27/03/2026

SHEAF NEURAL NETWORKS

INTRODUCTION

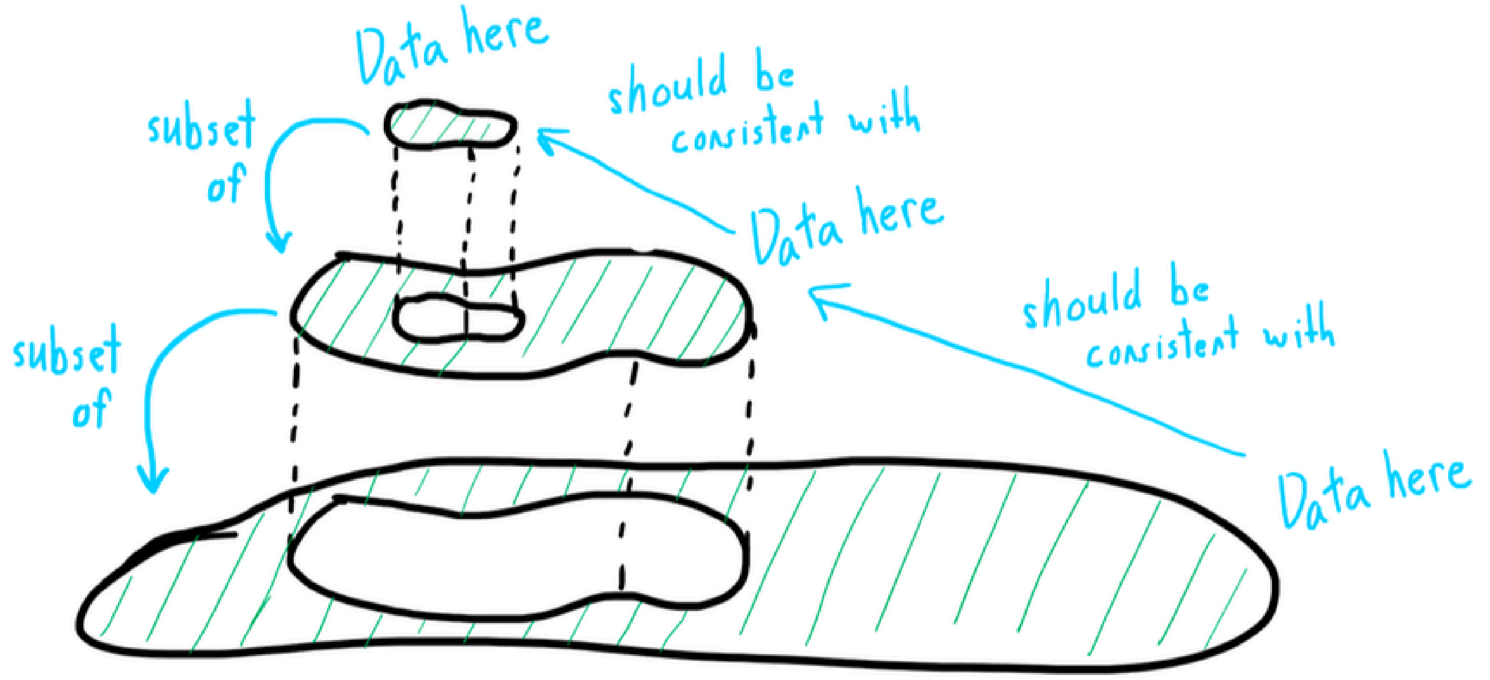
What is locally true everywhere, is not necessarily globally true.

Cauchy Problem

$$y^{(n)} = f(x, y, y', \ldots, y^{(n-1)})\ ,\quad y^{(i)}(x_0) = y_0^{(i)}$$

f continuous ⟶ local solutions

f Lipschitz ⟶ global solution

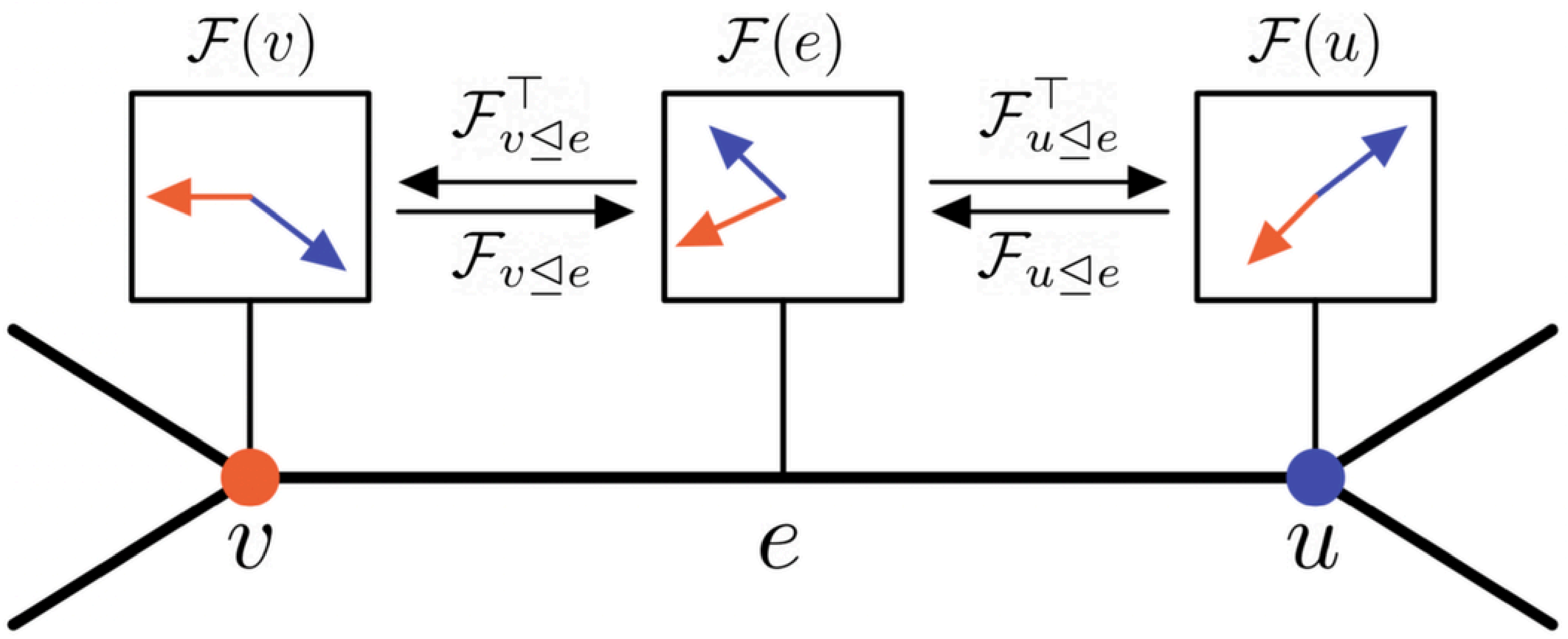


A sheaf enhances a mathematical object by attaching data to its pieces and organizing how this data behaves across the structure. [1]

Sheaf theory is meant to encode and study the passage from local to global.

SHEAVES

SHEAVES ON GRAPHS



DEFINITION.

A cellular sheaf $(G, \mathcal{F})$ on an undirected graph $G = (V, E)$ consists of:

1. a vector space $\mathcal{F}(v)$ for each $v \in V$
2. a vector space $\mathcal{F}(e)$ for each $e \in E$
3. a linear map $\mathcal{F}_{v \trianglelefteq e} : \mathcal{F}(v) \rightarrow \mathcal{F}(e)$

TERMINOLOGY

$\mathcal{F}(v)$

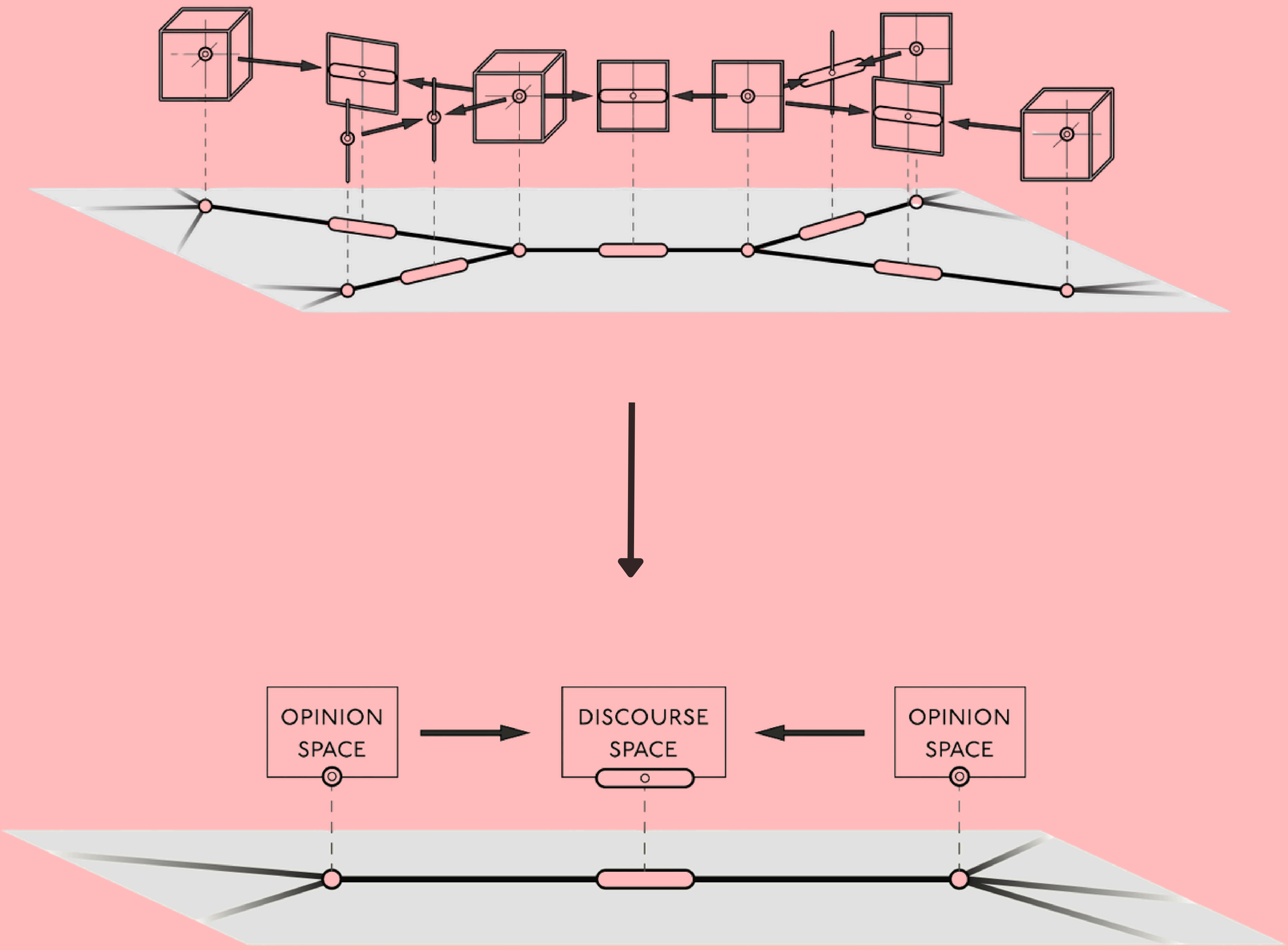
$\mathcal{F}(e)$ → Stalks

$\mathcal{F}_{v \trianglelefteq e} : \mathcal{F}(v) \rightarrow \mathcal{F}(e)$ → Restriction maps

The space formed by all the spaces associated with the nodes of the graph is called the space of 0-cochains $C^0(G; \mathcal{F}) := \bigoplus_{v \in V} \mathcal{F}(v)$, where $\bigoplus$ denotes the direct sum of vector spaces.

A particular important subspace of $C^0(G; \mathcal{F})$ is the space of global sections $H^0(G; \mathcal{F}) := \{\mathbf{x} \in C^0(G; \mathcal{F}) : \mathcal{F}_{v \trianglelefteq e}\mathbf{x}_v = \mathcal{F}_{u \trianglelefteq e}\mathbf{x}_u\}$.

OPINION DYNAMICS



Opinion dynamics provides a nice mental picture of cellular sheaves.

"Stalks over vertices are individual opinion spaces, stalks over edges are discourse spaces, and restriction maps are expressions of opinions on the topics of discourse, formulated linearly from basis opinions." [2]

OPINION DYNAMICS ON DISCOURSE SHEAVES*

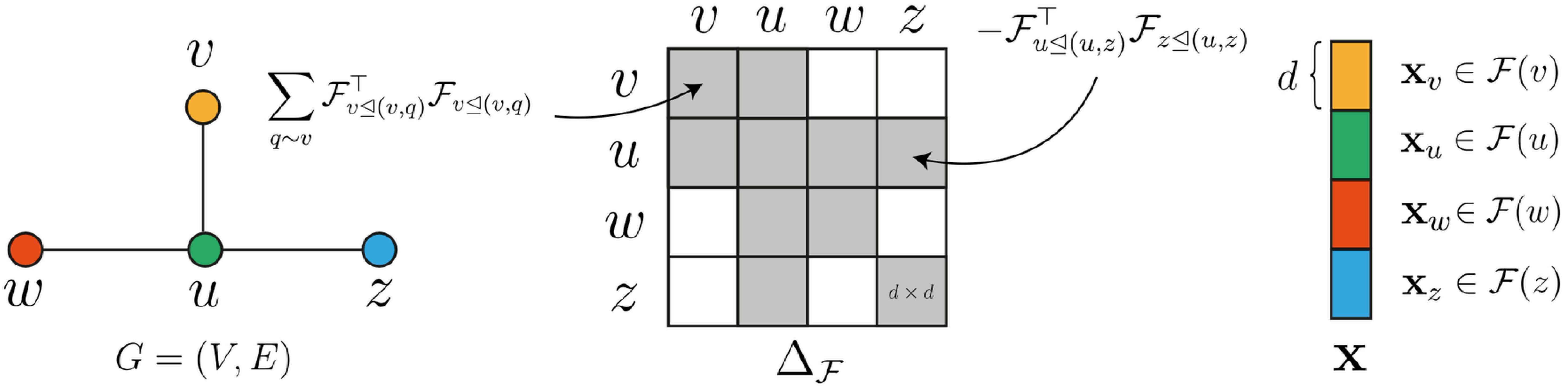
JAKOB HANSEN† AND ROBERT GHRIST‡

Abstract. We introduce a novel class of Laplacians and diffusion dynamics on *discourse sheaves* as a model for network dynamics, with application to opinion dynamics on social networks. These *sheaves* are algebraic data structures tethered to a network (or more general space) that can represent various modes of communication, including selective opinion modulation and lying. After introducing the sheaf model, we develop a sheaf Laplacian in this context and show how to evolve both opinions and communications with diffusion dynamics over the network. Issues of controllability, reachability, bounded confidence, and harmonic extension are addressed using this framework.

SHEAF LAPLACIAN

DEFINITION. The Laplacian of a sheaf $(G, \mathcal{F})$ is a linear map $L_\mathcal{F} : C^0(G,\mathcal{F}) \to C^0(G,\mathcal{F})$ defined node-wise as

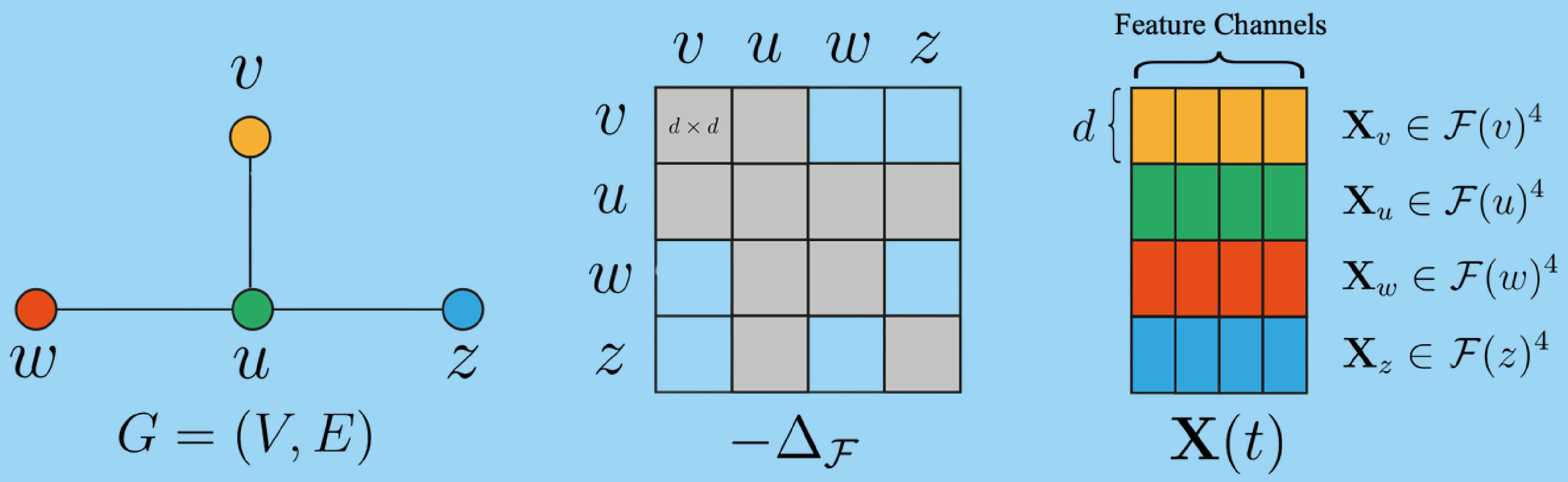
$$L_\mathcal{F}(\mathbf{x})_v := \sum_{v,u \trianglelefteq e} \mathcal{F}^\top_{v \trianglelefteq e}(\mathcal{F}_{v \trianglelefteq e}\mathbf{x}_v - \mathcal{F}_{u \trianglelefteq e}\mathbf{x}_u)$$



SHEAF DIFFUSION

We now consider the sheaf diffusion process governed by the PDE

$$\mathbf{X}(0) = \mathbf{X}, \quad \dot{\mathbf{X}}(t) = -\Delta_{\mathcal{F}}\mathbf{X}(t)$$



THEOREM (HODGE THEOREM)

As $t \to \infty$, the features converge to the projection of $\mathbf{X}(0)$ into $\ker(\Delta_{\mathcal{F}}) \cong H^0(G; \mathcal{F})$.

Let G be a graph with self-loops, degree matrix $\mathbf{D}$, adjacency matrix $\mathbf{A}$, normalised Laplacian $\Delta_0 := \mathbf{I} - \mathbf{D}^{-1/2}\mathbf{A}\mathbf{D}^{-1/2}$ and node features $\mathbf{X} \in \mathbb{R}^{n \times d}$.

Heat diffusion is described as:

$$\dot{\mathbf{X}}(t) = -\Delta_0\mathbf{X}(t) \rightsquigarrow \mathbf{X}(t+1) = \mathbf{X}(t) - \Delta_0\mathbf{X}(t) = (\mathbf{I} - \Delta_0)\mathbf{X}(t)$$

Heat diffusion converges to the kernel of the Laplacian, which for a connected graph is the space of constant signals.

HEAT KERNEL ACTS AS A LOW PASS FILTER

HEAT DIFFUSION ON GRAPHS

Standard diffusion converges to the space of constant signals, while sheaf diffusion converges to the space of global sections, which can encode non-trivial structure.

GRAPH CONVOLUTIONAL NETWORK

$\mathbf{X} \in \mathbb{R}^{\mathbf{N \times f}}$ Node Features

$\mathbf{A} \in \mathbb{R}^{N \times N}$ Adjacency Matrix

$$\mathbf{X}_{t+1} = \sigma\left(\mathbf{D}^{-1/2}\mathbf{A}\mathbf{D}^{-1/2}\mathbf{X}_t\mathbf{W}_t\right) = \sigma\left(\left(\mathbf{I} - \mathbf{\Delta}_0\right)\mathbf{X}_t\mathbf{W}_t\right)$$

Oversmoothing Problem

Heterophilic Problem

Any node classification task can be reduced to performing diffusion with the right sheaf.

$$\mathbf{X}_{t+1} = \mathbf{X}_t - \sigma\left(\Delta_{\mathcal{F}(t)}\left(\mathbf{I} \otimes \mathbf{W}_1^t\right)\mathbf{X}_t\mathbf{W}_2^t\right)$$

$\mathbf{x} \in \mathbb{R}^{\mathbf{f}}$ → $\mathbf{x} \in \mathbb{R}^{(\mathbf{d} \times \mathbf{f}')}$

The node signal is lifted to a higher-dimensional geometric space, where each node carries a d-dimensional fiber (stalk), enabling richer interactions.

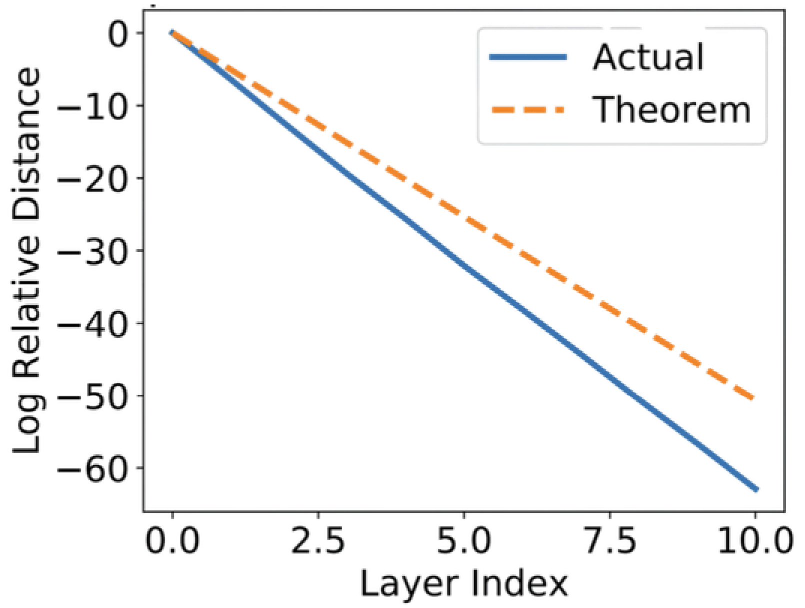
$\mathbf{X} \in \mathbb{R}^{\mathbf{N \times f}}$ Raw Node Features

$\mathbf{A} \in \mathbb{R}^{N \times N}$ Adjacency matrix

NEURAL SHEAF DIFFUSION

GRAPH CONVOLUTIONAL NETWORK

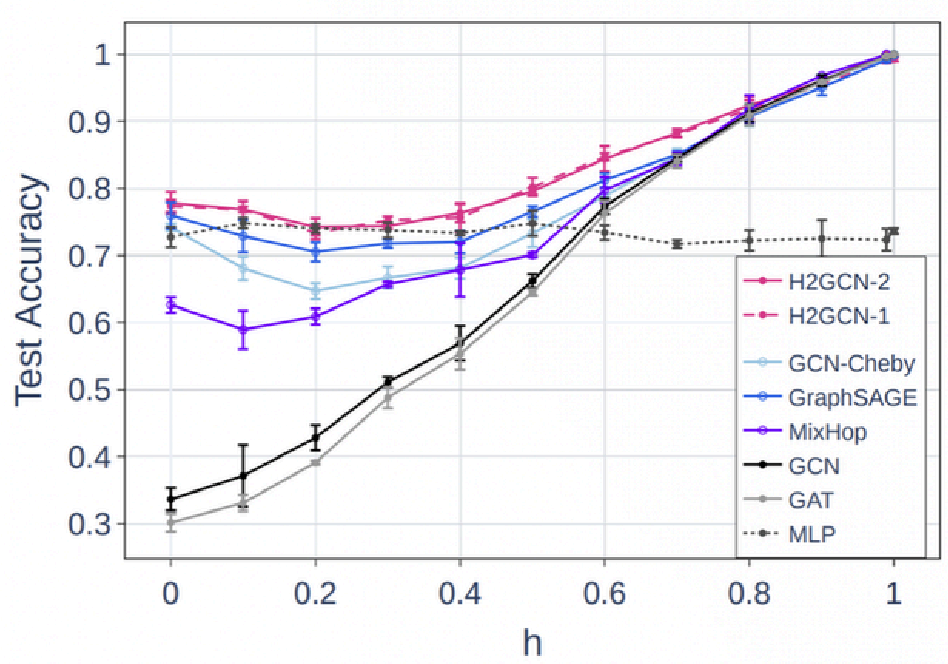
Oversmoothing Problem



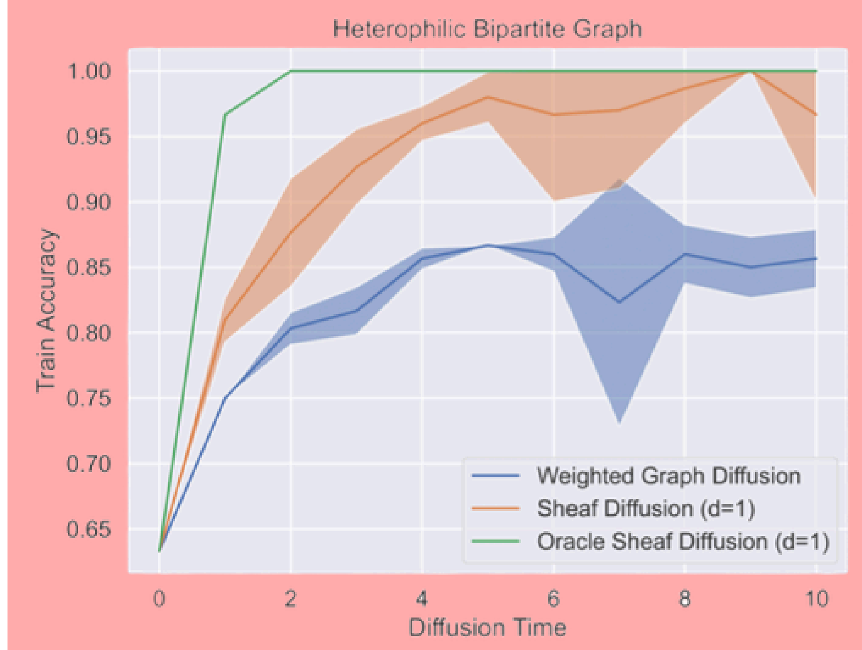
With more layers, GCN approaches a "smooth" subspace where all the node features are constant. [3]

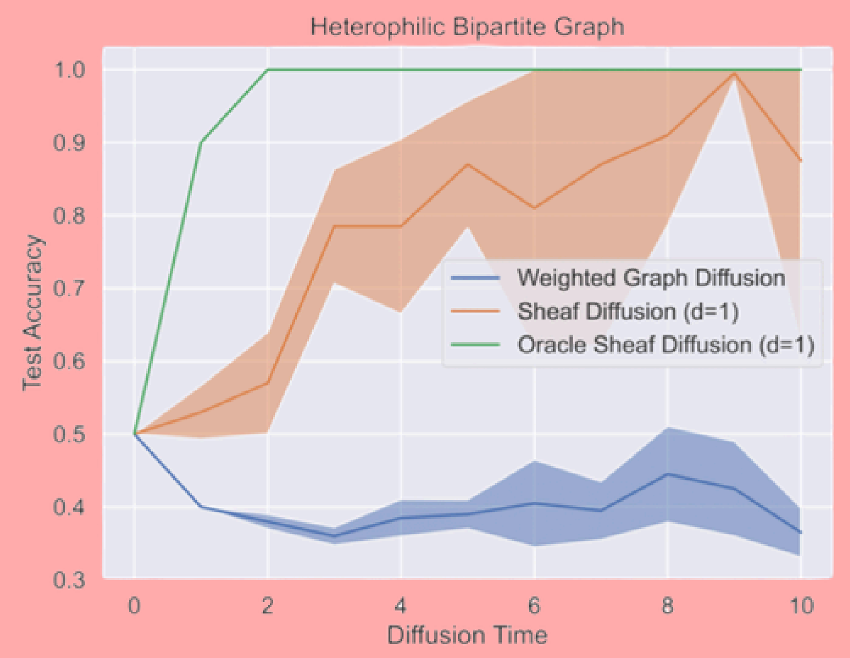
Heterophilic Problem

The performance of GNNs is strongly correlated to the homophily level of a graph. [4]



Different classes of sheaves induce diffusion processes with different capabilities. Furthermore, any node-classification problem can be reduced to performing diffusion with the right sheaf. [5]



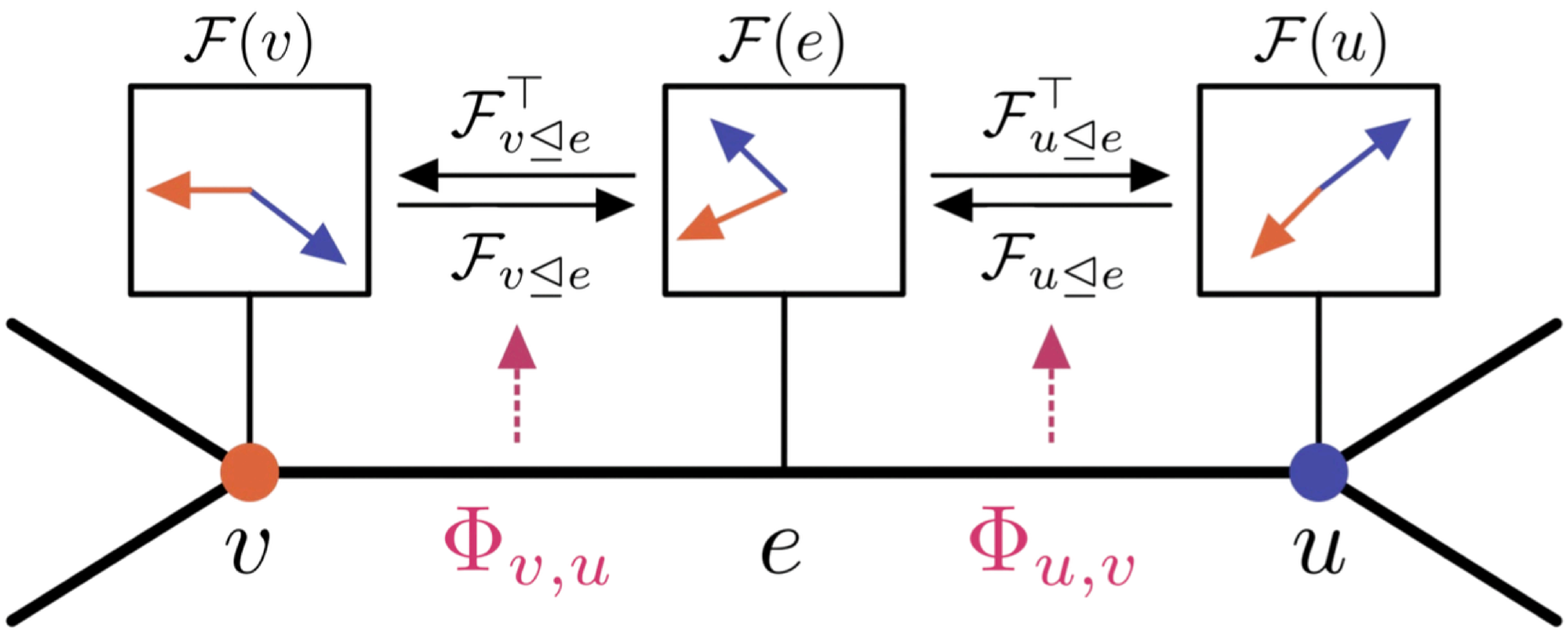


NEURAL SHEAF DIFFUSION

LEARNING SHEAVES

Each $d \times d$ matrix $\mathcal{F}_{v \trianglelefteq e}$ is learned via a parametric function $\Phi : \mathbb{R}^{d\times 2} \to \mathbb{R}^{d\times d}$:

$$\mathcal{F}_{v \trianglelefteq e:=(v,u)} = \Phi(\mathbf{x}_v, \mathbf{x}_u)$$



REAL-WORLD EVALUATION

	Texas	Wisconsin	Film	Squirrel	Chameleon	Cornell	Citeseer	Pubmed	Cora
Hom level	**0.11**	**0.21**	**0.22**	**0.22**	**0.23**	**0.30**	**0.74**	**0.80**	**0.81**
#Nodes	183	251	7,600	5,201	2,277	183	3,327	18,717	2,708
#Edges	295	466	26,752	198,493	31,421	280	4,676	44,327	5,278
#Classes	5	5	5	5	5	5	7	3	6
Diag-NSD	**85.67**±6.95	**88.63**±2.75	**37.79**±1.01	**54.78**±1.81	**68.68**±1.73	**86.49**±7.35	**77.14**±1.85	89.42±0.43	87.14±1.06
O(d)-NSD	**85.95**±5.51	**89.41**±4.74	**37.81**±1.15	**56.34**±1.32	**68.04**±1.58	**84.86**±4.71	76.70±1.57	**89.49**±0.40	86.90±1.13
Gen-NSD	82.97±5.13	**89.21**±3.84	**37.80**±1.22	53.17±1.31	67.93±1.58	**85.68**±6.51	76.32±1.65	89.33±0.35	87.30±1.15
GGCN	**84.86**±4.55	86.86±3.29	37.54±1.56	**55.17**±1.58	**71.14**±1.84	**85.68**±6.63	**77.14**±1.45	89.15±0.37	**87.95**±1.05
H2GCN	**84.86**±7.23	87.65±4.98	35.70±1.00	36.48±1.86	60.11±2.15	82.70±5.28	77.11±1.57	**89.49**±0.38	**87.87**±1.20
GPRGNN	78.38±4.36	82.94±4.21	34.63±1.22	31.61±1.24	46.58±1.71	80.27±8.11	77.13±1.67	87.54±0.38	**87.95**±1.18
FAGCN	82.43±6.89	82.94±7.95	34.87±1.25	42.59±0.79	55.22±3.19	79.19±9.79	N/A	N/A	N/A
MixHop	77.84±7.73	75.88±4.90	32.22±2.34	43.80±1.48	60.50±2.53	73.51±6.34	76.26±1.33	85.31±0.61	87.61±0.85
GCNII	77.57±3.83	80.39±3.40	37.44±1.30	38.47±1.58	63.86±3.04	77.86±3.79	**77.33**±1.48	**90.15**±0.43	**88.37**±1.25
Geom-GCN	66.76±2.72	64.51±3.66	31.59±1.15	38.15±0.92	60.00±2.81	60.54±3.67	**78.02**±1.15	**89.95**±0.47	85.35±1.57
PairNorm	60.27±4.34	48.43±6.14	27.40±1.24	50.44±2.04	62.74±2.82	58.92±3.15	73.59±1.47	87.53±0.44	85.79±1.01
GraphSAGE	82.43±6.14	81.18±5.56	34.23±0.99	41.61±0.74	58.73±1.68	75.95±5.01	76.04±1.30	88.45±0.50	86.90±1.04
GCN	55.14±5.16	51.76±3.06	27.32±1.10	53.43±2.01	64.82±2.24	60.54±5.30	76.50±1.36	88.42±0.50	86.98±1.27
GAT	52.16±6.63	49.41±4.09	27.44±0.89	40.72±1.55	60.26±2.50	61.89±5.05	76.55±1.23	87.30±1.10	86.33±0.48
MLP	80.81±4.75	85.29±3.31	36.53±0.70	28.77±1.56	46.21±2.99	81.89±6.40	74.02±1.90	75.69±2.00	87.16±0.37

REFERENCES

[1] Agrios, Mark. "A very elementary introduction to sheaves." arXiv preprint arXiv:2202.01379 (2022).

[2] Hansen, Jakob, and Robert Ghrist. "Opinion dynamics on discourse sheaves." SIAM Journal on Applied Mathematics 81.5 (2021): 2033–2060.

[3] Oono, Kenta, and Taiji Suzuki. "Graph neural networks exponentially lose expressive power for node classification." arXiv preprint arXiv:1905.10947 (2019).

[4] Zhu, Jiong, et al. "Beyond homophily in graph neural networks: Current limitations and effective designs." Advances in neural information processing systems 33 (2020): 7793–7804.

[5] Bodnar, Cristian, et al. "Neural sheaf diffusion: A topological perspective on heterophily and oversmoothing in gnns." Advances in Neural Information Processing Systems 35 (2022): 18527–18541.

THANK YOU!