Appendix A. Theoretical Proof

Theorem 1. *Suppose a set of base linear (or quadratic) filter functions* $F_i(\lambda)$ *with parameters* α_i*, where* $\alpha \in (0,1)$*. Let*

$$F(\lambda) = \sum_{i=1}^{K} p_i F_i(\lambda) = \hat{p} F(\lambda)$$

be a linear combination of these base filter functions, where $p_i > 0$*. Then,* $F(\lambda)$ *is a filter of the same type with parameters:*$\hat{p} = \sum_{i=1}^{K} p_i$ *and* $\hat{\alpha} = \frac{\sum_{i=1}^{K} p_i \alpha_i}{\sum_{i=1}^{K} p_i}$.

Proof. Assuming the base filter functions are low-pass filter functions.

$$\begin{aligned} F(\lambda) &= \sum p_i F_i(\lambda) && \text{(A.1)} \\ &= \sum p_i (1 - \alpha_i \lambda) && \text{(A.2)} \\ &= \sum p_i \times \left(1 - \frac{\sum p_i \alpha_i}{\sum p_i} \lambda\right). && \text{(A.3)} \end{aligned}$$

Let $\hat{p} = \sum p_i$, and $\hat{\alpha} = \sum p_i \alpha_i / \sum p_i$,

$$F(\lambda) = \hat{p}(1 - \hat{\alpha}\lambda). \tag{A.4}$$

We now prove $\hat{p} > 0$, and $\hat{\alpha} \in (0,1)$.

$$\hat{p} = \sum p_i > \min(p_i) > 0. \tag{A.5}$$

As $p_i > 0$,

$$\hat{a} = \frac{\sum p_i \alpha_i}{\sum p_i} \geq \frac{\sum p_i \min(\alpha_i)}{\sum p_i} = \min(\alpha_i) > 0. \tag{A.6}$$

$$\hat{a} = \frac{\sum p_i \alpha_i}{\sum p_i} \leq \frac{\sum p_i \max(\alpha_i)}{\sum p_i} = \max(\alpha_i) < 1. \tag{A.7}$$

Therefore, $0 < \hat{a} < 1$. With $p_i > 0$ and $0 < \hat{\alpha} < 1$, Equation A.4 fulfills the definition of our proposed low-pass linear filter. The demonstration is equally applicable to mid-pass and high-pass filter functions. □

Appendix B. Algorithm

Algorithm 1 ACF-HNN (Adaptive Cross-Frequency Hypergraph Neural Network)

Require: Hypergraph structure $\boldsymbol{H}, \boldsymbol{W}, \boldsymbol{D}_v, \boldsymbol{D}_e$, node features $\boldsymbol{X}$, layers L, basis filters K

Ensure: Node representations $\boldsymbol{X}^{(L)}$

Initialize: $\boldsymbol{X}^{(0)} \leftarrow \boldsymbol{X}$

Compute normalized Laplacian: $\boldsymbol{\Theta} \leftarrow \boldsymbol{D}_v^{-1/2} \boldsymbol{H} \boldsymbol{W} \boldsymbol{D}_e^{-1} \boldsymbol{H}^\top \boldsymbol{D}_v^{-1/2}$

for $l = 1$ **to** L **do**

    Multi-Frequency Filtering:

    Define filter operators $\boldsymbol{\Psi}_{\text{low}} = \boldsymbol{\Theta}$, $\boldsymbol{\Psi}_{\text{high}} = -\boldsymbol{\Theta}$, $\boldsymbol{\Psi}_{\text{mid}} = \boldsymbol{\Theta}^2$

    Define bias terms $\boldsymbol{\Gamma}_{\text{low}} = \boldsymbol{\Gamma}_{\text{high}} = \boldsymbol{I}$, $\boldsymbol{\Gamma}_{\text{mid}} = -\boldsymbol{I}$

    for $f \in \{\text{low}, \text{mid}, \text{high}\}$ **do**

        Initialize basis parameters $\{\alpha_{(f,i)}, p_{(f,i)}\}_{i=1}^K$ (learnable)

        Aggregate parameters: $\hat{\alpha}_f \leftarrow \frac{\sum_{i=1}^K p_{(f,i)} \alpha_{(f,i)}}{\sum_{i=1}^K p_{(f,i)}}$, $\hat{p}_f \leftarrow \sum_{i=1}^K p_{(f,i)}$

        Construct spatial kernel: $\boldsymbol{C}_f \leftarrow \hat{p}_f \left[(1 - \hat{\alpha}_f)\boldsymbol{\Gamma}_f + \hat{\alpha}_f \boldsymbol{\Psi}_f\right]$

        Feature propagation: $\boldsymbol{X}_f^{(l)} \leftarrow \boldsymbol{C}_f \boldsymbol{X}^{(l-1)} \boldsymbol{T}_f^{(l-1)}$ (parameter $\boldsymbol{T}_f^{(l-1)}$)

    end for

    Cross-Frequency Fusion: $\boldsymbol{X}^{(l)} \leftarrow \sigma\left(\boldsymbol{X}_{\text{low}}^{(l)} + \boldsymbol{X}_{\text{mid}}^{(l)} + \boldsymbol{X}_{\text{high}}^{(l)} + \boldsymbol{X}^{(l-1)}\right)$

end for

Appendix C. Experiments Details

Appendix C.1. Visual datasets details

The ModelNet40 dataset consists of 12,311 objects from 40 popular categories, and the same training/testing split is applied as introduced in (Wu et al. 2015), where 9,843 objects are used for training and 2,468 objects are used for testing. The NTU dataset is composed of 2,012 3D shapes from 67 categories, including car, chair, chess, chip, clock, cup, door, frame, pen, plant leaf and so on.In the NTU dataset, 80% data are used for training and the other 20% data are used for testing. The dataset statistics are summarized in Table C.1.

Table C.1: The detailed information of the ModelNet40 and the NTU datasets.

Dataset	**ModelNet40**	**NTU**
Objects	12311	2012
MVCNN Feature	4096	4096
GVCNN Feature	2048	2048
Training node	9843	1639
Testing node	2468	373
Classes	40	67

Appendix C.2. Hyperparameter Settings

The hyperparameter settings are summarized in Table C.2.

Table C.2: Hyperparameter settings for all experiments. L is the number of layers; hidden is the dimension of hidden features; init lr is the initial learning rate, weight decay is the weight decay rate, K is the number of base filter functions.

Dataset	L	**hidden**	**dropout**	K	**init lr**	**weight decay**
Cora	3	128	0.5	1	1e-3	5e-4
Citeseer	4	512	0.5	4	1e-3	2e-3
Pubmed	3	512	0.5	8	1e-3	5e-4
Cora-CA	3	128	0.5	1	1e-3	5e-4
DBLP-CA	4	256	0.5	2	1e-3	5e-4
Congress	4	128	0.5	1	1e-3	5e-4
Senate	2	16	0.5	16	1e-3	5e-4
Walmart	3	256	0.5	2	1e-3	5e-4
House	4	256	0.5	1	1e-3	9e-4