of LUMO, HOMO-1, HOMO) with MRS resolution $L = 3$. Two clusters are indicated by the red and violet boxes. Site 21 of stable bond formation can be observed as an outlier.

Observation and Results. We experiment with the Pt-CO orbital data for all possible combinations of HOMO, LUMO and HOMO-1 with three different resolutions of the MRSs ($L = 2, 3, 4$). As before, we consider equal weights of the attributes, i.e., $\omega_1 = \omega_2 = \omega_3 = \omega_4 = 0.25$. Fig. 11 shows the similarity measure plots for different combinations of scalar fields with three MRS resolutions. We observe that the similarity computed using a scalar field (HOMO) is unable to detect the correct site where the stable bond between Pt and C is formed, even with MRSs of increasing resolutions (first row of Fig. 11). We then tested with different bivariate data consisting of two of the orbital density data HOMO, LUMO and HOMO-1. We observe, the combinations (HOMO-1, HOMO) and (HOMO-1, LUMO) are again not able to show the correct site in the plots (see the second and third rows of Fig. 11). However, when we experiment the (HOMO, LUMO) bivariate field, a significant dip is observed at site 21 in the plot (fourth row of Fig. 11) for all the resolutions. Agarwal *et al.* [36] identified the same site of Pt-CO bond formation using this bivariate field. Finally, we experiment with the tri-variate field (HOMO, LUMO, HOMO-1) consists of all the scalar fields. We observe that the similarity plots for all the resolutions ($L = 2, 3, 4$) are able to detect the correct site (site 21), where a stable Pt-CO bond is formed (fifth row of Fig. 11).

7.4 Similarity Matrix

In addition to plotting the similarity measures between consecutive sites of the time-varying data, we also visualize a two-dimensional similarity matrix by computing the similarity measures between every pair of sites. Fig. 12a shows the similarity matrix for the Fe-258 atom data.The similarity measure, at each matrix entry, is between the corresponding MRSs where each MRS is computed using the fields p, n, t and resolution $L = 3$. The similarity matrix for the Fe-258 atom data indicates the presence of two clusters, consisting of the sites before and after the site of nuclear scission (site 260) respectively. We also note that the sites 280-330 and 330–390 are sub-clusters of the cluster consisting of the sites $280-390$. This indicates the detection of additional features with the increased MRS resolution ($L = 3$). Fig. 12b shows the similarity matrix of the Pt-CO bond detection data

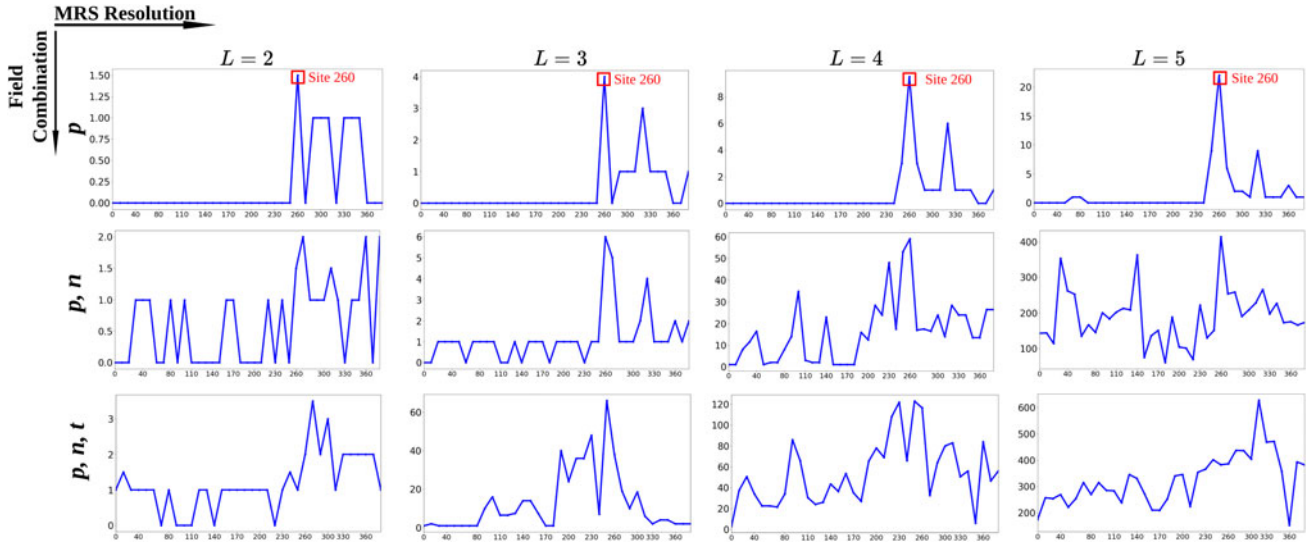


Fig. 13. Bottleneck distance (between MRSs) plots for Fe-258 data. Each row shows the plots using 4 different MRS resolutions ($L = 2, 3, 4, 5$). Each column shows the plots for 3 different field combinations. First row shows the plots for the scalar field (***p***), second row shows the plots for the bivariate field (***p, n***), and the third row shows the plots for the tri-variate field (***p, n, t***). The plots in the first row shows the most significant peak at site 260, which corresponds to the site of nuclear scission.

(using LUMO, HOMO-1, HOMO) with MRS resolution $L = 3$. In this similarity matrix, we note that the sites 1-12 form a cluster, which is indicated by the red box in Fig. 12b. These are the sites before the formation of the Pt-CO bond (site 13). We also see another cluster consisting of the sites 27-39 as highlighted by the violet box in Fig. 12b. Apart from the clusters, it can be seen that the site where the bond length stabilizes (site 21) is an outlier as most of the other sites have a low similarity value with the site 21.

7.5 Comparison With Bottleneck Distance Between MSMs

As we discussed, an MRS structure can be obtained by a particular consideration of the tower of covers in the MSM. The sequence of JCNs (from finer to coarser resolutions) in an MRS forms a tower of 1-dimensional simplicial complexes that are connected by simplicial maps. These connecting simplicial maps are obtained by the parent-child relationship of the nodes of JCNs at the consecutive resolutions (see Section 4 for more details). An algorithm for computing the persistence diagram corresponding to the tower of simplicial complexes connected by the simplicial maps was proposed by Dey *et al.* [35], [50]. The bottleneck distance between two such persistence diagrams, corresponding to two MRSs, can then be computed. We use the Sophia [51] and GUDHI [52] libraries, respectively, to compute the 0-dimensional persistence diagrams for simplicial maps and their bottleneck distance.

The bottleneck distance plots for Fe-258 data for various field combinations and with four MRS resolutions ($L = 2, 3, 4, 5$) is shown in Fig. 13. We observe that the site of nuclear scission is detected using the proton density scalar field (***p***) (see the first row of Fig. 13). However, the results are not consistent using other multi-field combinations. In contrast to this, the proposed technique detects the nuclear scission site using all the three field combinations (See Fig. 10). The bottleneck distance plots of Pt-CO bond detection data for various combination of fields and with three different MRS resolutions ($L = 2, 3, 4$) is shown in Fig. 14. We observe that the site of bond stabilization is captured clearly using the multi-field data of (LUMO, HOMO-1, HOMO) orbitals. However, using only one or two fields the results are not consistent, which is similar to the results obtained using our method. This matches with the results obtained using the proposed similarity measure. In Appendix C, available in the online supplemental material, we show the results of the bottleneck distance plots using the persistence diagrams of different dimensions for the component scalar fields of Fe-258 and Pt-CO data.

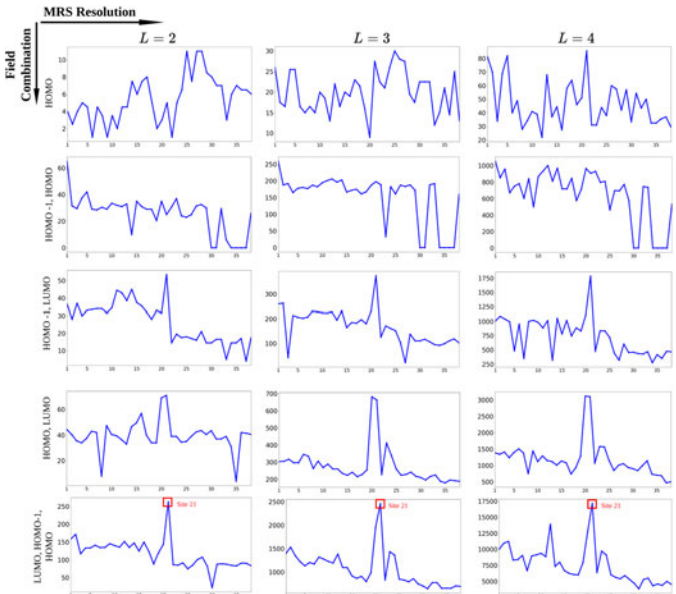
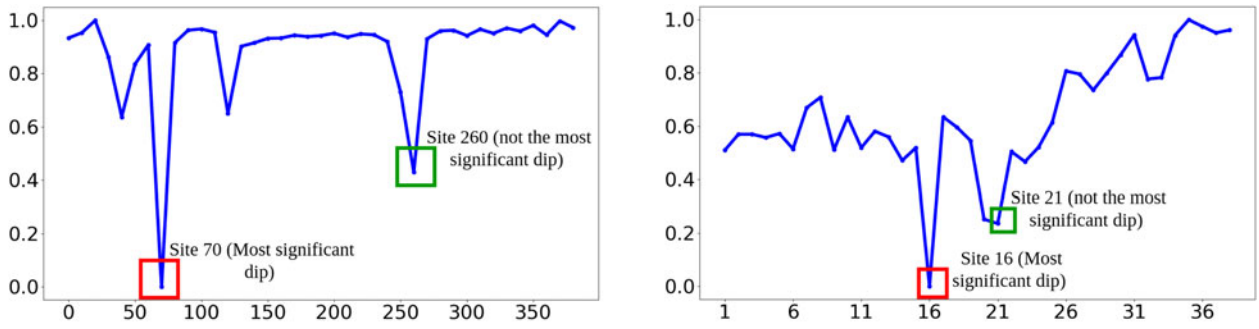


Fig. 14. Bottleneck distance (between MRSs) plots for Pt-CO data. Columns correspond to the plots for 3 different MRS resolutions ($L = 2, 3, 4$). The rows correspond to the plots for 5 different field combinations. The first row show the plots for the field (HOMO); the second and third rows show the plots for the fields (HOMO-1, HOMO) and (HOMO-1, LUMO) respectively; the fourth row shows the plots for the fields (HOMO, LUMO) and the fifth row shows the plots for the fields (LUMO, HOMO-1, HOMO). The plots of the third, fourth and fifth rows show the most significant peak at site 21, which corresponds to the stable bond formation between Pt and C atoms.



(a) Fe-258 fission data plot for ψ_{NRMS} (b) Pt-CO bond data plot for ψ_{NRMS}

Fig. 15. Root mean square based similarity plots: (a) Plot for Fe-258 atom data (p, n, t) shows the significant dip is at site 70, while the dip is expected at site 260. (b) Plot for Pt-CO bond detection data (HOMO, LUMO, HOMO-1) shows that the significant dip is obtained at site 16, while it is expected at site 21. Thus the plots using ψ_{NRMS} are unable to reveal the correct sites as most significant dips.

TABLE 1
Performance Results for Fe-258 and Pt-CO Datasets

Dataset	Sites	Dimension	L	$\lvert V\rvert, \lvert V'\rvert$	P, P' (in %)	$P_{\mathcal{V}}, P'_{\mathcal{V}}$ (in %)	T_{MRS}	$T_{\overline{\Phi}}$
Fe-258: $\boldsymbol{p}$	$250, 260$	$19 \times 19 \times 19$	5	$16, 18$	$90.3, 82.4$	$97.3, 97.4$	0.0013s	0.0041s
Fe-258: $\boldsymbol{p, n}$	$250, 260$	$19 \times 19 \times 19$	5	$428, 433$	$75.8, 71.2$	$96.3, 95.5$	0.0632s	0.1262s
Fe-258: $\boldsymbol{p, n, t}$	$250, 260$	$19 \times 19 \times 19$	5	$773, 748$	$65.3, 62.6$	$95.0, 94.2$	0.2478s	0.3620s
Pt-CO: HOMO-1	$20, 21$	$41 \times 41 \times 41$	4	$88, 84$	$66.0, 69.9$	$99.9, 99.8$	0.0069s	0.0086s
Pt-CO: HOMO-1, HOMO	$20, 21$	$41 \times 41 \times 41$	4	$1325, 1582$	$34.3, 28.9$	$98.5, 97.0$	1.1872s	0.2365s
HOMO-1, HOMO, LUMO	$20, 21$	$41 \times 41 \times 41$	4	$12937, 16339$	$4.9, 3.9$	$89.8, 85.0$	107.3628s	3.5520s

Here, Sites: Data sites or time-steps compared, Dimension: Data dimension in the volumetric domain of the component fields; L: MRS resolution; $|V|, |V'|$: Number of nodes in the finest resolution JCNs; P, P' : Percentage of nodes matched in each of the MRSs; $P_{\mathcal{V}}, P'_{\mathcal{V}}$: Percentage of data-volume matched in each of the MRSs; T_{MRS}: Time (in seconds) for constructing the corresponding MRSs and $T_{\overline{\Phi}}$: Time (in seconds) for computing the similarity between the MRSs.

7.6 Comparison With Root Mean Square Metric

To show the importance of the proposed topology-based similarity, we compute a similarity plot using the root mean square (RMS) distance metric. The RMS distance measure between two multi-fields $\mathbf{f}[i] = (f_1^{(i)}, f_2^{(i)}, \ldots, f_r^{(i)})$ and $\mathbf{f}[i+1] = (f_1^{(i+1)}, f_2^{(i+1)}, \ldots, f_r^{(i+1)})$ on a domain with m grid points is given as

$$d_{\text{RMS}}[i] = \sqrt{\frac{1}{m}\sum_{j=1}^{m}\sum_{k=1}^{r}\left(f_k^{(i)}(x_j) - f_k^{(i+1)}(x_j)\right)^2}, \tag{8}$$

for $t+1$ time-steps: $i = 0, 1, \ldots, t$. From this we obtain a normalized similarity measure as

$$\psi_{\text{NRMS}}[i] = 1 - \left\{\frac{d_{\text{RMS}}[i] - d_{\text{RMS}}^{\min}}{d_{\text{RMS}}^{\max} - d_{\text{RMS}}^{\min}}\right\}, \tag{9}$$

where, $d_{\text{RMS}}^{\max}$ and $d_{\text{RMS}}^{\min}$ are the maximum and the minimum of the set of values $\{d_{\text{RMS}}[i] : i = 0, 1, \ldots, t\}$, respectively.

Fig. 15 shows the RMS similarity measure plots, respectively for Fe-258 and Pt-CO bond data. We observe that for both the datasets, the RMS similarity measure plots are unable to detect the correct site of topological change (i.e., the nuclear scission or the Pt-CO stable bond formation).

7.7 Performance Results

Table 1 shows the performance results for computing the similarity between two multi-field data in Fe-258 and Pt-CO bond formation. All the experiments are performed on a Ubuntu 16.04 with 16GB memory and 20 core processor system. From the table, we observe that the time required for the construction of MRSs and the similarity calculation with different MRS resolutions.

8 Conclusion

In this article, we propose a novel Reeb space based method for measuring the topological similarity between two multi-fields. To capture the topology of a multi-field, we develop a multi-resolution Reeb space data-structure which converges to the actual Reeb space of the multi-field as the levels of resolution goes to infinity. The proposed technique has the following advantages:

- The proposed multi-resolution Reeb space is an aggregate of all the lower resolution JCNs and it captures the significant topological features that appear in all the lower resolution JCNs. Therefore, the multi-resolution Reeb space is a better data-structure for capturing the prominent topological features in a multi-field data compare to a JCN with a particular resolution.
- The proposed multi-field topology based technique is able to compute the similarity measure between two datasets more precisely than using a scalar topology. Therefore, topological dissimilarity using multiple fields can be detected compared to using one of the

fields, as illustrated in the experiment using Pt-CO dataset.
- In particular, the proposed measure between two MRSs satisfies the metric property, whereas the bottleneck distance between two MSMs provides a pseudo-metric.

However, the proposed method has its scope of improvement in the following points:

- The required levels of resolution of an MRS to detect the topological feature is dependent on the input data. Hence, a technique which depends only on the data and not on the levels of resolution needs to be developed.
- In the proposed technique, the approximation of a Reeb space is given by the MRS, which is a series of JCNs. The computational complexity of the JCN algorithm is high. Therefore, there is a need for a faster algorithm for computing an approximation of the Reeb space.
- Further research is required to obtain a theoretical proof of stability of the MRS structure with respect to the proposed distance measure. Some experimental results, in Appendix B, available in the online supplemental material, show that the proposed distance measure is upper bounded by the Root Mean Square distance.

The proposed method is able to detect the nuclear scission point in the Fe-258 atom and capture the stable bond formation in the Pt-CO data, which are applications in computational physics and computational chemistry, respectively. More experiments with data from other computational research domains, including climate data, for analyzing the effectiveness of the proposed method are required.

Acknowledgments

Amit Chattopadhyay would like to thank Prof. Osamu Saeki for his feedback on the initial draft of the article. The authors would like to thank the anonymous reviewers for their valuable suggestions and feedback to improve the article. This work was supported in part by the Science and Engineering Research Board (SERB), India, under Grant SERB/CRG/2018/000702 and in part by the International Institute of Information Technology (IIITB), Bangalore.

References

[1] M. Hilaga, Y. Shinagawa, T. Kohmura, and T. L. Kunii, "Topology matching for fully automatic similarity estimation of 3D shapes," in *Proc. Annu. Conf. Comput. Graph. Interactive Techn.*, 2001, pp. 203–212.

[2] D. Cohen-Steiner , H. Edelsbrunner, and J. Harer, "Stability of persistence diagrams," *Discret. Comput. Geom.*, vol. 37, no. 1, pp. 103–120, 2007.

[3] X. Zhang, C. L. Bajaj, and N. Baker, "Fast matching of volumetric functions using multi-resolution dual contour trees," *Texas Inst. Comput. Applied Math*, Tech. Rep. 2004.

[4] D. M. Thomas and V. Natarajan, "Multiscale symmetry detection in scalar fields by clustering contours," *IEEE Trans. Vis. Comput. Graph.*, vol. 20, no. 12, pp. 2427–2436, Dec. 2014.

[5] H. Saikia, H.-P. Seidel, and T. Weinkauf, "Extended branch decomposition graphs: Structural comparison of scalar data," *Comput. Graph. Forum*, vol. 33, no. 3, pp. 41–50, 2014.

[6] H. Edelsbrunner and J. Harer, "Jacobi sets of multiple morse functions," in *Foundations of Computational Matematics, Minneapolis*. Cambridge, U.K.: Cambridge Univ. Press, 2004.

[7] O. Saeki, *Topology of Singular Fibers of Differentiable Maps* . Berlin, Germany: Springer Heidelberg, 2004.

[8] H. Edelsbrunner, J. Harer, and A. K. Patel, "Reeb spaces of piecewise linear mappings," in *Proc. Annu. Symp. Comput. Geometry*, 2008, p. 242–250.

[9] G. E. Carlsson, G. Singh, and A. Zomorodian, "Computing multidimensional persistence," *J. Comput. Geom.*, vol. 1, no. 1, pp. 72–100, 2010.

[10] G. Singh, F. Memoli, and G. Carlsson, "Topological methods for the analysis of high dimensional data sets and 3d object recognition," in *Proc. Eurographics Symp. Point-Based Graph.*, M. Botsch, R. Pajarola, B. Chen, and M. Zwicker, Eds., 2007, pp. 91–100.

[11] H. Carr and D. Duke, "Joint contour nets," *IEEE Trans. Vis. Comput. Graph.*, vol. 20, no. 8, pp. 1100–1113, Aug. 2014.

[12] D. Duke, H. Carr, A. Knoll, N. Schunck, H. Nam, and A. Staszczak, "Visualizing nuclear scission through a multifield extension of topological analysis," *IEEE Trans. Vis. Comput. Graph.*, vol. 18, no. 14, pp. 2033–2040, Dec. 2012.

[13] S. Biasotti *et al.*, "Describing shapes by geometrical-topological properties of real functions," *ACM Comput. Surv.*, vol. 40, no. 4, pp. 1–87, Oct. 2008.

[14] S. Biasotti, S. Marini, M. Mortara, G. Patanè, M. Spagnuolo, and B. Falcidieno, "3D shape matching through topological structures," in *Discrete Geometry for Computer Imagery*, I. Nyström, G. Sanniti di Baja, and S. Svensson, Eds. Berlin, Germany: Springer, 2003, pp. 194–203.

[15] S. Biasotti, S. Marini, M. Spagnuolo, and B. Falcidieno, "Sub-part correspondence by structural descriptors of 3D shapes," *Comput.-Aided Des.*, vol. 38, no. 9, pp. 1002–1019, 2006.

[16] T. Tung and F. Schmitt, "Augmented reeb graphs for content-based retrieval of 3D mesh models," in *Proc. Shape Model. Appl.*, 2004, pp. 157–166.

[17] T. Tung and F. Schmitt, "The augmented multiresolution reeb graph approach for content-based retrieval of 3d shapes," *Int. J. Shape Model.*, vol. 11, pp. 91–120, 2005.

[18] J. Tierny, J.-P. Vandeborre, and M. Daoudi, "Partial 3D shape retrieval by reeb pattern unfolding," *Comput. Graph. Forum*, vol. 28, no. 1, pp. 41–55, 2009.

[19] H. Edelsbrunner, D. Letscher, and A. Zomorodian, "Topological persistence and simplification," *Discret. Comput. Geom.*, vol. 28, no. 4, pp. 511–533, 2002.

[20] T. K. Dey, D. Shi, and Y. Wang, "Comparing graphs via persistence distortion," 2015, *arXiv:1503.07414*.

[21] P. K. Agarwal, K. Fox, A. Nath, A. Sidiropoulos, and Y. Wang, "Computing the Gromov-Hausdorff distance for metric trees," *ACM Trans. Algorithms*, vol. 14, no. 2, pp. 1–20, Apr. 2018.

[22] D. Burago, Y. Burago, and S. Ivanov, "A course in metric geometry," *Graduate Studies in Math.* Providence, RI, USA: American Mathematical Society, 2001.

[23] G. Carlsson and A. Zomorodian, "The theory of multidimensional persistence," *Discrete Comput. Geometry*, vol. 42, pp. 71–93, 2007.

[24] V. Narayanan, D. M. Thomas, and V. Natarajan, "Distance between Extremum Graphs," in *Proc. IEEE Pacific Vis. Symp.*, 2015, pp. 263–270.

[25] K. Beketayev, D. Yeliussizov, D. Morozov, G. H. Weber, and B. Hamann, "Measuring the distance between merge trees," in *Proc. Topol. Methods Data Anal. Vis. III*, 2014, pp. 151–165.

[26] H. Saikia, H.-P. Seidel, and T. Weinkauf, "Fast similarity search in scalar fields using merging histograms," in *Proc. Topol. Methods Data Anal. Vis. IV : Theory, Algorithms, Appl.*, 2015, pp. 121–134.

[27] X. Gao, B. Xiao, D. Tao, and X. Li, "A survey of graph edit distance," *Pattern Anal. Appl.*, vol. 13, no. 1, pp. 113–129, 2010.

[28] B. Di Fabio and C. Landi, "The edit distance for reeb graphs of surfaces," *Discrete Comput. Geometry*, vol. 55, no. 2, pp. 423–461, 2016.

[29] R. Sridharamurthy, T. B. Masood, A. Kamakshidasan, and V. Natarajan, "Edit distance between merge trees," *IEEE Trans. Vis. Comput. Graph.*, vol. 26, no. 3, pp. 1518–1531, Mar. 2020.

[30] U. Bauer, X. Ge, and Y. Wang, "Measuring distance between reeb graphs," in *Proc. Annu. Symp. Comput. Geometry*, 2014, p. 464–473.

[31] V. De Silva , E. Munch, and A. Patel, "Categorified reeb graphs," *Discrete Comput. Geometry*, vol. 55, no. 4, p. 854–906, Jun. 2016.

[32] A. Chattopadhyay, H. Carr, D. Duke, Z. Geng, and O. Saeki, "Multivariate topology simplification," *Comput. Geometry, Theory Appl.*, vol. 58, pp. 1–24, Oct. 2016.
[33] U. Bauer, E. Munch, and Y. Wang, "Strong equivalence of the interleaving and functional distortion metrics for reeb graphs," 2015, *arXiv:1412.6646*.
[34] M. Carriere and S. Y. Oudot, "Local equivalence and intrinsic metrics between reeb graphs," in *Proc. Int. Symp. Comput. Geometry*, Jul. 2017, pp. 25:1–25:15.
[35] T. K. Dey, F. Mémoli, and Y. Wang, "Multiscale mapper: Topological summarization via codomain covers," in *Proc. Annu. ACM-SIAM Symp. Discrete Algorithms*, 2016, pp. 997–1013.
[36] T. Agarwal, A. Chattopadhyay, and V. Natarajan, "Topological feature search in time-varying multifield data," 2019, *arXiv: 1911.00687*.
[37] O. Saeki *et al.*, "The impact of applications on mathematics," *in Visualizing Multivariate Data Using Singularity Theory*. Tokyo, Japan: Springer Japan, 2014.
[38] O. Saeki, "Theory of singular fibers and reeb spaces for visualization," in *Topological Methods in Data Analysis and Vis. IV*, H. Carr, C. Garth, and T. Weinkauf, Eds. Cham, Switzerland: Springer, 2017.
[39] K. Cole-McLaughlin , H. Edelsbrunner, J. Harer, V. Natarajan, and V. Pascucci, "Loops in reeb graphs of 2-Manifolds," in *Proc. Annu. Symp. Comput. Geometry*, 2003, pp. 344–350.
[40] H. Carr, J. Snoeyink, and U. Axen, "Computing contour trees in all dimensions," *Comput. Geom. Theory Appl.*, vol. 24, no. 2, p. 75–94, Feb. 2003.
[41] A. Chattopadhyay, H. Carr, D. Duke, and Z. Geng, "Extracting Jacobi structures in reeb spaces," in *Proc. EuroVis - Short Papers*, N. Elmqvist, M. Hlawitschka, and J. Kennedy, Eds., 2014, pp. 1–4.
[42] R. E. Tarjan, "Efficiency of a good but not linear set union algorithm," *J. ACM*, vol. 22, no. 2, p. 215–225, Apr. 1975.
[43] J. Tu, M. Hajij, and P. Rosen, "Propagate and pair: A single-pass approach to critical point pairing in reeb graphs," in *Proc. Adv. Vis. Comput.*, G. Bebis, R. Boyle, B. Parvin, D. Koracin, D. Ushizima, S. Chai, S. Sueda, X. Lin, A. Lu, D. Thalmann, C. Wang, and P. Xu, Eds., 2019, pp. 99–113.
[44] W. Schroeder, H. Martin, and B. Lorensen, "The visualization toolkit: an object-oriented approach to 3d graphics," Kitware, 2004.
[45] A. Staszczak, J. Dobaczewski, and W. Nazarewicz, "Bimodal fission in the skyrme-hartree-fock approach," 2006, *arXiv:nucl-th/ 0612017*.
[46] I. Kendrick, D. Kumari, A. Yakaboski, N. Dimakis, and E. S. Smotkin, "Elucidating the ionomer-electrified metal interface," *J. Amer. Chem. Soc.*, vol. 132, no. 49, pp. 17 611–17 616, 2010.
[47] G. A. Somorjai and Y. Li, *Introduction to Surface Chemistry and Catalysis*. Hoboken, NJ, USA: Wiley, 2010.
[48] N. Dimakis, M. Cowan, G. Hanson, and E. S. Smotkin, "Attraction- repulsion mechanism for carbon monoxide adsorption on platinum and platinum- ruthenium alloys," *J. Phys. Chem. C*, vol. 113, no. 43, pp. 18 730–18 739, 2009.
[49] A. Patra, "Surface properties, adsorption, and phase transitions with a dispersion-corrected density functional," Ph.D. dissertation, Doctor Philosophy Phys., Temple University, Philadelphia, PA, USA, 2018.
[50] T. K. Dey, F. Fan, and Y. Wang, "Computing topological persistence for simplicial maps," in *Proc. Annu. Symp. Comput. Geometry*, 2014, pp. 345–354.
[51] M. Kerber and H. Schreiber. "Barcodes of towers and a streaming algorithm for persistent homology," *Discrete Comput. Geom*, vol. 61, pp. 852–879, 2019, doi: 10.1007/s00454-018-0030-0.
[52] The GUDHI Project, "GUDHI User and Reference Manual," 2020. [Online]. Available: https://gudhi.inria.fr/doc/3.4.0/

Yashwanth Ramamurthi received the MSc degree in theoretical computer science from the PSG College of Technology, Coimbatore. He is currently working toward the PhD degree with the International Institute of Information Technology, Bangalore. His research interests include computational topology, certified geometric computation, and topological data analysis.

Tripti Agarwal received the MSR degree in computer science from the International Institute of Information Technology, Bangalore, and the BTech degree in computer science and engineering from the College of Engineering Roorkee. She is currently working toward the PhD degree with SCI, University of Utah. Her research interests include topological data analysis, scientific visualization, and machine learning.

Amit Chattopadhyay (Member, IEEE) received the MTech degree in computer science from Indian Statistical Institute, Kolkata, the MSc degree in mathematics from the Indian Institute of Technology, Kharagpur, and the PhD degree in 2011 from the Johann Bernoulli Institute of Mathematics and Computer Science, University of Groningen, The Netherlands. He is currently a faculty with the International Institute of Information Technology, Bangalore. His research interests include computational topology and data analysis, certified geometric computation, scientific visualization, optimization on matrix manifolds, and level set method.