Pseudo-Multidimensional Persistence and Its Applications

Catalina Betancourt, Mathieu Chalifour, Rachel Neville, Matthew Pietrosanu, Mimi Tsuruga, Isabel Darcy, and Giseon Heo

Abstract While one-dimensional persistent homology can be an effective way to discriminate data, it has limitations. Multidimensional persistent homology is a technique amenable to data naturally described by more than a single parameter, and is able to encoding more robust information about the structure of the data. However, as indicated by Carlsson and Zomorodian (Discrete Comput Geom 42(1):71–271, 2009), no perfect higher-dimensional analogue of the one-dimensional persistence barcode exists for higher-dimensional filtrations. Xia and Wei (J Comput Chem 36:1502–1520, 2015) propose computing one-dimensional Betti number functions at various values of a second parameter and stacking these functions for each homological dimension. The aim of this visualization is to increase the discriminatory power of current one-dimensional persistence techniques, especially for datasets that have features more readily captured by a combination of two parameters. We apply this practical approach to three datasets, relating to (1) craniofacial shape and (2) Lissajous knots, both using parameters for scale and curvature; and (3) the Kuramoto–Sivashinsky partial differential equation, using parameters for both scale and time. This new approach is able to differentiate between topologically

C. Betancourt · I. Darcy
Department of Mathematics, The University of Iowa, Iowa City, IA, USA
e-mail: catalina-betancourt@uiowa.edu

M. Chalifour · M. Pietrosanu
Department of Mathematical & Statistical Sciences, University of Alberta, Edmonton, Alberta, Canada
e-mail: mrc@ualberta.ca; pietrosa@ualberta.ca

R. Neville
Department of Mathematics, University of Arizona, Tucson, AZ, USA
e-mail: raneville@math.arizona.edu

M. Tsuruga
Department of Mathematics, University of California, Davis, Davis, CA, USA
e-mail: mtsuruga@math.ucdavis.edu

G. Heo (✉)
School of Dentistry, University of Alberta, Edmonton, Alberta, Canada
e-mail: gheo@ualberta.ca

E. W. Chambers et al. (eds.), *Research in Computational Topology*, Association for Women in Mathematics Series 13, https://doi.org/10.1007/978-3-319-89593-2_10

equivalent geometric objects and offers insight into the study of the Kuramoto–Sivashinsky partial differential equation and Lissajous knots. We were unable to obtain meaningful results, however, in our applications to the screening of anomalous facial structures, although our method seems sensitive enough to identify patients at severe risk of a sleep disorder associated closely with craniofacial structure. This approach, though still in its infancy, presents new insights and avenues for the analysis of data with complex structure.

1 Introduction and Motivation

Persistent homology has emerged as a primary tool in topological data analysis, encoding both topological and geometric characteristics of data. This is done in the following way: a filtration of topological spaces is associated to data. Using standard methods from algebra and topology, the topological features are observed through the filtration and encoded in the invariant called a barcode. The most common filtration is built on point cloud data by forming a simplicial complex, on points within some proximity parameter, ϵ. Changes in homology of the simplicial complex are observed as the proximity parameter increases, which captures multiscale topological features. The reader is directed to [4, 12, 14] for a careful introduction to persistent homology.

Although persistence barcodes are useful for distinguishing many objects, there are limitations. For example, persistent homology alone cannot distinguish between data sampled from a circle or from an ellipse of similar size, as both shapes exhibit a topological hole. The discriminating power of persistent homology could benefit by exploiting multiple characteristics of the data. Collins et al. [6] suggested constructing simplicial complexes for a fixed proximity parameter ϵ by varying a curvature threshold parameter κ, approximating the curvature numerically using each point's nearest neighbors. The resulting filtration would start with a simplicial complex determined by the fixed value of ϵ and one would observe the changing homology of this complex as points with low curvature are filtered out of consideration. However, this method requires a choice of proximity parameter to be made initially. The removal of such choices is a strength of persistent homology.

Multidimensional persistent homology provides the mathematical framework to characterize datasets naturally described by more than one parameter. For instance, we may wish to study homological changes in protein unfolding patterns as a function of both time and scale, or even more simply perform several different filtrations on geometric objects incorporating scale, curvature, or torsion. Figure 1 shows the evolution of a simplicial complex under changes in both scale and curvature parameters ϵ and κ, respectively.

It has long been proven [5], however, that no perfect higher-dimensional analogue of the one-dimensional persistence barcode—that is, a complete discrete algebraic invariant—exists for higher-dimensional filtrations. The problem of generalizing a barcode from a single dimension to multiple dimensions remains a difficult

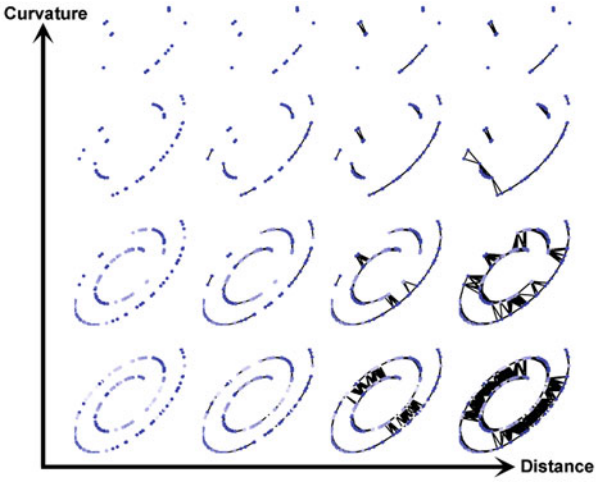

Fig. 1 Visualization of a bifiltration parametrized by curvature κ and distance, or scale ϵ for a simulated dataset sampled from an elliptical spiral. As the curvature parameter κ increases on the vertical axis, points with curvature less than κ are removed from the dataset. The scale parameter, ϵ, increases along the horizontal axis. At a pair of filtration values, ϵ_i and κ_j, a Vietoris–Rips (VR) complex is built on points in the dataset within a distance of ϵ_i, after removing points from the data with curvature less than κ_j (Edges, but not faces of the VR complex are shown)

mathematical problem [18]. The theoretical limitation of accessible representations of multidimensional persistence, however, does not stop researchers' heuristic applications to a wide range of fields including brain networks [17], Vicsek and D'Orsogna models [26], and biomolecular data [29], for example.

In this article, we consider a practical approach to approximating multidimensional persistent homology. Looking to the 2015 work of Xia and Wei [29], we further develop their methods of *pseudo-multidimensional persistence*. We are particularly motivated by the task of increasing the discriminatory power of current one-dimensional persistence techniques. To accomplish this goal, we use the pseudo-multidimensional persistence technique to incorporate additional parameters into the standard one-variable filtration and a computationally-feasible framework that admits the desired discriminatory power.

We close this introduction with a brief description of the following sections. Section 2 proposes a nonparametric method to compare the point-cloud datasets using pseudo-multidimensional persistence in two variables, which we apply to the simple problem of distinguishing between a circle and an ellipse using scale and curvature parameters. In Sects. 3–5, we apply methods for pseudo-multidimensional persistence to real-world and simulated datasets, including craniofacial shape, Lissajous knots, and the Kuramoto–Sivashinsky partial differential equation. We conclude this work in Sect. 6 with comments on these heatmap methods and suggestions for future work.

2 Illustration of Heatmap Method with Circle and Ellipse

We return to the challenge of distinguishing a circle from an ellipse to illustrate the extension of one-dimensional persistence to a two-dimensional heatmap built from filtrations of both scale and curvature. In general, an appropriate method for the calculation of curvature at each point of a point-cloud dataset may differ depending on the nature of the data. In this paper, for noisy data obtained in applied settings, such as the craniofacial data of Sect. 3, we employ a generalization of the hyper circle-fitting algorithm [2] to arbitrary dimensions to estimate curvature: The curvature at a given point will be defined as the reciprocal of the radius of the hypersphere fit locally to the dataset at that point. Our choice of algorithm is discussed further in Sect. 3. In pure settings, such as with the Lissajous knots of Sect. 4, it may be possible to calculate curvature analytically.

Consider a set of points sampled from a uniform random distribution on the boundary of a circle of radius 1 and the boundary of an ellipse with major axis 2 and minor axis 1, both with a small amount of added noise. We mix noise into the sample data in order to replicate real data that is generally composed of true signal which we wish to estimate, and random noise that we wish to ignore. We first set a dimension p as the dimension of the homology to examine, and further fix finite sequences $(\kappa_i)_{i=1}^K$ and $(\epsilon_j)_{j=1}^J$ of *curvature* and *distance thresholds*, respectively.

For each dataset and for each choice of curvature and distance thresholds κ_i and ϵ_j, we calculate the p-th Betti number, β_p at the specified thresholds. More specifically, only considering those points of the dataset with estimated curvature at least the chosen threshold κ_i, we calculate the number $\beta_p(\kappa_i, \epsilon_j)$ of p-th dimensional homological components at the scale parameter ϵ_j. In implementation, this is done by constructing a persistence barcode for the Vietoris–Rips filtration on each dataset after removing points with estimated curvature less than κ_i, and computing the number of p-th dimensional homological components at each distance threshold ε_j. The heatmap matrix formed by associating $\beta_p(\kappa_i, \epsilon_j)$ to each position (i, j). As shown in Fig. 2, the heatmaps for a circle and an ellipse visually differentiate between them. In particular, the heatmap representing the first homological dimension for a circle shows a topological hole of fairly uniform prominence (visualized as the values across a single row of the heatmap). This continues up to a certain curvature threshold, above which there is no longer a hole, as expected. There is a notable decrease in the prominence of the topological hole for higher curvature thresholds of the ellipse.

Motivated by the above result, we now propose a nonparametric test of dissimilarity between the heatmaps generated from two sample spaces S_A and S_B. To do this, we compare the heatmap derived from a single sample of S_B with heatmaps obtained from N samples of S_A. For our example, S_A and S_B represent, respectively, the sample space of points sampled from the circle and ellipse described above.

1. Fix a sequence of curvature and scale thresholds $(\kappa_i)_{i=1}^K$ and $(\epsilon_j)_{j=1}^J$, respectively. For $n = 1, 2, \cdots, N$, create the $K \times J$ *heatmap matrix* H_n^A, where each

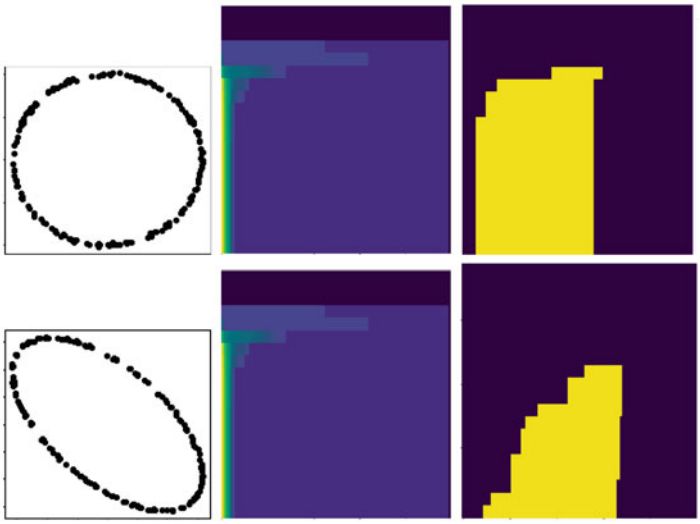

Fig. 2 (Top row) β_0 (center) and β_1 (right) heatmap plots of the circle on the left. (Bottom row) β_0 (center) and β_1 (right) heatmap plot of the ellipse on the left. The horizontal and vertical axes of each heatmap indicates values of the scale and curvature parameters ϵ and κ, respectively. In all heatmaps above, Betti numbers β are plotted as $\ln(\beta + 1)$, so that heatmap colors show better contrast for large differences in Betti number. We can see, visually, the differences in the circle and ellipse heatmaps for each dimension

of these heatmaps is generated from a sample of S_A. In general, we define the entry in the i-th row and j-th column of a heatmap matrix to be β_p at curvature and scale threshold parameters κ_i and ϵ_j, respectively, divided (or "normalized") by the total number of points in the sample used to obtain the heatmap.

2. Calculate the mean heatmap matrix $\overline{H}^A$ of the heatmap matrices obtained in the previous step. More precisely, $\overline{H}^A = \frac{1}{N}\sum_{n=1}^{N} H_n^A$.
3. For $n = 1, 2, \cdots, N$, calculate the dissimilarity d_n between each heatmap H_n^A and the mean heatmap $\overline{H}^A$: As an example, we could define such dissimilarity by

$$d_n = d(H_n^A, \overline{H}^A) = \sum_{i=1}^{K}\sum_{j=1}^{E}\left[\ln\left((H_n^A)_{ij} + 1\right) - \ln\left((\overline{H}^A)_{ij} + 1\right)\right]^2.$$

4. Calculate the dissimilarity d_{obs} between the heatmaps $\overline{H}^A$ and H^B, where H^B is the heatmap generated from a point-sampling of S_B.
5. (Step 4 may be iterated using repeated samples obtained from S_B, if available). One may take the p-value for this test to be the proportion of iterations of Step 4 where $\{d_n > d_{obs}\}$. A smaller p-value indicates stronger evidence that the heatmaps generated from each of the sample spaces S_A and S_B are different.

Table 1 Parameters and results for the proposed non-parametric test of dissimilarity for comparing point-cloud datasets

Test #	N	n_A	n_B	K	J	s	p_0	p_1
1 (circle, ellipse)	20	200	200	20	100	0.3	0	0
2 (ellipses)	20	200	200	20	100	0.3	1	1
3 (3-leaf, 4-leaf clovers)	20	250	250	20	50	0.3	0	0
4 (wedge product of two spheres)	20	600	600	20	100	0.3	0	0

The columns represent number of bootstrap samples (N), number of points in each point-cloud (n_A, n_B), number of curvature and scale thresholds used (K and J, respectively), radius of the hypersphere used for curvature estimation (s), and p-values for the proposed test in dimension 0 (p_0) and 1 (p_1)

To test the above procedure, we performed a number of experiments with generated data. These tests are summarized in Table 1, where the S_A and S_B columns describe the sample spaces being compared, N the number of samples drawn from S_A, n_A and n_B the number of points sampled from each sample space, K and J the number of curvature and scale parameter threshold values chosen, s the radius of the neighborhood used during curvature estimation, and p_0 and p_1 the p-values calculated by the test when comparing the zero- and one-dimensional homology groups of S_A and S_B.

Curvature thresholds are chosen to be $\{q_{k/K} | k = 0, \ldots, K-1\}$, where q_p is the p-th quantile for all computed curvature values, for $p \in [0, 1]$. Fifty scale thresholds were similarly chosen using quantiles of the death times of dimension-0 homological components obtained from univariate persistent homology applied to each of datasets under consideration. The sample spaces compared in each test, S_A and S_B, respectively, are as follows. Test #1 compares an ellipse with vertical axis of length 2 and horizontal axis of length 1, rotated 45 degrees CCW against the unit circle. Test #2 compares an ellipse with vertical axis of length 2 and horizontal axis of length 1, rotated 45 degrees counterclockwise against the same ellipse but rotated 45 degrees clockwise. Test #3 compares a"three-leaf clover" shape with a "four-leaf clover" shape, as shown in Fig. 3. Finally, Test #4 tests out method in three dimensions, and compares the wedge product of two spheres with the same radius of 1 against the wedge product of two spheres, one with radius 1 and the other with radius 2.

From the results of our tests, we see that the above approach is sensitive enough to distinguish between shapes that are homologically equivalent but have different curvatures, as discussed at the beginning of this section. Furthermore, this method is not overly sensitive to falsely distinguish between different point-samples of the same object transformed under rotation. The proposed technique is clearly reliant on the curvature estimates obtained for each point of the sampled dataset, particularly when the sampled data contains a moderate amount of noise. While the hyper circle-fitting algorithm is superior to numerous other circle-fitting methods—such as the Kasa, Pratt, or Taubin fits—in terms of essential bias [2], we note that the curvatures obtained are sensitive to moderate levels of noise in the sampled data. This is

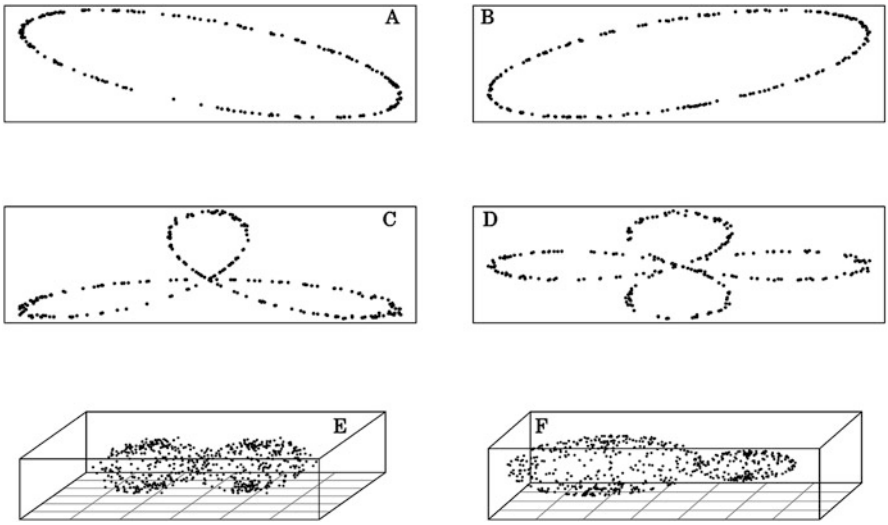

Fig. 3 (**a** and **b**) Test #2 datasets sampled from two ellipses identical up to rotation. (**c** and **d**) Test #3 dataset sampled from curves resembling a three- and four-leaf clover. (**e** and **f**) Test #4 dataset sampled from the wedge product of two spheres: only in one dataset are the two spheres of the same radius

particularly true when the size of the neighborhood used to locally fit a hypersphere to the given data (and, by extension, obtain the estimated curvature as the inverse of the radius of this hypersphere) is small, or when the number of points used to fit the hypersphere is otherwise small.

The remainder of this paper will detail several applications of heatmaps and is organized as follows. The first task is to predict a patient's risk of developing a specific sleep disorder from facial scans. Next, we distinguish specific types of knots. Third, we show how heatmaps may be used to distinguish between dynamic data of pattern evolution, driven by different parameter values. In these applications, some knowledge about the data and possible distinguishing characteristics drives the choice of the types of filtrations (for instance, curvature and distance in the previous example). It is our hope that through these examples, we demonstrate the added utility of the heatmap representation of a pseudo-multidimensional persistence in a variety of contexts.

3 Craniofacial Shape Analysis

We examine the use of heatmaps in an applied setting, namely, in the analysis of point-cloud scans obtained from pediatric patients for the screening of obstructive sleep apnea (OSA). Pediatric OSA is a sleep disorder with serious health problem that may lead to high blood pressure, behavioral challenges, or altered overall

growth. The gold standard for pediatric OSA diagnosis is overnight polysomnography in a hospital or sleep clinic [19]. In many countries, however, access to overnight polysomnography is severely limited and many children are unable to obtain a proper diagnosis before treatment. This absence of an accessible diagnostic method thus prompts the search for alternative screening methods. Evidence continues to demonstrate a link between craniofacial shape and pediatric OSA [13]. As part of a larger research initiative examining alternative OSA screening methods leaning on orthodontic expertise and statistical techniques for shape and high-dimensional data analysis, we examine the utility of the proposed heatmap method for identifying children at risk of developing OSA on the basis of 3D craniofacial scans.

The dataset used in our investigation is composed of 3D facial scans obtained from 31 children 2–17 years of age recruited from the Stollery Children's Hospital at the University of Alberta. Prior ethics approval has been granted by the University of Alberta's Research Ethics Board. All of the recruited subjects have undertaken overnight polysomnography and have had 3D photos of their face taken. Based on polysomnography results, an apnea–hypopnea index (AHI) for each subject was calculated. These AHI measurements obtained from polysomnography are commonly used to classify patients according to OSA risk severity into one of four categories: no likely risk ($AHI < 1$), mild risk ($1 \leq AHI < 5$), moderate risk ($5 \leq AHI < 10$), and severe risk ($AHI \geq 10$). In this preliminary analysis, we only seek to classify patients into one of two derived groups: no/mild risk ($AHI \leq 5$) and moderate/high risk ($AHI > 5$).

The 3D photo of each patient's face is itself a point-cloud in $\mathbb{R}^3$. For each point, a measure of curvature is computed (see Fig. 4). This task is more difficult than when working with analytic curves, such as the Lissajous knots in the following section. To approach the issue of curvature estimation in this applied setting, we seek to locally fit a sphere in the neighborhood of each point of the dataset. In general, we define the curvature of a dataset at a given point to be the reciprocal of the radius of a hypersphere fit to the dataset in a neighborhood of that point. In our case, we take this neighborhood to be a cube with a manually-selected side length centered at the point whose curvature is being estimated. The relative size of this neighborhood is shown in Fig. 5 for reference.

For the estimation of point-cloud curvature, we implement Al-Sharadqah and Chernov's hyper circle-fitting algorithm [2] in three dimensions. We choose to use this particular method due to its simplicity in generalizing to fit hyperspheres to data of arbitrary dimension and in its property of being essentially unbiased. In other words, while the hyperspheres fit by this method are not unbiased, the level of bias is proportional to the reciprocal of the number of points being fit to. As a result, we have control over the amount of bias in the parameters of the fitted hypersphere—an improvement over other estimation methods [2]. To create the heatmaps for a filtration in both scale and curvature, we must first choose a set of scale and curvature thresholds. In our case, we select curvature thresholds to be the quantiles of the curvature values computed for all points in all point-clouds from

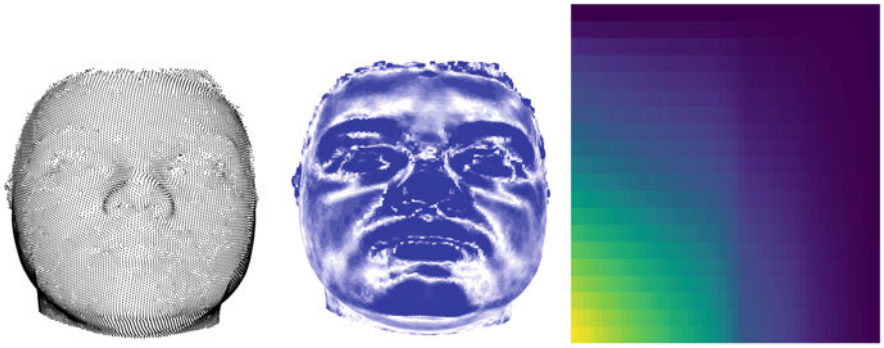

Fig. 4 (Left) An example of a raw point-cloud dataset. (Middle) The same point-cloud dataset as on the left, with points colored according to curvature. White represents low estimated curvature, while deep blue represents high estimated curvature. (Right) The β_0 heatmap corresponds to the two-parameter filtration of the dataset on the left by both scale and curvature

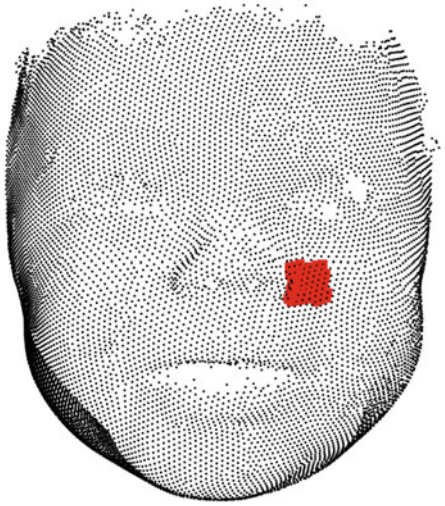

Fig. 5 A point-cloud extracted from a single patient. Shown in red is the neighborhood used to estimate the curvature of some point in the centre of the red-shaded region

all patients. Twenty quantiles were chosen, namely, $\{q_{k/19} | k = 0, \ldots, 19\}$, where q_p is the p-th quantile for all computed curvature values, for $p \in [0, 1]$. Fifty scale thresholds were similarly chosen using quantiles of the death times of dimension-0 homological components obtained from univariate persistent homology applied to each of the patient datasets. We stress that the existence of an "optimal" method of choosing these scale and curvature threshold values has not been investigated in any depth and remains a topic for future work.

Having obtained heatmaps for each patient, we now proceed to address the task of classifying patients into the no/low risk or the moderate/severe risk groups. Due to the study design, a number of the pediatric patients are designated as controls with little to no risk of OSA: Using the heatmaps obtained from these

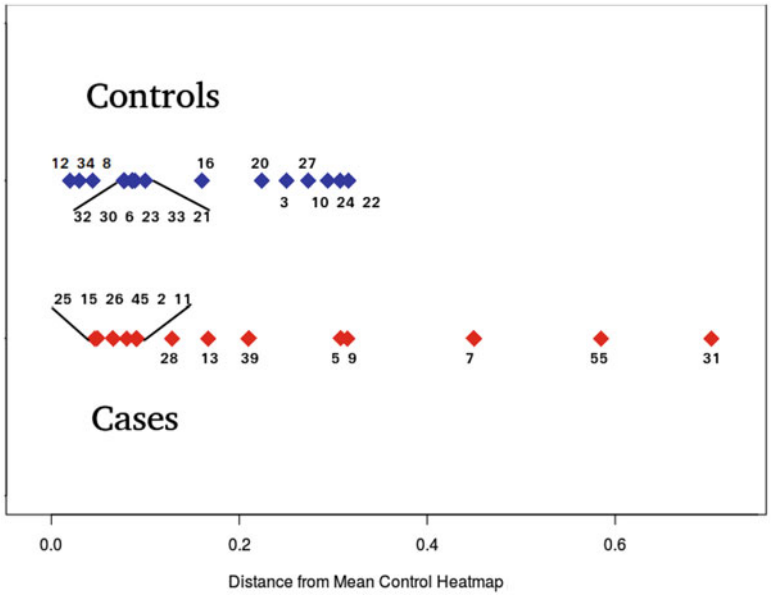

Fig. 6 A visual summary of the dissimilarity (x-axis) between each patient's computed heatmap and the average control heatmap. Controls are represented with blue points, while cases (that is, non-controls) are represented using red points. One case patient (numbered as patient #1), at a distance of 1.2 from the mean control heatmap, is not shown in this diagram to preserve scale. Clearly, the heatmap distance does not adequately distinguish cases from controls. Seven out of twelve cases were clustered with the most of control subjects. Subject #1 indicates a possible outlier; however, its actual AHI is 10.0, the 7th highest AHI. Subject #55 has the highest AHI of 40.2

control patients, we compute an average heatmap for the controls. From here, we calculate the dissimilarity between each patient's heatmap and the average control heatmap, according to the dissimilarity measure defined in the previous section. These dissimilarities, for both controls and cases (non-controls), are summarized in Fig. 6 and in Table 2, sorted by increasing AHI score.

It is immediately apparent that the distance from the mean control heatmap does not correlate well with AHI score. Indeed, patient 25, for example, has a heatmap closer to the mean control heatmap than nearly all control patients but has one of the highest AHI scores. Most case subjects (11 out of 15) were clustered together with control subjects. There are four cases (subjects 7, 55, 31, and 1 (not shown in the diagram)) that appear different from other subjects, however. Further investigating these subjects, we learn that subjects 1, 7, and 55 either have undergone or have scheduled a tonsillectomy and/or adenoidectomy surgery, suggesting that these patients have severe OSA symptoms. Subject 31 was recommended to an ears, nose, and throat specialist for further diagnosis.

Although our heatmap method is able to detect severe OSA cases, these results suggest that the approach used may not be appropriate in developing craniofacial

Table 2 Case information for the 31 children considered, sorted from left to right by increasing AHI score

ID	23	16	22	24	20	3	6	21	8	10	34	27	32	30	12	33
AHI	0.4	0.6	1.2	1.4	1.5	1.6	1.6	1.6	1.8	1.9	2.1	2.2	2.7	3.7	4	4.3
Cases	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No
OSA	1	1	2	2	2	2	2	2	2	2	2	2	2	2	2	2
ID	11	15	26	9	31	39	5	13	1	7	2	25	45	28	55	
AHI	5	5.4	5.6	6.7	7.5	7.6	8.8	9.4	10	10.4	11.7	11.7	13.6	13.8	40.2	
Cases	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	
OSA	2	3	3	3	3	3	3	3	3	4	4	4	4	4	4	

The rows represent, respectively, patient ID, apnea-hyponea index (AHI), identification as a case or control, and specific OSA severity classification (no(1), mild(2), moderate(3), and severe(4))

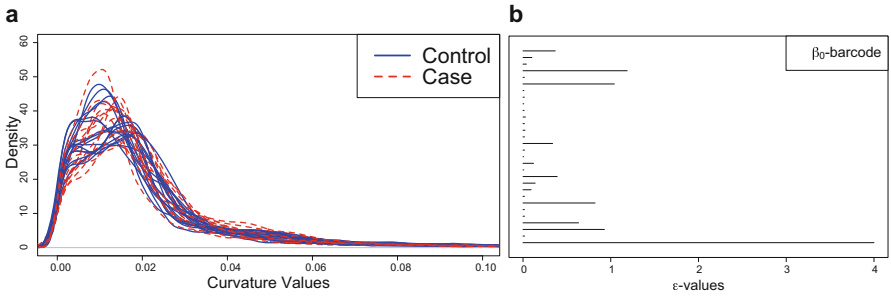

Fig. 7 (**a**) Estimated density function for the computed curvature values for each patient. Curves corresponding to control subjects are shown in blue, while those corresponding to cases are shown in red. (**b**) β_0-barcode based on the Kullback–Leibler measures between the density curves. β_0-barcode indicates six clusters among 31 subjects

shape analysis as a proxy for overnight polysomnography and the derived AHI scores. More generally, our results here leave much in the way of future work in this area, although demonstrated an ability to identify extreme OSA cases. It may be necessary to focus on a specific part of the face rather than the face in its entirely, or perhaps curvature and scale thresholds need to be selected in a way that will allow the resulting heatmaps to better highlight differences between the faces under study. Furthermore, it may be of use to apply different statistical techniques, such as those in machine learning, to this problem of heatmap classification.

As a means of comparison, we also apply one-dimensional persistent homology to craniofacial data. We estimate the probability density function of the curvature values computed for each patient (Fig. 7). We calculate Kullback–Leibler distance which measures the difference between two probability distributions over the same variable [16]. A β_0-barcode based on the Kullback–Leibler divergence between densities is shown in Fig. 7. Subjects corresponding to the six most persistent clusters are presented in Fig. 8 applying multidimensional scaling. Multidimensional scaling (MDS) is a nonlinear dimension reduction technique in which the data is assigned coordinates in a lower dimensional space in a way that most closely preserves the distance between points.

The largest cluster is formed with 17 subjects—8 cases and 9 controls. We also see a similar pattern in the second largest component composed of 3 cases and 2 controls. Univariate persistent homology is unable to differentiate between these two groups, while the heatmap approach seems to be able to do so for extreme cases. This result demonstrates that some benefits exist in choosing to use heatmaps over univariate persistent homology, particularly for the analysis of datasets described by multiple parameters.

Both the proposed heatmap and univariate persistent homology methods based on curvature fail to differentiate between obstructive sleep apnea patients and controls, except that bifiltration heatmap was capable of detecting a few extreme cases (see Fig. 6). We can think of three potential reasons for this: (1) curvature may not be

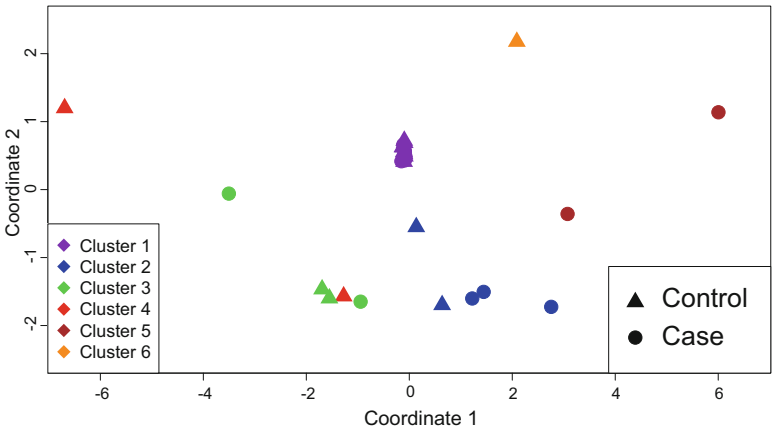

Fig. 8 Clustering patterns of subjects presented on first two MDS coordinates (coordinate 1 on X-axis and coordinate 2 on Y-axis). The subgroups (clusters) do not seem to relate to any common traits among the subjects. The largest cluster (in purple) consists of 17 subjects—8 cases and 9 controls. The second largest cluster (in blue) consists of five subjects—3 cases and 2 controls

the best way to measure craniofacial form, (2) our proposed heatmap method is not effective for complex data such as face shape, or (3) AHI may not be the most effective variable in determining the presence or absence of OSA as several researchers have suggested [30].

4 Lissajous Knots

In this section, we apply the heatmaps to the study of a geometric object, namely, Lissajous knots. The filtrations throughout this section use scale and curvature parameters. Our goal is to use known curves, Lissajous knots, to investigate how heatmaps can help us understand features of this data. In particular, as shown below, we deduce that the two flares seen in the heatmaps from these Lissajous knots correspond with the fact that the data points of highest curvature are separated into two disjoint sets. In the future, we would like to use this information to interpret the heatmaps of data sets from other one-dimensional curves (such as other types of knots and some time series data) where the density of data points is higher around curves.

A Lissajous knot [3] is a closed curve that is isotopic to some curve with parameterization of the form

$$x(t) = \cos(n_x t + \phi_x)$$

$$y(t) = \cos(n_y t + \phi_y)$$

$$z(t) = \cos(n_z t + \phi_z),$$

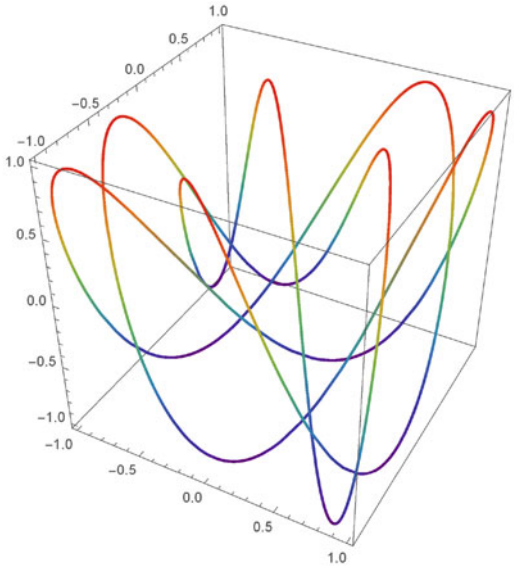

Fig. 9 The (3, 4, 7, 0.1, 0.7) Lissajous knot given by the parameterization $(x(t), y(t), z(t)) = (\cos(3t + 0.1), \cos(4t + 0.7), \cos(7t))$ is of type 8_{21} according to Rolfsen's Knot Table [22]. The knot is colored by z value (vertical axis). We immediately observe that the curvature of this knot is largest at its highest and lowest heights, that is, where minimum and maximum values of z are attained. Visually, we see that the knot's curvature oscillates periodically between large and small values as the path of the knot is traced out

where $0 \leq t \leq 2\pi$ and i and j are any distinct elements of $\{x, y, z\}$. We require the frequencies n_i to be pairwise relatively prime integers, ϕ_i a real number, and $n_i\phi_j - n_j\phi_i$ not a multiple of π. These conditions ensure that the resulting curve is closed and without self-intersection. To simplify we may, with a change of variable, assume that $\phi_z = 0$: this parameterization places the Lissajous knot inside the cube centered at the origin with edge length 2. As such, we may denote a Lissajous knot by $(n_x, n_y, n_z, \phi_x, \phi_y)$. For example, Fig. 9 shows the (3, 4, 7, 0.1, 0.7) Lissajous knot, given by $(x(t), y(t), z(t)) = (\cos(3t + 0.1), \cos(4t + 0.7), \cos(7t))$. Under Rolfsen's Knot Table classification [22], this knot is of type 8_{21}. According to Table 1 of [3], a given knot type may be associated with multiple values of $(n_x, n_y, n_z, \phi_x, \phi_y)$: As such, in this section, we restrict ourselves to exactly one parameterization for each of the knot types considered.

Lissajous knots have practical applications, particularly in the modelling of DNA [3], and have thus been well-studied. The knots in this family are highly symmetric, with curvature values changing predictably over a wide interval as a function of t. These properties make Lissajous knots an appropriate object of interest in our study of heatmaps where we may naturally incorporate both scale and curvature filtration parameters.

For each Lissajous knot considered, we create a dataset with size proportional to the knot's total arc length. The (3, 4, 7, 0.1, 0.7) Lissajous knot, for example, with the parameterization given previously, has an arc length of 36.8. We thus choose to create the dataset for this knot using 368 sampled points. Since the knot itself is parameterized by t, we create the dataset S by picking 368 equally-spaced t values between 0 and 2π, that is, $S = \{t_k = 2\pi k/368 \mid k = 0, \ldots, 367\}$ and find their spatial coordinates $(x(t_k), y(t_k), z(t_k))$, for $k = 0, \ldots, 367$.

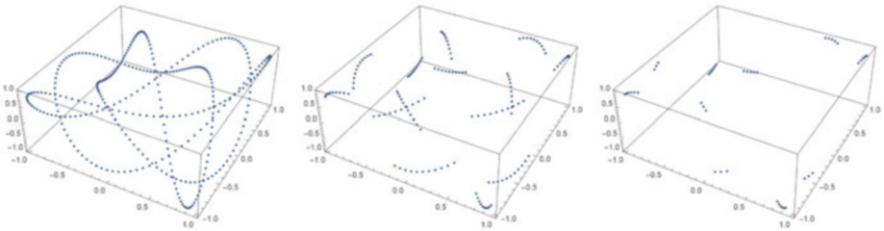

Fig. 10 Super-level set point-clouds from the (3, 4, 7, 0.1, 0.7) Lissajous knot dataset. (Left) All points with curvature greater than 0, that is, the entire dataset of 368 points. (Middle) All points with curvature greater than 1, consisting of 147 points. (Right) All points with curvature greater than 3, comprised of 47 points

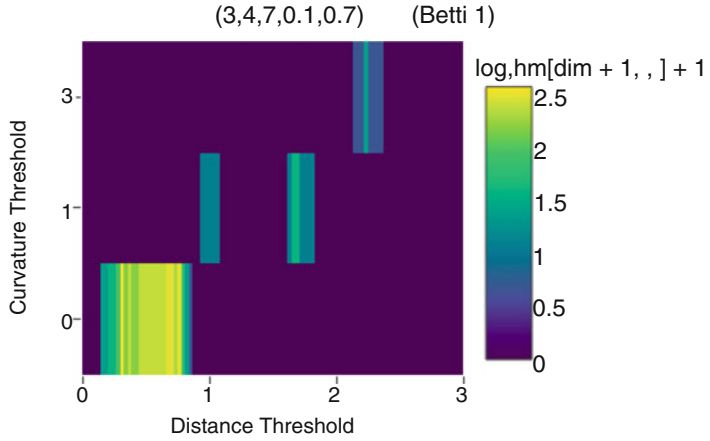

Fig. 11 Persistence heatmap for the first homology group of (3, 4, 7, 0.1, 0.7) Lissajous knot dataset. The one-dimensional heatmap for the entire dataset of 368 points is shown in the bottom row of the heatmap, with the one-dimensional heatmaps for the smaller datasets stacked on top in increasing order of curvature threshold used. The distance threshold increases along the x-axis from 0 to 3 using step size 0.3. The color scale on the right represents the natural log of the dimension-1 Betti numbers after being increased by 1

Using Mathematica, we then (analytically) compute the curvature, $\kappa(t)$, at each selected point, and create super-level sets $S_h = \{(x(t), y(t), z(t)) \in \mathbb{R}^3 : \kappa(t) > h\}$, for any real h, referred to as the curvature threshold. In short, S_h consists of all points of the dataset with curvature strictly greater than h.

Figure 10 shows three such super-level sets from the (3, 4, 7, 0.1, 0.7) Lissajous knot dataset. Figure 11 displays the persistence heatmap for the first homology group for this knot under various scale and curvature threshold values.

We can create a larger heatmap by choosing more curvature thresholds as in Fig. 12, which shows another heatmap for the (3, 4, 7, 0.1, 0.7) Lissajous knot. Along the x-axis is the scale parameter ranging from 0 to 3 with step size 0.3. Along the y-axis is the curvature parameter ranging from 0 to the maximum curvature by step size 0.1. The maximum curvature for each knot is different: For example, the we use maximum curvature 16.5 for the (3, 4, 7, 0.1, 0.7) knot.

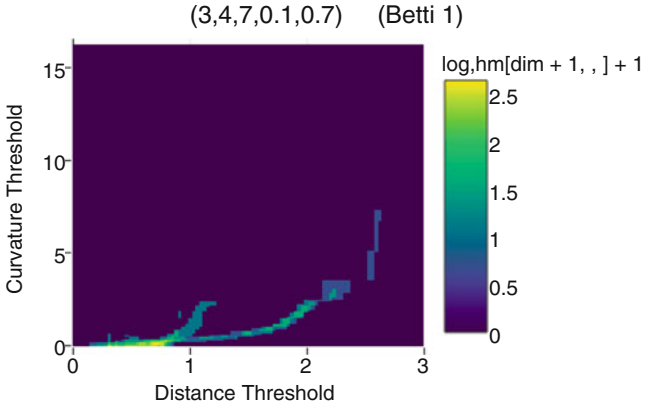

Fig. 12 Heatmap for the (3,4,7,0.1,0.7) Lissajous knot. The distance threshold increases along the x-axis from 0 to 3 by step size 0.3. The curvature threshold increases along the y-axis from 0 to its maximum curvature 16.5 by step size 0.1. The color scale on the right represents the natural log of the dimension-1 Betti numbers after being increased by 1

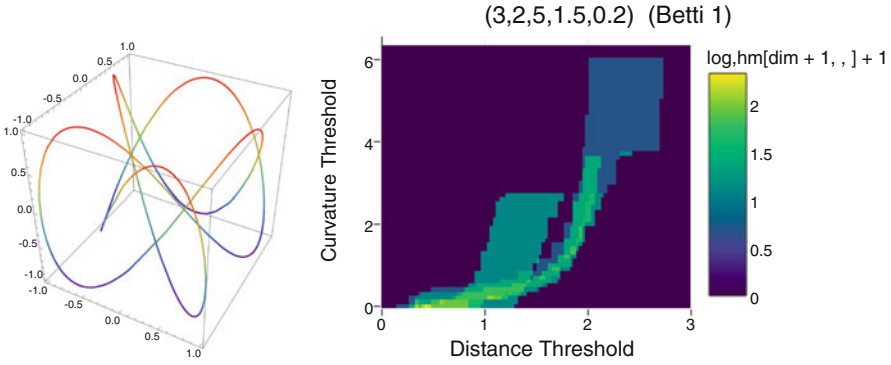

Fig. 13 Left: the (3, 2, 5, 1.5, 0.2) Lissajous knot with parametrization $(x, y, z) = (\cos(3t + 1.5), \cos(2t + 0.2), \cos(5t))$. Right: the heatmap for this knot. The distance threshold increases along the x-axis from 0 to 3 by step size 0.3. The curvature threshold increases along the y-axis from 0 to a maximum curvature 6.5 by step size 0.1. The color scale on the right represents the natural log of the dimension-1 Betti numbers after being increased by 1

Figures 13 and 14 show two more heatmaps for other Lissajous knots, namely, (3, 2, 5, 1.5, 0.2) and (2, 3, 11, 0.2, 0.7), respectively, along with their knot graphs. The (3, 2, 5, 1.5, 0.2) Lissajous knot is given by $(x, y, z) = (\cos(3t + 1.5), \cos(2t + 0.2), \cos(5t))$ and has total arc length of 26.4. We then choose to construct the dataset for this knot using 264 points. According to Rolfsen's Knot Table [22], this knot is of type 6_1. The (2, 3, 11, 0.2, 0.7) Lissajous knot is parameterized via $(x, y, z) = (\cos(2t + 0.2), \cos(3t + 0.7), \cos(11t))$ and has arc length of 47.8. We thus create the dataset for this knot using 478 points.

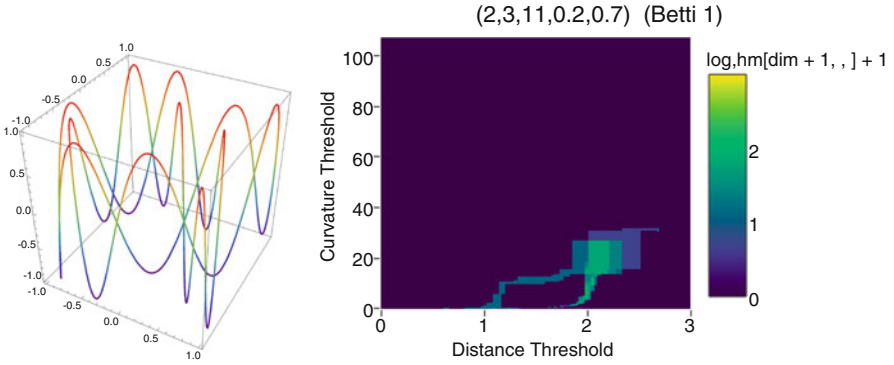

Fig. 14 Left: the (2,3,11,0.2,0.7) Lissajous knot with parametrization $(x, y, z) = (\cos(2t + 0.2), \cos(3t + 0.7), \cos(11t))$. (Right) Heatmap for this Lissajous knot. The distance threshold increases along the x-axis from 0 to 3 by step size 0.3. The curvature threshold increases along the y-axis from 0 to its maximum value 107 by step size 0.1

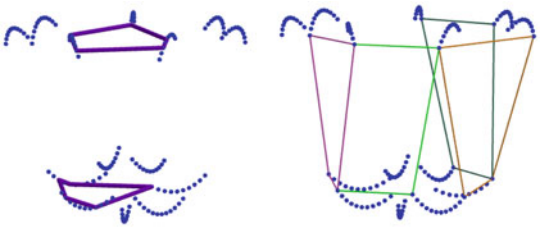

Fig. 15 Cycles in the super level set at curvature threshold 1 of the (3, 4, 7, 0.1, 0.7) Lissajous knot. (Left) Two cycles are present in the super-level set at distance threshold 1. (Right) Four cycles are present in the super-level set at distance threshold 1.7

A curiosity of Figs. 12, 13, and 14 is the two blue-green flares across each of the heatmaps. We investigate this further using the software ShortLoop[10] by visualizing the cycles in the first homology of our dataset. Figure 15 shows the cycles appearing in the super-level set of curvature threshold 1 for the (3, 4, 7, 0.1, 0.7) Lissajous knot at two different ε distance thresholds. We see two cycles at $\varepsilon = 1$ on the left, and four cycles at $\varepsilon = 1.7$ on the right.

As mentioned previously, Lissajous knots attain maximum curvature at their highest and lowest heights (z values). As a result, filtering our data by curvature may also be viewed as filtering by the absolute z-coordinate. For a given curvature threshold, our point-cloud can then, loosely speaking, be clustered into groups with a relatively high z and those with relatively low z. The flares in the heatmap represent these two clusters. When ε is smaller than the distance between the clusters, cycles in each cluster result, but no cycles exist connecting the two clusters. There is an interval of ε values in which there are still no cycles between the clusters, but the cycles within each cluster close up. This is the empty purple space between the flares. Once ε is large enough, cycles again form, now connecting the two clusters.

In addition to these two large clusters, we can see from the knots that we will have small clusters of points at each of the knot's local extrema. The number of local maxima is given by the n_z term in the parameterization. Knot (3, 2, 5, 1.5, 0.2) in Fig. 13, for example, has only five peaks, and the curvature at these peaks is less than that in other knots. As a result, the points in the mini-clusters at each local maximum and minimum are further spread apart, bringing these smaller clusters closer together. We see this by the wider and closer flares in Fig. 13 heatmap. On the other hand, knot (2, 3, 11, 0.2, 0.7) in Fig. 14 has eleven peaks and the points in the clusters at the peaks have a higher curvature than the other knots. The smaller clusters are thus denser, leading to larger gaps between these smaller clusters. The resulting heatmap for this knot has skinnier, more separated flares, as seen in Fig. 14.

The heatmaps we have studied thus far demonstrate that Lissajous knots with higher periodicity have skinnier flares and a larger range in curvature values that results in more open space at the top of the plot. This preliminary work suggests that this method may be useful in the study's other periodic time series datasets.

5 Classification of Anisotropic Kuramoto–Sivashinsky Solution

Indeed, there are many contexts where spatial structure of data is changing temporally. In this vein, we will explore a complex spatio-temporal pattern, using pseudo-multidimensional persistence to account for both the spatial and temporal variation of a pattern. Such patterns often occur in nonlinear systems that are driven from equilibrium by, for example, a gradient in temperature, concentration, or velocity [24]. Consider the coloration patterns of zebras or specific species of fish, ripples in sand dunes, or convection cell formations in clouds. Understanding these pattern-forming systems is important to a wide variety of fields in the scientific community such as biology, physics, engineering, and chemistry [7].

In modeling physical phenomena, it is often the case that poorly-resolved or poorly understood processes are parametrized rather than treated explicitly. Because of this, it becomes important to determine the influence of model parameters on the system. There are a variety of methods to do this, many of which require computationally expensive simulations [1]. Irregular time-varying structures and complexity of patterns, and sensitivity to initial conditions, among other things, make quantifying or even distinguishing patterns difficult [11]. Recently there has been much interest in using topological methods in pattern formation and pattern evolution, in particular in material sciences [28]. Computational topology has emerged as a tool that retains some essential information for studying patterns, while significantly reduces the dimensionality of the data [9]. For example, persistent homology has been used to distinguish between parameters for complex patterns formed through a phase separation process [11] in the Cahn–Hilliard equations. In this example, the patterns were studied at specific, static moments in time.

Pseudo-multidimensional persistence allows for the inclusion of the time evolution. We will apply this technique to simulations of the two-dimensional anisotropic Kuramoto–Sivashinsky equation.

The Kuramoto–Sivashinsky (KS) equation is a partial differential equation used to model systems driven from equilibrium [21]. It has found many applications in surface pattern-formation including flame front propagation [23], surface patterning by ion-beam erosion [8, 20], epitaxial growth and instabilities related to electromagnetism [27], the formation of suncups in snowfields [25], and solidification from a melt [15]. The solution $u(x, y, t)$ gives a patterned surface in two spatial variables that evolves in time. This equation arises in applications as surface nanopatterning by ion-beam erosion and solidification. The anisotropic Kuramoto–Sivashinsky (aKS) equation is given by

$$\frac{\partial u}{\partial t} = -\nabla^2 u - \nabla^2 \nabla^2 u + r\left(\frac{\partial u}{\partial x}\right)^2 + \left(\frac{\partial u}{\partial y}\right)^2, \tag{1}$$

where $\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}$. The parameter r controls the anisotropy in the nonlinear term. The goal of this experiment is to classify sets of solutions by parameter r. Numerical simulations of the aKS equation are generated for parameter values $r = 0.5, 0.75, 1, 1.25$, and 1.5. Thirty trials for each parameter value are generated using a low-amplitude white noise initial condition (Fig. 16). Persistent homology is computed using sublevel sets. At each threshold height, a cubical complex is built on neighboring points with values below the threshold. A filtration of cubical complexes is formed by increasing the threshold from below the surface to the maximum height of the surface. See [11] for an introduction to cubical homology [22].

This data set was initially investigated in the paper [1], with the goal of identifying parameter values of each example using persistent homology. Persistent homology was computed using a cubical complex on a sublevel set filtration at a single moment in time. In order to compare persistence diagrams, Adams et al. devised a stable vector representation of persistence diagrams called persistence images (PIs). Using this method, each persistence diagram is vectorized. Standard machine learning algorithms may then be applied to classify the vectors based on the parameter used to generate each surface. See the paper for full details on the method. Using a subspace discriminant ensemble, which uses the same classification algorithm repeatedly over randomly chosen subspaces of the data. The algorithm fits the data by building a model on the mean and variance of classes in a training set. Each example in the testing set is assigned a class for each iteration. A likely overall class is assigned at the end. In this example, the data was classified at several different time steps. The classification accuracies reported in their experiment are listed in Table 3. Using this method, one must choose a single moment in time to consider the data. We see the restriction to a single moment in time as a limiting factor because the temporal evolution of the pattern is ignored. The vectors representing several time steps may be concatenated into a larger vector to be used for classification, but this causes the size of the vectors to grow quickly,

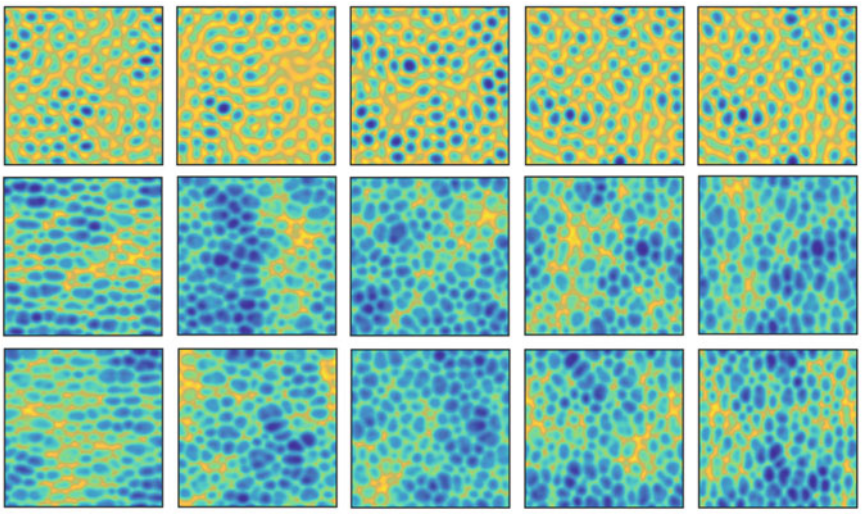

Fig. 16 Plots of the examples of the surfaces $u(x, y, \cdots)$ from the numerical simulations. Each column is generated using a different parameter value; $r = 0.5, 0.75, 1.0, 1.25$, and 1.5. Each row is a single point in time in the evolution of the pattern of the surface, the times shown are $t = 3, 5$ and 10. By $t = 5$ the elongation due to the anisotropy has stabilized, though the surface continues to evolve in time. r=1 (center column) is the isotropic case; there is no elongation in either direction

Table 3 Classification accuracies at different times of the aKS solution, using a fivefold cross-validated subspace discriminant ensemble on a vector representation of one-dimensional barcodes

Persistence representation	Time t=3 (%)	Time t=5 (%)	Time t=10 (%)
β_0 PIs	58.3	96.0	94.7
β_1 PIs	67.7	87.3	93.3
β_0 and β_1 PIs	72.7	95.3	97.3

Classification of times $t = 15$ and 20 results in accuracies similar to $t = 10$

which limits the machine learning techniques that may be efficiently applied. Even with a low resolution, generating persistence images at each time step results in a high-dimensional representation of the data. We consider heatmaps as an alternative representation to capture the evolution of the pattern over time in a way that lends itself to machine learning techniques.

A heatmap will be generated for each example. One dimension will be the height of the sublevel set filtration, represented along the horizontal axis, and the other dimension will be the time step, represented along the vertical axis. The top row represents the earliest time step. Each entry is given by $ln(\beta_i + 1)$ where β_i is the Betti number in dimension 0 or 1. A single heatmap then will give the homological feature information of the surface as it evolves in time. Examples of the heatmap in each homological dimension are shown in Fig. 17. The earliest time is shown on the top row. The pattern is just beginning to form at this stage, so topological features

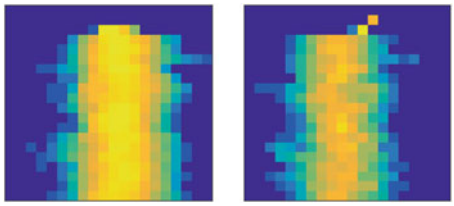

Fig. 17 Examples of heatmaps for the aKS equation, $r = 1$. The left shows zero-dimensional homology and the right is one-dimensional homology. The top row is the earliest time step. The pattern takes some time to emerge, so there in little by way of topological features at the early times

Table 4 Classification accuracies using the MDS reduction of heatmaps and the MDS reduction of PIs concatenated to incorporate several time steps

Homological dimension	Accuracy of MDS reduced heatmap (%)	Accuracy of MDS reduced PIs, t=5,10,15 (%)
β_0	69.3	72.0
β_1	100	98.7
β_0 and β_1	100	100

MDS reduces the overall dimension to 2. Classification was performed using a fivefold cross-validated linear discriminant

are still emerging as well. Interestingly, even though the pattern is dynamically changing, there are not significant changes from one row to the next for much of the time represented here. Once the heatmaps are formed, the L_2 distance is computed between each heatmap. As a natural dimension reduction step, we perform MDS with a chosen dimension of 2. At two-dimensions, 95% of the variance of the data is accounted for. The MDS representation is then used for classification. Similar to [1], classification is performed with a linear discriminant ensemble with fivefold cross-validation. In this case, there is no need to incorporate an ensemble of random subspaces since the dimension of the data is 2. The classification accuracy of the two-dimensional MDS for the heatmaps from β_0 and β_1 and the concatenated information are given in Table 4. We include a comparison with classification accuracies of PIs that have been concatenated to include several time steps, t=5,10, and 15. The PIs were also reduced to 2 dimensions with MDS. Classification accuracy is comparable when using the full time evolution (heatmaps), or several discrete time steps (PIs). It is clear that β_1 is much more discriminating than β_0. This is due to the nature of the pattern, which appears as raised bubbles. In a sublevel set filtration, β_1 will capture these features. Good classification results for the full time evolution are encouraging because it suggests that heatmaps will prove useful to capture more complete information on temporal evolution. In this case, enough information was captured in several time steps that a method such as PIs could be used, but in general this may not contain enough information to discriminate

between classes. Heatmaps allow users to make use of more temporal data that may be available to them. Heatmaps (combined with a dimension reduction technique) provided very low dimensional representations of dynamically changing patterns, that allowed for good parameter recovery.

6 Conclusion and Future Research

In this article, we proposed the heatmap as a tool to approximate multidimensional persistent homology and explored applications to a variety of topics, including shape analysis in 3D craniofacial imaging, Lissajous knots, and the solutions of the anisotropic Kuramoto–Sivashinsky equation. Our use of heatmaps in these areas allowed the incorporation of additional parameters such as scale, curvature, and time, into the regular univariate filtrations of persistent homology. The addition of a second parameter, particularly in our work with the Kuramoto–Sivashinsky equation, allowed a new avenue of insight into the datasets studied that are inherently described by multiple parameters.

As demonstrated in this paper, this technique for the approximation of multidimensional persistence has potential for a wide range of application in numerous fields. However, the techniques presented here are certainly still in their infancy and there remains much work to be done in the future to make them competitive with existing methods, particularly in terms of computational time and classification accuracy. These issues, including the determination of a dissimilarity appropriate for comparing heatmaps as well as the excessive computational time required by our own and other existing techniques for topological data analysis, will need to be addressed before these methods become widely applicable.

Acknowledgements We thank the Workshop for Women in Computational Topology (WinCompTop) held on August 12–19, 2016, Institute for Mathematics and its Applications (IMA). We would like to thank the Seed Grant from Women and Children's Health Research Institute, the National Sciences and Engineering Research Council of Canada (NSERC), the McIntyre Memorial fund from the School of Dentistry at the University of Alberta, and Biomedical Research Award from American Association of Orthodontists Foundation. We are grateful for sleep scholars, research coordinators, and postdoctoral fellows, and technicians in Stollery Children's Hospital, University of Alberta.

References

1. H. Adams, S. Chepushtanova, T. Emerson, E.M. Hanson, M. Kirby, F.C. Motta, R. Neville, C. Peterson, P.D. Shipman, L. Ziegelmeier, Persistence images: a stable vector representation of persistent homology. J. Mach. Learn. Res. **18**, 1–35 (2017)
2. A. Al-Sharadqah, N. Chernov, Error analysis for circle-fitting algorithms. Electron. J. Stat. **3**, 886–911 (2009)

3. M.G.V. Bogle, J.E. Hearst, V.F.R. Jones, L. Stoilov, Lissajous knots. J. Knot Theor. Ramif. **3**(02), 121–140 (1994)
4. P. Bubenik, Statistical topological data analysis using persistence landscapes. J. Mach. Learn. Res. **16**, 77–102 (2015)
5. G. Carlsson, A. Zomorodian, The theory of multidimensional persistence. Discrete Comput. Geom. **42**(1), 71–271 (2009)
6. A. Collins, A. Zomorodian, G. Carlsson, L.J. Guibas, A barcode shape descriptor for curve point cloud data. Comput. Graph. **28**, 881–894 (2004)
7. M. Cross, H. Greenside, *Pattern Formation and Dynamics in Nonequilibrium Systems* (Cambridge University Press, Cambridge, 2009)
8. R. Cuerno, A.-L. Barabási, Dynamic scaling of ion-sputtered surfaces. Phys. Rev. Lett. **74** 4746 (1995)
9. S. Day, W.D. Kalies, T. Wanner, Verified homology computations for nodal domains. Multiscale Model Simul. **7**(4), 1695–1726 (2009)
10. T.K. Dey, J. Sun, Y. Wang, Approximating cycles in a shortest basis of the first homology group from point data. Inverse Prob. **27**, 124004 (2011). An International Journal on the Theory and Practice of Inverse Problems, Inverse Methods and Computerized Inversion of Data
11. P. Dłotko, T. Wanner, Topological microstructure analysis using persistence landscapes. Phys. D **336**, 60–81 (2016)
12. H. Edelsbrunner, J. Harer, *Computational Topology: An Introduction* (American Mathematical Society, Providence, 2010)
13. C. Flores-Mir, M. Korayem, G. Heo, M. Witmans, M.P. Major, P.W. Major, Craniofacial morphological characteristics in children with obstructive sleep apnea syndrome: a systematic review and meta-analysis. J. Am. Dent. Assoc. **2013**, 269–277 (2013)
14. R. Ghrist, *Elementary and Applied Topology* (Createspace Independent Pub, Seattle, 2014)
15. A.A. Golovin, S.H. Davis, Effect of anisotropy on morphological instability in the freezing of a hypercooled melt. Phys. D 116 363–391 (1998)
16. S. Kullback, R.A. Leibler, On information and sufficiency. Ann. Math. Stat. **22**, 79–86 (1951)
17. H. Lee, H. Kang, M.K. Chung, S. Lim, B.N. Kim, D.S. Lee, Integrated multimodal network approach to PET and MRI based on multidimensional persistent homology. Hum. Brain Mapp. **38**, 3871402 (2017)
18. M. Lesnick, M. Wright, Interactive visualization of 2-D persistence modules (2015). arXiv:1512.00180
19. C.L. Marcus, L.J. Brooks, S.D. Ward, K.A. Draper, D. Gozal, A.C. Halbower, J. Jones, C. Lehmann, M.S. Schechter, S. Sheldon, R.N. Shiffman, K. Spruyt, Diagnosis and management of childhood obstructive sleep apnea syndrome. Pediatrics **130**, 576–584 (2012)
20. F.C. Motta, P.D. Shipman, R.M. Bradley, Highly ordered nanoscale surface ripples produced by ion bombardment of binary compounds. J. Phys. D Appl. Phys. **45**(12), 122001 (2012)
21. D. Papageorgiou, Y. Smyrlis, The route to chaos for the Kuramoto-Sivashinsky equation. Theor. Comput. Fluid Dyn. **3**(1), 15–42 (1991)
22. D. Rolfsen, *Knots and Links*. Sistema Librum 2.0 (SERBIULA, 2018), pp. 391–415
23. G.I. Sivashinsky, Instabilities, pattern formation, and turbulence in flames. Annu. Rev. Fluid Mech. **15**(1) 179–199 (1983)
24. H.L. Swinney, Emergence and evolution of patterns, in *Proceedings of AIP* (1999), pp. 3–22
25. T. Tiedje, K.A. Mitchell, B. Lau, A. Ballestad, E. Nodwell, Radiation transport model for ablation hollows on snowfields. J. Geophys. Res. **11** (F2) 2156–2202 (2006)
26. C.M. Topaz, L. Ziegelmeier, T. Halversion, Topological data analysis of biological aggregation models. PLoS ONE **10**(5), e0126383 (2015)
27. J. Villain, Continuum models of crystal growth from atomic beams with and without desorption. J. Phys. I Fr. **1** 19–42 (1991)
28. T. Wanner, Topological analysis of the diblock copolymer equation, in *Mathematical Challenges in a New Phase of Materials Science* (Springer, Berlin, 2016), pp. 27–51

29. K. Xia, G.-W. Wei, Multidimensional persistence in biomolecular data. J. Comput. Chem. **36**, 1502–1520 (2015)
30. M. Younes, W. Thompson, C. Leslie, T. Egan, E. Giannouli, Utility of technologist editing of polysomnography scoring performed by a validated automatic system. Ann. Am. Thorac. Soc. **12**, 1206–18 (2015)