



A Persistent Homology Approach to Heart Rate Variability Analysis With an Application to Sleep-Wake Classification

Yu-Min Chung^{1*}, Chuan-Shen Hu², Yu-Lun Lo³ and Hau-Tieng Wu^{4,5}

¹ Department of Mathematics and Statistics, University of North Carolina at Greensboro, Greensboro, NC, United States, ² Department of Mathematics, National Taiwan Normal University, Taipei, Taiwan, ³ Department of Thoracic Medicine, Chang Gung Memorial Hospital, Chang Gung University, School of Medicine, Taipei, Taiwan, ⁴ Department of Mathematics and Department of Statistical Science, Duke University, Durham, NC, United States, ⁵ Mathematics Division, National Center for Theoretical Sciences, Taipei, Taiwan

OPEN ACCESS

Edited by:

Fernando Soares Schindwein,
University of Leicester,
United Kingdom

Reviewed by:

Rocio Gonzalez-Diaz,
Sevilla University, Spain
Chengyuan Wu,
Agency for Science, Technology and
Research, Singapore
Jie Wu,

Hebei Normal University, China
Jonathan Jaquette,
Boston University, United States

*Correspondence:

Yu-Min Chung
y_chung2@uncg.edu

Specialty section:

This article was submitted to
Computational Physiology and
Medicine,
a section of the journal
Frontiers in Physiology

Received: 04 December 2020

Accepted: 05 February 2021

Published: 01 March 2021

Citation:

Chung Y-M, Hu C-S, Lo Y-L and Wu H-T (2021) A Persistent Homology Approach to Heart Rate Variability Analysis With an Application to Sleep-Wake Classification. *Front. Physiol.* 12:637684. doi: 10.3389/fphys.2021.637684

Persistent homology is a recently developed theory in the field of algebraic topology to study shapes of datasets. It is an effective data analysis tool that is robust to noise and has been widely applied. We demonstrate a general pipeline to apply persistent homology to study time series, particularly the instantaneous heart rate time series for the heart rate variability (HRV) analysis. The first step is capturing the shapes of time series from two different aspects—the persistent homologies and hence persistence diagrams of its sub-level set and Taken's lag map. Second, we propose a systematic and computationally efficient approach to summarize persistence diagrams, which we coined *persistence statistics*. To demonstrate our proposed method, we apply these tools to the HRV analysis and the sleep-wake, REM-NREM (rapid eyeball movement and non rapid eyeball movement) and sleep-REM-NREM classification problems. The proposed algorithm is evaluated on three different datasets via the cross-database validation scheme. The performance of our approach is better than the state-of-the-art algorithms, and the result is consistent throughout different datasets.

Keywords: persistent homology, persistence diagram, persistence statistics, sleep stage, heart rate variability

1. INTRODUCTION

Heart rate variability (HRV) is the physiological phenomenon of variation in the lengths of consecutive cardiac cycles, or the rhythm of heart rate (Draghici and Taylor, 2016). Interest in HRV has a long history (Billman, 2011), and there have been several theories describing how the heart rate rhythm, including, for example, the polyvagal theory (Porges, 2009) and the model of neurovisceral integration (Thayer and Sternberg, 2006). In short, HRV results from an integration of complicated interactions between various physiological systems and external stimuli (Vanderlei et al., 2009; Shaffer et al., 2014; Draghici and Taylor, 2016) on various scales, and the autonomic nervous system (ANS) plays a critical role (Thayer and Sternberg, 2006; Porges, 2009). A correct quantification of HRV yields dynamical information of various physiological systems and has various clinical applications (Stys and Stys, 1998), including improving diagnostic accuracy and treatment quality (Vanderlei et al., 2009).

In practice, the heart rhythm is quantified by the time series called *instantaneous heart rate* (IHR) coming from intervals between consecutive pairs of heart beats, which is usually determined from the R peak to R peak interval (RRI) by reading the electrocardiogram (ECG). See **Figure 1** for an illustration of the ECG, R peaks, and RRI. To quantify HRV, a common approach is studying various statistics of IHR. There have been a lot of efforts trying to quantify HRV, and proposed statistics could be briefly classified into four major categories—time domain approach, frequency domain approach (Task Force of the European Society of Cardiology and others, 1996), nonlinear geometric approach (Marwan et al., 2002; Voss et al., 2008), and information theory based approach (Costa et al., 2002). It is worth mentioning that while there have been a lot of researches in this direction with several proposed statistics, there is limited consensus and it is still an active research field due to the non-stationarity nature of the IHR time series (Pincus and Goldberger, 1994; Glass, 2009).

Topological data analysis (TDA) is a data analysis framework based on tools from algebraic topology (Carlsson, 2009; Epstein et al., 2011). In the past decades, its theoretical foundation has been actively established, and various algorithms have been proposed to study datasets from different fields. The basic idea underlying TDA is that the data organization can be well-captured by *counting holes*. Theoretically, the number of holes of different dimensions characterizes how the data is organized. Thus, researchers design useful statistics based on the information of holes. This simple yet powerful idea has been applied to different fields. Specifically, there have been several efforts applying TDA to analyze time series. For example, the Vietoris-Rips (VR) complex filtration and the bottleneck or Wasserstein distances among persistence diagrams are applied to study voices and body motions (Seversky et al., 2016; Venkataraman et al., 2016). A transformation of the persistence diagram, called persistence landscapes (Bubenik, 2015), has been applied to study trading records (Gidea and Katz, 2018), electroencephalogram (EEG) signals (Piangerelli et al., 2018; Wang et al., 2018; Wang et al., 2019), and cryptocurrency trend forecasting (Kim et al., 2018). Sliding Windows and 1-Persistence Scoring (Perea, 2019) offers both theoretical and practical TDA method to detect the periodicity of a time series. A brief overview of common techniques on the usages of TDA to time series is summarized in a preprint (Ravishanker and Chen, 2019). Recently, the proposed TDA tool for HRV analysis has been applied to differentiating patients with the history of transient ischemic attack and hypertension (Graff et al.,

2020). However, existing TDA approaches usually suffer from computational issues, which limits its application to large scale database. Finding a computationally efficient TDA algorithm is thus critical.

In this article, motivated by the complicated interaction among different physiological systems over various scales and inter-individual variability, the need for a useful tool for the HRV analysis, and the numerical limitation of the recently developed TDA tools, we hypothesize that topological information could be useful to quantify the HRV, and propose a computationally efficient approach to analyze time series via TDA.

1.1. Our Contribution

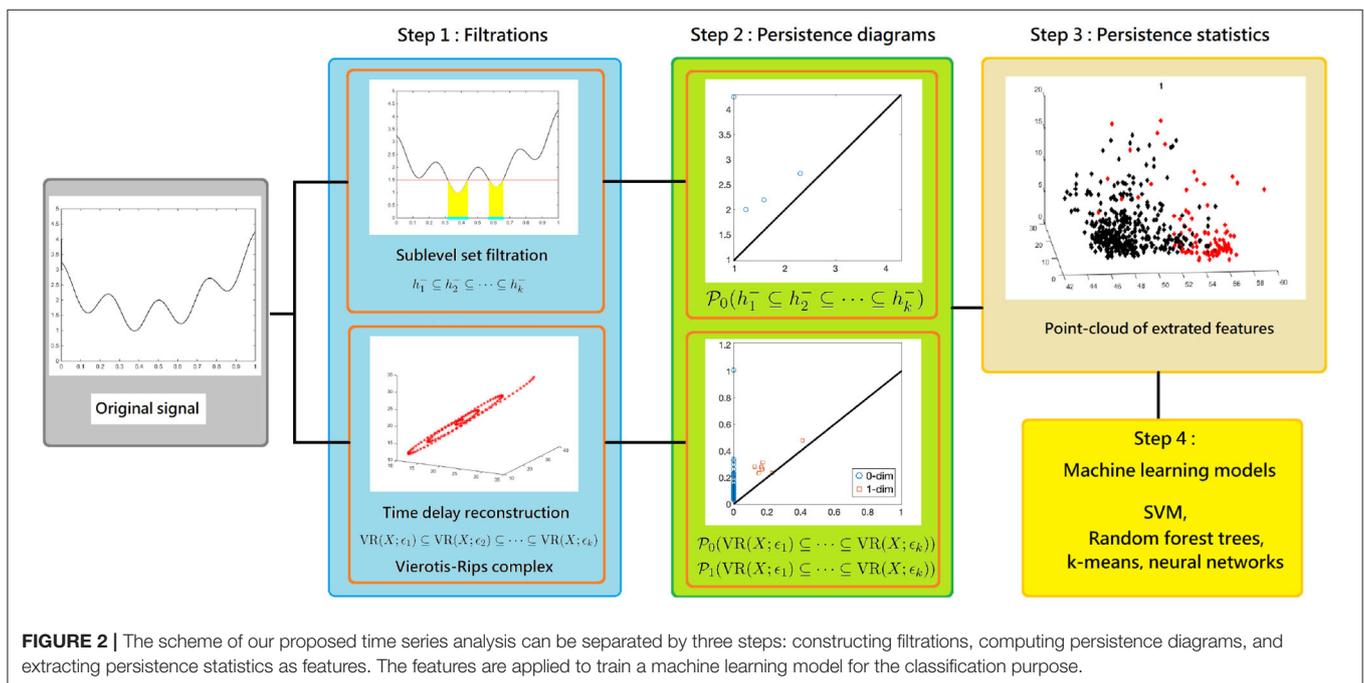
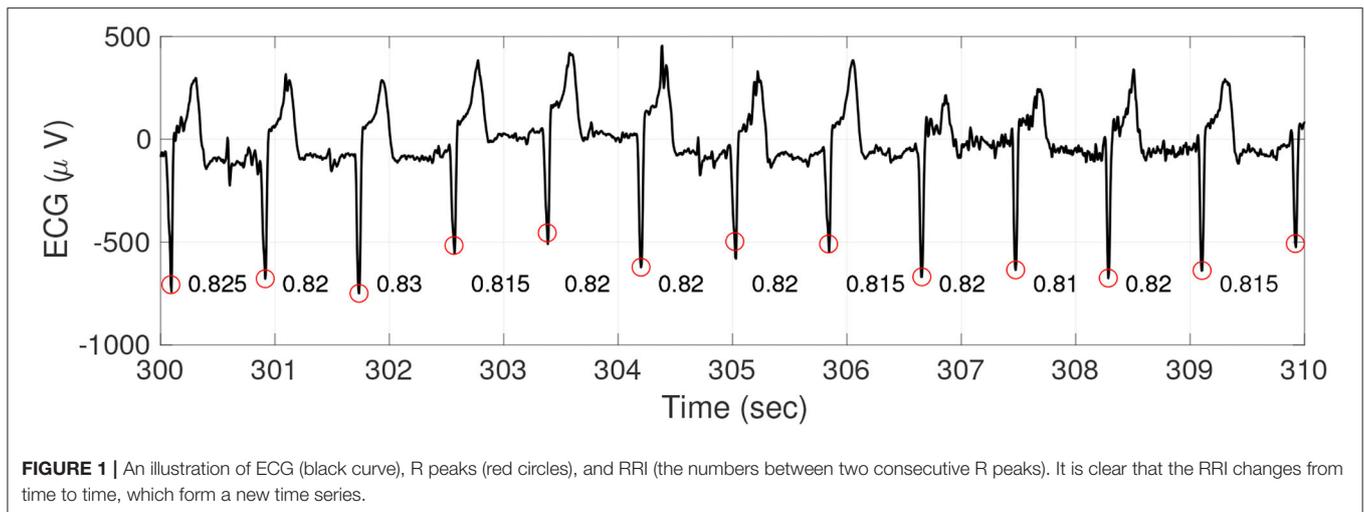
Based on the flexibility of TDA tools, and due to the non-stationarity of complicated time series we commonly encounter in real life, like the IHR, we propose a systematic, principled, and computationally efficient approach to study complicated time series by the TDA tools.

Our main scheme for studying a complicated time series is shown in **Figure 2**, which is divided into three steps that we will detail later. First, consider two *filtrations*, the Vietoris-Rips (VR) complex filtration of the Takens' lag map (Takens, 1981) and the sub-level set filtration of the time series, and *persistent homology*. Second, compute corresponding persistence diagrams. Finally, calculate *persistence statistics* (PS) as a novel statistic of the time series of interest. We mention that compared with existing TDA approach for time series analysis, our proposed persistence statistics features based on both sub-level set and Vietoris-Rips complexes filtrations are intuitive, straightforward to implement, and also computationally efficient.

1.2. Application—Sleep Dynamics

To demonstrate the usefulness of the proposed persistence statistics, we apply it to study IHR time series recorded during sleep, and use obtained statistics to classify sleep stages. Sleep is a universal recurrent physiological phenomenon. Sleep impacts the whole body, so we can read sleep via reading different physiological signals. Taking ECG into account is specifically attractive, since the ECG sensor is easy to install, and it is now widely available in mobile health devices. HRV of a subject is usually quantified by analyzing ECG, and it has been shown to be related to sleep dynamics (Zemaityte and Varoneckas, 1984; Vaughn et al., 1995; Toscani et al., 1996; Bonnet and Arand, 1997; Elsenbruch et al., 1999; Chouchou and Desseilles, 2014; Penzel et al., 2016). In other words, the heart rate rhythm provides a non-invasive window for researchers to study sleep. While there have been several studies trying to classify sleep stages based on HRV (Lewicke et al., 2008; Mendez and Matteucci, 2010; Long et al., 2012; Xiao et al., 2013; Aktaruzzaman et al., 2015; Ye et al., 2016; Malik et al., 2018), it still remains a challenging problem in the field. The challenge and difficulty of this mission can be appreciated from the reported results. In this article, we apply the proposed persistence statistics to quantify HRV during sleep, and propose a new prediction algorithm for the sleep stage; for example, an automatic classification of wake and sleep, REM and NREM, and wake, REM and NREM. We remark that while we focus on the HRV and sleep stage classification, the result

Abbreviations: AASM, American academy of sleep medicine; Acc, accuracy; AHI, apnea-hypopnea index; ANS, autonomic nervous system; AUC, area under curve; CGMH, Chang Gung Memorial Hospital; CV, cross validation; DREAMS, DREAMS subjects database; ECG, electrocardiogram; ECOC, error-correcting output codes; EEG, electroencephalogram; EOG, electrooculogram; EMG, electromyogram; HRV, heart rate variability; IHR, instantaneous heart rate; LOSOCV, leave-one-subject-out CV; NREM, non-rapid eyeball movement; PPG, photoplethysmography; PR, precision; PS, persistence statistics; PSG, polysomnogram; REM, rapid eyeball movement; RRI, R peak to R peak interval; SE, sensitivity; SP, specificity; SVM, support vector machine; TDA, topological data analysis; UCDSADB, St. Vincent's University Hospital/University College Dublin Sleep Apnea Database; VR, vietoris-rips.



indicates the potential of applying TDA-based approaches to study other complicated time series.

1.3. Organization

In section 2, we review the mathematical background of the persistent homology and persistence diagram. In section 3, we demonstrate two ways to use the persistent homology to study time series, and propose a new approach to summarize the persistence diagram, called the persistence statistics. The classification model based on the persistence statistics for the sleep stage classification will be discussed in detail in section 4. The discussion of our classification performance and a comparison with the state-of-arts results will be included in section 5. More technical details and numerical results are postponed to the **Supplementary Material**.

2. MATHEMATICAL BACKGROUND

In this section, we describe the mathematical background, including simplicial complex, homology, filtration of sets, and the persistent homology. Although these topics can be studied in an abstract and general way (see e.g., Munkres, 1993), to enhance the readability, we present them in a relatively concrete way without losing critical information.

2.1. Simplicial Complexes

To investigate the complicated structure of an object, an intuitive way is to use simple objects as building blocks to approximate the original object. In TDA, the main building block is the *simplicial complex*, which we briefly recall now. See **Supplementary Material** (section 1.1) for more detailed mathematical background and illustrative examples.

We start with the *simplex*. Intuitively, a simplex is a “triangle” of different dimension. Let x_0, x_1, \dots, x_q be affinely independent points in \mathbb{R}^d , where $d, q \in \mathbb{N}$ and $d \geq q$. The q -*simplex*, denoted by $\sigma := \langle x_0, x_1, \dots, x_q \rangle$, is defined to be the convex hull of x_0, x_1, \dots, x_q . Denote $\text{Vert}(\sigma) := \{x_0, x_1, \dots, x_q\}$. Any q -simplex is a q -dimensional object consisting of lower degree simplexes. We are interested in the relation among simplexes of different dimensions. Since any $V \subset \text{Vert}(\sigma)$ is also affinely independent, the convex hull of V , called a *face* of σ , forms a simplex of dimension $|V| \leq q$, where $|V|$ is the cardinality of V . If $|V| = k$, the face $\tau = \langle V \rangle$ is called a k -face of σ . A *simplicial complex* \mathcal{K} in \mathbb{R}^d is a collection of finite simplexes σ in \mathbb{R}^d so that any intersection of two arbitrary simplexes is a face to each of them; that is,

- If $\sigma \in \mathcal{K}$ and τ is a face of σ , then $\tau \in \mathcal{K}$;
- If $\sigma_1, \sigma_2 \in \mathcal{K}$, then $\sigma_1 \cap \sigma_2$ is a face of σ_1 and σ_2 . In particular, $\sigma_1 \cap \sigma_2 \in \mathcal{K}$.

2.2. Homology and Betti Numbers

In order to study the topological information of a given simplicial complex, we study relations among simplexes of different dimensions, and hence the “holes.” *Homology* and *Betti numbers* are classic subjects in the algebraic topology (Munkres, 1993), which capture “holes” of geometric objects of different dimensions. While we can discuss these topics in a more general setup, in this work, we mainly consider simplicial complexes as our target object. For example, the shape of the notation “ ∞ ” contains two 1-dimensional holes (Supplementary Figure 2C) and the empty void surrounded by the unit sphere $S^2 = \{(x, y, z) : x^2 + y^2 + z^2 = 1\}$ in \mathbb{R}^2 is a 2-dimensional hole (Supplementary Figure 2E). Moreover, the 0-dimensional holes of an object are defined to be its disjoint connected components (Supplementary Figure 2E). See Supplementary Material (section 1.2) for more information and illustrative examples.

We need an algebraic structure of simplexes. Given q -simplexes $\sigma_1, \sigma_2, \dots, \sigma_n$ in a simplicial complex \mathcal{K} , define the sum over \mathbb{Z}_2 as $c = \sum_{i=1}^n v_i \sigma_i$, where $v_i \in \mathbb{Z}_2$. This formal sum is commonly known as a q -chain. One could also define an addition operator as $\sum_{i=1}^n v_i \sigma_i + \sum_{i=1}^n \mu_i \sigma_i := \sum_{i=1}^n (v_i + \mu_i) \sigma_i$. We consider the collection of all q -chains, denoted as

$$C_q(\mathcal{K}) := \left\{ \sum_{i=1}^n v_i \sigma_i \mid v_i \in \mathbb{Z}_2, \sigma_i \in \mathcal{K}, \dim(\sigma_i) = q \right\}.$$

One could prove that $C_q(\mathcal{K})$ is actually a vector space over \mathbb{Z}_2 with the above addition. There is a natural relation between $C_q(\mathcal{K})$ and $C_{q-1}(\mathcal{K})$, called the *boundary map* (Munkres, 1993, section 1.5, p. 30). The q th *boundary map* $\partial_q : C_q(\mathcal{K}) \rightarrow C_{q-1}(\mathcal{K})$ over \mathbb{Z}_2 is defined by

$$\partial_q(\langle x_0, x_1, \dots, x_q \rangle) := \sum_{i=0}^q \langle x_0, \dots, \hat{x}_i, \dots, x_q \rangle,$$

where $\langle x_0, x_1, \dots, x_q \rangle \in \mathcal{K}$ and the $\hat{\bullet}$ denotes the drop-out operation. With the boundary maps, there is a nested relation

among chains

$$\dots \xrightarrow{\partial_{n+1}} C_n(\mathcal{K}) \xrightarrow{\partial_n} C_{n-1}(\mathcal{K}) \xrightarrow{\partial_{n-1}} \dots C_1(\mathcal{K}) \xrightarrow{\partial_1} C_0(\mathcal{K}).$$

This nested relation among chains is known as the *chain complex*, which is denoted as $\mathcal{C} = \{C_q, \partial_q\}_{q \in \mathbb{Z}}$.

A fundamental result in the homology theory (Munkres, 1993 Lemma 5.3 section 1.5, p. 30) is that the composition of any two consecutive boundary maps is trivial, i.e., $\partial_{q-1} \circ \partial_q = 0$. This result allows one to define the following quotient space. Denote *cycles* and *boundaries* by $Z_q(\mathcal{K})$ and $B_q(\mathcal{K})$, respectively, which are defined as

$$\begin{aligned} Z_q(\mathcal{K}) &:= \ker(\partial_q) = \{c \in C_q \mid \partial_q(c) = 0\}, B_q(\mathcal{K}) := \text{im}(\partial_{q+1}) \\ &= \{\partial_{q+1}(z) \in C_q \mid z \in C_{q+1}\}. \end{aligned}$$

Note that each $B_q(\mathcal{K})$ is a subspace of $Z_q(\mathcal{K})$ since $\partial_{q-1} \circ \partial_q = 0$. Therefore, we can define the q th *homology* to be the quotient space

$$H_q(\mathcal{K}) := \frac{Z_q(\mathcal{K})}{B_q(\mathcal{K})} = \frac{\ker(\partial_q)}{\text{im}(\partial_{q+1})}, \tag{1}$$

which is again a vector space. For instance, if $\mathcal{K} = \mathcal{K}_3$ in Figure 3, then $H_0(\mathcal{K}_3) \simeq \mathbb{Z}_2$ and $H_1(\mathcal{K}_3) \simeq \mathbb{Z}_2$ because it contains one connected component and one 1-dimensional hole. More precisely, the 1-dimensional hole in \mathcal{K}_3 is represented by the 1-cycle

$$c = \langle v_1, v_2 \rangle + \langle v_2, v_3 \rangle + \langle v_3, v_1 \rangle.$$

Actually, a q -hole may be represented by different q -cycles. For example, if $\mathcal{K} = \mathcal{K}_3$ in Figure 3, then the 1-cycles $c = \langle v_1, v_2 \rangle + \langle v_2, v_3 \rangle + \langle v_3, v_1 \rangle$ and $d = \langle v_1, v_4 \rangle + \langle v_4, v_2 \rangle + \langle v_2, v_3 \rangle + \langle v_3, v_1 \rangle$ represent the same 1-dimensional hole in \mathcal{K}_4 because $c + d = \langle v_1, v_2 \rangle + \langle v_1, v_4 \rangle + \langle v_4, v_2 \rangle \in \text{im}(\partial_2)$ is the boundary of the 2-simplex $\langle v_1, v_2, v_4 \rangle$. This gives us an intuition about the algebraic structure (1).

The q th *Betti number* is defined to be the dimension of the q th homology; that is,

$$\beta_q(\mathcal{K}) = \dim(H_q(\mathcal{K})), \tag{2}$$

which measures the number of q -dimensional holes. As a result, given a simplicial complex \mathcal{K} , the *homology* of \mathcal{K} is a collection of vector spaces $\{H_q(\mathcal{K})\}_{q=0}^\infty$, and its *Betti numbers* is denoted as $\beta(\mathcal{K}) := \{\beta_q(\mathcal{K})\}_{q=0}^\infty$.

2.3. Persistent Homology

We now introduce a natural generalization of homology, the *persistent homology*, that is suitable for data analysis. *Persistent homology* is more suitable for data analysis than homology due to this capability of dealing with inevitable noise in real world dataset. It depends on the notion of *filtration* to handle noise. In general, filtration is a sequence of simplicial complexes (see Figure 3 for an example). We are interested in how the “holes” vary in the filtration. Intuitively, if certain holes are “robust,” they will survive in the filtration.

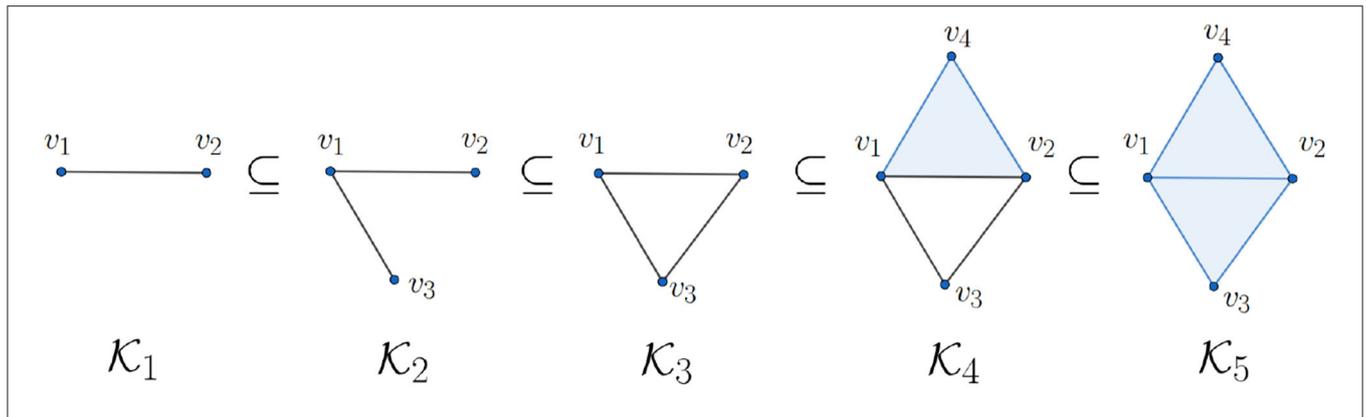


FIGURE 3 | Illustration of a filtration of simplicial complexes $\mathcal{K}_1, \mathcal{K}_2, \mathcal{K}_3, \mathcal{K}_4,$ and \mathcal{K}_5 .

Definition 1 (Edelsbrunner and Harer, 2010 section 3.4, p. 70). For an index set I , a filtration is a sequence of simplicial complexes, $\{\mathcal{K}_t\}_{t \in I}$, satisfying

$$\mathcal{K}_{t_1} \subseteq \mathcal{K}_{t_2}, \text{ whenever } t_1 \leq t_2. \tag{3}$$

From the previous discussion, for each \mathcal{K}_t in a filtration, one could compute its homology group and Betti number. Because of the nested subset relation in a filtration, there exist relations among simplicial complexes. This allows one to track and record the changes of the homology group and the Betti numbers, which we detail now. Given a fixed $q \geq 0$, each \mathcal{K}_i induces homology $H_q(\mathcal{K}_i)$. Denote $\iota_i: \mathcal{K}_i \rightarrow \mathcal{K}_{i+1}$ to be the inclusion map. Then $\iota_i(Z_q(\mathcal{K}_i)) \subseteq Z_q(\mathcal{K}_{i+1})$ and $\iota_i(B_q(\mathcal{K}_i)) \subseteq B_q(\mathcal{K}_{i+1})$ (Edelsbrunner and Harer, 2010 section 4.4, p. 95). Therefore, the mapping

$$\bar{\iota}_i: H_q(\mathcal{K}_i) \rightarrow H_q(\mathcal{K}_{i+1}), \bar{c} \mapsto \overline{\iota_i(c)} \tag{4}$$

induced by ι_i is a well-defined linear transformation over \mathbb{Z}_2 . We also define a linear transformation

$$\rho_q^{i,i+k} := \overline{\iota_{i+k-1}} \circ \dots \circ \overline{\iota_{i+1}} \circ \bar{\iota}_i, \tag{5}$$

which maps $H_q(\mathcal{K}_i)$ to $H_q(\mathcal{K}_{i+k})$. The following definition is crucial for defining *lifespans* of connected components or holes in homology theory.

Definition 2 (Edelsbrunner and Harer, 2010 section 7.1, p. 151). Let $\{\mathcal{K}_i\}_{i=0}^n$ be a filtration of simplicial complexes. For $q \in \mathbb{Z}_{\geq 0}$ and $i, j \in \mathbb{Z}_{\geq 0}$ with $i \leq j$, we define the persistent homology as

$$H_q^{i,j} := \frac{Z_q(\mathcal{K}_i)}{B_q(\mathcal{K}_j) \cap Z_q(\mathcal{K}_i)}. \tag{6}$$

Since $\mathcal{K}_0 \subseteq \mathcal{K}_1 \subseteq \dots \subseteq \mathcal{K}_n$, we have inclusions of q -chains: $C_q(\mathcal{K}_0) \subseteq C_q(\mathcal{K}_1) \subseteq \dots \subseteq C_q(\mathcal{K}_n)$ for all $q \geq 0$. Hence, the intersection $B_q(\mathcal{K}_j) \cap Z_q(\mathcal{K}_i)$ is a well-defined subspace of $Z_q(\mathcal{K}_i)$. Moreover, for $i \leq j$, the kernel of the linear transformation

$$\phi_q^{i,j}: Z_q(\mathcal{K}_i) \rightarrow \frac{Z_q(\mathcal{K}_j)}{B_q(\mathcal{K}_j)}, c \mapsto \bar{c} = c + B_q(\mathcal{K}_j)$$

induced by the inclusion map is $B_q(\mathcal{K}_j) \cap Z_q(\mathcal{K}_i)$. By the first isomorphism theorem, we obtain an injective linear transformation

$$\overline{\phi_q^{i,j}}: \frac{Z_q(\mathcal{K}_i)}{B_q(\mathcal{K}_j) \cap Z_q(\mathcal{K}_i)} \rightarrow \frac{Z_q(\mathcal{K}_j)}{B_q(\mathcal{K}_j)}.$$

Via the one-to-one linear mapping $\overline{\phi_q^{i,j}}$, the vector space $H_q^{i,j}$ may be viewed as a subspace of $H_q(\mathcal{K}_j)$. In particular, if $i = j$, then $H_q^{i,i} = H_q(\mathcal{K}_i) = H_q(\mathcal{K}_j)$, which means that the persistent homology is a generalization of the homology. With the inclusion $H_q^{i,j} \hookrightarrow H_q(\mathcal{K}_j)$, we define the moments of *birth* and *death* of a “hole” in the filtration.

Definition 3 (Edelsbrunner and Harer, 2010 section 7.1, p. 151). Let $\{\mathcal{K}_i\}_{i=0}^n$ be a filtration of simplicial complexes and $q \in \mathbb{Z}_{\geq 0}$.

- A q -hole \bar{c} ($c \in Z_q(\mathcal{K}_i)$) is **born** at \mathcal{K}_i if $\bar{c} \in H_q(\mathcal{K}_i) \setminus \{0\}$, but $\bar{c} \notin \text{im}(\rho_q^{i-1,i})$;
- A q -hole \bar{c} ($c \in Z_q(\mathcal{K}_i)$) **dies** at \mathcal{K}_j if $\rho_q^{i,j-1}(\bar{c}) \notin H_q^{i-1,j-1}$, but $\rho_q^{i,j}(\bar{c}) \in H_q^{i-1,j}$.

The death $d = \infty$ means that the q -hole never dies in the filtration.

We use **Figure 3** to explain the relation between these two abstract definitions. For instance, the non-trivial element \bar{c} represented by 1-chain $c = \langle v_1, v_2 \rangle + \langle v_2, v_3 \rangle + \langle v_3, v_1 \rangle$ in $H_1(\mathcal{K}_3)$ is born at \mathcal{K}_3 i.e., $\bar{c} \notin \text{im}(\rho_1^{2,3})$ because $H_1(\mathcal{K}_2) = \{0\}$ and $\mathbb{Z}_2 = H_1(\mathcal{K}_3) = \text{span}_{\mathbb{Z}_2}\{\bar{c}\}$. On the other hand, the fact $\{0\} \subseteq H_1^{2,5} \subseteq H_1(\mathcal{K}_5) = \{0\}$ shows that $\rho_1^{3,5}(\bar{c}) \in H_1(\mathcal{K}_5) = H_1^{2,5}$ and $\rho_1^{3,4}(\bar{c}) \notin H_1^{2,4}$ because $H_1^{2,4} = \{0\}$ (since $Z_1(\mathcal{K}_2) = \{0\}$) and $\mathbb{Z}_2 = H_1(\mathcal{K}_4) = \text{span}_{\mathbb{Z}_2}\{\rho_1^{3,4}(\bar{c})\}$, thus \bar{c} dies at \mathcal{K}_5 . We refer readers with interest to Edelsbrunner and Harer (2010) for more details in persistent homology.

2.4. Persistence Diagram

Persistence diagram proposed in Edelsbrunner et al. (2000) or equivalently persistence barcodes proposed in Carlsson et al. (2005) is a tool to visualize the complicated lifespans of holes in

a given filtration for data analysis. We use persistence diagram in this paper.

The persistence diagram possesses the desired *stability property* (Cohen-Steiner et al., 2007)—a bounded perturbation of a given filtration leads to a bounded perturbation of the corresponding persistence diagram. Due to the inevitable noise in real data, this stability property renders persistence diagram based approaches suitable for data analysis. The bottleneck and Wasserstein distances (Cohen-Steiner et al., 2007) are typical ways to measure differences among persistence diagrams. The formal statements of the stability property based on these two distances are provided in sections 3.1, 3.2. We refer readers with interest to Edelsbrunner and Harer (2010) for details in persistence diagram.

Definition 4 (Edelsbrunner and Harer, 2010 section 7.1, p. 152). Let $\{\mathcal{K}_i\}_{i=0}^n$ be a filtration of simplicial complexes and $q \in \mathbb{Z}_{\geq 0}$. The q th persistence diagram, denoted as $\mathcal{P}_q(\{\mathcal{K}_i\}_{i=0}^n)$, of the filtration is the multiset of q -dimension holes in the filtration.

In other words, a q -dimensional hole in a filtration is recorded by a pair (b, d) of integers where b and d are called the *birth* and *death* of the hole, respectively (Edelsbrunner and Harer, 2010). Although the above definition of persistence diagram seems technical, its interpretation is intuitive. For instance, consider the filtration shown in **Figure 3**. We look for the “changes” of topological structure (holes). Note that since a connected component is born at \mathcal{K}_1 (specifically, $\langle v_1, v_2 \rangle$), its birth value is $b = 1$; since it lives throughout the filtration, its death value is ∞ . We now turn our focus to the 1-dimensional hole. Note that a 1-dimensional hole (specifically, $\langle v_1, v_2 \rangle + \langle v_2, v_3 \rangle + \langle v_1, v_3 \rangle$) is formed at \mathcal{K}_3 , so its birth value is 3; note also that this hole is filled at \mathcal{K}_5 , so its death value is 5. Since there is no more change of holes, we have the persistence diagrams $\mathcal{P}_0(\{\mathcal{K}_i\}_{i=1}^5) = \{(1, \infty)\}$ and $\mathcal{P}_1(\{\mathcal{K}_i\}_{i=1}^5) = \{(3, 5)\}$.

Before closing this subsection, we illustrate how persistent homology and persistence diagram work by taking a noisy point cloud sampled from a circle contaminated by Gaussian noise shown in **Figure 4A**. If there is no noise, the 1st Betti number of the circle is $\beta_1 = 1$. In the noisy case, the Betti number information is contained in the form of the persistence diagram as shown in **Figure 4B**, where each point represents one 1-dimensional hole associated with its birth and death value. In **Figure 4B**, we observe that there is an outstanding point with long lifespan (located around birth value 0.05 and death value 0.25), while lifespans for other points are very small. This suggests that the noisy point cloud has a strong/robust 1-dimensional hole. This captures the main topology information, $\beta_1 = 1$, about this data.

2.5. Data Analysis With Persistence Diagram and Commonly Considered TDA Statistics

Usually, researchers design statistics on the persistence diagram of a given dataset via the chosen filtration. One basic result supporting this approach is Mileyko et al. (2011), where authors showed that the space of persistence diagrams with certain

metric is complete and separable. This result forms a theoretic foundation for any statistical methods. In Fasy et al. (2014) and Blumberg et al. (2014), authors derived confidence sets of persistence diagrams in order to separate the long lifespan holes from noisy ones, and also proposed four ways to estimate them. While these theoretical results shed light on applying TDA to analyze complex data, however, any operation in the space of persistence diagrams is complicated and difficult to compute. For example, computing bottleneck or Wasserstein distances among persistence diagrams is a difficult task and can be time consuming, even for the state-of-art algorithm (Kerber et al., 2017). Another result indicates that the mean in the space of persistence diagrams may not be unique (Turner et al., 2014a). This computational burden renders it less applicable to data analysis.

To get around the computational issue when working with those distances, one major approach is to “vectorize” persistence diagrams; that is, researchers map the space of persistence diagrams into another space. For example, persistence landscapes (Bubenik, 2015) map persistence diagrams into a Banach space, specifically L^p space. More examples include persistence image (Adams et al., 2017), generalized persistence landscapes (Berry et al., 2020), persistence path (Chevyrev et al., 2018), persistence codebook (Zelinski et al., 2020), persistence curves (Chung and Lawson, 2019), kernel based methods (Reininghaus et al., 2015; Kusano et al., 2016), and persistent entropy (Chintakunta et al., 2015; Atienza et al., 2019b). These methods have been studied and applied to different applications. In **Figure 5**, we provide a chart depicting the relationship among existing TDA tools. We mention that the proposed persistence statistics in section 3 could be viewed as a computationally efficient vectorization of persistence diagrams.

3. TDA FOR TIME SERIES ANALYSIS AND FEATURES EXTRACTION

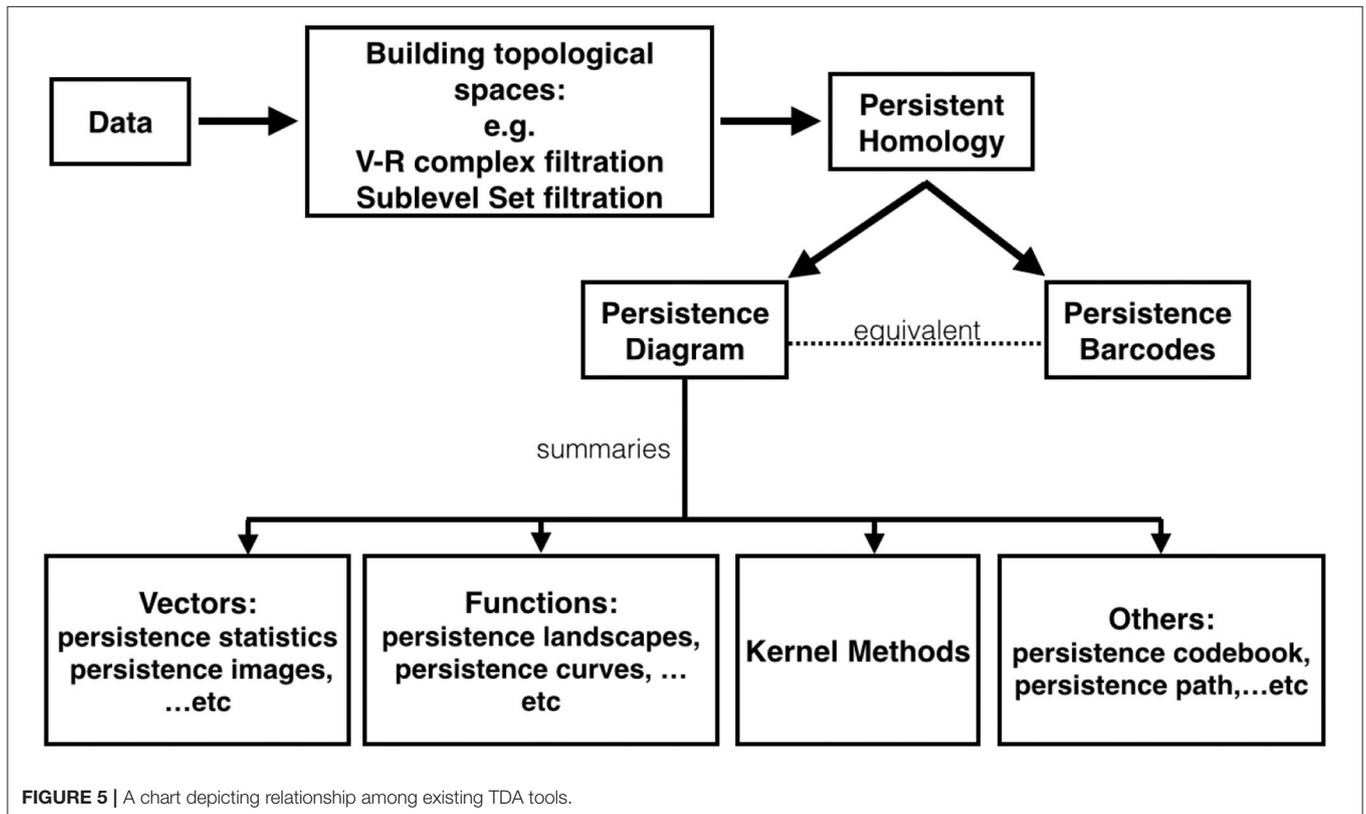
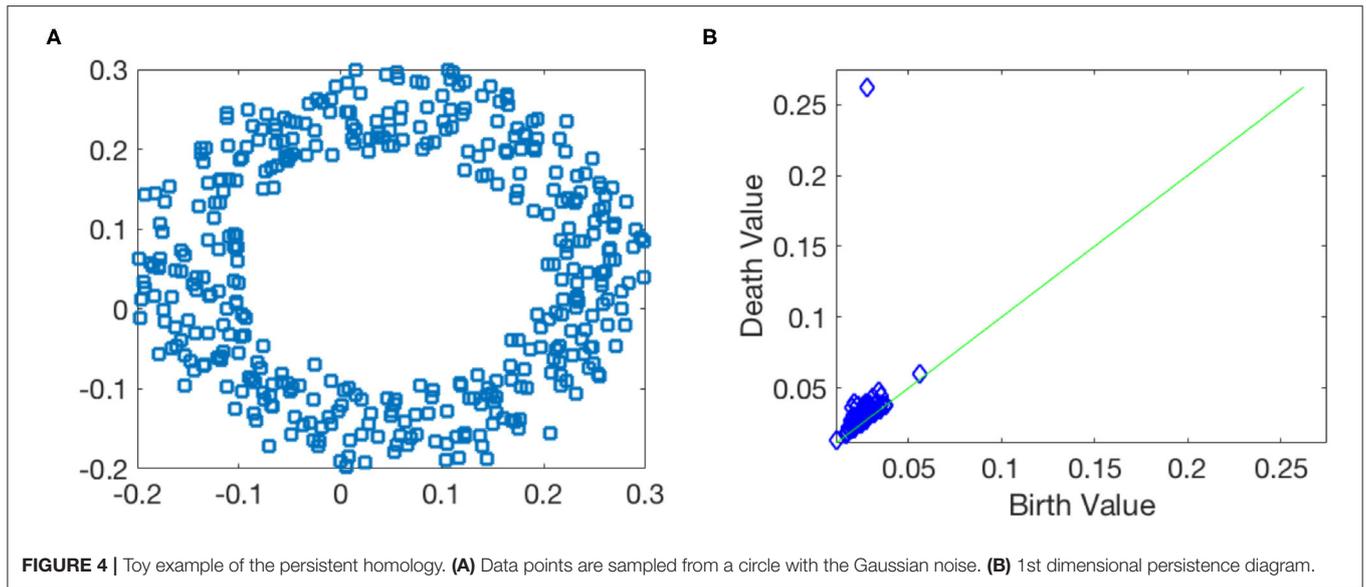
Armed with the theoretical background in section 2, we are ready to describe how to apply TDA for time series analysis. To apply the persistent homology to analyze complicated time series, we introduce two useful filtrations, the sub-level set filtration and the Vietoris-Rips complexes filtration. With these two filtrations, we introduce a novel features extraction methods, coined persistence statistics, based on the persistence diagrams of the sub-level set filtration and the Vietoris-Rips complexes filtration.

3.1. First Useful Filtration—Sub-Level Set Filtration

To simplify the discussion and illustrate the idea, we identify a time series as a discretization of a continuous function $f: [0, T] \rightarrow \mathbb{R}$, where T is some fixed constant. For each $h \in \mathbb{R}$, the *sub-level set* of f is defined as

$$f_h := f^{-1}((-\infty, h]) = \{t \in [0, T] \mid f(t) \leq h\}. \quad (7)$$

Clearly, $f_{h_1} \subseteq f_{h_2}$ whenever $h_1 \leq h_2$. Therefore, for any increasing sequence $\{h_i\}_i$, the collection of sub-level sets, $\{f_{h_i}\}_i$, forms a filtration. Intuitively, the sub-level set filtration reveals



the *oscillating information* of the functions. Since each f_h is a subset of $[0, T] \subseteq \mathbb{R}$, it only contains 0-dimensional structures, i.e., connected components. Hence, the only non-trivial persistence diagram in this case is \mathcal{P}_0 . For simplicity, when there is no danger of confusion, for a given function f , we use $\mathcal{P}_0(f)$ to denote $\mathcal{P}_0(\{f_{h_i}\})$, the persistence diagram associated with the sub-level sets filtration of f . As discussed

in Edelsbrunner and Harer (2010), each element in \mathcal{P}_0 is a min-max pair in the original function $f(t)$. The concept of this filtration is closely related to the size function theory (see Biasotti et al., 2008 and references therein) and is commonly used as a shape descriptor (Biasotti et al., 2008). In practice, persistence diagram is robust to noise under the *bottleneck distance*. This fact renders persistence diagram an useful

data analysis quantity. A precise statement of this *robustness* is below.

Theorem 3.1.1. *Let X be an n -dimensional rectangle in \mathbb{R}^n . Take two continuous functions $f, g: X \rightarrow \mathbb{R}$ with finitely many local extremums (minimums or maximums). Then, we have for $q \in \mathbb{N}$,*

$$d_B(\mathcal{P}_q(f), \mathcal{P}_q(g)) \leq \|f - g\|_\infty,$$

where d_B is the bottleneck distance defined as $d_B(\mathcal{P}_q(f), \mathcal{P}_q(g)) = \inf_\gamma \sup_{(b,d) \in \mathcal{P}_q(f)} \|(b,d) - \gamma(b,d)\|_\infty$, where γ ranges over all bijections from $\mathcal{P}_q(f)$ to $\mathcal{P}_q(g)$ considering the infinite points on the diagonal.

In fact, Theorem 3.1.1 is a special form of a stability theorem (Main Theorem in Cohen-Steiner et al., 2007, p. 109). See **Supplementary Material** (section 1.3) for an illustrative example of the sub-level sets filtration.

3.2. Second Useful Filtration—Vietoris-Rips Complexes Filtration

To introduce Vietoris-Rips (VR) complexes filtration for a given time series, we first embed the time series into a high dimension point cloud via *Takens's lag map* (Takens, 1981), which is constructed in the following way. Take $p \in \mathbb{N}$ to be the dimension of the embedding, and $\tau \in \mathbb{N}$ to be the lag step. For a given time series $\mathbf{x}: \mathbb{Z} \rightarrow \mathbb{R}$, the lag map with lag τ and dimension p is defined as

$$R_{p,\tau}(\mathbf{x}) = \{(\mathbf{x}(t), \mathbf{x}(t - \tau), \mathbf{x}(t - 2\tau), \dots, \mathbf{x}(t - (p - 1)\tau))^\top \mid t \in \mathbb{Z}\}, \tag{8}$$

which is a subset of \mathbb{R}^p . We postpone details of Takens's lag map to **Supplementary Material** (section 1.2). With the point cloud $R_{p,\tau}(\mathbf{x}) \subseteq \mathbb{R}^p$, we are ready to introduce the Vietoris-Rips complex.

In general, given a point cloud $\mathcal{X} = \{x_1, \dots, x_N\} \subset \mathbb{R}^p$, the main idea of Vietoris-Rips complex is to build simplicial complexes from \mathcal{X} if points in \mathcal{X} are closed enough. A formal definition is given below.

Definition 5 (Edelsbrunner and Harer, 2010 section 3.2, p. 61). *Let $\mathcal{X} = \{x_1, x_2, \dots, x_N\} \subseteq \mathbb{R}^p$ be a point cloud and take $\epsilon > 0$. The Vietoris-Rips complex is a collection of all simplices σ with vertices in \mathcal{X} with $\text{diam}(\sigma) \leq 2\epsilon$, where $\text{diam}(\sigma)$ is the diameter of σ ; that is,*

$$\text{VR}(\mathcal{X}; \epsilon) := \bigcup_{q=0}^p \{q\text{-simplex } \sigma \mid \text{diam}(\sigma) \leq 2\epsilon, \text{Vert}(\sigma) \subseteq \mathcal{X}\}. \tag{9}$$

Clearly, for an increasing sequence $\epsilon_1 < \epsilon_2 < \dots < \epsilon_N$, the corresponding sequence of Vietoris-Rips complexes forms a filtration:

$$\text{VR}(\mathcal{X}; \epsilon_1) \subseteq \text{VR}(\mathcal{X}; \epsilon_2) \subseteq \dots \subseteq \text{VR}(\mathcal{X}; \epsilon_N). \tag{10}$$

After determining the representation rules of connected components, the lifespan of holes of different dimensions can be computed easily. See **Supplementary Material** (section 1.3) for an illustrative example of the Vietoris-Rips filtration.

For simplicity, we denote the q -th persistence diagram associated with the Vietoris-Rips filtration as $\mathcal{P}_q(R_{p,\tau}(\mathbf{x})) := \mathcal{P}_q(\{\text{VR}(R_{p,\tau}(\mathbf{x}); \epsilon)\}_\epsilon)$. In parallel with Theorem 3.1.1, the stability of persistence diagrams extracted from a Vietoris-Rips filtration has been discussed in Chazal et al. (2014).

Theorem 3.2.1 (Chazal et al., 2014, Theorem 5.2). *For finite metric spaces (X, d_X) and (Y, d_Y) , then for $q \in \mathbb{N}$,*

$$d_B(\mathcal{P}_q(\text{VR}(X)), \mathcal{P}_q(\text{VR}(Y))) \leq 2d_{GH}(X, Y),$$

where d_B is the bottleneck distance and d_{GH} is the Gromov-Hausdorff distance.

The formal definition of Gromov-Hausdorff distance can be found in Chazal et al. (2014) and Burago et al. (2001), and conceptually, it measures the similarity between two metric spaces under distance-preserving transformations.

3.3. Persistence Statistics

We now introduce a set of new features to summarize persistence diagrams. It is computationally efficient and straightforward to implement. We propose to explore distributions of the birth b and the death d of all possible holes, and calculate their statistic measurements. This idea is considered one of the most straightforward way to extract features from persistence diagrams (Pun et al., 2018). Despite its simplicity, it has been used in several studies, such as skin lesions classification (Chung et al., 2018), bifurcations analysis in dynamical systems (Mittal and Gupta, 2017), and protein classification (Cang et al., 2015).

To be more specific, given a persistence diagram \mathcal{P} , we transform it into two multi-sets of numbers, M and L , defined as

$$M = \left\{ \frac{d+b}{2} \mid (b, d) \in \mathcal{P} \right\} \text{ and } L = \left\{ d - b \mid (b, d) \in \mathcal{P} \right\}. \tag{11}$$

Note that for the Vietoris-Rips complex filtration, $\frac{d+b}{2}$ captures the “size” of the associated hole, and $d - b$ captures the robustness of the associated hole. On the other hand, for the sublevel set filtration, $\frac{d+b}{2}$ reveals the locations of holes, and $d - b$ captures the differences between low and high peaks in a time series. Note that since the hole $(0, \infty)$ always exists in the persistence diagram as is shown in the previous section, it is omitted.

In this paper, for each persistence diagram, we consider eight summary statistics to represent the multi-set M , including mean, standard deviation, skewness, kurtosis, 25th, 50th, 75th percentile, and the *persistent entropy* (Chintakunta et al., 2015). We number them from 1 to 8. We consider the same summary statistics for the multi-set L , and number them from 9 to 16.

Definition 6 (Persistence Statistics). *Given a persistence diagram, the persistence statistics (PS) is defined as a map, $\Phi^{(PS)}$, that transforms the persistence diagram to a point in \mathbb{R}^{16} .*

As shown in Algorithm 1, for the Vietoris-Rips complex filtration, we consider 0-th and 1-th persistence diagrams; for the sub-level set filtration, we consider 0-th persistence diagram.

The persistent entropy of M and L , denoted as $E(M)$ (Chung and Lawson, 2019) and $E(L)$ (Atienza et al., 2020) respectively, describes the complexity of M and L . They are formally defined by

$$E(M) = \sum_{m \in M} \left[-\frac{m}{\sum_{m' \in M} m'} \log \frac{m}{\sum_{m' \in M} m'} \right] \text{ and } E(L) \\ = \sum_{l \in L} \left[-\frac{l}{\sum_{l' \in L} l'} \log \frac{l}{\sum_{l' \in L} l'} \right].$$

$E(L)$ has been used to study the cell arrangements (Atienza et al., 2019a), emotion recognition (Gonzalez-Diaz et al., 2019), and epileptic seizures detection in EEG signals (Piangerelli et al., 2018). From the theoretical perspective, $E(L)$ is a stable measurement (Theorem 3.12 in Atienza et al., 2020). $E(M)$ was first appeared in Chung et al. (2018) and the discussion about the stability of $E(M)$ can be found in Chung and Lawson (2019).

Note that while *intuitively*, holes with long lifespans are considered important features and those with short lifespans are considered noises, in our proposed features, we do not discriminate holes with long or short lifespans. In other words, we take all holes into consideration. This approach is supported by a recent discovery that those considered as noisy holes might actually contain important information. For example, in the drivers' behavior classification (Bendich et al., 2016), authors transformed the space of persistence diagrams into "binned" diagrams, and found that the main differences occurred in those short lifespan holes. Another work on the leave classification (Patrangenaru et al., 2018) also suggested that holes with short lifespans could better distinguish different types of leaves.

4. APPLICATION TO SLEEP STAGE CLASSIFICATION

In recent decades, a growing body of evidence shows that sleep is not only intimately related to personal health (Karni et al., 1994; Kang et al., 2009) but also has a direct impact on public health (Colten and Altevogt, 2006). In clinics, sleep experts score sleep stage by reading the electroencephalogram (EEG), electrooculogram (EOG), and electromyogram (EMG) based on the American Academy of Sleep Medicine (AASM) criteria (Iber et al., 2007; Berry et al., 2012). Sleep, however, impacts the whole body, and we can read sleep via reading physiological signals other than EEG, particularly ECG and HRV mentioned in section 1. The relationship between HRV and sleep dynamics has been widely studied in the physiology society (Zemaityte and Varoneckas, 1984; Vaughn et al., 1995; Toscani et al., 1996; Bonnet and Arand, 1997; Elsenbruch et al., 1999; Chouchou and Desseilles, 2014; Penzel et al., 2016). Specifically, when a subject

is awake, since the sympathetic tone of the ANS is dominant, he/she has a higher heart rate and a less stable heart rhythm due to external stimuli (Somers et al., 1993). When a subject is asleep, the heart rate is lower, and it reaches its lowest value during deep (slow wave) sleep (Snyder et al., 1964). During NREM (non-rapid eye movement) sleep, the parasympathetic nervous system dominates the sympathetic tone and the energy restoration and metabolic rates reach their lowest levels, so the heart rate decreases and the rhythm of the heart stabilizes (Somers et al., 1993).

The above physiological facts indicate that the heart rate rhythm provides a non-invasive window for researchers to study sleep. There have been several studies trying to classify sleep stages based *solely* on HRV. Most of them focus on classifying wake and sleep (Lewicke et al., 2008; Long et al., 2012; Aktaruzzaman et al., 2015; Ye et al., 2016; Malik et al., 2018), some focus on detecting drowsiness (Vicente et al., 2016), and some focus on classifying rapid eye movement (REM) and NREM (Mendez and Matteucci, 2010), or wake, REM, and NREM (Xiao et al., 2013). The challenge and difficulty of this mission can be appreciated from the reported results. In this section, we apply the TDA tool and the proposed persistence statistics to study this problem.

4.1. Datasets

The databases we use here are the same as those used in Malik et al. (2018). Here we summarize them and refer readers with interest in the database details to Malik et al. (2018). The *CGMH-training* database consists of standard polysomnogram (PSG) signals on patients suspicious of sleep apnea syndrome at the sleep center in Chang Gung Memorial Hospital (CGMH), Linkou, Taoyuan, Taiwan. The Institutional Review Board of CGMH approved the study protocol (No. 101-4968A3). All recordings were acquired on the Alice 5 data acquisition system (Philips Respironics, Murrysville, PA). Each recording lasts for at least 5 h. The sleep stages, including wake, REM, and NREM (REM and NREM constitute the *sleep* stage), were annotated by two experienced sleep specialists according to the AASM 2007 guidelines (Iber et al., 2007), and a consensus was reached. According to the protocol, the sleep specialists provide annotation for non-overlapping 30 s long epochs. In this study, we focus on the second lead of the ECG recording, which was sampled at 200 Hz. There are 90 participants without sleep apnea [each with apnea-hypopnea index (AHI) <5] in this database, among which we consider only 56 participants who have at least 10% epochs labeled as wake to avoid the imbalanced data issue.

We consider three validation databases. The first one is the *CGMH-validation database*. This database consists of 27 participants acquired independently of CGMH-training from the same sleep laboratory in CGMH under the same Institutional Review Board. The other two validation databases are publicly available. The DREAMS Subjects Database¹ (DREAMS), consists of 20 recordings from healthy participants, where the ECG recordings were acquired by the BrainnetTM system (Medatec, Brussels, Belgium). The sampling rate is 200 Hz, and the

¹doi: 10.5281/zenodo.2650142

minimum recording duration is 7 h. Although the race information is not provided, we may assume that its population constitution is different from that of the CGMH databases since it is collected from Belgium. This database is chosen to assess the model's performance on participants of a different race recorded from different recording machine. The third database is the St. Vincent's University Hospital/University College Dublin Sleep Apnea Database (UCDSADB) from Physionet (Goldberger et al., 2000)². It consists of 25 participants with sleep apnea of various severities. The ECG signal was recorded by Holter monitor at the sampling rate of 128 Hz. The minimum recording is 6 h long. We focus on the first ECG lead in this study. The UCDSADB is chosen to assess the model's performance on recordings which come from participants with sleep disorders. We remark the these validation databases are not used to tune the model's parameters, and no subject is rejected.

4.2. Time Series to Analyze—Instantaneous Heart Rate

The data preprocessing steps are the same as those shown in Malik et al. (2018). Here we summarize those steps and refer readers to Malik et al. (2018) for more details. First, apply a standard automatic R peak detection algorithm (Elgendi, 2013). Suppose there are n_k R peaks in the k -th subject's ECG recording. Denote $\{r_{k,i}\}_{i=1}^{n_k}$ the location in time (sec) of the detected R peaks of the k -th subject. We apply the 5-beat median filter to remove artifacts in the detected R peaks; that is, if a detected beat is too close or too far from their preceding beats, it is removed or interpolated. Then, the IHR of the k -th recording, denoted as \mathbf{x}_k , is determined by the shape-preserving piecewise cubic interpolation (Task Force of the European Society of Cardiology and others, 1996) over the nonuniform sampling

$$\mathbf{x}_k(r_{k,i}) = 60(r_{k,i} - r_{k,i-1})^{-1}. \quad (12)$$

\mathbf{x}_k describes the IHR at each time in beats-per-minute. The IHR is sampled at a sampling rate of 4 Hz. We break the IHR signal into 30-s epochs following the same epoch segmentation in the experts' annotations. We discard all epochs with fewer than 5 detected R peaks. This step is adjusted by physiological knowledge. For each labeled epoch, we build a time series of 90 s in length by concatenating the epoch with the preceding 2 epochs. For the sake of handling the inter-individual variance, each 90 s time series is normalized by subtracting its median value. Thus, for the j -th epoch of the k -th recording, the associated time series we consider is

$$\mathbf{x}^{(k,j)} := [\mathbf{x}_k(t_j - 359/4), \mathbf{x}_k(t_j - 358/4), \dots, \mathbf{x}_k(t_j - 1/4), \mathbf{x}_k(t_j)]^T - \text{median}\{\mathbf{x}_k(t_j - (q-1)/4) \mid q = 1, \dots, 360\} \in \mathbb{R}^{360}, \quad (13)$$

where t_j indicates the ending time of the j -th epoch.

4.3. IHR Time Series and Their Persistence Diagrams

Following the discussion in section 3, we apply TDA to IHR time series defined in (13), $\mathbf{x}^{(k,j)}$. More precisely, we consider $\mathcal{P}_0(\mathbf{x}^{(k,j)})$ via the sub-level set filtration, and $\mathcal{P}_i(R_{120,1}(\mathbf{x}^{(k,j)}))$ for $i = 0, 1$, via the Vietoris-Rips complex filtration. We extract persistence statistics from both $\mathcal{P}_0(\mathbf{x}^{(k,j)})$ and $\mathcal{P}_i(R_{120,1}(\mathbf{x}^{(k,j)}))$, where $i = 0, 1$. We summarize section 3 and highlight our approach in the following pseudocode. See also **Figure 2** for an illustration.

Algorithm 1: Feature Extraction Scheme

Input: A time series, $\mathbf{x}(t)$.

Output: Topological features used in this article.

1. Calculate $\mathcal{P}_0(\mathbf{x})$ via sublevel set filtration (as in section 3.1).
 2. Calculate $\mathcal{P}_i(R_{120,1}(\mathbf{x}))$ via VR complex filtration (as in section 3.2).
 3. Calculate PS features:

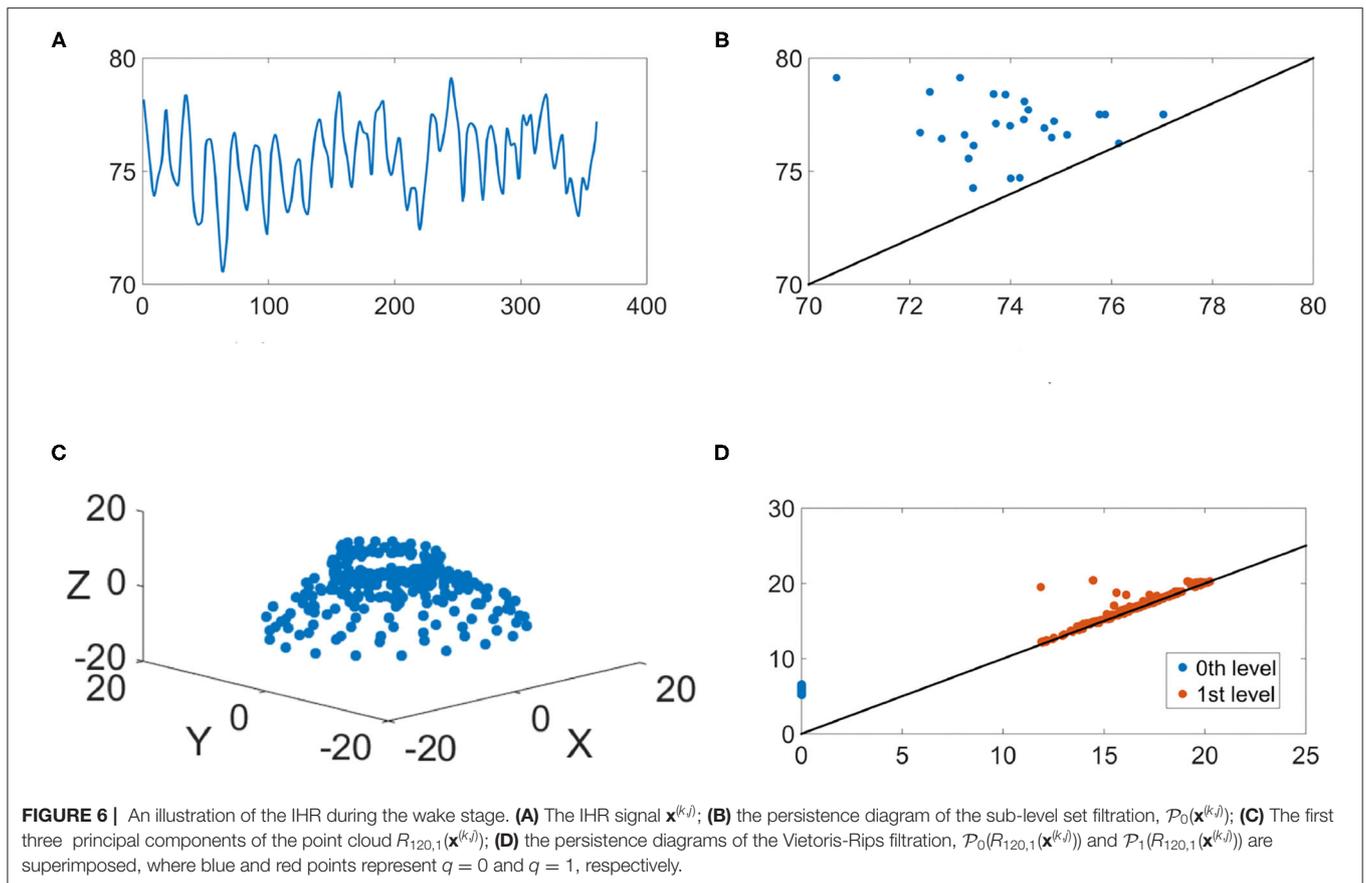
$[\Phi^{(\text{PS})}(\mathcal{P}_0(\mathbf{x})), \Phi^{(\text{PS})}(\mathcal{P}_0(R_{120,1}(\mathbf{x}))), \Phi^{(\text{PS})}(\mathcal{P}_1(R_{120,1}(\mathbf{x})))]$

 (as in section 3.3).
-

We illustrate the IHR time series and their persistence diagrams with different filtrations in **Figures 6, 7**. From a IHR time series during a wake (resp. sleep) epoch shown in **Figure 6A** (resp. **Figure 7**), we observe that these IHRs seem to be different: wake epoch seems to have more variability than sleep one does. Sub-level set filtration captures such variability in the form of the persistence diagram. As shown in **Figures 6B, 7B**, their persistence diagram's of sub-level set filtration are different. Points in **Figure 6B** spread widely while most points in **Figure 7B** are clustered around lower left portion of the diagram. Moreover, **Figure 6B** seems to have more long-lived points than **Figure 7B** does. Next, we examine the persistence diagrams of Vietoris-Rips complex filtration. In this work, we take $(p, \tau) = (120, 1)$, where $p = 120$ is equivalent to a 30 s long time series (since the sampling rate is 4 Hz). This set of parameters is motivated/guided by the AASM criteria where the sleep stage is assigned based on the 30-s readings. The parameters $(120, 1)$ can be thought as sliding a window of a 30-s long time series, and $\mathcal{P}(R_{120,1}(\mathbf{x}))$ stores information about changes of this 30-s sliding window over time. **Figures 6C, 7C** show examples of $R_{120,1}(\mathbf{x}^{(k,j)})$ projected onto their first three principal components. Visually, the point clouds of **Figures 6C, 7C** have different shapes (former seems to have a "lamp" shape while the latter does not), and their persistence diagrams shown in **Figures 6D, 7D** are also different. For instance, the red points in **Figure 6D** cluster around birth values $10 \sim 20$, the red points in **Figure 7D** have three clusters around the birth values $15 \sim 25$, $30 \sim 40$, and 55 . It is important to note that the computations on $\mathcal{P}_0(R_{120,1}(\mathbf{x}^{(k,j)}))$ are done on the \mathbb{R}^{120} space, and projection onto their first three principal components is merely for the visualization purpose.

As discussed in section 2.4, while it is possible to analyze the data via persistence diagrams, it is usually computationally challenging. The proposed persistence statistics allows us to further summarize the persistence diagrams and quantify the above observations. To examine the persistence statistics features,

²<https://archive.physionet.org/pn3/ucddb/>



take $\bigcup_k \{\Phi^{(PS)}(\mathcal{P}_0(\mathbf{x}^{(k,j)}))\}_{j=1}^{n_k}$ as an example. In order to compare them on the same scale, we perform the standard z-score normalization for each subject. We abuse the notation and use $\Phi^{(PS)}(\mathcal{P}_0(\mathbf{x}^{(k,j)}))$ to denote the normalized parameters. In **Figure 8A**, we show the boxplot of each normalized persistence statistics parameter, where blue (red) bars represent the persistence statistics associated with an IHR time series associated with the sleep (wake) stage. We performed a rank sum test with the null hypothesis that two samples have equal medians with a significance level of 0.05 with the Bonferroni correction. We found that there are significant differences between waking and sleeping features for all persistence statistics parameters, except for the kurtosis of M (labeled as 4 in **Figure 8A**), and the median of L (labeled as 14 in **Figure 8A**). The boxplot as shown in **Figure 8A** shows that the mean and standard deviation of M are the most distinguishable persistence statistics parameters between sleep and wake epochs. To further visualize these features, we apply the principle component analysis (PCA) to $\bigcup_k \{\Phi^{(PS)}(\mathcal{P}_0(\mathbf{x}^{(k,j)}))\}_{j=1}^{n_k}$, and plot the first three principal components in \mathbb{R}^3 as shown **Figure 8B**. We observe a separation between sleep and wake features. The visualization of $\Phi^{(PS)}(\mathcal{P}_i(R_{120,1}(\mathbf{x}^{(k,j)})))$, where $i = 0, 1$, is shown in **Supplementary Figure 5**.

Motivated by the above observation and discussion, we consider the following features for $\mathbf{x}^{(k,j)}$ to distinguish sleep and

wake epochs:

$$[\Phi^{(PS)}(\mathcal{P}_0(\mathbf{x}^{(k,j)})), \Phi^{(PS)}(\mathcal{P}_0(R_{120,1}(\mathbf{x}^{(k,j)}))), \Phi^{(PS)}(\mathcal{P}_1(R_{120,1}(\mathbf{x}^{(k,j)})))] \tag{14}$$

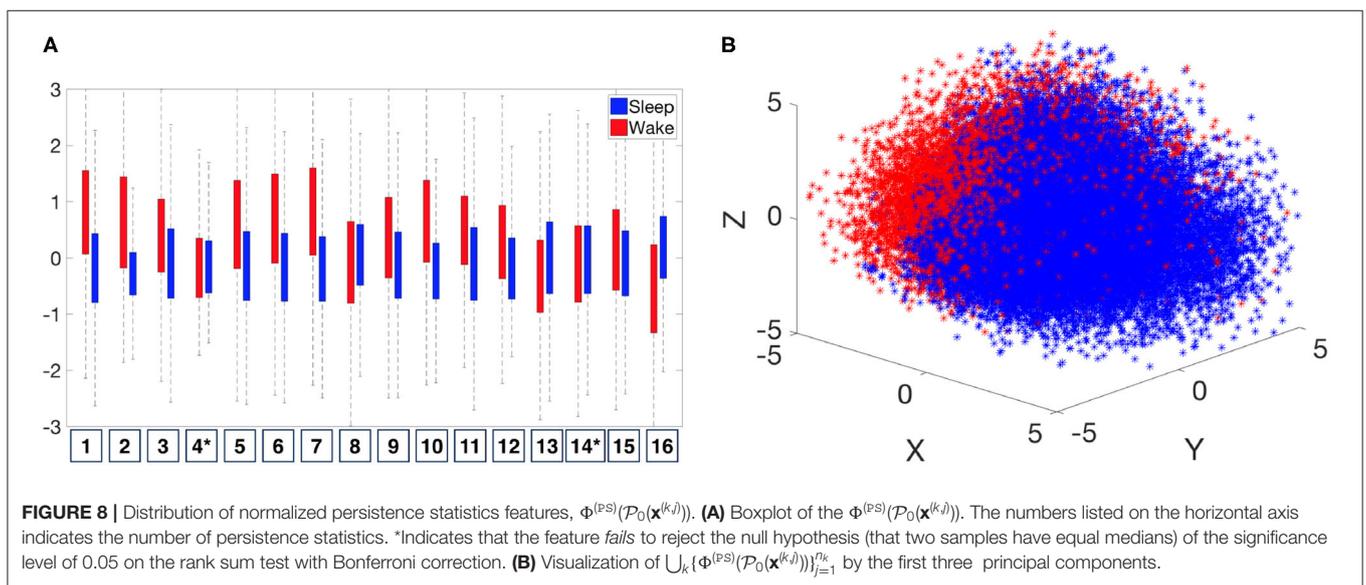
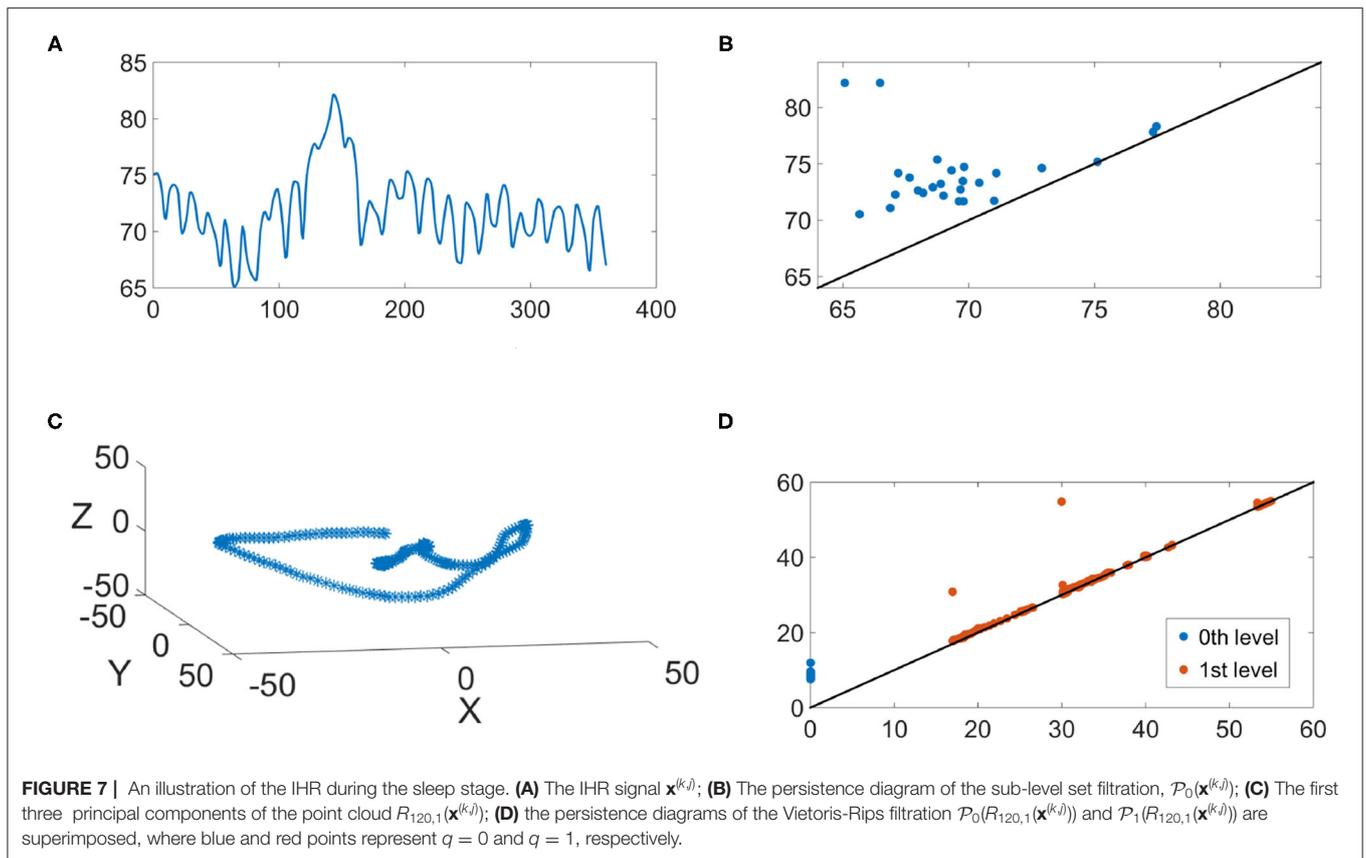
4.4. Automatic Sleep Stage Annotation System

4.4.1. Support Vector Machine as the Learning Model

We consider the widely applied classifier with a solid theoretical foundation, the support vector machine (SVM), to establish the heartbeat classification model. This is Step 4 (machine learning step) shown in **Figure 2**. The linear function kernel is considered in this work and we use the Matlab built-in function `fitcsvm` with default parameters. The input data are features shown in (14), which are calculated by the publicly available libraries DIPHA (<https://github.com/DIPHA/dipha>) and Ripser (<https://github.com/Ripser/ripser>). When there are more than 2 classes, we apply the Error-Correcting Output Codes (ECOC) Dietterich and Bakiri (1994) with one-versus-one design. Specifically, we use the Matlab built-in function `fitcecoc` with default parameters. The Matlab version is 2019b.

4.4.2. Statistics

We carry out the *cross-database validation*. Specifically, we train the SVM model on one database, and evaluate the performance



on the other databases. One of the main challenges in this automatic annotation problem is that the datasets are usually imbalanced; for example, the number of wake epochs is usually much smaller than that of sleep epochs (e.g., in the CGMH-training, the total number of wake epochs is 9,150, while the total number of sleep epochs is 54,547). Learning on imbalanced

datasets is one of challenging topics in machine learning (He and Ma, 2013; Kuhn and Johnson, 2013; Fernández et al., 2018). Typically, the accuracy would heavily skew toward the majority case. Taking the above CGMH-training as an example, if a model predicts all epochs as sleep, then its accuracy is $54547 / (9150 + 54547) \approx 85\%$, which may seem high at the first glance. However,

this model is clearly useless because it has no predictability for the wake, which can be seen through the sensitivity, $0/9150 = 0$. Therefore, in the case of imbalanced datasets, both sensitivity and specificity would be important indicators to evaluate the performance of a model. They both should be as high as possible, and they should be on the similar level.

In order to account for the imbalanced dataset, we adopt a down-sampling process. Let E_s and E_w be the collection of all sleep and wake epochs, respectively, across all subjects in the training set, and denote their cardinality by $|E_s|$ and $|E_w|$, respectively. We take all epochs in E_w , and randomly select $|E_w|$ epochs from E_s . The SVM model will then be built on these balanced epochs. Once the model is built on the training dataset, we test it on the entire testing dataset.

We report the following performance measurement indices. When there are m labels, denote $M \in \mathbb{R}^{m \times m}$ to be the confusion matrix of the automatic classification model, where M_{kl} represents the count of epochs whose known group labels are k and whose predicted group labels are l . The sensitivity (SE), positive predictivity (+P) and F1 for the k -th class, the Cohen kappa, and the overall accuracy (Acc) are defined as

$$SE_k = \frac{M_{kk}}{\sum_{l=1}^m M_{kl}}, \quad +P_k = \frac{M_{kk}}{\sum_{l=1}^m M_{lk}}, \quad F1_k = \frac{2(+P_k) \cdot SE_k}{(+P_k) + SE_k},$$

$$Acc = \frac{\sum_{k=1}^m M_{kk}}{\sum_{k=1}^m \sum_{l=1}^m M_{kl}}, \quad Kappa = \frac{Acc - EA}{1 - EA}, \tag{15}$$

respectively, where EA means the expected accuracy and is defined by

$$EA = \frac{\sum_{p=1}^m \left(\sum_{q=1}^m M_{pq} \right) \times \left(\sum_{q=1}^m M_{qp} \right)}{\left(\sum_{p,q=1}^m M_{pq} \right)^2}. \tag{16}$$

When $m = 3$, $k = 1$ means wake, $k = 2$ means REM, and $k = 3$ means NREM. When classifying wake and sleep stages, $k = 1$ means wake, and $k = 2$ means sleep; when classifying REM and NREM stages, $k = 1$ means REM, and $k = 2$ means NREM. When $m = 2$, SE_1 is reduced to the usual sensitivity (SE), SE_2 is reduced to the usual specificity (SP), and $+P_1$ is reduced to the precision (PR). For each database and each performance measurement, we report the mean \pm standard deviation of all subjects.

All experiments in this and next sections were done using Windows 7 operating system environment equipped with i5-4570 CPU and 32 GB RAM. Under this computational environment, given a random seed, the whole training process of an SVM model takes 5–7 min on average. For the reproducibility purpose, the Matlab code is available in the GitHub repository website³.

4.4.3. Automatic Sleep Stage Classification Result

We performed three classification tasks—sleep v.s. wake, REM v.s. NREM, and finally wake v.s. REM v.s. NREM. The random seed is fixed to 1 in all cases when we ran the subsampling scheme. The results are shown in **Tables 1–3**, where the SVM model was

TABLE 1 | SVM cross-database performance of subjects for Wake and Sleep classification with a single random seed.

	CGMH-training	CGMH-validation	DREAMS	UCDSADB
TP	76 \pm 43	76 \pm 44	101 \pm 55	85 \pm 44
FP	151 \pm 49	126 \pm 48	175 \pm 61	149 \pm 59
TN	462 \pm 68	449 \pm 102	592 \pm 112	448 \pm 110
FN	27 \pm 33	42 \pm 43	56 \pm 45	73 \pm 52
SE (%)	78.3 \pm 14.7	70.9 \pm 16.0	66.9 \pm 16.1	57.6 \pm 15.5
SP (%)	76.0 \pm 6.1	78.9 \pm 5.4	77.6 \pm 5.8	75.3 \pm 5.5
Acc (%)	75.2 \pm 5.4	75.8 \pm 4.4	74.7 \pm 5.0	70.6 \pm 5.4
PR (%)	34.0 \pm 17.0	38.1 \pm 19.6	37.0 \pm 18.8	35.6 \pm 17.3
F1	0.438 \pm 0.161	0.452 \pm 0.140	0.445 \pm 0.146	0.407 \pm 0.140
AUC	0.839 \pm 0.084	0.824 \pm 0.090	0.789 \pm 0.090	0.702 \pm 0.094
Kappa	0.320 \pm 0.146	0.322 \pm 0.123	0.308 \pm 0.148	0.238 \pm 0.133

The training database is CGMH-training. For each database and each performance measurement, we report the mean \pm standard deviation of all subjects.

TABLE 2 | SVM cross-database performance for REM and NREM classification with a single random seed in the training procedure.

	CGMH-training	CGMH-validation	DREAMS	UCDSADB
TP	75 \pm 31	68 \pm 30	93 \pm 34	64 \pm 34
FP	113 \pm 32	106 \pm 43	138 \pm 51	133 \pm 35
TN	400 \pm 68	391 \pm 94	490 \pm 97	373 \pm 68
FN	25 \pm 23	20 \pm 17	46 \pm 27	50 \pm 37
SE (%)	76.3 \pm 16.0	78.1 \pm 17.4	67.5 \pm 16.7	58.0 \pm 18.0
SP (%)	78.0 \pm 4.8	79.6 \pm 6.5	78.4 \pm 5.2	73.7 \pm 5.8
Acc (%)	77.4 \pm 5.6	77.8 \pm 8.3	76.3 \pm 6.4	70.4 \pm 6.4
PR (%)	39.4 \pm 13.6	41.4 \pm 19.0	41.0 \pm 15.4	31.9 \pm 16.0
F1	0.505 \pm 0.138	0.510 \pm 0.160	0.503 \pm 0.144	0.390 \pm 0.156
AUC	0.842 \pm 0.094	0.849 \pm 0.108	0.796 \pm 0.115	0.711 \pm 0.120
Kappa	0.382 \pm 0.150	0.393 \pm 0.175	0.312 \pm 0.175	0.227 \pm 0.162

The training database is CGMH-training. The subject #24 in CGMH-validation and subject #9 in UCDSADB were dropped because they do not have REM epochs. For each database and each performance measurement, we report the mean \pm standard deviation of all subjects.

trained on the CGMH-training dataset and tested on CGMH-validation, DREAMS, and UCDSADB, respectively. For the interested readers, we also include extensive experimental results with different settings in **Supplementary Material** (section 3), such as results of training on different datasets, and different random seeds. All results are similar to those reported in the main article.

Table 1 lists the result of classifying wake and sleep stages with different testing sets. For each testing database, we show the mean \pm standard deviation of each prediction outcome measurement of all subjects in that database. **Table 1** shows the performances of training the model on CGMH-training and testing it on CGMH-validation, DREAMS, and UCDSADB. When considering the CGMH database, the (SE, SP) pair for CGMH-training and CGMH-validation are (78.3 \pm 14.7%, 76.0 \pm 6.1%) and (70.9 \pm 16.0%, 78.9 \pm 5.4%), respectively. When testing

³https://github.com/peterbillhu/TDA_for_SleepWake_Classifications

TABLE 3 | SVM cross-database performance for Wake, REM, and NREM classification with a single random seed in the training procedure.

	CGMH-training	CGMH-validation	DREAMS	UCDSADB
SE (%) (Wake)	63.7 ± 15.3	61.1 ± 19.0	56.2 ± 14.5	39.5 ± 11.9
SE (%) (REM)	62.8 ± 17.4	67.1 ± 20.9	57.0 ± 18.0	48.4 ± 19.6
SE (%) (NREM)	71.9 ± 6.7	72.6 ± 6.4	72.5 ± 6.8	66.0 ± 7.6
+P (%) (Wake)	40.6 ± 19.1	43.7 ± 16.6	44.1 ± 18.3	39.3 ± 17.9
+P (%) (REM)	39.1 ± 14.9	40.0 ± 17.7	39.4 ± 15.5	28.5 ± 15.8
+P (%) (NREM)	89.3 ± 8.8	85.6 ± 16.0	83.7 ± 8.6	76.6 ± 9.2
Acc (%)	68.3 ± 6.4	67.6 ± 9.2	66.3 ± 6.4	57.1 ± 7.1
Kappa	0.401 ± 10.2	0.390 ± 0.117	0.372 ± 0.116	0.244 ± 0.108

The training database is CGMH-training. The subject #24 in CGMH-validation and the subject #9 in UCDSADB were dropped because they do not have REM epochs. For each database and each performance measurement, we report the mean ± standard deviation of all subjects.

on DREAMS, the (SE, SP) pair becomes (66.9 ± 16.1%, 77.6 ± 5.8%). SP remains in the range of 70%, although SE falls below 70%. This result of the cross database testing is similar to that of validation result. When tested on UCDSADB, the pair of (SE, SP) becomes (57.6 ± 15.5%, 75.3 ± 5.5%). The overall performance on UCDSADB drops as expected since it contains sleep apnea subjects, and their sleep dynamics is disturbed by the sleep apnea. Overall, the cross-database validation results suggest that our model does not overfit. Moreover, we found that the down-sampling scheme alleviates the imbalance database issue.

Table 2 shows the performance for the REM and NREM classification. In this task, since the number of NREM epochs is much more than that of REM epochs, we apply the same down-sampling process to NREM as discussed in section 4.4.2. **Tables 1, 2** have several similarities.

Finally, **Table 3** shows the performance for the wake, REM, and NREM classification. In this experiment, since the number of NREM is much more than those of wake and REM, the down-sampling scheme is applied to NREM. The Acc's in all cases are about 60%, except the UCDSADB. The SE's of wake, REM, and NREM are balanced and consistent across databases, except UCDSADB. Again, this result might be due to the fact that UCDSADB contains subjects with sleep apnea. On the other hand, note that the +P of NREM is higher than other classes, which is expected due to the dependence of +P on the database prevalence. In **Supplementary Material**, we provide more cross-database validation results.

5. DISCUSSION AND CONCLUSION

In this work, the TDA tools are considered to analyze time series. Specifically, we propose a set of novel persistence statistics features to quantify HRV by analyzing IHR time series by TDA tools. The proposed HRV features are applied to predict sleep stages, ranging from wake, REM, and NREM. In addition to being computationally efficient, the algorithm is theoretically sound supported by mathematical and statistical results. Note that while we focus on the HRV analysis for the sleep stage annotation, the

proposed algorithm has a potential to be applied to analyze other time series and study the HRV for other clinical problems.

5.1. Theoretical Supports and Open Problems

We find that empirically, M and L are simple yet effective representations of the persistence diagram and reveal signatures about the underlying object. It would be interesting to investigate, in theory, the probability distribution of M and L for a given simplicial complex. However, to the best of our knowledge, while there have been several works in this direction, it is still a relatively open problem. Recently, there has been some theoretical work toward this direction, namely the theory of random complexes (see e.g., survey papers Kahle, 2014; Bobrowski and Kahle, 2018 and references therein). In order to understand the role of noises in the persistence diagram, there have been studies on the topology of the noise. In the theory of random fields, authors in Mischaikow et al. (2010) used sub-level sets as filtration to study the number of components (β_0) with various random processes; in Adler et al. (2010), authors studied the relation between random fields and the persistent homology in general. In particular, as mentioned in Adler et al. (2010): "It would be interesting to know more about the real distributions lying behind the persistence diagram, but at this point we know very little." There is also a result in random cubical complexes (Hiraoka and Tsunoda, 2018), and a few work on the limiting theorem of total sum persistence (Owada, 2018) and persistence diagrams (Hiraoka and Tsunoda, 2018). It would also be interesting to study the stability of each persistence statistics. As of now, only the sum of L , the max of L , and the entropy of L have been shown to be stable (Cohen-Steiner et al., 2010; Atienza et al., 2020). However, the rest of persistence statistics is still unknown. Another interesting direction, instead of focusing on each persistence statistics, is to study the probability distributions of M and L . For instance, let ρ_M and ρ_L be the empirical probability density function of the sets M and L , respectively. For $S = M$ or $S = L$, could one establish $\|\rho_S(f) - \rho_S(g)\|_D \leq d_B(\mathcal{P}_q(f), \mathcal{P}_q(g))$, where $\|\cdot\|_D$ is some suitable statistical distance? We leave those interesting theoretical problems to future work.

5.2. Comparison With Existing Automatic Sleep Stage Annotation Results

There have been several results in automatic sleep stage annotation by taking *solely* the HRV into account. A common conclusion is that classifying sleep-wake by quantifying HRV is a challenging job. In general, due to the heterogeneity of the data sets, various evaluation criteria and different features and models used in these publications, it is difficult to have a direct comparison. But to be fair, below we summarize some related existing literatures for a discussion. To the best of our knowledge, except (Malik et al., 2018), there is no result reporting a cross-database validation. For those running validation on a single database, we shall distinguish two common cross validation (CV) schemes – *leave-one-subject-out* CV (LOSOCV) and *non-LOSOCV*. When the validation set and the training set come from different subjects, we call it the LOSOCV scheme;

otherwise we call it the non-LOSOCV scheme. The LOSOCV scheme is in general challenging due to the uncontrollable inter-individual variability, while the non-LOSOCV scheme tends to over-estimate. Therefore, for a fair comparison, below we only summarize papers considering *only* the IHR features and carrying out the LOSOCV scheme.

In Xiao et al. (2013), the database was composed of healthy participants aged 16 – 61 years. A random forest model was established to differentiate between the wake, REM, and NREM stages for those epochs labeled as “stationary.” Based on the confusion matrix provided in Xiao et al. (2013), the SE, SP, Acc, and F1 for detecting the wake stage are 51.2, 90.2, 84.0, and 0.50%. The authors also provided the SE of wake, REM, and NREM, which are 53.72 ± 20.15 , 59.01 ± 19.72 , and $79.50 \pm 7.82\%$ respectively. In **Table 3**, our validation on CGMH-validation is 61.1 ± 19.0 , 67.1 ± 20.9 , and $72.6 \pm 6.4\%$, respectively. Observe that our SE of wake and REM are better, and SE of NREM is on the similar level. In addition to the balance of all classes due to the sub-sampling scheme in our result, note that we focus on all epochs but not “stationary epochs,” and the subjects in CGMH-validation are not healthy but simply without sleep apnea.

In Mendez and Matteucci (2010), the database was composed of 24 participants aged 40 – 50 years with 0 AHI. The authors took the temporal information and the phase and magnitude of the “sleepy pole” as features to train a hidden Markov model to differentiate REM and NREM stages. The reported SE, SP, and Acc were 70.2, 85.1, and 79.3%, respectively. Our results outperform theirs. Our SE, SP and Acc of the REM and NREM classification in CGMH-validation shown in **Table 2** are 78.1 ± 17.4 , 79.6 ± 6.5 , and $77.8 \pm 8.3\%$, respectively. Observe that both Accs are similar which means portion of correct predictions are similar. Not only our SE is better, but SE and SP are also balanced.

In Lewicke et al. (2008), the database is composed of 190 infants. A variety of features and classification algorithms were considered and the wake and sleep classes were balanced for the analysis. The SE and SP of their multi-layer perceptron model without rejection was 79.0 and 77.5%, respectively. In **Table 1**, the SE and SP of our result on CGMH-validation is 70.9 ± 16.0 and $78.9 \pm 5.4\%$. Our performance is comparable to theirs. However, there is a fundamental difference between their experiments and ours—the sleep dynamics of infants and adults are different.

In Aktaruzzaman et al. (2015), the database is composed of 20 participants aged 49–68 years with varying degrees of sleep apnea. Detrended fluctuation analysis and a feed-forward neural network were applied to differentiate the wake and sleep stages. Various epoch lengths were considered, and the highest performance was recorded on an epoch length of 5 min. The Acc, SE, SP, and Cohen’s kappa were 71.9 ± 18.2 , 43.7 ± 27.3 , 89.0 ± 7.8 , and 0.29 ± 0.24 , respectively. We consider UCDSADB for a comparison. In **Table 1**, the Acc, SE, SP, and Cohen’s kappa of our testing result on UCDSADB is 70.6 ± 5.4 , 57.6 ± 15.5 , 75.3 ± 5.5 , and $0.238 \pm 0.133\%$. Their Acc and ours are on the same level, our SE is better than theirs, while their SP is better than ours. However, our SE and SP are balanced compared with theirs. A major difference is that our standard deviations for Acc, SE, SP are much smaller. Thus, our performance is comparable to theirs.

In Long et al. (2012), fifteen participants aged 31.0 ± 10.4 years with the Pittsburgh Sleep Quality Index less than 6 were considered. The linear discriminant-based classifier was trained with spectral HRV features. The SE, SP, Cohen’s kappa and AUC were $49.7 \pm 19.2\%$, $96 \pm 3.3\%$, 0.48 ± 0.24 and 0.54, respectively. As shown in **Table 1**, the SE, SP, Cohen’s kappa and AUC of our result on CGMH-validation is 70.9 ± 16.0 , 78.9 ± 5.4 , 0.322 ± 0.123 , and $0.824 \pm 0.090\%$, respectively. Again, compared with their results, our SE and SE are more balanced.

To make a conclusion, we emphasize that all those results under comparison are not carried out in the cross-database scheme. Also, usually the SE and SP are not balanced with high SP, which leads to the high accuracy. Therefore, the results suggest that the proposed persistence statistics features and chosen learning model lead a better, or at least similar, performance compared with the state-of-the-art results. The cross-database validation further suggests the usability of the persistence statistics features and the proposed learning scheme in clinical setups. Last but not the least, due to the numerical efficiency of the proposed persistence statistics features, it is potential to apply it to analyze large scale time series.

5.3. Technical Issues

We remark that although it is possible to include $\mathcal{P}_i(R_{120,1}(\mathbf{x}^{(k,j)}))$ for $i \geq 2$ in (14), in practice, it is a challenging task due to its computational complexity. Its computation is known to be poorly scalable in dimension and memory-intensive. We refer readers to Otter et al. (2017) for more details and comparisons among state-of-arts TDA packages and extensive benchmark. To get an idea of the computational cost, for any epoch, the computational time by the state-of-art package Ripser for $\mathcal{P}_1(R_{120,1}(\mathbf{x}^{(k,j)}))$, $\mathcal{P}_2(R_{120,1}(\mathbf{x}^{(k,j)}))$, and $\mathcal{P}_3(R_{120,1}(\mathbf{x}^{(k,j)}))$ are 0.06, 1.7, and 106 s in a standard laptop, respectively. This echos the fact that the computation of $\mathcal{P}_i(R_{120,1}(\mathbf{x}^{(k,j)}))$ does not scale well in i Otter et al. (2017). We demonstrated on adding features $\mathcal{P}_2(R_{120,1}(\mathbf{x}^{(k,j)}))$ and tested classification performance on DREAMS and UCDSADB datasets. The results are listed in **Supplementary Tables 18, 19**. Comparing **Supplementary Table 18** with **Supplementary Table 1** and **Supplementary Table 19** with **Supplementary Table 2**, we find that the improvements are too marginal to justify the additional computational time.

Thus, it would be inefficient to obtain the higher dimensional persistence features. A possible approach for tackling the computational inefficiency is to reduce the Taken’s embedding dimension. In Myers et al. (2019), the authors discussed how to find a proper embedding dimension p by considering the false nearest neighbors (Fraser and Swinney, 1986). It would be interesting to combining this embedding technique to our future works. Also, finding another reduction criterion for the Taken’s embedding is also an important future direction.

5.4. Limitations and Future Directions

In addition to the theoretical development discussed above, there are several interesting practical problems left untouched. While we systematically consider the inter-individual variance, the race, the machine, and the sleep disorder by taking three different

databases into account, we acknowledge the fact that the data is collected from the sleep lab. When the data is collected from the real-world mobile device, it is not clear if the algorithm could perform as well and run in real-time. Moreover, its performance for the home-based screening needs to be further evaluated. Yet, in the current mobile health market, the photoplethysmography (PPG) sensor has been widely applied, and its applicability for the sleep-wake classification has been reported in Malik et al. (2018). It is interesting to see how the TDA approach could be applied to analyze the HRV from the PPG for the sleep stage classification mission. From the data analysis perspective, it would be interesting to perform a more sophisticated analysis and take other features from the persistence diagram. For instance, the persistent homology transformation (Turner et al., 2014b) was recently developed and proven to be a sufficient statistic, and had been successfully applied to the shape analysis. It would be interesting to combine the persistent homology transformation and persistence statistics. IHR is a well-known non-stationary time series. Based on the encouraging results of applying the TDA, we suspect that the persistence statistics features could be applied to study other clinical problems related to HRV, and furthermore, analyze other physiological time series, including the multivariate ones. There has been some work using TDA tools to analyzing the multivariate time series (Merelli et al., 2015; Gidea and Katz, 2018; Wu and Hargreaves, 2021) where the critical step is to transform multivariate time series into a point cloud so that Vietoris-Rips complex persistent homology can be computed. It would also be interesting to investigate ways to use the sublevel set filtration in this context. We will explore those limitations/directions in our future work.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**supplementary material**, further inquiries can be directed to the corresponding author/s.

REFERENCES

- Adams, H., Emerson, T., Kirby, M., Neville, R., Peterson, C., Shipman, P., et al. (2017). Persistence images: a stable vector representation of persistent homology. *J. Mach. Learn. Res.* 18, 218–252. doi: 10.5555/3122009.3122017
- Adler, R. J., Bobrowski, O., Borman, M. S., Subag, E., Weinberger, S., et al. (2010). "Persistent homology for random fields and complexes," in *Borrowing Strength: Theory Powering Applications—a Festschrift for Lawrence D. Brown* (Institute of Mathematical Statistics), 124–143. doi: 10.1214/10-IMSCOLL609
- Aktaruzzaman, M., Migliorini, M., Tenhunen, M., Himanen, S. L., Bianchi, A. M., and Sassi, R. (2015). The addition of entropy-based regularity parameters improves sleep stage classification based on heart rate variability. *Med. Biol. Eng. Comput.* 53, 415–425. doi: 10.1007/s11517-015-1249-z
- Atienza, N., Escudero, L., Jimenez, M., and Soriano-Trigueros, M. (2019a). Persistent entropy: a scale-invariant topological statistic for analyzing cell arrangements. *arXiv [preprint]. arXiv:1902.06467*.
- Atienza, N., Gonzalez-Diaz, R., and Rucco, M. (2019b). Persistent entropy for separating topological features from noise in Vietoris-Rips complexes. *J. Intell. Inf. Syst.* 52, 637–655. doi: 10.1007/s10844-017-0473-4
- Atienza, N., Gonzalez-Diaz, R., and Soriano-Trigueros, M. (2020). On the stability of persistent entropy and new summary functions for topological data analysis. *Pattern Recogn.* 107:107509. doi: 10.1016/j.patcog.2020.107509
- Bendich, P., Chin, S. P., Clark, J., Desena, J., Harer, J., Munch, E., et al. (2016). Topological and statistical behavior classifiers for tracking applications. *IEEE Trans. Aero Elec. Syst.* 52, 2644–2661. doi: 10.1109/TAES.2016.160405
- Berry, E., Chen, Y.-C., Cisewski, J., and Fasy, B. (2020). Functional summaries of persistence diagrams. *J. Appl. Comput. Topol.* 4, 211–262. doi: 10.1007/s41468-020-00048-w
- Berry, R. B., Budhiraja, D. G., Gottlieb, D. J., Gozal, D., Iber, C., Kapur, V. K., et al. (2012). Rules for scoring respiratory events in sleep: update of the 2007 AASM manual for the scoring of sleep and associated events. *J. Clin. Sleep Med.* 8, 597–619. doi: 10.5664/jcsm.2172
- Biasotti, S., De Floriani, L., Falcidieno, B., Frosini, P., Giorgi, D., Landi, C., et al. (2008). Describing shapes by geometrical-topological properties of real functions. *ACM Comput. Surveys* 40:12. doi: 10.1145/1391729.1391731
- Billman, G. E. (2011). Heart rate variability—a historical perspective. *Front. Physiol.* 2:86. doi: 10.3389/fphys.2011.00086
- Blumberg, A. J., Gal, I., Mandell, M. A., and Pancia, M. (2014). Robust statistics, hypothesis testing, and confidence intervals for persistent homology on metric measure spaces. *Found. Comput. Math.* 14, 745–789. doi: 10.1007/s10208-014-9201-4
- Bobrowski, O., and Kahle, M. (2018). Topology of random geometric complexes: a survey. *J. Appl. Comput. Topol.* 1, 331–364. doi: 10.1007/s41468-017-0010-0

AUTHOR CONTRIBUTIONS

Y-MC initiated the project, devised the main pipeline, prototyped the code, and wrote the manuscripts. C-SH implemented the codes, ran all empirical results, produced all the tables, and wrote the manuscripts. Y-LL processed the datasets, interpreted the results, and provided the feedback on physiological aspects. H-TW initiated the project, interpreted the results, wrote the manuscripts, and provided the feedback on all aspects of the project. All authors contributed to the article and approved the submitted version.

FUNDING

C-SH was supported by the project MOST108-2119-M-002-031 hosted by the Ministry of Science and Technology in Taiwan.

ACKNOWLEDGMENTS

The authors acknowledge the hospitality of National Center for Theoretical Sciences (NCTS), Taipei, Taiwan during summer, 2019, when finishing this manuscript. The authors would like to thank Mr. Dominic Tanzillo for his help of proofreading. C-SH want to thank Prof. Jung-Kai Chen (NCTS), Prof. Chun-Chi Lin (NTNU), and Prof. Mao-Pei Tsui (NTU) for the kindly financial support for the work.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fphys.2021.637684/full#supplementary-material>

- Bonnet, M. H., and Arand, D. L. (1997). Heart rate variability: sleep stage, time of night, and arousal influences. *Electroencephalogr. Clin. Neurophysiol.* 102, 390–396. doi: 10.1016/S0921-884X(96)96070-1
- Bubenik, P. (2015). Statistical topological data analysis using persistence landscapes. *J. Mach. Learn. Res.* 16, 77–102. doi: 10.5555/2789272.2789275
- Burago, D., Burago, Y., and Ivanov, S. (2001). *A Course in Metric Geometry*. CRM Proceedings & Lecture Notes. Providence, RI: American Mathematical Society. doi: 10.1090/gsm/033
- Cang, Z., Mu, L., Wu, K., Opron, K., Xia, K., and Wei, G.-W. (30 Nov. 2015). A topological approach for protein classification. *Comput. Math. Biophys.* 3, 140–162. doi: 10.1515/mlbmb-2015-0009
- Carlsson, G. (2009). Topology and data. *Bull. Am. Math. Soc.* 46, 255–308. doi: 10.1090/S0273-0979-09-01249-X
- Carlsson, G., Zomorodian, A., Collins, A., and Guibas, L. J. (2005). Persistence barcodes for shapes. *Int. J. Shape Model.* 11, 149–187. doi: 10.1142/S0218654305000761
- Chazal, F., de Silva, V., and Oudot, S. (2014). Persistence stability for geometric complexes. *Geometr. Dedic.* 173, 193–214. doi: 10.1007/s10711-013-9937-z
- Chevyrev, I., Nanda, V., and Oberhauser, H. (2018). Persistence paths and signature features in topological data analysis. *IEEE Trans Pattern Anal.* 42, 192–202. doi: 10.1109/TPAMI.2018.2885516
- Chintakunta, H., Gentimis, T., Gonzalez-Diaz, R., Jimenez, M.-J., and Krim, H. (2015). An entropy-based persistence barcode. *Pattern Recogn.* 48, 391–401. doi: 10.1016/j.patcog.2014.06.023
- Chouchou, F., and Desselles, M. (2014). Heart rate variability: a tool to explore the sleeping brain? *Front. Neurosci.* 8:402. doi: 10.3389/fnins.2014.00402
- Chung, Y., Hu, C., Lawson, A., and Smyth, C. (2018). “Topological approaches to skin disease image analysis,” in *2018 IEEE International Conference on Big Data (Big Data)* (Seattle, WA), 100–105. doi: 10.1109/BigData.2018.8622175
- Chung, Y.-M., and Lawson, A. (2019). Persistence curves: a canonical framework for summarizing persistence diagrams. *arXiv [preprint]*. arXiv:1904.07768.
- Cohen-Steiner, D., Edelsbrunner, H., and Harer, J. (2007). Stability of persistence diagrams. *Discrete Comput. Geomet.* 37, 103–120. doi: 10.1007/s00454-006-1276-5
- Cohen-Steiner, D., Edelsbrunner, H., Harer, J., and Mileyko, Y. (2010). Lipschitz functions have Lp-stable persistence. *Found. Comput. Math.* 10, 127–139. doi: 10.1007/s10208-010-9060-6
- Colten, H. R., and Altevogt, B. M. (2006). “Functional and economic impact of sleep loss and sleep-related disorders,” in *Sleep Disorders and Sleep Deprivation: An Unmet Public Health Problem*, eds H. R. Colten and B. M. Altevogt (Washington, DC: National Academies Press), 137–163.
- Costa, M., Goldberger, A. L., and Peng, C.-K. (2002). Multiscale entropy analysis of complex physiologic time series. *Phys. Rev. Lett.* 89:068102. doi: 10.1103/PhysRevLett.89.068102
- Dietterich, T. G., and Bakiri, G. (1994). Solving multiclass learning problems via error-correcting output codes. *J. Artif. Intell. Res.* 2, 263–286. doi: 10.1613/jair.105
- Draghici, A. E., and Taylor, J. A. (2016). The physiological basis and measurement of heart rate variability in humans. *J. Physiol. Anthropol.* 35:22. doi: 10.1186/s40101-016-0113-7
- Edelsbrunner, H., and Harer, J. (2010). *Computational Topology. An Introduction*. Rhode Island: American Mathematical Society. doi: 10.1090/mkb/069
- Edelsbrunner, H., Letscher, D., and Zomorodian, A. (2000). “Topological persistence and simplification,” in *Proceedings 41st Annual Symposium on Foundations of Computer Science* (Redondo Beach, CA), 454–463. doi: 10.1109/SFCS.2000.892133
- Elgendi, M. (2013). Fast QRS detection with an optimized knowledge-based method: evaluation on 11 standard ECG databases. *PLoS ONE* 8:e73557. doi: 10.1371/journal.pone.0073557
- Elsenbruch, S., Harnish, M., and Orr, W. (1999). Heart rate variability during waking and sleep in healthy males and females. *Sleep* 22, 1067–1071. doi: 10.1093/sleep/22.8.1067
- Epstein, C., Carlsson, G., and Edelsbrunner, H. (2011). Topological data analysis. *Inverse Probl.* 27:120201. doi: 10.1088/0266-5611/27/12/120201
- Fasy, B. T., Lecci, F., Rinaldo, A., Wasserman, L., Balakrishnan, S., Singh, A., et al. (2014). Confidence sets for persistence diagrams. *Ann. Stat.* 42, 2301–2339. doi: 10.1214/14-AOS1252
- Fernández, A., García, S., Galar, M., Prati, R., Krawczyk, B., and Herrera, F. (2018). *Learning From Imbalanced Data Sets*. Berlin: Springer International Publishing. doi: 10.1007/978-3-319-98074-4
- Fraser, A. M., and Swinney, H. L. (1986). Independent coordinates for strange attractors from mutual information. *Phys. Rev. A* 33, 1134–1140. doi: 10.1103/PhysRevA.33.1134
- Gidea, M., and Katz, Y. (2018). Topological data analysis of financial time series: landscapes of crashes. *Phys. A* 491, 820–834. doi: 10.1016/j.physa.2017.09.028
- Glass, L. (2009). Introduction to controversial topics in nonlinear science: Is the normal heart rate chaotic? *Chaos* 19:028501. doi: 10.1063/1.3156832
- Goldberger, A., Amaral, L., Glass, L., Hausdorff, J., Ivanov, P., Mark, R., et al. (2000). Physiobank, physiotoolkit, and physionet: components of a new research resource for complex physiologic signals. *Circulation* 101, e215–e220. doi: 10.1161/01.CIR.101.23.e215
- Gonzalez-Diaz, R., Paluzo-Hidalgo, E., and Quesada, J. F. (2019). “Towards emotion recognition: a persistent entropy application,” in *International Workshop on Computational Topology in Image Context* (Marseille: Springer), 96–109. doi: 10.1007/978-3-030-10828-1_8
- Graff, G., Graff, B., Jabłoński, G., and Narkiewicz, K. (2020). “The application of persistent homology in the analysis of heart rate variability,” in *2020 11th Conference of the European Study Group on Cardiovascular Oscillations (ESGCO)* (Pisa), 1–2. doi: 10.1109/ESGCO49734.2020.9158054
- He, H., and Ma, Y. (2013). *Imbalanced Learning: Foundations, Algorithms, and Applications*. Hoboken, NJ: Wiley. doi: 10.1002/9781118646106
- Hiraoka, Y., and Tsunoda, K. (2018). Limit theorems for random cubical homology. *Discrete Comput. Geomet.* 60, 665–687. doi: 10.1007/s00454-018-0007-z
- Iber, C., Ancoli-Israel, S., Chesson, A. Jr, and Quan, S. (2007). *The AASM Manual for Scoring of Sleep and Associated Events-Rules: Terminology and Technical Specification*. Darien, IL: American Academy of Sleep Medicine.
- Kahle, M. (2014). Topology of random simplicial complexes: a survey. *AMS Contemp. Math* 620, 201–222. doi: 10.1090/conm/620/12367
- Kang, J.-E., Lim, M. M., Bateman, R. J., Lee, J. J., Smyth, L. P., Cirrito, J. R., et al. (2009). Amyloid- β Dynamics are regulated by Orexin and the sleep-wake cycle. *Science* 326, 1005–1007. doi: 10.1126/science.1180962
- Karni, A., Tanne, D., Rubenstein, B. S., Askenasy, J. J., and Sagi, D. (1994). Dependence on REM sleep of overnight improvement of a perceptual skill. *Science* 265, 679–682. doi: 10.1126/science.8036518
- Kerber, M., Morozov, D., and Nigmatov, A. (2017). Geometry helps to compare persistence diagrams. *J. Exp. Algorithm.* 22, 1–4. doi: 10.1145/3064175
- Kim, K., Kim, J., and Rinaldo, A. (2018). Time series featurization via topological data analysis: an application to cryptocurrency trend forecasting. *arXiv [preprint]*. arXiv:1812.02987.
- Kuhn, M., and Johnson, K. (2013). *Applied Predictive Modeling*. New York, NY: Springer. doi: 10.1007/978-1-4614-6849-3
- Kusano, G., Hiraoka, Y., and Fukumizu, K. (2016). “Persistence weighted Gaussian kernel for topological data analysis,” in *International Conference on Machine Learning* (New York, NY), 2004–2013.
- Lewicke, A., Sazonov, E., Corwin, M. J., Neuman, M., and Schuckers, S. (2008). Sleep versus wake classification from heart rate variability using computational intelligence: consideration of rejection in classification models. *IEEE Trans. Biomed. Eng.* 55, 108–118. doi: 10.1109/TBME.2007.900558
- Long, X., Fonseca, P., Haakma, R., Aarts, R. M., and Foussier, J. (2012). “Time-frequency analysis of heart rate variability for sleep and wake classification,” in *BIBE* (Larnaca), 85–90. doi: 10.1109/BIBE.2012.6399712
- Malik, J., Lo, Y.-L., and Wu, H.-T. (2018). Sleep-wake classification via quantifying heart rate variability by convolutional neural network. *Physiol. Meas.* 39:085004. doi: 10.1088/1361-6579/aad5a9
- Marwan, N., Wessel, N., Meyerfeldt, U., Schirdewan, A., and Kurths, J. (2002). Recurrence-plot-based measures of complexity and their application to heart-rate-variability data. *Phys. Rev. E* 66:026702. doi: 10.1103/PhysRevE.66.026702
- Mendez, M., and Matteucci, M. (2010). Sleep staging from Heart Rate Variability: time-varying spectral features and Hidden Markov Models. *Int. J. Biomed. Eng. Technol.* 3, 246–263. doi: 10.1504/IJBET.2010.032695
- Merelli, E., Piangerelli, M., Rucco, M., and Toller, D. (2015). “A topological approach for multivariate time series characterization: the epileptic brain,” in *BICT'15: Proceedings of the 9th EAI International Conference on Bio-Inspired*

- Information and Communications Technologies (formerly BIONETICS)* (New York, NY). doi: 10.4108/eai.3-12-2015.2262525
- Mileyko, Y., Mukherjee, S., and Harer, J. (2011). Probability measures on the space of persistence diagrams. *Inverse Probl.* 27:124007. doi: 10.1088/0266-5611/27/12/124007
- Mischaikow, K., and Wanner, T. (2010). Topology-guided sampling of nonhomogeneous random processes. *Ann. Appl. Probabil.* 20, 1068–1097. doi: 10.1214/09-AAP652
- Mittal, K., and Gupta, S. (2017). Topological characterization and early detection of bifurcation and chaos in complex systems using persistent homology. *Chaos* 27:051102:1-9. doi: 10.1063/1.4983840
- Munkres, J. R. (1993). *Elements of Algebraic Topology*. Boston, MA: Westview Press.
- Myers, A., Munch, E., and Khasawneh, F. A. (2019). Persistent homology of complex networks for dynamic state detection. *Phys. Rev. E* 100:022314. doi: 10.1103/PhysRevE.100.022314
- Otter, N., Porter, M. A., Tillmann, U., Grindrod, P., and Harrington, H. A. (2017). A roadmap for the computation of persistent homology. *EPJ Data Sci.* 6:17. doi: 10.1140/epjds/s13688-017-0109-5
- Owada, T. (2018). Limit theorems for betti numbers of extreme sample clouds with application to persistence barcodes. *Ann. Appl. Probabil.* 28, 2814–2854. doi: 10.1214/17-AAP1375
- Patrangnanu, V., Bubenik, P., Paige, R. L., and Osborne, D. (2018). Challenges in topological object data analysis. *Sankhya A.* 81, 244–271. doi: 10.1007/s13171-018-0137-7
- Penzel, T., Kantelhardt, J. W., Bartsch, R. P., Riedl, M., Kraemer, J. F., Wessel, N., et al. (2016). Modulations of heart rate, ECG, and cardio-respiratory coupling observed in polysomnography. *Front. Physiol.* 7:460. doi: 10.3389/fphys.2016.00460
- Perea, J. (2019). Topological time series analysis. *arXiv:1812.05143*. doi: 10.1090/noti1869
- Piangerelli, M., Rucco, M., Tesei, L., and Merelli, E. (2018). Topological classifier for detecting the emergence of epileptic seizures. *BMC Res. Notes* 11:392. doi: 10.1186/s13104-018-3482-7
- Pincus, S. M., and Goldberger, A. L. (1994). Physiological time-series analysis: what does regularity quantify? *Am. J. Physiol. Heart Circ. Physiol.* 266, 1643–1656. doi: 10.1152/ajpheart.1994.266.4.H1643
- Porges, S. W. (2009). The polyvagal theory: new insights into adaptive reactions of the autonomic nervous system. *Clev. Clin. J. Med.* 76(Suppl. 2):S86. doi: 10.3949/ccjm.76.s2.17
- Pun, C. S., Xia, K., and Lee, S. X. (2018). Persistent-homology-based machine learning and its applications—a survey. *arXiv:1811.00252*. doi: 10.2139/ssrn.3275996
- Ravishanker, N., and Chen, R. (2019). Topological data analysis (TDA) for time series. *arXiv:1909.10604*.
- Reininghaus, J., Huber, S., Bauer, U., and Kwitt, R. (2015). “A stable multi-scale kernel for topological machine learning,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (Boston, MA), 4741–4748. doi: 10.1109/CVPR.2015.7299106
- Seversky, L. M., Davis, S., and Berger, M. (2016). “On time-series topological data analysis: new data and opportunities,” in *Workshop paper on IEEE Conference on Computer Vision and Pattern Recognition* (Las Vegas, NV). doi: 10.1109/CVPRW.2016.131
- Shaffer, F., McCraty, R., and Zerr, C. L. (2014). A healthy heart is not a metronome: an integrative review of the heart’s anatomy and heart rate variability. *Front. Psychol.* 5:1040. doi: 10.3389/fpsyg.2014.01040
- Snyder, F., Hobson, J. A., Morrison, D. F., and Goldfrank, F. (1964). Changes in respiration, heart rate, and systolic blood pressure in human sleep. *J. Appl. Physiol.* 19, 417–422. doi: 10.1152/jappl.1964.19.3.417
- Somers, V. K., Dyken, M. E., Mark, A. L., and Abboud, F. M. (1993). Sympathetic-nerve activity during sleep in normal subjects. *N. Engl. J. Med.* 328, 303–307. doi: 10.1056/NEJM199302043280502
- Stys, A., and Stys, T. (1998). Current clinical applications of heart rate variability. *Clin. Cardiol.* 21, 719–724. doi: 10.1002/clc.4960211005
- Takens, F. (1981). “Detecting strange attractors in turbulence,” in *Dynamical Systems and Turbulence, Vol. 898 of Lecture Notes in Mathematics*, eds D. Rand and L. S. Young (Berlin; Heidelberg: Springer), 366–381. doi: 10.1007/BFb0091924
- Task Force of the European Society of Cardiology and others (1996). Heart rate variability, standards of measurement, physiological interpretation, and clinical use. *Circulation* 93, 1043–1065. doi: 10.1161/01.CIR.93.5.1043
- Thayer, J. F., and Sternberg, E. (2006). Beyond heart rate variability: vagal regulation of allostatic systems. *Ann. N. Y. Acad. Sci.* 1088, 361–372. doi: 10.1196/annals.1366.014
- Toscani, L., Gangemi, P. F., Parigi, A., Silipo, R., Ragghianti, P., Sirabella, E., et al. (1996). Human heart rate variability and sleep stages. *Ital. J. Neurol. Sci.* 17, 437–439. doi: 10.1007/BF01997720
- Turner, K., Mileyko, Y., Mukherjee, S., and Harer, J. (2014a). Fréchet means for distributions of persistence diagrams. *Discr. Comput. Geom.* 52, 44–70. doi: 10.1007/s00454-014-9604-7
- Turner, K., Mukherjee, S., and Boyer, D. M. (2014b). Persistent homology transform for modeling shapes and surfaces. *Inf. Inference* 3, 310–344. doi: 10.1093/imaia/iau011
- Vanderlei, L., Pastre, C., Hoshi, R., Carvalho, T., and Godoy, M. (2009). Basic notions of heart rate variability and its clinical applicability. *Rev. Bras. Cir. Cardiovasc.* 24, 205–217. doi: 10.1590/S0102-76382009000200018
- Vaughn, B. V., Quint, S. R., Messenheimer, J. A., and Robertson, K. R. (1995). Heart period variability in sleep. *Electroencephalogr. Clin. Neurophysiol.* 94, 155–162. doi: 10.1016/0013-4694(94)00270-U
- Venkataraman, V., Ramamurthy, K. N., and Turaga, P. (2016). “Persistent homology of attractors for action recognition,” in *IEEE International Conference on Image Processing* (Phoenix, AZ). doi: 10.1109/ICIP.2016.7533141
- Vicente, J., Laguna, P., Bartra, A., and Bailón, R. (2016). Drowsiness detection using heart rate variability. *Med. Biol. Eng. Comput.* 54, 927–937. doi: 10.1007/s11517-015-1448-7
- Voss, A., Schulz, S., Schroeder, R., Baumert, M., and Caminal, P. (2008). Methods derived from nonlinear dynamics for analysing heart rate variability. *Philos. Trans. R. Soc. A* 367, 277–296. doi: 10.1098/rsta.2008.0232
- Wang, Y., Ombao, H., and Chung, M. K. (2018). Topological data analysis of single-trial electroencephalographic signals. *Ann. Appl. Stat.* 12, 1506–1534. doi: 10.1214/17-AOAS1119
- Wang, Y., Ombao, H., and Chung, M. K. (2019). “Statistical persistent homology of brain signals,” in *ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (Brighton), 1125–1129. doi: 10.1109/ICASSP.2019.8682978
- Wu, C., and Hargreaves, C. (2021). Topological machine learning for multivariate time series topological machine learning for multivariate time series. *J. Exp. Theor. Artif. Intell.* doi: 10.1080/0952813X.2021.1871971
- Xiao, M., Yan, H., Song, J., Yang, Y., and Yang, X. (2013). Sleep stages classification based on heart rate variability and random forest. *Biomed. Signal Process Control* 8, 624–633. doi: 10.1016/j.bspc.2013.06.001
- Ye, Y., Yang, K., Jiang, J., and Ge, B. (2016). “Automatic sleep and wake classifier with heart rate and pulse oximetry: derived dynamic time warping features and logistic model,” in *10th Annu. Int. Syst. Conf. SysCon 2016 - Proc. (Orlando, FL)*, 1–6. doi: 10.1109/SYSCON.2016.7490623
- Zelinski, B., Juda, M., and Zeppelzauer, M. (2020). Persistence codebooks for topological data analysis. *Artif. Intell. Rev.* doi: 10.1007/s10462-020-09897-4. Available online at: <https://link.springer.com/article/10.1007/s10462-020-09897-4>
- Zemaityte, D., and Varoneckas, G. (1984). Heart rhythm control during sleep. *Psychophysiology* 2, 279–290. doi: 10.1111/j.1469-8986.1984.tb02935.x

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Chung, Hu, Lo and Wu. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.