From Latent Graph to Latent Topology Inference: Differentiable Cell Complex Module

Claudio Battiloro*,[1] Indro Spinelli*,[2] Lev Telyatnikov[2]
Michael Bronstein[3] Simone Scardapane[2] Paolo Di Lorenzo[2]

* *Equal Contribution*
[1] *Harvard University* [2] *Sapienza University of Rome* [3] *University of Oxford*

ICLR

International Conference on Learning Representations 2024, May 7-11, Vienna, Austria

Finanziato dall'Unione europea NextGenerationEU | Ministero dell'Università e della Ricerca | Italiadomani PIANO NAZIONALE DI RIPRESA E RESILIENZA | FAIR Future Artificial Intelligence Research

Contribution

We introduce the concept of Latent Topology Inference (LTI), whose goal is not (only) learning a graph but rather learning a higher-order combinatorial topological space describing multi-way interactions among data points

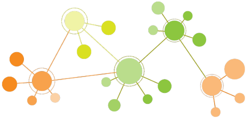
As a first instance of LTI, we introduce the Differentiable Cell Complex Module (DCM), a novel architecture that dynamically learns a latent regular cell complex to improve the downstream task

The DCM is tested on several homophilic and heterophilic datasets

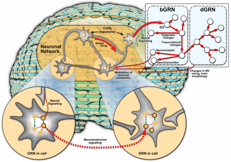
Significant accuracy gains on heterophilic benchmarks with provided input graphs show that the DCM leads to robust performance even when the input graph does not fit the data well

Topological Deep Learning
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

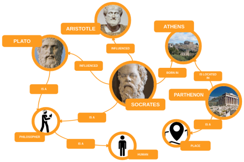


In many systems (biological, brain, social networks,...) the complex interactions among data cannot be reduced to dyadic relationships



(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



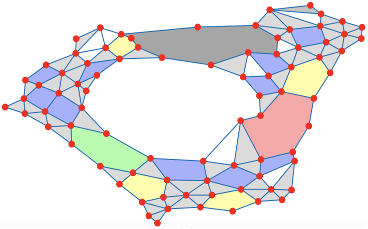
(c) In Knowledge Graphs, higher-order relationship could provide further insight and analysis

Topological Deep Learning
Regular Cell Complexes

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

Go beyond graphs: Simplicial Complexes, Cell Complexes, Cellular Sheaves, ...

A graph is a cell complex of order 1, containing only cells of order 0 (nodes, in red) and 1 (edges, in blue). Here we consider a cell complex $\mathcal{C}_{\mathcal{G}}$ of order 2 as a graph augmented with order 2 cells (***polygons***, in various colors) being some of its induced cycles.

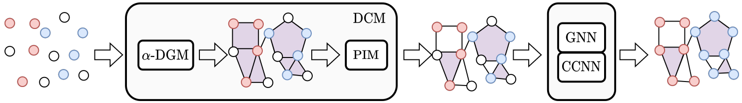


The Differentiable Cell Complex Module (DCM) is a function that first learns a graph describing the pairwise interactions among data points

Then, it leverages the graph as the skeleton of a regular cell complex describing multi-way interactions among data points

The inferred topology, i.e. the inferred edges and polygons, 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

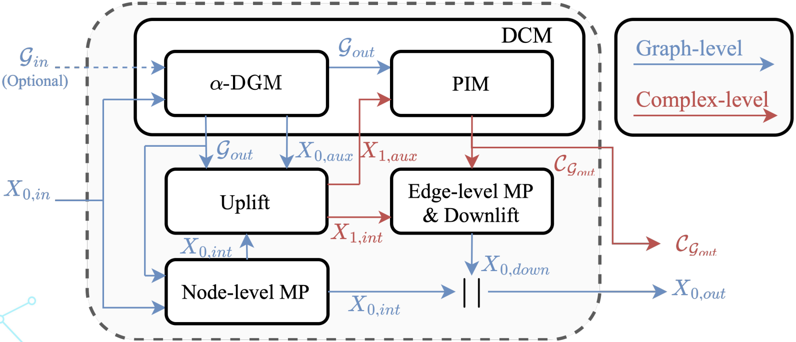


From Latent Graph To Latent Topology Inference

Differentiable Cell Complex Module 2

Both edge and polygon sampling is realized through the α-entmax class of functions

In this way, we obtain flexible, non-regular, and sparse topologies



Homophilic Numerical Results

Table: Homophilic-graph node classification (Test acc. in % avg.ed over 10 splits).

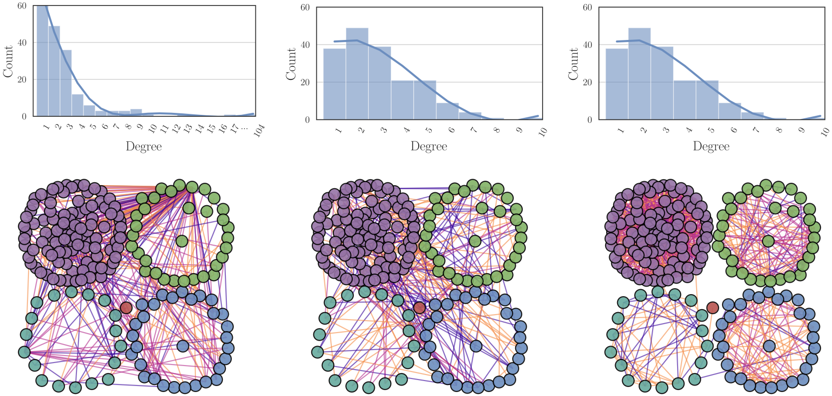
		Cora	CiteSeer	PubMed	Physics	CS
	Model/Hom. level	0.81	0.74	0.80	0.93	0.80
w/o graph	MLP	58.92 ± 3.28	59.48 ± 2.14	85.75 ± 1.02	94.91 ± 0.28	87.80 ± 1.54
	KCM	78.47 ± 2.09	75.20 ± 2.41	86.66 ± 0.91	95.61 ± 0.18	95.14 ± 0.32
	DGM-E	62.48 ± 3.24	62.47 ± 3.20	83.89 ± 0.70	94.03 ± 0.45	76.05 ± 6.89
	DGM-M	70.85 ± 4.30	68.86 ± 2.97	**87.43** ± 0.40	95.25 ± 0.36	92.22 ± 1.09
	DCM	**78.80** ± 1.84	**76.47** ± 2.45	87.38 ± 0.91	**96.45** ± 0.12	**95.40** ± 0.40
	DCM ($\alpha = 1$)	**78.73** ± 1.99	**76.32** ± 2.75	**87.47** ± 0.77	**96.22** ± 0.27	**95.35** ± 0.37
w graph	GCN	83.11 ± 2.29	69.97± 2.00	85.75 ± 1.01	95.51 ± 0.34	87.28 ± 1.54
	GCN2	**87.85** ± 1.41	78.53 ± 2.66	**89.60** ± 0.70	**97.41** ± 0.34	95.05 ± 0.38
	GAT	**89.81** ± 1.77	78.18 ± 2.31	88.53 ± 0.61	**98.87** ± 0.30	94.42 ± 0.70
	KCM	78.43 ± 2.11	75.23 ± 2.45	86.61 ± 0.95	96.16 ± 0.17	95.46 ± 0.36
	CWN	88.63 ± 1.91	75.53 ± 2.13	87.97 ± 0.77	96.23 ± 0.24	93.52 ± 0.59
	GCN+CCCN	86.09 ± 1.82	78.36 ± 3.33	88.59 ± 0.67	96.90 ± 0.30	95.31 ± 0.49
	DGM-E	82.11 ± 4.24	72.35 ± 1.92	87.69 ± 0.67	95.96 ± 0.40	87.17 ± 3.82
	DGM-M	86.63 ± 3.25	75.42 ± 2.39	87.82 ± 0.59	96.21 ± 0.44	92.86 ± 0.96
	DCM	85.78 ± 1.71	**78.72** ± 2.84	88.49 ± 0.62	96.99 ± 0.44	**95.79** ± 0.48
	DCM ($\alpha = 1$)	85.97 ± 1.86	**78.60** ± 3.16	**88.61** ± 0.69	96.69 ± 0.46	**95.78** ± 0.49

Heterophilic Numerical Results

Table: Heterophilic-graph node classification (Test acc. in % avg.ed over 10 splits).

		Texas	Wisconsin	Squirrel	Chameleon
	Model/Hom. level	0.11	0.21	0.22	0.23
w/o graph	MLP	77.78 ± 10.24	85.33 ± 4.99	30.44 ± 2.55	40.35 ± 3.37
	KCM	84.12 ± 11.37	87.10 ± 5.15	35.15 ± 1.38	52.12 ± 2.02
	DGM-E	80.00 ± 8.31	**88.00** ± 5.65	34.35 ± 2.34	48.90 ± 3.61
	DGM-M	81.67 ± 7.05	**89.33** ± 1.89	35.00 ± 2.35	48.90 ± 3.61
	DCM	**85.71** ± 7.87	87.49 ± 5.94	**35.55** ± 2.24	**53.63** ± 3.07
	DCM ($\alpha = 1$)	**84.96** ± 10.24	86.72 ± 6.02	**35.25** ± 2.22	**53.67** ± 3.19
w graph	GCN	41.66 ± 11.72	47.20 ± 9.76	24.19 ± 2.56	32.56 ± 3.53
	GCN2	75.50 ± 7.81	74.57 ± 5.38	33.09 ± 1.76	49.50 ± 3.02
	GAT	66.72 ± 11.22	60.52 ± 9.23	**35.07** ± 2.13	50.73 ± 3.12
	KCM	83.92 ± 11.15	84.92 ± 5.21	34.47 ± 1.49	**53.12** ± 2.02
	CWN	65.87 ± 6.33	64.57 ± 7.12	32.44 ± 2.75	43.86 ± 2.51
	GCN+CCCN	84.43 ± 9.11	84.03 ± 5.42	OOM	OOM
	DGM-E	60.56 ± 8.03	70.67 ± 10.49	29.87 ± 2.46	44.19 ± 3.85
	DGM-M	62.78 ± 9.31	76.00 ± 3.26	30.44 ± 2.38	45.68 ± 2.66
	DCM	**84.87** ± 10.04	**86.33** ± 5.14	34.95 ± 2.59	53.05 ± 3.00
	DCM ($\alpha = 1$)	**84.96** ± 5.60	**85.36** ± 5.05	**35.13** ± 2.27	**53.76** ± 3.72

LTI's Visualization



Evolution of the latent complex for the Texas dataset. Edges in orange, triangles in lilac, squares in purple ($K_{max} = 4$). Homophily levels $h = [0.11, 0.44, 0.99]$.

Contacts

Claudio's Linkedin https://www.linkedin.com/in/claudio-battiloro-b4390b175/ and X https://twitter.com/ClaBat9:

Indro's Linkedin https://www.linkedin.com/in/indro-spinelli-770380123/ and X https://twitter.com/IndroSpinelli: