DIRAC CONVOLUTIONAL NEURAL NETWORKS FOR SIMPLICIAL COMPLEXES

Anonymous authors
Paper under double-blind review

ABSTRACT

Higher-order systems capture interactions that go beyond the pairwise relationships represented by graphs and arise naturally in models of biological and physical systems. Simplicial complexes provide a common framework for representing such higher-order interactions, motivating the development of neural architectures that operate on them. In this work, we introduce *Dirac Convolutional Networks* (DCNs), a family of simplicial neural networks (SNNs) based on *Dirac convolutions* we introduce. Our construction is derived from a spectral filter formulation for multichannel, graded signals on simplicial complexes. We instantiate our framework with *normalized root-to-leaf Dirac operator* and propose *Root-to-Leaf DCNs* (R2L-DCNs) that inherit the benefits of normalized operators. First, normalization fixes the spectral scale of the Dirac operator and Hodge Laplacians, allowing filters to transfer across simplicial complexes. Second, it improves the stability of convolutional SNNs, allowing deep neural architectures. We establish this by proving that continuous, linear versions of these networks can preserve both the Dirichlet energy and the norm with right weights, and that their Euler discretizations are more stable when normalized operators are used. Finally, we empirically validate these properties across multiple benchmarks, showing that R2L-DCNs outperform existing SNN architectures on trajectory prediction and mesh classification benchmarks, and can maintain performance as network depth increases.

1 INTRODUCTION

Graph representation learning (GRL) (Hamilton, 2020) has proven to be useful in domains where signals naturally lie on the nodes of a graph, extending methods of deep learning (Goodfellow et al., 2016) and graph signal processing (GSP) (Leus et al., 2023; Dong et al., 2020) to graph structured data. More recently, the need for higher-order models that go beyond pairwise interactions has been emphasized, to accurately model certain systems in biology, physics and social sciences (Bianconi, 2021a; Battiston et al., 2020; Bick et al., 2023). This gave rise to *topological deep learning* (Hajij et al., 2023b; Papamarkou et al., 2024), generalizing methods of GRL and GSP to higher-order topological domains, specifically *simplicial complexes* (SCs), which allows for signals that represent k-fold interactions between members of a system that are not supported by graphs. Different perspectives on GRL have come up throughout the years focusing on graph neural networks, such as spectral GNNs (Bruna et al., 2014; Defferrard et al., 2017; Kipf & Welling, 2017) inspired by convolutions on graphs (Gama et al., 2020), message-passing paradigm that arises from the permutation equivariance (Gilmer et al., 2017), or dynamical perspectives that view GNNs as a discretization or solution of some parametrized diffusion equation (Chamberlain et al., 2021; Giovanni et al., 2023; Han et al., 2023); and some of these approaches have been generalized to SCs (Ebli et al., 2020; Yang et al., 2025; Bodnar et al., 2021; Einizade et al., 2025). Similarly, bottlenecks in GNNs, such as oversmoothing (Li et al., 2018; Rusch et al., 2023) and oversquashing (Alon & Yahav, 2021; Topping et al., 2022), have also been carried to and analyzed in simplicial neural networks (SNNs) (Yang et al., 2025; Einizade et al., 2025; Taha et al., 2025).

Unlike standard graph architectures, existing SNN architectures are based on unnormalized notions of higher-order Laplacians (Lim, 2019), impacting their stability and consistency across SCs and hindering the construction of deeper SNNs. Moreover, the asymptotic behaviour of SNNs remains poorly understood and has received less attention compared to oversmoothing and oversquashing in graphs. To address these limitations, we introduce *Dirac Convolutional Networks* (DCNs) a

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

generalized family of SNNs derived from Dirac convolutions for graded, multi-channel signals on SCs, which recovers existing convolutional SNNs as special cases. Furthermore, we leverage, for the first time, the *normalized root-to-leaf Dirac operator* and *Hodge Laplacians* introduced in (Viganò et al., 2026) to develop SNNs that are more stable and transferable across SCs. We rigorously study the stability of DCNs, proving that continuous-time DCNs preserve the Dirichlet energy and the norm, while normalization yields better stability bounds for their forward-Euler discretizations by bounding the spectrum.

Our contributions are:

- We introduce *Dirac convolutions* (DCs), a generalized framework for spectral filtering on SCs that account for inter-rank interactions.
- We formulate DCNs, generalizing existing convolutional SNNs into a common framework.
- We instantiate DCNs with root-to-leaf normalization and Chebyshev polynomials, yielding *Root-to-Leaf DCNs* (R2L-DCNs) and demonstrate its significance for filter transferability.
- We rigorously characterize the stability of linear continuous-time and forward-Euler DCNs, proving Dirichlet energy and norm preservation for continuous-time DCNs with skew-symmetric weights, and derive eigenvalue-dependent stability bounds, showing benefits of normalization.

Our experiments demonstrate that R2L-DCNs consistently outperform state of the art SNNs in trajectory prediction and mesh-classification task, while achieving comparable or superior performance in node and graph classification. We will make our code publicly available upon publication.

2 Background

We provide the necessary background for simplicial complexes (Nanda, 2021), the Dirac operator (Bianconi, 2021a; Calmon et al., 2023), and root-to-leaf normalization (Viganò et al., 2026). For further details, see Appendix A.

Simplicial complexes. Let $\mathcal{X}_0 = \{v_0, \dots, v_n\}$ be a finite vertex set. A *simplicial complex* $\mathcal{X}$ is a collection of nonempty subsets of $\mathcal{X}_0$ that contains all singletons and is closed under taking nonempty subsets (Nanda, 2021). An element $\sigma \in \mathcal{X}$ is a *simplex* of *dimension* $\dim \sigma := |\sigma| - 1$. Writing $\mathcal{X}_k$ for the set of k-simplices, we have $\mathcal{X} = \mathcal{X}_0 \sqcup \cdots \sqcup \mathcal{X}_{\dim \mathcal{X}}$, where $\dim \mathcal{X}$ is the largest dimension of a simplex. If $\sigma \in \mathcal{X}_k$ and $\sigma' \subset \sigma$ with $\sigma' \in \mathcal{X}_{k-1}$, then σ' is a *face* of σ and σ is a *coface* of σ'. To account for orientations, signals on $\mathcal{X}_k$ are modeled as *k-cochains*, which are alternating functions on oriented k-simplices by fixing an ordering of the vertices. Fixing an orientation identifies the *cochain space* C_k with $\mathbb{R}^{|\mathcal{X}_k|}$. The *boundary operator* $\mathbf{B}_k : C_k \to C_{k-1}$ is the signed incidence matrix between k-simplices and their $(k-1)$-faces, with signs determined by the chosen orientations. Its transpose $\mathbf{B}_k^\top$ is the corresponding *coboundary operator*. (Co)boundary operators satisfy the relation $\mathbf{B}_k \mathbf{B}_{k+1} = \mathbf{0}$, which is crucial for studying algebraic and spectral properties of SCs. The *down* and *up Hodge Laplacians* at rank k are respectively:

$$\mathbf{L}_{k,d} = \mathbf{B}_k^\top \mathbf{B}_k, \qquad \mathbf{L}_{k,u} = \mathbf{B}_{k+1} \mathbf{B}_{k+1}^\top, \tag{1}$$

and their sum $\mathbf{L}_k = \mathbf{L}_{k,d} + \mathbf{L}_{k,u}$ is the k-th *Hodge Laplacian*, acting on rank-k signals. For $k = 0$, $\mathbf{L}_0$ reduces to the graph Laplacian (Godsil & Royle, 2001), and for higher $k \geq 1$ the Laplacians encode information about the interactions of adjacent ranks and the topology of $\mathcal{X}_k$ (Lim, 2019).

Dirac operator. The *Dirac operator* acts on all cochains at once, $\mathbf{D} : \bigoplus_k C_k \to \bigoplus_k C_k$, by applying all boundary and coboundary maps simultaneously:

$$\mathbf{D} = \begin{pmatrix} 0 & \mathbf{B}_1 & 0 & \cdots \\ \mathbf{B}_1^\top & 0 & \mathbf{B}_2 & \cdots \\ 0 & \mathbf{B}_2^\top & 0 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \quad \mathbf{D}_{[1]} = \begin{pmatrix} 0 & \mathbf{B}_1 & 0 & \cdots \\ \mathbf{B}_1^\top & 0 & 0 & \cdots \\ 0 & 0 & 0 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \quad \mathbf{D}_{[2]} = \begin{pmatrix} 0 & 0 & 0 & \cdots \\ 0 & 0 & \mathbf{B}_2 & \cdots \\ 0 & \mathbf{B}_2^\top & 0 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}. \tag{2}$$

It satisfies $\mathbf{D}^2 = \mathbf{L}$, the block-diagonal matrix with blocks $\mathbf{L}_0, \dots, \mathbf{L}_{\dim \mathcal{X}}$, and $\mathbf{D}$ can be read as a signed adjacency matrix of the incidence graph of all simplices. It also decomposes as $\mathbf{D} = \sum_{k=1}^{\dim \mathcal{X}} \mathbf{D}_{[k]}$, where $\mathbf{D}_{[k]}$ is the operator on $\bigoplus_k C_k$ whose only nonzero blocks are $\mathbf{B}_k$ and $\mathbf{B}_k^\top$, i.e. the Dirac operator of the single pair $C_{k-1} \oplus C_k$, which all satisfy $\mathbf{D}_{[i]} \mathbf{D}_{[j]} = 0$ for $i \neq j$ (Calmon et al., 2023). This shows that the images (hence eigenvectors) of all distinct $\mathbf{D}_{[i]}$ are orthogonal, allowing independent spectral analysis. Each $\mathbf{D}_{[i]}$ relates to the Laplacians via $\mathbf{D}_{[i]}^2 = \mathbf{L}_{i-1,u} \oplus \mathbf{L}_{i,d}$.

Root-to-Leaf normalization. Root-to-leaf (R2L) normalization in (Viganò et al., 2026) is derived from a random walk on the Hasse diagram of the simplicial complex $\mathcal{X}$ ordered by inclusion (Figure 1). Its maximal elements, i.e. simplices with no cofaces, are called *leaves* and minimal elements, i.e. vertices $\sigma \in \mathcal{X}_0$, are called *roots*, with the assumption that the Hasse diagram is connected. An *increasing* (resp. *decreasing*) *path* follows incidences, raising (resp. lowering) the rank by exactly 1 at each step. For $\sigma \in \mathcal{X}$, let L_σ be the set of increasing paths from σ to a leaf and R_σ the set of decreasing paths from σ to a root, counting the length-zero path, so that $|L_\sigma| = 1$ when σ is itself a leaf and $|R_\sigma| = 1$ when it is a root. Set $h_\sigma := |L_\sigma|/|R_\sigma|$, and for each rank k collect these into the diagonal weight matrix $\mathbf{W}_k := \mathrm{diag}\big((h_\sigma)_{\sigma\in\mathcal{X}_k}\big)$, with a fixed ordering of $\mathcal{X}_k$. The *normalized boundary operator* is

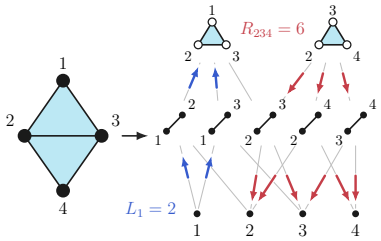


Figure 1: Example of root and leaf paths.

$$\widetilde{\mathbf{B}}_k := \mathbf{W}_{k-1}^{-1/2}\,\mathbf{B}_k\,\mathbf{W}_k^{1/2}, \tag{3}$$

with transpose $\widetilde{\mathbf{B}}_k^\top$ the *normalized coboundary operator*. The opposite exponents on the two sides make the weights cancel in $\widetilde{\mathbf{B}}_k\widetilde{\mathbf{B}}_{k+1} = \mathbf{W}_{k-1}^{-1/2}\mathbf{B}_k\mathbf{B}_{k+1}\mathbf{W}_{k+1}^{1/2} = 0$, so the normalized operators still form a chain complex. Substituting them into the definitions above yields the *normalized Dirac operator* $\widetilde{\mathbf{D}}$ and the *normalized down, up, and Hodge Laplacians* $\widetilde{\mathbf{L}}_{k,d}, \widetilde{\mathbf{L}}_{k,u}, \widetilde{\mathbf{L}}_k$, which satisfy $\widetilde{\mathbf{D}}^2 = \widetilde{\mathbf{L}}$ as before. The normalization bounds all of these spectra uniformly: $\mathrm{spec}(\widetilde{\mathbf{D}}) \subseteq [-1, 1]$ and hence $\mathrm{spec}(\widetilde{\mathbf{L}}_{k,d}), \mathrm{spec}(\widetilde{\mathbf{L}}_{k,u}), \mathrm{spec}(\widetilde{\mathbf{L}}_k) \subseteq [0, 1]$. Different normalization methods for Hodge Laplacians have been discussed for SCs (Horak & Jost, 2011; Schaub et al., 2020; Jost & Zhang, 2024), and we choose R2L as it works on all up and down Laplacians of all ranks, and specifically gives normalized (co)boundary operators.

Let $\mathcal{G} = (V, E)$ be a graph with adjacency matrix $\mathbf{A}$ and degree matrix $\mathbf{K} := \mathrm{diag}(\mathbf{A1})$. Its graph Laplacian is $\mathbf{\Delta} := \mathbf{K} - \mathbf{A}$, with normalized variant $\widetilde{\mathbf{\Delta}} := \mathbf{K}^{-1/2}\mathbf{\Delta}\mathbf{K}^{-1/2}$ (Chung, 1997). Viewing $\mathcal{G}$ as a simplicial complex of dimension 1, every vertex is a root with $h_v = \deg(v)$ and every edge is a leaf with $h_e = 1/2$, giving $\mathbf{W}_0 = \mathbf{K}$ and $\mathbf{W}_1 = \frac{1}{2}\mathbf{I}$. Consequently, $\widetilde{\mathbf{L}}_0 = \widetilde{\mathbf{L}}_{0,u} = \frac{1}{2}\widetilde{\mathbf{\Delta}}$, where the factor $\frac{1}{2}$ arises because each edge has two vertices below it. Thus, root-to-leaf normalization reduces to degree-based normalization on graphs.

3 Dirac Convolutional Networks

We introduce *Dirac convolutions* (DCs) for multi-channel, graded signals on SCs, and develop *Dirac Convolutional Networks* (DCNs) as a general framework for convolutional SNNs with inter-rank coupling. We instantiate DCNs with the normalized R2L Dirac operator and truncated Chebyshev filters, and demonstrate the resulting filter transferability across a family of SCs. Proofs of the propositions in this section are in Appendix B.

3.1 Multi-channel and Graded Convolutions on Simplicial Complexes

Consider the simplicial complex $\mathcal{X}$, with a corresponding Dirac operator $\mathbf{D} = \sum_{i=1}^{\dim\mathcal{X}} \mathbf{D}_{[i]}$ which can be normalized or unnormalized, we use $\mathbf{D}$ to denote either. Let $\mathbf{X} \in \oplus_k C_k$ denote a signal on all ranks of the complex, with $\mathbf{X}_k \in \mathbb{R}^{N_k\times d_k}$ denoting the d_k-dimensional signal on rank k. In our notation, d'_k will denote the output dimensions for rank-k. Monomial polynomial basis refers to $\{T_i(x) = x^i\}_{i\geq 0}$, and Chebyshev polynomials refer to polynomial defined by the recursion $T_0(\lambda) = 1, T_1(\lambda) = \lambda, T_i(\lambda) = 2\lambda T_{i-1}(\lambda) - T_{i-2}$ (Hammond et al., 2009).

Definition 3.1 (Coupled linear spectral filter)**.** *For a fixed $\mathbf{D}_{[i]}$ on $\mathcal{X}$, a coupled linear spectral filter is a linear map $H_i : spec(\mathbf{D}_{[i]}) \to \mathbb{R}^{(d'_{i-1}+d'_i)\times(d_{i-1}+d_i)}$.*

Dirac convolutions. $\mathbf{D}_{[i]}$ acts on the simplices $\mathcal{X}_{i-1} \sqcup \mathcal{X}_i$, and the corresponding coupled signal space is $C_{i-1} \oplus C_i$. Each output of the filter $H_i(\lambda)$ acts on the coupled signals $\mathbf{X}_{i-1} \oplus \mathbf{X}_i$. In the matrix forms, the constraints become more apparent:

$$\mathbf{X}_{i-1} \oplus \mathbf{X}_i = \begin{pmatrix} \mathbf{X}_{i-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{X}_i \end{pmatrix}, \qquad H_i(\lambda) = \begin{pmatrix} H_i(\lambda)_{i-1,i-1} & H_i(\lambda)_{i,i-1} \\ H_i(\lambda)_{i-1,i} & H_i(\lambda)_{i,i} \end{pmatrix}. \tag{4}$$

As $\mathbf{D}_{[i]}$ is symmetric, it admits an orthogonal eigendecomposition $\mathbf{D}_{[i]} = \mathbf{U}_i \mathbf{\Lambda}_i \mathbf{U}_i^\top$. This leads to the following definitions of $\mathbf{D}_{[i]}$-*convolutions* and *Dirac convolutions*:

Definition 3.2 ($\mathbf{D}_{[i]}$-convolutions)**.** *Given $\mathbf{D}_{[i]}$ and a filter H_i, the corresponding $\mathbf{D}_{[i]}$-convolution is a map $H_i(\mathbf{D}_{[i]}) : \mathbb{R}^{(N_{i-1}+N_i)\times(d_{i-1}+d_i)} \to \mathbb{R}^{(N_{i-1}+N_i)\times(d'_{i-1}+d'_i)}$ defined via:*

$$H_i\big(\mathbf{D}_{[i]}\big)\big(\mathbf{X}_{i-1} \oplus \mathbf{X}_i\big) := \sum_{j=1}^{N_{i-1}+N_i} \mathbf{u}_{i,j}\mathbf{u}_{i,j}^\top \big(\mathbf{X}_{i-1} \oplus \mathbf{X}_i\big) H_i(\lambda_{i,j})^\top, \tag{5}$$

where $\mathbf{u}_{i,j}$ and $\lambda_{i,j}$ corresponds to j-th eigenvector and eigenvalue of $\mathbf{D}_{[i]}$, respectively.

As the eigenspaces of distinct $\mathbf{D}_{[i]}, \mathbf{D}_{[j]}$ are orthogonal, $\mathbf{D}_{[i]}$-convolutions act on distinct eigenspaces. This allows us to extend to *Dirac convolutions* on the whole signal $\mathbf{X}$:

Definition 3.3 (Dirac convolutions)**.** *Given $\mathbf{D} = \sum_{i=1}^{\dim \mathcal{X}} \mathbf{D}_{[i]}$ on a simplicial complex $\mathcal{X}$, and a collection of coupled linear spectral filters $(H_i)_{i=1}^{\dim \mathcal{X}}$, the Dirac convolution on a signal $\mathbf{X} \in \oplus_k C_k$ is defined via the action*

$$H(\mathbf{D})\mathbf{X} := \sum_{i=1}^{\dim \mathcal{X}} H_i\big(\mathbf{D}_{[i]}\big)\big(\mathbf{X}_{i-1} \oplus \mathbf{X}_i\big). \tag{6}$$

Intuitively, convolutions can be viewed as a method to process the signal independently on each eigenspace, where the processing is defined via the output of the spectral filter on the corresponding eigenvalue. Our formulation enforces that each $\mathbf{D}_{[i]}$ has an independent filter. As the eigenvalues in different ranks might correspond to different topological quantities (Grande & Schaub, 2024; 2025), this independence allows for flexibility and is desired. We now show that Dirac convolutions can be written as a polynomial of the Dirac operator:

Proposition 3.4. *For given frequency responses $H = \{H_1, H_2, \ldots, H_{\dim \mathcal{X}}\}$ and a simplical signal $\mathbf{X}$, the Dirac convolution can be written as a polynomial of $\mathbf{D}_{[1]}, \mathbf{D}_{[2]}, \ldots, \mathbf{D}_{[\dim \mathcal{X}]}$ as*

$$H(\mathbf{D})\mathbf{X} = \sum_{i=1}^{\dim \mathcal{X}} \sum_{k=0}^{R_i} T_k(\mathbf{D}_{[i]})\big(\mathbf{X}_{i-1} \oplus \mathbf{X}_i\big)\mathbf{\Theta}_{i,k}^\top \tag{7}$$

where $R_i \leq |spec(\mathbf{D}_{[i]})|$, and $\mathbf{\Theta}_{i,k} \in \mathbb{R}^{(d'_{i-1}+d'_i)\times(d_{i-1}+d_i)}$, where T_k denote any basis for polynomials of degree $\leq R_i$ (monomial, Chebyshev). Furthermore, if $T_k(-\lambda) = (-1)^k T_k(\lambda)$ and if we require the output to be in $C_{i-1} \oplus C_i$, then $\mathbf{\Theta}_{i,k}$ has block-wise diagonal structure for even k, and block-wise anti-diagonal structure for odd k:

$$\mathbf{\Theta}_{i,k} = \begin{pmatrix} \mathbf{\Theta}_{i,k}^{\downarrow} & \mathbf{0} \\ \mathbf{0} & \mathbf{\Theta}_{i,k}^{\uparrow} \end{pmatrix} \quad k \text{ is even}, \qquad \mathbf{\Theta}_{i,k} = \begin{pmatrix} \mathbf{0} & \mathbf{\Phi}_{i,k}^{\downarrow} \\ \mathbf{\Phi}_{i,k}^{\uparrow} & \mathbf{0} \end{pmatrix} \quad k \text{ is odd}, \tag{8}$$

with $\mathbf{\Theta}_{i,k}^{\uparrow} \in \mathbb{R}^{d'_i \times d_i}$, $\mathbf{\Theta}_{i,k}^{\downarrow} \in \mathbb{R}^{d'_{i-1} \times d_{i-1}}$, $\mathbf{\Phi}_{i,k}^{\uparrow} \in \mathbb{R}^{d'_i \times d_{i-1}}$, and $\mathbf{\Phi}_{i,k}^{\downarrow} \in \mathbb{R}^{d'_{i-1} \times d_i}$.

Chiral eigenpairs and parameter reduction. For any nonzero eigenvalue λ of $\mathbf{D}_{[i]}$, $-\lambda$ is also an eigenvalue, with corresponding eigenvectors related by *chirality*, $\mathbf{e}_{-\lambda} = \mathbf{\Gamma}_{[i]}\mathbf{e}_\lambda$, where $\mathbf{\Gamma}_{[i]} = \mathrm{diag}(\mathbf{I}_{N_{i-1}}, -\mathbf{I}_{N_i})$. Chiral pairing enforces the disappearance of diagonal and anti-diagonal entries in the odd and even powers of the polynomial. Consequently, (7) requires significantly fewer parameters than directly parametrizing the frequency response in (4). Moreover, since even powers of the Dirac operator correspond to powers of the Hodge Laplacians, the odd and even powers admit a natural interpretation as inter-rank and rank-wise processing, respectively.

Corollary 3.5. *With T_k as the Chebyshev basis, Dirac convolutions defined by $\widetilde{\mathbf{D}}$ have the following rank-wise update for rank $i \in \{1, 2, \ldots, \dim \mathcal{X} - 1\}$:*

$$\begin{aligned}
[H(\widetilde{\mathbf{D}})\mathbf{X}]_i = &\sum_{k=0}^{\lfloor \frac{L_i}{2} \rfloor} T_k(2\widetilde{\mathbf{L}}_{i,d} - \mathbf{I}_{N_i})\mathbf{X}_i(\mathbf{\Theta}_{i,k}^{\uparrow})^\top + \sum_{k=0}^{\lfloor \frac{L_i-1}{2} \rfloor} V_k(2\widetilde{\mathbf{L}}_{i,d} - \mathbf{I}_{N_i})\widetilde{\mathbf{B}}_i^\top \mathbf{X}_{i-1}(\mathbf{\Phi}_{i,k}^{\uparrow})^\top \\
&+ \sum_{k=0}^{\lfloor \frac{U_i}{2} \rfloor} T_k(2\widetilde{\mathbf{L}}_{i,u} - \mathbf{I}_{N_i})\mathbf{X}_i(\mathbf{\Theta}_{i+1,k}^{\downarrow})^\top + \sum_{k=0}^{\lfloor \frac{U_i-1}{2} \rfloor} V_k(2\widetilde{\mathbf{L}}_{i,u} - \mathbf{I}_{N_i})\widetilde{\mathbf{B}}_{i+1}\mathbf{X}_{i+1}(\mathbf{\Phi}_{i+1,k}^{\downarrow})^\top,
\end{aligned} \tag{9}$$

where $L_i = N_i + N_{i-1}, U_i = N_i + N_{i+1}$ and V_k denotes the Chebyshev polynomials of third kind defined by the recursion $V_0(\lambda) = 1, V_1(\lambda) = 2\lambda - 1$, and $V_k(\lambda) = 2\lambda V_{k-1}(\lambda) - V_{k-2}(\lambda)$.

Truncated Dirac Convolutions. Since computing the convolutions become computationally challenging as L_j and U_j grows, we choose to truncate the polynomials in (9). Chebyshev polynomials are orthogonal and more stable in the domain $[-1, 1]$, and have been previously utilized in spectral GNNs (Defferrard et al., 2017; Kipf & Welling, 2017). Normalization is therefore needed to leverage Chebyshev polynomials. Further detail on Chebyshev polynomials of third kind can be found in (Doman, 2015).

3.2 Dirac Convolutional Networks (DCNs)

We define *Dirac Convolutional Networks* (DCNs) as a family of architectures obtained by applying the polynomial Dirac convolution $H(\mathbf{D})\mathbf{X}$ in (7), where $\mathbf{D}$ may be normalized or unnormalized, $H(\mathbf{D})$ is any chosen polynomial basis. We allow the choice of R_i in (7) to truncate the polynomial. For a non-linearity ρ, learnable parameters $\mathbf{\Theta}, \mathbf{\Phi}$ and bias $\mathbf{b}^{(t)}$, an L-layer DCN is given by:

$$\mathbf{X}^{(t+1)} = \rho\Big(H(\mathbf{D})\mathbf{X}^{(t)} + \mathbf{b}^{(t)}\Big) \qquad t \in [L], \tag{10}$$

We instantiate DCNs as *Root-to-Leaf Dirac Convolutional Networks* (R2L-DCNs) using the normalized Dirac operator, a Chebyshev polynomial basis, and $R_i = 2$. The resulting convolution is given by (9) with $L_i = R_i = 2$, enabling both inter-rank messages through the (co)boundary operators and within-rank messages through the Hodge Laplacians, while remaining computationally tractable.

Recovering Existing Convolutional SNNs and GNNs. Existing convolutional SNNs can be recovered as a special instances of DCNs. DCNs recover SCCNNs (Yang et al., 2025) by using the unnormalized Dirac operator with monomial basis, SCNNs (Yang et al., 2022a) setting the coupling parameters to $\mathbf{0}$, (Ebli et al., 2020) by parameter sharing between up and down Laplacian messages.

Relation to message-passing. Message-passing (Gilmer et al., 2017) is often presented as an alternative to spectral methods in GNNs, and has been extended to SCs (Bodnar et al., 2021). Recent work shows that the two views have largely equivalent expressive power, differing primarily in the analysis tools they enable (Vasileiou et al., 2026). We therefore adopt the spectral view, and analyse asymptotic properties more naturally expressed in this framework.

3.3 Filter Transferability

In tasks involving multiple simplicial complexes, or when a learned model is used for different simplicial complexes, interpreting an SNN layer as a filter implicitly assumes that the corresponding informative eigenspaces occur at comparable frequencies across complexes. This does not necessarily hold for unnormalized operators. To illustrate, we consider the family of SCs, shown on the right, which we call *book complexes* $\{\mathcal{B}_k\}_{k\in\mathbb{N}^+}$ (Figure 2), where $\mathcal{B}_k$ consists of k triangles sharing a fixed edge, and all vertices and edges included in them. For the rank-1 up-Laplacian, the highest eigenvalue of $\mathcal{B}_k$ is $k + 2$, which corresponds to the geometric notion of *balancedness* of the signed incidence graph between edges and triangles (Fan et al., 2025; Viganò et al., 2026). Thus, a high-pass filter H_{high} designed to extract the balanced eigenspace of a 1-cochain on $\mathcal{B}_3$ by passing frequency $\lambda = 5$, will not extract the corresponding eigenspace of $\mathcal{B}_k$ for $k > 3$. However, under normalized operators, the balanced eigenspace will always occur at frequency $\widetilde{\lambda} = 1$, allowing a fixed filter $\widetilde{H}_{\text{high}}$ to pass this eigenspace across all complexes.

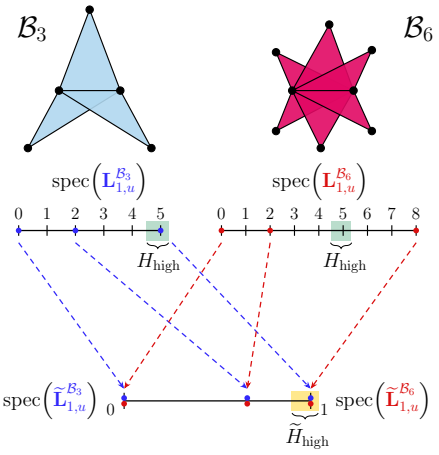


Figure 2: Unnormalized and normalized spectra of book complexes $\mathcal{B}_3$ and $\mathcal{B}_6$. H_{high} extracts the balanced eigenspace in $\mathcal{B}_3$ but not $\mathcal{B}_6$, while $\widetilde{H}_{\text{high}}$ extracts it in both.

More generally, normalization places the spectrum on a common scale across all simplicial complexes, allowing learned spectral filters (or SNN layers) on one complex to be applied to others without changing the frequencies they target. This provides a consistent spectral inductive bias across complexes, including unseen complexes whose unnormalized spectra might differ in scale from those observed during training. Similar notions of filter transferability have been studied for graph filters and GNNs (Maskey et al., 2023; Levie et al., 2021; 2019; Ruiz et al., 2021; Zhang et al., 2025).

4 Asymptotics of Dirichlet Energy in DCNs

We now focus on the linear term of the filter in equation 7. We view DCNs with residual connection as a forward-Euler discretization of a linear ODE and characterize its possible asymptotic behaviors. We show that in the continuous case, DCNs can preserve the norm and the Dirichlet energy with skew-symmetric weight matrices. In the discretization, we show that the normalized Dirichlet energy always tends to the maximum eigenvalue of the Hodge Laplacians and provide the rates of convergence. Results in this section work for both $\mathbf{D}$ and $\widetilde{\mathbf{D}}$, and proofs are in Appendix C.

4.1 Continuous Linear DCNs Can Preserve Dirichlet Energy

We focus our attention to a fixed $i \in \{1, 2, \dots, \dim \mathcal{X}\}$, and for a signal $\mathbf{X} \in \oplus_k C_k$, we denote $\mathbf{X}_{[i]} := \mathbf{X}_{i-1} \oplus \mathbf{X}_i$. As we focus on the linear term of (7), we drop the index k and use $\mathbf{\Theta}_{[i]}$ to denote the matrix with $\mathbf{\Phi}_i^{\downarrow}$ and $\mathbf{\Phi}_i^{\uparrow}$ in the anti-diagonals, as in (8), and assume the channel dimensions of each rank are fixed throughout the evolution. Hence we define a family of *continuous linear DCNs* and their Euler discretizations with the following equations:

$$\dot{\mathbf{X}}_{[i]}(t) = \mathbf{D}_{[i]}\mathbf{X}_{[i]}(t)\mathbf{\Theta}_{[i]}, \quad (11) \qquad \mathbf{X}_{[i]}^{(n+1)} = \mathbf{X}_{[i]}^{(n)} + \tau \mathbf{D}_{[i]}\mathbf{X}_{[i]}^{(n)}\mathbf{\Theta}_{[i]}, \quad \tau > 0, \quad (12)$$

with initial conditions $\mathbf{X}_{[i]}(0)$ and $\mathbf{X}_{[i]}^{(0)}$ respectively. Eq. (12) with $\tau = 1$ corresponds to a layer of DCN with residual connections and no non-linearities. Instead of defining energy rank-wise, we define it per coupling using the $\mathbf{L}_{[i]} := \mathbf{D}_{[i]}^2 = \mathbf{L}_{i-1,u} \oplus \mathbf{L}_{i,d}$. The i^{th} *Dirichlet energy* $\mathcal{E}_{\mathbf{L}_{[i]}}$ and the corresponding *normalized* i^{th} *Dirichlet energy* or the *Rayleigh quotient* $\mathcal{R}_{[\mathbf{L}_{[i]}]}$ are defined as follows:

$$\mathcal{E}_{\mathbf{L}_{[i]}}(\mathbf{X}_{[i]}) := \mathrm{Tr}(\mathbf{X}_{[i]}^\top \mathbf{L}_{[i]} \mathbf{X}_{[i]}), \quad (13) \qquad \mathcal{R}_{[\mathbf{L}_{[i]}]}(\mathbf{X}_{[i]}) := \frac{\mathcal{E}_{\mathbf{L}_{[i]}}(\mathbf{X}_{[i]})}{\|\mathbf{X}_{[i]}\|_F^2}. \quad (14)$$

We use the Rayleigh quotient in our discussions instead of the unnormalized energy, as it makes the analysis scale invariant and independent of the norm of the signal. We write $u \asymp v$ if there exists constants $0 < a \leq b$ and $t_0 \geq 0$ such that $au(t) \leq v(t) \leq bu(t)$ for $t \geq t_0$. We say f converges to A at rate r if $r(t) \to 0$ and $|f(t) - A| \asymp r(t)$. Throughout the statements, use μ_{max} to denote the largest eigenvalue of $\mathbf{L}_{[i]}$ and μ_{max-1} as the second-largest eigenvalue distinct from μ_{max}; and θ_{max} is the eigenvalue of $\mathbf{\Theta}_{[i]}$ with largest norm. We now characterize the asymptotic behaviour of continuous DCNs with symmetric and skew-symmetric weight matrices:

Proposition 4.1 (Continuous asymptotics). *If $\mathbf{\Theta}_{[i]}^\top = -\mathbf{\Theta}_{[i]}$, then for all initial conditions $\mathbf{X}_{[i]}(0)$ the solution of (11) satisfies:*

$$\|\mathbf{X}_{[i]}(t)\|_F^2 = \|\mathbf{X}_{[i]}(0)\|_F^2, \qquad \mathcal{E}_{\mathbf{L}_{[i]}}\big(\mathbf{X}_{[i]}(t)\big) = \mathcal{E}_{\mathbf{L}_{[i]}}\big(\mathbf{X}_{[i]}(0)\big).$$

Consequently, $\mathcal{R}_{[\mathbf{L}_{[i]}]}(\mathbf{X}_{[i]}(t))$ is constant. Alternatively, if $\mathbf{\Theta}_{[i]}^\top = \mathbf{\Theta}_{[i]} \neq \mathbf{0}$. Then for almost all initial conditions $\mathbf{X}_{[i]}(0)$ we have $\mathcal{R}_{[\mathbf{L}_{[i]}]}(\mathbf{X}_{[i]}(t)) \longrightarrow \mu_{max}$ and

$$\mu_{max} - \mathcal{R}_{[\mathbf{L}_{[i]}]}(\mathbf{X}_{[i]}(t)) \asymp e^{-2\theta_{max}(\sqrt{\mu_{max}} - \sqrt{\mu_{max-1}})t}.$$

Comparison to the graph case. In GNNs, oversmoothing is characterized by convergence of the graph Rayleigh quotient to 0, while continuous GCNs with symmetric weights converge towards extremal eigenspaces (Giovanni et al., 2023). Our result shows that the linear terms of continuous DCNs preserve the energy of each eigenspace, mitigating oversmoothing. However, since oversmoothing can be beneficial in homophilic settings, we emphasize that this applies only to the linear terms and even-order filter terms can induce different dynamics.

4.2 Euler Discretization Converges to Highest Eigenvalue

We now focus on the forward-Euler discretization as in (12) and characterize its asymptotic behaviour for symmetric and skew-symmetric $\mathbf{\Theta}_{[i]}$ as well as the convergence rates:

Proposition 4.2 (Forward-Euler asymptotics)**.** *Let $\mathbf{\Theta}_{[i]} \neq \mathbf{0}$, and suppose either of $\mathbf{\Theta}_{[i]}^\top = \mathbf{\Theta}_{[i]}$ or $\mathbf{\Theta}_{[i]}^\top = -\mathbf{\Theta}_{[i]}$ is true. Then for almost all initial conditions the solution of (12) satisfies:*

$$\mathcal{R}_{[\mathbf{L}_{[i]}]}(\mathbf{X}_{[i]}^{(n)}) \longrightarrow \mu_{max}. \tag{15}$$

The convergence rate is given by:

$$\mu_{max} - \mathcal{R}_{[\mathbf{L}_{[i]}]}(\mathbf{X}_{[i]}^{(n)}) \asymp \begin{cases} \left(\frac{1+\tau\theta_{max}\sqrt{\mu_{max}-1}}{1+\tau\theta_{max}\sqrt{\mu_{max}}}\right)^{2n} & \text{if} \quad \mathbf{\Theta}_{[i]}^\top = \mathbf{\Theta}_{[i]} \\ \left(\frac{1+\tau^2\theta_{max}^2\mu_{max}-1}{1+\tau^2\theta_{max}^2\mu_{max}}\right)^{n} & \text{if} \quad \mathbf{\Theta}_{[i]}^\top = -\mathbf{\Theta}_{[i]}. \end{cases} \tag{16}$$

Effects of Normalization on Convergence Rate. Both Prop. 4.1 and 4.2 apply to $\mathbf{D}_{[i]}$ and $\widetilde{\mathbf{D}}_{[i]}$, and the main difference is the boundedness of μ_{max}. In $\widetilde{\mathbf{D}}_{[i]}$, as the eigenvalues are upper-bounded by 1, convergence happens slower, while in $\mathbf{D}$ μ_{max} can be very large depending on the structure of $\mathcal{X}$. The convergence rates for unnormalized Laplacians can be improved by reducing the time-step; however, this is implicitly equivalent to normalizing the Dirac operator by dividing by a constant. Such normalization is not geometry-aware, while R2L normalization does this respecting the topology of the simplicial complex, making it preferable.

5 Empirical Results

We evaluate R2L-DCNs in the tasks of trajectory prediction, mesh, node and graph classification, and validate the convergence results in deep networks. We benchmark against a variety of SNNs and GNNs, specifically COSIMO (Einizade et al., 2025), SCCNN (Yang et al., 2025), S2CNN (Bunch et al., 2020), SCNN (Yang et al., 2021), SCoNe (Roddenberry et al., 2021), SNN (Ebli et al., 2020), SCCN (Yang et al., 2022b), SCA-CMPS (Hajij et al., 2022b), SAN (Giusti et al., 2022), HSN (Hajij et al., 2022a), GCN (Kipf & Welling, 2017), GAT (Veličković et al., 2018) and GIN (Xu et al., 2019). We also consider R2L-DCN$^{(\text{even})}$, which ignores the coupling and only passes messages rank-wise using the normalized Hodge Laplacians.

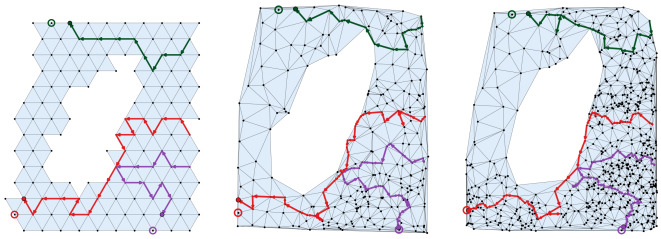


Figure 3: Sample paths from ocean drifters, with hexagonal grids (*left*) and irregular meshes (*center-right*) with 150 and 600 additional vertices. Empty circles indicate the next step in the trajectory.

5.1 Real World and Synthetic Datasets

Trajectory Prediction. Trajectories on meshes naturally form 1-cochains (i.e. edge-flows), with ocean drifters providing a standard benchmark (Schaub et al., 2020; Roddenberry et al., 2021). In this dataset, the task is to predict the next step of the trajectories of buoys near Madagascar, given the trajectory up until that point. We generate irregular re-meshings by adding points around selected regions to validate the importance of normalization in uneven edge-degree distributions (Figure 3). R2L-DCN consistently outperforms their unnormalized counterparts on both regular and irregular meshes (Table 1), highlighting the benefit of normalization under irregular degree distributions, which is established in the graph case (von Luxburg, 2007). R2L-DCN$^{(\text{even})}$ shows highest performance in the irregularized datasets, likely due to the fact that the signal only lies on the edge-level and no signal coupling is involved.

Table 1: Next-step prediction accuracy in ocean-drifters and its irregularizations (%) (↑). Best in **bold**, second best underlined.

Method	ocean	ocean$^{(+150)}$	ocean$^{(+600)}$
SNN	52.5 ±6.0	31.2 ±6.2	31.4 ±4.1
SCoNe	49.0 ±8.0	43.2 ±6.4	37.3 ±5.3
SCNN	53.0 ±8.0	40.1 ±3.0	34.2 ±4.1
S2CNN	46.0 ±6.0	36.9 ±6.4	22.7 ±8.4
SCCNN	54.5 ±8.0	45.6 ±7.1	40.2 ±5.1
COSIMO	55.0 ±6.0	35.2 ±6.0	27.6 ±6.9
R2L-DCN	**56.7** ±5.3	46.8 ±8.3	41.1 ±4.0
R2L-DCN$^{(even)}$	53.2 ±5.4	**48.0** ±6.0	**45.4** ±5.0

Table 2: Mesh-classification accuracy (%) (↑), in full and small SHREC'11 dataset. Best in **bold**, second best underlined.

Method	full	small
HSN	54.1 ±5.7	32.0 ±10.9
SAN	68.8 ±8.9	27.5 ±12.1
SCA-CMPS	62.0 ±6.7	33.0 ±8.9
SCCN	64.7 ±12.4	31.0 ±11.0
SCCNN	70.8 ±5.7	37.0 ±10.6
COSIMO	52.2 ±7.9	31.5 ±6.3
R2L-DCN	**72.7** ±5.8	**38.5** ±9.1

Table 3: Node and graph classification accuracy (%)(↑). ROC-AUC for minesweeper.

Method	roman-empire	amazon-ratings	minesweeper	texas	wisconsin	cornell	proteins
GCN	80.6 ±0.6	49.7 ±1.2	90.0 ±0.6	73.2 ±5.7	82.2 ±6.3	65.4 ±6.2	71.1 ±3.2
GAT	83.0 ±0.6	**49.9** ±1.0	90.6 ±0.6	78.4 ±7.0	84.5 ±7.8	**75.4** ±6.2	71.1 ±3.0
GIN	80.7 ±0.7	49.8 ±0.7	89.9 ±0.8	73.5 ±8.0	81.4 ±7.7	69.2 ±8.8	71.1 ±3.4
SCCN	85.6 ±0.7	46.4 ±0.7	90.3 ±1.1	66.5 ±8.4	57.6 ±5.7	44.6 ±8.3	69.6 ±2.4
SCCNN	85.5 ±0.4	45.3 ±0.6	88.9 ±0.8	73.0 ±7.5	65.7 ±7.1	52.7 ±8.8	73.1 ±3.1
COSIMO	84.4 ±0.5	40.4 ±0.5	87.6 ±0.6	**80.8** ±6.0	83.5 ±4.9	71.1 ±6.4	75.1 ±1.6
R2L-DCN	**86.8** ±0.7	44.2 ±0.9	**91.4** ±0.6	79.2 ±7.4	**84.7** ±6.0	71.6 ±10.0	**75.2** ±1.4

Mesh classification. Triangle meshes naturally form a 2-dimensional simplicial complex with features on all ranks, making it suitable to benchmark SNNs. Hence we further evaluate R2L-DCNs on the SHREC'11 non-rigid shape classification benchmark (Lian et al., 2011), which consists of 600 triangle meshes of deformable objects belonging to 30 different categories. We also create a smaller version of the dataset by sampling 100 of the 600 meshes, similar to (Einizade et al., 2025). R2L-DCNs outperform the competing models on this benchmark, indicating that normalization can help leverage the higher-order structure better compared to unnormalised models.

Graph Datasets. Finally, we evaluate R2L-DCNs on node and graph classification. For node classification, we use roman-empire, amazon-ratings, and minesweeper from Platonov et al. (2024), together with the Texas, Wisconsin, and Cornell datasets used by Pei et al. (2020). For graph classification, we use PROTEINS from the TUDataset collection (Morris et al., 2020). In the node classification tasks, we lift node features to the edge level using the coboundary map, as vertex-edge interaction is the simplest inter-rank coupling where Dirac operators can be defined. For PROTEINS, we lift each graph to a 2-dimensional complex by filling in all 3-cycles as triangles. R2L-DCNs achieve the best performance on 4 of the 7 datasets. We also observe that incorporating higher-order features does not always help, on some datasets it improves performance, while on others it degrades it. R2L-DCNs perform particularly well on the datasets where higher-order models perform better than GNNs, suggesting that R2L-DCNs can effectively exploit inter-rank coupling when such structure is informative.

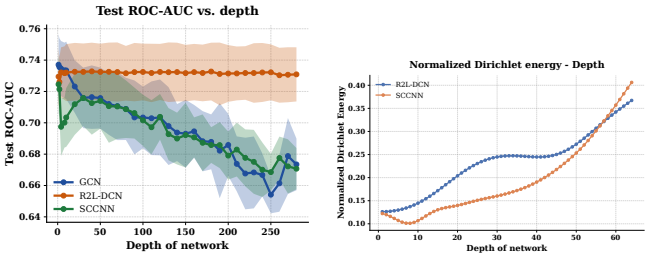


Figure 4: Minesweeper ROC-AUC per model depth, for GCN, R2L-DCN, SCCNN. Results over 5 splits (left). Normalized Dirichlet energy throughout the layers of the models (right).

5.2 Asymptotic Behaviour with Deep SNNs

In order to see the effect of the maximum eigenvalues, we rewire the minesweeper dataset to have uneven degree distribution. We do this by removing around 20000 edges while ensuring the graph

remains connected and every node has degree greater than 3, then adding 200000 new edges based on 2 and 3-hop connections. Models still perform near 70%, showing they learn (predicting the most frequent class is 50%). To be consistent with the theory, we consider models with no non-linearity and a weight matrix shared through layers. We observe that R2L-DCN keeps its performance while the performance of SCCNN, that uses unnormalized operators, degrades as models get deeper, which can be seen in Figure 4.

We further evaluate the convergence rates by computing the normalized Dirichlet energy at each layer for trained R2L-DCNs and SCCNNs, shown in Figure 4. As the theory suggests, the energies are increasing. One interesting observation is that, R2L-DCNs seem to converge faster than SCCNNs. This can be explained via the dependence of the rates on θ_{max} as well as the Dirac eigenvalues. We observe that, the final weight matrix learned by SCCNNs has $\|\mathbf{\Theta}\|_{op} \approx 0.11$, while for R2L-DCNs we have $\|\mathbf{\Theta}\|_{op} \approx 0.25$. This suggests, unnormalized models can compensate for the instability by learning smaller weight matrices, eventually degrading the performance, while normalized architectures do not suffer from this.

6 Related Work

Spectral Methods on Graphs and Simplicial Complexes. Overview of graph Fourier transform, choices of shift operators, and spectral filters on graphs have been discussed in (Tremblay et al., 2017; Domingos & Moura, 2020; Marques et al., 2016). Extensions of GSP to simplicial complexes have been done via Hodge Laplacians (Schaub et al., 2022; Petruso et al., 2026; Yang et al., 2022a; Isufi & Yang, 2022) and via Dirac operator (Calmon et al., 2023; Bianconi, 2021b; Wang et al., 2025; Nino et al., 2026). Spectra of Hodge Laplacians has also been interpreted and used for clustering (Grande & Schaub, 2024; 2025).

Learning on Simplicial Complexes and Graphs. Convolution based SNNs have been developed in (Ebli et al., 2020; Bunch et al., 2020; Yang et al., 2025), and various other SNNs have been proposed in literature (Bodnar et al., 2021; Goh et al., 2022; Roddenberry et al., 2021; Gurugubelli & Chepuri, 2024; Papillon et al., 2024). A Dirac operator based neural network architecture has been proposed in (Nauck et al., 2024). Dynamics-inspired analysis and architectures have been explored in GNNs (Chamberlain et al., 2021; Giovanni et al., 2023; Han et al., 2023) and extended to simplicial complexes in (Einizade et al., 2025). GNNs based on energy preserving dynamics and skew-symmetry have been proposed in (Trenta et al., 2025; Gravina et al., 2023).

Higher-order Deep Learning. Different higher-order structures can be used to represent relational data beyond simplicial complexes, such as cell complexes (Hajij et al., 2021; Bodnar et al., 2022), combinatorial complexes (Hajij et al., 2023a;b), and hypergraphs (Kim et al., 2024; Abiad et al., 2026). Several neural architectures are proposed to operate on these complexes (Hajij et al., 2025; Wang et al., 2026; Bastian et al., 2026; Barsbey et al., 2025). Our focus is on SCs, but similar methods can be developed and discussed for other domains.

Normalized Operators on Simplicial Complexes. Different normalization methods for Hodge Laplacians have been proposed in literature. (Schaub et al., 2020) also constructs normalized Hodge 1-Laplacians based on random walks. (Horak & Jost, 2011) bound the up-Laplacian spectra rank-wise between $[0, k+2]$, for a given rank k. We choose R2L normalization (Viganò et al., 2026), since it bounds both up and down Hodge Laplacians and the Dirac operator spectra independent of rank or dimension, preserving the desired geometric and topological properties.

7 Conclusion

We proposed *Dirac Convolutional Networks* (DCNs) as a common framework for learning across the ranks of a simplicial complex, with existing convolutional architectures arising as special cases. Through root-to-leaf normalization we instantiated DCN filters on a shared spectral scale and proved a slow-down in the concentration of deep linear models in the highest-frequency modes, by studying the asymptotics. Our experiments indicate strong results for our inter-rank coupled R2L-DCNs for signals with higher-order information.

Limitations and future work. Our stability proofs concern linear dynamics with restricted weights, and the experiments cover a limited set of tasks and complex constructions. Extending the analysis to nonlinear layers, testing transfer to unseen complexes, and learning when higher-order interactions are useful are natural future work.

AI USE STATEMENT

In this work, we used generative AI tools for programming, debugging and validating experimental pipelines, creating visualizations, literature research, ideation and language editing. We have reviewed all AI-assisted code, visuals, texts and resources. We take responsibility for the final content of this work, including text, claims or artifacts produced with the aid of generative AI.

ETHICS AND REPRODUCIBILITY.

Our work is mainly of theoretical nature, with a proposal of a new simplicial neural network architecture. We do not expect any negative impacts from our proposed architectures. All the data used is publicly available, and for the datasets we generated we have specified the generation details. We explain our training pipeline and therefore the experiments can be reproduced, and we will release our code upon publication.

REFERENCES

Aida Abiad, Alex Arenas, Ágnes Backhausz, József Balogh, Christopher R S Banerji, Sergio Barbarossa, Ginestra Bianconi, Christian Bick, Magnus Botnan, Timoteo Carletti, Lucia Cavallaro, Andrea Civilini, Tina Eliassi-Rad, Xue Gong, Krystal Guo, Heather Harrington, Jürgen Jost, P L Krapivsky, Pietro Liò, Ben D MacArthur, Carolina Mattsson, Pedro A M Mediano, Ana P Millan, Raffaella Mulas, Alice Patania, Giovanni Petri, Cerene Rathilal, Rubén J Sánchez-García, Martina Scolamiero, Michael T Schaub, Hanlin Sun, Yu Tian, Francesco Vaccarino, and Kelin Xia. Hypergraphs and simplicial complexes in focus: a roadmap for future research in higher-order interactions. *Journal of Physics: Complexity*, 7(2):022501, apr 2026. doi: 10.1088/2632-072X/ae3c4e. URL https://doi.org/10.1088/2632-072X/ae3c4e.

Uri Alon and Eran Yahav. On the bottleneck of graph neural networks and its practical implications, 2021. URL https://arxiv.org/abs/2006.05205.

Melih Barsbey, Rubén Ballester, Andac Demir, Carles Casacuberta, Pablo Hernández-García, David Pujol-Perich, Sarper Yurtseven, Sergio Escalera, Claudio Battiloro, Mustafa Hajij, et al. Higher-order molecular learning: The cellular transformer. In *ICLR 2025 Workshop on Generative and Experimental Perspectives for Biomolecular Design*, 2025.

Lennart Bastian, Samuel Leventhal, Mustafa Hajij, and Tolga Birdal. Topological neural operators, 2026. URL https://arxiv.org/abs/2606.09806.

Federico Battiston, Giulia Cencetti, Iacopo Iacopini, Vito Latora, Maxime Lucas, Alice Patania, Jean-Gabriel Young, and Giovanni Petri. Networks beyond pairwise interactions: Structure and dynamics. *Physics Reports*, 874:1–92, August 2020. ISSN 0370-1573. doi: 10.1016/j.physrep.2020.05.004. URL http://dx.doi.org/10.1016/j.physrep.2020.05.004.

Ginestra Bianconi. *Higher-Order Networks*. Elements in the Structure and Dynamics of Complex Networks. Cambridge University Press, 2021a.

Ginestra Bianconi. The topological dirac equation of networks and simplicial complexes. *Journal of Physics: Complexity*, 2(3):035022, 2021b. ISSN 2632-072X. doi: 10.1088/2632-072x/ac19be. URL http://dx.doi.org/10.1088/2632-072X/ac19be.

Christian Bick, Elizabeth Gross, Heather A. Harrington, and Michael T. Schaub. What are higher-order networks? *SIAM Review*, 65(3):686–731, August 2023. ISSN 1095-7200. doi: 10.1137/21m1414024. URL http://dx.doi.org/10.1137/21M1414024.

Cristian Bodnar, Fabrizio Frasca, Yu Guang Wang, Nina Otter, Guido Montúfar, Pietro Liò, and Michael Bronstein. Weisfeiler and lehman go topological: Message passing simplicial networks, 2021. URL https://arxiv.org/abs/2103.03212.

Cristian Bodnar, Fabrizio Frasca, Nina Otter, Yu Guang Wang, Pietro Liò, Guido Montúfar, and Michael Bronstein. Weisfeiler and lehman go cellular: Cw networks, 2022. URL https://arxiv.org/abs/2106.12575.

Joan Bruna, Wojciech Zaremba, Arthur Szlam, and Yann LeCun. Spectral networks and locally connected networks on graphs, 2014. URL https://arxiv.org/abs/1312.6203.

Eric Bunch, Qian You, Glenn Fung, and Vikas Singh. Simplicial 2-complex convolutional neural nets, 2020. URL https://arxiv.org/abs/2012.06010.

Lucille Calmon, Michael T Schaub, and Ginestra Bianconi. Dirac signal processing of higher-order topological signals. *New Journal of Physics*, 25(9):093013, 2023. ISSN 1367-2630. doi: 10.1088/1367-2630/acf33c. URL http://dx.doi.org/10.1088/1367-2630/acf33c.

Benjamin Paul Chamberlain, James Rowbottom, Maria Gorinova, Stefan Webb, Emanuele Rossi, and Michael M. Bronstein. Grand: Graph neural diffusion, 2021. URL https://arxiv.org/abs/2106.10934.

F. R. K. Chung. *Spectral Graph Theory*. American Mathematical Society, 1997.

Michaël Defferrard, Xavier Bresson, and Pierre Vandergheynst. Convolutional neural networks on graphs with fast localized spectral filtering, 2017. URL https://arxiv.org/abs/1606.09375.

B.G.S. Doman. *The Classical Orthogonal Polynomials*. World Scientific Publishing Company, 2015. ISBN 9789814704052. URL https://books.google.co.uk/books?id=91cGCwAAQBAJ.

Joao Domingos and Jose M. F. Moura. Graph fourier transform: A stable approximation. *IEEE Transactions on Signal Processing*, 68:4422–4437, 2020. ISSN 1941-0476. doi: 10.1109/tsp.2020.3009645. URL http://dx.doi.org/10.1109/TSP.2020.3009645.

X Dong, D Thanou, L Toni, M Bronstein, and P Frossard. Graph signal processing for machine learning: a review and new perspectives. *IEEE Signal Processing Magazine*, 37(6), 2020.

Stefania Ebli, Michaël Defferrard, and Gard Spreemann. Simplicial neural networks, 2020. URL https://arxiv.org/abs/2010.03633.

Aref Einizade, Dorina Thanou, Fragkiskos D. Malliaros, and Jhony H. Giraldo. Continuous simplicial neural networks, 2025. URL https://arxiv.org/abs/2503.12919.

Yi-Zheng Fan, Hui-Feng Wu, and Yi Wang. The largest laplacian eigenvalue and the balancedness of simplicial complexes. *Journal of Algebraic Combinatorics*, 61(4), 2025. ISSN 1572-9192. doi: 10.1007/s10801-025-01419-1. URL http://dx.doi.org/10.1007/s10801-025-01419-1.

Fernando Gama, Elvin Isufi, Geert Leus, and Alejandro Ribeiro. Graphs, convolutions, and neural networks: From graph filters to graph neural networks. *IEEE Signal Processing Magazine*, 37(6):128–138, November 2020. ISSN 1558-0792. doi: 10.1109/msp.2020.3016143. URL http://dx.doi.org/10.1109/MSP.2020.3016143.

Justin Gilmer, Samuel S. Schoenholz, Patrick F. Riley, Oriol Vinyals, and George E. Dahl. Neural message passing for quantum chemistry, 2017. URL https://arxiv.org/abs/1704.01212.

Francesco Di Giovanni, James Rowbottom, Benjamin P. Chamberlain, Thomas Markovich, and Michael M. Bronstein. Understanding convolution on graphs via energies, 2023. URL https://arxiv.org/abs/2206.10991.

L. Giusti, C. Battiloro, P. Di Lorenzo, S. Sardellitti, and S. Barbarossa. Simplicial attention neural networks, 2022. URL https://arxiv.org/abs/2203.07485.

Chris Godsil and Gordon F. Royle. *Algebraic Graph Theory*. Number Book 207 in Graduate Texts in Mathematics. Springer, 2001. ISBN 9781461301639 1461301637. doi: 10.1007/978-1-4613-0163-9.

Christopher Wei Jin Goh, Cristian Bodnar, and Pietro Liò. Simplicial attention networks, 2022. URL https://arxiv.org/abs/2204.09455.

Ian Goodfellow, Yoshua Bengio, Aaron Courville, and Yoshua Bengio. *Deep learning*, volume 1. MIT Press, 2016.

Vincent P. Grande and Michael T. Schaub. Disentangling the spectral properties of the hodge laplacian: not all small eigenvalues are equal. In *ICASSP 2024 - 2024 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 9896–9900, 2024. doi: 10.1109/ICASSP48485.2024.10446051.

Vincent P. Grande and Michael T. Schaub. Topological point cloud clustering, 2025. URL https://arxiv.org/abs/2303.16716.

Alessio Gravina, Davide Bacciu, and Claudio Gallicchio. Anti-symmetric dgn: a stable architecture for deep graph networks, 2023. URL https://arxiv.org/abs/2210.09789.

Sravanthi Gurugubelli and Sundeep Prabhakar Chepuri. SaNN: Simple yet powerful simplicial-aware neural networks. In *The Twelfth International Conference on Learning Representations*, 2024. URL https://openreview.net/forum?id=eUgS9Ig8JG.

Mustafa Hajij, Kyle Istvan, and Ghada Zamzmi. Cell complex neural networks, 2021. URL https://arxiv.org/abs/2010.00743.

Mustafa Hajij, Karthikeyan Natesan Ramamurthy, Aldo Guzmán-Sáenz, and Ghada Za. HIGH SKIP NETWORKS: A HIGHER ORDER GENERALIZATION OF SKIP CONNECTIONS. In *ICLR 2022 Workshop on Geometrical and Topological Representation Learning*, 2022a. URL https://openreview.net/forum?id=Sc8glB-k6e9.

Mustafa Hajij, Ghada Zamzmi, Theodore Papamarkou, Vasileios Maroulas, and Xuanting Cai. Simplicial complex representation learning, 2022b. URL https://arxiv.org/abs/2103.04046.

Mustafa Hajij, Ghada Zamzmi, Theodore Papamarkou, AIdo Guzman-Saenz, ToIga Birdal, and Michael T Schaub. Combinatorial complexes: bridging the gap between cell complexes and hypergraphs. In *2023 57th Asilomar Conference on Signals, Systems, and Computers*, pp. 799–803. IEEE, 2023a.

Mustafa Hajij, Ghada Zamzmi, Theodore Papamarkou, Nina Miolane, Aldo Guzmán-Sáenz, Karthikeyan Natesan Ramamurthy, Tolga Birdal, Tamal K. Dey, Soham Mukherjee, Shreyas N. Samaga, Neal Livesay, Robin Walters, Paul Rosen, and Michael T. Schaub. Topological deep learning: Going beyond graph data, 2023b. URL https://arxiv.org/abs/2206.00606.

Mustafa Hajij, Mathilde Papillon, Florian Frantzen, Jens Agerberg, Ibrahem AlJabea, Rubén Ballester, Claudio Battiloro, Guillermo Bernárdez, Tolga Birdal, Aiden Brent, Peter Chin, Sergio Escalera, Simone Fiorellino, Odin Hoff Gardaa, Gurusankar Gopalakrishnan, Devendra Govil, Josef Hoppe, Maneel Reddy Karri, Jude Khouja, Manuel Lecha, Neal Livesay, Jan Meißner, Soham Mukherjee, Alexander Nikitin, Theodore Papamarkou, Jaro Prílepok, Karthikeyan Natesan Ramamurthy, Paul Rosen, Aldo Guzmán-Sáenz, Alessandro Salatiello, Shreyas N. Samaga, Simone Scardapane, Michael T. Schaub, Luca Scofano, Indro Spinelli, Lev Telyatnikov, Quang Truong, Robin Walters, Maosheng Yang, Olga Zaghen, Ghada Zamzmi, Ali Zia, and Nina Miolane. Topox: A suite of python packages for machine learning on topological domains, 2024. URL https://arxiv.org/abs/2402.02441.

Mustafa Hajij, Lennart Bastian, Sarah Osentoski, Hardik Kabaria, John L. Davenport, Sheik Dawood, Balaji Cherukuri, Joseph G. Kocheemoolayil, Nastaran Shahmansouri, Adrian Lew, Theodore Papamarkou, and Tolga Birdal. Copresheaf topological neural networks: A generalized deep learning framework, 2025. URL https://arxiv.org/abs/2505.21251.

William L. Hamilton. Graph representation learning. *Synthesis Lectures on Artificial Intelligence and Machine Learning*, 14(3):1–159, 2020.

David K Hammond, Pierre Vandergheynst, and Rémi Gribonval. Wavelets on graphs via spectral graph theory, 2009. URL https://arxiv.org/abs/0912.3848.

Andi Han, Dai Shi, Lequan Lin, and Junbin Gao. From continuous dynamics to graph neural networks: Neural diffusion and beyond, 2023. URL https://arxiv.org/abs/2310.10121.

Danijela Horak and Jürgen Jost. Spectra of combinatorial laplace operators on simplicial complexes, 2011. URL https://arxiv.org/abs/1105.2712.

Elvin Isufi and Maosheng Yang. Convolutional filtering in simplicial complexes, 2022. URL https://arxiv.org/abs/2201.12584.

Jürgen Jost and Dong Zhang. Cheeger inequalities on simplicial complexes. *ANNALI SCUOLA NORMALE SUPERIORE - CLASSE DI SCIENZE*, pp. 30, December 2024. ISSN 0391-173X. doi: 10.2422/2036-2145.202307_009. URL http://dx.doi.org/10.2422/2036-2145.202307_009.

Sunwoo Kim, Soo Yong Lee, Yue Gao, Alessia Antelmi, Mirko Polato, and Kijung Shin. A survey on hypergraph neural networks: An in-depth and step-by-step guide. In *Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, KDD '24, pp. 6534–6544. ACM, August 2024. doi: 10.1145/3637528.3671457. URL http://dx.doi.org/10.1145/3637528.3671457.

Thomas N. Kipf and Max Welling. Semi-supervised classification with graph convolutional networks, 2017. URL https://arxiv.org/abs/1609.02907.

Geert Leus, Antonio G. Marques, José M.F. Moura, Antonio Ortega, and David I Shuman. Graph signal processing: History, development, impact, and outlook. *IEEE Signal Processing Magazine*, 40(4):49–60, 2023. ISSN 1558-0792. doi: 10.1109/msp.2023.3262906. URL http://dx.doi.org/10.1109/MSP.2023.3262906.

Ron Levie, Elvin Isufi, and Gitta Kutyniok. On the transferability of spectral graph filters, 2019. URL https://arxiv.org/abs/1901.10524.

Ron Levie, Wei Huang, Lorenzo Bucci, Michael Bronstein, and Gitta Kutyniok. Transferability of spectral graph convolutional neural networks. *Journal of Machine Learning Research*, 22(272): 1–59, 2021. URL http://jmlr.org/papers/v22/20-213.html.

David A. Levin, Yuval Peres, and Elizabeth L. Wilmer. *Markov chains and mixing times*. American Mathematical Society, 2006. URL http://scholar.google.com/scholar.bib?q=info:3wf9IU94tyMJ:scholar.google.com/&output=citation&hl=en&as_sdt=2000&ct=citation&cd=0.

Qimai Li, Zhichao Han, and Xiao-Ming Wu. Deeper insights into graph convolutional networks for semi-supervised learning, 2018. URL https://arxiv.org/abs/1801.07606.

Z. Lian, A. Godil, B. Bustos, M. Daoudi, J. Hermans, S. Kawamura, Y. Kurita, G. Lavoué, H. V. Nguyen, R. Ohbuchi, Y. Ohkita, Y. Ohishi, F. Porikli, M. Reuter, I. Sipiran, D. Smeets, P. Suetens, H. Tabia, and D. Vandermeulen. Shrec '11 track: Shape retrieval on non-rigid 3d watertight meshes. In *Eurographics Workshop on 3D Object Retrieval*, pp. 79–88. The Eurographics Association, 2011. doi: 10.2312/3DOR/3DOR11/079-088.

Lek-Heng Lim. Hodge laplacians on graphs, 2019. URL https://arxiv.org/abs/1507.05379.

Antonio G. Marques, Santiago Segarra, Geert Leus, and Alejandro Ribeiro. Sampling of graph signals with successive local aggregations. *IEEE Transactions on Signal Processing*, 64(7):1832–1843, 2016. doi: 10.1109/TSP.2015.2507546.

Sohir Maskey, Ron Levie, and Gitta Kutyniok. Transferability of graph neural networks: An extended graphon approach. *Applied and Computational Harmonic Analysis*, 63:48–83, 2023. ISSN 1063-5203. doi: https://doi.org/10.1016/j.acha.2022.11.008. URL https://www.sciencedirect.com/science/article/pii/S1063520322000987.

J.C. Mason and D.C. Handscomb. *Chebyshev Polynomials*. CRC Press, 2002. ISBN 9781420036114. URL https://books.google.co.uk/books?id=g1DMBQAAQBAJ.

Christopher Morris, Nils M. Kriege, Franka Bause, Kristian Kersting, Petra Mutzel, and Marion Neumann. Tudataset: A collection of benchmark datasets for learning with graphs, 2020. URL https://arxiv.org/abs/2007.08663.

Vidit Nanda. Computational algebraic topology, 2021. URL https://people.maths.ox.ac.uk/nanda/cat/TDANotes.pdf. Lecture notes, University of Oxford.

Christian Nauck, Rohan Gorantla, Michael Lindner, Konstantin Schürholt, Antonia S. J. S. Mey, and Frank Hellmann. Dirac–bianconi graph neural networks – enabling non-diffusive long-range graph predictions, 2024. URL https://arxiv.org/abs/2407.12419.

Leonardo Di Nino, Tiziana Cattai, Sergio Barbarossa, Ginestra Bianconi, and Paolo Di Lorenzo. Learning dirac spectral transforms for topological signals, 2026. URL https://arxiv.org/abs/2602.14590.

Theodore Papamarkou, Tolga Birdal, Michael M. Bronstein, Gunnar E. Carlsson, Justin Curry, Yue Gao, Mustafa Hajij, Roland Kwitt, Pietro Lio, Paolo Di Lorenzo, Vasileios Maroulas, Nina Miolane, Farzana Nasrin, Karthikeyan Natesan Ramamurthy, Bastian Rieck, Simone Scardapane, Michael T. Schaub, Petar Veličković, Bei Wang, Yusu Wang, Guowei Wei, and Ghada Zamzmi. Position: Topological deep learning is the new frontier for relational learning. In *Proceedings of the 41st International Conference on Machine Learning*, volume 235 of *Proceedings of Machine Learning Research*, pp. 39529–39555. PMLR, 2024. URL https://proceedings.mlr.press/v235/papamarkou24a.html.

Mathilde Papillon, Sophia Sanborn, Mustafa Hajij, and Nina Miolane. Architectures of topological deep learning: A survey of message-passing topological neural networks, 2024. URL https://arxiv.org/abs/2304.10031.

Hongbin Pei, Bingzhe Wei, Kevin Chen-Chuan Chang, Yu Lei, and Bo Yang. Geom-gcn: Geometric graph convolutional networks, 2020. URL https://arxiv.org/abs/2002.05287.

Flavia Petruso, Maria Giulia Preti, and Dimitri Van De Ville. Topological signal processing: An application-oriented tutorial, 2026. URL https://arxiv.org/abs/2605.22853.

Oleg Platonov, Denis Kuznedelev, Michael Diskin, Artem Babenko, and Liudmila Prokhorenkova. A critical look at the evaluation of gnns under heterophily: Are we really making progress?, 2024. URL https://arxiv.org/abs/2302.11640.

T. Mitchell Roddenberry, Nicholas Glaze, and Santiago Segarra. Principled simplicial neural networks for trajectory prediction, 2021. URL https://arxiv.org/abs/2102.10058.

Luana Ruiz, Fernando Gama, and Alejandro Ribeiro. Graph neural networks: Architectures, stability and transferability, 2021. URL https://arxiv.org/abs/2008.01767.

T. Konstantin Rusch, Michael M. Bronstein, and Siddhartha Mishra. A survey on oversmoothing in graph neural networks, 2023. URL https://arxiv.org/abs/2303.10993.

Michael T. Schaub, Austin R. Benson, Paul Horn, Gabor Lippner, and Ali Jadbabaie. Random walks on simplicial complexes and the normalized hodge 1-laplacian. *SIAM Review*, 62(2):353–391, January 2020. ISSN 1095-7200. doi: 10.1137/18m1201019. URL http://dx.doi.org/10.1137/18M1201019.

Michael T. Schaub, Jean-Baptiste Seby, Florian Frantzen, T. Mitchell Roddenberry, Yu Zhu, and Santiago Segarra. Signal processing on simplicial complexes, 2022. URL https://arxiv.org/abs/2106.07471.

Diaaeldin Taha, James Chapman, Marzieh Eidi, Karel Devriendt, and Guido Montúfar. Demystifying topological message-passing with relational structures: A case study on oversquashing in simplicial message-passing, 2025. URL https://arxiv.org/abs/2506.06582.

Lev Telyatnikov, Guillermo Bernardez, Marco Montagna, Mustafa Hajij, Martin Carrasco, Pavlo Vasylenko, Mathilde Papillon, Ghada Zamzmi, Michael T. Schaub, Jonas Verhellen, Pavel Snopov, Bertran Miquel-Oliver, Manel Gil-Sorribes, Alexis Molina, Victor Guallar, Theodore Long, Julian Suk, Patryk Rygiel, Alexander Nikitin, Giordan Escalona, Michael Banf, Dominik Filipiak, Max Schattauer, Liliya Imasheva, Alvaro Martinez, Halley Fritze, Marissa Masden, Valentina Sánchez, Manuel Lecha, Andrea Cavallo, Claudio Battiloro, Matt Piekenbrock, Mauricio Tec, George Dasoulas, Nina Miolane, Simone Scardapane, and Theodore Papamarkou. Topobench: A framework for benchmarking topological deep learning, 2025. URL https://arxiv.org/abs/2406.06642.

Jake Topping, Francesco Di Giovanni, Benjamin Paul Chamberlain, Xiaowen Dong, and Michael M. Bronstein. Understanding over-squashing and bottlenecks on graphs via curvature, 2022. URL https://arxiv.org/abs/2111.14522.

Nicolas Tremblay, Paulo Gonçalves, and Pierre Borgnat. Design of graph filters and filterbanks, 2017. URL https://arxiv.org/abs/1711.02046.

Alessandro Trenta, Alessio Gravina, and Davide Bacciu. SONAR: Long-range graph propagation through information waves. In *The Thirty-ninth Annual Conference on Neural Information Processing Systems*, 2025. URL https://openreview.net/forum?id=Hxfjmc95rl.

Antonis Vasileiou, Juan Cervino, Pascal Frossard, Charilaos I. Kanatsoulis, Christopher Morris, Michael T. Schaub, Pierre Vandergheynst, Zhiyang Wang, Guy Wolf, and Ron Levie. Graph learning should move beyond restrictive views of spectral and message-passing gnns, 2026. URL https://arxiv.org/abs/2602.10031.

Petar Veličković, Guillem Cucurull, Arantxa Casanova, Adriana Romero, Pietro Liò, and Yoshua Bengio. Graph attention networks, 2018. URL https://arxiv.org/abs/1710.10903.

Francesco Viganò, Tolga Birdal, Michael T. Schaub, and Mauricio Barahona. Root-to-leaf path random walks, normalized hodge laplacians, and cheeger inequalities on simplicial complexes, 2026. URL https://arxiv.org/abs/2604.27241.

Ulrike von Luxburg. A tutorial on spectral clustering, 2007. URL https://arxiv.org/abs/0711.0189.

Runyue Wang, Yu Tian, Pietro Liò, and Ginestra Bianconi. Dirac-equation signal processing: Physics boosts topological machine learning. *PNAS Nexus*, 4(5), April 2025. ISSN 2752-6542. doi: 10.1093/pnasnexus/pgaf139. URL http://dx.doi.org/10.1093/pnasnexus/pgaf139.

Tony Danjun Wang, Ka Young Kim, Tolga Birdal, Nassir Navab, and Lennart Bastian. Topoor: A unified topological scene representation for the operating room, 2026. URL https://arxiv.org/abs/2603.09466.

Keyulu Xu, Weihua Hu, Jure Leskovec, and Stefanie Jegelka. How powerful are graph neural networks?, 2019. URL https://arxiv.org/abs/1810.00826.

Maosheng Yang, Elvin Isufi, and Geert Leus. Simplicial convolutional neural networks, 2021. URL https://arxiv.org/abs/2110.02585.

Maosheng Yang, Elvin Isufi, Michael T. Schaub, and Geert Leus. Simplicial convolutional filters. *IEEE Transactions on Signal Processing*, 70:4633–4648, 2022a. ISSN 1941-0476. doi: 10.1109/tsp.2022.3207045. URL http://dx.doi.org/10.1109/TSP.2022.3207045.

Maosheng Yang, Geert Leus, and Elvin Isufi. Hodge-aware convolutional learning on simplicial complexes. *Transactions on Machine Learning Research*, 2025. ISSN 2835-8856. URL https://openreview.net/forum?id=Nm5sp09Q25.

Ruochen Yang, Frederic Sala, and Paul Bogdan. Efficient representation learning for higher-order data with simplicial complexes. In Bastian Rieck and Razvan Pascanu (eds.), *Proceedings of the First Learning on Graphs Conference*, volume 198 of *Proceedings of Machine Learning Research*, pp. 13:1–13:21. PMLR, 09–12 Dec 2022b. URL https://proceedings.mlr.press/v198/yang22a.html.

Haoyu Zhang, Yuxuan Cheng, Wenqi Fan, Yulong Chen, and Yifan Zhang. *Rethinking Graph Domain Adaptation: A Spectral Contrastive Perspective*, pp. 448–464. Springer Nature Switzerland, October 2025. ISBN 9783032061065. doi: 10.1007/978-3-032-06106-5_26. URL http://dx.doi.org/10.1007/978-3-032-06106-5_26.

A Further Details on Simplicial Complexes

Here, we provide a more detailed explanation of how orientations are defined in simplicial complexes and briefly define Hodge decomposition. We also explain the normalised graph Laplacians and their uses more in detail.

Orientations and signals. A simplex is *oriented* by fixing an ordering of its vertices, written $\sigma = [v_{i_0}, \dots, v_{i_k}]$; two orderings give the same orientation iff they differ by an even permutation, so each simplex has exactly two, and we write $-\sigma$ for the flip. The signals of rank k form the *cochain group* C_k of alternating functions f on oriented k-simplices, i.e. $f(-\sigma) = -f(\sigma)$. Fixing a reference orientation for every simplex (in practice the one induced by $v_0 < \cdots < v_n$) gives $C_k \cong \mathbb{R}^{|\mathcal{X}_k|}$, with one basis vector per simplex. A different choice of reference orientations is recorded by a diagonal sign matrix $\mathbf{S}_k \in \{-1, +1\}^{|\mathcal{X}_k| \times |\mathcal{X}_k|}$ with $\mathbf{S}_k^2 = \mathbf{I}$ and $\mathbf{S}_0 = \mathbf{I}$.

Boundaries and Cohomology. If the face σ' of $\sigma = [v_{i_0}, \dots, v_{i_k}]$ is obtained by deleting v_{i_j}, set $[\sigma' : \sigma] = (-1)^j$. The *boundary operator* $\mathbf{B}_k : C_k \to C_{k-1}$ is the matrix with entries $(\mathbf{B}_k)_{\sigma'\sigma} = [\sigma' : \sigma]$ when σ' is a face of σ and 0 otherwise, and its adjoint $\mathbf{B}_k^\top$ is the *coboundary operator*. Consecutive coboundaries compose to zero, $\mathbf{B}_{k+1}^\top \mathbf{B}_k^\top = 0$, so $\operatorname{im} \mathbf{B}_k^\top \subseteq \ker \mathbf{B}_{k+1}^\top$ and the *cohomology* groups are the quotient vector spaces $H_k = \ker \mathbf{B}_{k+1}^\top / \operatorname{im} \mathbf{B}_k^\top$. Their dimensions count k-dimensional holes, which are connected components for $k = 0$, independent cycles for $k = 1$, cavities for $k = 2$, and so on. A change of orientation acts by $\mathbf{B}_k^\top \mapsto \mathbf{S}_k \mathbf{B}_k^\top \mathbf{S}_{k-1}$, which preserves $\mathbf{B}_{k+1}^\top \mathbf{B}_k^\top = 0$ and leaves everything below unchanged up to this change of basis.

Hodge Laplacians. The *down* and *up Laplacians* at rank k are $\mathbf{L}_{k,d} = \mathbf{B}_k^\top \mathbf{B}_k$ and $\mathbf{L}_{k,u} = \mathbf{B}_{k+1} \mathbf{B}_{k+1}^\top$, and the k-th *Hodge Laplacian* is $\mathbf{L}_k = \mathbf{L}_{k,d} + \mathbf{L}_{k,u}$, all acting on C_k. One has $\ker \mathbf{L}_k \cong H_k$, and the *Hodge decomposition* splits the signal space into orthogonal subspaces,

$$\mathbb{R}^{|\mathcal{X}_k|} \cong \operatorname{im}(\mathbf{B}_{k+1}) \oplus \ker(\mathbf{L}_k) \oplus \operatorname{im}(\mathbf{B}_k^\top), \tag{17}$$

which gives an interpretable decomposition of any rank-k signal. For $k = 0$ this recovers the graph Laplacian, so $\mathbf{L}_k$ generalizes it to higher order (Lim, 2019). Hodge decomposition depends on the fact that $\mathbf{L}_{k,d}\mathbf{L}_{k,u} = \mathbf{L}_{k,u}\mathbf{L}_{k,d} = \mathbf{0}$. Normalized root-to-leaf Hodge Laplacians satisfy this property, hence provide a valid normalized Hodge decomposition.

Normalized graph Laplacians. Let $\mathcal{G} = (V, E)$ be a graph on $V = \{v_1, \dots, v_n\}$, with *adjacency matrix* $\mathbf{A} \in \mathbb{R}^{n \times n}$ given by $A_{ij} = 1$ if $\{v_i, v_j\} \in E$ and 0 otherwise, and *degree matrix* $\mathbf{K} := \operatorname{diag}(\mathbf{A}\mathbf{1})$. The *graph Laplacian* $\mathbf{\Delta} := \mathbf{K} - \mathbf{A}$ is the standard operator acting on functions on vertices. Their *normalized* variants are $\widetilde{\mathbf{A}} := \mathbf{K}^{-1/2}\mathbf{A}\mathbf{K}^{-1/2}$ and $\widetilde{\mathbf{\Delta}} := \mathbf{K}^{-1/2}\mathbf{\Delta}\mathbf{K}^{-1/2} = \mathbf{I} - \widetilde{\mathbf{A}}$, whose main effect is to bound the spectrum (i.e. the multiset of eigenvalues) uniformly across graphs: $\operatorname{spec}(\widetilde{\mathbf{A}}) \subseteq [-1, 1]$ and $\operatorname{spec}(\widetilde{\mathbf{\Delta}}) \subseteq [0, 2]$. The spectrum of $\widetilde{\mathbf{\Delta}}$ carries geometric and dynamical information: the multiplicity of the eigenvalue 0 is the number of connected components, the multiplicity of 2 is the number of bipartite connected components, and the smallest nonzero eigenvalue bounds the edge expansion of $\mathcal{G}$ and controls the mixing time of the random walk on it (Levin et al., 2006; Chung, 1997).

B Proofs and Additional Details of Section 3

In this section we provide further details of the proofs in section 3.

Proof of Proposition 3.4. The spectral filter H_i is a function defined on $\operatorname{spec}(\mathbf{D}_i)$, which is a finite set of at most $|\operatorname{spec}(\mathbf{D}_{[i]})|$ elements, therefore H_i only takes finitely many values in $\mathbb{R}^{(d'_{i-1}+d'_i)\times(d_{i-1}+d_i)}$. Via any polynomial interpolation method, we can find $R_i \leq |\operatorname{spec}(\mathbf{D}_{[i]})|$ such that there is a polynomial p of degree $\leq R_i$, with matrix coefficients, satisfying $H_i(\lambda) = p(\lambda)$ for all $\lambda \in \operatorname{spec}(\mathbf{D}_{[i]})$. Consider any basis $\{T_k\}_k$ for all polynomials of degree R_i, which is $R_i + 1$ dimensional. Then there are $R_i + 1$ matrices $\{\mathbf{\Theta}_{i,k}\}_{k=0}^{R_i}$ for which:

$$p(\lambda) = \sum_{k=0}^{R_i} \mathbf{\Theta}_{i,k} T_k(\lambda) \tag{18}$$

Substituting this polynomial into equation 5, we get:

$$\begin{aligned} H_i\big(\mathbf{D}_{[i]}\big)\big(\mathbf{X}_{i-1}\oplus\mathbf{X}_i\big) &:= \sum_{j=1}^{N_{i-1}+N_i} \mathbf{u}_{i,j}\mathbf{u}_{i,j}^\top\big(\mathbf{X}_{i-1}\oplus\mathbf{X}_i\big)H_i(\lambda_{i,j})^\top \\ &= \sum_{j=1}^{N_{i-1}+N_i} \mathbf{u}_{i,j}\mathbf{u}_{i,j}^\top\big(\mathbf{X}_{i-1}\oplus\mathbf{X}_i\big)\sum_{k=0}^{R_i}\mathbf{\Theta}_{i,k}^\top T_k(\lambda_{i,j}) \\ &= \sum_{k=0}^{R_i}\left(\sum_{j=1}^{N_{i-1}+N_i}\mathbf{u}_{i,j}T_k(\lambda_{i,j})\mathbf{u}_{i,j}^\top\big(\mathbf{X}_{i-1}\oplus\mathbf{X}_i\big)\mathbf{\Theta}_{i,k}^\top\right). \end{aligned} \tag{19}$$

As $\mathbf{D}_{[i]}$ is symmetric, its eigenvectors are orthogonal and every polynomial directly acts onto its eigenvalues:

$$T_k(\mathbf{D}_{[i]}) = T_k(\mathbf{U}_i\mathbf{\Lambda}_i\mathbf{U}_i^\top) = \mathbf{U}_i T_k(\mathbf{\Lambda}_i)\mathbf{U}_i^\top = \sum_{j=1}^{N_{i-1}+N_i}\mathbf{u}_{i,j}T_k(\lambda_{i,j})\mathbf{u}_{i,j}^\top \tag{20}$$

and substituting it into equation 19 gives:

$$H_i\big(\mathbf{D}_{[i]}\big)\big(\mathbf{X}_{i-1}\oplus\mathbf{X}_i\big) = \sum_{k=0}^{R_i} T_k(\mathbf{D}_{[i]})\big(\mathbf{X}_{i-1}\oplus\mathbf{X}_i\big)\mathbf{\Theta}_{i,k}^\top. \tag{21}$$

For the second part of the proof, define:

$$\mathbf{\Gamma}_{[i]} := \begin{pmatrix}\mathbf{I}_{N_{i-1}} & \mathbf{0}\\ \mathbf{0} & -\mathbf{I}_{N_i}\end{pmatrix}, \qquad \gamma_{[i]} := \begin{pmatrix}\mathbf{I}_{d_{i-1}} & \mathbf{0}\\ \mathbf{0} & -\mathbf{I}_{d_i}\end{pmatrix}, \qquad \gamma_{[i']} := \begin{pmatrix}\mathbf{I}_{d'_{i-1}} & \mathbf{0}\\ \mathbf{0} & -\mathbf{I}_{d'_i}\end{pmatrix}, \tag{22}$$

The matrix $\mathbf{\Gamma}_{[i]}$ has the property that a matrix $\mathbf{M}$ is block-diagonal if and only if $\mathbf{\Gamma}_{[i]}\mathbf{M}\mathbf{\Gamma}_{[i]} = \mathbf{M}$, and $\mathbf{M}$ is block-anti-diagonal if and only if $\mathbf{\Gamma}_{[i]}\mathbf{M}\mathbf{\Gamma}_{[i]} = -\mathbf{M}$. Since $\mathbf{D}_{[i]}$ is block-anti-diagonal, $\mathbf{\Gamma}_{[i]}\mathbf{D}_{[i]}\mathbf{\Gamma}_{[i]} = -\mathbf{D}_{[i]}$. Suppose now $\{T_k\}_k$ satisfies $T_k(-\lambda) = (-1)^k T_k(\lambda)$, which are satisfied by the monomial and Chebyshev bases.

Using the fact that $\mathbf{\Gamma}_{[i]}^2 = \mathbf{I}$, for all k we have that $\mathbf{\Gamma}_{[i]}(\mathbf{D}_{[i]})^k\mathbf{\Gamma}_{[i]} = (\mathbf{\Gamma}_{[i]}\mathbf{D}_{[i]}\mathbf{\Gamma}_{[i]})^k$. This allows us to take the conjugation with $\mathbf{\Gamma}_{[i]}$ into the argument of the polynomials.

$$\mathbf{\Gamma}_{[i]}\,T_k\big(\mathbf{D}_{[i]}\big)\,\mathbf{\Gamma}_{[i]} = T_k\big(\mathbf{\Gamma}_{[i]}\mathbf{D}_{[i]}\mathbf{\Gamma}_{[i]}\big) = T_k\big(-\mathbf{D}_{[i]}\big) = (-1)^k\,T_k\big(\mathbf{D}_{[i]}\big). \tag{23}$$

Using the previous fact, this shows that $T_k(\mathbf{D}_{[i]})$ is block-diagonal for even k and block-anti-diagonal for odd k. A simplicial signal is block-diagonal by definition, therefore it satisfies $\mathbf{\Gamma}_{[i]}\big(\mathbf{X}_{i-1}\oplus\mathbf{X}_i\big)\gamma_{[i]} = \mathbf{X}_{i-1}\oplus\mathbf{X}_i$.

Now we can decompose each coefficient into its block-diagonal and block-anti-diagonal parts, $\mathbf{\Theta}_{i,k} = \mathbf{\Theta}_{i,k}^{(0)} + \mathbf{\Theta}_{i,k}^{(1)}$. With the same argument as before, we get $\gamma'_{[i]}\mathbf{\Theta}_{i,k}^{(0)}\gamma_{[i]} = \mathbf{\Theta}_{i,k}^{(0)}$ and $\gamma'_{[i]}\mathbf{\Theta}_{i,k}^{(1)}\gamma_{[i]} = -\mathbf{\Theta}_{i,k}^{(1)}$. Inserting $\mathbf{\Gamma}_{[i]}^2 = \mathbf{I}$ and $\gamma_{[i]}^2 = \mathbf{I}$ and using equation 23, we get the following for $l \in \{0,1\}$:

$$\begin{aligned} \mathbf{\Gamma}_{[i]}\Big(T_k\big(\mathbf{D}_{[i]}\big)\big(\mathbf{X}_{i-1}\oplus\mathbf{X}_i\big)\mathbf{\Theta}_{i,k}^{(l)\top}\Big)\gamma'_{[i]} &= \mathbf{\Gamma}_{[i]}T_k\big(\mathbf{D}_{[i]}\big)\mathbf{\Gamma}_{[i]}\mathbf{\Gamma}_{[i]}\big(\mathbf{X}_{i-1}\oplus\mathbf{X}_i\big)\gamma_{[i]}\gamma_{[i]}\mathbf{\Theta}_{i,k}^{(l)\top}\gamma'_{[i]} \\ &= (-1)^k\,T_k\big(\mathbf{D}_{[i]}\big)\big(\mathbf{X}_{i-1}\oplus\mathbf{X}_i\big)(-1)^l\mathbf{\Theta}_{i,k}^{(l)\top} \\ &= (-1)^{k+l}\,T_k\big(\mathbf{D}_{[i]}\big)\big(\mathbf{X}_{i-1}\oplus\mathbf{X}_i\big)\mathbf{\Theta}_{i,k}^{(l)\top} \end{aligned} \tag{24}$$

Since we enforce the output to be in $C_{i-1}\oplus C_i$, it must be block diagonal. In equation 24, the contribution to the block-diagonals come only when $k+l$ is even, meaning for T_k, l must be odd

when k is odd (meaning only the anti-block-diagonal of $\mathbf{\Theta}_{i,k}$ contributes to the signal), and l must be even when k is even (meaning only the block-diagonal of $\mathbf{\Theta}_{i,k}$ contributes to the signal). So we can replace $\mathbf{\Theta}_{i,k}$ with its block-diagonal component for even k, and its block-anti-diagonal term for odd k, and the output is unchanged. This shows the second part of the proof.

□

Proof of Corollary 3.5. For Chebyshev polynomials T_k, we first need to show the property that $T_k(-\lambda) = (-1)^k T_k(\lambda)$. We show this by induction. For $T_0(\lambda) = 1$ and $T_1(\lambda) = \lambda$, they hold trivially. For $k \geq 2$, we get $T_k(-\lambda) = -2\lambda T_{k-1}(-\lambda) - T_{k-2}(-\lambda) = -(-1)^{k-1} 2\lambda T_{k-1}(\lambda) - (-1)^{k-2} T_{k-2}(\lambda)$. We have $-(-1)^{k-1} = (-1)^k$, and $(-1)^{k-2} = (-1)^2(-1)^{k-2} = (-1)^k$. This gives $T_k(-\lambda) = (-1)^k T_k(\lambda)$, completing the first part.

For the second part, we use the identity from (Mason & Handscomb, 2002), that $T_{2k+1}(\lambda) = \lambda V_k(2\lambda^2 - 1)$. Substituting in $\mathbf{D}_{[i]}$ gives $T_{2k+1}(\mathbf{D}_{[i]}) = \mathbf{D}_{[i]} V_k(2\mathbf{D}_{[i]}^2 - \mathbf{I}) = \mathbf{D}_{[i]} V_k(2\mathbf{L}_{[i]} - \mathbf{I})$.

□

C Proofs and Additional Details of Section 4

In this section we provide further details of the proofs in section 4. For this section we use $\mathbf{u}_j, \lambda_j$ to denote the j-th eigenvector and eigenvalue of $\mathbf{D}_{[i]}$, omitting the i index as it is fixed. Similarly, we use $\mathbf{v}_j, \theta_j$ to denote the j-th eigenvector and eigenvalue of $\mathbf{\Theta}_{[i]}$, and $\mu_\mathbf{j}$ to denote the eigenvalues of $\mathbf{L}_{[i]}$. Note that $\mu_j = \lambda_j^2$, and they share eigenvectors. We take the convention for λ_j to be in increasing order.

Furthermore, the matrix dynamics reduce to scalar ODEs, making the analysis simpler.

Lemma C.1. *Define $c_{j,k}(t), c_{j,k}^{(n)}$ as follows:*

$$c_{j,k}(t) := \mathbf{u}_j^\top \mathbf{X}_{[i]}(t) \mathbf{v}_k, \qquad c_{j,k}^{(n)} := \mathbf{u}_j^\top \mathbf{X}_{[i]}^{(n)} \mathbf{v}_k. \tag{25}$$

Then:

- *in the continuous case, $\dot{c}_{j,k}(t) = \lambda_j \theta_k \, c_{j,k}(t)$, hence $c_{j,k}(t) = c_{j,k}(0)\, e^{\lambda_j \theta_k t}$,*
- *in the discretization, $c_{j,k}^{(n+1)} = (1 + \tau \lambda_j \theta_k)\, c_{j,k}^{(n)}$, hence $c_{j,k}^{(n)} = (1 + \tau \lambda_j \theta_k)^n c_{j,k}^{(0)}$.*

Furthermore, the Frobenius norm and the i-th Dirichlet energy of the signal $\mathbf{X}_{[i]}$ is given via $\|\mathbf{X}_{[i]}\|_F^2 = \sum_{j,k} |c_{j,k}|^2$ and $\mathcal{E}_{\mathbf{L}_{[i]}}(\mathbf{X}_{[i]}) = \sum_{j,k} \mu_j |c_{j,k}|^2$ respectively. Therefore:

$$\mathcal{R}_{[\mathbf{L}_{[i]}]}(\mathbf{X}_{[i]}) = \frac{\sum_{j,k} \mu_j |c_{j,k}|^2}{\sum_{j,k} |c_{j,k}|^2} \tag{26}$$

Proof. Differentiating $c_{j,k}(t)$ gives:

$$\dot{c}_{j,k} = \mathbf{u}_j^\top \dot{\mathbf{X}}_{[i]} \mathbf{v}_k = \mathbf{u}_j^\top \mathbf{D}_{[i]} \, \mathbf{X}_{[i]} \, \mathbf{\Theta}_{[i]} \mathbf{v}_k = (\mathbf{D}_{[i]} \mathbf{u}_j)^\top \mathbf{X}_{[i]} (\mathbf{\Theta}_{[i]} \mathbf{v}_k) = \lambda_j \theta_k \, c_{j,k}. \tag{27}$$

giving the solution $c_{j,k}(t) = c_{j,k}(0)\, e^{\lambda_j \theta_k t}$. Same argument applies for the Euler discretization. For the last part, expanding $\mathbf{X}_{[i]} = \sum_{j,k} c_{j,k} \mathbf{u}_j \mathbf{v}_k^*$ gives $\|\mathbf{X}_{[i]}\|_F^2 = \sum_{j,k} |c_{j,k}|^2$, and since $\mathbf{L}_{[i]} \mathbf{u}_j = \mu_j \mathbf{u}_j$,

$$\mathcal{E}_{\mathbf{L}_{[i]}}(\mathbf{X}_{[i]}) = \mathrm{Tr}(\mathbf{X}_{[i]}^* \mathbf{L}_{[i]} \mathbf{X}_{[i]}) = \sum_{j,k} \mu_j |c_{j,k}|^2. \tag{28}$$

The result for the Rayleigh quotient follows from dividing the two.

□

Proof of Proposition 4.1. **Case 1.** Let $\mathbf{\Theta}_{[i]}^\top = -\mathbf{\Theta}_{[i]}$. Using $\dot{\mathbf{X}}_{[i]} = \mathbf{D}_{[i]} \mathbf{X}_{[i]} \mathbf{\Theta}_{[i]}$, $\mathbf{D}_{[i]}^\top = \mathbf{D}_{[i]}$ and the product rule:

$$\frac{d}{dt}\|\mathbf{X}_{[i]}\|_F^2 = \mathrm{Tr}(\dot{\mathbf{X}}_{[i]}^\top \mathbf{X}_{[i]}) + \mathrm{Tr}(\mathbf{X}_{[i]}^\top \dot{\mathbf{X}}_{[i]}) \tag{29}$$

$$= \mathrm{Tr}(\mathbf{\Theta}_{[i]}^\top \mathbf{X}_{[i]}^\top \mathbf{D}_{[i]} \mathbf{X}_{[i]}) + \mathrm{Tr}(\mathbf{X}_{[i]}^\top \mathbf{D}_{[i]} \mathbf{X}_{[i]} \mathbf{\Theta}_{[i]}) \tag{30}$$

$$= -\mathrm{Tr}(\mathbf{\Theta}_{[i]} \mathbf{X}_{[i]}^\top \mathbf{D}_{[i]} \mathbf{X}_{[i]}) + \mathrm{Tr}(\mathbf{X}_{[i]}^\top \mathbf{D}_{[i]} \mathbf{X}_{[i]} \mathbf{\Theta}_{[i]}) = 0, \tag{31}$$

where cyclicity of trace gives $\mathrm{Tr}(\mathbf{\Theta}_{[i]} \mathbf{X}_{[i]}^\top \mathbf{D}_{[i]} \mathbf{X}_{[i]}) = \mathrm{Tr}(\mathbf{X}_{[i]}^\top \mathbf{D}_{[i]} \mathbf{X}_{[i]} \mathbf{\Theta}_{[i]})$.

Since $\mathbf{L}_{[i]} = \mathbf{D}_{[i]}^2$, we have $\mathbf{D}_{[i]} \mathbf{L}_{[i]} = \mathbf{L}_{[i]} \mathbf{D}_{[i]}$. Hence:

$$\frac{d}{dt}\mathcal{E}_{\mathbf{L}_{[i]}}(\mathbf{X}_{[i]}) = \mathrm{Tr}(\dot{\mathbf{X}}_{[i]}^\top \mathbf{L}_{[i]} \mathbf{X}_{[i]}) + \mathrm{Tr}(\mathbf{X}_{[i]}^\top \mathbf{L}_{[i]} \dot{\mathbf{X}}_{[i]}) \tag{32}$$

$$= \mathrm{Tr}(\mathbf{\Theta}_{[i]}^\top \mathbf{X}_{[i]}^\top \mathbf{D}_{[i]} \mathbf{L}_{[i]} \mathbf{X}_{[i]}) + \mathrm{Tr}(\mathbf{X}_{[i]}^\top \mathbf{L}_{[i]} \mathbf{D}_{[i]} \mathbf{X}_{[i]} \mathbf{\Theta}_{[i]}) \tag{33}$$

$$= -\mathrm{Tr}(\mathbf{\Theta}_{[i]} \mathbf{X}_{[i]}^\top \mathbf{D}_{[i]} \mathbf{L}_{[i]} \mathbf{X}_{[i]}) + \mathrm{Tr}(\mathbf{X}_{[i]}^\top \mathbf{L}_{[i]} \mathbf{D}_{[i]} \mathbf{X}_{[i]} \mathbf{\Theta}_{[i]}) = 0, \tag{34}$$

where the last equality again follows from the cyclicity of trace, and that $\mathbf{L}_{[i]}$ and $\mathbf{D}_{[i]}$ commute.

Case 2. Let $\mathbf{\Theta}_{[i]}^\top = \mathbf{\Theta}_{[i]}$.

By Lemma C.1, we have $|c_{j,k}(t)|^2 = e^{2\lambda_j \theta_k t}\, a_{j,k},$. Bounding, $\lambda_j \theta_k \le |\lambda_j||\theta_k| \le \sqrt{\mu_{max}}\, \theta_{max}$ and since $\mathrm{spec}(\mathbf{D}_{[i]})$ is symmetric about 0, both $\pm\sqrt{\mu_{max}}$ occur, so for any k with $|\theta_k| = \theta_{max}$ some j attains $\lambda_j \theta_k = +\sqrt{\mu_{max}}\, \theta_{max}$ whatever the sign of θ_k. The bound is therefore attained, and equality forces $\mu_j = \mu_{max}$. Continuous time case this is

$$\mathcal{R}_{[\mathbf{L}_{[i]}]}\big(\mathbf{X}_{[i]}(t)\big) = \frac{\sum_{j,k} \mu_j\, a_{j,k}\, e^{2\lambda_j \theta_k t}}{\sum_{j,k} a_{j,k}\, e^{2\lambda_j \theta_k t}}. \tag{35}$$

Dividing numerator and denominator by $e^{2\sqrt{\mu_{max}}\theta_{max} t}$, every mode with $\lambda_j \theta_k < \sqrt{\mu_{max}}\theta_{max}$ picks up a decaying exponential and drops out, leaving only maximizing modes, all of which have $\mu_j = \mu_{max}$. $\mathcal{R}_{[\mathbf{L}_{[i]}]}(\mathbf{X}_{[i]}(t)) \to \mu_{max}$.

For the rate, write

$$\mu_{max} - \mathcal{R}_{[\mathbf{L}_{[i]}]}\big(\mathbf{X}_{[i]}(t)\big) = \frac{\sum_{j,k} (\mu_{max} - \mu_j)\, a_{j,k}\, e^{2\lambda_j \theta_k t}}{\sum_{j,k} a_{j,k}\, e^{2\lambda_j \theta_k t}}. \tag{36}$$

Now the largest exponent in the numerator becomes $2\sqrt{\mu_{max-1}}\, \theta_{max}$ by the same argument, while the denominator still grows like $e^{2\sqrt{\mu_{max}}\theta_{max} t}$. Therefore

$$\mu_{max} - \mathcal{R}_{[\mathbf{L}_{[i]}]}\big(\mathbf{X}_{[i]}(t)\big) \asymp e^{-2\theta_{max}\left(\sqrt{\mu_{max}} - \sqrt{\mu_{max-1}}\right)t}, \tag{37}$$

so the convergence rate is set by the gap between the two largest distinct eigenvalues of $\mathbf{L}_{[i]}$.

□

Now the proof for the discretization.

Proof of Proposition 4.2. By Lemma C.1, $|c_{j,k}^{(n)}|^2 = |1 + \tau\lambda_j\theta_k|^{2n} a_{j,k}$ with $a_{j,k} := |c_{j,k}^{(0)}|^2$, so in both cases the Rayleigh quotient is a ratio of sums of n-th powers, and the modes with the largest base dominate.

Case 1. Let $\mathbf{\Theta}_{[i]}^\top = -\mathbf{\Theta}_{[i]}$. Then each θ_k is purely imaginary, hence so is $\lambda_j \theta_k$, and

$$|1 + \tau\lambda_j\theta_k|^2 = 1 + \tau^2\lambda_j^2|\theta_k|^2 = 1 + \tau^2\mu_j|\theta_k|^2. \tag{38}$$

Therefore the quotient is

$$\mathcal{R}_{[\mathbf{L}_{[i]}]}\big(\mathbf{X}_{[i]}^{(n)}\big) = \frac{\sum_{j,k} \mu_j\, a_{j,k}\big(1 + \tau^2\mu_j|\theta_k|^2\big)^n}{\sum_{j,k} a_{j,k}\big(1 + \tau^2\mu_j|\theta_k|^2\big)^n}. \tag{39}$$

The base is increasing in μ_j and in $|\theta_k|$ separately, so it is maximized at $\mu_j = \mu_{max}$ and $|\theta_k| = \theta_{max}$, with value $1+\tau^2\theta_{max}^2\mu_{max}$, and the maximum is attained only by modes with $\mu_j = \mu_{max}$. Dividing through by this maximal base to the n-th power, the same argument as in the continuous case gives $\mathcal{R}_{[\mathbf{L}_{[i]}]}(\mathbf{X}_{[i]}^{(n)}) \to \mu_{max}$, with the gap controlled by the largest base among modes with $\mu_j < \mu_{max}$, namely $1 + \tau^2\theta_{max}^2\mu_{max-1}$:

$$\mu_{max} - \mathcal{R}_{[\mathbf{L}_{[i]}]}\big(\mathbf{X}_{[i]}^{(n)}\big) \asymp \left(\frac{1+\tau^2\theta_{max}^2\mu_{max-1}}{1+\tau^2\theta_{max}^2\mu_{max}}\right)^n. \tag{40}$$

Case 2. Let $\mathbf{\Theta}_{[i]}^\top = \mathbf{\Theta}_{[i]}$ Now $\lambda_j\theta_k$, is real, and

$$\mathcal{R}_{[\mathbf{L}_{[i]}]}\big(\mathbf{X}_{[i]}^{(n)}\big) = \frac{\sum_{j,k}\mu_j\, a_{j,k}\,|1+\tau\lambda_j\theta_k|^{2n}}{\sum_{j,k} a_{j,k}\,|1+\tau\lambda_j\theta_k|^{2n}}. \tag{41}$$

The value $\lambda_j\theta_k = +\sqrt{\mu_{max}}\,\theta_{max}$ is attained by the symmetry of $\mathrm{spec}(\mathbf{D}_{[i]})$ same as in the continuous case. So the maximal base is $\big(1+\tau\theta_{max}\sqrt{\mu_{max}}\big)^2$, the largest base among modes with $\mu_j < \mu_{max}$ is $\big(1+\tau\theta_{max}\sqrt{\mu_{max-1}}\big)^2$, and the same argument gives $\mathcal{R}_{[\mathbf{L}_{[i]}]}(\mathbf{X}_{[i]}^{(n)}) \to \mu_{max}$ with

$$\mu_{max} - \mathcal{R}_{[\mathbf{L}_{[i]}]}\big(\mathbf{X}_{[i]}^{(n)}\big) \asymp \left(\frac{1+\tau\theta_{max}\sqrt{\mu_{max-1}}}{1+\tau\theta_{max}\sqrt{\mu_{max}}}\right)^{2n}. \tag{42}$$

□

D Additional Details on Experiments and Implementation

The normalized root-to-leaf (co)boundary operators and Hodge Laplacians are implemented in TopoNetX (Hajij et al., 2024). For a k-simplex, the number of root paths will always be $(k+1)!$, which makes the computation more tractable, however the number of leaf paths are calculated via iterating through the maximal elements of the simplicial complex. In the datasets we consider, computation of the normalized operators is not a computational bottleneck. We leave implementations of more efficient algorithms for computing the normalized operators for later.

All experiments on trajectory prediction, mesh classification, node classification and graph classification are results over 10 train/test splits, with the best model chosen by separating 20% of the training set for validation. Learning rates and batch sizes are chosen via a hyperparameter sweep for every model. For the graph tasks, we use implementations given in Topobench (Telyatnikov et al., 2025) and TopoModelX (Hajij et al., 2024).

For trajectory prediction, we generate the irregular meshes via 3 hyperparameters: n, which defines the number of regions to be irregularized; r, which is the radius of irregularization; and k, the number of points to be added to that region. We then randomly choose n center points in the mesh, and add k points randomly chosen from within r distance of the chosen centers. For ocean$^{(150)}$, we choose $n = 3, r = 0.2, k = 50$, and for ocean$^{(600)}$, we choose $n = 6, r = 0.1, k = 100$.

For node classification tasks, we lift the node features to edge level by using the unnormalized coboundary map. In our experiments, we find lifting with the normalised coboundary operator to perform similarly. This can be due to the fact that the graphs we consider are not very irregular, and their degree variance is not high. We restrict lifting to edge-level, especially to emphasize the fact that one can use the simplest inter-rank coupling, that is the vertex-edge levels coupling in order to improve the performance. We also observe that this is not always the case, and further investigation is needed.

For PROTEINS, we lift the vertex features to edges with the coboundary matrix. For traingles, we use both random features with a fixed seed, and also the degrees of the triangles, specifically how many other triangles they are connected to via edges. We find that random lifting performs better in this dataset. This can possibly be explained by viewing the coupling as not just including higher-order information, but also a way to allow for more versatile dynamics that can carry information more efficiently. We recommend more experiments to be done to investigate this avenue.