Leveraging the Power of Signatures for the Construction of Topological Complexes for the Analysis of Multivariate Complex Dynamics

Stéphane Chrétien, Ben Gao, Astrid Thébault Guiochon, Rémi Vaucher

To cite this version:

Stéphane Chrétien, Ben Gao, Astrid Thébault Guiochon, Rémi Vaucher. Leveraging the Power of Signatures for the Construction of Topological Complexes for the Analysis of Multivariate Complex Dynamics. The Twelfth International Conference on Complex Networks and their Applications, Nov 2023, Menton Riviera, France. pp.283-294, 10.1007/978-3-031-53468-3_24 . hal-04590221

HAL Id: hal-04590221
https://hal.univ-lyon2.fr/hal-04590221v1

Submitted on 30 May 2024

Leveraging the power of Signatures for the construction of topological complexes for the analysis of multivariate complex dynamics

Stéphane Chrétien[1,4], Ben Gao[1,2], Astrid Thébault Guiochon[3], and Rémi Vaucher[1,2]

[1] Laboratoire ERIC, Université Lyon 2,
`stephane.chretien,ben.gao,remi.vaucher@univ-lyon2.fr`,
WWW home page: `https://sites.google.com/site/stephanegchretien/home`
[2] HALIAS,
`ben.gao,remi.vaucher@halias.fr`,
WWW home page: `https://remivaucher.github.io`
[3] Université Lyon 2,
`a.thebaultguiochon@univ-lyon2.fr`,
WWW home page: `https://linktr.ee/astridthebaultg`
[4] The Alan Turing Institute, London, UK.

Abstract. Topological Data Analysis is a field of great interest in many applications such as finance or neuroscience. The goal of the present paper is to propose a novel approach to building simplicial complexes that capture the multiway ordered interactions in the components of high-dimensional time series using the theory of Signatures. Signatures represent one of the most powerful transforms for extracting group-wise structural features and we put them to work in the task of discovering statistically meaningful simplices from a complex that we estimate sequentially. Numerical experiments on an fMRI dataset illustrates the efficiency and relevance of our approach.

Keywords: Topological Data Analysis, Signatures, simplicial complex, multiway interactions, high-dimensional time series, fMRI data analysis

1 Introduction

Topological Data Analysis (TDA) is a new field with a wide range of application in fields such as finance, neuroscience, medicine, etc. TDA addresses the problem of accounting for groupwise interactions in the data and therefore opens very promising prospects to better apprehend complex phenomena than models relying on pairwise interactions only. Several tools from algebraic topology, such as homology groups, homotopy groups, Betti numbers, etc can be put to work in building a set of relevant features that can capture the intricate nature of dependencies.

Some compelling examples of the benefits of using topological features appear regularly in the literature. In [17] the homological features of brain functional networks are shown to take different values in two states depending on the absorption of some drug. More generally, it is shown in [21] that homological cycles in structural brain networks finds connections between regions of early and late evolutionary origin. TDA can also be efficiently used in dynamical settings as well. One very intriguing example is change detection as illustrated in the study of functional brain networks conditioned in different tasks [22]. In [12] it is investigated how speech-related brain regions connectivity changes in different scenarios of speech perception.

In the present paper, we focus on the the analysis of dynamical high dimensional phenomena and on the problem of constructing associated relevant topological structures with the aim of proposing news computational tools for deepening our understanding of the higher order structures hidden in time series data. Our main contribution is a new approach to building statistically informed simplicial complexes and possibly more general structures.

Our two main tools will be the basic objects of TDA and Signature theory. Signatures were recently proposed as a very powerfull feature map for time series and dynamical systems in [5,7,14]. The introduction of Signatures for building topological structures for high dimensional dynamical phenomena is new and appears as a key and very natural ingredient that can accurately account for orientations of the various simplices in the complex at hand while capturing the main shape features from the dynamics. Using Signature in a statistical/machine learning context is an approach which is adopted in a growing number of applications nowadays [7] and our work is also intended to illustrate the relevance of Signature theory combined with statistics/machine learning for building a higher order interaction modelling framework.

In mathematical terms, our proposal is based on the assumption that k-simplices are simply sets of k nodes with there time series attached to them, with an orientation prescribed by the ordering of the nodes in computing their associated k-Signature. Recall that the orientation encoded in the computation of the associated signature carries potentially very interesting interpretation about the causal dependencies of the times series [11]. In the next step, the relevance of incorporating a simplex into our simplicial complex is assessed using a purely statistical procedure: each oriented k-simplex is associated with a corresponding k-Signature that is included into a set of multivariate features that is used to predict the Signatures of all the other potential simplices. More precisely, our construction is a generalisation of the approach developed by Meinshausen and Bühlmann in [16] for Gaussian Graphical Models. Simplices that are selected from the set of potential Simplices whose Signature can predict the Signature of a target simplex in terms of confidently regressing or predicting[5] the Signature associated with this target are included as candidates for being considered as adjacent to this target. Using this procedure, we obtain a construction of a simplicial complex that accurately incorporates the statistical relationships between all the simplices in terms of regression or prediction, while keeping track of the inherent orientations of the simplices.

---

[5] for time dependent Signatures

The plan of the paper is as follows. In Section 2, we recall the necessary background on topological data analysis and Signature theory. In Section 3, we present our method for constructing the simplicial complex using the Signatures of the simplices and the LASSO algorithm. In Section 4, we present our numerical experiments on real datasets. A conclusion section completes the paper.

2 Background on Signatures and topology

In this section, we summarise the mathematical prerequisites from topology and the theory of Signatures.

2.1 Recalls on the theory of Signatures

The theory of Signatures is a new topic of growing interest that emerged as a sub-branch of the theory of rough paths [15,10,9] which has a long history in mathematics and control that may have started with the work of Chen [4]. Rough paths provide a new framework for the analysis of stochastic processes and permitted to resolve various open problems, including the existence of a solution to the KPZ equation in mathematical physics, a result for which Martin Heirer was awarded the Fields medal [6]. Recently, this theory developed as a new tool for the analysis of signals in the area of Machine Learning [5,7,14] where remarkable performance was achieved for a series of difficult pratical problems including the analysis of financial data, medical data and textual data [14,13]. Lately, an intriguing relationship with recurrent neural networks was exhibited using the viewpoint of control theory [8].

Let us now turn to the definition and some interesting properties of Signatures. Consider a d-dimensional path $X = (X^1, X^2, ..., X^d) : \mathbb{R} \to \mathbb{R}^d$. Then, the (truncated)[6] **signature** of X on $[a,b]$ is an object in $\mathcal{T}(\mathbb{R}^d) = \mathbb{R}^d \oplus \mathbb{R}^{d\times d} \oplus \mathbb{R}^{d\times d\times d} \oplus \dots$, defined, for $j \in \mathbb{N}^*$ by

$$\begin{aligned}(S_{[a,b]}(X))_{i_1,i_2,\dots,i_j} &:= S^{i_1 i_2 \dots i_j}_{[a,b]}(X) \\ &= \underset{a\leq s_1\leq s_2\leq\cdots\leq s_j\leq b}{\int\cdots\int} dX^{i_1}_{s_1} dX^{i_2}_{s_2} \dots dX^{i_j}_{s_j} \end{aligned} \tag{1}$$

which lies in $\mathbb{R}^{\overbrace{d \times d \times \cdots \times d}^{j \text{ times}}}$.

Chen's identity is a very useful result that allows to compute the Signature recursively based on linear interpolation of observed values of a trajectory.

Theorem 1 (Chen's identity). *Let $X : [a,b] \to \mathbb{R}^d$ and $Y : [b,c] \to \mathbb{R}^d$. Consider the* ***concatenation*** *of X and Y (noted $X * Y$) defined by:*

$$\begin{aligned}(X * Y) &: [a,c] \to \mathbb{R}^d \\ & t \mapsto \begin{cases} X(t) & , \quad t \in [a,b] \\ X(b) - Y(b) + Y(t) \ , & t \in [b,c]. \end{cases}\end{aligned}$$

---

[6] The k-truncated version of the signature is $S^{(1)}(X) \oplus S^{(2)}(X) \oplus \cdots \oplus S^{(k)}(X)$

Then:

$$S_{[a,c]}(X * Y) = S_{[a,b]}(X) \otimes S_{[b,c]}(Y)$$

Augmentation of a path The Signature defines X uniquely on $[a, b]$ close to *tree-like equivalence* (i.e. there exist $I, J \subset [a, b]$, such that $X|_I(t) = X|_J(b-t)$).

Proposition 1. *$S_{[a,b]}(X)$ define X uniquely on $[a, b]$ if there exists $1 \leq i \leq d$ such that X^i is strictly monotonic on $[a, b]$.*

This result leads to consider the **time-augmented path** $\tilde{X}$ associated with X, defined as $\tilde{X} = (t, X^1, X^2, \ldots, X^d)$ in order to ensure the unicity of $S(X)$. Another augmentation will be useful for our work, namely the Lead-Lag augmentation.

Definition 1. *Consider a d-dimensional path X with $T+1$ timesteps. The **lead-lag augmentation** of X is a $2d$-dimensional path $X_{lead,lag} = (X^{Lead}, X^{Lag})$ with $2T + 1$ time steps such that:*

$$\begin{aligned} X^{Lead} &= \{X(0), X(1), X(1), X(2), \ldots, X(T), X(T)\} \\ X^{Lag} &= \{X(0), X(0), X(1), X(1), \ldots, X(T-1), X(T)\} \end{aligned}$$

The Lead-Lag augmentation was found to play an important role in many machine learning applications [7].

2.2 Recalls on topology

We now turn to some useful definitions from topology. Consider a set of n vertices $V = \{v_1, \ldots, v_n\}$.

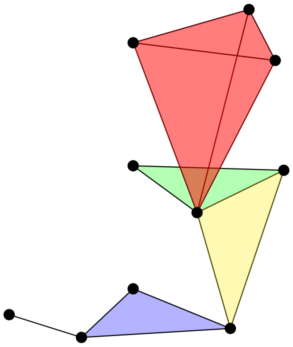
Definition 2. *For $k < n$, a k-**simplex** σ_k of V is the collection of a subset V_k of length $k+1$ and all its subsets. The **geometric realization** of a k-simplex is the convex hull C of $k+1$ points, such that $\dim(C) = k$. A face (of dimension l) σ_k is a collection of set in σ_k that form a l-simplex ($l \leq k$).*

Definition 3. *A **simplicial complex** $\mathcal{C}$ on V is a collection of simplices (of V) such that for every $\sigma^i \in \mathcal{C}$, there exists j with $\sigma^i \cap \sigma^j$ a sub-simplex of both σ^i and σ^j.*

Definition 4. *The **dimension** of $\mathcal{C}$ is the dimension of the highest simplex in $\mathcal{C}$ (i.e. the highest k such that there exist a k-simplex in $\mathcal{C}$).*

Definition 5. *An **orientation** on $\mathcal{C}$ is a order of the vertices for every simplicies in $\mathcal{C}$. We use the notation $[v_1, v_2, v_3]$ to denote an oriented 3-simplex.*

From a geometric point of view, $\mathcal{C}$ is constructed by attaching a group of simplices to each other by binding them with a shared face.



Definition 6. *Let Δ be a simplicial complex and σ one of its k-sub-simplices. The **neighborhood** of v in Δ is the set $Lk(\sigma, \Delta)$ of all the k-sub-simplices τ of Δ such that:*

- $\sigma \cap \tau = \emptyset$
- $\sigma \cup \tau$ *is a face of* Δ.

Especially, the neighborhood of a vertex v is all the vertices v_i such that the edge $\{v, v_i\} \in \Delta$.

The goal of our work is to build a simplicial complex encoding the groupwise interactions explaining the dependencies inside high dimensional time series. The main ingredient in our approach is to make use of relevance measures in predicting *simplex indexed*-groups of time series using other *simplex indexed*-groups of time series as a criterion for selecting potential higher and higher dimensional simplices for (or for the sake of mitigating the computational complexity, sequential greedy) aggregations. We turn to the description of our approach in the next section.

3 Building simplicial complexes between time series

3.1 Presentation of the method

In the same spirit as for Gaussian graphical models [16], our goal is to infer a higher order model from data using a model selection based numerical procedure. One brute force method is to use sparse solutions of LASSO type estimators that can be employed to predict all sub-groups of times series based on all other groups of time series.

In the present paper, our goal is to propose a better structured solution to the problem of capturing interesting structures in the dependency relationship among the various components of high dimensional time series. Interesting types of structure often come from Topological Data Analysis (TDA) as presented in e.g. [3]. Inferring such structures is however extremely computationally extensive, and even more so if we account for the necessity of using Cross-Validation types of hyper parameter calibration procedures.

Our proposal is to adopt a principled sequential approach to simplicial complexes estimation. In our approach, the simplices that will be incorporated into the simplicial complex are chosen among completions of existing simplices into one order higher simplices, hence allowing to stratify the construction by the complexity of the interactions. The selection is performed using the LASSO, complemented by inspection of the R^2 criterion. One key ingredient of our construction is the use of the Signature transform as features for prediction. Signatures bring an essential information to the construction of our simplicial complex, namely orientation, since, Signatures encode the order in which the times series are integrated in (1). Using the orientation can be instrumental for the interpretation of the interactions among the various components of high dimensional time series, and noticing any difference in the prediction capabilities of two different orderings might be extremely useful in practice.

We now turn to the details of the implementation.

3.2 Our greedy order stratified algorithm:

Consider a group of d (augmented) time series $\mathcal{G} = \{X^1, X^2, \ldots, X^d\}$ with $T+1$ timesteps each of the form $X^i = [(0, X^i(0)), (t_1, X^i(t_1)), \ldots, (t_T, X^i(t_T))]$.

In this section, we introduce our main contribution, namely the construction of a simplicial complex that encodes the multiway dependencies of the various dimensions in a high-dimensional time series. We also discuss a greedy algorithm for sequentially building the sought for simplicial complex that mitigates the computational complexity. As mentioned earlier in the introduction, the main principle of the our algorithm is to build consistent simplices within groups of time series using regression or prediction error measures. In our implementation, we chose to present the *Signature prediction* version, consisting of predicting the Signature at a simplex as a linear function of the signatures at other simplices of various orders.

In order to keep control on the computational complexity of the method, we now present a lighter greedy algorithm. For any $t \in [\![0, T-L]\!]$ with L to be specified.

Closure of a simplicial complex.

In our framework, we need to introduce the following definition.

Definition 7. *Let $\mathcal{C}$ denote a simplicial complex. We denote by $\overline{\mathcal{C}}$ the simplicial complex consisting in appending all k-simplices whose **single** incorporation results in creating a simplex of order $k+1$ using the simplices already present in $\mathcal{C}$ only.*

The sequential algorithm.

We now define the steps of our sequential greedy method as follows.

```
**Data:** Set $\ell = 1$ and set $\mathcal{C}^{(1)} = \{1, \dots, d\}$.
**Result:** The time series interaction simplicial complex
while No more simplex is selected do
    Select an (augmented) $k$-subset of nodes $C^J = \{X^{j_1}, X^{j_2}, \dots, X^{j_k}\}$, with
      $J = \{j_1, \dots, j_k\}$ in $\mathcal{C}^{(\ell)}$, and compute $S_{[t,t+L]}(C^J)$.;
    For every (augmented) $k'$-combination $C^{J'} = \{X^{j'_1}, X^{j'_2}, \dots, X^{j'_{k'}}\}$ with
      $J' = \{j'_1, \dots, j'_{k'}\}$ in $\overline{\mathcal{C}^{(l)}}$, compute $S_{[t,t+L]}(C^{J'})$;
    Predict $S_{[t,t+L]}(C^J)$ from $\left(S_{[t,t+L]}(C^J)\right)_{1 \leq j \leq \binom{d-k-1}{k+1}}$ with LASSO;
    if $R^2 > 0,67$ then
        Select all non-zero $\beta_j$ LASSO coefficients
    else
        Set $\beta_j = 0$ for all $j$.
    end
end
```

Algorithm 1: Sequential construction of the simplicial complex

Each non-zero β_j coefficient represents a k-simplex whose vertices are $\{X_i, X_{j_1}, X_{j_2}, \dots, X_{j_{k+1}}\}$. This simplex comes with a natural weight w. As this k-simplex can be produced multiple time (by predicting X_i with $\{X_{j_1}, \dots, X_{j_{k+1}}\}$ or X_{j_l} by $\{X_i, X_{j_1}, \dots, X_{j_{k+1}}\}_{k \neq l}$), this weight is produced as the sum of all the non-zero LASSO coefficients obtained.

By iterating the procedure for every possible simplices whose signature is a statistically interesting quantity to predict, and every dimension of simplex $k \leq K$ (for a chosen k), one can produce a simplicial complex among $\mathcal{G}$.

Remarks

Let us now address some technical question that arise from the proposed construction.

Time dependancy: The algorithm is applied to evolving time series for which the computation of the Signatures is updated incrementally and prediction is performed using these updated Signatures as time increases.

Orientation: This algorithm gives a natural orientation on every simplex, as $S(X_i, X_j) \neq S(X_j, X_i)$ which is often of great potential use for interpretability.

4 Numerical experiments and Applications

Multivariate times series are omnipresent and high order correlation occurred frequently in e.g. the domains of finance and neuroscience. We evaluated our method on two public data sets analysed by [19] : the fMRI resting-state data from the HCP https://www.humanconnectome.org/.

4.1 Practical choices and hyper parameters

We consider only simplices up to triangles $k \leq 2$ for now. Due to the concerns about complexity, the depth of signature is set to $depth = 2$ for the construction of both 1-simplex and 2-simplex. More precisely, we use time-augmented path to calculate 2-truncated signatures of 0-simplices, in order to construct 1-simplices. As a design choice, for 2-simplices the predictors (1-simplices composed of two times series) are not augmented when applying LASSO regression. Clearly, many different choices could be imposed on how we model the dependencies between subgroups of time series.

The regularisation term of LASSO is crucial for our method since it directly controls the sparsity in prediction using the linear models on signature features. Recall that the selected groups of time series will immediately be translated into new simplices in our sequentially growing simplicial complex.

In the present numerical experiments, we show that coarsely selected values for these hyperparameters already provide interesting results on a real dataset. In practice, $\lambda_{1-simplex}$ and $\lambda_{2-simplex}$ have been empirically tuned to the values $\lambda_{1-simplex} = 1000$ and $\lambda_{2-simplex} = 3$ for the fMRI dataset. More experiments based on extensive comparisons over a refined grid will be tested in an extended version of the present paper. The latest version of our implementation is available on our GitHub page :

https://github.com/ben2022lo/conf-complex-network

4.2 Modelling interactions in Functional MRI datasets

Functional connectivity is a neuroscience approach aimed at understanding the organization of the human brain based not solely on spatial proximity and structural factors, but rather on its functionality, i.e. its connectivity patterns between different brain regions and networks. For instance, even seemingly routine tasks such as paying attention during a lecture have been found to activate regions like the pulvinar (within the thalamus), the superior colliculus (in the midbrain), and the posterior parietal cortex [18]. In this perspective, and given that functional brain imaging data can be regarded as time series, the theory of Signatures could prove to be particularly useful.

In the absence of specific tasks or external stimulation (resting, meditating, sleeping, etc.), the brain enters what is known as resting-state. The Default Mode Network (DMN) becomes prominently active during this resting state. This neural network includes key regions such as the medial prefrontal cortex (mPFC), the posterior parietal lobe (PTL), the posterior cingulate cortex (PCC), and the precuneus [2].

We tested our method on resting-state fMRI(rs-fMRI) data [7] preprocessed by the same pipeline in [19]. The dataset contains 100 cortical (Schaefer100 [20]) and 19 subcortical ROIs (Regions of Interest). In order to evaluate the quality of identified interactions, we have selected a subset of 15 ROIs of which

---

[7] HCP, http://www.humanconnectome.org/

functional connectivities during resting-state are well known. We constructed simplicial complex on all 1200 timesteps, and analysed the top 10 1-simplices and 2-simplices that are the most persistent, i.e.that occurred on most time-steps. Besides, we observed that the life duration distribution of 1-simplices is centred and symmetric, whereas the distribution of 2-simplices is positively skewed.

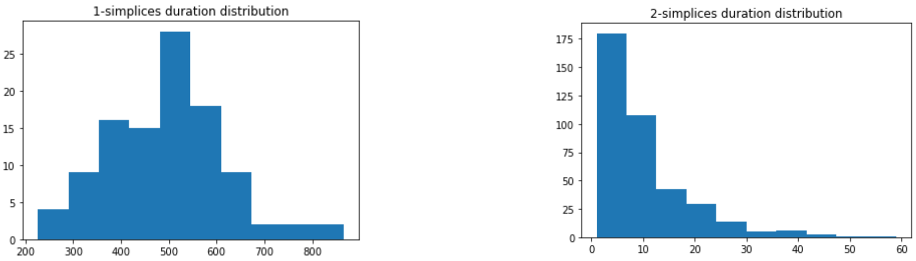


Fig. 1: Histograms of the observed duration for the discovered 1 and 2-simplices

The 1-simplex representing the interaction between *7Networks LH Vis 9* and *7Networks LH SomMot 4* occurred on most occasions (865 time steps). The most persistent 2-simplex (59 time steps) represents the interaction among *LH Cont Par 1*, *RH Default PFCdPFCm 2* and *RH Default pCunPCC 2*.

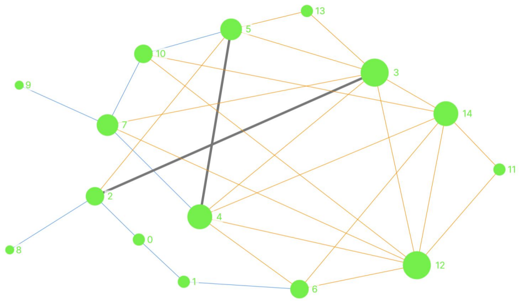


Fig. 2: Simplicial complex constructed with persistent simplicies. 1-simplices are blue, 2-simplicies are defined by their orange 1-simplex faces. The gray signifies the coincidence of a 1-simplex and one face of a 2-simplex. The selected ROI and numerated 0-simplices are matched by the following dictionary: 0 - *LH Vis 9*, 1 - *LH SomMot 4*, 2 - *LH DorsAttn Post 4*, 3 - *Cont Par 1*, 4 - *LH Cont pCun 1*, 5 - *LH Default pCunPCC 1*, 6 - *LH Default pCunPCC 2*, 7 - *RH Cont Par 1*, 8 - *RH Cont PFCl 1*, 9 - *RH Cont pCun 1*, 10 - *RH Default Par 1*, 11 - *RH Default PFCdPFCm 1*, 12 - *RH Default PFCdPFCm 2*, 13 - *RH Default PFCdPFCm 3*, 14 - *RH Default pCunPCC 2*.

Most of the persistent 1-simplices involve the prefrontal cortex, the parietal lobe and the precuneus, which is consistent with literature as all three regions are active during resting-state [2]. The most persistent interaction include subregions of the left hemisphere's visual and somatomotor networks. Component *LH SomMot 4* has previously been associated with components of the left (*LH Default PFC*) and right (*RH Default PFCv 2*) PFC in a study proposing an age prediction pipeline Ayu using rs-fMRI data [1]. In the same study, several visual areas (*RH Vis 1, 3* and *4*) are linked to the PFC areas during rs-fMRI (*LH Default PFC 1, 2,* and *3*), although they are located in the right hemisphere.

The top 10 interactions that occurred the most include components from the same three recurrent brain areas that are the PFC, parietal regions and the precuneus, with the latter taking par in all 10 of them. This aligns with previous work [2] as all three are indeed involved in the Default Mode Network which is active during rs-fMRI.

The most persistent simplicies and matching ROIs are represented in Figure 3. In particular, the interactions discovered using our approach show excellent coherence with well identified spatial activity zones.

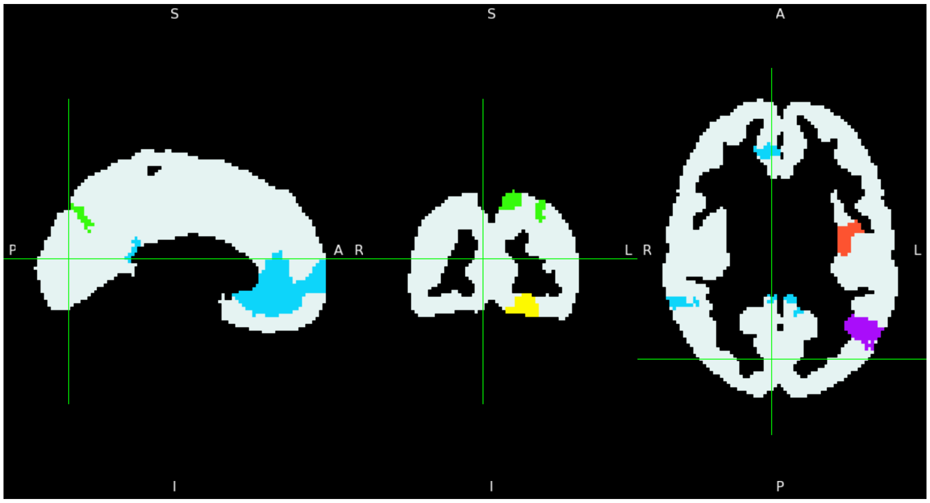


Fig. 3: Anatomical representation of the 10 most persistent 1-simplices and 2-simplices matched to their corresponding network in the 7-network parcellation by [23]. Only parcels that are part of the most persistent simplices are colored. Each color corresponds to a network as per the following color legend: yellow - *Visual*, red - *Somatomor*, purple - *Dorsal Attention*, green - *Control*, blue - *Default*.

5 Discussion and future work

The qualification of high-order interaction of time series is a relatively new research area. Previous work, such as [19], tried to estimate the higher order interactions in high dimensional signals. Our method, based on the theory of Signatures that captures higher dimensional interactions together with what can be encoded as a simplex orientations, is able to leverage much more refined information about the mutual behaviour of the observed phenomena.

From a theoretical point of view, although the orientation of the various simplices discovered in the sequential construction was not used proper, it could be fruitfully exploited in the future. Secondly, the problem of dimension consistency could be appropriately tackled using group LASSO types of techniques or SLOPE-based approaches.

To conclude, we mention that the method we just presented could also be handily put to work on nonstationary problems for e.g. change point detection such as in early detection of epidemics, using human in the loop validation steps. The simplicial complex could be sequentially updated as a function of time as well, leading to a dynamical topological structure whose characteristics and abrupt potential changes could help extract valuable information about the emergence of certain interesting phenomena.

References

1. Kushal Borkar, Anusha Chaturvedi, P K Vinod, and Raju Bapi. Ayu-characterization of healthy aging from neuroimaging data with deep learning and rsfmri. *Frontiers in Computational Neuroscience*, 16, 09 2022.
2. Samantha Broyd, Charmaine Demanuele, Stefan Debener, Suzannah Helps, Christopher James, and Edmund Sonuga-Barke. Default-mode brain dysfunction in mental disorders: A systematic review. *Neuroscience and biobehavioral reviews*, 33:279–96, 10 2008.
3. Frédéric Chazal and Bertrand Michel. An introduction to topological data analysis: fundamental and practical aspects for data scientists. *Frontiers in artificial intelligence*, 4:108, 2021.
4. Kuo-Tsai Chen. Integration of paths, geometric invariants and a generalized baker-hausdorff formula. *Annals of Mathematics*, 65(1):163–178, 1957.
5. Ilya Chevyrev and Andrey Kormilitzin. A primer on the signature method in machine learning. *arXiv preprint arXiv:1603.03788*, 2016.
6. Jean-Pierre Eckmann and Uni Genève. Martin hairer got the fields medal for his study of the kpz equation.
7. Adeline Fermanian. Embedding and learning with signatures. *Computational Statistics & Data Analysis*, 157:107148, 2021.
8. Adeline Fermanian, Pierre Marion, Jean-Philippe Vert, and Gérard Biau. Framing rnn as a kernel method: A neural ode approach. *Advances in Neural Information Processing Systems*, 34:3121–3134, 2021.
9. Peter K Friz and Martin Hairer. *A course on rough paths.* Springer, 2020.
10. Peter K Friz and Nicolas B Victoir. *Multidimensional stochastic processes as rough paths: theory and applications*, volume 120. Cambridge University Press, 2010.

11. Chad Giusti and Darrick Lee. Iterated integrals and population time series analysis. In *Topological Data Analysis: The Abel Symposium 2018*, pages 219–246. Springer, 2020.
12. Heejung Kim, Jarang Hahm, Hyekyoung Lee, Eunjoo Kang, Hyejin Kang, and Dong Soo Lee. Brain networks engaged in audiovisual integration during speech perception revealed by persistent homology-based network filtration. *Brain connectivity*, 5(4):245–258, 2015.
13. Andrey Kormilitzin, Nemanja Vaci, Qiang Liu, Hao Ni, Goran Nenadic, and Alejo Nevado-Holgado. An efficient representation of chronological events in medical texts. *arXiv preprint arXiv:2010.08433*, 2020.
14. Terry Lyons and Andrew D McLeod. Signature methods in machine learning. *arXiv preprint arXiv:2206.14674*, 2022.
15. Terry Lyons and Zhongmin Qian. *System control and rough paths*. Oxford University Press, 2002.
16. Nicolai Meinshausen and Peter Bühlmann. High-dimensional graphs and variable selection with the lasso. 2006.
17. Giovanni Petri, Paul Expert, Federico Turkheimer, Robin Carhart-Harris, David Nutt, Peter J Hellyer, and Francesco Vaccarino. Homological scaffolds of brain functional networks. *Journal of The Royal Society Interface*, 11(101):20140873, 2014.
18. Michael Posner and Steve Petersen. The attention system of the human brain. *Annual review of neuroscience*, 13:25–42, 02 1990.
19. Andrea Santoro, Federico Battiston, Giovanni Petri, and Enrico Amico. Higher-order organization of multivariate time series. *Nature Physics*, 19(2):221–229, 2023.
20. Alexander Schaefer, Ru Kong, Evan Gordon, Timothy Laumann, Xi-Nian Zuo, Avram Holmes, Simon Eickhoff, and B. Yeo. Local-global parcellation of the human cerebral cortex from intrinsic functional connectivity mri. 07 2017.
21. Ann E Sizemore, Chad Giusti, Ari Kahn, Jean M Vettel, Richard F Betzel, and Danielle S Bassett. Cliques and cavities in the human connectome. *Journal of computational neuroscience*, 44:115–145, 2018.
22. Bernadette J Stolz, Heather A Harrington, and Mason A Porter. Persistent homology of time-dependent functional networks constructed from coupled time series. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 27(4), 2017.
23. B.T. Thomas Yeo, Fenna Krienen, Jorge Sepulcre, Mert Sabuncu, Danial Lashkari, Marisa Hollinshead, Joshua Roffman, Jordan Smoller, Lilla Zollei, Jonathan Polimeni, Bruce Fischl, Hesheng Liu, and Randy Buckner. The organization of the human cerebral cortex estimated by functional correlation. *Journal of neurophysiology*, 106:1125–65, 06 2011.