Towards a Quantitative Theory of Digraph-Based Complexes and its Applications in Brain Network Analysis

Heitor Baldo

Adviser: Prof. Koichi Sameshima
Co-Adviser: Prof. André Fujita

PhD Thesis Defense

Interunit Bioinformatics Graduate Program
University of São Paulo

July 01, 2024

Outline

Overview - Motivation

- Brain network analysis $\rightarrow$ **Organization and dynamics of brain activity** $\rightarrow$ Study of brain disorders such as **epilepsy**.
- **Information partial directed coherence** (iPDC) (based on Granger causality; related to mutual information) $\rightarrow$ Brain connectivity digraphs $\rightarrow$ Reflects the **causal relationships** between brain structures.
- Usual graph measures cannot always provide relevant insights into the network topology $\rightarrow$ The analysis of the **clique topology** allows us to **go beyond usual graph theoretical analysis** $\rightarrow$ Reveals **new properties** associated with the network.

Overview - Goals

The goal of the thesis is twofold:

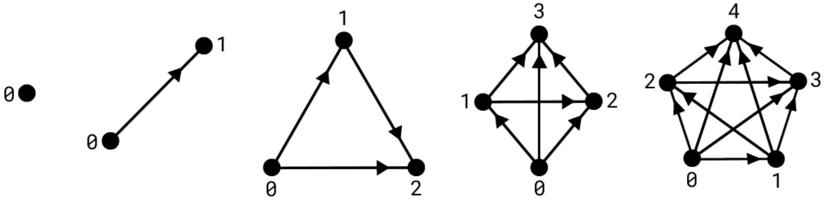
1. To develop rigorously a **new quantitative theory for digraph-based complexes**, with special emphasis on directed higher-order connectivity (higher-order relations) between directed cliques (including the software).

2. To apply the methods of the new theory to **epileptic brain networks** obtained through iPDC to quantitatively investigate their higher-order topologies and **search for new biomarkers based on their directed higher-order structures and connectivities**, thus pointing out potential applications of the theory in network neuroscience.

Part I: Towards a Quantitative Theory of Digraph-Based Complexes

Directed Cliques

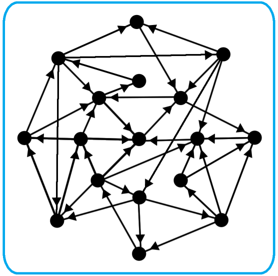
A **directed** $(k+1)$-**clique** (Reimann, 2017) is a digraph $G = (V, E)$, $V = \{v_0, ..., v_k\}$, whose underlying undirected graph is a $(k+1)$-clique and for each $0 \leq i < j \leq k$, $v_i \in V$, and there is a directed edge from v_i to v_j, i.e. $(v_i, v_j) \in E$, for all $i < j$.



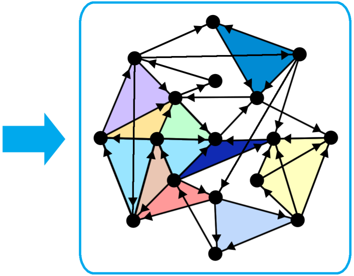
Notice that, by definition, every directed $(k+1)$-clique has a source (v_0) and a sink (v_k) and is a directed acyclic graph (DAG).

Directed Flag Complexes

Directed Flag Complexes = Directed Clique Complexes



Original network



Directed flag complex

The directed k-simplices of a directed flag complex span the corresponding directed $(k+1)$-cliques in the corresponding digraph.

Directed Higher-Order Connectivity between Directed Simplices

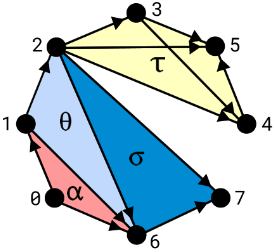
$\mathrm{dFl}(G) \to$ the directed flag complex of a digraph G, and $\sigma^{(n)} = [v_0, ..., v_n] \in \mathrm{dFl}(G)$.

- The i-th face map $\hat{d}_i$ is defined by

$$\hat{d}_i(\sigma^{(n)}) = \begin{cases} (v_0, ..., \hat{v}_i, ..., v_n), & \text{if } i \leq n, \\ (v_0, ..., v_{n-1}, \hat{v}_n), & \text{if } i > n. \end{cases} \tag{1}$$

- Let $\sigma = \sigma^{(n)}, \tau = \tau^{(m)} \in \mathrm{dFl}(G)$, with $n, m \geq q$. σ is lower $(q, \hat{d}_i, \hat{d}_j)$-near to τ if either of the following conditions is true:
  1. $\sigma \subseteq \tau$;
  2. $\hat{d}_i(\sigma) \supseteq \alpha^{(q)} \subseteq \hat{d}_j(\tau)$, for some q-simplex $\alpha^{(q)}$ (i.e. if they share a q-face).

Directed Higher-Order Connectivity - New Nomenclatures and Notations



$\sigma \sim^{\pm}_{L_0} \tau$, since $\hat{d}_1(\sigma) \supseteq [2] \subseteq \hat{d}_2(\tau)$ and $\hat{d}_2(\sigma) \supseteq [2] \subseteq \hat{d}_1(\tau)$

$\theta \sim^{+}_{L_1} \sigma$, since $\hat{d}_0(\theta) \supseteq [2,6] \subseteq \hat{d}_2(\sigma)$

$\alpha \sim^{+}_{L_1} \theta$, since $\hat{d}_0(\alpha) \supseteq [1,6] \subseteq \hat{d}_1(\theta)$

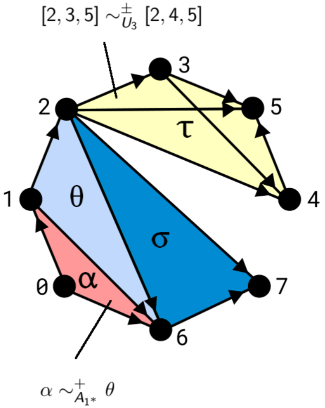
$\mathrm{dFl}(G) \to$ the directed flag complex of a given simple digraph G **without double-edges**.

For $\sigma, \tau \in \mathrm{dFl}(G)$, we have the following definitions:

- σ is said to be lower $(-)$-q-near to $\tau^{(m)}$ if they are lower $(q, d_i(\sigma), d_j(\tau))$-near with $i \geq j$. Notation: $\sigma \sim^{-}_{L_q} \tau$ $(\sigma \leftarrow \tau)$.
- σ is said to be lower $(+)$-q-near to τ if they are lower $(q, d_i(\sigma), d_j(\tau))$-near with $i \leq j$. Notation: $\sigma \sim^{+}_{L_q} \tau$ $(\sigma \rightarrow \tau)$.
- σ is said to be lower $(\pm)$-q-near to τ if $\sigma^{(n)} \sim^{-}_{L_q} \tau^{(m)}$ and $\sigma \sim^{+}_{L_q} \tau^{(m)}$. Notation: $\sigma \sim^{\pm}_{L_q} \tau = \tau \sim^{\pm}_{L_q} \sigma$ $(\sigma \leftrightarrow \tau)$.
- We define strictly lower $(\bullet)$-q-adjacency, where $\bullet \in \{-, +, \pm\}$, as:

$$\sigma^{(n)} \sim^{\bullet}_{L_{q^*}} \tau^{(m)} \iff \sigma^{(n)} \sim^{\bullet}_{L_q} \tau^{(m)} \text{ and } \sigma^{(n)} \not\sim^{\star}_{L_{q+1}} \tau^{(m)}, \ \forall \star .$$

Directed Higher-Order Connectivity - New Definitions I



Upper q-Adjacency

Definition. Let $\sigma^{(n)}, \tau^{(m)} \in \mathrm{dFl}(G)$. For $n, m \leq p', p'' \leq p \leq \dim \mathrm{dFl}(G)$, $\sigma^{(n)}$ is said to be upper $(p, \hat{d}_i, \hat{d}_j)$-near to $\tau^{(m)}$ the following condition is true:

$$\sigma^{(n)} = \hat{d}_i(\Theta^{(p')}) \subseteq \Theta^{(p)} \supseteq \hat{d}_j(\Theta^{(p'')}) = \tau^{(m)}, \text{ for some } \Theta^{(p')}, \Theta^{(p'')} \subseteq \Theta^{(p)} \in \mathrm{dFl}(G).$$

- $\sigma^{(n)}$ is $(\bullet)$-p-upper adjacent ($\bullet \in \{-,+,\pm\}$) to $\tau^{(m)}$ if they are upper (p, d_i, d_j)-near with $i \geq j$, if $\bullet = -$, $i \leq j$, if $\bullet = +$, and both, if $\bullet = \pm$; and we denote
$$\sigma^{(n)} \sim^{\bullet}_{U_p} \tau^{(m)}.$$

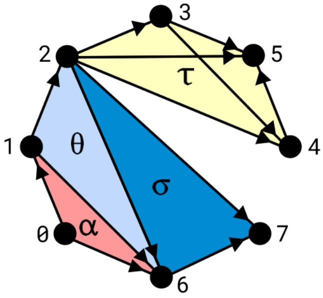
q-Adjacency

$$\sigma^{(n)} \sim^{\bullet}_{A_q} \tau^{(m)} \iff \sigma^{(n)} \sim^{\bullet}_{L_{q^*}} \tau^{(m)} \text{ and } \sigma^{(n)} \not\sim^{\star}_{U_p} \tau^{(m)}, \ \forall \star .$$

Maximal q-Adjacency

$$\sigma^{(n)} \sim^{\bullet}_{A_{q^*}} \tau^{(m)} \iff \sigma^{(n)} \sim^{\bullet}_{A_q} \tau^{(m)} \text{ and } \sigma^{(n)} \not\subset \sigma^{(r)}, \ \forall \sigma^{(r)} : \sigma^{(r)} \sim^{\star}_{A_q} \tau^{(m)}, \ \forall \star .$$

Directed Higher-Order Connectivity - New Definitions II



$\alpha \overset{+}{\sim}_{A_0^*} \tau$, since $\alpha \overset{+}{\sim}_{A_{1^*}} \theta \overset{+}{\sim}_{A_0^*} \tau$

Maximal q-Connectivity

$$\sigma^{(n)} \overset{\bullet}{\sim}_{A_{q_0^*}} \alpha_1^{(n_1)} \overset{\bullet}{\sim}_{A_{q_1^*}} \cdots \overset{\bullet}{\sim}_{A_{q_{l-1}^*}} \alpha_l^{(n_l)} \overset{\bullet}{\sim}_{A_{q_l^*}} \tau^{(m)}$$

with $0 \leq q \leq \min(n, m, n_1, \dots, n_l)$

Notation: $\sigma^{(n)} \overset{\bullet}{\sim}_{A_{q^*}} \tau^{(m)}$

Directed Simplicial q-Walk

There is a **directed simplicial q-walk** from $\sigma^{(n)}$ to $\tau^{(m)}$ if

$$\sigma^{(n)} \overset{+}{\sim}_{A_{q^*}} \tau^{(m)}$$

The directed q-distance is the length of the shortest directed q-walk between them.

Notation: $\vec{d}_q(\sigma, \tau)$,

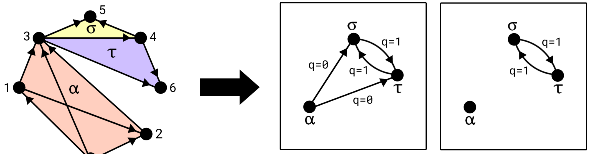
Maximal q-Digraphs

Definition. The (maximal) q-digraph of $\mathrm{dFl}(G)$ is the digraph $\mathcal{G}_q = (\mathcal{V}_q, \mathcal{E}_q)$ whose vertices are the maximal directed simplices of $\mathrm{dFl}(G)$ and the arc $\sigma \to \tau$ exists iff $\sigma \sim^{+}_{A_{q^*}} \tau$.

Definition. The maximal q-adjacency matrix of $\mathcal{G}_q$, denoted by $\mathcal{H}_q = \mathcal{H}_q(\mathcal{G}_q)$, is a real square matrix whose entries are given by

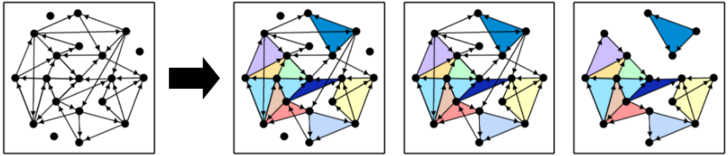
$$\left(\mathcal{H}_q\right)_{ij} = \begin{cases} 1, & \text{if } \sigma_i \sim^{+}_{A_{q^*}} \sigma_j, \\ 0, & \text{if } i = j \text{ or } \sigma_i \not\sim^{+}_{A_{q^*}} \sigma_j. \end{cases} \tag{2}$$

Example:

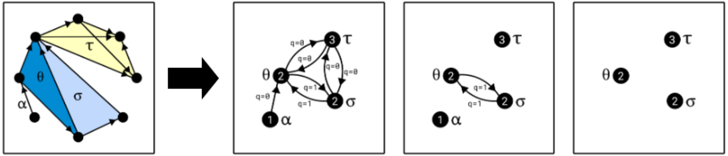


0-digraph　　1-digraph

Examples - Maximal Directed Simplices and Maximal q-Digraphs

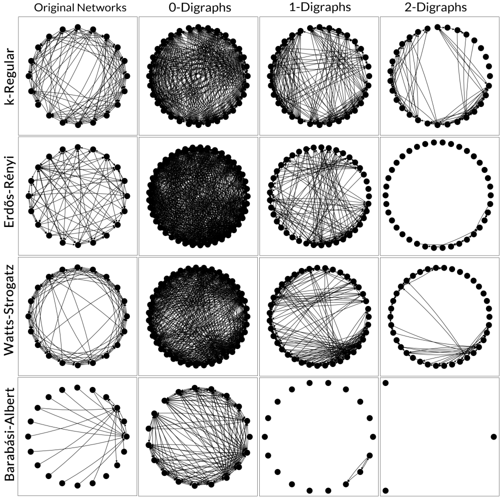


(a) Digraph G. (b) $\mathrm{dFl}_0^*(G)$. (c) $\mathrm{dFl}_1^*(G)$. (d) $\mathrm{dFl}_2^*(G)$.



(a) $\mathrm{dFl}(G)$. (b) $\mathcal{G}_0$. (c) $\mathcal{G}_1$. (d) $\mathcal{G}_2$.

Examples with Random Digraphs



The size (N) of the networks may change from one level q to another

Simplicial Characterization Measures I

Graph characterization is concerned with describing some property of the network through a local or global numerical graph invariant (properties that are invariant under graph isomorphism).

$$\mathcal{G}_q = (\mathcal{V}_q, \mathcal{E}_q) \rightarrow \text{ Simplicial Characterization Measures } \rightarrow \begin{cases} \text{Distance-Based Measures} \\ \text{Centrality Measures} \\ \text{Segregation Measures} \\ \text{Entropy Measures} \\ \text{Spectrum-Related Measures} \end{cases}$$

Simplicial Characterization Measures II

$$\vec{E}^{q}_{glob}(\mathcal{G}_q) = \frac{1}{|\mathcal{V}_q|} \sum_{\sigma \in \mathcal{V}_q} \frac{\sum_{\tau \in \mathcal{V}_q, \tau \neq \sigma} \vec{d}_q^{-1}(\sigma, \tau)}{|\mathcal{V}_q| - 1}$$

Directed Global q-Efficiency (global)

Interpretation: It captures how efficiently information is propagated through the network.

$$C^{-}_{\deg_q}(\sigma) = \frac{\deg^{-}_{A_{q^*}}(\sigma)}{|\mathcal{V}_q| - 1}$$

In-q-Degree Centrality (local)

Interpretation: It quantifies the importance or centrality of a node based on the number of incoming arcs.

$$C^{+}_{\deg_q}(\sigma) = \frac{\deg^{+}_{A_{q^*}}(\sigma)}{|\mathcal{V}_q| - 1}$$

Out-Degree Centrality (local)

Interpretation: It quantifies the importance or centrality of a node based on the number of outgoing arcs.

$$\vec{HC}_q(\sigma) = \sum_{\substack{\tau \in \mathcal{V}_q \\ \tau \neq \sigma}} \frac{1}{\vec{d}_q(\sigma, \tau)}$$

Directed q-Harmonic Centrality (local)

Interpretation: It identifies the nodes that can spread/ receive information in an efficient way.

$$GRC_q(\mathcal{G}_q) = \frac{\sum_{\sigma \in \mathcal{V}_q} [C^{max}_{R,q} - C_{R,q}(\sigma)]}{|\mathcal{V}_q| - 1}$$

Global q-Reaching Centrality (global)

Interpretation: It captures how the nodes influence the flow of information through the digraph.

Simplicial Characterization Measures III

$$\vec{\vec{C}}_q(\mathcal{G}_q) = \frac{1}{|\mathcal{V}_q|} \sum_{\sigma \in \mathcal{V}_q} \frac{\vec{T}_q(\sigma)}{\deg^{tot}_{A_{q^*}}(\sigma)(\deg^{tot}_{A_{q^*}}(\sigma) - 1) - 2\deg^{\pm}_{A_{q^*}}(\sigma)}$$

Average q-Clustering Coefficient
(global)

Interpretation: It quantifies the tendency of nodes to form clusters in the network.

$$H_q^-(\mathcal{G}_q) = -\sum_{k=1}^{n} p_q^-(k) \log_2 p_q^-(k)$$

In-q-Degree Distribution Entropy
(global)

Interpretation: It quantifies the "degree of disorder" or "degree of randomness" (in relation to the inner flux) of a network.

$$H_q^+(\mathcal{G}_q) = -\sum_{k=1}^{n} p_q^+(k) \log_2 p_q^+(k)$$

Out-q-Degree Distribution Entropy
(global)

Interpretation: It quantifies the "degree of disorder" or "degree of randomness" (in relation to the outer flux) of a network.

$$\varepsilon_q(\mathcal{G}_q) = ||(\mathcal{H}_q \mathcal{H}_q^T)^{1/2}||_* = Tr((\mathcal{H}_q \mathcal{H}_q^T)^{1/2})$$

q-Energy
(global)

Interpretation: It can be interpreted as a measure of connectivity of a network.

Simplicial Homology and Persistent Homology

p-th boundary map: $\partial_p : C_p(\mathcal{X};\mathbb{K}) \to C_{p-1}(\mathcal{X};\mathbb{K})$

$$\partial_i(\sigma) = \sum_{i=0}^{n}(-1)^i \hat{d}_i(\sigma) = \sum_{i=0}^{n}(-1)^i[v_0,\dots,\hat{v}_i,\dots,v_n].$$

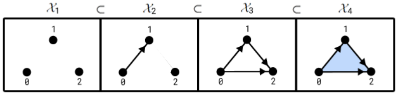
p-th homology:

$$\boxed{H_p(\mathcal{X}) = \ker \partial\big|_{C_p} / \operatorname{Im} \partial\big|_{C_{p+1}}}$$

p-th Betti number:

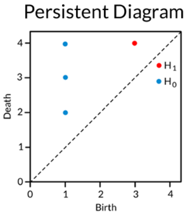
$$\boxed{\beta_p(\mathcal{X}) = \dim H_p(\mathcal{X})}$$

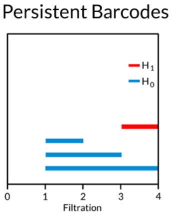
Filtration:

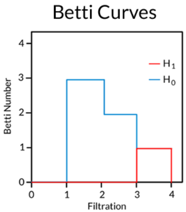


n-th persistent homology:

$$\boxed{H_n^{i,j} = \operatorname{Im} f_n^{i,j} = \ker \partial_n(\mathcal{X}_i)/(\operatorname{Im} \partial_{n+1}(\mathcal{X}_j) \cap \ker \partial_n(\mathcal{X}_i))}$$







Simplicial Similarity Comparison Distances I

Comparative graph analysis is concerned with methods to compare the structural similarity or structural distance between two or more graphs.

Consider two persitent diagrams, P and Q, and a bijection $\eta : P \to Q$.

$$d_{W_\infty}(P,Q) = \inf_{\eta:P\to Q} \sup_{x\in P} ||x-\eta(x)||_\infty$$

Bottleneck Distance

$$d_{W_p}(P,Q) = \inf_{\eta:P\to Q} \left(\sum_{x\in P} ||x-\eta(x)||_\infty^p \right)^{1/p}$$

p-Wasserstein Distance

$$d_{\mathcal{B}}(\mathcal{B}_P,\mathcal{B}_Q) = \left(\int_{\mathbb{R}} |\mathcal{B}_P(x)-\mathcal{B}_Q(x)|^p dx \right)^{1/p}$$

p-Betti Distance

We can compare the topology of two different directed flag complexes (and therefore the topology of their underlying digraphs) by comparing the similarity of their persistence diagrams and/or their Betti curves.

Simplicial Similarity Comparison Distances II

Definition. Considering a directed flag complex $\mathcal{X}$, we have the following definitions:

- First topological structure vector: $Str_1(\mathcal{X}) = (c_0, ..., c_n)$, where c_k is the number of $(k+1)$-cliques.
- Fourth topological structure vector: $Str_4(\mathcal{X}) = (\beta_0, ..., \beta_n)$, where β_k is the k-th Betti number.
- Fifth topological structure vector: $Str_5(\mathcal{X}) = (l_0, ..., l_n)$, where l_i is the length of the k-th barcode of the persistence barcode.

Definition. Consider two directed flag complexes $\mathcal{X}_1$ and $\mathcal{X}_2$, with $\mathcal{X}_1 \neq \emptyset$ or $\mathcal{X}_2 \neq \emptyset$. The n-th topological structure distance between $\mathcal{X}_1$ and $\mathcal{X}_2$ is:

$$\boxed{T^n_{td}(\mathcal{X}_1, \mathcal{X}_2) = \frac{||Str_n(\mathcal{X}_1) - Str^*_n(\mathcal{X}_2)||_2}{||Str_n(\mathcal{X}_1)||_2 + ||Str^*_n(\mathcal{X}_2)||_2}.}$$

Simplicial Similarity Comparison Distances III

Definition. The histogram cosine kernel (HCK) between $\mathcal{X}_1$ and $\mathcal{X}_2$, is defined as

$$\boxed{K_{HCK}(\mathcal{X}_1, \mathcal{X}_2) = \frac{\langle Str_1(\mathcal{X}_1), Str_1^*(\mathcal{X}_2)\rangle}{\sqrt{\langle Str_1(\mathcal{X}_1), Str_1(\mathcal{X}_1)\rangle}\sqrt{\langle Str_1^*(\mathcal{X}_2), Str_1^*(\mathcal{X}_2)\rangle}}.}$$

Definition. The Jaccard kernel between $\mathcal{X}_1$ and $\mathcal{X}_2$, is defined as 1 minus the ratio between the cardinality of their intersection and the cardinality of their union, i.e.

$$\boxed{K_J(\mathcal{X}_1, \mathcal{X}_2) = 1 - \frac{|\mathcal{X}_1 \cap \mathcal{X}_2|}{|\mathcal{X}_1 \cup \mathcal{X}_2|}.}$$

Part II: Brain Connectivity Networks and a Quantitative Graph/Simplicial Analysis of Epileptic Brain Networks

Brain Connectivity and Brain Connectivity Estimation

Structural (anatomical) connectivity refers to a set of anatomical connections that link neural elements.

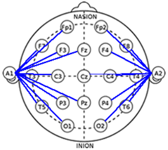
Functional connectivity refers to the temporal interdependence of activation patterns of anatomically separate brain areas.

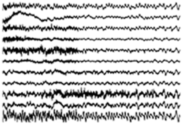
Effective connectivity reflects the causal relationships between activated brain regions by characterizing the influence that one brain structure has on another.

iPDC (information partial directed coherence)

- Uses Multivariate AutoRegressive (MAR) model;
- Based on Granger causality;
- Analysis in frequency-domain;
- Related to mutual information.

Electroencephalography (EEG)





MVAR $\left\{ X(n) = \sum_{r=1} A_r X(n-r) + W(n) \right.$

frequency domain

$$\left\{ \begin{array}{l} \bar{A}(\omega) = I - A(\omega) = [\bar{a}_1(\omega)\bar{a}_2(\omega)\ldots\bar{a}_m(\omega)] \\ \bar{A}_{ij}(\omega) = \begin{cases} 1 - \sum_{r=1}^{+\infty} a_{ij}(r)e^{-i\omega r}, & \text{if i = j,} \\ -\sum_{r=1}^{+\infty} a_{ij}(r)e^{-i\omega r}, & \text{otherwise.} \end{cases} \end{array} \right.$$

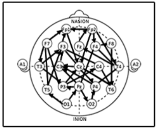
iPDC

$$\boxed{\iota\pi_{ij}(\omega) := \frac{\sigma_i^{-1/2}\bar{A}_{ij}(\omega)}{\sqrt{\bar{a}_j^H(\omega)\Sigma_W^{-1}\bar{a}_j(\omega)}}}$$

$$\boxed{H_0 : \iota\pi_{ij}(\omega) = 0}$$

Its rejection implies the existence of a directed connection from xj(n) to xi(n).

Brain connectivity network



iPDC produces a weighted digraph. The statistically significant connections (arcs) are estimated asymptotically.

Epilepsy as a Disorder of Brain Connectivity

- Epilepsy is notoriously a brain connectivity network disorder, characterized by a direct relation between **abnormal network dynamics and clinical symptoms**.
- Fact constantly observed in the literature: patients diagnosed with epilepsy present **alterations in brain connectivity networks** in the **ictal phase (seizure)** compared to the **interictal phase**.

Example:

Motivation: Alterations in the brain connectivity networks can be identified through measures associated with the network topology.

(Hu et al., 2019) estimated G-connectivity networks through PDC from EEG data obtained from 10 patients with focal epilepsy.

Results:
- The clustering coefficients were significantly increased in the frontal, parietal, and temporal lobes during the ictal phase compared compared with the interictal period.

Analysis of Epileptic Brain Networks - Research Questions

We want to analyze the dynamics of epileptic seizures through quantitative analysis of brain (higher-order) network topology:

- How do the topological and functional properties of G-connectivity networks and their respective q-digraphs **change during the seizure** in each hemisphere and in each frequency band, both at the node level and at the various **clique topology** levels?
- Does the analysis of **higher-order structures** of brain connectivity networks provide novel and better **biomarkers** for seizure dynamics and also for the laterality of the seizure focus than the usual theoretical graph analyses?

Materials: EEG Data (Siena Scalp EEG Database)

EEG data were obtained from the Siena Scalp EEG Database (Detti et al., 2020). We chose eight patients with left temporal lobe epilepsy (TLE) based on the quality of the signal.

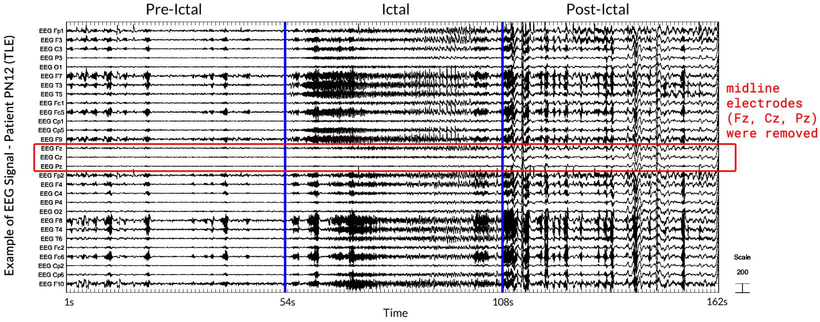
Public available at `https://physionet.org/content/siena-scalp-eeg/1.0.0/`.

Table: Patient information.

Pat. id	Age	Gender	Localization	Lateralization	Time (min)
PN01	46	Male	Temporal Lobe	Left	809
PN06	36	Male	Temporal Lobe	Left	722
PN07	20	Female	Temporal Lobe	Left	523
PN09	27	Female	Temporal Lobe	Left	410
PN12	71	Male	Temporal Lobe	Left	246
PN13	34	Female	Temporal Lobe	Left	519
PN14	49	Male	Temporal Lobe	Left	1408
PN16	41	Female	Temporal Lobe	Left	303

Methodology - Data Preprocessing

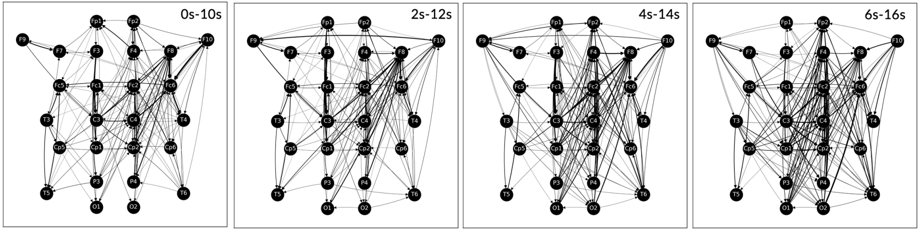
- **1)** Downsampling: from 512 Hz to 256 Hz. **2)** 1 Hz high-pass filter; notch filter at 50 Hz. **3)** The midline electrodes (Fz, Cz, and Pz) were removed (final number of channels: 26). **4)** An ICA was performed to remove artifactual components.
- Next, we identified the **pre-ictal, ictal, and post-ictal phases**, and then divided the signals from each phase into 30s epochs: 30s immediately before the seizures, 30s starting on the seizures onset, and 30s immediately after the seizures.



Methodology - Dynamic iPDC Networks Estimated from the EEG Data

- iPDC $\rightarrow$ **delta** [1 - 4 Hz), **theta** [4 - 8 Hz), and **alpha** [8 - 13 Hz);
- iPDC $\rightarrow$ **Sliding window technique** for each epoch, sliding the window in 10s cuts; overlap of 8s (11 digraphs on each seizure phase).
- VAR autoregression coefficients: estimated through the **Nuttall-Strand's method**. Order for the VAR models: **fixed order $\boldsymbol{p = 12}$**.
- Only the statistically significant connections were considered; estimated asymptotically with a **significance level of 0.1%**. Arc weights and double-edges were removed.

iPDC networks - Patient PN01 - Delta band [1 Hz, 4Hz) (seizure onset)



Methodology - Simplicial Characterization Measures

Simplicial Characterization Measures:

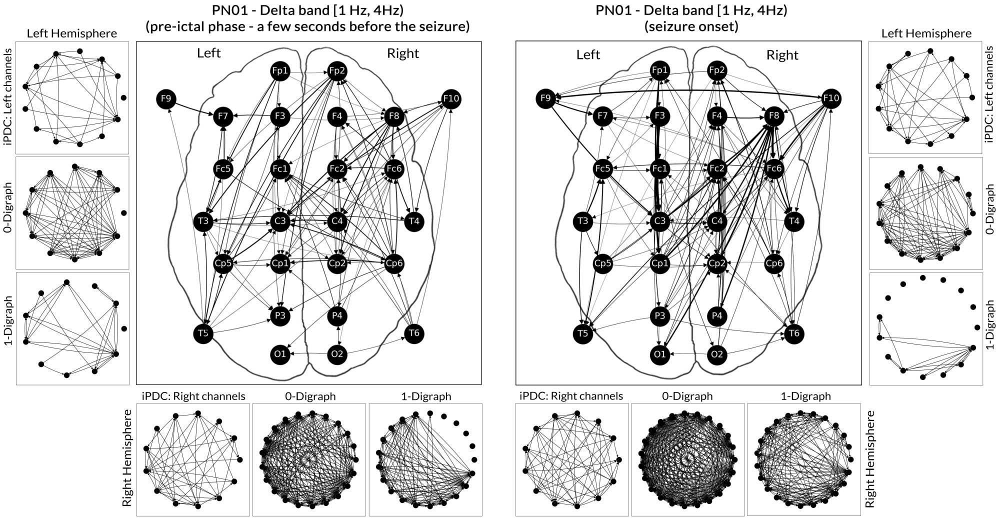
- In-q-degree centrality
- Out-q-degree centrality
- q-Harmonic centrality
- Global q-efficiency
- Global q-reaching centrality
- Average q-clustering coefficient
- In-q-degree distribution entropy
- Out-q-degree distribution entropy
- q-Energy

- q-Digraphs, for q = 0, 1, 2, 3, 4.
- Right and left hemispheres.
- Each seizure phase (pre-ictal, ictal, post-ictal).
- Each frequency band (delta, theta, alpha).
- We used a fixed N.

Statistical Analysis: pre-ictal vs ictal / ictal vs post-ictal

For each simplicial measure:
- Wilcoxon paired test, at a significance level $\alpha = 0.05$ in each frequency band (delta, theta, alpha) and at each level $q = -1, 0, 1, 2, 3, 4$.
- Right and left hemispheres.

Methodology - q-Digraphs - Pre-Ictal Phase vs Ictal Phase (Delta band)



Methodology - Simplicial Similarity Comparison Distances

Simplicial Similarity Comparison Distances:

- Bottleneck Distance
- p-Wasserstein Distance
- p-Betti Distance
- 1st Topological Structure Distance
- 4th Topological Structure Distance
- 5th Topological Structure Distance
- Histogram Cosine Kernel
- Jaccard Kernel

Let d be any of the eight chosen simplicial distances.

Seizure phases:

- $d(G^L_{ic}, G^L_{pre})$ vs $d(G^L_{pre}, G^L_{pre})$
- $d(G^L_{ic}, G^L_{pos})$ vs $d(G^L_{pos}, G^L_{pos})$
- $d(G^R_{ic}, G^R_{pre})$ vs $d(G^R_{pre}, G^R_{pre})$
- $d(G^R_{ic}, G^R_{pos})$ vs $d(G^R_{pos}, G^R_{pos})$

Laterality (ictal phase):

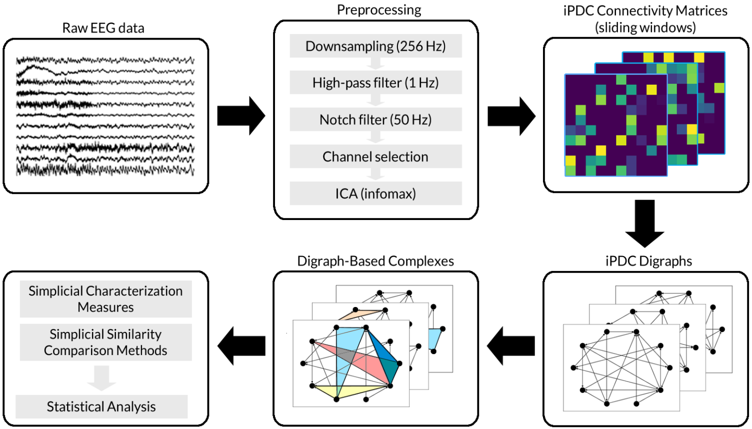
$d(G^L_{ic}, G^R_{ic})$ vs $d(G^R_{ic}, G^R_{ic})$

Statistical Analysis: Comparison between the means of the distributions

For each simplicial distance:
- A Wilcoxon paired test, at a significance level $\alpha = 0.05$, was performed to verify the statistical differences between the means of the previous distribution pairs, for each frequency band.

Analysis Pipeline

The raw EEG data was pre-processed in `EEGLAB` and the iPDC networks were obtained using the MATLAB package `asympPDC` (version 3.0).

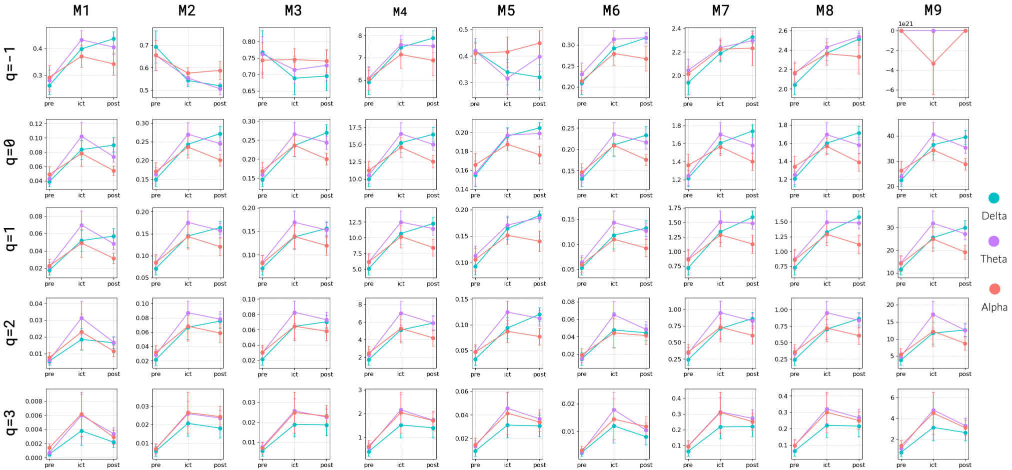


Statistical Analysis Results - Simplicial Characterization Measures - Left Hemisphere

W-statistics of the levels q and frequencies at which the measures showed a statistically significant increase ($p < 0.05$) in their magnitude in the **ictal phase** compared with the **pre-ictal phase**:

		q (Delta)					q (Theta)				
id	**Measure**	**-1**	**0**	**1**	**2**	**3**	**-1**	**0**	**1**	**2**	**3**
M1	Global q-efficiency	3	2	2	2	0	-	-	-	2	3
M2	In-q-degree centrality	-	2	3	1	0	-	-	-	2	3
M3	Out-q-degree centrality	-	2	2	1	0	-	-	-	2	3
M4	Harmonic q-closeness centrality	-	2	2	1	0	-	-	-	2	3
M5	Global q-reaching centrality	-	2	2	1	0	-	-	-	3	-
M6	Average q-clustering coefficient	2	2	1	1	-	-	-	2	1	1
M7	In-q-degree distribution entropy	-	-	2	1	0	-	-	-	3	-
M8	Out-q-degree distribution entropy	-	-	2	1	0	-	-	-	3	-
M9	q-Energy	3	2	2	1	0	-	-	-	2	3

Means and Standard Deviations - Simplicial Characterization Measures - Left Hemisphere

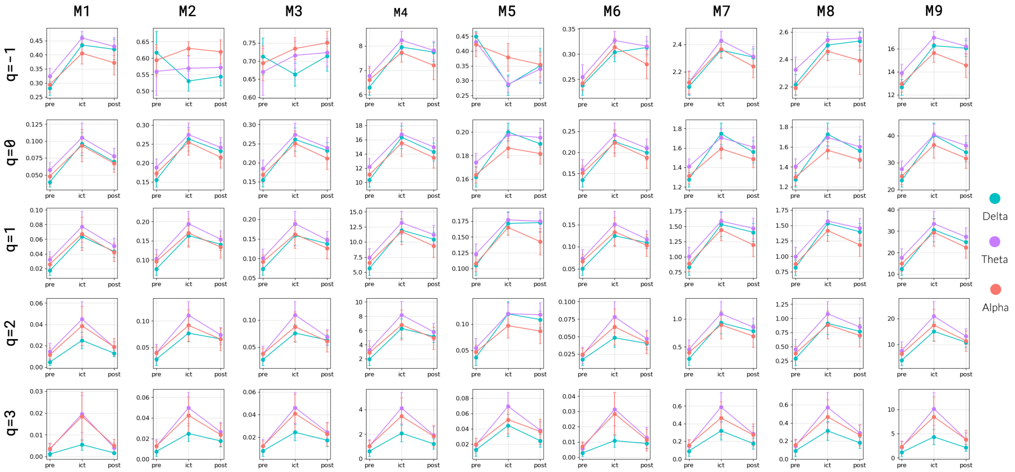


Statistical Analysis Results - Simplicial Characterization Measures - Right Hemisphere

W-statistics of the levels q and frequencies at which the measures showed a statistically significant increase ($p < 0.05$) in their magnitude in the **ictal phase** compared with the **pre-ictal phase**:

		q (Delta)					q (Theta)				
id	**Measure**	**-1**	**0**	**1**	**2**	**3**	**-1**	**0**	**1**	**2**	**3**
M1	Global q-efficiency	**2**	3	3	0	-	**0**	-	-	-	-
M2	In-q-degree centrality	-	-	3	0	-	-	-	3	1	1
M3	Out-q-degree centrality	-	-	0	0	-	-	-	3	0	1
M4	Harmonic q-closeness centrality	**1**	2	-	0	-	**3**	-	-	3	-
M5	Global q-reaching centrality	**1**	1	-	0	2	**3**	1	0	1	-
M6	Average q-clustering coefficient	-	-	0	3	-	**0**	-	1	1	1
M7	In-q-degree distribution entropy	**2**	3	-	0	2	**0**	-	3	3	1
M8	Out-q-degree distribution entropy	**3**	3	-	0	2	**1**	-	2	2	1
M9	q-Energy	**2**	2	-	0	3	**0**	-	-	-	1

Means and Standard Deviations - Simplicial Characterization Measures - Right Hemisphere



Discussion of the Results - Simpicial Characterization Measures

- All simpicial characterization measures showed statistically significant increases in their magnitudes from the **pre-ictal phase to the ictal phase**, for various levels q, for both hemispheres, especially in the δ and θ bands, suggesting an increase in **clustering, efficiency, and connectivity**, and a shift towards a "more random" organization (both in relation to inner and outer flow) in higher-order networks.
- No statistically significant changes were observed from the **ictal phase to the post-ictal phase**.
- This may suggest that several topological and functional aspects of the brain networks change from the pre-ictal phase to the ictal phase, **at various higher-order levels of topological organization** (clique organization), especially in the δ and θ bands.
- Most of the usual graph measures did not detect significant differences in the left hemisphere, which may suggest that changes in the network topology at the node level ($q = -1$) do not undergo as many changes as in the higher-order network topology.

Statistical Analysis Results - Simplicial Similarity Comparison Methods

- No statistically significant differences ($p < 0.05$) between the means of the distributions $d(G^L_{ic}, G^R_{ic})$ and $d(G^R_{ic}, G^R_{ic})$ were observed.
- W-statistics for simplicial distances showing significant differences ($p < 0.05$) between the distributions: $\boldsymbol{w_1}$ **:** $d(G^L_{ic}, G^L_{pre})$ vs $d(G^L_{pre}, G^L_{pre})$; $\boldsymbol{w_2}$ **:** $d(G^L_{ic}, G^L_{pos})$ vs $d(G^L_{pos}, G^L_{pos})$; $\boldsymbol{w_3}$ **:** $d(G^R_{ic}, G^R_{pre})$ vs $d(G^R_{pre}, G^R_{pre})$; $\boldsymbol{w_4}$ **:** $d(G^R_{ic}, G^R_{pos})$ vs $d(G^R_{pos}, G^R_{pos})$:

Distance	**Delta**				**Theta**				**Alpha**			
	w_1	w_2	w_3	w_4	w_1	w_2	w_3	w_4	w_1	w_2	w_3	w_4
Bottleneck Distance	-	-	-	-	-	-	3	-	-	-	-	-
p-Wasserstein Distance	-	-	2	-	0	1	0	-	-	-	-	2
Betti Distance	-	-	-	-	2	-	2	-	-	-	-	-
1st Topological Structure Distance	-	-	-	0	3	-	-	1	-	-	-	1
4th Topological Structure Distance	-	-	-	-	-	-	-	-	-	-	-	-
5th Topological Structure Distance	-	-	-	2	-	2	-	-	-	-	-	-
Histogram Cosine Kernel	-	-	-	2	-	-	-	0	-	-	-	1
Jaccard Kernel	0	0	0	0	0	0	0	0	0	0	0	0

Discussion of the Results

- Most of the simplicial distances revealed changes in the clique topology of the networks of both hemispheres, from the **pre-ictal phase to the ictal phase**, especially in the θ band, which reinforces the findings obtained by the simplicial characterization measures.
- Regarding the laterality of the seizure focus, the analysis through simplicial distances did not find any statistically significant difference between the **left and right hemispheres clique topology in the ictal phase**.

Conclusions

- In General: Directed higher-order connectivity analysis of digraphs can be used to reveal differences between digraphs with apparently similar topology as we increase the level of organization/connection of cliques.
- Epileptic Brain Networks: We found evidence that the analysis of higher-order structures represented by q-digraphs obtained from G-connectivity networks may be a reliable way to find biomarkers associated with epileptic networks, but its establishment as a viable rigorous method will depend on future work.

Codes are available for Python in the DigplexQ package

```
pip3 install digplexq
```

https://github.com/heitorbaldo/DigplexQ

DigplexQ (Python package)

Soon available for Julia

DigplexQ.jl (Julia package)

Acknowledgements

Universidade de São Paulo

- Thanks to CAPES for the financial support.
- Many thanks to Prof. Koichi Sameshima, Prof. André Fujita, and Prof. Luiz Baccalá for guidance, insights, and support for this work.
- Thanks to my family and friends for support and encouragement throughout this journey.
- Thanks to the defense committee members, Prof. Daniel Takahashi, Prof. Mitre Dourado, and Prof. Carlos Hoppen.

References I

[1] Detti, P., Vatti, G., & Zabalo Manrique de Lara, G. (2020). **EEG Synchronization Analysis for Seizure Prediction: A Study on Data of Noninvasive Recordings.** *Processes*, 8(7), 846.

[2] Hu, Y., Zhang, Q., Li, R., Potter, T., & Zhang, Y. (2019). **Graph-based Brain Network Analysis in Epilepsy: an EEG Study**, *2019 9th International IEEE/EMBS Conference on Neural Engineering (NER)*, San Francisco, CA, USA, 130-133.

[3] Reimann, M. W., Nolte, M., Scolamiero, M., Turner, K., Perin, R., Chindemi, G., Dłotko, P., Levi, R., Hess, K., & Markram, H. (2017). **Cliques of Neurons Bound into Cavities Provide a Missing Link between Structure and Function.** *Frontiers in computational neuroscience*, 11, 48.