Directional Sheaf Hypergraph Networks: Unifying Learning on Directed and Undirected Hypergraphs

Emanuele Mule[1], Stefano Fiorini, Antonio Purificato[1], Federico Siciliano[1], Stefano Coniglio[2], Fabrizio Silvestri[1]

Sapienza University of Rome[1]

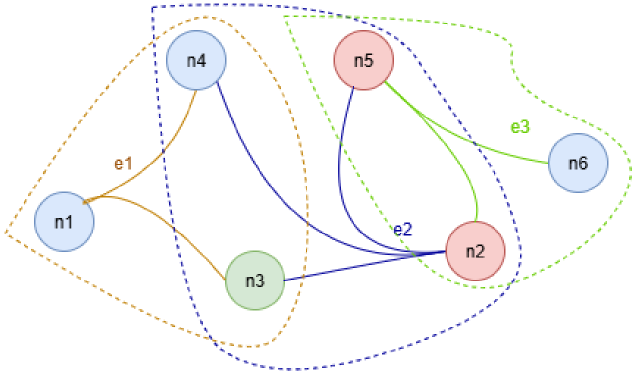
University of Bergamo[2]

Learning on Hypergraphs

- **Hypergraph**: generalization of a graph where each hyperedge can connect $k \geq 2$ nodes simultaneously

- **Hypergraph Neural Networks**: standard for learning on higher-order relational structures, representing **hyperedges as set of nodes**

- Most existing methods **ignore directionality**, treating such relations as **symmetric**



Undirected Hypergraph

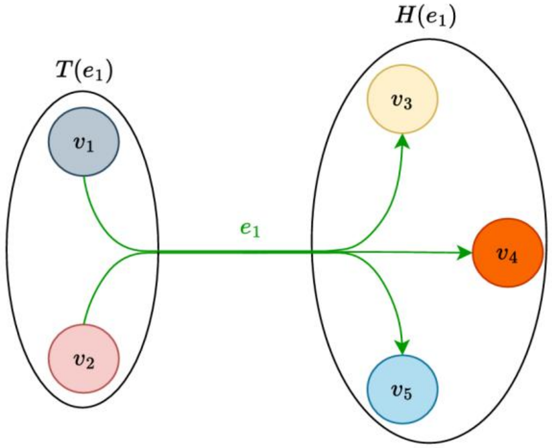
Directed Hypergraphs

(Definition) Directed Hypergraph

$$\mathcal{H} = (V, E)$$

V denotes the vertex set, E denotes the directed hyperedges set.

$$\forall e \in E,\ T(e), H(e) \subseteq V, \quad T(e) \cap H(e) = \emptyset$$

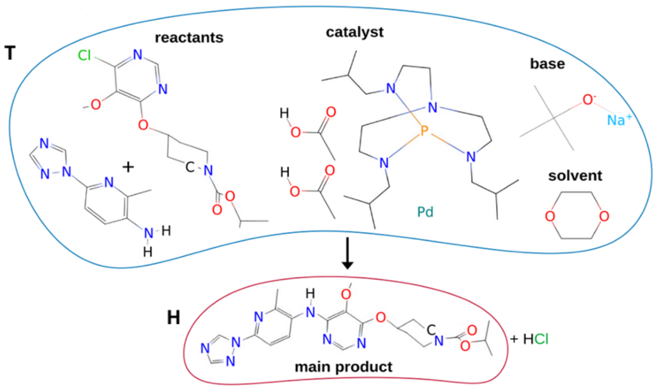


Directed Hyperedge

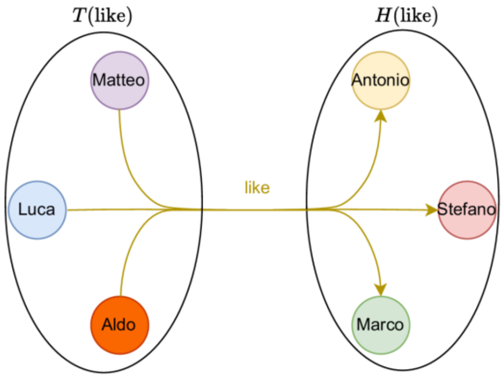
Applications

Many potential **applications** to real-world scenarios, just to name a few:

Chemical Reactions



Social Networks



Sheaf Hypergraph Networks

- Sheaf Hypergraph Networks[1] extend Sheaf Neural Networks[2] to **higher-order structures**
- However, the proposed Sheaf Hypergraph Laplacian is **not positive semidefinite (PSD)**, breaking its connection with standard spectral convolution operators
- Additionally, it does not model **hyperedge directionality**, limiting expressiveness in many applications

$$(\mathbf{L}_{\mathcal{F}})_{uu} = \sum_{e:\, u \in e} \frac{1}{\delta_e} \mathcal{F}_{u \trianglelefteq e}^{\top} \mathcal{F}_{u \trianglelefteq e}$$

$$(\mathbf{L}_{\mathcal{F}})_{uv} = -\sum_{\substack{e:\, u,v \in e \\ v \neq u}} \frac{1}{\delta_e} \mathcal{F}_{u \trianglelefteq e}^{\top} \mathcal{F}_{v \trianglelefteq e}$$

Sheaf Hypergraph Laplacian

[1] Duta et al., Sheaf Hypergraph Networks. In NeurIPS 2023.
[2] Bodnar et al., Neural Sheaf Diffusion: A Topological Perspective on Heterophily and Oversmoothing in GNNs. In NeurIPS 2022.

Directed Hypergraph Cellular Sheaf

A Directed Hypergraph Cellular Sheaf $\vec{\mathcal{F}}$ over $\mathcal{H}$ consists of the following:

1. A complex-valued and directional matrix $\mathcal{S}^{(q)} \in \mathbb{C}^{m \times n}$ with $q \in \mathbb{R}$ defined for each hyperedge-node incidence as follows:

$$\mathcal{S}^{(q)}_{u \trianglelefteq e} = \begin{cases} 1 & \text{if } u \in H(e) \\ e^{-2\pi i q} & \text{if } u \in T(e) \\ 0 & \text{otherwise} \end{cases}$$

2. A vector space $\mathcal{F}(u) \in \mathbb{C}^d$
3. A vector space $\mathcal{F}(e) \in \mathbb{C}^d$
4. For each incidence $u \trianglelefteq e$, a restriction map $\vec{\mathcal{F}}_{u \trianglelefteq e}$: $\mathcal{F}(u) \to \mathcal{F}(e)$ defined as: $\vec{\mathcal{F}}_{u \trianglelefteq e} = \mathcal{S}^{(q)}_{u \trianglelefteq e} \cdot \mathcal{F}_{u \trianglelefteq e}$, where $\mathcal{F}_{u \trianglelefteq e} \in \mathbb{R}^{d \times d}$ is a real, directionless, restriction map.

Directed Sheaf Hypergraph Laplacian: Definition

We define the Directed Sheaf Hypergraph Laplacian is defined as follows:

$$(L^{\vec{\mathcal{F}}})_{uv} = \begin{cases} D_u - \sum_{e:u\in e} \frac{1}{\delta_e} \mathcal{F}_{u\trianglelefteq e}^\top \mathcal{F}_{u\trianglelefteq e} = \sum_{e:u\in e} (1 - \frac{1}{\delta_e}) \mathcal{F}_{u\trianglelefteq e}^\top \mathcal{F}_{u\trianglelefteq e} & u = v \\ - \sum_{e:u,v\in e} \frac{1}{\delta_e} \vec{\mathcal{F}}_{u\trianglelefteq e}^\dagger \vec{\mathcal{F}}_{v\trianglelefteq e} = - \sum_{e:u,v\in e} \frac{1}{\delta_e} (\mathcal{S}_{u\trianglelefteq e}^{(q)})^\dagger (\mathcal{S}_{v\trianglelefteq e}^{(q)}) \mathcal{F}_{u\trianglelefteq e}^\top \mathcal{F}_{v\trianglelefteq e} & u \neq v. \end{cases}$$

In the special case $q = \frac{1}{4} \rightarrow e^{-2\pi i q} = -i$ reads:

$$(L^{\vec{\mathcal{F}}})_{uv} = \begin{cases} \sum_{e:u\in e} \left(1 - \frac{1}{\delta_e}\right) \mathcal{F}_{u\trianglelefteq e}^\top \mathcal{F}_{u\trianglelefteq e}, & u = v, \\ -\sum_{\substack{e\in E \\ u,v\in H(e) \\ \vee u,v\in T(e)}} \frac{1}{\delta_e} \mathcal{F}_{u\trianglelefteq e}^\top \mathcal{F}_{v\trianglelefteq e} - \mathbf{i} \left(\sum_{\substack{e\in E \\ u\in T(e) \\ \wedge v\in H(e)}} \frac{1}{\delta_e} \mathcal{F}_{u\trianglelefteq e}^\top \mathcal{F}_{v\trianglelefteq e} - \sum_{\substack{e\in E \\ u\in H(e) \\ \wedge v\in T(e)}} \frac{1}{\delta_e} \mathcal{F}_{u\trianglelefteq e}^\top \mathcal{F}_{v\trianglelefteq e} \right), & u \neq v. \end{cases}$$

Directed Sheaf Hypergraph Laplacian: Definition

We define the Directed Sheaf Hypergraph Laplacian is defined as follows:

$$(L^{\vec{\mathcal{F}}})_{uv} = \begin{cases} D_u - \sum_{e:u\in e} \frac{1}{\delta_e} \mathcal{F}^\top_{u\trianglelefteq e}\mathcal{F}_{u\trianglelefteq e} = \sum_{e:u\in e}(1-\frac{1}{\delta_e})\mathcal{F}^\top_{u\trianglelefteq e}\mathcal{F}_{u\trianglelefteq e} & u=v \\ -\sum_{e:u,v\in e}\frac{1}{\delta_e}\vec{\mathcal{F}}^\dagger_{u\trianglelefteq e}\vec{\mathcal{F}}_{v\trianglelefteq e} = -\sum_{e:u,v\in e}\frac{1}{\delta_e}(\mathcal{S}^{(q)}_{u\trianglelefteq e})^\dagger(\mathcal{S}^{(q)}_{v\trianglelefteq e})\mathcal{F}^\top_{u\trianglelefteq e}\mathcal{F}_{v\trianglelefteq e} & u\neq v.\end{cases}$$

In the special case $q = \frac{1}{4} \rightarrow e^{-2\pi i q} = -i$ reads:

$$(L^{\vec{\mathcal{F}}})_{uv} = \begin{cases} \sum_{e:u\in e}\left(1-\frac{1}{\delta_e}\right)\mathcal{F}^\top_{u\trianglelefteq e}\mathcal{F}_{u\trianglelefteq e}, & u=v, \\ -\sum_{\substack{e\in E \\ u,v\in H(e) \\ \vee u,v\in T(e)}}\frac{1}{\delta_e}\mathcal{F}^\top_{u\trianglelefteq e}\mathcal{F}_{v\trianglelefteq e} - \mathbf{i}\left(\sum_{\substack{e\in E \\ u\in T(e) \\ \wedge v\in H(e)}}\frac{1}{\delta_e}\mathcal{F}^\top_{u\trianglelefteq e}\mathcal{F}_{v\trianglelefteq e} - \sum_{\substack{e\in E \\ u\in H(e) \\ \wedge v\in T(e)}}\frac{1}{\delta_e}\mathcal{F}^\top_{u\trianglelefteq e}\mathcal{F}_{v\trianglelefteq e}\right), & u\neq v.\end{cases}$$

Only contribution in the undirected case

Net-flow behavior

Directed Sheaf Hypergraph Laplacian: Properties

Spectral Properties

- Always admits a spectral decomposition with real eigenvalues
- Is PSD, hence $\lambda_i \geq 0; \forall i$
- In its normalized form its spectrum is bounded: $\lambda_{max} = 1$

Generalization Properties

- Magnetic Laplacian for 2-uniform directed hypergraphs
- Generalized Directed Laplacian for q = $\frac{1}{4}$
- Sheaf Laplacian
- Undirected Hypergraph Laplacian[5]

[5] Zhou et al., Learning with Hypergraphs: Clustering, Classification, and Embedding. In NeurIPS 2006.

Directional Sheaf Hypergraph Networks

$$\mathbf{X}_{t+1} = \sigma\Big(\big(\mathbf{I}_{nd} - \mathbf{L}_N^{\mathcal{F}}\big)\left(\mathbf{I}_n \otimes \mathbf{W}_1\right)\mathbf{X}_t\mathbf{W}_2\Big) \in \mathbb{C}^{nd \times f}$$

The directionless restriction maps are predicted by using an MLP:

$$\mathcal{F}_{v \trianglelefteq e} = \Phi\big(\mathbf{x}_v \,\|\, \mathbf{x}_e\big) \in \mathbb{R}^{d \times d}$$

Directionality is encoded using the directional matrix

$$\mathcal{S}^{(q)}_{u \trianglelefteq e} = \begin{cases} 1 & \text{if } u \in H(e) \\ e^{-2\pi i q} & \text{if } u \in T(e) \\ 0 & \text{otherwise} \end{cases}$$

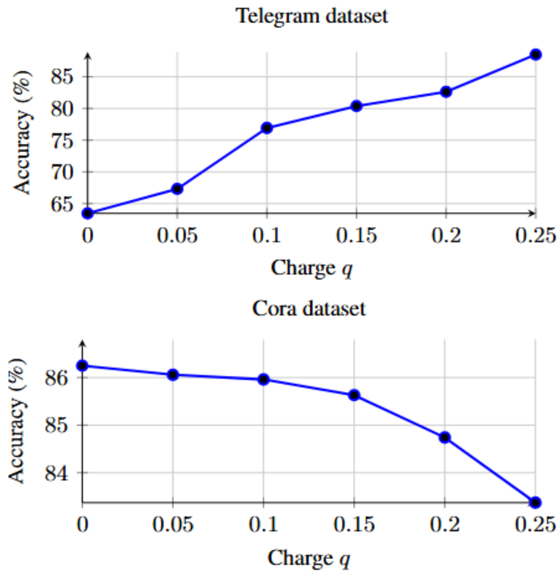
Results: Real-world Datasets

Table 1: Mean accuracy ± standard deviation on node classification datasets. For each dataset, the best result is shown in **bold**, and the second best is underlined.

	Roman-empire	Squirrel	email-EU	Telegram	Chameleon	email-Enron	Cora
HGNN	38.44 ± 0.44	35.47 ± 1.44	48.91 ± 3.11	51.73 ± 3.38	39.98 ± 2.28	52.85 ± 7.27	87.25 ± 1.01
HNHN	46.07 ± 1.22	35.62 ± 1.30	29.68 ± 1.68	38.22 ± 6.95	35.81 ± 3.23	18.64 ± 6.90	78.16 ± 0.98
UniGCNII	78.89 ± 0.51	38.28 ± 2.56	44.98 ± 2.69	51.73 ± 5.05	39.85 ± 3.19	47.43 ± 7.47	87.53 ± 1.06
LEGCN	65.60 ± 0.41	39.18 ± 1.54	32.91 ± 1.83	45.38 ± 4.23	39.29 ± 2.04	37.03 ± 7.16	74.96 ± 0.94
HyperND	68.31 ± 0.69	40.13 ± 1.85	32.79 ± 2.90	44.62 ± 5.49	44.95 ± 3.20	38.11 ± 7.69	78.48 ± 1.02
AllDeepSets	81.79 ± 0.72	40.69 ± 1.90	37.37 ± 6.29	49.19 ± 6.73	42.97 ± 3.60	37.29 ± 7.90	86.86 ± 0.85
AllSetTransformer	83.53 ± 0.64	40.53 ± 1.33	38.26 ± 3.57	66.92 ± 4.36	43.85 ± 5.42	63.78 ± 3.66	86.73 ± 1.13
ED-HNN	83.82 ± 0.31	39.85 ± 1.79	68.91 ± 4.00	60.38 ± 3.86	44.67 ± 2.33	51.35 ± 6.04	86.94 ± 1.25
SheafHyperGNN	74.50 ± 0.57	42.01 ± 1.11	52.78 ± 9.13	70.00 ± 5.32	41.06 ± 4.94	63.51 ± 5.95	87.15 ± 0.64
PhenomNN	71.22 ± 0.45	39.45 ± 2.19	37.69 ± 4.40	47.69 ± 6.59	43.62 ± 4.29	47.02 ± 6.75	**88.12 ± 0.86**
GeDi-HNN	$\underline{83.87 \pm 0.63}$	43.02 ± 3.00	52.31 ± 2.84	77.12 ± 4.82	39.29 ± 2.04	50.54 ± 5.80	85.16 ± 0.94
DHGNN	77.58 ± 0.54	39.85 ± 1.79	32.35 ± 2.93	79.62 ± 5.78	44.08 ± 4.11	42.16 ± 8.04	83.16 ± 1.33
DHGNN (w/ emb.)	22.50 ± 0.81	40.33 ± 1.42	55.10 ± 3.48	80.58 ± 3.89	40.85 ± 2.76	58.38 ± 7.57	73.12 ± 1.04
DSHN	OOM	$\underline{43.55 \pm 2.87}$	$\underline{78.62 \pm 2.50}$	**88.65 ± 5.54**	**47.02 ± 4.35**	$\underline{75.68 \pm 3.42}$	87.84 ± 0.90
DSHNLight	**89.24 ± 0.57**	**44.09 ± 2.36**	**82.67 ± 1.29**	$\underline{81.15 \pm 4.19}$	$\underline{46.50 \pm 4.09}$	**76.76 ± 2.48**	$\underline{88.02 \pm 1.11}$

Our models achieve consistent improvements on 6/7 datasets.

Results: Charge Parameter q



- In highly homophilic networks **direction is detrimental**
- In more heterophilic networks it **impacts positively**
- This is **in line with findings** in the directed graphs literature

Conclusions

1. Introduced Directed Hypergraph Cellular Sheaves and the Directed Sheaf Hypergraph Laplacian
2. Encodes hyperedge direction through a complex-valued inductive bias
3. Achieves consistent performance gains across benchmark datasets

Future works:

- Extend directional encoding through quaternion-valued Sheaf Neural Networks
- Make the charge parameter q learnable, enabling each layer to adapt its diffusion process dynamically

Check out our Github

Emanuele Mule[1], Stefano Fiorini, Antonio Purificato[1], Federico Siciliano[1], Stefano Coniglio[2], Fabrizio Silvestri[1]

SAPIENZA
UNIVERSITÀ DI ROMA

Sapienza University of Rome[1]

UNIVERSITÀ
DEGLI STUDI
DI BERGAMO

University of Bergamo[2]