frontiers | Frontiers in Artificial Intelligence

TYPE Methods
PUBLISHED 14 December 2023
DOI 10.3389/frai.2023.1293504

OPEN ACCESS

EDITED BY
Mohammad Zavid Parvez,
Charles Sturt University, Australia

REVIEWED BY
Jarosław Paweł Drapała,
Wrocław University of Technology, Poland
Muhammad Iqbal Hossain,
BRAC University, Bangladesh
Jia Uddin,
Woosong University, Republic of Korea

*CORRESPONDENCE
Anass B. El-Yaagoubi
✉ anass.bourakna@kaust.edu.sa

RECEIVED 13 September 2023
ACCEPTED 22 November 2023
PUBLISHED 14 December 2023

CITATION
El-Yaagoubi AB, Chung MK and Ombao H (2023) Statistical inference for dependence networks in topological data analysis. *Front. Artif. Intell.* 6:1293504. doi: 10.3389/frai.2023.1293504

# Statistical inference for dependence networks in topological data analysis

Anass B. El-Yaagoubi[1]*, Moo K. Chung[2] and Hernando Ombao[1]

[1]Statistics Program, King Abdullah University of Science and Technology, Thuwal, Saudi Arabia,
[2]Department of Biostatistics and Medical Informatics, University of Wisconsin-Madison, Madison, WI, United States

Topological data analysis (TDA) provide tools that are becoming increasingly popular for analyzing multivariate time series data. One key aspect in analyzing multivariate time series is dependence between components. One application is on brain signal analysis. In particular, various dependence patterns in brain networks may be linked to specific tasks and cognitive processes. These dependence patterns may be altered by various neurological and cognitive impairments such as Alzheimer's and Parkinson's diseases, as well as attention deficit hyperactivity disorder (ADHD). Because there is no ground-truth with known dependence patterns in real brain signals, testing new TDA methods on multivariate time series is still a challenge. Our goal here is to develop novel statistical inference procedures via simulations. Simulations are useful for generating some null distributions of a test statistic (for hypothesis testing), forming confidence regions, and for evaluating the performance of proposed TDA methods. To the best of our knowledge, there are no methods that simulate multivariate time series data with potentially complex user-specified connectivity patterns. In this paper we present a novel approach to simulate multivariate time series with specific number of cycles/holes in its dependence network. Furthermore, we also provide a procedure for generating higher dimensional topological features.

KEYWORDS

**topological data analysis, time series analysis, simulating topological dependence patterns, spectral analysis, simulation-based inference**

## 1 Introduction

Topological data analysis (TDA) has witnessed many important advances over the last twenty years that aim to unravel and provide insight to the "shape" of the data (Edelsbrunner et al., 2002; Edelsbrunner and Harer, 2008; Wasserman, 2018; Chazal and Michel, 2021). The development of TDA tools such as barcodes and persistence diagrams (Ghrist, 2008; Bubenik, 2015; Adams et al., 2017) have opened many new perspectives for analyzing various types of data (Umeda, 2017; Gholizadeh and Zadrozny, 2018; Motta, 2018; Xu et al., 2021; Leykam and Angelakis, 2023). These tools enable practitioners to grasp the topological characteristics inherent in high-dimensional data, which often remain beyond the reach of classical data analysis methods. However, a primary constraint of TDA tools is the absence of robust statistical inference techniques. Our goal in this paper is to introduce a simulation-based inference approach to address this limitation.

Many data sets exhibit a temporal structure (e.g., brain signals, economic data, climate data). In recent years, there has been a noticeable transition from primarily utilizing TDA techniques on clouds of to increasingly applying them on dependence networks of multivariate time series data, particularly for multivariate brain signals such as electroencephalograms (EEG) and local field potentials (LFP) (El-Yaagoubi et al., 2023). Rather than using TDA techniques on a cloud of points via a time delay embedding transformation, it is suggested that the multivariate time series be transformed to its dependence network where the nodes correspond to the time series components and the weight on the edges depend on the intensity of the statistical dependency between any given pair of time series in a network. There is currently no systematic method nor statistical model for conducting simulations (for the purpose of statistical inference or evaluating TDA methods) on networks with complex dependence structure. This is a serious limitation because simulations can form a basis for inference as well as evaluation of TDA methods for sensitivity, specificity, predictive ability. In this paper, we will develop an easily implementable method for simulating data with complex dependence patterns. Thus, the main contributions of the proposed simulation method are the following: (a) they enable a rigorous evaluation and comparison of data analytic methods and (b) provide tools for conducting proper statistical inference (in particular, hypothesis testing).

In the literature, the topology of the brain network (structural and functional) is believed to be organized according to principles that maximize the flow of information and minimize the energy expenditure for maintaining the entire network, such as small world networks (Sporns, 2013; Pessoa, 2014; Muldoon et al., 2016; Henry et al., 2020; Fathian et al., 2022). This topological structure of the brain network can be altered by various conditions such as attention deficit hyperactivity disorder (ADHD), Alzheimer's and Parkinson's diseases. Topological tools have been developed to assess and analyze the topological patterns of different groups' brain networks (e.g., healthy control vs. pathological), as well as quantifying the impact of these disorders on brain organization.

The goal of topological data analysis for time series is to provide computational tools that can assess the topological features present in the dependence network of a multivariate time series. Through the use of persistent homology theory, TDA provides a framework for analyzing the topological features, such as connected components, holes, cavities, etc. that are present in the network (Wasserman, 2018; Chazal and Michel, 2021).

In order to analyze the topological features present in the various dependence networks, we consider the homology of the filtration (increasing sequence of thresholded networks) obtained from these dependence networks, also known as Vietoris-Rips filtration. Let $X(t) = [X_1(t), \ldots, X_P(t)]^T$ be the observed brain signals at $P$ different locations at time $t \in \{1, \ldots T\}$. One can define the dependence network to be a weighted graph with weights between nodes $p$ and $q$ to be some dependence measure between the observed time series components $X_p(t)$ and $X_q(t)$ as seen in Figure 1.

Since the Vietoris-Rips filtration relies on the notion of distance, we then define distance between brain channels at locations $p$ and $q$ to be a decreasing function of the strength of a dependence metric between this pair of channels. For example, using coherence or correlation as dependence measures, one can use the transform $g(x) = 1 - |x|$ (where $x$ is correlation or coherence). A comprehensive discussion on dependence in the spectral domain (e.g., coherence) is given in Ombao and Pinto (2022). Therefore, using this measure of distance, the Vietoris-Rips filtration (see Hausmann, 2016) is constructed by connecting nodes

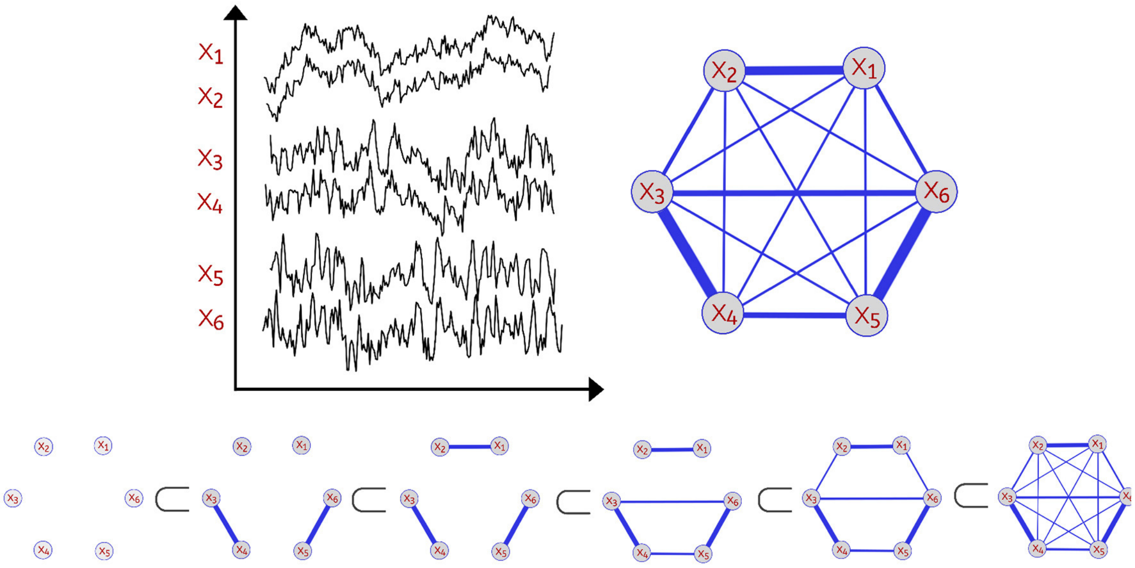

FIGURE 1
Example of a multivariate time series with P=6 channels **(Top Left)**. Dependence network of the multivariate time series **(Top Right)**. Considering this weighted network with $P$ nodes (brain regions) with corresponding time series $X_1(t), \ldots, X_P(t)$, we define the distance between two time series components as a decreasing function of the strength of dependence. Thus, a pair of channels with weak dependence has a large value of the distance measure. For every threshold $\epsilon$, a network with edges with weights not exceeding $\epsilon$ is constructed leading to an increasing sequence also known as Vietoris-Rips filtration **(Bottom)**. As the threshold $\epsilon$ grows, there's birth and death of various topological features.

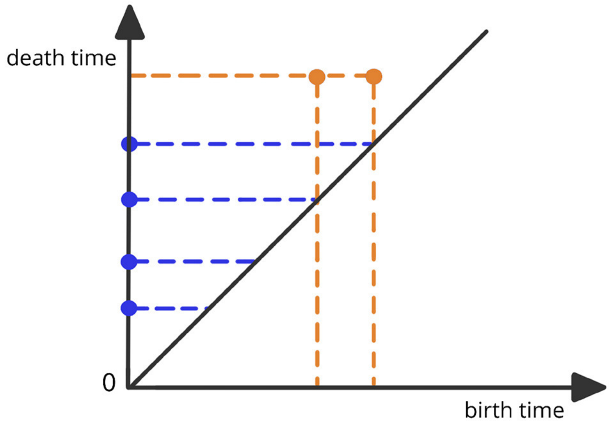

FIGURE 2
Example of a persistence diagram corresponding to the previous Vietoris-Rips filtration. The various dots correspond to different birth-death pairs ($b_\ell, d_\ell$) with birth time in the x-axis and death time in the y-axis.

that have a distance less or equal to some given threshold $\epsilon$, which results in the following filtration:

$$\mathcal{X}_{\epsilon_1} \subset \mathcal{X}_{\epsilon_2} \subset \cdots \subset \mathcal{X}_{\epsilon_n}, \tag{1}$$

where $\mathcal{X}\epsilon$ represents the simplicial complex at the threshold level $\epsilon$. This complex is defined to be the combination of all $k$-simplices (nodes, edges, triangles etc.) of brain channels that are within a maximum distance of $\epsilon$ from each other. The thresholds for distance are defined as $0 < \epsilon_1 < \epsilon_2 < \cdots < \epsilon_{n-1} < \epsilon_n$. For a visual illustration, refer to Figure 1. The objective of this approach is to assess the scales at which topological features (connected components, cycles, holes, and so on) appear (birth time) and then vanish (death time) (Wasserman, 2018; Chazal and Michel, 2021). The Vietoris-Rips filtration can be a complex object. Therefore, the most common topological summary being utilized is the persistence diagram (PD) which is a diagram that represents the times of births and deaths of the topological features in the VR filtration (see Figure 2). Every birth-death pair is represented by a point in the diagram, e.g., $(b_1, d_1)$, $(b_2, d_2)$, ..., where $b_\ell$ is the birth time of the $\ell$-th feature and $d_\ell$ is the death time of the $\ell$-th feature. The points in the PD are colored based on the dimension of the feature they correspond to (e.g., one color for the connected components, another color for the cycles etc.).

There are multiple approaches that are available for modeling and generating multivariate time series data, each offering distinct advantages and limitations. For instance:

- **Parametric VARMA models:** Parametric Vector Autoregressive Moving Average (VARMA) models exhibit flexibility by accommodating dependencies on lagged values of each variable and interactions between terms or variables. This versatility enables the representation of contemporaneous and lagged dependencies, providing a comprehensive modeling framework (Shumway and Stoffer, 2017, Gorrostieta et al., 2018).
- **Gaussian process-based models:** GP provide a flexible framework for capturing complex and non-linear relationships. These models accommodate non-stationary processes, allowing changes in mean and covariance structures over time. Moreover, they allow the incorporation of prior knowledge through the selection of covariance functions (kernels) (Mohammadi et al., 2019).
- **Copula-based models:** These model are effective in modeling tail dependence, particularly in extreme events where variables exhibit dependencies in the tails of their distributions. This capability is crucial for understanding rare and extreme events and enables the modeling of non-Gaussian marginal distributions, addressing real-world scenarios where individual time series do not follow Gaussian distributions (Brechmann and Czado, 2014).
- **Machine learning models:** These data-driven models, such as Generative Adversarial Networks (GANs) offer innovative and powerful methods for generating synthetic multivariate time series, showcasing high flexibility and versatility. They excel in learning complex patterns and dependencies, including non-linearities and temporal dynamics (Snow, 2020).

However, a notable limitation shared by these existing models is their inability to generate multivariate time series data with predetermined patterns in the dependence network, such as a specific number of cycles. Cycles represent situations where nearby channels exhibit correlations or dependencies, while channels further apart may not have direct connections but are linked through intermediary channels in a cyclic manner. The number of cycles can vary, ranging from none to multiple cycles. Driven by the challenges at hand, this paper introduces an innovative method for simulating multivariate time series data with predetermined dependence patterns. The dual purpose is to enable formal statistical tests on complex topological networks and to evaluate the effectiveness of Topological Data Analysis (TDA) methods in the realm of multivariate time series data. Table 1 provides a comprehensive overview, summarizing both the strengths and limitations of existing approaches alongside our proposed method.

In this paper, we introduce an innovative method that leverages mixtures of latent second-order autoregressive processes to generate multivariate time series data showcasing diverse connectivity patterns within the dependence network. In Section 2, we provide a concise overview of AR(2) processes, detailing the generation of various dependence patterns by mixing these AR(2) processes using carefully selected weights. We also demonstrate how the resulting persistence diagram effectively identifies such patterns. In Section 3, we explain how to generate multivariate time series data with more general patterns, and investigate the sensitivity of this approach at various signal-to-noise ratios. Finally, in Section 4, we use our approach to carry out simulation-based inference based on the notion of total persistence.

## 2 Modeling dependence patterns in multivariate time series

In neuroscience, the concept of the nervous system as a (structural and functional) network of interconnected neurons is

TABLE 1 **Comparison between our approach and existing approaches in multivariate time series modeling.**

| Approach | Strengths and advantages | Limitations and challenges |
|---|---|---|
| Parametric VARMA | • Captures lead-lag dependencies<br>• Easily interpretable | • Assumes stationarity<br>• Limited to linear dependencies |
| Gaussian processes | • Captures non-linear dependencies<br>• Allows incorporation of prior knowledge | • Difficulty in imposing arbitrary dependence patterns<br>• Limited scalability |
| Copula-based | • Models tail dependence<br>• Captures non-linear dependencies<br>• Accommodates a wide range of marginal distributions | • Difficulty in imposing arbitrary dependence patterns<br>• Handling time dependence can be challenging |
| Machine learning | • Captures complex and non-linear dependencies<br>• Captures lead-lag dependencies<br>• No stationarity assumption | • Requires training data<br>• Prone to overfitting<br>• Limited interpretability |
| Our approach | • Captures any dependence pattern<br>• Easily interpretable | • Requires definition of a dependence graph<br>• Choice of weight decay function |

now well established (Friston, 2011; Sporns, 2013; Nakagawa and Deco, 2015; Fan et al., 2016). Many brain investigations have led to countless discoveries concerning the brain's anatomical and functional organization. The ongoing scientific endeavor in neuroscience to map the complicated networks of the human brain with increasing accuracy has been primarily due to the technological advances in brain imaging techniques that have resulted in new statistical techniques that aim to study and analyze various patterns in these complex networks. Such methods not only help neuroscientists understand the segregation of brain functions but also the integration of information processing. As a result, the validity of such novel techniques must be evaluated in terms of various metrics such as false positive or false negative rates, type I and type II errors, power of the test (ability of the test to detect differences in patterns between groups when they truly exist). It is impossible to evaluate/assess such approaches without a proper method for generating multivariate time series data with ground truth patterns in its dependence network. This is readily accomplished via extensive computer simulations under various settings of the truth (user-specified ground truth for the dependence networks).

Given the intrinsic complexity of brain signals, which are considered to be a superposition of random oscillations at specific frequencies or frequency bands, it can be challenging to discover and analyze the interrelationships between distinct time series components. As a result, this paper will adopt a frequency-specific strategy to generate meaningful simulations. We will develop a method where the multivariate time series data with dependency connections that are allowed to vary across frequency bands [again, we refer the reader to Ombao and Pinto (2022) for a discussion on spectral metrics for dependence]. For this reason, we will consider coherence as our frequency-specific dependence measure, since it can capture specific oscillations that are common to components in a network of signals. The typical spectral approach for analyzing brain data is to first estimate the spectral matrix, then construct the connectivity network using a spectral dependence measure, usually, coherence or partial coherence (Bowyer, 2016; Hu et al., 2019).

In the following subsections, we will generate multivariate time series data with specific dependence patterns. While numerous methods exist for generating multivariate time series as mixtures of random oscillations, our primary focus will be on utilizing second-order autoregressive processes that are concentrated around a specific frequency band.

## 2.1 Autoregressive processes of order 2

Electrophysiological signals are modeled as mixtures of many random oscillations. Here, each random oscillation with a desired power spectrum will be modeled as a second-order autoregressive process [AR(2)]. One advantage of using AR(2) processes as building blocks for a time series is their ability to represent oscillations at precise frequency bands. See Prado et al. (2001) and Granados-Garcia et al. (2021).

A linear mixture of second order autoregressive processes [AR(2) processes] can be used to simulate the brain oscillatory activity at specific frequency bands. An AR(2) process with a spectral peak at pre-specified frequency and bandwidth can be used to describe a latent process as follows:

$$Z(t) = \phi_1 Z(t-1) + \phi_2 Z(t-2) + W(t) \tag{2}$$

where $W(t)$ is white noise process with $\mathbb{E}\big(W(t)\big) = 0$ and $\mathbb{V}\big(W(t)\big) = \sigma^2$; the relationship between the AR(2) model parameters $\phi_1$ and $\phi_2$ and the spectral properties, namely the frequency peak and bandwidth, will be derived as follows. Note that Equation 2 can be rewritten as $W(t) = (1-\phi_1 B^1 - \phi_2 B^2)Z(t)$ where the back backshift operator $B$ is defined as $B^k Z(t) = Z(t-k)$ for $k = 1, 2$. The AR(2) characteristic polynomial function is:

$$\Phi(r) = 1 - \phi_1 r^1 - \phi_2 r^2. \tag{3}$$

Denote the roots of the $\Phi(r)$ to be $r_1$ and $r_2$. To fulfill the conditions of stationarity and casuality, the roots should lie outside of the unit circle on the complex plane, i.e., $|r_1| > 1$ and $|r_2| > 1$. Consider the case when $r_1$ and $r_2$ are (non-real) complex-valued. Hence these can be expressed as $r_1 = M\exp(i2\pi\psi)$ and $r_2 = M\exp(-i2\pi\psi)$ where the phase $\psi \in (0, 0.5)$ and the magnitude $M > 1$ to satisfy causality (Shumway and Stoffer, 2017). For this latent process $Z(t)$, suppose that the sampling rate is denoted by $SR$ and the peak frequency is $f \in (f_{\min}; f_{\max})$. Then the roots of the AR(2) latent process must have the phase $\psi = f/SR$. In practice, if the sampling rate $SR$ is 100 Hz and we wish to simulate an alpha-band latent process where the peak is at 10 Hz, then it is necessary to set $\psi = 10/100$ and the root magnitude $M$ to some number greater than 1 but "close" to 1 so that the spectrum of $Z(t)$ is mostly concentrated on the frequency band $f$-Hz. The corresponding AR(2) coefficients are derived to be $\phi_1 = \frac{2}{M}\cos(2\pi\psi)$ and $\phi_2 = -\frac{1}{M^2}$. Some examples of such stationary AR(2) processes as well as their corresponding spectrum is given in Figure 3.

## 2.2 Mixtures of AR(2) processes

In TDA applications, Rips-Vietoris filtrations are often applied to multivariate time series data (Umeda, 2017; Gholizadeh and Zadrozny, 2018; El-Yaagoubi et al., 2023). These filtrations are often constructed from clouds of points, or from a weighted network. Traditionally, due to their stochastic nature brain signals have often been modeled using their underlying dependence networks. For instance in Bullmore and Sporns (2009), Henry et al. (2020), and Fathian et al. (2022), the authors use graph theoretical methods on complex brain networks.

To replicate a specific dependency pattern in the dependence network of a multivariate time series, it is necessary to emulate the decay in the dependence structure as time series components get farther away from each other. First sample a graph $G = (N, E)$ with the desired structure (i.e., cycles or holes etc.), second define the observed time series components as mixtures of the latent processes, such that components near to each other in the graph share more latent processes, which makes them more dependent on one another, while components far away in the graph will share fewer latent components, resulting in lower interdependence. Let $Z_p(t)$ be the latent iid AR(2) processes centered around a specific frequency band. Therefore, to generate multivariate time series with a desired dependence patterns (as defined by the graph $G$) the following model is suggested:

$$Y_p(t) = \sum_{q=1}^{P} W_{p,q} Z_q(t) + \epsilon_p(t) \tag{4}$$

$$W_{p,q} = \begin{cases} \frac{1}{1+d_G(p,q)}, & \text{if } d_G(p,q) \leq K, \\ 0, & \text{if } d_G(p,q) > K \end{cases} \tag{5}$$

We generate a $P$-dimensional vector of observations $Y(t) = [Y_1(t), ..., Y_P(t)]^T \in \mathbb{R}^P$ that is a linear mixture of $P$ latent iid AR(2) processes $Z_1(t), ..., Z_P(t)$ according to Equation 4, whith $\mathbb{E}\big(Z_p(t)\big) = 0$, and $\mathbb{V}\big(Z_p(t)\big) = 1$. The mixing weights $W_{p,q}$ contain the information about the importance of the $q$-th latent AR(2) process $Z_q(t)$ in the $p$-th observed component $Y_p(t)$, and as defined by Equation 5, the weights are chosen to be inversely proportional to the distance $d_G(p,q)$ between the nodes in the graph, and $K$ being the maximum distance threshold that is considered, in practice we take $K = 2$ or $K = 3$. Theoretically, any distance-decreasing function might be used. However, selecting a faster decay (such as exponential decay) could result in a too-sharp decrease in the dependence based distance, making it more challenging to identify the topological features in the filtration.

It is essential to consider that the complexity of the dependence pattern directly impacts the required size of the sampled graph (denoted by $P$). This necessity arises from the need to attain higher resolution, ensuring that the sampled nodes adequately cover the entire manifold. Consequently, our approach may encounter limitations when generating intricate patterns within the dependence network of multivariate time series with small dimensions. Moreover, depending on the specific characteristics of the sampled graph and the number of nodes, alternative decaying functions, such as $\frac{1}{1+\sqrt{x}}$, $\frac{1}{1+x^2}$, $\exp(-x)$, need to be tested as they may yield improved results. These functions exhibit varying decay rates for the mixture weights, offering flexibility and adaptability in different scenarios.

Let the $P$-dimensional observed vector $Y(t) = [Y_1(t), \dots, Y_P(t)]^T$. Then we have the following Cramer representation:

$$Y(t) = \int_{-1/2}^{1/2} \exp(i2\pi\omega t) dX(\omega), \tag{6}$$

where the $X(\omega)$ is a $P$-variate random process whose mean is zero with orthogonal increments having the following covariance:

$$Cov(dX(\omega), dX(\lambda)) = \begin{cases} f(\omega)d\omega d\lambda \text{ if } \omega = \lambda + 2\pi k, \; k \text{ an integer}, \\ 0, \text{ otherwise}. \end{cases} \tag{7}$$

and $f(\omega)$ is the spectral density matrix. If we define the filtered components at band $\Omega$ to be:

$$Y_{1,\Omega}(t) = \sum_{\ell} \Psi_\ell Y_1(t-\ell), \tag{8}$$

$$\vdots$$

$$Y_{P,\Omega}(t) = \sum_{\ell} \Psi_\ell Y_P(t-\ell), \tag{9}$$

where the filter $\Psi$ is the band pass filter centered around frequency band $\Omega$. In Ombao and Van Bellegem (2008), coherence between $Y_1(.)$ and $Y_2(.)$ at frequency band $\Omega$ is derived to be the squared correlation between the phase-adjusted $Y_{1,\Omega}(t)$ and $Y_{2,\Omega}(t)$. Coherence will then be used (via a decreasing transformation) to define frequency-specific distance between time series components $Y_p(.)$ and $Y_q(.)$.

Consider the observed data $\{Y(t), t = 1, \dots, T\}$. The spectral matrix $f(\omega)$ can be estimated parametrically (e.g., by fitting a VARMA model), non-parametrically (by smoothing the periodogram) or semi-parametrically. In our case we will be using

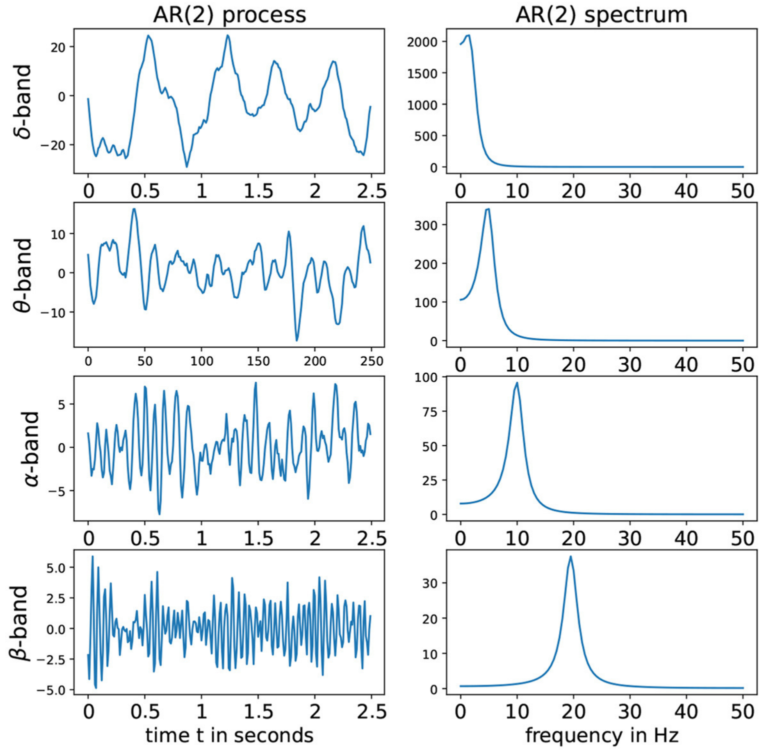

FIGURE 3
**(Left)** AR(2) processes for different frequency bands. **(Right)** Corresponding spectra. From top to bottom we have: delta-band with peak frequency at 2 Hz, theta-band with peak frequency at 5 Hz, alpha-band with peak frequency at 10 Hz, alpha-band with peak frequency at 2Hz and beta-band with peak frequency at 19.5 Hz with sampling rate (SR) of 100*Hz*.

the smoothed periodogram approach. The Fourier $P$-dimensional coefficient at frequency $\omega_k$ are defined as:

$$d(\omega_k) = \frac{1}{\sqrt{T}} \sum_{t=1}^{T} Y(t) \exp\left(-i\omega_k t\right), \tag{10}$$

then the Fourier periodogram is defined to be:

$$I(\omega_k) = d(\omega_k) d(\omega_k)^*, \tag{11}$$

where the $^*$ operator in represents the conjugate transpose. Consequently, $I(\omega_k)$ is a $P \times P$ matrix. It can be shown that the periodogram $I(\omega_k)$ is asymptotically unbiased. However, it is not a consistent estimator of the spectral matrix as the asymptotic variance does not decrease to zero even when we get more and more observations (i.e., $T \rightarrow \infty$). Hence, we construct a mean-squared consistent estimator to be:

$$\widehat{f}(\omega_k) = \sum_{\omega} k_h(\omega - \omega_k) I(\omega) \tag{12}$$

where $k_h(\omega - \omega_k)$ is a non-negative smoothing kernel centered around $\omega_k$ and $h$ is the bandwidth parameter. In order to derive our distance function, first we define coherence as follows:

$$\mathcal{C}\Big(Y_p(.), Y_q(.), \omega\Big) = \frac{|\widehat{f}_{p,q}(\omega)|^2}{\widehat{f}_{p,p}(\omega)\widehat{f}_{q,q}(\omega)} \in [0, 1], \tag{13}$$

then we define the dependence-based frequency-specific distance function to be a decreasing function, for example, $\mathcal{G}(x) = 1 - x$:

$$\mathcal{D}\Big(Y_p(.), Y_q(.), \omega\Big) = \mathcal{G}\Big(\mathcal{C}(Y_p(.), Y_q(.), \omega)\Big). \tag{14}$$

In the following, using the ideas explained previously we start by generating multivariate time series data with specified dependence patterns. We explain how dependence information contained in the graph $G$ can be encoded in the homology structure of connectivity network, using a first example with one main cycle then a second example with two main cycles in the dependence network.

## 2.3 Multivariate time series with cyclic patterns

Our aim is to simulate multivariate time series data with predefined patterns in the dependence network. These simulations serve multiple purposes: (1) conducting a statistical test by obtaining the distribution of a predefined test statistic under the null hypothesis through simulations, and (2) evaluating the performance of TDA methods, including mean-squared error of the estimator and the test's power for group comparisons as sample size increases or the discrepancy between group parameters widens.

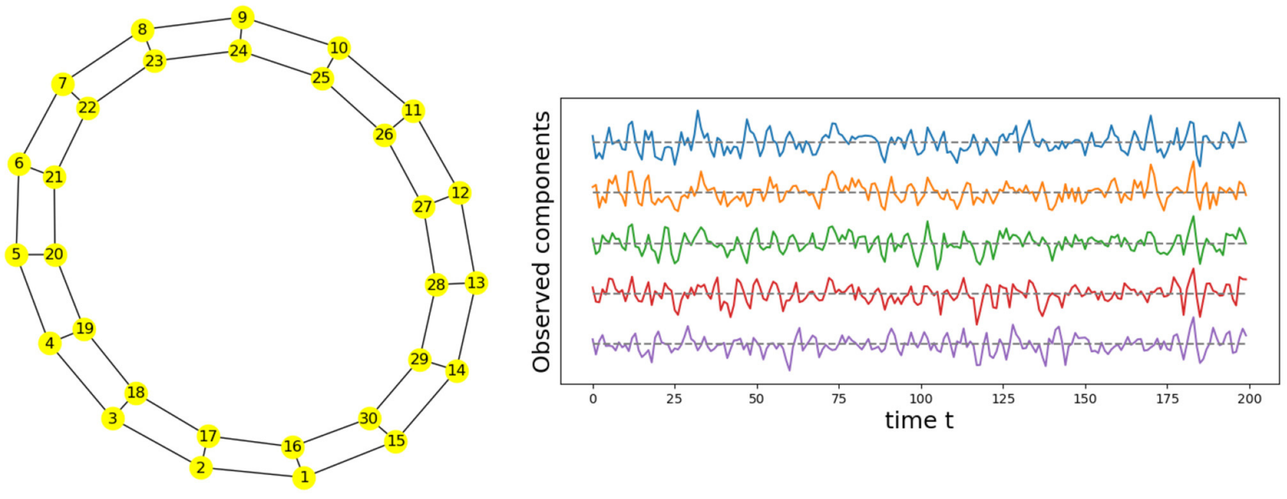

FIGURE 4
Cyclic dependence graph - Circular Ladder model **(Left)**. Generated time series with $P = 30$. The following components are plotted from top to bottom: $Y_1(t)$, $Y_2(t)$, $Y_3(t)$, $Y_4(t)$, $Y_5(t)$ **(Right)**.

The topology of the brain network is known to be organized according to principles that maximize the flow of information and minimize the energy cost for maintaining the entire network (Sporns, 2013; Pessoa, 2014; Hilgetag and Goulas, 2015; Muldoon et al., 2016). However, neurological or mental diseases may affect that organization by degrading the structural and functional connectivity of the brain (e.g., Alzheimer's disease, ADHD etc.) (Bassett and Bullmore, 2009; Henry et al., 2020; Fathian et al., 2022). Therefore, it is important to model these alterations in functional connectivity using time series models that can capture dependencies beyond pairwise nodes in a brain network. Here, we will develop a procedure for simulating multivariate time series with a given number of cycles that may reflect complexity in the brain functional network.

### 2.3.1 One main cycle pattern

In this first example we generate a multivariate time series with one cycle in the dependence structure. In this setting, we will impose the time series components that are relatively close to each other to be more strongly dependent than components that are farther apart.

Given the previous network definition of the circular ladder model, as displayed in Figure 4, we write the expression for $K = 2$, using Equation 4 for the observed time series $Y(t) = [Y_1(t), \ldots, Y_{30}(t)]^T$ and latent process $Z(t) = [Z_1(t), \ldots, Z_{30}(t)]^T$ as follows:

$$Y(t) = WZ(t) + \epsilon(t), \tag{15}$$

where $W_{p,q}$ is the contribution of the latent process $Z_q(t)$ in the observed process $Y_p(t)$ and is equal to $\frac{1}{1+d_G(p,q)}$ if the distance $d_G(p,q)$ between nodes $p$ and $q$ is less than or equal to 2 and 0 otherwise. Thus, the weight matrix $W$ has dimension $30 \times 30$; the latent process vector $Z(t)$ $30 \times 1$; and $\epsilon(t)$ is a $30 \times 1$ noise vector. Using this circular ladder model we can generate and visualize the multivariate time series data as follows in Figure 4. Given the model described above, we calculate pair-wise correlations for the following pairs: $Y_1(t) - Y_2(t)$, $Y_5(t) - Y_6(t)$, and $Y_1(t) - Y_6(t)$. This calculation results in $Corr\Big(Y_1(t), Y_2(t)\Big) = Corr\Big(Y_5(t), Y_6(t)\Big)$, which simplifies to $\frac{10/6}{76/36} \sim 0.76$. Furthermore, $Corr\Big(Y_1(t), Y_6(t)\Big) = 0$. Therefore, the correlation-distance between pairs 1-2 and 5-6 is the same $1 - 0.76$, this is due to symmetry in the graph in Figure 4 and correlation-distance between component 1 and 6 is 1 because they do not share any latent components. Therefore, the time series components 1 and 24 are farther apart (based on the dependence distance) than 1 and 2 or 5 and 6, which is exactly the desired property.

Having developed intuition behind the mechanism that generates the time series components, we now directly compute the coherence matrices for various frequency bands and analyze the topological patterns present in the resulting network. After computing, at each frequency band, the coherence matrix, we also consequently compute the distance matrix, we now build the Rips-Vietoris filtration and visualize the results in the persistence diagram as can be seen in Figure 5.

The orange point in the middle frequency persistence diagram, far from the diagonal, represents the main cycle in the dependence structure. Whereas the orange dots near the diagonal represent the secondary cycles that are present all around the network, see in Figure 5.

### 2.3.2 Two main cycles pattern

We now develop a model for generating multivariate time series with two cycles in the dependence structure. Similarly, from Equation 4, one can generate the multivariate time series with the double circular ladder model as defined in Figure 6. Using this

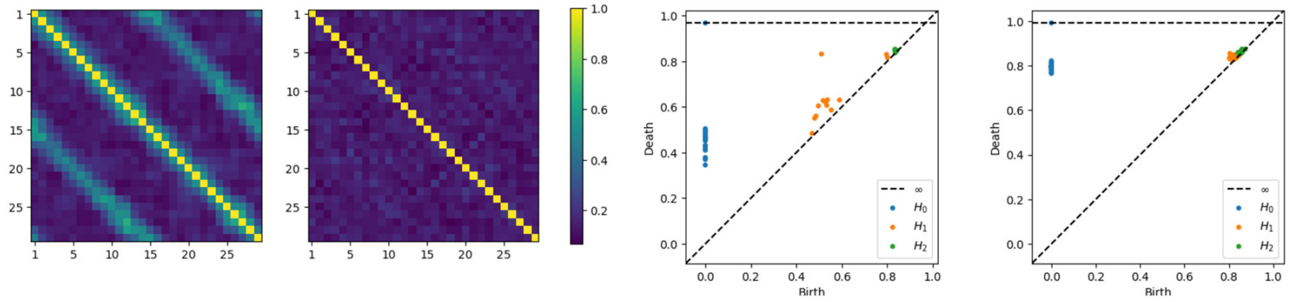

FIGURE 5
Coherence estimates for a one-cycle model. Leftmost: middle-frequency band. Second from the left: high-frequency band. Persistence diagrams for the same one-cycle model, with the middle-frequency band shown as the third from the left and the high-frequency band as the fourth from the left.

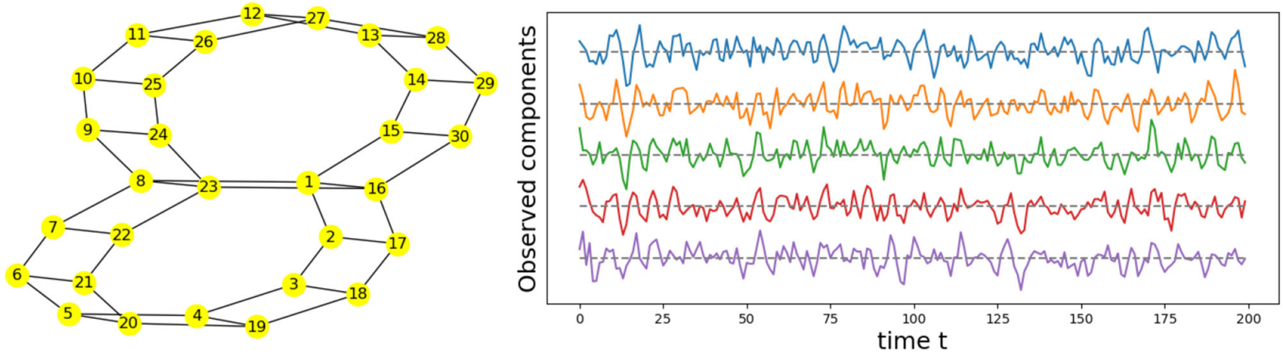

FIGURE 6
Cyclic dependence graph with two main cycles—double circular ladder model. Every node corresponds to a time series component **(Left)**. Generated time series with $P = 30$. The following components are plotted from top to bottom: $Y_1(t)$, $Y_2(t)$, $Y_3(t)$, $Y_4(t)$, $Y_5(t)$, derived from the double circular ladder model **(Right)**.

mechanism we can generate and visualize the multivariate time series as follows (see Figure 6).

Without delving into the computational details of the coherence expression, it's important to note that in this new example, the coherence between any pair of channels within a subnetwork (one of the main cycles) will not reach zero, given the smaller diameter of the subcycles. Intuitively, components that are farther apart tend to exhibit weaker dependence. Therefore, having more connections shortens the path between nodes, resulting in increased dependence, which makes sense since more connections also mean more latent processes are being shared. On the other hand, components that are closer will exhibit stronger dependence as they share more latent processes. We can directly compute the coherence matrices for middle and high frequency bands and analyze the topological patterns present in the resulting network. After estimating the coherence matrices and therefore the distance matrices, the next step is to apply the tools of TDA, i.e., building the Rips-Vietoris filtration and visualize the results in the persistence diagram, as can be seen in Figure 7.

The x-axis represents the birth time (denoted $d_i$), while the y-axis represents the death time (denoted $b_i$), all the point representing valid features in the diagram have to lay above the diagonal line since the death time is larger than the birth time, i.e., $d_i > b_i$. The orange points far from the diagonal represent the two main cycles in the dependence structure. Whereas the orange dots near the diagonal represent the secondary cycles that are present all around the network, as can be seen in Figure 6.

Irrespective of the specific pattern of interest, the approach detailed in Section 2.2, which leverages mixtures of AR(2) processes, proves to be versatile. It allows for the generation of multivariate time series data exhibiting diverse patterns within the dependence network, as long as the pattern can be represented by a graph. The following flowchart (Figure 8) provides an insightful summary of the procedural steps involved, contributing to a more accessible and intuitive understanding of our approach.

## 3 Generating multivariate time series with general patterns in its dependence network

Depending on the application of interest, the simulated patterns presented above may not be sufficient. However, the methodology is general and can be used to define many patterns in the dependence network of a multivariate time series. The goal in this section is to develop a novel robust procedure for simulating multivariate time series with complex dependence structures. Suppose that the

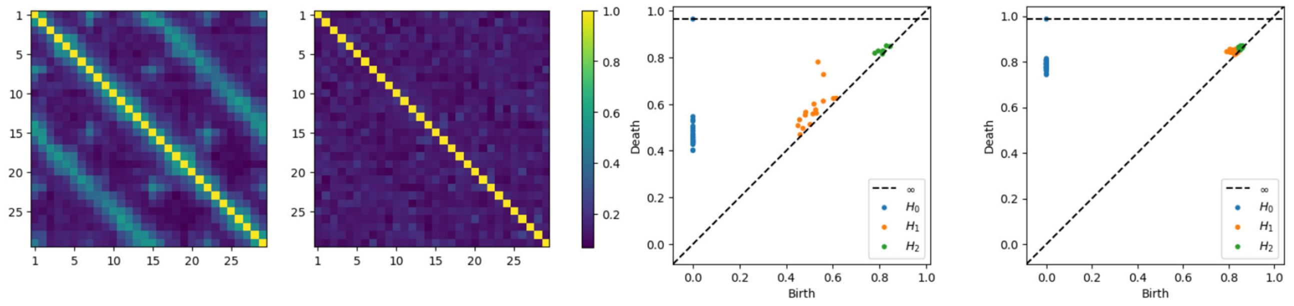

FIGURE 7
Coherence estimates for a two-cycle model. Leftmost: middle-frequency band. Second from the left: high-frequency band. Persistence diagrams for the same one-cycle model, with the middle-frequency band shown as the third from the left and the high-frequency band as the fourth from the left.

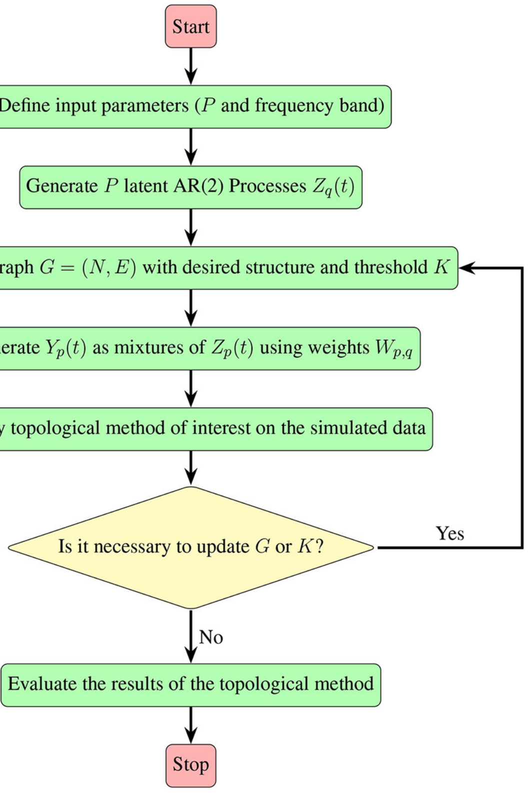
Start
Define input parameters ($P$ and frequency band)
Generate $P$ latent AR(2) Processes $Z_q(t)$
Define graph $G = (N, E)$ with desired structure and threshold $K$
Generate $Y_p(t)$ as mixtures of $Z_p(t)$ using weights $W_{p,q}$
Apply topological method of interest on the simulated data
Is it necessary to update $G$ or $K$?
Yes
No
Evaluate the results of the topological method
Stop

FIGURE 8
A visual guide demonstrating the simulation steps for multivariate time series, detailed in Section 2.2, along with the subsequent evaluation of the topological method of interest.

interest is on a specific connectivity pattern, such as a torus or a double torus, see left hand side of Figure 9.

## 3.1 Defining the graph structure

After defining the shape of the manifold of reference $\mathcal{M}$ for the dependence structure, we need to define a graph $G_{\mathcal{M}} = (N_{\mathcal{M}}, E_{\mathcal{M}})$ by sampling points $x_i$ from such a topological structure (i.e., $x_i \in N_{\mathcal{M}} \subset \mathcal{M}$), as can be seen in red dots in right hand side of Figure 9. It is necessary to properly define this graph, because it is needed to compute the mixing weights in Equation 5, which are a (decaying) function of the graph distance between nodes. The right hand side of Figure 9 display a Voronoi tessellation (green lines) over the manifold. For every pair of nodes (region centers) that share a common border, we add an edge in the set of edges $E_{\mathcal{M}}$ in the graph $G_{\mathcal{M}}$. Once the graph is properly defined, as demonstrated in Figure 12, our next step involves generating a latent AR(2) process for each node within the graph. Subsequently, as shown in Figure 11, we create a new multivariate time series by employing the weighted approach described in Equation 4.

## 3.2 Sampling points from a manifold

In order to define the graph structure, it is necessary to have a mechanism to sample points uniformly from a manifold. Multiple sampling procedures have been proposed in the literature, for example (Diaconis et al., 2013; Soize and Ghanem, 2016; Baggenstoss, 2017; Prado and Ritto, 2021). It may be relatively straightforward to sample from simple manifolds, such as circles or spheres, because it is simple to parameterize the entire manifold, for instance using polar or spherical coordinates. However, generally speaking, sampling from more intricate manifolds can be rather difficult. Uniformly sampling from manifolds extends beyond synthetic data generation and has broader implications in various domains. For instance, in the context of physical simulations, many systems have state spaces represented as manifolds. Achieving uniform sampling from these manifolds is critical for the effective study of system dynamics. Moreover, the relevance of uniform manifold sampling is apparent in biological and medical data analysis. Manifold representations are commonly used for complex data, including DNA structures, protein conformations, and brain functional connectivity. In these scenarios, non-uniform sampling can introduce biases into the analysis, impacting the quality and accuracy of results.

Our paper primarily focuses on two-dimensional surfaces, offering an novel approach based on quotient group representation for graph sampling with predefined patterns. This methodology contributes to solving the broader challenge of uniformly graph

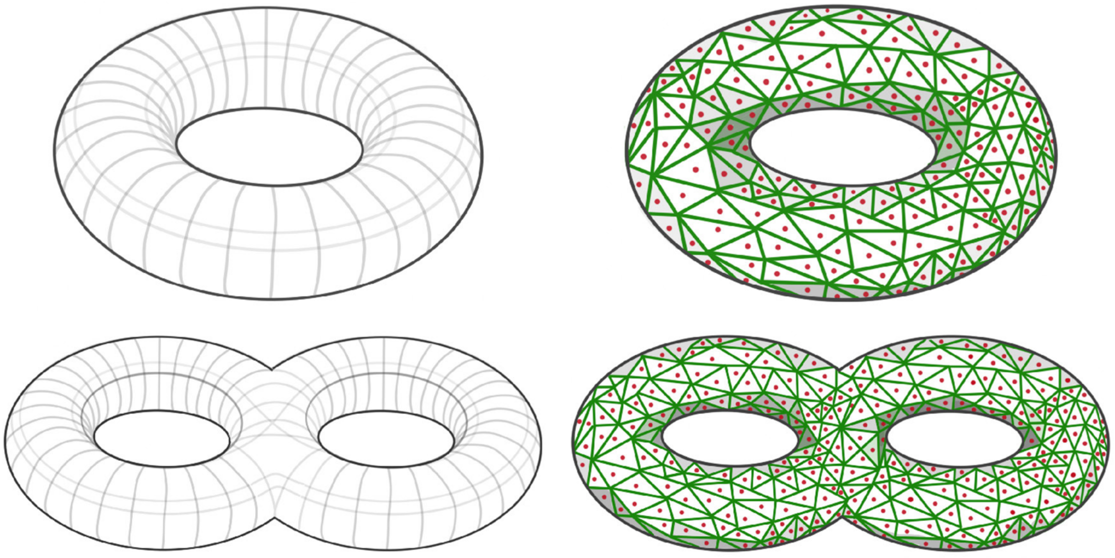

FIGURE 9
**(Top Left)** Simple torus dependence pattern; **(Top Right)** Sampled simple torus dependence pattern; **(Bottom Left)** Double torus dependence pattern; **(Bottom Right)** Sampled double torus dependence pattern.

sampling from manifolds, which finds relevance in fields beyond synthetic data generation. The sampling problem (from a given manifold) is closely related to Bertrand's Paradox and the principle of indifference. Indeed, for such problem to display a unique solution one has to properly define the problem at hand and what is meant by sampling in a non-ambiguous way (Jaynes, 1973; Marinoff, 1994). For instance, considering the one dimensional circle embedded in $\mathbb{R}^2$, every point $p_1 \in \mathcal{M}_1$ of the manifold $\mathcal{M}_1$ can be represented by an angle $\theta$:

$$\mathcal{M}_1 = \{(x,y)|x^2+y^2=r^2\}, \tag{16}$$

$$p_1 = \big(r\cos(\theta), r\sin(\theta)\big), \quad \theta \in [0, 2\pi]. \tag{17}$$

Similarly, considering the two dimensional sphere embedded in $\mathbb{R}^3$, every point $p_2 \in \mathcal{M}_2$ can be represented by a pair of coordinates:

$$\mathcal{M}_2 = \{(x,y,z)|x^2+y^2+z^2=r^2\}, \tag{18}$$

$$p_2 = \big(r\sin(\theta)\cos(\phi), r\sin(\theta)\sin(\phi), r\cos(\theta)\big), \phi \in [0, 2\pi],$$
$$\theta \in [-\pi/2, \pi/2]. \tag{19}$$

For instance, the parameterizations in Equation 17 correctly characterizes the circle. Hence, it is possible to sample points $p_i$ from the manifold $\mathcal{M}_1$ using the following procedure:

$$\theta_1 \sim \mathcal{U}(0, 2\pi), \tag{20}$$

$$p_1 = \big(r\cos(\theta_1), r\sin(\theta_1)\big). \tag{21}$$

Similarly, the parameterization in Equation 19 correctly characterizes the two dimensional sphere of radius $r$. Hence, to sample points $p_i$ from $\mathcal{M}_2$ we can use the following procedure:

$$\theta_2 \sim \mathcal{U}(0, 2\pi), \tag{22}$$

$$\phi_2 \sim \mathcal{U}(-\pi/2, \pi/2), \tag{23}$$

$$p_2 = \big(r\sin(\theta_2)\cos(\phi_2), r\sin(\theta_2)\sin(\phi_2), r\cos(\theta_2)\big). \tag{24}$$

Both examples presented above rely on parameterized immersions. When the chosen parameterization $f : S \to \mathcal{M}$ is not volume-preserving, the resulting sample will not be uniform. Indeed, this approach will lead to compressed regions being oversampled. Moreover, expanded regions can be undersampled, i.e., based on uniform sampling in the parameter space $S$ the sampled points in $\mathcal{M}$ are denser in regions where the parameterization $f$ has higher curvature (Diaconis et al., 2013). For example, in the first example the sample is uniform, however, in the second example the sample will not be uniform as there are compressed regions around the poles and expanded regions farther away from the poles. To remedy this issue, one potential approach is to generate a large sample using the previous approach then discarding some of the samples to correct for the compressed and expanded regions (see Diaconis et al., 2013). The rejection rate is chosen as a function of the determinant of the Jacobian of the parameterization $f : S \to \mathcal{M}$. Other interesting approaches have been proposed in the literature, such as Soize and Ghanem (2016) and Prado and Ritto (2021). However, these approaches provides tools for sampling only for simple manifolds.

When the manifold of interest is not very simple, such as a double torus in Figure 9 or even more complicated surfaces, it can be quite challenging to generate a sample using the above mentioned approach, since in some cases there may not be a global parameterization. Indeed, for smooth manifolds the parameterization is only guaranteed locally, to parameterize the entire manifold it is necessary to look at what is known as an Atlas

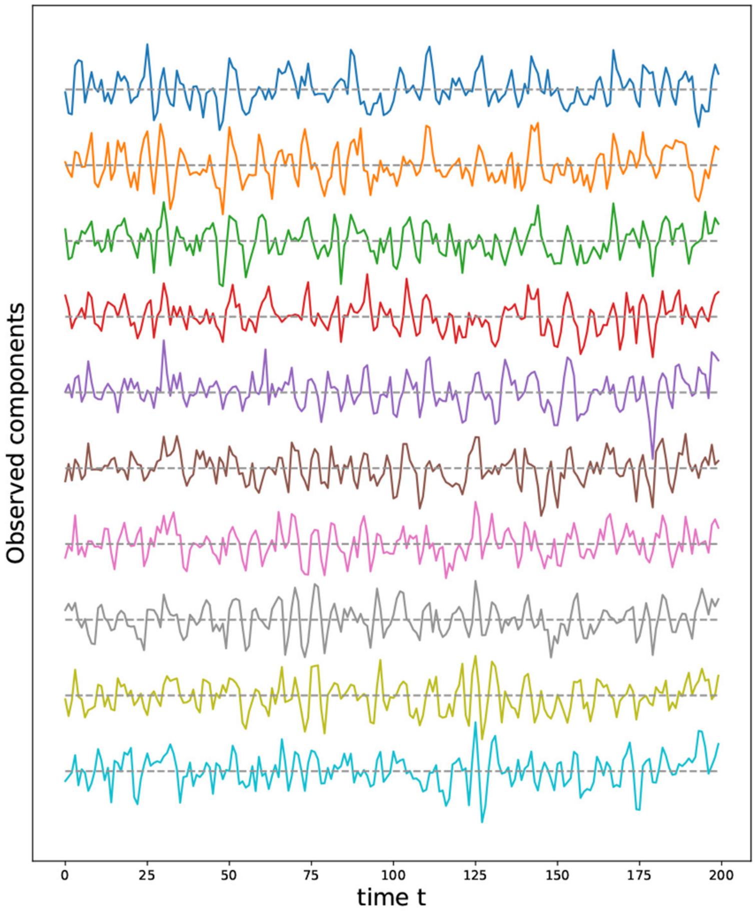

FIGURE 10
**(Left)** Polygonal quotient group representation for a cylinder, sphere, torus and a double torus. **(Right)** Corresponding 3D visualization.

representation of the manifold, refer to Tu (2008), for more details regarding the parameterization of manifolds.

For these reasons, we propose the following method to sample from a certain set of two-dimensional manifolds (surfaces), such as the sphere, torus, and double torus, which is based on the representation of these manifolds using quotient space of polygons (see Figure 10).

The advantage behind this proposed representation lies in the simplicity with which we can sample from the corresponding manifold. Indeed, given a polygonal representation, one can sample uniformly from the flat polygons, then identify the nodes present on equivalent edges. In Figure 12, we illustrate the process of constructing the graph representing a Torus manifold from the initial sample, taken from a rectangle, to the graph after node identification.

Using the same approach as described in Equation 4, we generate the torus multivariate time series, as seen in Figure 11. After estimating the coherence matrix for this multivariate time series at middle and high frequency bands we compute the persistence diagrams and hence produce the following results, as displayed in Figure 12. This figure clearly shows the topological features of the targeted torus structure. Indeed, the two off-diagonal orange dots represent the two one dimensional wholes in a torus (circles surrounding each of the wholes) and one off-diagonal green dot representing the two-dimensional whole (or cavity inside of the torus), indeed it is known that the first three Betti numbers of a torus are $\beta_0 = 1$, $\beta_1 = 2, \beta_2 = 1$.

The number of points to sample from the manifold depends on the target topological feature that is being investigated. Generally speaking, in order to detect features in the persistence diagram, the diameter of every subgraph surrounding the topological feature of interest needs to be at least of the same magnitude or larger than twice the constant $K$ in the mixing equation. This is an important point to keep in mind. In Section 2.3.1, to detect the main cycle we need the diameter $(P/4)$ to be larger than $2K$, if $K = 2$ then we need to chose $P \geq 16$, if $K = 3$ then we need to chose $P \geq 24$ etc. In Figure 2.3.2, to detect the main cycles we need the diameter of the smallest subgraph surrounding one of the main cycles (roughly $P/8$ if both cycles are of comparable size) to be larger than $2K$, i.e., if $K = 2$ then we need to chose at least $P \sim 32$. For this reason, we can detect only the main cycles and the secondary cycles appear like noise in the persistence diagrams (see Figures 5, 7).

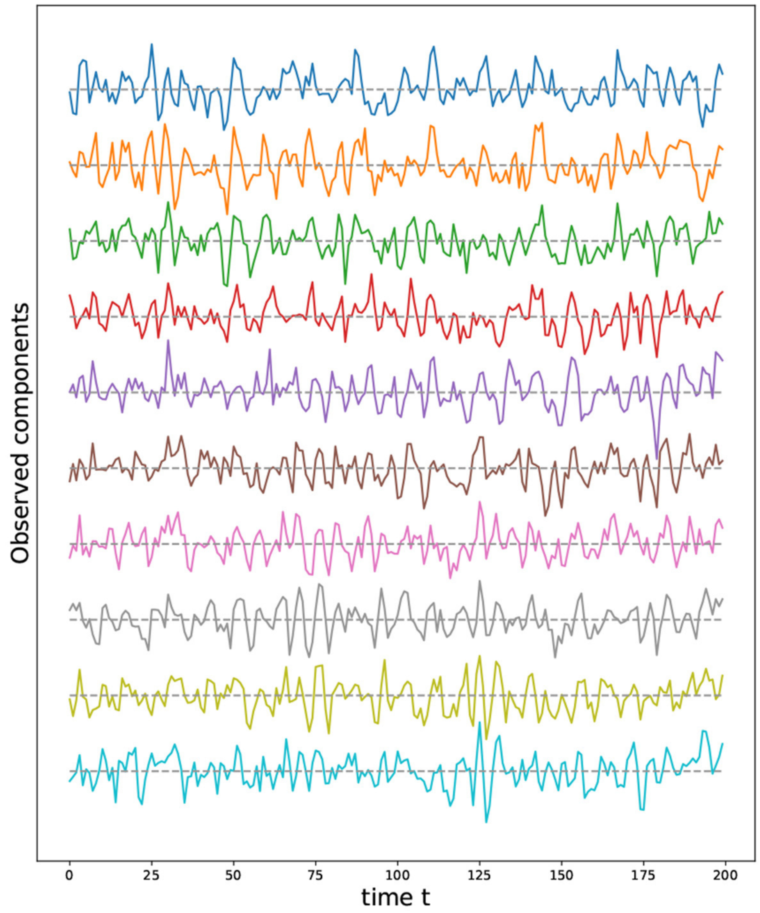

FIGURE 11
First ten time series components from a multivariate time series model generated using a torus structure with an initial grid of 9 by 17 nodes (i.e., $P = 153$).

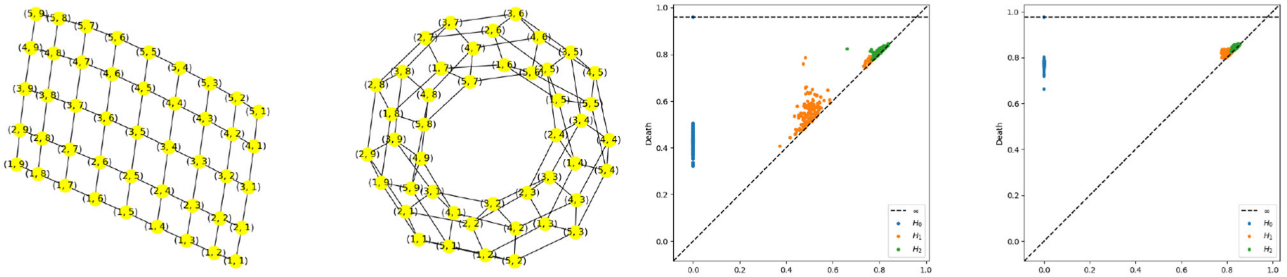

FIGURE 12
Torus graph pattern based on polygonal quotient group representation. Left most: Regular sample from a rectangle. Second from the left: Torus graph after node identification. Persistence diagram based on a multivariate time series generated using a torus structure with an initial grid of 9 by 17 nodes,i.e., $P = 153$, and $K = 3$, with the middle-frequency band shown as the third from the left and the high-frequency band as the fourth from the left.

## 3.3 A robustness study: navigating noise effects

Our aim here is to study the sensitivity of our approach to noise. The observed signal, denoted as $Y(t)$, is composed of two components: the underlying signal or stochastic process, $S(t)$, and additive noise, $N(t)$. While $Y(t)$ is what we directly observe, $S(t)$ remains hidden from our measurements, characterized by a variance of $\sigma_S^2$. In contrast, $N(t)$ is independent of $S(t)$ and introduces noise with a variance of $\sigma_N^2$. To assess this sensitivity, we utilize the signal-to-noise ratio ($SNR$), defined as $SNR = \frac{\sigma_S^2}{\sigma_N^2}$, which quantifies the relative strength of the underlying signal to the

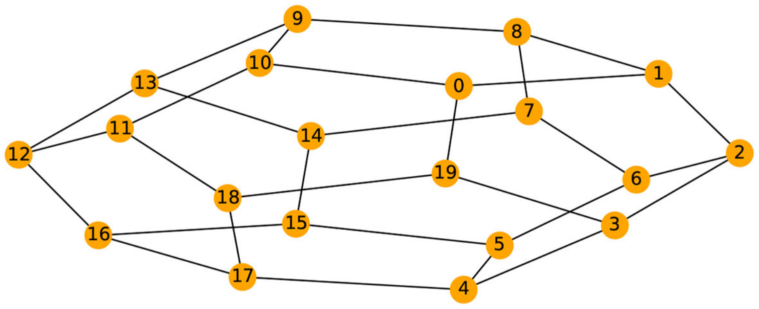

FIGURE 13
Example of a spherical dependence pattern known as a Dodecahedron.

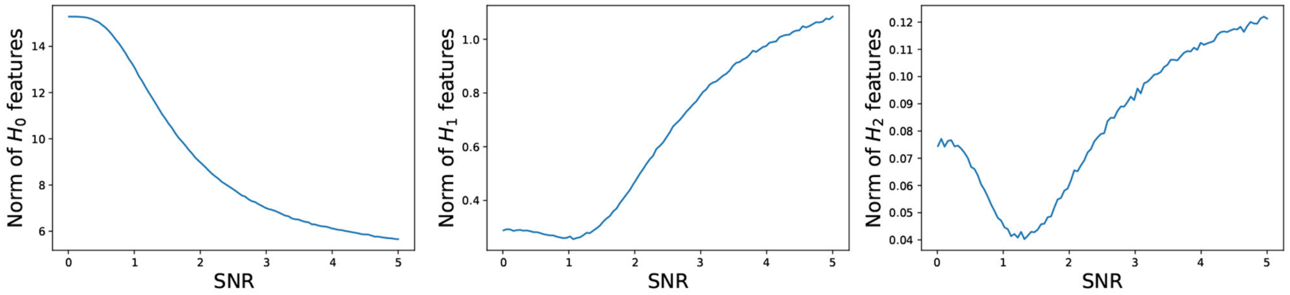

FIGURE 14
Total persistence as a function of the signal to noise ratio. **(Left)** 0-dimensional homology; **(Middle)** 1-dimensional homology; **(Right)** 2-dimensional homology. Total persistence in the $y$-axis and signal to noise ratio $SNR$ in the $x$-axis. The plots are based on the average total persistence for 1,000 replicates.

additive noise. To assess the effect of the noise on the topological features of the dependence pattern in the underlying signal, we generate multivariate times series data from a structure that has two dimensional feature. i.e., a spherical structure (see Figure 13). Define the total persistence to be the norm of the persistence diagram' features as follows:

$$P_k = \sum_{i \in PD_k} (d_i^k - b_i^k) \tag{25}$$

where $b_i^k$ and $d_i^k$ represent, respectively, the birth and death of the $i$-th $k$-dimensional topological feature in the persistence diagram. For every dimension $k$, the total persistence $P_k$ is defined to be the sum of the persistence of all $k$-dimensional features in the persistence diagram. In what follows, we study the behavior of the total persistence $P_k$ as a function of the signal to noise ratio (see Figure 14).

The persistence $P_0$ of the 0-dimensional features decreases as the signal to noise ratio grows, which is to be expected because at low $SNR$, the time series components are mostly independent, resulting in large mutual distances and many unconnected components and at high $SNR$, the time series components are mostly dependent, resulting in smaller mutual distances, i.e., fewer connected components. On the other hand, the persistence $P_k$ of the 1- and 2-dimensional features increases as the signal to noise ratio grows, which makes sense. At low $SNR$, the time series components are independent, and the connectivity pattern is not visible, but at high $SNR$, the time series components are mostly dependent according to the spherical structure, i.e., more 1- and 2-dimensional features.

## 4 Statistical inference in TDA—A simulation approach

There are many disorders that can alter the connectivity of the brain such as Alzheimer's disease, Parkinson's disease, ADHD. These conditions are known to alter the topology of the brain's connectivity structure by creating holes, cavities or other patterns in the connectivity network. We will develop a statistical inference method, via simulations of multivariate time series, for differentiating between two topological patterns that differ in their one dimensional homology structure. Based on the idea developed in Sections 2.3.1 and 2.3.2, we generate $N = 50$ samples from one model $\mathcal{M}_1$ with one main cycle in its dependence pattern, and $N = 50$ samples from another model $\mathcal{M}_2$ with two main cycles in its dependence

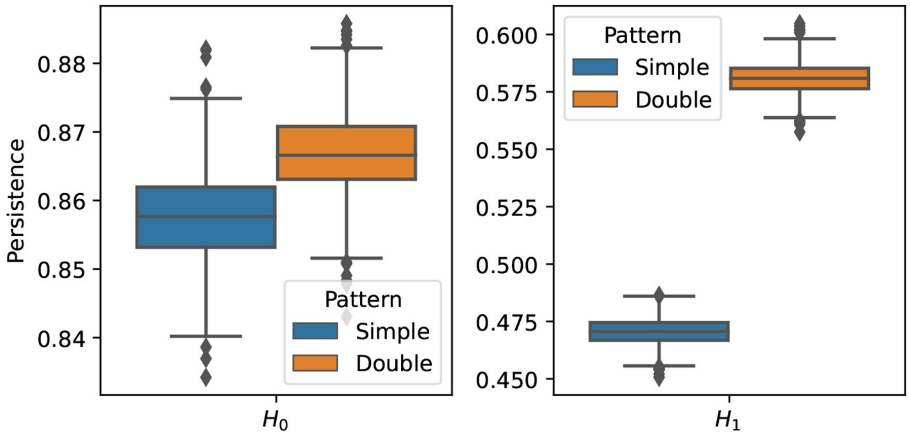

FIGURE 15
Boxplot of the one **(Left)** and two **(Right)** homology groups topological summaries for both group simple main cycle (blue) and double main cycle (orange) based on $B = 1,000$ bootstrap samples.

pattern:

$$Y^{(1,i)}(t) = W^{(1)}Z^{1,i}(t) + \epsilon^{1,i}(t), i = 1, \dots, N \tag{26}$$

$$Y^{(2,i)}(t) = W^{(2)}Z^{1,i}(t) + \epsilon^{2,i}(t), i = 1, \dots, N \tag{27}$$

where $W^{(1)}$ and $W^{(2)}$ are respectively the mixing weights for model one and two as defined in Sections 2.3.1 and 2.3.2, $Z^{1,i}(t)$ and $Z^{2,i}(t)$ are the iid latent processes, $\epsilon^{1,i}(t)$ and $\epsilon^{2,i}(t)$ are the additive Gaussian noise.

After generating the time series for both groups, we compute the corresponding persistence diagrams then we compute a topological summary, total persistence as described in the previous section, i.e., $T_i^1$ and $T_i^2$ for $i = 1, \dots, N$. In order to compare the topologies of both groups we compute the group mean of these summaries for the one/two dimensional homology etc., and then assess the variability using a bootstrap approach:

1. Draw $T_1^{1*}, \dots, T_N^{1*}$ and $T_1^{2*}, \dots, T_N^{2*}$ from the empirical distribution based on $T_1^1, \dots, T_N^1$ and $T_1^2, \dots, T_N^2$
2. Compute the group mean $\widehat{T}_b^{(1*)} = \frac{1}{N}\sum_{i=1}^N T_i^{1*}$ and $\widehat{T}_b^{(2*)} = \frac{1}{N}\sum_{i=1}^N T_i^{2*}$ for the one and two homology groups.
3. Repeat B times the previous two steps.
4. Visualize the boxplot of the bootstrap samples.

The results of the procedure above is displayed in Figure 15. It can be seen that the two groups differ mainly in their cyclic structure (1-dimensional homology), high orange boxplot means more persistence of such features but not in their connected components structure (0-dimensional homology). In conclusion, based on the simulated data sets generated from models $\mathcal{M}_1$ and $\mathcal{M}_2$, it is possible to generate multivariate time series data with varying cyclic behavior in its dependence patterns.

## 5 Conclusion

This article presents an innovative approach for simulating multivariate time series data with predetermined cyclic dependency structures, which is crucial for evaluating the effectiveness of proposed Topological Data Analysis (TDA) techniques. To the best of our knowledge, our proposed method is the first to utilize mixtures of AR(2) processes to create frequency-specific dependency structures. Since our method is fairly general, it may be applied in a wide variety of situations. It can also be utilized to produce higher dimensional topological features. The proposed ideas were illustrated on examples with different cycle counts. A novel procedure based on the quotient group representation to create even more complex dependency patterns such as a torus is presented. To investigate the effect of the variance of the additive noise on the topological features, we conducted a thorough sensitivity analysis. Finally, we gave a demonstration of how our method can be applied to make simulation-based inference.

## Data availability statement

The data utilized in this study are simulated and generated for the purpose of experimentation. As such, they do not represent real-world observations. The simulation details and code used to generate the data are available upon request from the corresponding author.

## Author contributions

AE-Y: Conceptualization, Methodology, Software, Visualization, Writing – original draft, Formal analysis, Investigation, Validation. MC: Supervision, Writing – review & editing. HO: Funding acquisition, Resources, Supervision, Writing – review & editing.

## Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This research was supported by funding from King Abdullah University of Science and Technology (CRG) and the US National Institutes of Health (NIH MH133614).

## Acknowledgments

The authors gratefully acknowledge Sarah Aracid (KAUST and University of the Philippines) for her invaluable help with the figures and artwork.

## 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.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## References

Adams, H., Emerson, T., Kirby, M., Neville, R., Peterson, C., and Shipman, P. (2017). Persistence images: A stable vector representation of persistent homology. *J. Mach. Learn. Res.* 18, 1–35.

Baggenstoss, P. M. (2017). Uniform manifold sampling (ums): sampling the maximum entropy pdf. *IEEE Trans. Signal Process.* 65, 2455–2470. doi: 10.1109/TSP.2017.2659645

Bassett, D., and Bullmore, E. (2009). Human brain networks in health and disease. *Curr. Opin. Neurol.* 22, 340–347. doi: 10.1097/WCO.0b013e32832d93dd

Bowyer, S. M. (2016). Coherence a measure of the brain networks: past and present. *Neuropsychiat. Electrophysiol.* 2, 1–12. doi: 10.1186/s40810-015-0015-7

Brechmann, E. C., and Czado, C. (2014). Copar–multivariate time series modeling using the copula autoregressive model. *Appl. Stoch. Models Bus. Ind.* 31, 425–437. doi: 10.1002/asmb.2043

Bubenik, P. (2015). Statistical topological data analysis using persistence landscapes. *J. Mach. Learn. Res.* 16, 77–102.

Bullmore, E., and Sporns, O. (2009). Complex brain networks: graph theoretical analysis of structural and functional systems. Nature reviews. *Neuroscience* 10, 186–198. doi: 10.1038/nrn2575

Chazal, F., and Michel, B. (2021). An introduction to topological data analysis: fundamental and practical aspects for data scientists. *Front. Artif. Intell.* 4, 667963. doi: 10.3389/frai.2021.667963

Diaconis, P., Holmes, S., and Shahshahani, M. (2013). Sampling from a manifold. *Institute of Mathe. Stat.* 10:102–125. doi: 10.1214/12-IMSCOLL1006

Edelsbrunner, H., and Harer, J. (2008). Persistent homology–a survey. *Discrete Computat. Geo.* 453:257–282. doi: 10.1090/conm/453/08802

Edelsbrunner, H., Letscher, D., and Zomorodian, A. (2002). Topological persistence and simplification. *Discrete Computat. Geo.* 28, 511–533. doi: 10.1007/s00454-002-2885-2

El-Yaagoubi, A. B., Chung, M. K., and Ombao, H. (2023). Topological data analysis for multivariate time series data. *Entropy* 25, 1509. doi: 10.3390/e25111509

Fan, L., Li, H., Zhuo, J., Zhang, Y., Wang, J., Chen, L., et al. (2016). The human brainnetome atlas: a new brain atlas based on connectional architecture. *Cerebral Cortex* 26, 3508–3526. doi: 10.1093/cercor/bhw157

Fathian, A., Jamali, Y., Raoufy, M. R., and Alzheimer's Disease Neuroimaging Initiative. (2022). The trend of disruption in the functional brain network topology of alzheimer's disease. *Sci. Rep.* 12, 14998. doi: 10.21203/rs.3.rs-440491/v1

Friston, K. J. (2011). Functional and effective connectivity: a review. *Brain Connect.* 1, 13–36. doi: 10.1089/brain.2011.0008

Gholizadeh, S., and Zadrozny, W. (2018). A short survey of topological data analysis in time series and systems analysis. *ArXiv.* [preprint] ArXiv:abs/1809.10745.

Ghrist, R. (2008). Barcodes: The persistent topology of data. *Bull. New Ser. Am. Math. Soc.* 45, 61–75. doi: 10.1090/S0273-0979-07-01191-3

Gorrostieta, C., Ombao, H., and Sachs, R. V. (2018). Time-dependent dual frequency coherence in multivariate non-stationary time series. *J. Time Series Analy.* 40, 3–22. doi: 10.1111/jtsa.12408

Granados-Garcia, G., Fiecas, M., Babak, S., Fortin, N. J., and Ombao, H. (2021). Brain waves analysis via a non-parametric bayesian mixture of autoregressive kernels. *Computat. Statist. Data Analy.* 2021, 107409. doi: 10.1016/j.csda.2021.107409

Hausmann, J.-C. (2016). *On the Vietoris-Rips Complexes and a Cohomology Theory for Metric Spaces*. Princeton, NJ: Princeton University Press.

Henry, T. R., Duffy, K. A., Rudolph, M. D., Nebel, M. B., Mostofsky, S. H., and Cohen, J. R. (2020). Bridging global and local topology in whole-brain networks using the network statistic jackknife. *Network Neurosci.* 4, 70–88. doi: 10.1162/netn_a_00109

Hilgetag, C., and Goulas, A. (2015). Is the brain really a small-world network? *Brain Struct. Funct.* 221, 6. doi: 10.1007/s00429-015-1035-6

Hu, L., Fortin, N. J., and Ombao, H. (2019). Modeling high-dimensional multichannel brain signals. *Stat. Biosci.* 11:91–126. doi: 10.1007/s12561-017-9210-3

Jaynes, E. T. (1973). The well-posed problem. *Foundat. Phys.* 3, 477–493. doi: 10.1007/BF00709116

Leykam, D., and Angelakis, D. G. (2023). Topological data analysis and machine learning. *Adv. Phys.* 8, 2202331. doi: 10.1080/23746149.2023.2202331

Marinoff, L. (1994). A resolution of bertrand's paradox. *Philos. Sci.* 61, 1–24. doi: 10.1086/289777

Mohammadi, H., Challenor, P., and Goodfellow, M. (2019). Emulating dynamic non-linear simulators using gaussian processes. *Comput. Statist. Data Analy.*139, 178–196. doi: 10.1016/j.csda.2019.05.006

Motta, F. (2018). *Topological Data Analysis: Developments and Applications*, 369–391.

Muldoon, S., Bridgeford, E., and Bassett, D. (2016). Small-world propensity and weighted brain networks. *Sci. Rep.* 6, 22057. doi: 10.1038/srep22057

Nakagawa, T., and Deco, G. (2015). *Multiscale Brain Connectivity*. New York: Springer, 1768-1770.

Ombao, H., and Pinto, M. (2022). "Spectral dependence," in *Econometrics and Statistics*.

Ombao, H., and Van Bellegem, S. (2008). Evolutionary coherence of nonstationary signals. *IEEE Trans. Signal Proc.* 56, 2259–2266. doi: 10.1109/TSP.2007.914341

Pessoa, L. (2014). Understanding brain networks and brain organization. *Phys. Life Rev.* 11, 400–435. doi: 10.1016/j.plrev.2014.03.005

Prado, L. S., and Ritto, T. G. (2021). Data driven dirichlet sampling on manifolds. *J. Comput. Phys.* 444, 110583. doi: 10.1016/j.jcp.2021.110583

Prado, R., West, M., and Krystal, A. (2001). Multichannel electroencephalographic analyses via dynamic regression models with time-varying lag-lead structure. *J. Royal Statist. Soc. Series C* 50, 95–109. doi: 10.1111/1467-9876.00222

Shumway, R. H. and Stoffer, D. S. (2017). *Time Series Analysis and Its Applications*. New York: Springer.

Snow, D. (2020). "Mtss-gan: Multivariate time series simulation generative adversarial networks," in *Social Science Research Network*.

Soize, C., and Ghanem, R. (2016). Data-driven probability concentration and sampling on manifold. *J. Comput. Phys.* 321, 242–258. doi: 10.1016/j.jcp.2016.05.044

Sporns, O. (2013). Structure and function of complex brain networks. *Dialogues Clin. Neurosci.* 15, 247–262. doi: 10.31887/DCNS.2013.15.3/osporns

Tu, L. (2008). *An Introduction to Manifolds.*

Umeda, Y. (2017). Time series classification via topological data analysis. *Trans. Japan. Soc. Artif. Intelli.* 32, D-G72_1-12. doi: 10.1527/tjsai.D-G72

Wasserman, L. (2018). Topological data analysis. *Ann. Rev. Statist. Appl.* 5, 501–532. doi: 10.1146/annurev-statistics-031017-100045

Xu, X., Drougard, N., and Roy, R. N. (2021). Topological data analysis as a new tool for eeg processing. *Front. Neurosci.* 15, 761703. doi: 10.3389/fnins.2021.761703