Signal Processing and Learning over Topological Spaces

Claudio Battiloro

Harvard University, Cambridge, MA, USA

Best Ph.D. Thesis Award, GTTI Annual Meeting, Pisa, September 5 2024
Thesis Supervisor: Prof. Paolo Di Lorenzo

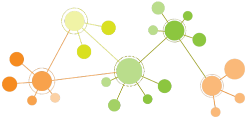
Thesis Outline

- The aim of this thesis is to introduce a variety of signal processing methodologies specifically designed to model, interpret, and learn from data defined on topological spaces
- The primary motivation is addressing the constraints encountered with traditional graph-based representations
- This thesis emphasizes the necessity to account for sophisticated, multiway, and geometry-sensitive interactions that are not captured by conventional graph models
- The implications of these developments are potentially profound for the signal processing and machine learning communities

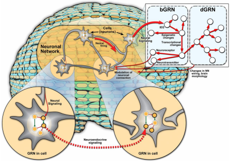
Topological Signal Processing

Goals and Motivation

- Graph-based representation: data are associated with the vertices of a graph to capture pairwise relations encoded by the presence of links

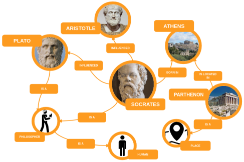


- In many systems (biological, brain, social networks,...) the complex interactions among data cannot be reduced to dyadic relationships



(a) In Gene Regulatory Networks, some reactions occur when a set of genes interact

(b) In Social Networks, agents can interact in a group without having pairwise connetions



(c) In Knowledge Graphs, higher-order relationship could provide further insight and analysis

What combinatorial topological spaces do we need to incorporate higher-order relationships?

Go beyond graphs: Simplicial Complexes and Cell complexes

- In this presentation I will focus on Simplicial Complexes for the sake of simplicity, Cell Complexes can be seen as a further generalization

- Simplicial complex: Given a finite set of vertices $\mathcal{V}$, a k-simplex is a subset of $\mathcal{V}$ with cardinality $k+1$. A simplicial Complex $\mathcal{X}^{(K)}$ of order K, is a collection of k-simplices up to order K closed under inclusion

- **Example: co-authorships networks**
  - vertices A, B,C, D are authors
  - there is an edge if two authors have co-authored at least one paper (e.g. A-B but not B-D)
  - there is a triangle between three authors if they have co-authored at least one paper (e.g. A-B-C but not A-D-C)

Topological Signal Processing
Simplicial and Cell Complexes

What combinatorial topological spaces do we need to incorporate higher-order relationships?

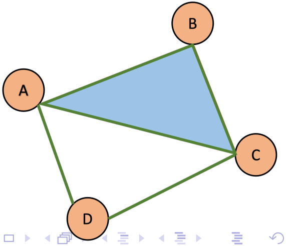
Go beyond graphs: Simplicial Complexes and Cell complexes

- In this presentation I will focus on Simplicial Complexes for the sake of simplicity, Cell Complexes can be seen as a further generalization
- Simplicial complex: Given a finite set of vertices $\mathcal{V}$, a k-simplex is a subset of $\mathcal{V}$ with cardinality $k+1$. A simplicial Complex $\mathcal{X}^{(K)}$ of order K, is a collection of k-simplices up to order K closed under inclusion
- **Example: co-authorships networks**
  - vertices A, B,C, D are authors
  - there is an edge if two authors have co-authored at least one paper (e.g. A-B but not B-D)
  - there is a triangle between three authors if they have co-authored at least one paper (e.g. A-B-C but not A-D-C)

What combinatorial topological spaces do we need to incorporate higher-order relationships?

Go beyond graphs: Simplicial Complexes and Cell complexes

- In this presentation I will focus on Simplicial Complexes for the sake of simplicity, Cell Complexes can be seen as a further generalization
- Simplicial complex: Given a finite set of vertices $\mathcal{V}$, a k-simplex is a subset of $\mathcal{V}$ with cardinality $k+1$. A simplicial Complex $\mathcal{X}^{(K)}$ of order K, is a collection of k-simplices up to order K closed under inclusion
- **Example: co-authorships networks**
  - vertices A, B,C, D are authors
  - there is an edge if two authors have co-authored at least one paper (e.g. A-B but not B-D)
  - there is a triangle between three authors if they have co-authored at least one paper (e.g. A-B-C but not A-D-C)



Simplicial Signals

- Simplicial signals: A k-simplicial signal is defined as a mapping from the set of all k-simplices contained in the complex to real-valued vectors

- We focus w.l.o.g on complexes $\mathcal{X}^{(2)}$ of order up to two, thus a set of vertices $\mathcal{V}$ with $|\mathcal{V}| = V$, a set of edges $\mathcal{E}$ with $|\mathcal{E}| = E$ and a set of triangles $\mathcal{T}$ with $|\mathcal{T}| = T$ are considered. The corresponding signals are defined as:

$$\mathbf{x}^{(0)} : \mathcal{V} \rightarrow \mathbb{R}^V, \qquad \mathbf{x}^{(1)} : \mathcal{E} \rightarrow \mathbb{R}^E, \qquad \mathbf{x}^{(2)} : \mathcal{T} \rightarrow \mathbb{R}^T,$$

  thus graph, edge and triangle signals, respectively

- Example: co-authorships networks
  - The graph signal $\mathbf{x}^{(0)} = [9, 2, 4, 8]$ collects the number of papers written by single authors (nodes)
  - The edge signal $\mathbf{x}^{(1)} = [1, 2, 3, 3, 6]$ collects the number of papers jointly written by pairs of authors (edges)
  - The triangle signal $\mathbf{x}^{(2)} = [3]$ collects the number of papers jointly written by triplets of authors (triangles)

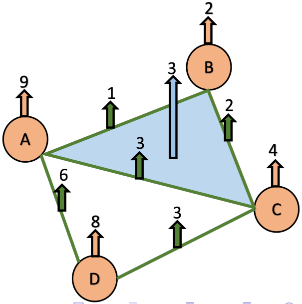
- Simplicial signals: A k-simplicial signal is defined as a mapping from the set of all k-simplices contained in the complex to real-valued vectors

- We focus w.l.o.g on complexes $\mathcal{X}^{(2)}$ of order up to two, thus a set of vertices $\mathcal{V}$ with $|\mathcal{V}| = V$, a set of edges $\mathcal{E}$ with $|\mathcal{E}| = E$ and a set of triangles $\mathcal{T}$ with $|\mathcal{T}| = T$ are considered. The corresponding signals are defined as:

$$\mathbf{x}^{(0)} : \mathcal{V} \to \mathbb{R}^V, \qquad \mathbf{x}^{(1)} : \mathcal{E} \to \mathbb{R}^E, \qquad \mathbf{x}^{(2)} : \mathcal{T} \to \mathbb{R}^T,$$

thus graph, edge and triangle signals, respectively

- **Example: co-authorships networks**
  - The graph signal $\mathbf{x}^{(0)} = [9, 2, 4, 8]$ collects the number of papers written by single authors (nodes)
  - The edge signal $\mathbf{x}^{(1)} = [1, 2, 3, 3, 6]$ collects the number of papers jointly written by pairs of authors (edges)
  - The triangle signal $\mathbf{x}^{(2)} = [3]$ collects the number of papers jointly written by triplets of authors (triangles)



- The structure of $\mathcal{X}^{(2)}$ is fully described by the set of its incidence matrices $\mathbf{B}_k$, $k = 1, 2$, that establish which k-simplices are incident to which $(k-1)$-simplices:

$$[\mathbf{B}_k]_{i,j} = \begin{cases} 0, & \text{if } \mathcal{H}_{k-1,i} \not\subset \mathcal{H}_{k,j} \\ 1, & \text{if } \mathcal{H}_{k-1,i} \subset \mathcal{H}_{k,j} \text{ and } \mathcal{H}_{k-1,i} \sim \mathcal{H}_{k,j} \\ -1, & \text{if } \mathcal{H}_{k-1,i} \subset \mathcal{H}_{k,j} \text{ and } \mathcal{H}_{k-1,i} \nsim \mathcal{H}_{k,j} \end{cases}$$

- From the incidence information, we can build the combinatorial Laplacian matrices:

$$\mathbf{L}_0 = \mathbf{B}_1\mathbf{B}_1^T,$$
$$\mathbf{L}_1 = \underbrace{\mathbf{B}_1^T\mathbf{B}_1}_{\mathbf{L}_d} + \underbrace{\mathbf{B}_2\mathbf{B}_2^T}_{\mathbf{L}_u}$$
$$\mathbf{L}_2 = \mathbf{B}_2^T\mathbf{B}_2$$

- The term $\mathbf{L}_d$ is called lower Laplacian and it encodes the lower adjacency $\mathcal{N}^d$ of edges $\rightarrow$ Two edges are lower adjacent if they share a common vertex

- The term $\mathbf{L}_u$ is called upper Laplacian and it encodes the upper adjacency $\mathcal{N}^u$ of edges $\rightarrow$ Two edges are upper adjacent if they are faces of the same triangle

- The structure of $\mathcal{X}^{(2)}$ is fully described by the set of its incidence matrices $\mathbf{B}_k$, $k = 1, 2$, that establish which k-simplices are incident to which $(k-1)$-simplices:

$$[\mathbf{B}_k]_{i,j} = \begin{cases} 0, & \text{if } \mathcal{H}_{k-1,i} \not\subset \mathcal{H}_{k,j} \\ 1, & \text{if } \mathcal{H}_{k-1,i} \subset \mathcal{H}_{k,j} \text{ and } \mathcal{H}_{k-1,i} \sim \mathcal{H}_{k,j} \\ -1, & \text{if } \mathcal{H}_{k-1,i} \subset \mathcal{H}_{k,j} \text{ and } \mathcal{H}_{k-1,i} \not\sim \mathcal{H}_{k,j} \end{cases}$$

- From the incidence information, we can build the combinatorial Laplacian matrices:

$$\mathbf{L}_0 = \mathbf{B}_1\mathbf{B}_1^T,$$
$$\mathbf{L}_1 = \underbrace{\mathbf{B}_1^T\mathbf{B}_1}_{\mathbf{L}_d} + \underbrace{\mathbf{B}_2\mathbf{B}_2^T}_{\mathbf{L}_u}$$
$$\mathbf{L}_2 = \mathbf{B}_2^T\mathbf{B}_2$$

- The term $\mathbf{L}_d$ is called lower Laplacian and it encodes the lower adjacency $\mathcal{N}^d$ of edges $\rightarrow$ Two edges are lower adjacent if they share a common vertex

- The term $\mathbf{L}_u$ is called upper Laplacian and it encodes the upper adjacency $\mathcal{N}^u$ of edges $\rightarrow$ Two edges are upper adjacent if they are faces of the same triangle

Sparse Signal Representation on Combinatorial Topological Spaces

Related publications:

Topological Slepians: Maximally Localized Representations of Signals Over Simplicial Complexes, C. Battiloro et al., IEEE ICASSP 2023

Parametric Dictionary Learning for Topological Signal Representation, C. Battiloro et al., EURASIP EUSIPCO 2023

Sparse Signal Representation
Motivations and state of the art

- Desiderata: Novel techniques for sparse signal representation over simplicial complexes
  - In the context of Topological Signal Processing, a natural basis for signal representation is given by the topological Fourier modes, generally leading to inefficients and non-sparse signal representations (as in classical SP, DSP and GSP)

- Simplicial Signal Processing (TSP):
  - Simplicial FIR filters [Isufi21]
  - Graph Slepians [Tsitsvero16]
  - Simplicial Wavelets (Hodgelets) [Roddenberry22]

- Contribution: We introduce novel model-based and data-driven techniques to design overcomplete dictionaries for signals over simplicial complexes.
  - We introduce Topological Slepians, a novel model-based class of signals that are maximally concentrated on the topological domain and perfectly bandlimited
  - We introduce a novel data-driven dictionary learning algorithm with guaranteed topology-awareness and locality
  - We test the proposed methods on a sparse representation task of real traffic data, showing superior performance w.r.t. other state-of-the-art methods

- A simplicial complex FIR filter acting on edge signals is a polynomial of the Laplacian defined as:

$$\mathbf{S} = \sum_{i=1}^{J} h_{u,i}\mathbf{L}_u^i + \sum_{i=1}^{J} h_{d,i}\mathbf{L}_d^i + h\mathbf{I}, \quad (1)$$

  where J is a positive integer and $h_{u,i},\ h_{d,i},\ h \in \mathbb{R}$

- We build a novel class of overcomplete topological dictionaries:

$$\mathbf{D} = [\mathbf{S}_1, ..., \mathbf{S}_P] \in \mathbb{R}^{N \times PN},$$

  where each $\mathbf{S}_p$, $p = 1, \ldots, P$ is defined as in (1) and has a different set of coefficients

- We collect the coefficients in a vector $\mathbf{h} \in \mathbb{R}^{(2J+1)P}$

- (Localization Guarantees) The $v-$th atom of the $p-$th sub-dictionary will have:
  - A component localized on the J-hop lower neighborhood of the $v-$th simplex
  - A component localized on the J-hop upper neighborhood of the $v-$th simplex

Simplicial Complex Filters and Dictionary Structure

- A simplicial complex FIR filter acting on edge signals is a polynomial of the Laplacian defined as:

$$\mathbf{S} = \sum_{i=1}^{J} h_{u,i}\mathbf{L}_u^i + \sum_{i=1}^{J} h_{d,i}\mathbf{L}_d^i + h\mathbf{I}, \qquad (1)$$

  where J is a positive integer and $h_{u,i},\ h_{d,i},\ h \in \mathbb{R}$

- We build a novel class of overcomplete topological dictionaries:

$$\mathbf{D} = [\mathbf{S}_1, ..., \mathbf{S}_P] \in \mathbb{R}^{N \times PN},$$

  where each $\mathbf{S}_p$, $p = 1, \ldots, P$ is defined as in (1) and has a different set of coefficients

- We collect the coefficients in a vector $\mathbf{h} \in \mathbb{R}^{(2J+1)P}$

- (Localization Guarantees) The $v-$th atom of the $p-$th sub-dictionary will have:
  - A component localized on the J-hop lower neighborhood of the $v-$th simplex
  - A component localized on the J-hop upper neighborhood of the $v-$th simplex

Parametric Topological Dictionary Learning

Problem Formulation

- A training set of M k-topological signals $\mathbf{Y} = [\mathbf{y}_1, \ldots, \mathbf{y}_M] \in \mathbb{R}^{N \times M}$ is given
- The aim is learning a dictionary which can represent the training signals as a sparse linear combination of the atoms (its columns) $\rightarrow$ We need to learn the filters coefficients $\mathbf{h}$
- The problem is cast as the joint optimization of the dictionary coefficients and the sparse signal representation:

$$(\mathbf{h}^*, \mathbf{X}^*) = \underset{\mathbf{h},\mathbf{X}}{\arg\min} \|\mathbf{Y} - \mathbf{D}\mathbf{X}\|_F^2 + \gamma \|\mathbf{h}\|_2^2$$

subject to:

$a)$ $\|\mathbf{x}_i\|_0 \leq K_0, \; i = 1, \ldots, M \rightarrow$ sparsity requirement

$b)$ $\mathbf{0} \preccurlyeq \mathbf{S}_p \preccurlyeq d\mathbf{I}, \; p = 1, \ldots, S \rightarrow$ non-negative & bounded spectra

$c)$ $(d - \epsilon)\mathbf{I} \preccurlyeq \sum_{p=1}^{S} \mathbf{S}_p \preccurlyeq (d + \epsilon)\mathbf{I} \rightarrow$ whole spectrum coverage

$d)$ $\mathbf{S}_p$ as in (1), $p = 1, \ldots, P \rightarrow$ parametric dictionary

where $\mathbf{x}_i$ is the $i-$th column of $\mathbf{X} \in \mathbb{R}^{PN \times M}$, i.e. the sparse signal representation of the $i-$th training signal

Parametric Topological Dictionary Learning

Problem Formulation

- A training set of M k-topological signals $\mathbf{Y} = [\mathbf{y}_1, \dots, \mathbf{y}_M] \in \mathbb{R}^{N \times M}$ is given
- The aim is learning a dictionary which can represent the training signals as a sparse linear combination of the atoms (its columns) $\rightarrow$ We need to learn the filters coefficients $\mathbf{h}$
- The problem is cast as the joint optimization of the dictionary coefficients and the sparse signal representation:

$$(\mathbf{h}^*, \mathbf{X}^*) = \underset{\mathbf{h}, \mathbf{X}}{\arg\min} \, \|\mathbf{Y} - \mathbf{D}\mathbf{X}\|_F^2 + \gamma \, \|\mathbf{h}\|_2^2$$

subject to:

$a)$ $\|\mathbf{x}_i\|_0 \leq K_0, \; i = 1, \dots, M$ $\rightarrow$ sparsity requirement

$b)$ $\mathbf{0} \preccurlyeq \mathbf{S}_p \preccurlyeq d\mathbf{I}, \; p = 1, \dots, S$ $\rightarrow$ non-negative & bounded spectra

$c)$ $(d - \epsilon)\mathbf{I} \preccurlyeq \sum_{p=1}^{S} \mathbf{S}_p \preccurlyeq (d + \epsilon)\mathbf{I}$ $\rightarrow$ whole spectrum coverage

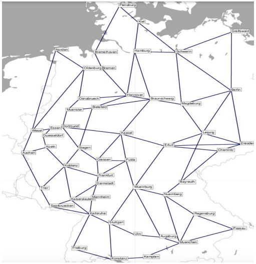
$d)$ $\mathbf{S}_p$ as in (1), $p = 1, \dots, P$ $\rightarrow$ parametric dictionary

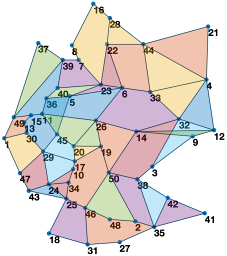
where $\mathbf{x}_i$ is the $i-$th column of $\mathbf{X} \in \mathbb{R}^{PN \times M}$, i.e. the sparse signal representation of the $i-$th training signal

Sparse Signal Representation

Real-World Numerical Results 1

- We consider the German National Research and Education Network operated by the German DFN-Verein (DFN)
- The complex consists of 50 nodes, 89 edges and 39 2-cells





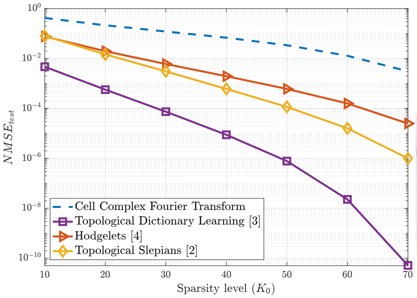
- The data traffic is aggregated daily over February 2005
- The data measurements are expressed in Mbit/sec and collected on each link

Sparse Signal Representation
Real-World Numerical Results 2

- We then have a collection of edge signals $\mathbf{Y} \in \mathbb{R}^{89\times 28}$
- For dictionary learning, we use 16 and 12 signals as training $\mathbf{Y}_{train}$ and test $\mathbf{Y}_{test}$ sets
- As evaluation metric,we use the test Normalized Mean Squared Error (NMSE):

$$NMSE_{test} = \frac{1}{16}\sum_{m=1}^{16} \frac{\|\mathbf{y}_{test,m} - \mathbf{D}\mathbf{x}_{test,m}\|_2^2}{\|\mathbf{y}_{test,m}\|_2^2}$$

- The Slepians dictionary shows superior performance w.r.t. Topological Wavelets (Hodgelets)
- The proposed dictionary learning algorithm achieves the best result
- This is expectable, because it estimates the underlying unknown generation model
- However, Topological Slepians can be leveraged even if no data are available



Topological Attention Neural Networks

Related publications:

Generalized Simplicial Attention Neural Networks, C. Battiloro et al., Sub. to IEEE TSIPN

Simplicial Attention Neural Networks, L. Giusti*, C. Battiloro* et al., ArXiv preprint 2022

- We saw that the (FIR) filtering of a simplicial (edge) signal $\mathbf{z}^{in}$ is then defined as:

$$\mathbf{z}^{out} = \sum_{j=1}^{J_d} w_{d,j} \mathbf{L}_d^j \mathbf{z}^{in} + \sum_{j=1}^{J_u} w_{u,j} \mathbf{L}_u^j \mathbf{z}^{in}$$

where $\mathbf{w}_d = \left[w_{d,1}, ..., w_{d,J_d}\right]^T \in \mathbb{R}^{J_d}$ and $\mathbf{w}_u = \left[w_{u,1}, ..., w_{u,J_u}\right]^T \in \mathbb{R}^{J_u}$ are the filter weights, $J_d \in \mathbb{N}$ is the lower filter order and $J_u \in \mathbb{N}$ is the upper filter order

- A Simplicial Convolutional Neural Network (SCN) layer is defined as a bank of simplicial filters followed by a point-wise non-linearity $\sigma(\cdot)$:

$$\mathbf{Z}^{out} = \sigma\left(\sum_{j=1}^{J_d} \mathbf{L}_d^j \mathbf{Z}^{in} \mathbf{W}_{d,j} + \sum_{j=1}^{J_u} \mathbf{L}_u^j \mathbf{Z}^{in} \mathbf{W}_{u,j} + \mathbf{Z}^{in} \mathbf{W}_h\right)$$

where $\mathbf{Z}^{out} \in \mathbb{R}^{E \times F'}$, $\mathbf{W}_h$,$\mathbf{W}_{d,j}$s and $\mathbf{W}_{u,j}$s $\in \mathbb{R}^{F \times F'}$ are learnable (filters) weights

Simplicial Attention Neural Networks

Motivations and state of the art

- Desiderata: Developing Simplcial Neural Networks (SNNs) architectures equipped with attention mechanisms
  - In the Deep Learning community, attention mechanisms are a class of techniques which allow to enhance some parts of the input data while diminishing other part
  - Therefore, topology-aware attention re-weights neighbours in a task-oriented fashion
- Simplicial Neural Networks (SNNs):
  - First SNN architecture [Ebli21]
  - Message Passing SNN [Bodnar21]
  - Hodge-Based SNN [Yang22]
- Attention Networks:
  - Valuable Attention Models [Bahdanau15][Vaswani17]
  - First Graph Attention Networks [Velickovic17]
- Contribution: We introduce the first Simplicial Attention Network (SAN) architecture. We generalize the original graph-attention mechanism in order to process simplex structured data.

Simplicial Attention Neural Networks

Layer Definition

- Re-weighting the neighbours translates in learning the Laplacian entries. A SAN layer is then defined as:

$$\mathbf{Z}^{out} = \sigma\left(\sum_{j=1}^{J_d} \mathbf{L}_{d,a}^{j} \mathbf{Z}^{in} \mathbf{W}_{d,j} + \sum_{j=1}^{J_u} \mathbf{L}_{u,a}^{j} \mathbf{Z}^{in} \mathbf{W}_{u,j} + \mathbf{Z}^{in} \mathbf{W}_h\right)$$

- Simplcial Attentional Mechanisms: the entries of the (attentional) Laplacians $\mathbf{L}_{u,a}$ and $\mathbf{L}_{d,a}$ are learned via an upper and a lower attentional mechanisms $a_u(\cdot)$ and $a_d(\cdot)$:

$$a_u : \mathbb{R}^{F'} \times \mathbb{R}^{F'} \times \mathbb{R}^{J_u} \to \mathbb{R} \qquad a_d : \mathbb{R}^{F'} \times \mathbb{R}^{F'} \times \mathbb{R}^{J_d} \to \mathbb{R}$$

- Under this setting, the entries of $\mathbf{L}_{u,a}$ and $\mathbf{L}_{d,a}$ are computed as:

$$\left[\mathbf{L}_{u,a}\right]_{i,j} = \text{softmax}_j\left(a_u\left(\left\{\left[\mathbf{Z}^{in}\right]_i \mathbf{W}_{u,k}\right\}_{k=1}^{J_u} \left\{\left[\mathbf{Z}^{in}\right]_j \mathbf{W}_{u,k}\right\}_{k=1}^{J_u}\right)\mathbb{I}(j \in \mathcal{N}_{u,i})\right)$$

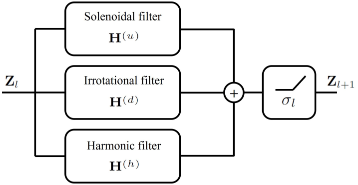
$$\left[\mathbf{L}_{d,a}\right]_{i,j} = \text{softmax}_j\left(a_d\left(\left\{\left[\mathbf{Z}^{in}\right]_i \mathbf{W}_{d,k}\right\}_{k=1}^{J_d}, \left\{\left[\mathbf{Z}^{in}\right]_j \mathbf{W}_{d,k}\right\}_{k=1}^{J_d}\right)\mathbb{I}(j \in \mathcal{N}_{d,i})\right)$$

where $\left[\mathbf{Z}^{in}\right]_j$ is the j-th row of $\mathbf{Z}^{in}$

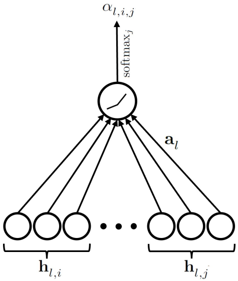
Simplicial Attention Networks

Layer Illustration and Example of Attention

- The SAN layer is made of three parallel branches followed by an aggreagation step →



- An possible attention mechanism is a single-layer feedforward neural network with LeakyReLU nonlinearity →



Simplicial Attention Neural Networks

Numerical Results

- Inductive (Supervised) Task. Trajectory Classification:

Architecture	Activation	Synthetic Flow (%)	Ocean Drifters (%)
MPSN [26]	Id	82.6 ± 3.0	73.0 ± 2.7
	ReLU	50.0 ± 0.0	46.5 ± 5.7
	Tanh	95.2 ± 1.8	72.5 ± 0.0
SCNN [32]	Id	66.5 ± 0.16	98.1 ± 0.01
	ReLU	100 ± 0.0	97.0 ± 0.01
	Tanh	67.2 ± 0.16	97.0 ± 0.16
SAT [33]	Id	99.7 ± 0.0	97.0 ± 0.01
	ReLU	100 ± 0.0	95.0 ± 0.00
	Tanh	100 ± 0.0	95.0 ± 0.01
SAN $\left(J^{(h)} = 0\right)$	Id	100 ± 0.0	97.1 ± 0.02
	ReLU	100 ± 0.0	97.0 ± 0.2
	Tanh	100 ± 0.0	97.5 ± 0.02
SAN	Id	100 ± 0.0	**99.0 ± 0.01**
	ReLU	100 ± 0.0	98.5 ± 0.01
	Tanh	**100 ± 0.0**	98.5 ± 0.01

- Transductive (Semisupervised) Task. Missing Data Imputation:

%Miss/Order N_k	Method	0 352	1 1474	2 3285	3 5019	4 5559	5 4547
10%	SNN [25]	91 ± 0.3	91 ± 0.2	91 ± 0.2	91 ± 0.2	91 ± 0.2	90 ± 0.4
	SCNN [32]	91 ± 0.4	91 ± 0.2	91 ± 0.2	91 ± 0.2	91 ± 0.2	91 ± 0.2
	SCNN (ours)	90 ± 0.3	91 ± 0.3	91 ± 0.3	93 ± 0.2	92 ± 0.2	94 ± 0.1
	SAT [33]	18 ± 0.0	31 ± 0.0	28 ± 0.1	34 ± 0.1	53 ± 0.1	55 ± 0.1z
	SAN	**91** ± 0.4	**95** ± **1.9**	**95** ± **1.9**	**97** ± **1.6**	**98** ± **0.9**	**98** ± **0.7**
20%	SNN [25]	81 ± 0.6	82 ± 0.3	81 ± 0.6	82 ± 0.3	81 ± 0.6	82 ± 0.5
	SCNN [32]	81 ± 0.7	82 ± 0.3	81 ± 0.7	82 ± 0.3	81 ± 0.7	83 ± 0.3
	SCNN (ours)	81 ± 0.6	83 ± 0.7	81 ± 0.6	88 ± 0.4	86 ± 0.7	89 ± 0.6
	SAT [33]	18 ± 0.0	30 ± 0.0	29 ± 0.1	35 ± 0.1	50 ± 0.1	58 ± 0.1
	SAN	**82** ± **0.8**	**91** ± **2.4**	**82** ± **0.8**	**96** ± **0.4**	**96** ± **1.3**	**97** ± **0.9**
30%	SNN [25]	72 ± 0.6	73 ± 0.4	81 ± 0.6	82 ± 0.3	81 ± 0.6	73 ± 0.5
	SCNN [32]	72 ± 0.5	73 ± 0.4	81 ± 0.7	82 ± 0.3	81 ± 0.7	74 ± 0.3
	SCNN (ours)	72 ± 0.6	76 ± 0.6	81 ± 0.6	82 ± 1.2	80 ± 0.7	86 ± 0.8
	SAT [33]	19 ± 0.0	33 ± 0.1	25 ± 0.1	33 ± 0.0	47 ± 0.1	53 ± 0.1
	SAN	**75** ± **2.1**	**89** ± **2.1**	**82** ± **0.8**	**94** ± **0.4**	**95** ± **0.5**	**96** ± **0.5**
40%	SNN [25]	63 ± 0.7	64 ± 0.3	81 ± 0.6	82 ± 0.3	81 ± 0.6	65 ± 0.3
	SCNN [32]	63 ± 0.6	64 ± 0.3	81 ± 0.7	82 ± 0.3	81 ± 0.7	65 ± 0.2
	SCNN (ours)	63 ± 0.7	67 ± 1.1	81 ± 0.6	79 ± 1.0	74 ± 1.1	83 ± 0.9
	SAT [33]	20 ± 0.0	29 ± 0.0	22 ± 0.0	43 ± 0.1	51 ± 0.1	50 ± 0.1
	SAN	**67** ± **1.9**	**85** ± **2.8**	**82** ± **0.8**	**91** ± **0.9**	**93** ± **1.1**	**95** ± **1.6**
50%	SNN [25]	54 ± 0.7	55 ± 0.5	81 ± 0.6	82 ± 0.3	81 ± 0.6	56 ± 0.3
	SCNN [32]	54 ± 0.6	55 ± 0.4	81 ± 0.7	82 ± 0.3	81 ± 0.7	56 ± 0.3
	SCNN (ours)	55 ± 0.9	60 ± 1.1	81 ± 0.6	71 ± 1.3	68 ± 1.3	79 ± 2.0
	SAT [33]	19 ± 0.0	30 ± 0.1	22 ± 0.0	32 ± 0.1	43 ± 0.0	48 ± 0.1
	SAN	**61** ± **1.9**	**79** ± **4.3**	**82** ± **0.8**	**88** ± **1.5**	**92** ± **0.7**	**94** ± **1.1**

From Latent Graph to Latent Topology Inference

UNIVERSITY OF OXFORD

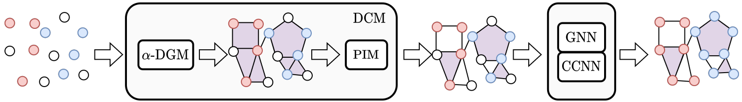
UNIVERSITY OF CAMBRIDGE

Related publications:

From Latent Graph to Latent Topology Inference: Differentiable Cell Complex Module, C. Battiloro*, Indro Spinelli* et al., ICLR 2024

Differentiable Cell Comple Module

- The Differentiable Cell Complex Module (DCM) is a function that first learns a graph describing the pairwise interactions among data points
- Then, it leverages the graph as the 1-skeleton of a regular cell complex describing multi-way interactions among data points
- The inferred topology is then used in two message-passing networks, at node and edge levels to solve the downstream task
- The whole architecture is trained in an end-to-end fashion



Signal Processing and Learning over Tangent Bundles

Related publications:

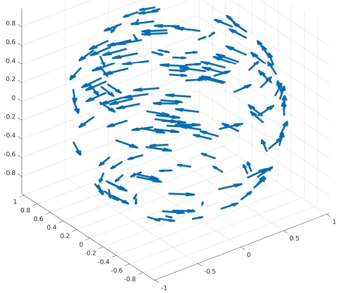
Tangent Bundle Convolutional Learning: from Manifolds to Cellular Sheaves and Back, C. Battiloro et al., IEEE Transaction on Signal Processing

Tangent bundle filters and neural networks: From manifolds to cellular sheaves and back, C.

Battiloro et al., IEEE ICASSP 2023

Tangent Bundle Convolutional Learning
From Manifolds to Cellular Sheaves and Back

- We introduce a novel convolution operation for tangent bundle signals, i.e. vector fields over Riemann manifolds
- We define tangent bundle filters and tangent bundle neural networks (TNNs)
- The proposed convolution generalizes most of the well-known convolutions
- We show, for the first time, that Sheaf Neural Networks (a generalization of Graph Neural Networks) converge to TNNs as the number of nodes goes to infinity
- We numerically evaluate the effectiveness of TNNs on various learning tasks



Conclusions

- In this thesis, we have shown that the exploration and exploitation of topological signal processing methods can unveil transformative potential in understanding complex data structures and extracting meaningful insights

- The marriage of topology and signal processing offers a robust framework for analyzing non-trivial data configurations, capturing intricate multi-way patterns often overlooked by traditional and graph-based methods

- As the field continues to evolve, it is anticipated that topological signal processing techniques will become an indispensable tool in the arsenal of modern data analysis and processing.

- My Linkedin `https://www.linkedin.com/in/claudio-battiloro-b4390b175/` and X `https://twitter.com/ClaBat9`:

Simplicial Signal Processing

Frequency Domain

- Simplicial signals of various order can be represented over the bases of the eigenvectors of the high order Laplacians
- Using the eigendecomposition $\mathbf{L}_1 = \mathbf{U}\boldsymbol{\Lambda}\mathbf{U}^T$, the Simplicial Fourier Transform (SFT) of order 1 of a simplicial (edge) signal $\mathbf{x}$ is defined as:

$$\widetilde{\mathbf{x}} \triangleq \mathbf{U}^T \mathbf{x}$$

- We refer to the eigenvalue domain (set) $\mathcal{A}$ of the SFT as the frequency domain
- High order Laplacians admit a Hodge decomposition, e.g. the 1-simplicial (edge) signal space can be decomposed as:

$$\mathbb{R}^E = \textbf{im}(\mathbf{B}_1^T) \bigoplus \textbf{im}(\mathbf{B}_2) \bigoplus \textbf{ker}(\mathbf{L}_1),$$

- Therefore, the eigenvalues of $\mathbf{L}_1$ are the union of the non-zero eigenvalues $\mathcal{F}^u$ of $\mathbf{L}_u$, the non-zero eigenvalues $\mathcal{F}^d$ of $\mathbf{L}_d$ and the zero eigenvalue of multiplicity $\dim(\text{Ker}(\mathbf{L}_1))$

- Simplicial signals of various order can be represented over the bases of the eigenvectors of the high order Laplacians
- Using the eigendecomposition $\mathbf{L}_1 = \mathbf{U}\boldsymbol{\Lambda}\mathbf{U}^T$, the Simplicial Fourier Transform (SFT) of order 1 of a simplicial (edge) signal $\mathbf{x}$ is defined as:

$$\widetilde{\mathbf{x}} \triangleq \mathbf{U}^T \mathbf{x}$$

- We refer to the eigenvalue domain (set) $\mathcal{A}$ of the SFT as the frequency domain
- High order Laplacians admit a Hodge decomposition, e.g. the 1-simplicial (edge) signal space can be decomposed as:

$$\mathbb{R}^E = \textbf{im}(\mathbf{B}_1^T) \bigoplus \textbf{im}(\mathbf{B}_2) \bigoplus \textbf{ker}(\mathbf{L}_1),$$

- Therefore, the eigenvalues of $\mathbf{L}_1$ are the union of the non-zero eigenvalues $\mathcal{F}^u$ of $\mathbf{L}_u$, the non-zero eigenvalues $\mathcal{F}^d$ of $\mathbf{L}_d$ and the zero eigenvalue of multiplicity $\dim(\text{Ker}(\mathbf{L}_1))$

Topological Slepians
Localization and Concentration Sets

- To formalize the concept of localization, we need two concentration operators
- Fix an edge concentration set $\mathcal{S} \subset \mathcal{E}$. The edge-limiting operator on $\mathcal{S}$ is defined as:

$$\mathbf{C}_{\mathcal{S}} = \mathrm{diag}(\mathbf{1}_{\mathcal{S}})$$

  where $\mathbf{1}_{\mathcal{S}} \in \mathbb{R}^E$ is a vector having ones in the indices specified in $\mathcal{S}$, and zero otherwise. An edge signal $\mathbf{x}$ is perfectly localized onto the set $\mathcal{S}$ if $\mathbf{C}_{\mathcal{S}}\mathbf{x} = \mathbf{x}$

- Fix a frequency concentration set $\mathcal{F} \subset \mathcal{A}$. The frequency-limiting operator on $\mathcal{F}$ is defined as:

$$\mathbf{B}_{\mathcal{F}} = \mathbf{U}\mathrm{diag}(\mathbf{1}_{\mathcal{F}})\mathbf{U}^T,$$

  An edge signal $\mathbf{x}$ is perfectly localized over the bandwidth $\mathcal{F}$ if $\mathbf{B}_{\mathcal{F}}\mathbf{x} = \mathbf{x}$

- To formalize the concept of localization, we need two concentration operators
- Fix an edge concentration set $\mathcal{S} \subset \mathcal{E}$. The edge-limiting operator on $\mathcal{S}$ is defined as:

$$\mathbf{C}_{\mathcal{S}} = \mathrm{diag}(\mathbf{1}_{\mathcal{S}})$$

where $\mathbf{1}_{\mathcal{S}} \in \mathbb{R}^E$ is a vector having ones in the indices specified in $\mathcal{S}$, and zero otherwise. An edge signal $\mathbf{x}$ is perfectly localized onto the set $\mathcal{S}$ if $\mathbf{C}_{\mathcal{S}}\mathbf{x} = \mathbf{x}$

- Fix a frequency concentration set $\mathcal{F} \subset \mathcal{A}$. The frequency-limiting operator on $\mathcal{F}$ is defined as:

$$\mathbf{B}_{\mathcal{F}} = \mathbf{U}\mathrm{diag}(\mathbf{1}_{\mathcal{F}})\mathbf{U}^T,$$

An edge signal $\mathbf{x}$ is perfectly localized over the bandwidth $\mathcal{F}$ if $\mathbf{B}_{\mathcal{F}}\mathbf{x} = \mathbf{x}$

- To formalize the concept of localization, we need two concentration operators
- Fix an edge concentration set $\mathcal{S} \subset \mathcal{E}$. The edge-limiting operator on $\mathcal{S}$ is defined as:

$$\mathbf{C}_{\mathcal{S}} = \mathrm{diag}(\mathbf{1}_{\mathcal{S}})$$

where $\mathbf{1}_{\mathcal{S}} \in \mathbb{R}^E$ is a vector having ones in the indices specified in $\mathcal{S}$, and zero otherwise. An edge signal $\mathbf{x}$ is perfectly localized onto the set $\mathcal{S}$ if $\mathbf{C}_{\mathcal{S}}\mathbf{x} = \mathbf{x}$

- Fix a frequency concentration set $\mathcal{F} \subset \mathcal{A}$. The frequency-limiting operator on $\mathcal{F}$ is defined as:

$$\mathbf{B}_{\mathcal{F}} = \mathbf{U}\mathrm{diag}(\mathbf{1}_{\mathcal{F}})\mathbf{U}^T,$$

An edge signal $\mathbf{x}$ is perfectly localized over the bandwidth $\mathcal{F}$ if $\mathbf{B}_{\mathcal{F}}\mathbf{x} = \mathbf{x}$

Topological Slepians
Problem Formulation

- Topological Slepians are the set of orthonormal vectors that are maximally concentrated over $\mathcal{S}$, and perfectly localized onto $\mathcal{F}$

- Formally, they are the set of vectors solving the problem:

$$\begin{aligned}\boldsymbol{\psi}_i = &\arg\max_{\boldsymbol{\psi}_i} \; ||\mathbf{C}_{\mathcal{S}}\boldsymbol{\psi}_i||_2^2 \\ &\text{subject to} \\ &\text{(a) } ||\boldsymbol{\psi}_i|| = 1, \text{ (b) } \mathbf{B}_{\mathcal{F}}\boldsymbol{\psi}_i = \boldsymbol{\psi}_i, \text{ (c) } <\boldsymbol{\psi}_i, \boldsymbol{\psi}_j>= 0\end{aligned}$$

- The solution are the eigenvectors of the operator $\mathbf{B}_{\mathcal{F}}\mathbf{C}_{\mathcal{S}}\mathbf{B}_{\mathcal{F}}$, i.e.:

$$\mathbf{B}_{\mathcal{F}}\mathbf{C}_{\mathcal{S}}\mathbf{B}_{\mathcal{F}}\psi_i = \lambda_i\psi_i, \; \lambda_1 \geq \lambda_2 \geq ... \geq \lambda_C > 0$$

- Let $\mathbf{\Psi}_{\mathcal{S},\mathcal{F}}$ be the set of slepians corresponding to the concentration sets $\mathcal{S}$ and $\mathcal{F}$. An (overcomplete) dictionary of topological slepians is of the form:

$$\mathbf{D}_{\mathcal{C}} = \left[\mathbf{\Psi}_{\mathcal{S}_1,\mathcal{F}_1}, ..., \mathbf{\Psi}_{\mathcal{S}_i,\mathcal{F}_i}, ..., \mathbf{\Psi}_{\mathcal{S}_M,\mathcal{F}_M}\right],$$

collecting M sets of slepians obtained from the pairs of concentration sets $\{\mathcal{S}_i, \mathcal{F}_i\}_{i=1}^M$

- Topological Slepians are the set of orthonormal vectors that are maximally concentrated over $\mathcal{S}$, and perfectly localized onto $\mathcal{F}$

- Formally, they are the set of vectors solving the problem:

$$\boldsymbol{\psi}_i = \underset{\boldsymbol{\psi}_i}{\arg\max} \; ||\mathbf{C}_{\mathcal{S}}\boldsymbol{\psi}_i||_2^2$$
$$\text{subject to}$$
$$\text{(a) } ||\boldsymbol{\psi}_i|| = 1, \text{ (b) } \mathbf{B}_{\mathcal{F}}\boldsymbol{\psi}_i = \boldsymbol{\psi}_i, \text{ (c) } <\boldsymbol{\psi}_i, \boldsymbol{\psi}_j>= 0$$

- The solution are the eigenvectors of the operator $\mathbf{B}_{\mathcal{F}}\mathbf{C}_{\mathcal{S}}\mathbf{B}_{\mathcal{F}}$, i.e.:

$$\mathbf{B}_{\mathcal{F}}\mathbf{C}_{\mathcal{S}}\mathbf{B}_{\mathcal{F}}\boldsymbol{\psi}_i = \lambda_i\boldsymbol{\psi}_i, \; \lambda_1 \geq \lambda_2 \geq ... \geq \lambda_C > 0$$

- Let $\boldsymbol{\Psi}_{\mathcal{S},\mathcal{F}}$ be the set of slepians corresponding to the concentration sets $\mathcal{S}$ and $\mathcal{F}$. An (overcomplete) dictionary of topological slepians is of the form:

$$\mathbf{D}_{\mathcal{C}} = \left[\boldsymbol{\Psi}_{\mathcal{S}_1,\mathcal{F}_1}, ..., \boldsymbol{\Psi}_{\mathcal{S}_i,\mathcal{F}_i}, ..., \boldsymbol{\Psi}_{\mathcal{S}_M,\mathcal{F}_M}\right],$$

collecting M sets of slepians obtained from the pairs of concentration sets $\{\mathcal{S}_i, \mathcal{F}_i\}_{i=1}^M$

- Topological Slepians are the set of orthonormal vectors that are maximally concentrated over $\mathcal{S}$, and perfectly localized onto $\mathcal{F}$

- Formally, they are the set of vectors solving the problem:

$$\boldsymbol{\psi}_i = \underset{\boldsymbol{\psi}_i}{\arg\max} \ ||\mathbf{C}_{\mathcal{S}}\boldsymbol{\psi}_i||_2^2$$

$$\text{subject to}$$

$$\text{(a) } ||\boldsymbol{\psi}_i|| = 1, \text{ (b) } \mathbf{B}_{\mathcal{F}}\boldsymbol{\psi}_i = \boldsymbol{\psi}_i, \text{ (c) } <\boldsymbol{\psi}_i, \boldsymbol{\psi}_j>= 0$$

- The solution are the eigenvectors of the operator $\mathbf{B}_{\mathcal{F}}\mathbf{C}_{\mathcal{S}}\mathbf{B}_{\mathcal{F}}$, i.e.:

$$\mathbf{B}_{\mathcal{F}}\mathbf{C}_{\mathcal{S}}\mathbf{B}_{\mathcal{F}}\boldsymbol{\psi}_i = \lambda_i\boldsymbol{\psi}_i, \ \lambda_1 \geq \lambda_2 \geq ... \geq \lambda_C > 0$$

- Let $\boldsymbol{\Psi}_{\mathcal{S},\mathcal{F}}$ be the set of slepians corresponding to the concentration sets $\mathcal{S}$ and $\mathcal{F}$. An (overcomplete) dictionary of topological slepians is of the form:

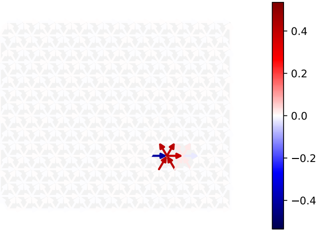
$$\mathbf{D}_{\mathcal{C}} = \Big[\boldsymbol{\Psi}_{\mathcal{S}_1,\mathcal{F}_1}, ..., \boldsymbol{\Psi}_{\mathcal{S}_i,\mathcal{F}_i}, ..., \boldsymbol{\Psi}_{\mathcal{S}_M,\mathcal{F}_M}\Big],$$

collecting M sets of slepians obtained from the pairs of concentration sets $\{\mathcal{S}_i, \mathcal{F}_i\}_{i=1}^M$

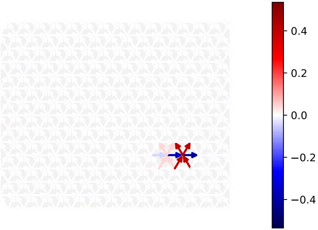
Topological Slepians

Example of Slepians

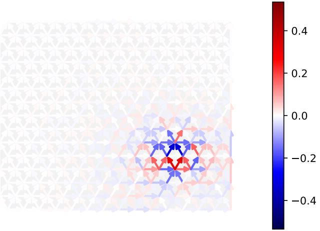
(d) Concentration Set



(e) 1st Slepian



(f) 2nd Slepian



(g) 3rd Slepian