HODGE SPECTRAL CONCENTRATION IN SIMPLICIAL COMPLEX NEURAL NETWORKS

Anonymous authors
Paper under double-blind review

ABSTRACT

Simplicial neural networks extend learning on graphs to higher-order topological structures, but feature propagation in these networks remains less well understood. We develop a spectral framework and characterize these dynamics through the Hodge decomposition. Rather than summarizing feature evolution with a single energy, as used to characterize over-smoothing and -sharpening in graph neural networks, we track the feature-normalized lower and upper Hodge Rayleigh quotients separately. This decouples spectral concentration from changes in feature magnitude and retains information about the gradient and curl subspaces, allowing us to distinguish three forms of *spectral concentration*: harmonic-dominant, gradient-dominant, and curl-dominant dynamics. For linear simplicial convolutions with independent, symmetric lower- and upper-channel mixing weights, we show that the dynamics are a gradient flow of a learnable energy and derive three competing spectral growth rates associated with the harmonic and highest-frequency gradient and curl eigenspaces. We then show that these rates determine the asymptotic regime for almost every input. Since they are computed from the learned weights and Hodge frequencies, these rates provide a post-training diagnostic of which eigenspace the normalized features concentrate in. We instantiate our framework with normalized propagation choices, including an adjacency-like form that recovers graph convolution at rank zero. Finally, experiments on synthetic edge classification and ocean-drifter trajectory prediction tasks validate the spectral predictions and illustrate how supervision and the choice of propagation operator select among the qualitatively distinct spectral concentration regimes.

1 INTRODUCTION

Modeling structured data in domains such as molecular property prediction, social networks, and knowledge bases benefits from inductive biases that inherently preserve relationships (Gilmer et al., 2017; Ying et al., 2018). While mainstay methods for relational learning represent interactions using pairwise graphs (e.g., GNNs), scholars have recently proposed numerous architectures that faithfully represent group-wise interactions through higher-order topological structures such as simplicial and cell complexes (Bunch et al., 2020; Bodnar et al., 2021; Papamarkou et al., 2024). A growing family of topological neural networks learns from signals supported on nodes, edges, faces, and higher-dimensional cells (Hajij et al., 2022; Yang et al., 2022a; Yang & Isufi, 2023).

The theoretical understanding of these higher-order architectures, however, remains considerably less mature than that of their graph counterparts. Simplicial neural networks (SNNs) propagate information using operators derived from the Hodge Laplacian, which captures interactions both below and above a given simplicial rank. For edge signals, the lowest rank at which both types of interactions occur, Hodge theory provides an orthogonal decomposition into gradient, harmonic, and curl components, representing qualitatively different forms of information flow (Lim, 2020). Recent work has begun to use this structure to analyze SNN expressivity and over-smoothing (Yang et al., 2025; Einizade et al., 2025), but there is not yet a comparably established set of definitions and analytical tools for understanding deep feature propagation in SNNs as for GNNs. This limits both our understanding of existing architectures and the principled design of higher-order networks.

Contributions. In this work, we develop a framework for understanding these dynamics. In particular, we characterize deep feature propagation in SNNs through a pair of feature-normalized

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.

lower and upper Hodge-theoretic Rayleigh quotients, whose joint behavior distinguishes harmonic-dominant (HD), gradient-dominant (GD), and curl-dominant (CD) regimes (props. 3.1 and 3.3). For linear edge-feature flows, we establish a gradient-flow formulation (thm. 3.4) and show that the depth-asymptotic analysis reduces to three competing spectral growth rates associated with the harmonic, gradient, and curl components. When the harmonic space is nontrivial and one rate is strictly largest, their comparison determines the asymptotic regime for almost every input (thm. 3.6). These rates can be computed from the trained channel weights and eigenvalues of the lower and upper Hodge Laplacians, yielding a post-training diagnostic of the Hodge structure amplified by the network. Experiments on synthetic edge classification and ocean-drifter prediction support these predictions and show how supervision and the propagation operator select among the different regimes, as illustrated by the phase portrait in fig. 2. Code will be made publicly available upon publication.

2 Preliminaries

We review the necessary background on simplicial complexes and Hodge theory, referring to Hatcher (2002) for extended background in algebraic topology and Eckmann (1944) for the classical treatment of combinatorial Hodge theory.

2.1 Simplicial Complexes and Hodge Theory

Definition 2.1 (Simplicial complex (Eckmann, 1944))**.** A finite abstract simplicial complex X on a finite vertex set V is a collection of non-empty subsets of V such that $\{v\} \in X$ for every $v \in V$ and $\varnothing \neq \tau \subseteq \sigma \in X$ implies $\tau \in X$. We omit the empty simplex. A simplex with $k+1$ vertices is called a k-simplex, or k-face; denote the set of k-faces by X_k and let $N_k = |X_k|$. In the case of $k = 1$ we can interpret the simplex as a graph.

Boundary matrices encode how simplices of adjacent ranks are incident (see fig. 1). These form the Hodge Laplacian:

Definition 2.2 (Hodge Laplacian (Hatcher, 2002))**.** For $k \geq 1$, let $\mathbf{B}_k \in \mathbb{R}^{N_{k-1} \times N_k}$ denote the boundary matrix, mapping an oriented k-simplex to its $(k-1)$-faces, with $\mathbf{B}_0 = 0$. Since $\mathbf{B}_k \mathbf{B}_{k+1} = 0$, the k-th combinatorial Hodge Laplacian is

$$\mathbf{L}_k = \underbrace{\mathbf{B}_k^\top \mathbf{B}_k}_{\mathbf{L}_k^{\text{down}}} + \underbrace{\mathbf{B}_{k+1} \mathbf{B}_{k+1}^\top}_{\mathbf{L}_k^{\text{up}}}.$$

The lower term couples k-simplices sharing a $(k-1)$-face, while the upper term couples those contained in a common $(k+1)$-simplex (fig. 1). Both are symmetric positive semidefinite and satisfy $\mathbf{L}_k^{\text{down}} \mathbf{L}_k^{\text{up}} = \mathbf{L}_k^{\text{up}} \mathbf{L}_k^{\text{down}} = 0$.

Based on the Hodge Laplacian, we can decompose a signal space $\mathbb{R}^{N_k}$ on X into distinct subspaces.

Definition 2.3 (Hodge decomposition)**.** The Hodge decomposition of the space of real signals on k-faces is the orthogonal direct sum

$$\mathbb{R}^{N_k} = \operatorname{im}(\mathbf{B}_k^\top) \oplus \ker(\mathbf{L}_k) \oplus \operatorname{im}(\mathbf{B}_{k+1}).$$

Consequently, every $\mathbf{s}^k \in \mathbb{R}^{N_k}$ decomposes uniquely as $\mathbf{s}^k = \mathbf{s}_G^k + \mathbf{s}_C^k + \mathbf{s}_H^k$ where $\mathbf{s}_G^k$, $\mathbf{s}_C^k$, and $\mathbf{s}_H^k$ are its gradient, curl, and harmonic components, respectively.

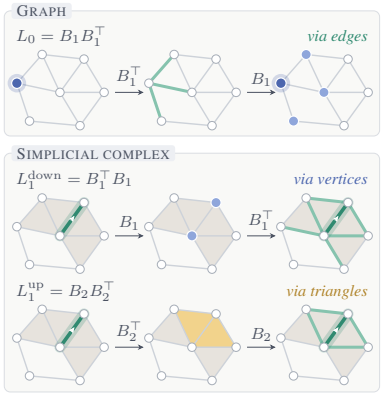


Figure 1: **Feature propagation in graphs and simplicial complexes (SCs).** On a graph, $\mathbf{L}_0 = \mathbf{B}_1 \mathbf{B}_1^\top$ sends a node signal to its incident edges and back to the neighboring nodes. On a simplicial complex, an edge signal has two channels: $\mathbf{L}_1^{\text{down}} = \mathbf{B}_1^\top \mathbf{B}_1$ couples edges that share a vertex (gradient), while $\mathbf{L}_1^{\text{up}} = \mathbf{B}_2 \mathbf{B}_2^\top$ couples edges that share a filled triangle (curl). The unfilled polygon is a hole ($\beta_1 = 1$), which supports a harmonic signal.

The kernel of $\mathbf{L}_k$ captures the *homology* over $\mathcal{X}$. $\beta_k = \dim \ker(\mathbf{L}_k)$ is the k-th Betti number, which counts independent topological features: β_0 is the number of connected components, β_1 the number of loops that do not bound filled triangles, and β_2 the number of enclosed voids.

Remark 2.4 (Simplicial frequencies)**.** Since $\mathbf{L}_k^{\text{down}}\mathbf{L}_k^{\text{up}} = \mathbf{0}$, the lower and upper Hodge Laplacians admit a common orthonormal eigenbasis. Each joint eigenvector $\mathbf{v}$, satisfying

$$\mathbf{L}_k^{\text{down}}\mathbf{v} = \nu_g\mathbf{v}, \qquad \mathbf{L}_k^{\text{up}}\mathbf{v} = \nu_c\mathbf{v},$$

has an eigenvalue pair of one of the three forms

$$\underbrace{(\nu_g, 0),\ \nu_g > 0}_{\text{gradient}} \qquad \underbrace{(0, \nu_c),\ \nu_c > 0}_{\text{curl}} \qquad \underbrace{(0,0)}_{\text{harmonic}}.$$

Thus, the nonzero eigenvalues of $\mathbf{L}_k^{\text{down}}$ and $\mathbf{L}_k^{\text{up}}$ define the gradient and curl frequency families, respectively (Yang et al., 2022b).

For edge signals ($k = 1$), the decomposition admits a physical interpretation: the gradient component is curl-free, the curl component is divergence-free, and the harmonic component is both divergence- and curl-free (Yang et al., 2022b; Lim, 2020). On a connected graph, harmonic node signals are constant. Harmonic edge signals, however, can describe circulation around holes without being constant. Whenever the corresponding Hodge subspace is nontrivial, we write $\nu_g^{\min} \le \nu_g^{\max}$ for the smallest and largest gradient frequencies, i.e. the eigenvalues of $\mathbf{L}_k^{\text{down}}$ restricted to $\text{im}(\mathbf{B}_k^\top)$, and $\nu_c^{\min} \le \nu_c^{\max}$ analogously for the curl frequencies of $\mathbf{L}_k^{\text{up}}$ restricted to $\text{im}(\mathbf{B}_{k+1})$. These frequencies are strictly positive. If $\beta_k > 0$, then $\lambda_{\min}(\mathbf{L}_k) = 0$ and the corresponding eigenspace is harmonic. If $\beta_k = 0$, then all eigenvalues of $\mathbf{L}_k$ are positive and $\lambda_{\min}(\mathbf{L}_k) = \min\{\nu_g^{\min}, \nu_c^{\min}\}$, with the minimum taken over the Hodge families that are present.

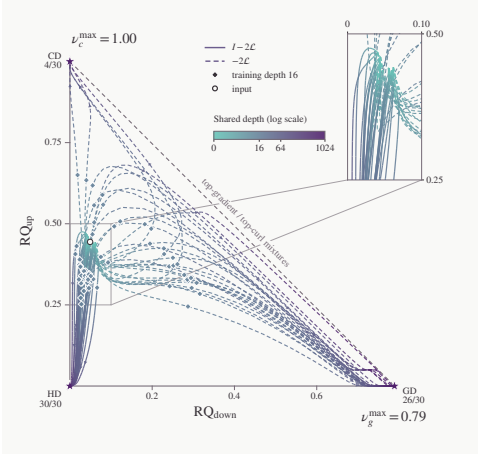


Figure 2: **Phase diagram depicting transition states during gradient flow of an SNN.** We start from an initial input and evolve over time. We plot the relative spectral concentration measured via up and down Hodge Rayleigh quotients prop. 3.1 for inference runs on an Ocean drifter Roddenberry et al. (2021) dataset show eventual convergence to one of three regimes (see prop. 3.3).

2.2 Learning on Simplicial Complexes

We study the gradient flow of a simple SNN.

Proposition 2.5 (Rankwise Simplicial Convolutional Neural Network Layer)**.** *For a fixed rank k and feature dimension d we consider a first-order layer that updates features $\mathbf{F}_m \in \mathbb{R}^{N_k \times d}$ by*

$$\mathbf{F}_{m+1} = \sigma\Big(\mathbf{F}_m + \tau\big[\mathbf{F}_m\mathbf{W}_0^{(m)} + \mathbf{L}_k^{\text{down}}\mathbf{F}_m\mathbf{W}_1^{(m)} + \mathbf{L}_k^{\text{up}}\mathbf{F}_m\mathbf{W}_2^{(m)}\big]\Big), \tag{1}$$

where m indexes layers, $\tau > 0$ is a step size, σ acts elementwise, and the learnable matrices $\mathbf{W}_i^{(m)} \in \mathbb{R}^{d\times d}$ for $i \in \{0,1,2\}$ mix channels. The three learned terms act on each k-simplex itself, on other k-simplices sharing a $(k-1)$-face, and on those sharing a $(k+1)$-coface, respectively.

This construction connects closely to widely adopted architectures in the learning literature. For instance, with combinatorial Hodge operators, matching activations and constant width, eq. (1) has the hidden-layer form of SCoNe at rank 1 (Roddenberry et al., 2021), degree-one SCNN (Yang et al., 2022a), and degree-one SCCNN without inter-rank terms (Yang & Isufi, 2023). Our theory studies the linear, symmetric, depth-shared restriction defined below. Linearity preserves every Hodge subspace, symmetric weights yield a gradient flow with real growth rates, and degree-one filters reduce the asymptotic regime to three endpoint rates (sec. 3.3). Eq. (1) is the single-rank, degree-one instance of a common layer formulation that contains all three architectures (dfn. C.1). We detail the exact correspondence in App. C.

Proposition 2.6 (Linear simplicial gradient flow)**.** *In* (1)*, set $\sigma = \text{id}$ and share symmetric weights across depth: $\mathbf{W}_i^{(m)} = \mathbf{W}_i = \mathbf{W}_i^\top$ for $i = 0,1,2$. The resulting network is exactly the forward Euler discretization, with step size τ, of*

$$\dot{\mathbf{F}} = \mathbf{F}\mathbf{W}_0 + \mathbf{L}_k^{\text{down}}\mathbf{F}\mathbf{W}_1 + \mathbf{L}_k^{\text{up}}\mathbf{F}\mathbf{W}_2, \tag{2}$$

where $\mathbf{F}\colon t \mapsto \mathbf{F}(t) \in \mathbb{R}^{N_k\times d}$ gives the features at continuous depth $t \ge 0$ (argument omitted when clear), $\dot{\mathbf{F}} = \mathrm{d}\mathbf{F}/\mathrm{d}t$, and $\mathbf{F}(0) = \mathbf{F}_0$. Layer m approximates the flow at time $t_m = m\tau$.

We call eq. (2) our *general affine model*: its propagation operator is affine in the lower and upper Laplacians, and it only depends linearly on features. Because the channel weights are symmetric, this yields a negative gradient flow (thm. 3.4). In this manner, we analyze a moderately restricted case of the network in eq. (1), with each Hodge subspace invariant. The continuous and discrete models have the same asymptotic regime under the step-size condition in prop. H.1.

3 A Hodge Spectral Concentration Framework

Before analyzing the flow in eq. (2), we first develop a general framework for spectral concentration of edge-feature dynamics that applies to any propagation on edges, including the nonlinear layer eq. (1). Throughout, we specialize to rank $k = 1$ (features on edges), the lowest rank at which both lower and upper interactions occur, and write $\mathbf{F}(t) \in \mathbb{R}^{N_1 \times d}$ for the edge features.

We also establish a notational convention that carries over to the instantiation in sec. 4.1 and the experiments. Our analysis leverages root-to-leaf normalized Laplacian $\boldsymbol{\mathcal{L}}_k^{\text{down}}, \boldsymbol{\mathcal{L}}_k^{\text{up}}$ of Viganò et al. (2026), which uniformly normalize the spectra at every rank of a simplicial complex to $[0, 1]$. This provides a common spectral scale across all ranks and different complexes, for instance, across datasets. We provide the construction of this normalized Laplacian in Apx. B and discuss the importance of normalization, as well as alternative choices in Apx. B.2. $\hat{\mathbf{B}}_1, \hat{\mathbf{B}}_2$ denote the boundary maps and may refer to combinatorial or normalized versions, subject only to $\hat{\mathbf{B}}_1\hat{\mathbf{B}}_2 = 0$, and

$$\boldsymbol{\mathcal{L}}_{\text{down}} := \hat{\mathbf{B}}_1^\top \hat{\mathbf{B}}_1, \qquad \boldsymbol{\mathcal{L}}_{\text{up}} := \hat{\mathbf{B}}_2 \hat{\mathbf{B}}_2^\top, \qquad \boldsymbol{\mathcal{L}}_1 := \boldsymbol{\mathcal{L}}_{\text{down}} + \boldsymbol{\mathcal{L}}_{\text{up}}.$$

Note that the analysis only depends on symmetry, positive semi-definiteness, and $\boldsymbol{\mathcal{L}}_{\text{down}}\boldsymbol{\mathcal{L}}_{\text{up}} = 0$.

3.1 Normalized Dirichlet Energy Definitions of Spectral Concentration

We first introduce a *pair* of Rayleigh quotients, one for each of the up and down Hodge Laplacians, with the aim of respecting the natural structure that arises from the Hodge decomposition. The pair, as opposed to a single summed quantity, retains information about the gradient and curl subspaces separately and allows us to define three distinct limiting spectral concentration regimes.

Proposition 3.1 (Hodge-theoretic Rayleigh quotients)**.** *For $\mathbf{F} \in \mathbb{R}^{N_1 \times d}$, $\mathbf{F} \neq 0$, define*

$$\mathrm{RQ}_{\text{down}}(\mathbf{F}) := \frac{\mathrm{Tr}\big(\mathbf{F}^\top \boldsymbol{\mathcal{L}}_{\text{down}}\, \mathbf{F}\big)}{\|\mathbf{F}\|^2}, \qquad \mathrm{RQ}_{\text{up}}(\mathbf{F}) := \frac{\mathrm{Tr}\big(\mathbf{F}^\top \boldsymbol{\mathcal{L}}_{\text{up}}\, \mathbf{F}\big)}{\|\mathbf{F}\|^2}, \tag{3}$$

with

$$\mathrm{RQ}(\mathbf{F}) := \mathrm{RQ}_{\text{down}}(\mathbf{F}) + \mathrm{RQ}_{\text{up}}(\mathbf{F}) = \frac{\mathrm{Tr}\big(\mathbf{F}^\top \boldsymbol{\mathcal{L}}_1 \mathbf{F}\big)}{\|\mathbf{F}\|^2}, \tag{4}$$

with $\|\cdot\|$ the Frobenius norm. $\mathrm{RQ}(\mathbf{F})$ is the Dirichlet energy normalized by the feature norm, i.e. the Rayleigh quotient of $\boldsymbol{\mathcal{L}}_1$.

These are channel-summed Rayleigh quotients of symmetric positive semi-definite operators. Whenever the corresponding Hodge family is nontrivial, $0 \leq \mathrm{RQ}_{\text{down}}(\mathbf{F}) \leq \nu_g^{\max}$, $0 \leq \mathrm{RQ}_{\text{up}}(\mathbf{F}) \leq \nu_c^{\max}$, while in all cases $\lambda_{\min}(\boldsymbol{\mathcal{L}}_1) \leq \mathrm{RQ}(\mathbf{F}) \leq \lambda_{\max}(\boldsymbol{\mathcal{L}}_1)$ (see 18).

Assumption 3.2 (Nondegeneracy)**.** Both the gradient and the curl frequency families are nontrivial, i.e. the complex has at least one edge and one triangle, and $\nu_g^{\max} > \lambda_{\min}(\boldsymbol{\mathcal{L}}_1)$ and $\nu_c^{\max} > \lambda_{\min}(\boldsymbol{\mathcal{L}}_1)$.

Proposition 3.3 (Harmonic-, gradient-, and curl-dominant dynamics)**.** *Let $t \mapsto \mathbf{F}(t) \in \mathbb{R}^{N_1 \times d} \setminus \{0\}$ be any edge-feature dynamics, in continuous time or in discrete depth (layers). Under asmp. 3.2, the dynamics is*

- *(i)* harmonic-dominant *(HD) if* $\mathrm{RQ}(\mathbf{F}(t)) \to 0$ *as* $t \to \infty$*;*
- *(ii)* gradient-dominant *(GD) if* $\mathrm{RQ}_{\text{down}}(\mathbf{F}(t)) \to \nu_g^{\max}$*;*
- *(iii)* curl-dominant *(CD) if* $\mathrm{RQ}_{\text{up}}(\mathbf{F}(t)) \to \nu_c^{\max}$.

These limits describe *spectral concentration*: the fraction of squared feature mass outside the selected eigenspace tends to zero. HD concentrates the harmonic subspace, while GD and CD concentrate in the highest-frequency gradient and curl eigenspaces, respectively. This gives a direct connection to graph over-smoothing and over-sharpening, where low- and high-frequency dominance

(LFD/HFD) describe concentration in the lowest and highest graph Laplacian eigenspaces (Di Giovanni et al., 2023) which we describe in App. A. At rank 0, the harmonic criterion recovers graph LFD, where features converge to a degree-scaled constant eigenvector. For edge features, however, harmonic dominance retains information about global topological holes, which may be helpful or harmful depending on the learning task. Gradient and curl dominance are also distinct behaviors, which do not have natural analogs in the graph case. Because each asymptotic regime corresponds to convergence to a Hodge subspace with distinct characteristics, the over-smoothing and over-sharpening dichotomy no longer applies. We henceforth use the terms HD, GD, and CD as the three generic forms of spectral concentration.

These are *model-agnostic* and apply to arbitrary continuous- or discrete-depth dynamics. The same construction extends to arbitrary rank, with any absent lower or upper Hodge family omitted at the boundary ranks. Fundamental properties of the RQ pair, including the feasible region, a bottom-eigenspace characterization of HD, and the mutual exclusion of GD and CD, are in App. D.

3.2 Energy and Gradient Flow

We now study the family of linear models with symmetric weights introduced in sec. 2.2 as the gradient flow of a parametric energy and show that Hodge subspace suppression and amplification can be induced by the positive and negative eigenvalues of the learnable weights. We then derive the exact conditions for each type of behavior. We define the parametric multi-particle energy over the edge features as

$$\mathcal{E}(\mathbf{F}) = -\tfrac{1}{2}\operatorname{Tr}\big(\mathbf{F}^\top \mathbf{F}\,\mathbf{W}_0\big) - \tfrac{1}{2}\operatorname{Tr}\big(\mathbf{F}^\top \boldsymbol{\mathcal{L}}_{\text{down}}\mathbf{F}\,\mathbf{W}_1\big) - \tfrac{1}{2}\operatorname{Tr}\big(\mathbf{F}^\top \boldsymbol{\mathcal{L}}_{\text{up}}\mathbf{F}\,\mathbf{W}_2\big). \tag{5}$$

Theorem 3.4. *If $\mathbf{W}_0$, $\mathbf{W}_1$ and $\mathbf{W}_2$ are symmetric, then the dynamical system in eq.* (2) *is the negative gradient flow of $\mathcal{E}(\mathbf{F})$, and consequently $\frac{d}{dt}\mathcal{E}(\mathbf{F}(t)) \leq 0$.*

This gradient-flow property holds for each affine (indeed polynomial) propagation operator built from symmetric Laplacians, whether normalized or not (proof in App. F). The last two terms of $\mathcal{E}$ are quadratic forms in the divergence $\hat{\mathbf{B}}_1\mathbf{F}$ at each node and in the circulation $\hat{\mathbf{B}}_2^\top\mathbf{F}$ around each 2-simplex. At a node, divergence is the signed, normalization-weighted sum of the features on its incident edges. A negative eigenvalue of $\mathbf{W}_1$ (resp. $\mathbf{W}_2$) penalizes divergence (resp. circulation) and acts attractively, while a positive one acts repulsively. App. E develops this interpretation.

3.3 Asymptotic Regimes of the Flow

The regimes of prop. 3.3 apply to any edge-feature dynamics. For the flow in eq. (2), we now show that exactly one of HD, GD, or CD occurs for almost every input and present the conditions on the channel weights and Hodge frequencies that determine each regime. Throughout this section asmp. 3.2 is in force and we assume that $\beta_1 > 0$. The case $\beta_1 = 0$ is treated in App. G.1. We first decouple the flow into a d-dimensional linear system for each joint eigenvector of $\boldsymbol{\mathcal{L}}_{\text{down}}$ and $\boldsymbol{\mathcal{L}}_{\text{up}}$, and observe that the normalized features concentrate on the fastest-growing eigenvectors of the flow.

Eigenvectors of the flow. Since $\boldsymbol{\mathcal{L}}_{\text{down}}\boldsymbol{\mathcal{L}}_{\text{up}} = 0$ and both operators are symmetric, they share an orthonormal eigenbasis $\{\mathbf{u}_j\}_{j=1}^{N_1}$,

$$\boldsymbol{\mathcal{L}}_{\text{down}}\mathbf{u}_j = \nu_j^{\text{down}}\,\mathbf{u}_j, \qquad \boldsymbol{\mathcal{L}}_{\text{up}}\mathbf{u}_j = \nu_j^{\text{up}}\,\mathbf{u}_j, \tag{6}$$

where $(\nu_j^{\text{down}}, \nu_j^{\text{up}})$ is $(\nu, 0)$, $(0, \nu)$ or $(0, 0)$ with $\nu > 0$ for a gradient, curl or harmonic eigenvector (rem. 2.4). Since the Laplacians act on the left and the weights on the right, the channel coefficients $\mathbf{c}_j := \mathbf{F}^\top\mathbf{u}_j \in \mathbb{R}^d$ of $\mathbf{F}$ along each $\mathbf{u}_j$ evolve independently,

$$\dot{\mathbf{c}}_j = \mathbf{G}_j\,\mathbf{c}_j, \qquad \mathbf{G}_j := \mathbf{W}_0 + \nu_j^{\text{down}}\mathbf{W}_1 + \nu_j^{\text{up}}\mathbf{W}_2, \tag{7}$$

with a *channel generator* $\mathbf{G}_j$, symmetric because the weights are, that depends on $\mathbf{u}_j$ only through its frequency pair. Let $\boldsymbol{\psi}_{j,r}$, $r = 1, \ldots, d$, be orthonormal eigenvectors of $\mathbf{G}_j$ with eigenvalues $\mu_{j,r}$. The $N_1 d$ rank-one matrices $\mathbf{u}_j\boldsymbol{\psi}_{j,r}^\top$, each pairing one spatial eigenvector with one channel eigenvector, form an orthonormal eigenbasis of the linear map $\mathbf{F} \mapsto \mathbf{F}\mathbf{W}_0 + \boldsymbol{\mathcal{L}}_{\text{down}}\mathbf{F}\mathbf{W}_1 + \boldsymbol{\mathcal{L}}_{\text{up}}\mathbf{F}\mathbf{W}_2$ that generates eq. (2), with eigenvalues $\mu_{j,r}$. Expanding the input in this basis, $\mathbf{F}(0) = \sum_{j,r} a_{j,r}\,\mathbf{u}_j\boldsymbol{\psi}_{j,r}^\top$ with $a_{j,r} := \mathbf{u}_j^\top\mathbf{F}(0)\boldsymbol{\psi}_{j,r}$, the solution of eq. (2) is

$$\mathbf{F}(t) = \sum_{j,r} e^{\mu_{j,r}t}\,a_{j,r}\,\mathbf{u}_j\boldsymbol{\psi}_{j,r}^\top. \tag{8}$$

The flow rescales each coefficient of the input by $e^{\mu_{j,r}t}$. Let $\mu_* := \max_{j,r} \mu_{j,r}$ be the top rate, $T_* := \{(j,r) : \mu_{j,r} = \mu_*\}$, and $E_* := \operatorname{span}\{\mathbf{u}_j \boldsymbol{\psi}_{j,r}^\top : (j,r) \in T_*\}$ the top eigenspace.

Proposition 3.5 (Normalized limit)**.** *By orthonormality, $\|\mathbf{F}(t)\|^2 = \sum_{j,r} e^{2\mu_{j,r}t} a_{j,r}^2$, and for every $\mathbf{F}(0)$ not orthogonal to E_*, i.e. with $\sum_{(j,r)\in T_*} a_{j,r}^2 > 0$, the normalized features converge to the normalized orthogonal projection of $\mathbf{F}(0)$ onto E_*:*

$$\begin{aligned} \frac{\mathbf{F}(t)}{\|\mathbf{F}(t)\|} &\longrightarrow \frac{\sum_{(j,r)\in T_*} a_{j,r}\, \mathbf{u}_j \boldsymbol{\psi}_{j,r}^\top}{\left(\sum_{(j,r)\in T_*} a_{j,r}^2\right)^{1/2}}, \\ \left(\mathrm{RQ}_{\mathrm{down}}, \mathrm{RQ}_{\mathrm{up}}\right)(\mathbf{F}(t)) &\longrightarrow \frac{\sum_{(j,r)\in T_*} a_{j,r}^2 \left(\nu_j^{\mathrm{down}}, \nu_j^{\mathrm{up}}\right)}{\sum_{(j,r)\in T_*} a_{j,r}^2}. \end{aligned} \tag{9}$$

Every eigenvector with rate below μ_* is suppressed relative to $\|\mathbf{F}(t)\|$ by a factor $e^{(\mu_{j,r}-\mu_*)t}$, so the normalized features concentrate on E_*, and the Rayleigh-quotient pair converges to the mass-weighted average of the frequency pairs in T_* (proof in App. F.2). The regime therefore corresponds to which eigenvectors $\mathbf{u}_j$ appear in E_*.

Endpoint rates. Define the three *endpoint rates*

$$\rho_H := \lambda_{\max}(\mathbf{W}_0), \qquad \rho_G := \lambda_{\max}\big(\mathbf{W}_0 + \nu_g^{\max}\mathbf{W}_1\big), \qquad \rho_C := \lambda_{\max}\big(\mathbf{W}_0 + \nu_c^{\max}\mathbf{W}_2\big), \tag{10}$$

the largest eigenvalues of the channel generators of a harmonic mode and of the top-frequency gradient and curl modes. Because $\lambda_{\max}(\mathbf{G}_j)$ is a convex function of the frequency of $\mathbf{u}_j$ within each Hodge family, and $\beta_1 > 0$ ensures all three endpoint frequencies occur, the top rate is $\mu_* = \max\{\rho_H, \rho_G, \rho_C\}$ (lem. F.1, App. F).

Theorem 3.6 (Spectral concentration trichotomy)**.** *Suppose the maximum of $\{\rho_H, \rho_G, \rho_C\}$ is attained by exactly one of them. Then, for almost every initial condition $\mathbf{F}(0)$,*

(i) if $\rho_H > \max\{\rho_G, \rho_C\}$, the flow is HD: $\left(\mathrm{RQ}_{\mathrm{down}}, \mathrm{RQ}_{\mathrm{up}}\right)(\mathbf{F}(t)) \longrightarrow (0, 0)$;

(ii) if $\rho_G > \max\{\rho_H, \rho_C\}$, the flow is GD: $\left(\mathrm{RQ}_{\mathrm{down}}, \mathrm{RQ}_{\mathrm{up}}\right)(\mathbf{F}(t)) \longrightarrow (\nu_g^{\max}, 0)$;

(iii) if $\rho_C > \max\{\rho_H, \rho_G\}$, the flow is CD: $\left(\mathrm{RQ}_{\mathrm{down}}, \mathrm{RQ}_{\mathrm{up}}\right)(\mathbf{F}(t)) \longrightarrow (0, \nu_c^{\max})$.

Here "almost every" excludes the initial conditions orthogonal to the top eigenspace E_*. The proof is in App. F.

Consequences. These results require only the trained weights and the two eigenvalues $\nu_g^{\max}, \nu_c^{\max}$, so thm. 3.6 can serve as a post-training diagnostic of the Hodge subspace selected by network to amplify. If $\mathbf{W}_1 \preceq 0$ then $\mathbf{W}_0 + \nu\mathbf{W}_1 \preceq \mathbf{W}_0$, so $\lambda_{\max}(\mathbf{W}_0 + \nu\mathbf{W}_1) \le \rho_H$ for every $\nu > 0$ and gradient dominance is impossible: GD requires a positive eigenvalue of $\mathbf{W}_1$, and likewise CD requires a positive eigenvalue of $\mathbf{W}_2$, that is, a repulsive channel direction in the sense of sec. 3.2. Ties between rates are discussed in rem. G.3. Finally, by prop. 2.6 the layer update (1) with $\sigma = \mathrm{id}$ and shared symmetric weights is the forward Euler step of (2), which multiplies the component along $\mathbf{u}_j \boldsymbol{\psi}_{j,r}^\top$ by $1 + \tau\mu_{j,r}$ per layer instead of $e^{\mu_{j,r}t}$. Let $\mu_{\min} := \min_{j,r} \mu_{j,r}$ be the smallest rate (see lem. F.1):

$$\mu_{\min} = \min\big\{\lambda_{\min}(\mathbf{W}_0),\ \lambda_{\min}(\mathbf{W}_0 + \nu_g^{\max}\mathbf{W}_1),\ \lambda_{\min}(\mathbf{W}_0 + \nu_c^{\max}\mathbf{W}_2)\big\}. \tag{11}$$

The discrete iteration is dominated by the same top-rate eigenspace, and hence has the same HD, GD or CD regime, whenever $2 + \tau(\mu_* + \mu_{\min}) > 0$; for larger steps the bottom-rate eigenspace dominates instead, with a sign alternating from layer to layer (prop. H.1, App. H). We impose this step-size restriction throughout our experiments.

4 Experiments

Our experiments aim to demonstrate, both on synthetic simplicial complexes and real-world data, that the conditions on the spectra of the learned weights and Laplacian introduced in the theoretical analysis can indeed predict the HD, GD, and CD behavior as measured by our Hodge-theoretic Rayleigh quotient definitions. A secondary, more exploratory aim is to identify whether the trained

model selects the most helpful limiting behavior for the dataset topology and supervision. We present two main sets of experiments: (1) an edge classification task using a synthetic simplicial complex, and (2) a trajectory-prediction task on a real-word dataset of ocean drifters. For the main flows of each model (i.e. before the readout layer, which differ per task), we use the flow models introduced in sec. 4.1.

4.1 Architecture and Propagation Operators

For our experiments, we define two propagation variants of the SNN in eq. (1) using normalized Laplacians and two parameter choices in its linear flow eq. (2). Both variants use two symmetric channel-mixing matrices.

We consider two parameter choices, each expressed in terms of two symmetric matrices $\mathbf{W}_{\mathrm{down}}, \mathbf{W}_{\mathrm{up}} \in \mathbb{R}^{d\times d}$. Our principal instantiation uses $\mathbf{W}_0 = \mathbf{W}_{\mathrm{down}} + \mathbf{W}_{\mathrm{up}}$, $\mathbf{W}_1 = -2\mathbf{W}_{\mathrm{down}}$, $\mathbf{W}_2 = -2\mathbf{W}_{\mathrm{up}}$, giving the adjacency analogue

$$\dot{\mathbf{F}} = \boldsymbol{\mathcal{A}}_{\mathrm{down}}\mathbf{F}\mathbf{W}_{\mathrm{down}} + \boldsymbol{\mathcal{A}}_{\mathrm{up}}\mathbf{F}\mathbf{W}_{\mathrm{up}}, \tag{12}$$

with $\boldsymbol{\mathcal{A}}_{\mathrm{down}} := \mathbf{I} - 2\boldsymbol{\mathcal{L}}_{\mathrm{down}}$ and $\boldsymbol{\mathcal{A}}_{\mathrm{up}} := \mathbf{I} - 2\boldsymbol{\mathcal{L}}_{\mathrm{up}}$. For comparison, setting $\mathbf{W}_0 = 0$, $\mathbf{W}_1 = -2\mathbf{W}_{\mathrm{down}}$, $\mathbf{W}_2 = -2\mathbf{W}_{\mathrm{up}}$ yields pure Hodge Laplacian propagation,

$$\dot{\mathbf{F}} = -2\boldsymbol{\mathcal{L}}_{\mathrm{down}}\mathbf{F}\mathbf{W}_{\mathrm{down}} - 2\boldsymbol{\mathcal{L}}_{\mathrm{up}}\mathbf{F}\mathbf{W}_{\mathrm{up}}. \tag{13}$$

The adjacency analogue operator maps the bounded frequency interval of the normalized Laplacian $[0, 1]$ onto $[-1, 1]$. So the harmonic frequency 0 has multiplier 1, while the maximal possible normalized frequency 1 has multiplier -1. With normalization, all complexes are on the same scale, allowing the regime conditions derived above to be compared without additional complex-dependent spectral rescaling. Under this choice, the feature-space generators of Section 3.3 are $\mathbf{G}^{\mathrm{harm}} = \mathbf{W}_{\mathrm{down}} + \mathbf{W}_{\mathrm{up}}$, $\mathbf{G}^{\mathrm{grad}}_{\nu} = (1-2\nu)\mathbf{W}_{\mathrm{down}} + \mathbf{W}_{\mathrm{up}}$ and $\mathbf{G}^{\mathrm{curl}}_{\nu} = \mathbf{W}_{\mathrm{down}} + (1-2\nu)\mathbf{W}_{\mathrm{up}}$, and all of the preceding energy and regime results apply directly after substituting these generators. We show recovery of graph-level analogs in App. B.3.

Implementation. All models propagate edge features at rank 1. Our principal model uses the adjacency-like flow above, discretized by forward Euler with step τ as in (1), i.e. $\mathbf{F}_{m+1} = \mathbf{F}_m + \tau\big(\boldsymbol{\mathcal{A}}_{\mathrm{down}}\mathbf{F}_m\mathbf{W}_{\mathrm{down}} + \boldsymbol{\mathcal{A}}_{\mathrm{up}}\mathbf{F}_m\mathbf{W}_{\mathrm{up}}\big)$. The pure-Laplacian flow, which differs only in the identity term $\mathbf{W}_0$, serves as a comparison. Both experiments use independent symmetric channel matrices shared across depth, without a commutation constraint. The four-disk models use eight-channel edge inputs directly, width $d = 8$, training depth 32 and step size $\tau = 0.25$, with unconstrained symmetric channel weights. For ocean drifters, a learned lift $F_0 = xa^\top$ maps the scalar edge input to $d = 16$ channels; training uses depth 16 and $\tau = 0.1$, with each channel matrix bounded in spectral norm by one. Full training details appear in App. J.3.

4.2 Spectral Edge-Role Classification

Setup. We test whether task-supervision can select the asymptotic regime of a linear flow with symmetric, depth-shared weights. On a seeded reconstruction of the four-disk complex of Grande & Schaub (2024) (1,600 vertices, 4,671 edges, 3,071 triangles, $\beta_1 = 1$), we define three edge-classification tasks with a shared, noisy feature input that mixes low- and high-frequency gradient, curl, and harmonic components. Only the labels varied across the three tasks. Following the Hodge spectral clustering of Grande & Schaub (2024), labels are k-means clusters of the harmonic eigenvector, of the top gradient eigenvectors, and of the top curl eigenvectors, respectively (construction, splits and training details in App. J.2), thereby assigning edge "roles" according to the Hodge decomposition. After training we compute the endpoint generators $\mathbf{G}_H = \mathbf{W}_{\mathrm{down}} + \mathbf{W}_{\mathrm{up}}$, $\mathbf{G}_G = (1 - 2\nu_g^{\max})\mathbf{W}_{\mathrm{down}} + \mathbf{W}_{\mathrm{up}}$ and $\mathbf{G}_C = \mathbf{W}_{\mathrm{down}} + (1 - 2\nu_c^{\max})\mathbf{W}_{\mathrm{up}}$, predict the regime from the largest of $\lambda_{\max}(\mathbf{G}_S)$, $S \in \{H, G, C\}$, and then tracking the Rayleigh-quotient pair during inference. We aim to (1) empirically validate our spectral condititions and (2) explore whether task supervision helps the model select the most "useful" subspace to amplify.

Results. The predicted regime matches the task in every case (Tab. 1): the harmonic, upper-gradient and upper-curl tasks yield η_H, η_G and η_C as the leading rate, respectively, and the Rayleigh-quotient pairs move from the first layers toward the predicted limits $(0, 0)$, $(\nu_g^{\max}, 0)$ and $(0, \nu_c^{\max})$, reaching them to three digits by depth 4,096 (Fig. 3, right), while the learned classifications track the target labels (Fig. 3, left). Since the labels are built from the same normalized Laplacian's eigenvectors that the flow uses, this is a controlled test that the three regimes are *selectable* by supervision.

Table 1: Leading eigenvalues η_S of the learned endpoint generators (seed 0); bold marks the largest per row, which gives the predicted regime. RQ: Rayleigh quotients after 4,096 Euler steps on the first held-out noise realization ($\nu_g^{\max} = 0.7413$, $\nu_c^{\max} = 0.9836$).

	Eigenvalue				RQ	
Task	η_H	η_G	η_C	Predicted	down	up
Harmonic	**1.493**	0.449	0.339	Harmonic	2.5×10^{-7}	9.6×10^{-8}
Upper-gradient	0.520	**0.726**	0.540	Gradient	0.7412	≈ 0
Upper-curl	0.600	0.385	**0.652**	Curl	≈ 0	0.9833

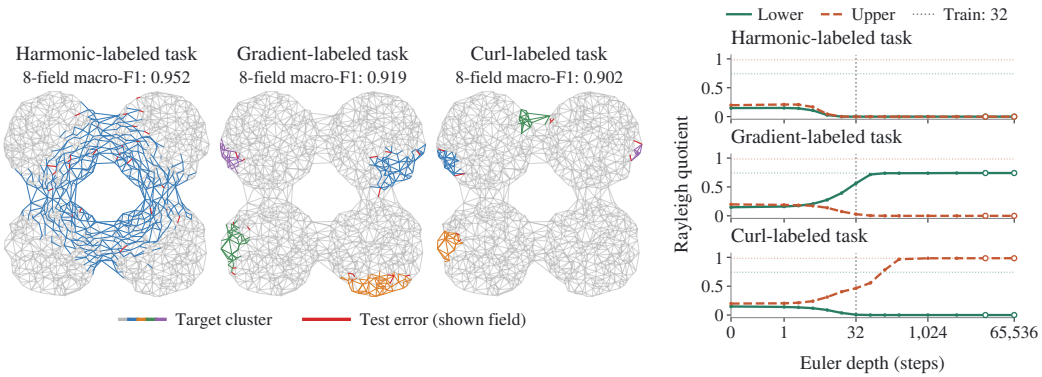


Figure 3: **(left) Target labels and prediction error for the three edge-prediction tasks**. All tasks use the same simplicial complex and input features, but labels vary according to the edge roles of harmonic, gradient, and curl. Colors identify classes within each task's target roles, with gray denoting the remainder, which a supervised class rather than unlabeled edges. **(right)** Lower and upper Rayleigh quotients converging to their limits along inference trajectories, with horizontal dotted lines indicating the corresponding maximum Laplacian eigenvalues.

4.3 Trajectory Prediction on Ocean Drifters

Setup. We use the ocean-drifter dataset of SCoNe (Roddenberry et al., 2021): 200 trajectories on a fixed oriented complex around Madagascar with 133 vertices, 320 edges and 186 triangles, whose harmonic space has dimension two (the Madagascar hole and a smaller neighboring island). Each trajectory prefix is encoded as its signed net edge-traversal vector $\mathbf{x} \in \mathbb{R}^{320}$ and the task is to predict the next vertex among the neighbors of its endpoint. We train the adjacency-like and pure-Laplacian flows of sec. 4.1 with a nonlinear decoder, then freeze each model, predict its regime from the learned generators, and extend propagation to depth 1,024 on the held-out examples, tracking the Rayleigh-quotient pair and the feature mass in the predicted eigenspace (details in App. J).

Results. The two architectures select different regimes (Fig. 4). Every adjacency-like checkpoint is predicted harmonic-dominant, and along the frozen trajectories both Rayleigh quotients decay toward zero. The pure-Laplacian flow, whose identity-free generator cannot act on the harmonic sector (App. I), is instead predicted top-gradient in 26 checkpoints and top-curl in the remaining four, and its Rayleigh quotients approach curl and grad endpoints. Because actual drifter trajectories circulate around the two holes, the task-relevant signal has a large projection onto the harmonic space (Schaub et al., 2020), which the adjacency-like flow, the only one able to amplify the harmonic subspace, converges to. This suggests that harmonic information may help for this topology.

5 Related Works

Graph convolution and over-smoothing. Deep graph neural networks may drive node representations toward low-dimensional or low-frequency subspaces, a phenomenon commonly called over-smoothing (Oono & Suzuki, 2020; Cai & Wang, 2020; Rusch et al., 2023). From a dynamical-systems perspective, linear graph convolutions can be interpreted as gradient flows whose asymptotic behavior depends jointly on graph frequencies and learned channel-mixing weights, allowing either low- or high-frequency dominance, corresponding to over-smoothing and over-sharpening (Di Giovanni et al., 2023). Other works on graph over-smoothing are summarized in App. A.

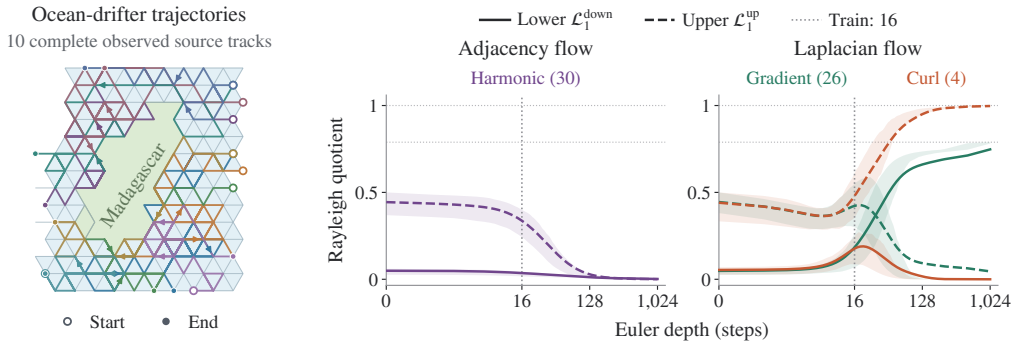


Figure 4: **(left) Sample trajectories in the ocean drifter dataset Roddenberry et al. (2021). (right) Frozen-flow spectral evolution on the ocean-drifter complex for adjacency-based and Laplacian flow**, grouped by its predicted top-gradient or top-curl endpoint. Lines indicate means, and shading indicates interquartile ranges across test occurrences. Vertical dotted lines mark training depth 16. Horizontal dotted lines in the right panel indicate the predicted limiting Rayleigh quotients. All adjacency trained flows (30) learned to amplify the harmonic subspace while Laplacian flows selected either gradient (26) or curl (4).

Topological deep learning (TDL) extends representation learning from graphs to simplicial, cellular, and combinatorial complexes (Hajij et al., 2023), enabling information exchange across cells of multiple ranks (Papamarkou et al., 2024; Hajij et al., 2022; Bodnar et al., 2021). Representative approaches include higher-order message passing, Hodge convolutions (Ebli et al., 2020; Bunch et al., 2020), sheaf or copresheaf architectures (Hansen & Gebhart, 2020; Bodnar et al., 2022; Hajij et al., 2026), generative diffusion models (Huang & Birdal, 2026), and recently neural operators on topological domains (Bastian et al., 2026). Recent work develops more sophisticated architectures and unified evaluation frameworks (Telyatnikov et al., 2024; Papillon et al., 2025).

Over-smoothing in simplicial networks. Prior work studies *simplicial over-smoothing* through decay of the unnormalized Hodge Dirichlet energy, showing that decoupled lower–upper interactions or continuous-time parameters can control this decay (Yang et al., 2025; Einizade et al., 2025). Our analysis expands these notions using feature-normalized Hodge Rayleigh quotients, separating spectral concentration from changes in feature magnitude and distinguishing asymptotic regimes.

6 Concluding Remarks

Discussion. Our results suggest that over-smoothing on simplicial complexes is most naturally understood through the Hodge decomposition. At rank one, the relevant asymptotic behavior separates into harmonic, gradient, and curl sectors, and for the linear symmetric-weight models studied here, these regimes can be predicted directly from the spectra of the learned channel-mixing generators. This provides a simple post-training diagnostic for determining which type of simplicial structure is ultimately amplified by the learned dynamics.

Conclusion. This work developed a spectral framework for characterizing feature propagation in Hodge-Laplacian-based simplicial neural networks. Using a pair of feature-normalized lower and upper Hodge Rayleigh quotients, we established three asymptotic regimes: harmonic dominance, gradient dominance, and curl dominance. For linear edge-feature flows with independent lower and upper channel mixing, we showed that the asymptotic behavior is governed by three competing spectral growth rates, which can be computed from the trained weights and Hodge eigenvalues to provide a post-training interpretability tool. Our experiments support these predictions and illustrate how supervision and the propagation operator influence the regime selected at depth.

Limitations and future work. Our main theoretical results are deliberately limited to linear SNN dynamics with depth-shared, symmetric channel-mixing operators. We also focus on asymptotic results, leaving characterization of finite-depth behavior to future work. In particular, analyzing our spectral framework as a dynamical system could recast the HD, GD, and CD limits as attractors, allowing us to study their stability, basins of attraction, and transitions, and potentially providing insight into finite-depth dynamics and the underlying mechanisms driving SNNs to these regimes. Finally, extending our results from simplicial complexes to more general cell complexes would broaden the framework to a larger class of higher-order relational domains.

AI use statement. In this work, we used generative AI tools for language editing, code editing, debugging, validating our synthetic data simulation, and visualization. All AI-assisted text and code were reviewed by the authors. The authors take full responsibility for the final content.

Ethics and reproducibility. Our work is theoretical in nature; we thus do not anticipate any direct adverse consequences from the insights into the spectral concentration of SNNs. Experiments use public data, thus no human subjects or personal data are involved. Training protocols are documented in App. J; all implementations will be published upon acceptance.

REFERENCES

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

Cristian Bodnar, Fabrizio Frasca, Yuguang Wang, Nina Otter, Guido F Montufar, Pietro Lió, and Michael Bronstein. Weisfeiler and lehman go topological: Message passing simplicial networks. In Marina Meila and Tong Zhang (eds.), *Proceedings of the 38th International Conference on Machine Learning*, volume 139 of *Proceedings of Machine Learning Research*, pp. 1026–1037. PMLR, 18–24 Jul 2021.

Cristian Bodnar, Francesco Di Giovanni, Benjamin Chamberlain, Pietro Lio, and Michael Bronstein. Neural sheaf diffusion: A topological perspective on heterophily and oversmoothing in gnns. *Advances in Neural Information Processing Systems*, 35:18527–18541, 2022.

Eric Bunch, Qian You, Glenn Fung, and Vikas Singh. Simplicial 2-complex convolutional neural networks. In *NeurIPS 2020 Workshop on Topological Data Analysis and Beyond*, 2020. URL https://openreview.net/forum?id=TLbnsKrt6J-.

Chen Cai and Yusu Wang. A note on over-smoothing for graph neural networks. *arXiv preprint arXiv:2006.13318*, 2020. doi: 10.48550/arXiv.2006.13318. URL https://arxiv.org/abs/2006.13318.

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.

Stefania Ebli, Michaël Defferrard, and Gard Spreemann. Simplicial neural networks. *arXiv preprint arXiv:2010.03633*, 2020. doi: 10.48550/arXiv.2010.03633. URL https://arxiv.org/abs/2010.03633.

Beno Eckmann. Harmonische Funktionen und Randwertaufgaben in einem Komplex. *Commentarii Mathematici Helvetici*, 17(1):240–255, 1944. doi: 10.1007/BF02566245.

Aref Einizade, Dorina Thanou, Fragkiskos D. Malliaros, and Jhony H. Giraldo. Continuous simplicial neural networks. In *The Thirty-ninth Annual Conference on Neural Information Processing Systems*, 2025. URL https://openreview.net/forum?id=fPPfFMVTZo.

Bastian Epping, Alexandre René, Moritz Helias, and Michael T. Schaub. Graph neural networks do not always oversmooth. In *Advances in Neural Information Processing Systems*, volume 37, pp. 48164–48188, 2024. URL https://proceedings.neurips.cc/paper_files/paper/2024/hash/5623c35f3ab5e2c72aeb3abce27dc28f-Abstract-Conference.html.

Justin Gilmer, Samuel S. Schoenholz, Patrick F. Riley, Oriol Vinyals, and George E. Dahl. Neural message passing for quantum chemistry. In *Proceedings of the 34th International Conference on Machine Learning*, volume 70 of *Proceedings of Machine Learning Research*, pp. 1263–1272. PMLR, 2017. URL https://proceedings.mlr.press/v70/gilmer17a.html.

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

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, 2022. URL https://arxiv.org/abs/2206.00606.

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, 2023.

Mustafa Hajij, Lennart Bastian, Sarah Osentoski, Hardik Kabaria, John Davenport, Balaji Cherukuri, Joseph Kocheemoolayil, Nastaran Shahmansouri, Adrian Lew, Theodore Papamarkou, et al. Copresheaf topological neural networks: A generalized deep learning framework. *Advances in Neural Information Processing Systems*, 38:149759–149803, 2026.

Jakob Hansen and Thomas Gebhart. Sheaf neural networks. *arXiv preprint arXiv:2012.06333*, 2020.

Allen Hatcher. *Algebraic Topology*. Cambridge University Press, 2002. URL https://pi.math.cornell.edu/~hatcher/AT/ATpage.html.

Yiming Huang and Tolga Birdal. Hog-diff: Higher-order guided diffusion for graph generation. In *International Conference on Learning Representations*, volume 2026, pp. 111572–111614, 2026.

Nicolas Keriven. Not too little, not too much: a theoretical analysis of graph (over) smoothing. *Advances in Neural Information Processing Systems*, 35:2268–2281, 2022.

Thomas N. Kipf and Max Welling. Semi-supervised classification with graph convolutional networks. In *International Conference on Learning Representations (ICLR)*, 2017. URL https://arxiv.org/abs/1609.02907.

Qimai Li, Zhichao Han, and Xiao-Ming Wu. Deeper insights into graph convolutional networks for semi-supervised learning. In *Proceedings of the AAAI Conference on Artificial Intelligence*, 2018. URL https://arxiv.org/abs/1801.07606.

Lek-Heng Lim. Hodge laplacians on graphs. *SIAM Review*, 62(3):685–715, 2020. doi: 10.1137/18M1223101. URL https://arxiv.org/abs/1507.05379.

Kenta Oono and Taiji Suzuki. Graph neural networks exponentially lose expressive power for node classification. In *International Conference on Learning Representations*, 2020. URL https://openreview.net/forum?id=S1ldO2EFPr.

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, Guillermo Bernardez, Claudio Battiloro, and Nina Miolane. Topotune: A framework for generalized combinatorial complex neural networks. In *Forty-second International Conference on Machine Learning*, 2025. URL https://openreview.net/forum?id=S5njonQdBf.

T. Mitchell Roddenberry, Nicholas Glaze, and Santiago Segarra. Principled simplicial neural networks for trajectory prediction. In *Proceedings of the 38th International Conference on Machine Learning*, volume 139 of *Proceedings of Machine Learning Research*, pp. 9020–9029. PMLR, 2021. URL https://proceedings.mlr.press/v139/roddenberry21a.html.

T. Konstantin Rusch, Michael M. Bronstein, and Siddhartha Mishra. A survey on oversmoothing in graph neural networks. *arXiv preprint arXiv:2303.10993*, 2023. doi: 10.48550/arXiv.2303.10993. 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, 2020.

Diaaeldin Taha, James Chapman, Marzieh Eidi, Karel Devriendt, and Guido Montufar. Demystifying topological message-passing with relational structures: A case study on oversquashing in simplicial message-passing. In *The Thirteenth International Conference on Learning Representations*, 2025. URL https://openreview.net/forum?id=QC2qE1tcmd.

Lev Telyatnikov, Guillermo Bernardez, Marco Montagna, Mustafa Hajij, Martin Carrasco, Pavlo Vasylenko, Mathilde Papillon, Ghada Zamzmi, Michael T Schaub, Jonas Verhellen, et al. Topobench: A framework for benchmarking topological deep learning. *arXiv preprint arXiv:2406.06642*, 2024.

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. *arXiv preprint arXiv:2604.27241*, 2026.

Maosheng Yang and Elvin Isufi. Convolutional learning on simplicial complexes. *arXiv preprint arXiv:2301.11163*, 2023. doi: 10.48550/arXiv.2301.11163. URL https://arxiv.org/abs/2301.11163.

Maosheng Yang, Elvin Isufi, and Geert Leus. Simplicial convolutional neural networks. In *2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 8847–8851. IEEE, 2022a. doi: 10.1109/ICASSP43922.2022.9746017.

Maosheng Yang, Elvin Isufi, Michael T. Schaub, and Geert Leus. Simplicial convolutional filters. *IEEE Transactions on Signal Processing*, 70:4633–4648, 2022b. doi: 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.

Rex Ying, Ruining He, Kaifeng Chen, Pong Eksombatchai, William L. Hamilton, and Jure Leskovec. Graph convolutional neural networks for web-scale recommender systems. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 974–983, 2018. doi: 10.1145/3219819.3219890. URL https://arxiv.org/abs/1806.01973.

Kaicheng Zhang, Piero Deidda, Desmond Higham, and Francesco Tudisco. Are we measuring oversmoothing in graph neural networks correctly? In *International Conference on Learning Representations*, 2026. URL https://openreview.net/forum?id=r3DOISnvHD.

Zhijian Zhuo, Yifei Wang, Jinwen Ma, and Yisen Wang. Graph neural networks (with proper weights) can escape oversmoothing. In Vu Nguyen and Hsuan-Tien Lin (eds.), *Proceedings of the 16th Asian Conference on Machine Learning*, volume 260 of *Proceedings of Machine Learning Research*, pp. 17–32. PMLR, 2025. URL https://proceedings.mlr.press/v260/zhuo25a.html.

OVERVIEW OF THE APPENDIX

In the appendix, we provide additional background, the deferred proofs, and extensions of the results in the main paper.

A Background: over-smoothing and over-sharpening on graphs . . . 13

B The root-to-leaf normalization . . . 15
- B.1 Definition and basic properties . . . 15
- B.2 Why root-to-leaf normalization? . . . 15
- B.3 Recovery of the graph setting . . . 16

C Connections to existing simplicial architectures . . . 16
- C.1 A common layer formulation . . . 16
- C.2 SCoNe, SCNN, and SCCNN . . . 17
- C.3 Restriction to the analyzed flow . . . 17

D Properties of the Rayleigh-quotient pair . . . 18
- D.1 Eigenspace characterizations of the three regimes . . . 18
- D.2 Feasible region and mutual exclusion . . . 19
- D.3 Why HD is over-smoothing, and GD and CD are not . . . 19

E Energy, attraction and repulsion . . . 20

F Proofs of the results in Section 3 . . . 22
- F.1 Proof of Theorem 3.4 . . . 22
- F.2 Proof of Proposition 3.5 . . . 23
- F.3 Endpoint reduction . . . 23
- F.4 Proof of Theorem 3.6 . . . 24

G Extensions of the spectral concentration trichotomy . . . 25
- G.1 The case $\beta_1 = 0$: no harmonic anchor . . . 25
- G.2 Commuting channel weights . . . 26
- G.3 Ties between growth rates . . . 26

H Discretization of the flow . . . 27

I Adjacency-like versus Laplacian propagation . . . 29

J Experimental details . . . 30
- J.1 Datasets and input signals . . . 30
- J.2 Four-disk target labels . . . 30
- J.3 Models and training . . . 30
- J.4 Spectral diagnostics . . . 31

A BACKGROUND: OVER-SMOOTHING AND OVER-SHARPENING ON GRAPHS

Much research has been devoted to understanding message passing in neural networks on graph-structured data, and more recently on simplicial and cellular complexes (Taha et al., 2025). In Graph Convolutional Networks (GCNs), node features are transformed by a shared, learned mapping and propagated across edges (Kipf & Welling, 2017). Despite their success on various tasks involving graph-structured data, GCNs tend to struggle on *heterophilic graphs* and suffer from *over-smoothing* as the number of layers in the neural network increases. Heterophily (homophily) is a property of labeled graphs that measures the tendency for nodes with different (the same) labels to share edges. Over-smoothing refers to node features becoming increasingly indistinguishable with depth because each propagation step acts as a form of Laplacian smoothing (Li et al., 2018), so that features converge exponentially to a subspace determined by the connected components of the graph (Oono & Suzuki, 2020) and their Dirichlet energy decays to zero (Cai & Wang, 2020; Rusch et al., 2023). Subsequent works refine and further characterize these dynamics. In particular, Keriven (2022) show that a moderate amount of smoothing can be beneficial before it becomes harmful; Epping et al. (2024) show, in the infinite-width (Gaussian-process) limit, that GCNs initialized with sufficiently large weight variance enter a non-over-smoothing phase in which node features remain informative at depth; and Zhuo et al. (2025) attribute over-smoothing to inadequately learned weights rather than to the architecture itself, showing that suitable weight matrices allow vanilla GNNs to avoid it even after infinitely many propagation steps. Other measures of over-smoothing, such as the numerical rank of the feature matrix, have also been proposed Zhang et al. (2026). Di Giovanni et al. (2023) investigate these behaviors through a dynamical systems framework, where node features evolve

according to the gradient flow of an energy functional related to the Dirichlet energy. We restate their framework here.

The graph Laplacian and adjacency. For a graph $G = (V, E)$ with normalized adjacency $\tilde{\mathbf{A}}$, the normalized graph Laplacian is $\mathbf{\Delta} = \mathbf{I} - \tilde{\mathbf{A}}$. Low frequencies of $\mathbf{\Delta}$ correspond to smooth signals over the graph, while high frequencies correspond to rapid variation between neighboring nodes. Many graph-based message-passing architectures use these operators to propagate information across edges; GCNs in particular use the normalized adjacency.

Dirichlet energy and gradient flows. The graph Dirichlet energy measures the smoothness of the feature signal across edges:

$$E_{\mathrm{Dir}}(\mathbf{F}) = \mathrm{trace}(\mathbf{F}^\top \mathbf{\Delta} \mathbf{F}).$$

Since low graph frequencies correspond to smoother signals, graph convolutions are often seen as low-pass filters that progressively reduce the Dirichlet energy and smooth features. To quantify and study this behavior, Di Giovanni et al. (2023) introduce a parametric energy functional incorporating symmetric learnable channel-mixing weights $\mathbf{\Omega}, \mathbf{W}$:

$$E_\theta(\mathbf{F}) = \sum_i \langle \mathbf{f}_i, \mathbf{\Omega} \mathbf{f}_i \rangle - \sum_{i,j} A_{ij} \langle \mathbf{f}_i, \mathbf{W} \mathbf{f}_j \rangle + \phi^0(\mathbf{F}, \mathbf{F}(0)).$$

Positive eigenvalues of $\mathbf{W}$ induce attractive forces (smoothing), while negative eigenvalues induce repulsive forces (sharpening). Graph convolutional networks can be interpreted as gradient flows of this functional:

$$\dot{\mathbf{F}}(t) = -\mathbf{F}(t)\mathbf{\Omega} + \mathbf{A}\mathbf{F}(t)\mathbf{W} - \mathbf{F}(0)\tilde{\mathbf{W}}.$$

The above energy formulation allows for both over-smoothing and the opposite behavior, over-sharpening, which is noteworthy since GCNs were thought to only exhibit over-smoothing.

Over-smoothing and over-sharpening: LFD and HFD. For completeness, we restate the criteria of Di Giovanni et al. (2023, Def. 4.2). For nonzero features $\mathbf{F}(t)$, with t denoting continuous depth or a discrete layer index, *low-frequency dominance (LFD)* is characterized by

$$\lim_{t \to \infty} \frac{E_{\mathrm{Dir}}(\mathbf{F}(t))}{\|\mathbf{F}(t)\|_F^2} = 0.$$

Conversely, *high-frequency dominance (HFD)* indicates over-sharpening:

$$\lim_{t \to \infty} \frac{E_{\mathrm{Dir}}(\mathbf{F}(t))}{\|\mathbf{F}(t)\|_F^2} = \lambda_{n-1},$$

where $\| \cdot \|_F$ is the Frobenius norm and λ_{n-1} is the largest eigenvalue of $\mathbf{\Delta}$. These criteria give the spectral characterizations of over-smoothing and over-sharpening discussed in sec. 3.1; at rank 0 on a connected graph, LFD coincides with the harmonic-dominant regime of prop. 3.3.

For the simplified continuous-time graph convolution considered by Di Giovanni et al.,

$$\dot{\mathbf{F}}(t) = \mathbf{A}\mathbf{F}(t)\mathbf{W}, \qquad \mathbf{A} = \mathbf{D}^{-1/2}\mathbf{A}_{\mathrm{graph}}\mathbf{D}^{-1/2} = \mathbf{I} - \mathbf{\Delta},$$

assume that the graph is simple, undirected, and connected, and that the channel-mixing matrix $\mathbf{W}$ is symmetric, with eigenvalues

$$\mu_0 < \mu_1 \leq \cdots \leq \mu_{d-1}.$$

Let

$$0 = \lambda_0 \leq \cdots \leq \lambda_{n-1}$$

denote the eigenvalues of the normalized graph Laplacian $\mathbf{\Delta}$. Then, for almost every initial feature matrix $\mathbf{F}(0)$,

$$|\mu_0|(\lambda_{n-1} - 1) > \mu_{d-1}$$

implies HFD dynamics, whereas

$$|\mu_0|(\lambda_{n-1} - 1) < \mu_{d-1}$$

implies LFD dynamics (Di Giovanni et al., 2023, Thm. 4.3). App. B.3 shows that this dichotomy is the rank-0 case of our trichotomy under the root-to-leaf normalization.

B THE ROOT-TO-LEAF NORMALIZATION

Most existing simplicial convolutional architectures are analyzed with unnormalized Hodge Laplacians, whereas the graph convolutional networks studied by Kipf & Welling (2017) and Di Giovanni et al. (2023) propagate with the symmetrically normalized adjacency. Since there is no canonical choice of normalization for higher-order Laplacians, we choose one that preserves the Hodge structure on which the analysis of sec. 3 rests: the root-to-leaf normalization of Viganò et al. (2026), which produces normalized Hodge Laplacians whose spectra are bounded uniformly across all complexes, the higher-order analogue of the $[0, 2]$ bound for the normalized graph Laplacian.

B.1 DEFINITION AND BASIC PROPERTIES

For a k-face σ, the *leaf-path function* $\mathrm{LP}(\sigma)$ is the number of ascending chains $\sigma = \sigma_k \subset \sigma_{k+1} \subset \cdots \subset \sigma_m$ of faces of consecutive dimensions that end at a maximal face of X, and the *root-path function* $\mathrm{RP}(\sigma)$ is the number of descending chains from σ to a vertex; since such a chain removes one vertex at a time, $\mathrm{RP}(\sigma) = (k+1)!$ on every k-face. Equivalently, $\mathrm{LP}(\sigma) = \sum_\phi (\dim \phi - \dim \sigma)!$, the sum running over the maximal faces $\phi \supseteq \sigma$ (Viganò et al., 2026, Lemma 2.1). We fix for each dimension k the diagonal, strictly positive weight matrix $\mathbf{M}_k \in \mathbb{R}^{N_k \times N_k}$ (denoted $\mathbf{W}_k$ in Viganò et al. (2026); we write $\mathbf{M}_k$ to avoid a clash with the channel weights) whose (σ, σ) entry is the ratio of the leaf-path to the root-path function,

$$(\mathbf{M}_k)_{\sigma\sigma} = H(\sigma) = \frac{\mathrm{LP}(\sigma)}{\mathrm{RP}(\sigma)} = \frac{\mathrm{LP}(\sigma)}{(k+1)!}.$$

We define the *normalized* boundary operators by conjugating the combinatorial boundaries with the square roots of these weights,

$$\hat{\mathbf{B}}_k = \mathbf{M}_{k-1}^{-1/2}\, \mathbf{B}_k\, \mathbf{M}_k^{1/2},$$

equivalently the normalized coboundary operators $\delta_{k-1} = \hat{\mathbf{B}}_k^\top = \mathbf{M}_k^{1/2}\, \partial_{k-1}^*\, \mathbf{M}_{k-1}^{-1/2}$ of Viganò et al. (2026), with $\partial_{k-1}^* = \mathbf{B}_k^\top$.

Hodge preservation.. When multiplying boundary matrices, the inner weight matrices cancel and the combinatorial homology is preserved:

$$\hat{\mathbf{B}}_k \hat{\mathbf{B}}_{k+1} = \big(\mathbf{M}_{k-1}^{-1/2} \mathbf{B}_k \mathbf{M}_k^{1/2}\big)\big(\mathbf{M}_k^{-1/2} \mathbf{B}_{k+1} \mathbf{M}_{k+1}^{1/2}\big) = \mathbf{M}_{k-1}^{-1/2} (\mathbf{B}_k \mathbf{B}_{k+1}) \mathbf{M}_{k+1}^{1/2} = 0.$$

The *normalized* Hodge Laplacians are then

$$\boldsymbol{\mathcal{L}}_k^{\mathrm{down}} = \hat{\mathbf{B}}_k^\top \hat{\mathbf{B}}_k = \mathbf{M}_k^{1/2} \mathbf{B}_k^\top \mathbf{M}_{k-1}^{-1} \mathbf{B}_k \mathbf{M}_k^{1/2},$$

$$\boldsymbol{\mathcal{L}}_k^{\mathrm{up}} = \hat{\mathbf{B}}_{k+1} \hat{\mathbf{B}}_{k+1}^\top = \mathbf{M}_k^{-1/2} \mathbf{B}_{k+1} \mathbf{M}_{k+1} \mathbf{B}_{k+1}^\top \mathbf{M}_k^{-1/2},$$

which inherit properties from combinatorial Hodge theory:

- **(NH1)** $\boldsymbol{\mathcal{L}}_k^{\mathrm{down}}$ and $\boldsymbol{\mathcal{L}}_k^{\mathrm{up}}$ are symmetric and positive semi-definite; and $\hat{\mathbf{B}}_k \hat{\mathbf{B}}_{k+1} = 0$ gives $\boldsymbol{\mathcal{L}}_k^{\mathrm{down}} \boldsymbol{\mathcal{L}}_k^{\mathrm{up}} = 0$, so they are simultaneously diagonalizable, admit a normalized Hodge decomposition, and have harmonic space isomorphic to the k-th cohomology (Viganò et al., 2026).
- **(NH5)** Their spectra are bounded uniformly over all complexes by 1: $\mathrm{Spec}(\boldsymbol{\mathcal{L}}_k^{\mathrm{down}}), \mathrm{Spec}(\boldsymbol{\mathcal{L}}_k^{\mathrm{up}}) \subseteq [0, 1]$ (Viganò et al., 2026).

B.2 WHY ROOT-TO-LEAF NORMALIZATION?

Normalization fixes the spectral scale of the Hodge operators, allowing frequency responses and asymptotic thresholds to be compared across complexes. We use the root-to-leaf normalization of Viganò et al. (2026) because it preserves the basic Hodge structure (in particular, $\hat{\mathbf{B}}_k \hat{\mathbf{B}}_{k+1} = 0$, the Hodge decomposition, and the nonzero spectral correspondence between adjacent up- and down-Laplacians) while producing symmetric positive-semidefinite operators with spectra uniformly contained in $[0, 1]$. Random-walk normalizations such as that of Schaub et al. (2020) instead arise

naturally in a degree-weighted geometry (and admit symmetric representatives after a change of coordinates). The root-to-leaf construction is therefore more convenient here due to its symmetry, which lets us use standard Euclidean gradient flows and Rayleigh quotients for our analysis directly. Because the spectra of all ranks lie in the same interval $[0, 1]$, the regime conditions of thm. 3.6 and the limits of the Rayleigh-quotient pair are moreover on a common scale across ranks and across complexes; with unnormalized operators the spectra of different ranks and different complexes need not be comparable, so a notion of dominance "at maximal frequency" would have no complex-independent meaning.

B.3 Recovery of the graph setting

When the simplicial complex is a connected graph $G = (V, E)$ viewed as a one-dimensional complex, every edge is a maximal face, so $\mathrm{LP}(e) = 1$ and $\mathrm{LP}(v) = \deg(v)$, and the root-to-leaf weights are

$$\mathbf{M}_0 = \mathbf{D}, \qquad \mathbf{M}_1 = \tfrac{1}{2}\mathbf{I},$$

with $\mathbf{D}$ the degree matrix. Hence

$$\boldsymbol{\mathcal{L}}_0^{\mathrm{up}} = \hat{\mathbf{B}}_1\hat{\mathbf{B}}_1^{\top} = \tfrac{1}{2}\mathbf{D}^{-1/2}\mathbf{B}_1\mathbf{B}_1^{\top}\mathbf{D}^{-1/2} = \tfrac{1}{2}\boldsymbol{\Delta}, \qquad \boldsymbol{\Delta} = \mathbf{I} - \mathbf{D}^{-1/2}\mathbf{A}_{\mathrm{graph}}\mathbf{D}^{-1/2},$$

is one half of the symmetric normalized graph Laplacian, and the adjacency-like operator of sec. 4.1 is

$$\boldsymbol{\mathcal{A}}_0^{\mathrm{up}} = \mathbf{I} - 2\boldsymbol{\mathcal{L}}_0^{\mathrm{up}} = \mathbf{I} - \boldsymbol{\Delta} = \mathbf{D}^{-1/2}\mathbf{A}_{\mathrm{graph}}\mathbf{D}^{-1/2}.$$

At rank 0 there is no lower term, since $\mathbf{B}_0 = 0$, so the adjacency-like flow of sec. 4.1 has a single channel matrix $\mathbf{W} := \mathbf{W}_{0,\mathrm{up}}$ (in the notation of sec. 3.3, $\mathbf{W}_0 = \mathbf{W}$ and $\mathbf{W}_2 = -2\mathbf{W}$) and reads

$$\dot{\mathbf{F}}_0 = \mathbf{D}^{-1/2}\mathbf{A}_{\mathrm{graph}}\mathbf{D}^{-1/2}\,\mathbf{F}_0\,\mathbf{W},$$

which is exactly the simplified graph convolution studied by Di Giovanni et al. (2023) and restated in App. A.

The spectral conditions reduce accordingly. If $0 = \lambda_0 \le \cdots \le \lambda_{n-1}$ are the eigenvalues of $\boldsymbol{\Delta}$, the root-to-leaf frequencies of the rank-0 upper family are $\nu = \lambda/2$, so that $1 - 2\nu = 1 - \lambda$, and the channel generator (7) of a mode of frequency ν is $\mathbf{W}_0 + \nu\mathbf{W}_2 = (1 - 2\nu)\mathbf{W} = (1 - \lambda)\mathbf{W}$. The harmonic generator is therefore $\mathbf{W}$ and the top-frequency generator is $(1 - \lambda_{n-1})\mathbf{W}$. Since G is connected with $n \ge 2$ nodes, $\lambda_{n-1} > 1$: the eigenvalues of $\boldsymbol{\Delta}$ sum to $\mathrm{Tr}\,\boldsymbol{\Delta} = n$ while $\lambda_0 = 0$, so the largest is at least $n/(n-1)$. Writing $\mu_0 < \mu_1 \le \cdots \le \mu_{d-1}$ for the eigenvalues of $\mathbf{W} \neq 0$, the two endpoint rates (10) are therefore

$$\rho_H = \lambda_{\max}(\mathbf{W}) = \mu_{d-1}, \qquad \rho_{\mathrm{top}} = \lambda_{\max}\big((1 - \lambda_{n-1})\mathbf{W}\big) = (\lambda_{n-1} - 1)\,(-\mu_0).$$

If $\mu_0 \ge 0$ then $\rho_{\mathrm{top}} \le 0 < \rho_H$ and the flow is LFD. If $\mu_0 < 0$ then $\rho_{\mathrm{top}} = |\mu_0|(\lambda_{n-1} - 1)$, and the comparison of $|\mu_0|(\lambda_{n-1} - 1)$ with μ_{d-1} recovers the LFD/HFD dichotomy of Di Giovanni et al. (2023, Thm. 4.3): on a connected graph the harmonic space of $\boldsymbol{\Delta}$ is $\mathrm{span}(\mathbf{D}^{1/2}\mathbf{1})$, the degree-scaled constant signal (constant only for regular graphs), so harmonic dominance is LFD, and dominance of the top upper frequency is HFD.

C Connections to existing simplicial architectures

We compare hidden propagation layers using the combinatorial boundary maps and Hodge Laplacians of sec. 2. When comparing architectures, we assume that activations and feature sizes are matched when ignoring task-specific prediction heads. Normalized variants require compatible boundary maps and symmetric lower and upper Laplacians whose product is zero, which the root-to-leaf operators of App. B provide.

C.1 A common layer formulation

Definition C.1 (Simplicial neural network layer)**.** Let $\mathbf{F}_k^{(m)} \in \mathbb{R}^{N_k \times d_m}$ be the rank-k features at layer m, σ an elementwise activation, and $T_d, T_u \ge 0$ filter degrees. Omit the layer index on the

weights, and consider

$$\mathbf{F}_k^{(m+1)} = \sigma\Big(\mathbf{F}_k^{(m)}\mathbf{W}_{k,0} + \sum_{r=1}^{T_d}(\mathbf{L}_k^{\mathrm{down}})^r\mathbf{F}_k^{(m)}\mathbf{W}_{k,d,r} + \sum_{s=1}^{T_u}(\mathbf{L}_k^{\mathrm{up}})^s\mathbf{F}_k^{(m)}\mathbf{W}_{k,u,s} + \sum_{r=0}^{T_d}(\mathbf{L}_k^{\mathrm{down}})^r\mathbf{B}_k^\top\mathbf{F}_{k-1}^{(m)}\mathbf{V}_{k,d,r} + \sum_{s=0}^{T_u}(\mathbf{L}_k^{\mathrm{up}})^s\mathbf{B}_{k+1}\mathbf{F}_{k+1}^{(m)}\mathbf{V}_{k,u,s}\Big), \tag{14}$$

with learnable channel matrices $\mathbf{W}, \mathbf{V} \in \mathbb{R}^{d_m \times d_{m+1}}$. Nonexistent rank inputs are omitted.

The first line filters features within rank k; the second filters adjacent-rank features after boundary projection, which needs only one-sided filters since $\mathbf{L}_k^{\mathrm{up}}\mathbf{B}_k^\top = 0$ and $\mathbf{L}_k^{\mathrm{down}}\mathbf{B}_{k+1} = 0$. On a fixed complex, any spectral filter $\phi(\mathbf{L}_k^{\mathrm{down}}, \mathbf{L}_k^{\mathrm{up}})$ equals $c\mathbf{I} + P(\mathbf{L}_k^{\mathrm{down}}) + Q(\mathbf{L}_k^{\mathrm{up}})$ for polynomials P, Q without constant term (interpolate on the joint spectrum), so the polynomial form loses no generality.

The main-text layer in eq. (1) is the case $\mathbf{V} = 0$, $T_d = T_u = 1$, constant width d, and

$$\mathbf{W}_{k,0} = \mathbf{I}_d + \tau\mathbf{W}_0^{(m)}, \qquad \mathbf{W}_{k,d,1} = \tau\mathbf{W}_1^{(m)}, \qquad \mathbf{W}_{k,u,1} = \tau\mathbf{W}_2^{(m)}.$$

Since $\tau > 0$, this reparameterization is invertible, so every single-rank, degree-one instance with square channel matrices is of the form eq. (1). The correspondence requires neither symmetric nor depth-shared weights.

C.2 SCoNe, SCNN, and SCCNN

SCoNe. The edge hidden layer of Roddenberry et al. (2021, Eq. (7)) is the case $k = 1, T_d = T_u = 1$, $\mathbf{V} = 0$, and hence, precisely an instance of eq. (1). Their weights $\mathbf{W}_1, \mathbf{W}_0, \mathbf{W}_2$ correspond to our identity, lower, and upper coefficients, respectively. Their final node readout is separate; with identity activation it loses triangle dependence, as their Eq. (9) shows.

SCNN. The filter bank of Yang et al. (2022a, Eqs. (1)–(4)) is the case $\mathbf{V} = 0$, with each channel matrix collecting the corresponding scalar filter coefficients across input–output channel pairs. At rank 1, degree one recovers SCoNe's hidden-layer form.

SCCNN. The multichannel layer of Yang & Isufi (2023, Eq. (B.1)) is the full form (14); its two intra-rank zeroth-order weights sum to $\mathbf{W}_{k,0}$, and setting $\mathbf{V} = 0$ recovers SCNN. The same architecture is presented in Yang et al. (2025, Eqs. (4)–(5)).

C.3 Restriction to the analyzed flow

Taking identity activation and symmetric channel weights shared across depth in (1) gives

$$\mathbf{F}_{m+1} = \mathbf{F}_m + \tau\big(\mathbf{F}_m\mathbf{W}_0 + \mathbf{L}_k^{\mathrm{down}}\mathbf{F}_m\mathbf{W}_1 + \mathbf{L}_k^{\mathrm{up}}\mathbf{F}_m\mathbf{W}_2\big), \tag{15}$$

the forward Euler step of eq. (2). These restrictions are additional to the source architectures. Continuous growth is governed by generator eigenvalues; discrete growth by the magnitudes of the Euler multipliers. The condition in prop. H.1 ensures the same dominant eigenspace.

The broader polynomial family. Higher-degree filters lead to the fixed-rank linear flow

$$\dot{\mathbf{F}} = \sum_{i=0}^{p}\mathbf{A}_i\mathbf{F}\mathbf{W}_i, \qquad \mathbf{A}_i = \alpha_i\mathbf{I} + P_i(\mathbf{L}_k^{\mathrm{down}}) + Q_i(\mathbf{L}_k^{\mathrm{up}}), \qquad \mathbf{W}_i = \mathbf{W}_i^\top, \tag{16}$$

with real scalars α_i, polynomials P_i, Q_i having zero constant term, and channel weights fixed across depth. Equation (2) is the case $p = 2$ with $\mathbf{A}_0 = \mathbf{I}$, $\mathbf{A}_1 = \mathbf{L}_k^{\mathrm{down}}$, and $\mathbf{A}_2 = \mathbf{L}_k^{\mathrm{up}}$. More generally, each $\mathbf{A}_i$ is diagonal in the joint eigenbasis: on frequency (ν_d, ν_u) its multiplier is $\alpha_i + P_i(\nu_d) + Q_i(\nu_u)$. The decoupling into one d-dimensional system per joint eigenmode uses only this diagonal structure, so it holds for any linear layer of this form, even with non-symmetric or depth-dependent weights. The gradient-flow property (thm. 3.4) and the normalized limit (prop. 3.5) additionally use the symmetry of the weights, and carry over verbatim. The endpoint reduction of lem. F.1 requires both symmetry and degree one, and fails without either: the spectral abscissa of

non-symmetric weights need not be convex in the frequency, and higher-degree responses can peak at an interior frequency (e.g. $\dot{\mathbf{F}} = 2a\,\mathbf{L}_k^{\mathrm{down}}\mathbf{F} - (\mathbf{L}_k^{\mathrm{down}})^2\mathbf{F}$ selects the gradient frequency closest to a). A pointwise nonlinearity breaks the Hodge-subspace invariance itself, so none of these results apply to it directly.

D Properties of the Rayleigh-quotient pair

Throughout this appendix $\boldsymbol{\mathcal{L}}_{\mathrm{down}}$, $\boldsymbol{\mathcal{L}}_{\mathrm{up}}$ and $\boldsymbol{\mathcal{L}}_1$ are the rank-1 operators of sec. 3, and asmp. 3.2 is in force. Write

$$\Sigma_g := \mathrm{Spec}(\boldsymbol{\mathcal{L}}_{\mathrm{down}}) \setminus \{0\}, \qquad \Sigma_c := \mathrm{Spec}(\boldsymbol{\mathcal{L}}_{\mathrm{up}}) \setminus \{0\}$$

for the sets of gradient and curl frequencies, and fix the common orthonormal eigenbasis $\{\mathbf{u}_j\}_{j=1}^{N_1}$ of (6) with frequency pairs $(\nu_j^{\mathrm{down}}, \nu_j^{\mathrm{up}})$. For a nonzero feature matrix $\mathbf{F}$, let $\mathbf{c}_j := \mathbf{F}^\top \mathbf{u}_j \in \mathbb{R}^d$ be its channel coefficients along $\mathbf{u}_j$, so that $\mathbf{F} = \sum_j \mathbf{u}_j \mathbf{c}_j^\top$, and define the *mode masses*

$$m_j(\mathbf{F}) := \frac{\|\mathbf{c}_j\|_2^2}{\|\mathbf{F}\|_F^2} \geq 0, \qquad \sum_{j=1}^{N_1} m_j(\mathbf{F}) = 1, \tag{17}$$

where the normalization follows from orthonormality of the $\mathbf{u}_j$. Since $\boldsymbol{\mathcal{L}}_{\mathrm{down}}\mathbf{F} = \sum_j \nu_j^{\mathrm{down}}\,\mathbf{u}_j\mathbf{c}_j^\top$ and $\boldsymbol{\mathcal{L}}_{\mathrm{up}}\mathbf{F} = \sum_j \nu_j^{\mathrm{up}}\,\mathbf{u}_j\mathbf{c}_j^\top$, the same expansion gives

$$\mathrm{RQ}_{\mathrm{down}}(\mathbf{F}) = \sum_j \nu_j^{\mathrm{down}}\, m_j(\mathbf{F}), \mathrm{RQ}_{\mathrm{up}}(\mathbf{F}) = \sum_j \nu_j^{\mathrm{up}}\, m_j(\mathbf{F}), \mathrm{RQ}(\mathbf{F}) = \sum_j \big(\nu_j^{\mathrm{down}} + \nu_j^{\mathrm{up}}\big)\, m_j(\mathbf{F}), \tag{18}$$

so that each Rayleigh quotient is the mass-weighted average of the corresponding frequencies. Both lemmas below are direct consequences of (18).

D.1 Eigenspace characterizations of the three regimes

Each regime of prop. 3.3 is defined through the limit of a scalar Rayleigh quotient. The following lemma shows that each such limit is equivalent to concentration of the normalized feature mass on a single eigenspace, and quantifies the equivalence. Let $\mathbf{P}_{\mathrm{harm}}$ be the orthogonal projector onto $\mathcal{H}_1$, and let $\mathbf{P}_g$ and $\mathbf{P}_c$ be the orthogonal projectors onto the top gradient and top curl eigenspaces $\ker(\boldsymbol{\mathcal{L}}_{\mathrm{down}} - \nu_g^{\max}\mathbf{I})$ and $\ker(\boldsymbol{\mathcal{L}}_{\mathrm{up}} - \nu_c^{\max}\mathbf{I})$. Define the spectral gaps

$$\gamma_H := \min\big(\Sigma_g \cup \Sigma_c\big),$$
$$\gamma_g := \nu_g^{\max} - \max\big(\{\nu \in \Sigma_g : \nu < \nu_g^{\max}\} \cup \{0\}\big),$$
$$\gamma_c := \nu_c^{\max} - \max\big(\{\nu \in \Sigma_c : \nu < \nu_c^{\max}\} \cup \{0\}\big),$$

all of which are strictly positive.

Lemma D.1 (Eigenspace characterizations of HD, GD and CD)**.** *Let $\beta_1 > 0$, so that $\lambda_{\min}(\boldsymbol{\mathcal{L}}_1) = 0$ with eigenspace $\mathcal{H}_1$. For every $\mathbf{F} \neq 0$,*

$$\gamma_H \frac{\|(\mathbf{I} - \mathbf{P}_{\mathrm{harm}})\mathbf{F}\|_F^2}{\|\mathbf{F}\|_F^2} \;\leq\; \mathrm{RQ}(\mathbf{F}) \;\leq\; \lambda_{\max}(\boldsymbol{\mathcal{L}}_1)\, \frac{\|(\mathbf{I} - \mathbf{P}_{\mathrm{harm}})\mathbf{F}\|_F^2}{\|\mathbf{F}\|_F^2}, \tag{19}$$

$$\gamma_g \frac{\|(\mathbf{I} - \mathbf{P}_g)\mathbf{F}\|_F^2}{\|\mathbf{F}\|_F^2} \;\leq\; \nu_g^{\max} - \mathrm{RQ}_{\mathrm{down}}(\mathbf{F}) \;\leq\; \nu_g^{\max}\, \frac{\|(\mathbf{I} - \mathbf{P}_g)\mathbf{F}\|_F^2}{\|\mathbf{F}\|_F^2}, \tag{20}$$

and the analogue of (20) *holds for $\nu_c^{\max} - \mathrm{RQ}_{\mathrm{up}}(\mathbf{F})$ with γ_c and $\mathbf{P}_c$. In particular $\mathrm{RQ}(\mathbf{F}) \geq 0$, with equality if and only if every column of $\mathbf{F}$ lies in $\mathcal{H}_1$, and a dynamics $\mathbf{F}(t) \neq 0$ is HD, GD or CD if and only if, respectively,*

$$\frac{\|(\mathbf{I} - \mathbf{P}_{\mathrm{harm}})\mathbf{F}(t)\|_F}{\|\mathbf{F}(t)\|_F} \longrightarrow 0, \qquad \frac{\|(\mathbf{I} - \mathbf{P}_g)\mathbf{F}(t)\|_F}{\|\mathbf{F}(t)\|_F} \longrightarrow 0, \qquad \frac{\|(\mathbf{I} - \mathbf{P}_c)\mathbf{F}(t)\|_F}{\|\mathbf{F}(t)\|_F} \longrightarrow 0.$$

Proof. By (18), $\mathrm{RQ}(\mathbf{F}) = \sum_j \lambda_j m_j$ with $\lambda_j := \nu_j^{\mathrm{down}} + \nu_j^{\mathrm{up}}$ the eigenvalue of $\boldsymbol{\mathcal{L}}_1$ on $\mathbf{u}_j$. Since $(\mathbf{I} - \mathbf{P}_{\mathrm{harm}})\mathbf{F} = \sum_{j:\lambda_j>0} \mathbf{u}_j \mathbf{c}_j^\top$, orthonormality gives $\|(\mathbf{I} - \mathbf{P}_{\mathrm{harm}})\mathbf{F}\|_F^2 / \|\mathbf{F}\|_F^2 = \sum_{j:\lambda_j>0} m_j$, and every $\lambda_j > 0$ lies in $[\gamma_H, \lambda_{\max}(\boldsymbol{\mathcal{L}}_1)]$ from (19). For (20), write $\nu_g^{\max} - \mathrm{RQ}_{\mathrm{down}}(\mathbf{F}) = \sum_j (\nu_g^{\max} - \nu_j^{\mathrm{down}})\, m_j$. The coefficient vanishes when $\mathbf{u}_j$ lies in the top gradient eigenspace, and for every other j it lies in $[\gamma_g, \nu_g^{\max}]$, because ν_j^{down} is then either 0 (harmonic or curl mode) or a gradient frequency below $\nu_g^{\max}$; summing as before gives (20). The curl case is identical. The equality case and the three equivalences follow by letting the sandwiched quantities tend to zero. □

Remark D.2 (Bottom eigenspace for general β_1)**.** The proof of (19) uses only that $\mathcal{H}_1$ is the eigenspace of the smallest eigenvalue of $\boldsymbol{\mathcal{L}}_1$. For arbitrary β_1, let $E_{\min} := \ker(\boldsymbol{\mathcal{L}}_1 - \lambda_{\min}(\boldsymbol{\mathcal{L}}_1)\mathbf{I})$ with orthogonal projector $\mathbf{P}_{\min}$, and let $\gamma_{\min} := \min\{\lambda \in \mathrm{Spec}(\boldsymbol{\mathcal{L}}_1) : \lambda > \lambda_{\min}(\boldsymbol{\mathcal{L}}_1)\} - \lambda_{\min}(\boldsymbol{\mathcal{L}}_1)$. The same argument gives

$$\gamma_{\min} \frac{\|(\mathbf{I} - \mathbf{P}_{\min})\mathbf{F}\|_F^2}{\|\mathbf{F}\|_F^2} \le \mathrm{RQ}(\mathbf{F}) - \lambda_{\min}(\boldsymbol{\mathcal{L}}_1) \le \big(\lambda_{\max}(\boldsymbol{\mathcal{L}}_1) - \lambda_{\min}(\boldsymbol{\mathcal{L}}_1)\big) \frac{\|(\mathbf{I} - \mathbf{P}_{\min})\mathbf{F}\|_F^2}{\|\mathbf{F}\|_F^2},$$

so $\mathrm{RQ}(\mathbf{F}(t)) \to \lambda_{\min}(\boldsymbol{\mathcal{L}}_1)$ if and only if the normalized mass concentrates on $E_{\min}$. This is the form used for $\beta_1 = 0$ in App. G.1.

D.2 Feasible region and mutual exclusion

Lemma D.3 (Feasible polygon; mutual exclusion of the regimes)**.** *For every* $\mathbf{F} \neq 0$*,*

$$\big(\mathrm{RQ}_{\mathrm{down}}(\mathbf{F}), \mathrm{RQ}_{\mathrm{up}}(\mathbf{F})\big) = \sum_{j=1}^{N_1} m_j(\mathbf{F}) \left(\nu_j^{\mathrm{down}}, \nu_j^{\mathrm{up}}\right) \in \mathcal{P}$$

with

$$\mathcal{P} := \mathrm{conv}\Big(\{(\nu, 0) : \nu \in \Sigma_g\} \cup \{(0, \nu) : \nu \in \Sigma_c\} \cup \{(0,0) \text{ if } \beta_1 > 0\}\Big).$$

If $\beta_1 > 0$*,* $\mathcal{P}$ *is the triangle with vertices* $(0,0)$*,* $(\nu_g^{\max}, 0)$*,* $(0, \nu_c^{\max})$*; if* $\beta_1 = 0$*, it is the convex hull of* $(\nu_g^{\min}, 0)$*,* $(\nu_g^{\max}, 0)$*,* $(0, \nu_c^{\min})$*,* $(0, \nu_c^{\max})$*. In either case*

$$\frac{\mathrm{RQ}_{\mathrm{down}}(\mathbf{F})}{\nu_g^{\max}} + \frac{\mathrm{RQ}_{\mathrm{up}}(\mathbf{F})}{\nu_c^{\max}} \le 1, \tag{21}$$

equivalently

$$\mathrm{RQ}_{\mathrm{up}}(\mathbf{F}) \le \frac{\nu_c^{\max}}{\nu_g^{\max}} \big(\nu_g^{\max} - \mathrm{RQ}_{\mathrm{down}}(\mathbf{F})\big) \qquad \text{and} \qquad \mathrm{RQ}_{\mathrm{down}}(\mathbf{F}) \le \frac{\nu_g^{\max}}{\nu_c^{\max}} \big(\nu_c^{\max} - \mathrm{RQ}_{\mathrm{up}}(\mathbf{F})\big).$$

Consequently, GD implies $(\mathrm{RQ}_{\mathrm{down}}, \mathrm{RQ}_{\mathrm{up}}) \to (\nu_g^{\max}, 0)$*, CD implies* $(\mathrm{RQ}_{\mathrm{down}}, \mathrm{RQ}_{\mathrm{up}}) \to (0, \nu_c^{\max})$*, and the regimes HD, GD and CD are pairwise mutually exclusive.*

Proof. The identity (18) exhibits the pair as a convex combination of the points $(\nu_j^{\mathrm{down}}, \nu_j^{\mathrm{up}})$, each of which is $(\nu, 0)$ with $\nu \in \Sigma_g$, $(0, \nu)$ with $\nu \in \Sigma_c$, or $(0,0)$, the last only if $\beta_1 > 0$, so are elements of $\mathcal{P}$. For every j at most one of $\nu_j^{\mathrm{down}}, \nu_j^{\mathrm{up}}$ is nonzero, so $\nu_j^{\mathrm{down}}/\nu_g^{\max} + \nu_j^{\mathrm{up}}/\nu_c^{\max} \le 1$; averaging against the masses gives (21), and multiplying by $\nu_c^{\max}$ or by $\nu_g^{\max}$ gives the two equivalent forms. If the dynamics is GD then $\mathrm{RQ}_{\mathrm{down}} \to \nu_g^{\max}$, hence $\mathrm{RQ}_{\mathrm{up}} \to 0$ by the first form; the CD case is symmetric, and the two limit points are distinct, so GD and CD exclude each other. Finally, GD forces $\mathrm{RQ} \to \nu_g^{\max} > 0$ and CD forces $\mathrm{RQ} \to \nu_c^{\max} > 0$, whereas HD requires $\mathrm{RQ} \to 0$. □

D.3 Why HD is over-smoothing, and GD and CD are not

Why HD recovers over-smoothing on graphs By lem. D.1, HD means that the normalized feature mass concentrates on the harmonic space $\mathcal{H}_1$, the eigenspace of the smallest eigenvalue of $\boldsymbol{\mathcal{L}}_1$: asymptotically the normalized divergence $\hat{\mathbf{B}}_1 \mathbf{F}(t)/\|\mathbf{F}(t)\|_F$ and curl $\hat{\mathbf{B}}_2^\top \mathbf{F}(t)/\|\mathbf{F}(t)\|_F$ both vanish, and every spectral component above the lowest Hodge frequency is suppressed relative to $\|\mathbf{F}(t)\|_F$. This is the direct analogue of low-frequency dominance on graphs (App. A), where the

lowest eigenspace of the symmetric normalized Laplacian is spanned by the degree-scaled constant $\mathbf{D}^{1/2}\mathbf{1}$ and the limiting node features therefore carry no information beyond node degree. The difference lies in what the limit looks like: a harmonic edge signal is not determined by local degrees and may retain global circulation determined by the topology of the complex, so HD erases local divergence and curl structure while preserving the β_1 topological modes. When $\beta_1 = 0$ there is no harmonic space and HD cannot occur; App. G.1 shows that the flow then concentrates on one of the four endpoint frequencies instead.

Why GD and CD split the notion of over-sharpening GD and CD mean that the features remain dominated by the most oscillatory mode of the gradient or of the curl family: divergence, respectively curl, persists relative to $\|\mathbf{F}(t)\|_F$ at maximal frequency. There is an exact correspondence with the graph case: $\boldsymbol{\mathcal{L}}_{\text{down}} = \hat{\mathbf{B}}_1^\top \hat{\mathbf{B}}_1$ and $\boldsymbol{\mathcal{L}}_0^{\text{up}} = \hat{\mathbf{B}}_1 \hat{\mathbf{B}}_1^\top$ share their nonzero spectra, so gradient dominance means that the normalized features align with the coboundary lift of the top graph frequency, that is, over-sharpening in the sense of Di Giovanni et al. (2023) seen from edges rather than nodes. Curl dominance is its face-level counterpart and has no rank-0 analogue. Thm. 3.6 shows that for the flow (2) only the three endpoint limits occur generically.

E Energy, attraction and repulsion

We now use the physical particle attraction and repulsion interpretation of the gradient flow of the parametric energy to explain why both smoothing and sharpening are possible. Recall

$$\boldsymbol{\mathcal{L}}_{\text{down}} = \hat{\mathbf{B}}_1^\top \hat{\mathbf{B}}_1, \qquad \boldsymbol{\mathcal{L}}_{\text{up}} = \hat{\mathbf{B}}_2 \hat{\mathbf{B}}_2^\top,$$

with $\hat{\mathbf{B}}_1$ and $\hat{\mathbf{B}}_2$ the normalized boundary operators. We interpret each edge-feature vector $\mathbf{f}_e \in \mathbb{R}^d$ as a particle in feature space. Unlike on a graph, however, these particles interact through two different higher-order structures: node stars and 2-simplices.

Starting from the energy

$$\mathcal{E}(\mathbf{F}) = \underbrace{-\frac{1}{2}\mathrm{Tr}\big(\mathbf{F}^\top \mathbf{F}\, \mathbf{W}_0\big)}_{\text{external term}} \underbrace{-\frac{1}{2}\mathrm{Tr}\big(\mathbf{F}^\top \boldsymbol{\mathcal{L}}_{\text{down}} \mathbf{F}\mathbf{W}_1\big)}_{\text{node term}} \underbrace{-\frac{1}{2}\mathrm{Tr}\big(\mathbf{F}^\top \boldsymbol{\mathcal{L}}_{\text{up}} \mathbf{F}\mathbf{W}_2\big)}_{\text{face term}},$$

the external term is independent of the simplicial incidence structure. Defining

$$\mathbf{W}_{\text{ext}} := \mathbf{W}_0,$$

we have

$$\mathrm{Tr}(\mathbf{F}^\top \mathbf{F}\mathbf{W}_{\text{ext}}) = \sum_e \langle \mathbf{f}_e, \mathbf{W}_{\text{ext}} \mathbf{f}_e \rangle,$$

and therefore

$$\mathcal{E}_{\text{ext}}(\mathbf{F}) = -\frac{1}{2} \sum_e \langle \mathbf{f}_e, \mathbf{W}_{\text{ext}} \mathbf{f}_e \rangle.$$

This is the analogue of the graph-independent external energy in Di Giovanni et al. (2023).

We begin by looking at the node term, corresponding to divergence. Insert $\boldsymbol{\mathcal{L}}_{\text{down}} = \hat{\mathbf{B}}_1^\top \hat{\mathbf{B}}_1$ and regroup:

$$\mathrm{Tr}\big(\mathbf{F}^\top \hat{\mathbf{B}}_1^\top \hat{\mathbf{B}}_1 \mathbf{F}\mathbf{W}_1\big) = \mathrm{Tr}\big((\hat{\mathbf{B}}_1\mathbf{F})^\top (\hat{\mathbf{B}}_1\mathbf{F})\mathbf{W}_1\big).$$

Define the normalized divergence field

$$\operatorname{div} \mathbf{F} := \hat{\mathbf{B}}_1 \mathbf{F} \in \mathbb{R}^{N_0 \times d},$$

whose row at node v is

$$\mathbf{g}_v := (\hat{\mathbf{B}}_1\mathbf{F})_{v,:} = \sum_e (\hat{\mathbf{B}}_1)_{ve} \mathbf{f}_e \in \mathbb{R}^d.$$

Only edges incident to v contribute to this sum. Thus $\mathbf{g}_v$ is the signed, normalization-weighted aggregate of the edge features in the star of v, or equivalently the normalized net in/outflow at that node, channel by channel. Applying the same trace identity gives

$$\mathrm{Tr}\big(\mathbf{F}^\top \boldsymbol{\mathcal{L}}_{\text{down}} \mathbf{F}\mathbf{W}_1\big) = \sum_v \langle \mathbf{g}_v, \mathbf{W}_1 \mathbf{g}_v \rangle.$$

Hence the node contribution is

$$\mathcal{E}_{\mathrm{down}}(\mathbf{F}) = -\frac{1}{2}\sum_{v}\langle \mathbf{g}_v, \mathbf{W}_1\mathbf{g}_v\rangle.$$

We now look at the face term, corresponding to curl. Similarly, using $\boldsymbol{\mathcal{L}}_{\mathrm{up}} = \hat{\mathbf{B}}_2\hat{\mathbf{B}}_2^\top$,

$$\mathrm{Tr}\big(\mathbf{F}^\top\hat{\mathbf{B}}_2\hat{\mathbf{B}}_2^\top\mathbf{F}\mathbf{W}_2\big) = \mathrm{Tr}\big((\hat{\mathbf{B}}_2^\top\mathbf{F})^\top(\hat{\mathbf{B}}_2^\top\mathbf{F})\mathbf{W}_2\big).$$

Define the normalized curl field

$$\mathrm{curl}\,\mathbf{F} := \hat{\mathbf{B}}_2^\top\mathbf{F} \in \mathbb{R}^{N_2\times d},$$

whose row at a 2-simplex $\sigma \in X_2$ is

$$\mathbf{h}_\sigma := (\hat{\mathbf{B}}_2^\top\mathbf{F})_{\sigma,:} = \sum_{e}(\hat{\mathbf{B}}_2)_{e\sigma}\mathbf{f}_e \in \mathbb{R}^d.$$

The nonzero coefficients occur on the three oriented edges of σ, so $\mathbf{h}_\sigma$ is the signed, normalization-weighted circulation of the feature signal around that face, channel by channel. Therefore

$$\mathrm{Tr}\big(\mathbf{F}^\top\boldsymbol{\mathcal{L}}_{\mathrm{up}}\mathbf{F}\mathbf{W}_2\big) = \sum_{\sigma\in X_2}\langle \mathbf{h}_\sigma, \mathbf{W}_2\mathbf{h}_\sigma\rangle,$$

and the face contribution is

$$\mathcal{E}_{\mathrm{up}}(\mathbf{F}) = -\frac{1}{2}\sum_{\sigma\in X_2}\langle \mathbf{h}_\sigma, \mathbf{W}_2\mathbf{h}_\sigma\rangle.$$

Decomposition by Hodge sector. Write $\mathbf{F} = \mathbf{F}_H + \mathbf{F}_G + \mathbf{F}_C$ for the column-wise orthogonal projections of $\mathbf{F}$ onto the harmonic, gradient and curl subspaces. Because the three subspaces are mutually orthogonal, $\mathbf{F}_s^\top\mathbf{F}_{s'} = 0$ for $s \neq s'$, so the external term splits as $\mathcal{E}_{\mathrm{ext}}(\mathbf{F}) = \mathcal{E}_{\mathrm{ext}}(\mathbf{F}_H) + \mathcal{E}_{\mathrm{ext}}(\mathbf{F}_G) + \mathcal{E}_{\mathrm{ext}}(\mathbf{F}_C)$. Moreover $\hat{\mathbf{B}}_1\mathbf{F} = \hat{\mathbf{B}}_1\mathbf{F}_G$, since $\hat{\mathbf{B}}_1$ vanishes on $\mathcal{H}_1$ and on $\mathrm{im}\,\hat{\mathbf{B}}_2$, and likewise $\hat{\mathbf{B}}_2^\top\mathbf{F} = \hat{\mathbf{B}}_2^\top\mathbf{F}_C$. Hence

$$\mathcal{E}(\mathbf{F}) = \mathcal{E}_{\mathrm{ext}}(\mathbf{F}_H) + \big[\mathcal{E}_{\mathrm{ext}}(\mathbf{F}_G) + \mathcal{E}_{\mathrm{down}}(\mathbf{F}_G)\big] + \big[\mathcal{E}_{\mathrm{ext}}(\mathbf{F}_C) + \mathcal{E}_{\mathrm{up}}(\mathbf{F}_C)\big] :$$

the energy is a sum of three energies, one per Hodge sector; the node term sees only the gradient part and the face term sees only the curl part. This is the energetic counterpart of the decoupling of the flow into independent joint eigenmodes in sec. 3.3.

Attraction and repulsion. We now split each channel-mixing matrix according to the signs of its eigenvalues. As in Di Giovanni et al. (2023, Sec. 4.2.2), for $s \in \{1, 2\}$, symmetry allows us to write

$$\mathbf{W}_s = \boldsymbol{\Theta}_{s,+}^\top\boldsymbol{\Theta}_{s,+} - \boldsymbol{\Theta}_{s,-}^\top\boldsymbol{\Theta}_{s,-},$$

so that, for any $\mathbf{z} \in \mathbb{R}^d$,

$$\langle \mathbf{z}, \mathbf{W}_s\mathbf{z}\rangle = \|\boldsymbol{\Theta}_{s,+}\mathbf{z}\|_2^2 - \|\boldsymbol{\Theta}_{s,-}\mathbf{z}\|_2^2.$$

Substituting these decompositions into the energy gives

$$\begin{aligned}\mathcal{E}(\mathbf{F}) = &-\frac{1}{2}\sum_{e}\langle \mathbf{f}_e, \mathbf{W}_{\mathrm{ext}}\mathbf{f}_e\rangle \\ &-\frac{1}{2}\sum_{v}\left(\|\boldsymbol{\Theta}_{1,+}\mathbf{g}_v\|_2^2 - \|\boldsymbol{\Theta}_{1,-}\mathbf{g}_v\|_2^2\right) \\ &-\frac{1}{2}\sum_{\sigma\in X_2}\left(\|\boldsymbol{\Theta}_{2,+}\mathbf{h}_\sigma\|_2^2 - \|\boldsymbol{\Theta}_{2,-}\mathbf{h}_\sigma\|_2^2\right).\end{aligned}$$

The flow follows the direction of steepest descent of the total energy. Whether a particular divergence or circulation component is attractive or repulsive therefore depends on the sign of the corresponding channel eigenvalue. More precisely, let μ be an eigenvalue of $\mathbf{W}_1$. Along that channel direction, the node contribution has coefficient

$$-\frac{1}{2}\mu$$

multiplying the squared divergence. Hence

$$\mu < 0 \quad \Longrightarrow \quad \text{positive divergence penalty},$$

whose gradient contribution suppresses that divergence component and therefore induces a down-smoothing, or attractive, tendency. Conversely,

$$\mu > 0 \quad \Longrightarrow \quad \text{negative divergence contribution},$$

whose gradient contribution favors increasing that component and therefore induces a down-sharpening, or repulsive, tendency.

The same interpretation holds for the upper interaction. If μ is an eigenvalue of $\mathbf{W}_2$, then the corresponding squared circulation appears with coefficient

$$-\frac{1}{2}\mu.$$

Thus $\mu < 0$ gives an attractive, up-smoothing tendency, while $\mu > 0$ gives a repulsive, up-sharpening tendency. The lower and upper families have independent channel-mixing matrices and can therefore induce different behaviors simultaneously.

Here the external term acts independently on each edge feature and does not measure either divergence or circulation. In particular, on the harmonic subspace,

$$\hat{\mathbf{B}}_1\mathbf{F} = 0, \qquad \hat{\mathbf{B}}_2^\top\mathbf{F} = 0,$$

so both structural terms vanish. The harmonic component then evolves only through the external channel mixing,

$$\mathbf{P}_{\mathrm{harm}}\mathbf{F}(t) = \mathbf{P}_{\mathrm{harm}}\mathbf{F}(0)\exp(t\mathbf{W}_0)\,,$$

where $\mathbf{P}_{\mathrm{harm}}$ is the orthogonal projector onto the harmonic subspace; App. I discusses the consequences for the two tyings of sec. 4.1.

These interpretations concern the contribution of each signed quadratic term to the gradient. They do not imply that the individual terms are themselves monotone along the full coupled flow; only the total energy $\mathcal{E}(\mathbf{F}(t))$ is guaranteed to decrease. Likewise, an attractive or repulsive tendency in one interaction family does not by itself determine the asymptotic regime. The eventual HD, GD, or CD behavior depends on the comparison between the full spectral growth rates derived in sec. 3.3.

Differences from the graph case. There are two important differences from the graph case in Di Giovanni et al. (2023). First, their structural forces are pairwise, acting between node-feature particles across graph edges. Here the particles are edge features and the interactions are multi-way. The node term couples all edges incident to a common node through $\mathbf{g}_v$, while the face term couples the three boundary edges of a 2-simplex through $\mathbf{h}_\sigma$. Second, the lower and upper interaction families have independent channel-mixing matrices, so attraction and repulsion can occur differently in the gradient and curl sectors.

F Proofs of the results in Section 3

We use the notation of App. D: the joint eigenbasis $\{\mathbf{u}_j\}_{j=1}^{N_1}$ of (6), the channel coefficients $\mathbf{c}_j = \mathbf{F}^\top\mathbf{u}_j$, and the mode masses (17). The vectorized generator $\boldsymbol{\mathcal{M}}$ introduced in the first proof below is used throughout the remaining appendices.

F.1 Proof of Theorem 3.4

Proof. Let $\mathbf{f} = \mathrm{vec}(\mathbf{F}) \in \mathbb{R}^{N_1 d}$ stack the columns of $\mathbf{F}$. Using the trace–Kronecker identity $\mathrm{Tr}(\mathbf{X}^\top\mathbf{A}\mathbf{X}\mathbf{B}) = \mathrm{vec}(\mathbf{X})^\top(\mathbf{B}^\top \otimes \mathbf{A})\,\mathrm{vec}(\mathbf{X})$ and the symmetry of $\mathbf{W}_0, \mathbf{W}_1, \mathbf{W}_2$, the energy is the quadratic form

$$\mathcal{E}(\mathbf{F}) = -\tfrac{1}{2}\,\mathbf{f}^\top\boldsymbol{\mathcal{M}}\,\mathbf{f}, \qquad \boldsymbol{\mathcal{M}} := \mathbf{W}_0 \otimes \mathbf{I} + \mathbf{W}_1 \otimes \boldsymbol{\mathcal{L}}_{\mathrm{down}} + \mathbf{W}_2 \otimes \boldsymbol{\mathcal{L}}_{\mathrm{up}}. \tag{22}$$

Each factor is symmetric and the Kronecker product of symmetric matrices is symmetric, so $\boldsymbol{\mathcal{M}} = \boldsymbol{\mathcal{M}}^\top$ and $\nabla_{\mathbf{f}}\mathcal{E} = -\boldsymbol{\mathcal{M}}\mathbf{f}$. The negative gradient flow $\dot{\mathbf{f}} = -\nabla_{\mathbf{f}}\mathcal{E} = \boldsymbol{\mathcal{M}}\mathbf{f}$ is, by $(\mathbf{B}^\top \otimes \mathbf{A})\,\mathrm{vec}(\mathbf{X}) = \mathrm{vec}(\mathbf{A}\mathbf{X}\mathbf{B})$ applied to each term, exactly (2) in matrix form. Finally $\frac{d}{dt}\mathcal{E}(\mathbf{F}(t)) = \langle\nabla_{\mathbf{f}}\mathcal{E}, \dot{\mathbf{f}}\rangle = -\|\boldsymbol{\mathcal{M}}\mathbf{f}\|_2^2 \le 0$. □

The proof uses only that each factor is symmetric, so the gradient-flow property holds for every affine (indeed every polynomial) propagation operator built from symmetric Laplacians, normalized or not. We call $\boldsymbol{\mathcal{M}}$ the *vectorized generator* of the flow.

F.2 Proof of Proposition 3.5

Proof. For the Frobenius inner product, $\langle \mathbf{u}_i\boldsymbol{\psi}^\top, \mathbf{u}_j\boldsymbol{\phi}^\top\rangle = \mathrm{Tr}(\boldsymbol{\psi}\mathbf{u}_i^\top\mathbf{u}_j\boldsymbol{\phi}^\top) = (\mathbf{u}_i^\top\mathbf{u}_j)(\boldsymbol{\psi}^\top\boldsymbol{\phi})$, which is 1 if $(i,\boldsymbol{\psi}) = (j,\boldsymbol{\phi})$ and 0 for distinct elements of $\{\mathbf{u}_j\boldsymbol{\psi}_{j,r}^\top\}$; hence these $N_1 d$ matrices are orthonormal, and the norm identity $\|\mathbf{F}(t)\|^2 = \sum_{j,r} e^{2\mu_{j,r}t}a_{j,r}^2$ is the sum of squared coefficients in (8). Factoring $e^{\mu_* t}$ out of (8) and out of the norm identity, every term with $\mu_{j,r} < \mu_*$ carries a factor $e^{(\mu_{j,r}-\mu_*)t} \to 0$, which gives the first limit in (9). For the second, left-multiplication by $\boldsymbol{\mathcal{L}}_{\mathrm{down}}$ (resp. $\boldsymbol{\mathcal{L}}_{\mathrm{up}}$) maps $\mathbf{u}_j\boldsymbol{\psi}_{j,r}^\top$ to $\nu_j^{\mathrm{down}}\,\mathbf{u}_j\boldsymbol{\psi}_{j,r}^\top$ (resp. $\nu_j^{\mathrm{up}}\,\mathbf{u}_j\boldsymbol{\psi}_{j,r}^\top$) by (6), and the terms are orthogonal, so $\mathrm{Tr}(\mathbf{F}^\top\boldsymbol{\mathcal{L}}_{\mathrm{down}}\mathbf{F}) = \sum_{j,r} e^{2\mu_{j,r}t}a_{j,r}^2\,\nu_j^{\mathrm{down}}$ and likewise for $\boldsymbol{\mathcal{L}}_{\mathrm{up}}$; dividing by the norm identity and letting $t\to\infty$ leaves the stated weighted averages over T_*. The excluded set $\{\sum_{(j,r)\in T_*} a_{j,r}^2 = 0\}$ is the orthogonal complement $E_*^\perp$ of the top eigenspace, a proper subspace since $T_* \neq \varnothing$. That the $\mathbf{u}_j\boldsymbol{\psi}_{j,r}^\top$ are eigenvectors of the map $\mathbf{F}\mapsto \mathbf{F}\mathbf{W}_0 + \boldsymbol{\mathcal{L}}_{\mathrm{down}}\mathbf{F}\mathbf{W}_1 + \boldsymbol{\mathcal{L}}_{\mathrm{up}}\mathbf{F}\mathbf{W}_2$ with eigenvalues $\mu_{j,r}$ is the identity $\boldsymbol{\mathcal{M}}(\boldsymbol{\psi}_{j,r}\otimes\mathbf{u}_j) = \mu_{j,r}(\boldsymbol{\psi}_{j,r}\otimes\mathbf{u}_j)$ established next. □

Spectrum of the vectorized generator. Since $\mathrm{vec}(\mathbf{u}_j\boldsymbol{\psi}^\top) = \boldsymbol{\psi}\otimes\mathbf{u}_j$ and $(\mathbf{W}\otimes\mathbf{A})(\boldsymbol{\psi}\otimes\mathbf{u}) = (\mathbf{W}\boldsymbol{\psi})\otimes(\mathbf{A}\mathbf{u})$, (6) gives $\boldsymbol{\mathcal{M}}(\boldsymbol{\psi}\otimes\mathbf{u}_j) = (\mathbf{G}_j\boldsymbol{\psi})\otimes\mathbf{u}_j$. The subspaces $\mathbb{R}^d\otimes\mathrm{span}\{\mathbf{u}_j\}$ are therefore invariant under $\boldsymbol{\mathcal{M}}$, mutually orthogonal, and span $\mathbb{R}^{N_1 d}$, and $\boldsymbol{\mathcal{M}}$ is block diagonal on them with the symmetric blocks $\mathbf{G}_j$. Its eigenpairs are exactly $(\mu_{j,r}, \boldsymbol{\psi}_{j,r}\otimes\mathbf{u}_j)$, so

$$\mathrm{Spec}(\boldsymbol{\mathcal{M}}) = \bigcup_{j=1}^{N_1} \mathrm{Spec}(\mathbf{G}_j), \tag{23}$$

$$\lambda_{\max}(\boldsymbol{\mathcal{M}}) = \max_j \lambda_{\max}(\mathbf{G}_j) = \mu_*, \lambda_{\min}(\boldsymbol{\mathcal{M}}) = \min_j \lambda_{\min}(\mathbf{G}_j) = \mu_{\min}. \tag{24}$$

Because $\mathbf{G}_j$ depends on $\mathbf{u}_j$ only through its frequency pair, the index set T_* is a union of full joint eigenspaces and does not depend on the chosen basis, and prop. 3.5 says that $\mathbf{f}(t)/\|\mathbf{f}(t)\|_2$ converges to the normalized orthogonal projection of $\mathbf{f}(0)$ onto the top eigenspace $\ker(\boldsymbol{\mathcal{M}} - \mu_*\mathbf{I})$. The extreme rates μ_* and $\mu_{\min}$ of sec. 3.3 are thus the extreme eigenvalues of $\boldsymbol{\mathcal{M}}$, which is the form used in App. H.

F.3 Endpoint reduction

For a real parameter ν, write

$$\mathbf{G}^{\mathrm{harm}} := \mathbf{W}_0, \qquad \mathbf{G}_\nu^{\mathrm{grad}} := \mathbf{W}_0 + \nu\mathbf{W}_1, \qquad \mathbf{G}_\nu^{\mathrm{curl}} := \mathbf{W}_0 + \nu\mathbf{W}_2, \tag{25}$$

so that the channel generator (7) of a harmonic mode, a gradient mode of frequency ν, or a curl mode of frequency ν is $\mathbf{G}^{\mathrm{harm}}$, $\mathbf{G}_\nu^{\mathrm{grad}}$ or $\mathbf{G}_\nu^{\mathrm{curl}}$ respectively, and the endpoint rates (10) are $\rho_H = \lambda_{\max}(\mathbf{G}^{\mathrm{harm}})$, $\rho_G = \lambda_{\max}(\mathbf{G}_{\nu_g^{\max}}^{\mathrm{grad}})$ and $\rho_C = \lambda_{\max}(\mathbf{G}_{\nu_c^{\max}}^{\mathrm{curl}})$. The affine families (25) are defined for every real ν, not only for the frequencies that occur in the complex. The following lemma is the convexity step used in sec. 3.3.

Lemma F.1 (Endpoint reduction)**.** *The map $\nu \mapsto \lambda_{\max}(\mathbf{G}_\nu^{\mathrm{grad}})$ is convex on $\mathbb{R}$, and hence for $0 \le \nu \le \nu_g^{\max}$,*

$$\lambda_{\max}\big(\mathbf{G}_\nu^{\mathrm{grad}}\big) \le \Big(1 - \frac{\nu}{\nu_g^{\max}}\Big)\rho_H + \frac{\nu}{\nu_g^{\max}}\,\rho_G. \tag{26}$$

The same holds for the curl family with ρ_C in place of ρ_G. Symmetrically, $\nu \mapsto \lambda_{\min}(\mathbf{G}_\nu^{\mathrm{grad}})$ is concave, and for $0 \le \nu \le \nu_g^{\max}$,

$$\lambda_{\min}\big(\mathbf{G}_\nu^{\mathrm{grad}}\big) \ge \Big(1 - \frac{\nu}{\nu_g^{\max}}\Big)\lambda_{\min}(\mathbf{W}_0) + \frac{\nu}{\nu_g^{\max}}\,\lambda_{\min}\Big(\mathbf{G}_{\nu_g^{\max}}^{\mathrm{grad}}\Big), \tag{27}$$

again with the curl analogue. Consequently, if $\beta_1 > 0$, the top and bottom rates of the flow are

$$\mu_* = \max\{\rho_H, \rho_G, \rho_C\}, \qquad \mu_{\min} = \min\big\{\lambda_{\min}(\mathbf{W}_0),\ \lambda_{\min}(\mathbf{G}_{\nu_g^{\max}}^{\mathrm{grad}}),\ \lambda_{\min}(\mathbf{G}_{\nu_c^{\max}}^{\mathrm{curl}})\big\},$$

the second being (11).

Proof. Write $r_g(\nu) := \lambda_{\max}(\mathbf{G}_\nu^{\mathrm{grad}})$ and $r_c(\nu) := \lambda_{\max}(\mathbf{G}_\nu^{\mathrm{curl}})$. We use the continuous parameter to obtain bounds and then evaluate them at the frequencies present.

Convexity. For a real symmetric $\mathbf{S}$, $\lambda_{\max}(\mathbf{S}) = \max_{\|\boldsymbol{\psi}\|_2=1} \boldsymbol{\psi}^\top \mathbf{S}\boldsymbol{\psi}$. For each fixed unit $\boldsymbol{\psi}$, the function $q_\psi(\nu) := \boldsymbol{\psi}^\top \mathbf{G}_\nu^{\mathrm{grad}} \boldsymbol{\psi} = \boldsymbol{\psi}^\top \mathbf{W}_0 \boldsymbol{\psi} + \nu\, \boldsymbol{\psi}^\top \mathbf{W}_1 \boldsymbol{\psi}$ is affine in ν, and $r_g(\nu) = \max_{\|\boldsymbol{\psi}\|_2=1} q_\psi(\nu)$ is a pointwise maximum of affine functions, hence convex: for $\theta \in [0,1]$,

$$r_g\big((1-\theta)\nu_0 + \theta\nu_1\big) = \max_{\|\boldsymbol{\psi}\|_2=1} \big[(1-\theta) q_\psi(\nu_0) + \theta q_\psi(\nu_1)\big] \le (1-\theta)\, r_g(\nu_0) + \theta\, r_g(\nu_1).$$

The maximizing $\boldsymbol{\psi}$ may change with ν; no common eigenbasis of the channel weights is required. The same argument applies to r_c.

Endpoint bound. At $\nu = 0$ both affine families reduce to $\mathbf{G}^{\mathrm{harm}}$, so $r_g(0) = r_c(0) = \rho_H$, while $r_g(\nu_g^{\max}) = \rho_G$ and $r_c(\nu_c^{\max}) = \rho_C$. For $0 \le \nu \le \nu_g^{\max}$, convexity with $\theta = \nu/\nu_g^{\max}$ gives (26); since its right-hand side is a convex combination, $r_g(\nu) \le \max\{\rho_H, \rho_G\}$. Likewise $r_c(\nu) \le \max\{\rho_H, \rho_C\}$ on $[0, \nu_c^{\max}]$. These bounds do not require a harmonic mode to exist.

Top rate. Assume $\beta_1 > 0$. By (23), $\mu_* = \max_j \lambda_{\max}(\mathbf{G}_j)$, where each $\mathbf{G}_j$ equals $\mathbf{G}^{\mathrm{harm}}$, $\mathbf{G}^{\mathrm{grad}}_{\nu_j^{\mathrm{down}}}$ or $\mathbf{G}^{\mathrm{curl}}_{\nu_j^{\mathrm{up}}}$ according to the type of $\mathbf{u}_j$; by the endpoint bound every block is at most $\max\{\rho_H, \rho_G, \rho_C\}$. Conversely each of the three rates is attained: $\beta_1 > 0$ provides a harmonic $\mathbf{u}_j$ with $\mathbf{G}_j = \mathbf{G}^{\mathrm{harm}}$, and the finite nonempty families Σ_g and Σ_c contain their maxima, providing modes with $\mathbf{G}_j = \mathbf{G}^{\mathrm{grad}}_{\nu_g^{\max}}$ and $\mathbf{G}_j = \mathbf{G}^{\mathrm{curl}}_{\nu_c^{\max}}$. Hence $\mu_* = \max\{\rho_H, \rho_G, \rho_C\}$.

Bottom rate. The same three steps apply to $\lambda_{\min}$ with every inequality reversed: $\lambda_{\min}(\mathbf{S}) = \min_{\|\boldsymbol{\psi}\|_2=1} \boldsymbol{\psi}^\top \mathbf{S}\boldsymbol{\psi}$ is a pointwise minimum of the affine functions q_ψ, hence concave, which gives (27) and its curl analogue; the right-hand side is a convex combination, so on each family $\lambda_{\min}(\mathbf{G}_j)$ is at least the smaller of its two endpoint values; by (23), $\mu_{\min} = \lambda_{\min}(\boldsymbol{\mathcal{M}}) = \min_j \lambda_{\min}(\mathbf{G}_j)$; and the three endpoint values are attained by the same three modes as above. Hence $\mu_{\min} = \min\{\lambda_{\min}(\mathbf{W}_0), \lambda_{\min}(\mathbf{G}^{\mathrm{grad}}_{\nu_g^{\max}}), \lambda_{\min}(\mathbf{G}^{\mathrm{curl}}_{\nu_c^{\max}})\}$. □

F.4 Proof of Theorem 3.6

Proof. By prop. 3.5, for every $\mathbf{F}(0)$ outside the proper subspace $\{\sum_{(j,r)\in T_*} a_{j,r}^2 = 0\}$, which consists of the initial conditions with zero projection onto the top-rate eigenspace and has Lebesgue measure zero, the normalized features converge and the Rayleigh-quotient pair converges to the weighted average (9) of the frequency pairs $(\nu_j^{\mathrm{down}}, \nu_j^{\mathrm{up}})$ over $T_* = \{(j,r) : \mu_{j,r} = \mu_*\}$. It therefore suffices to show that, under the strict-maximum hypothesis, every j appearing in T_* has the same frequency pair, namely $(0,0)$, $(\nu_g^{\max}, 0)$ or $(0, \nu_c^{\max})$ in the three cases; the limit is then that pair. Recall from lem. F.1 that $\mu_* = \max\{\rho_H, \rho_G, \rho_C\}$, and that $\mu_{j,r} \le \lambda_{\max}(\mathbf{G}_j)$ with $\mathbf{G}_j$ determined by the Hodge type and frequency of $\mathbf{u}_j$.

Case (i): $\rho_H > \max\{\rho_G, \rho_C\}$. For a gradient mode $\mathbf{u}_j$ of frequency $0 < \nu \le \nu_g^{\max}$, (26) gives

$$\lambda_{\max}(\mathbf{G}_j) \le \Big(1 - \frac{\nu}{\nu_g^{\max}}\Big)\rho_H + \frac{\nu}{\nu_g^{\max}}\rho_G < \rho_H,$$

strictly because the coefficient of $\rho_G < \rho_H$ is positive; the same holds for curl modes with ρ_C. Hence $\mu_{j,r} = \mu_* = \rho_H$ forces $\mathbf{u}_j$ to be harmonic, every pair in T_* is $(0,0)$, and $(\mathrm{RQ}_{\mathrm{down}}, \mathrm{RQ}_{\mathrm{up}}) \to (0,0)$: the flow is HD.

Case (ii): $\rho_G > \max\{\rho_H, \rho_C\}$. Harmonic modes have $\lambda_{\max}(\mathbf{G}_j) = \rho_H < \rho_G$, and curl modes have $\lambda_{\max}(\mathbf{G}_j) \le \max\{\rho_H, \rho_C\} < \rho_G$ by the endpoint bound. For a gradient mode of frequency $\nu < \nu_g^{\max}$, (26) gives $\lambda_{\max}(\mathbf{G}_j) < \rho_G$, strictly because the coefficient of $\rho_H < \rho_G$ is positive. Hence $\mu_{j,r} = \mu_*$ forces $\mathbf{u}_j$ to be a gradient mode of frequency exactly $\nu_g^{\max}$, every pair in T_* is $(\nu_g^{\max}, 0)$, and $(\mathrm{RQ}_{\mathrm{down}}, \mathrm{RQ}_{\mathrm{up}}) \to (\nu_g^{\max}, 0)$: the flow is GD.

Case (iii): $\rho_C > \max\{\rho_H, \rho_G\}$ is case (ii) with the roles of the gradient and curl families exchanged, giving the limit $(0, \nu_c^{\max})$ and CD.

Neither a simple top eigenvalue nor a positive value of μ_* is needed: only differences of growth rates enter. If two rates tied at the maximum, T_* could contain modes of different types and the limit would be the corresponding weighted combination of their frequency pairs; see rem. G.3. □

G Extensions of the spectral concentration trichotomy

G.1 The case $\beta_1 = 0$: no harmonic anchor

Sec. 3.3 assumes $\beta_1 > 0$ throughout, so that the harmonic sector supplies the rate ρ_H at frequency zero. When $\beta_1 = 0$ there is no harmonic mode, $\lambda_{\min}(\boldsymbol{\mathcal{L}}_1) = \min\{\nu_g^{\min}, \nu_c^{\min}\} > 0$, and $\mathrm{RQ}(\mathbf{F}) \geq \lambda_{\min}(\boldsymbol{\mathcal{L}}_1) > 0$ for every $\mathbf{F} \neq 0$, so HD in the sense of prop. 3.3 cannot occur; the natural replacement of the harmonic sector is the bottom eigenspace of $\boldsymbol{\mathcal{L}}_1$ (rem. D.2). The endpoint reduction of lem. F.1 still applies within each family, but the relevant endpoints are now the smallest and largest frequencies actually present. Define

$$\rho_G^- := \lambda_{\max}\Big(\mathbf{G}^{\mathrm{grad}}_{\nu_g^{\min}}\Big), \qquad \rho_C^- := \lambda_{\max}\Big(\mathbf{G}^{\mathrm{curl}}_{\nu_c^{\min}}\Big).$$

Proposition G.1 (No harmonic anchor)**.** *Suppose $\beta_1 = 0$ and asmp. 3.2 holds. Then the top rate of the flow* (2) *is*

$$\mu_* = \max\{\rho_G^-, \rho_G, \rho_C^-, \rho_C\}.$$

If the maximum is attained by exactly one of the four rates, then for almost every $\mathbf{F}(0)$ the normalized features concentrate on the corresponding family and endpoint frequency, and

$$(\mathrm{RQ}_{\mathrm{down}}, \mathrm{RQ}_{\mathrm{up}}) \longrightarrow \begin{cases} (\nu_g^{\max}, 0), & \rho_G \textit{ largest (GD)}, \\ (0, \nu_c^{\max}), & \rho_C \textit{ largest (CD)}, \\ (\nu_g^{\min}, 0), & \rho_G^- \textit{ largest}, \\ (0, \nu_c^{\min}), & \rho_C^- \textit{ largest}. \end{cases}$$

In the language of lem. D.3, the limit is the vertex of $\mathcal{P}$ whose endpoint rate is largest. A bottom-endpoint limit concentrates on the bottom eigenspace of $\boldsymbol{\mathcal{L}}_1$, i.e. $\mathrm{RQ} \to \lambda_{\min}(\boldsymbol{\mathcal{L}}_1)$, exactly when its frequency equals $\min\{\nu_g^{\min}, \nu_c^{\min}\}$. If $\nu_g^{\min} \neq \nu_c^{\min}$ and the larger bottom frequency wins, the limiting frequency lies strictly between $\lambda_{\min}(\boldsymbol{\mathcal{L}}_1)$ and $\lambda_{\max}(\boldsymbol{\mathcal{L}}_1)$, and the flow is neither GD, nor CD, nor concentrated on the bottom eigenspace.

Proof. Top rate. With r_g, r_c as in the proof of lem. F.1, $\rho_G^- = r_g(\nu_g^{\min})$, $\rho_G = r_g(\nu_g^{\max})$, and similarly for the curl family. If $\nu_g^{\min} < \nu_g^{\max}$, convexity of r_g gives, for $\nu \in [\nu_g^{\min}, \nu_g^{\max}]$ and $\theta_g(\nu) := (\nu - \nu_g^{\min})/(\nu_g^{\max} - \nu_g^{\min}) \in [0, 1]$,

$$r_g(\nu) \leq \big(1 - \theta_g(\nu)\big)\rho_G^- + \theta_g(\nu)\,\rho_G \leq \max\{\rho_G^-, \rho_G\}; \tag{28}$$

if $\nu_g^{\min} = \nu_g^{\max}$ the family has a single frequency and $\rho_G^- = \rho_G$. Both endpoints belong to Σ_g, so $\max_{\nu \in \Sigma_g} r_g(\nu) = \max\{\rho_G^-, \rho_G\}$, and likewise $\max_{\nu \in \Sigma_c} r_c(\nu) = \max\{\rho_C^-, \rho_C\}$. Since $\beta_1 = 0$, every $\mathbf{u}_j$ is a gradient or a curl mode, so by (23) $\mu_* = \max_j \lambda_{\max}(\mathbf{G}_j) = \max\{\rho_G^-, \rho_G, \rho_C^-, \rho_C\}$. This proves the first assertion without any uniqueness assumption.

Sector identification. Suppose ρ_G^- is the unique maximum. Then $\rho_G^- > \rho_G$ forces $\nu_g^{\min} < \nu_g^{\max}$, since equal frequencies have equal generators. For every gradient frequency $\nu > \nu_g^{\min}$ we have $\theta_g(\nu) > 0$, so (28) gives $r_g(\nu) < \rho_G^-$, and every curl mode has $\lambda_{\max}(\mathbf{G}_j) \leq \max\{\rho_C^-, \rho_C\} < \rho_G^-$. Hence $\mu_{j,r} = \mu_*$ forces $\mathbf{u}_j$ to be a gradient mode of frequency exactly $\nu_g^{\min}$. If ρ_G is the unique maximum the same argument applies with $1 - \theta_g(\nu) > 0$ for $\nu < \nu_g^{\max}$, and the two curl cases exchange the roles of the families. In each case all indices in T_* share one frequency pair, so prop. 3.5 gives, for almost every $\mathbf{F}(0)$, concentration on that eigenspace and the stated limit of the Rayleigh-quotient pair. No simultaneous diagonalization of the channel weights is required.

Classification. The first two limits are GD and CD by prop. 3.3. Since $\beta_1 = 0$, every joint eigenmode is a gradient or curl mode and is an eigenvector of $\boldsymbol{\mathcal{L}}_1$ with eigenvalue equal to its frequency, so $\mathrm{Spec}(\boldsymbol{\mathcal{L}}_1) = \Sigma_g \cup \Sigma_c$ and $\lambda_{\min}(\boldsymbol{\mathcal{L}}_1) = \min\{\nu_g^{\min}, \nu_c^{\min}\}$. In a bottom-gradient limit $\mathrm{RQ} \to \nu_g^{\min}$, which equals $\lambda_{\min}(\boldsymbol{\mathcal{L}}_1)$ exactly when $\nu_g^{\min} \leq \nu_c^{\min}$, and rem. D.2 then gives concentration on the bottom eigenspace; the bottom-curl case is symmetric. If instead $\nu_g^{\min} > \nu_c^{\min}$ and ρ_G^- wins, then $\mathrm{RQ} \to \nu_g^{\min} > \lambda_{\min}(\boldsymbol{\mathcal{L}}_1)$, while $\mathrm{RQ}_{\mathrm{down}} \to \nu_g^{\min} < \nu_g^{\max}$ and $\mathrm{RQ}_{\mathrm{up}} \to 0 < \nu_c^{\max}$, so the flow is neither GD nor CD, and its limiting frequency satisfies $\lambda_{\min}(\boldsymbol{\mathcal{L}}_1) < \nu_g^{\min} < \nu_g^{\max} \leq \lambda_{\max}(\boldsymbol{\mathcal{L}}_1)$. □

Prop. G.1 shows why concentration on the harmonic subspace alone is not a sufficient description of asymptotic behavior in simplicial complexes: when $\ker \boldsymbol{\mathcal{L}}_1 = \{0\}$ the dynamics can still concentrate on the lowest available Hodge frequency, and it can also settle at a frequency that is extreme within one Hodge family but intermediate in the full spectrum.

G.2 Commuting channel weights

Corollary G.2 (Commuting channel weights)**.** *Suppose $\mathbf{W}_0$, $\mathbf{W}_1$ and $\mathbf{W}_2$ commute pairwise, and let $\{\boldsymbol{\phi}_r\}_{r=1}^d$ be a common orthonormal eigenbasis with eigenvalue triples $(\mu_r^{(0)}, \mu_r^{(1)}, \mu_r^{(2)})$. Then*

$$\rho_H = \max_r \mu_r^{(0)}, \qquad \rho_G = \max_r \big(\mu_r^{(0)} + \nu_g^{\max}\mu_r^{(1)}\big), \qquad \rho_C = \max_r \big(\mu_r^{(0)} + \nu_c^{\max}\mu_r^{(2)}\big),$$

where in each maximum the eigenvalues with index r belong to the same eigenvector and are not maximized independently. For the adjacency-like tying of sec. 4.1 with commuting $\mathbf{W}_{\mathrm{down}}, \mathbf{W}_{\mathrm{up}}$ and joint eigenvalue pairs $(\mu_r^{\mathrm{down}}, \mu_r^{\mathrm{up}})$, this reads

$$\rho_H = \max_r \big(\mu_r^{\mathrm{down}} + \mu_r^{\mathrm{up}}\big), \qquad \rho_G = \max_r \big((1 - 2\nu_g^{\max})\mu_r^{\mathrm{down}} + \mu_r^{\mathrm{up}}\big),$$

$$\rho_C = \max_r \big(\mu_r^{\mathrm{down}} + (1 - 2\nu_c^{\max})\mu_r^{\mathrm{up}}\big),$$

The trichotomy conditions thus reduce to explicit inequalities between the channel eigenvalues and the two extreme Hodge frequencies.

Proof. Commuting symmetric matrices are simultaneously diagonalizable by an orthonormal basis. In that basis $\mathbf{G}_\nu^{\mathrm{grad}}\boldsymbol{\phi}_r = (\mu_r^{(0)} + \nu\mu_r^{(1)})\boldsymbol{\phi}_r$, so $\lambda_{\max}(\mathbf{G}_\nu^{\mathrm{grad}}) = \max_r(\mu_r^{(0)} + \nu\mu_r^{(1)})$; evaluating at $\nu = 0$ and $\nu = \nu_g^{\max}$ gives ρ_H and ρ_G, and ρ_C follows in the same way from $\mathbf{G}_\nu^{\mathrm{curl}}$. The adjacency-like form substitutes $\mathbf{W}_0 = \mathbf{W}_{\mathrm{down}} + \mathbf{W}_{\mathrm{up}}$, $\mathbf{W}_1 = -2\mathbf{W}_{\mathrm{down}}$, $\mathbf{W}_2 = -2\mathbf{W}_{\mathrm{up}}$, which commute pairwise whenever $\mathbf{W}_{\mathrm{down}}$ and $\mathbf{W}_{\mathrm{up}}$ do. □

G.3 Ties between growth rates

Remark G.3 (Ties)**.** Thm. 3.6 and prop. G.1 assume a strict maximum. Without it, prop. 3.5 applies verbatim: the normalized features converge to the normalized projection of $\mathbf{F}(0)$ onto the top eigenspace E_*, and the limiting Rayleigh-quotient pair is the mass-weighted average (9) of the frequency pairs of *all* modes in T_*, which now come from more than one Hodge subspace or frequency. Convexity determines which modes these are. If $\rho_G = \rho_C > \rho_H$, the bound (26) and its curl analogue are strict at every interior frequency, so T_* consists of top-gradient and top-curl modes only, and the limit lies on the segment from $(\nu_g^{\max}, 0)$ to $(0, \nu_c^{\max})$. If $\rho_H = \rho_G \geq \rho_C$, the convex map $\nu \mapsto \lambda_{\max}(\mathbf{G}_\nu^{\mathrm{grad}})$ equals ρ_H at both ends of $[0, \nu_g^{\max}]$, so either it is strictly smaller at every interior frequency, and T_* contains only harmonic and top-gradient modes, or it is constant on the whole interval, and every gradient mode of the complex belongs to T_*. The second case occurs, for instance, for $\mathbf{W}_0 = \mathbf{W}_1 = 0$, $\mathbf{W}_2 = -\mathbf{I}$, where all harmonic and gradient modes have rate 0 and all curl modes have negative rate; the limit is then $\big(\sum_j \nu_j^{\mathrm{down}}\,\tilde{m}_j,\, 0\big)$ with $\tilde{m}_j$ the masses of $\mathbf{F}(0)$ on the harmonic and gradient modes renormalized to sum to one, a point of the segment from $(0, 0)$ to $(\nu_g^{\max}, 0)$ that depends on the input's mass at interior gradient frequencies. The tie $\rho_H = \rho_C$ is symmetric, and a three-way tie combines the two cases. In every case the limit is a point of $\mathcal{P}$ that depends on the projection of the initial features onto the tied modes.

H DISCRETIZATION OF THE FLOW

We return from the continuous flow to the discrete layer update. By App. F.2, after vectorization the flow reads $\dot{\mathbf{f}} = \boldsymbol{\mathcal{M}}\mathbf{f}$ with the symmetric generator $\boldsymbol{\mathcal{M}}$ of (22), whose eigenpairs are $(\mu_{j,r}, \boldsymbol{\psi}_{j,r}\otimes\mathbf{u}_j)$. The layer update (1) with $\sigma = \mathrm{id}$ and shared symmetric weights is the forward Euler step with step size $\tau > 0$,

$$\mathbf{f}_{m+1} = (\mathbf{I} + \tau\boldsymbol{\mathcal{M}})\mathbf{f}_m,$$

so a component along an eigenvector with $\boldsymbol{\mathcal{M}}\boldsymbol{\psi} = \eta\boldsymbol{\psi}$ evolves as $e^{\eta t}$ in continuous time and as $(1+\tau\eta)^m$ in discrete depth. The continuous flow is therefore governed by the largest eigenvalue of $\boldsymbol{\mathcal{M}}$, whereas the discrete iteration is governed by the eigenvalue for which $|1+\tau\eta|$ is largest. Let

$$\eta_{\min} := \lambda_{\min}(\boldsymbol{\mathcal{M}}) = \mu_{\min}, \qquad \eta_{\max} := \lambda_{\max}(\boldsymbol{\mathcal{M}}) = \mu_*,$$

the smallest and largest rates of sec. 3.3 (by (23)), and assume $\eta_{\min} < \eta_{\max}$. Since $\eta \mapsto |1+\tau\eta|$ is convex, its maximum over $\mathrm{Spec}(\boldsymbol{\mathcal{M}}) \subseteq [\eta_{\min}, \eta_{\max}]$ is attained at one of these two endpoints. Comparing their moduli gives the following result.

Proposition H.1 (Step size condition)**.** *For almost every initial condition* $\mathbf{F}_0$*:*

(i) If

$$\tau(\eta_{\max} + \eta_{\min}) > -2,$$

then

$$|1+\tau\eta_{\max}| > |1+\tau\eta_{\min}|,$$

and in fact $1+\tau\eta_{\max} > 0$*. The discrete iteration is therefore asymptotically dominated by the top eigenspace of* $\boldsymbol{\mathcal{M}}$*, just as in continuous time. Hence it has the same HD, GD, or CD regime predicted by the continuous-time analysis.*

(ii) If

$$\tau(\eta_{\max} + \eta_{\min}) < -2,$$

then

$$|1+\tau\eta_{\min}| > |1+\tau\eta_{\max}|, \qquad 1+\tau\eta_{\min} < 0.$$

The bottom eigenspace of $\boldsymbol{\mathcal{M}}$ *instead dominates. Its direction therefore alternates sign from one layer to the next while its magnitude dominates the remaining components.*

At equality,

$$\tau(\eta_{\max} + \eta_{\min}) = -2,$$

the two endpoint multipliers have equal modulus and the limiting behavior depends on the initial projections onto both eigenspaces.

Proof. Let $\mathbf{f}_m = \mathrm{vec}(\mathbf{F}_m)$ and write

$$\mathbf{T}_\tau := \mathbf{I} + \tau\boldsymbol{\mathcal{M}}, \qquad \mathbf{f}_m = \mathbf{T}_\tau^m \mathbf{f}_0.$$

Throughout, $\eta_{\min} = \lambda_{\min}(\boldsymbol{\mathcal{M}}) < \lambda_{\max}(\boldsymbol{\mathcal{M}}) = \eta_{\max}$ and $\tau > 0$. All top and bottom eigenspaces below are eigenspaces of the full feature generator $\boldsymbol{\mathcal{M}}$, not of the Hodge Laplacian. For each distinct eigenvalue η of $\boldsymbol{\mathcal{M}}$, let $\boldsymbol{\Pi}_\eta$ denote its orthogonal spectral projector, and set

$$\boldsymbol{\Pi}_+ := \boldsymbol{\Pi}_{\eta_{\max}}, \qquad \boldsymbol{\Pi}_- := \boldsymbol{\Pi}_{\eta_{\min}}, \qquad \mathbf{v}_+ := \boldsymbol{\Pi}_+\mathbf{f}_0, \qquad \mathbf{v}_- := \boldsymbol{\Pi}_-\mathbf{f}_0.$$

Moduli of the Euler multipliers. If $\boldsymbol{\mathcal{M}}\boldsymbol{\psi} = \eta\boldsymbol{\psi}$, then $\mathbf{T}_\tau\boldsymbol{\psi} = (1+\tau\eta)\boldsymbol{\psi}$. Thus $\mathbf{T}_\tau$ has the same eigenspaces as $\boldsymbol{\mathcal{M}}$, with eigenvalue $\alpha_\eta := 1+\tau\eta$ on the η-eigenspace, and for every $m \geq 1$,

$$\mathbf{f}_m = \sum_{\eta\in\mathrm{Spec}(\boldsymbol{\mathcal{M}})} \alpha_\eta^m \, \boldsymbol{\Pi}_\eta \mathbf{f}_0.$$

The magnitude of each component is multiplied by $|\alpha_\eta|^m$, so we must maximize the modulus, not the signed multiplier. Write $\alpha_+ := 1+\tau\eta_{\max}$ and $\alpha_- := 1+\tau\eta_{\min}$. The squared modulus $h(\eta) := (1+\tau\eta)^2$ is strictly convex, since $h''(\eta) = 2\tau^2 > 0$. For an interior value $\eta = (1-\theta)\eta_{\min} + \theta\eta_{\max}$ with $0 < \theta < 1$,

$$h(\eta) < (1-\theta)h(\eta_{\min}) + \theta h(\eta_{\max}) \leq \max\{h(\eta_{\min}), h(\eta_{\max})\}.$$

Both endpoints belong to the spectrum of $\boldsymbol{\mathcal{M}}$. Therefore $\max_{\eta\in\mathrm{Spec}(\boldsymbol{\mathcal{M}})}|\alpha_\eta| = \max\{|\alpha_-|, |\alpha_+|\}$, and no interior eigenvalue can attain this maximum. To compare the two endpoints, factor the difference of their squares:

$$|\alpha_+|^2 - |\alpha_-|^2 = \tau(\eta_{\max} - \eta_{\min})\big[2 + \tau(\eta_{\max} + \eta_{\min})\big].$$

The first factor is strictly positive, so the sign of the difference is exactly the sign of $2 + \tau(\eta_{\max} + \eta_{\min})$. This proves the modulus comparisons in parts (i) and (ii), as well as equality of the moduli at the stated threshold. For the signs of the winning multipliers, note that

$$\alpha_+ - \alpha_- = \tau(\eta_{\max} - \eta_{\min}) > 0, \qquad \alpha_+ + \alpha_- = 2 + \tau(\eta_{\max} + \eta_{\min}).$$

Top-eigenspace dominance in part (i). Suppose $\tau(\eta_{\max} + \eta_{\min}) > -2$. Then $|\alpha_+| > |\alpha_-|$, and $\alpha_+ = \frac{1}{2}\big[(\alpha_+ + \alpha_-) + (\alpha_+ - \alpha_-)\big] > 0$. Every other eigenvalue $\eta < \eta_{\max}$ satisfies $|\alpha_\eta| < \alpha_+$. Factoring out the winning multiplier gives

$$\mathbf{f}_m = \alpha_+^m\big(\mathbf{v}_+ + \mathbf{r}_m\big), \qquad \mathbf{r}_m := \sum_{\eta<\eta_{\max}} \Big(\frac{\alpha_\eta}{\alpha_+}\Big)^m \mathbf{\Pi}_\eta \mathbf{f}_0 \to 0.$$

If $\mathbf{v}_+ \neq 0$, orthogonality of the spectral eigenspaces yields $\mathbf{f}_m/\|\mathbf{f}_m\|_2 \to \mathbf{v}_+/\|\mathbf{v}_+\|_2$. This is the same normalized limit as in continuous time (prop. 3.5), since $e^{-\eta_{\max}t}e^{t\boldsymbol{\mathcal{M}}}\mathbf{f}_0 \to \mathbf{v}_+$. Continuity of the quadratic forms defining the Rayleigh quotients then gives the same limiting Rayleigh quotients. Consequently, whenever the continuous-time analysis establishes an HD, GD, or CD regime, the discrete iteration has that regime as well.

Alternating bottom-eigenspace dominance in part (ii). Suppose $\tau(\eta_{\max} + \eta_{\min}) < -2$. Then $|\alpha_-| > |\alpha_+|$, and $\alpha_- = \frac{1}{2}\big[(\alpha_+ + \alpha_-) - (\alpha_+ - \alpha_-)\big] < 0$. Set $R := |\alpha_-| = -\alpha_- > 0$. Every other eigenvalue $\eta > \eta_{\min}$ has $|\alpha_\eta| < R$, so

$$\mathbf{f}_m = R^m\big((-1)^m\mathbf{v}_- + \mathbf{r}_m\big), \qquad \mathbf{r}_m := \sum_{\eta>\eta_{\min}} \Big(\frac{\alpha_\eta}{R}\Big)^m \mathbf{\Pi}_\eta \mathbf{f}_0 \to 0.$$

If $\mathbf{v}_- \neq 0$, orthogonality gives $(-1)^m\mathbf{f}_m/\|\mathbf{f}_m\|_2 \to \mathbf{v}_-/\|\mathbf{v}_-\|_2$. Thus the bottom eigenspace of $\boldsymbol{\mathcal{M}}$ dominates relative to all remaining components, while the normalized even and odd subsequences approach opposite directions. This proves part (ii), regardless of whether the unnormalized feature norm grows or decays.

Equality. At $\tau(\eta_{\max} + \eta_{\min}) = -2$, we have $\alpha_+ = -\alpha_- = \frac{\tau}{2}(\eta_{\max} - \eta_{\min}) =: R > 0$, and all interior eigenvalues have multiplier modulus strictly smaller than R. Hence

$$\mathbf{f}_m = R^m\big(\mathbf{v}_+ + (-1)^m\mathbf{v}_- + \mathbf{r}_m\big), \qquad \mathbf{r}_m \to 0,$$

and, with $D := \sqrt{\|\mathbf{v}_+\|_2^2 + \|\mathbf{v}_-\|_2^2} > 0$,

$$\frac{\mathbf{f}_{2k}}{\|\mathbf{f}_{2k}\|_2} \longrightarrow \frac{\mathbf{v}_+ + \mathbf{v}_-}{D}, \qquad \frac{\mathbf{f}_{2k+1}}{\|\mathbf{f}_{2k+1}\|_2} \longrightarrow \frac{\mathbf{v}_+ - \mathbf{v}_-}{D}.$$

These formulas describe the dependence on both initial projections.

Finally, the nonzero-projection conditions used above hold for almost every initial condition: the excluded sets are $\ker \mathbf{\Pi}_+$, $\ker \mathbf{\Pi}_-$, and $\ker \mathbf{\Pi}_+ \cap \ker \mathbf{\Pi}_-$ respectively, each a proper linear subspace of $\mathbb{R}^{N_1 d}$ and therefore of Lebesgue measure zero. The step-size formulation follows by rearranging part (i): if $\eta_{\max} + \eta_{\min} < 0$, its condition is equivalent to $0 < \tau < 2/\big(-(\eta_{\max} + \eta_{\min})\big)$; if $\eta_{\max} + \eta_{\min} \geq 0$, it holds for every $\tau > 0$. □

This is a consequence of discretization with no continuous-time counterpart. A very negative continuous rate decays rapidly under $e^{\eta t}$, but if the Euler step is sufficiently large then $1+\tau\eta < -1$, so the same mode grows in magnitude while alternating sign. Therefore we impose a step-size restriction in the discretized flow model, just as Di Giovanni et al. (Di Giovanni et al., 2023) do to exclude this behavior in the graph setting. Equivalently, if $\eta_{\max} + \eta_{\min} < 0$, preservation of the continuous-time asymptotic regime requires

$$\tau < \frac{2}{-(\eta_{\max} + \eta_{\min})},$$

and if $\eta_{\max} + \eta_{\min} \geq 0$ the condition in prop. H.1(i) holds for every $\tau > 0$. The proof uses only the symmetry of the generator, so the proposition holds verbatim for the general flow (16).

I ADJACENCY-LIKE VERSUS LAPLACIAN PROPAGATION

The two tyings of sec. 4.1 differ in how they treat the harmonic sector. Recall that a harmonic mode has

$$\nu_d = \nu_u = 0,$$

and therefore evolves according to the feature-space generator

$$\mathbf{G}^{\mathrm{harm}} = \mathbf{W}_0.$$

Thus the identity weight $\mathbf{W}_0$ determines whether the harmonic component is part of the learned dynamics at all.

For pure Laplacian propagation,

$$\mathbf{W}_0 = 0,$$

so

$$\mathbf{G}^{\mathrm{harm}} = 0.$$

Then,

$$\mathbf{P}_{\mathrm{harm}}\mathbf{F}(t) = \mathbf{P}_{\mathrm{harm}}\mathbf{F}(0)$$

and the harmonic component is preserved exactly and cannot be amplified, suppressed, or mixed across feature channels by the learned weights. This is the higher-order analogue of the observation that pure graph-Laplacian propagation leaves the lowest-frequency component unchanged (Di Giovanni et al., 2023).

By contrast, for the adjacency-like choice

$$\mathbf{W}_0 = \mathbf{W}_{\mathrm{down}} + \mathbf{W}_{\mathrm{up}},$$

we have

$$\mathbf{G}^{\mathrm{harm}} = \mathbf{W}_{\mathrm{down}} + \mathbf{W}_{\mathrm{up}},$$

and hence

$$\mathbf{P}_{\mathrm{harm}}\mathbf{F}(t) = \mathbf{P}_{\mathrm{harm}}\mathbf{F}(0)\exp\big(t(\mathbf{W}_{\mathrm{down}} + \mathbf{W}_{\mathrm{up}})\big).$$

The harmonic component can therefore be amplified, suppressed, and mixed across feature channels. In this way, retaining the identity term allows the learned dynamics to control the harmonic subspace of the features. Recalling the growth rates and regime conditions of sec. 3.3, under pure Laplacian propagation the harmonic growth rate is fixed at

$$\rho_H = 0,$$

while adjacency-like propagation makes

$$\rho_H = \lambda_{\max}\big(\mathbf{W}_{\mathrm{down}} + \mathbf{W}_{\mathrm{up}}\big)$$

learnable. The identity term therefore changes not only the evolution within the harmonic subspace, but also whether harmonic modes can dominate or be suppressed relative to the gradient and curl sectors.

Remark I.1 (Structural bias of pure Laplacian propagation)**.** Let $\beta_1 > 0$. Under pure Laplacian propagation, $\rho_H = 0$, $\rho_G = 2\nu_g^{\max}\big(-\lambda_{\min}(\mathbf{W}_{\mathrm{down}})\big)$ and $\rho_C = 2\nu_c^{\max}\big(-\lambda_{\min}(\mathbf{W}_{\mathrm{up}})\big)$. Hence the flow is HD if and only if $\mathbf{W}_{\mathrm{down}} \succ 0$ and $\mathbf{W}_{\mathrm{up}} \succ 0$ (ties aside), and it is GD or CD whenever either matrix has a negative eigenvalue. A generic symmetric initialization is indefinite, so pure Laplacian propagation is GD or CD at initialization with probability close to one, and harmonic dominance is reachable only if training makes both channel matrices positive definite. Even then the harmonic component is merely preserved, while the other sectors decay by at most a factor $e^{-2\nu\|\mathbf{W}\|_2 t}$ at frequency ν. The adjacency-like flow is free from this restriction: $\rho_H = \lambda_{\max}(\mathbf{W}_{\mathrm{down}} + \mathbf{W}_{\mathrm{up}})$ is learnable and generically nonzero, so all three regimes are available at initialization and the selected regime is determined by training.

J EXPERIMENTAL DETAILS

J.1 DATASETS AND INPUT SIGNALS

Four disks. We reconstruct the example of Grande & Schaub (2024) by sampling 400 points uniformly in each radius-2 disk centered at $(0, 0)$, $(0, 4.2)$, $(4.2, 0)$ and $(4.2, 4.2)$ (seed 0), and constructing a GUDHI alpha complex with $\alpha = 0.45$. It has 1600 vertices, 4671 edges, 3071 triangles and $\beta_1 = 1$. All tasks use this same complex and the normalized operators of App. B.

The shared eight-channel edge input mixes five spectral bands: the lowest three gradient modes, lowest four curl modes, the harmonic mode, and highest six modes of each nonzero family. For each band U_b, we draw a Gaussian mixing matrix R_b with eight columns, normalize $U_b R_b$ to equal Frobenius norm across bands, and rescale their sum to unit RMS (feature seed 20260918). Each observation adds independent Gaussian noise of standard deviation 0.1. Training uses two fresh fields per epoch; validation and testing use four and eight fixed fields (seeds 81101 and 82202).

Ocean drifters. We use the released GitHub data and outer splits for the SCoNe ocean-drifter benchmark of Roddenberry et al., cited in Section 4.3. The fixed complex has 133 vertices, 320 edges, 186 triangles and $\beta_1 = 2$, with 200 trajectory-prefix examples. Inputs are signed net edge-traversal counts. The target is the next vertex among the prefix endpoint's neighbors. No coordinate features or synthetic noise are added. We use the ten supplied 180/20 train/test splits unchanged. Our only additional partition is an inner 144/36 fitting/validation split within each training set, grouping identical input/endpoint pairs and using seed $7000 + s$ for split s.

J.2 FOUR-DISK TARGET LABELS

Following the spectral clustering construction of Grande & Schaub (2024), we cluster edge rows of three embeddings: the harmonic eigenvector, the six highest-frequency gradient eigenvectors, and the six highest-frequency curl eigenvectors from our normalized Hodge Laplacians. Upper-frequency selection is our extension of the original construction.

Before clustering, each row z is multiplied by $\operatorname{sign}(\mathbf{1}^\top z)$. For a zero sum, the first largest-magnitude entry sets the sign. Zero rows remain zero. Rows are not normalized. We apply k-means with 30 initializations and seed 0, using two harmonic classes and five classes for each other task. Label classes are ordered by centroid norm. We note that gray denotes a supervised class (the remainder of the edges after selecting out the clusters) that are included in the classification task.

All tasks share 2802 training, 930 validation and 939 test edges (seed 823). We conduct a transductive experiment. The full complex is available, but optimization only uses training-edge labels.

J.3 MODELS AND TRAINING

Both experiments evolve only edge features using the Euler flows of Section 4.1. Channel weights are independent, symmetric and shared across depth, with no commutation constraint. We do not include activation, bias or feature normalization in or between the propagation layers. A nonlinear readout layer is applied only at the end.

Four disks. We use adjacency-like propagation with width $d = 8$, depth 32 and $\tau = 0.25$, without a learned input lift. Weights are parameterized as $W = (Q + Q^\top)/2$, with initial entries of Q drawn from $\mathcal{N}(0, 0.01^2)$ and no spectral-norm bound. The final field is divided by its whole-field RMS. Each edge is then classified by an affine map of its 36 quadratic features f_i^2 and $\sqrt{2} f_i f_j$ $(i < j)$, retaining information from every Hodge sector.

We train with Adam for 250 epochs, learning rate 0.02, weight decay 10^{-4} and gradient-norm clipping at 5. Each epoch uses one batch of two noisy fields. Cross-entropy is restricted to training edges, with class weights $N_{\text{train}}/(Kn_k)$. We evaluate validation every five epochs and restore the minimum-loss checkpoint We use seeds $0, 1, 2$. Test macro-F1 pools the eight test fields. Figure 3 shows seed 0, with error overlays for the first test field.

Ocean drifters. Both flows use $F_0 = xa^\top$ with learned $a \in \mathbb{R}^{16}$, width 16, depth 16 and $\tau = 0.1$. Each symmetric channel matrix is parameterized as $W = Q/\max\{1, \|Q\|_2\}$. The candidate readout

concatenates the orientation-aligned candidate-edge feature with $(\widehat{B}_1 F)[u, :]$ at candidate vertex u, followed by an affine $32 \to 16$ layer, $\tanh$, and a linear $16 \to 1$ score. Invalid neighbors are masked before the softmax. Both models have 832 parameters and use paired initialization seeds.

For each split and architecture, Adam searches learning rates $\{0.001, 0.003, 0.01\}$ and explicit squared-parameter penalties $\{0, 10^{-4}\}$, using batches of at most 32 and gradient clipping at 1. Selection minimizes validation negative log-likelihood, with at most 1000 epochs and patience 100 (selection seed $12000 + s$). We then refit the chosen configuration on all 180 training examples for the selected epoch count with seeds $0, 1, 2$, giving 30 checkpoints per architecture.

Implementation. Four-disk training uses float32 and two CPU threads; ocean training uses float64 and one thread per worker. Both use PyTorch 2.12.0, NumPy 2.4.6 and Python 3.13.7. Ocean runs used an eight-logical-CPU ARM64 host, with no GPU; summed fitting times, including selection, were 625.0 s for adjacency-like and 587.0 s for Laplacian flow.

J.4 Spectral diagnostics

For each frozen model and joint spatial frequency $(\nu_j^{\mathrm{down}}, \nu_j^{\mathrm{up}})$, we diagonalize

$$G_j = (s - 2\nu_j^{\mathrm{down}})W_{\mathrm{down}} + (s - 2\nu_j^{\mathrm{up}})W_{\mathrm{up}},$$

where $s = 1$ for adjacency-like and $s = 0$ for pure-Laplacian propagation. We compare the three endpoint rates and verify $2 + \tau(\eta_{\max} + \eta_{\min}) > 0$ over all modal eigenvalues. The Euler step-size check passes for every reported checkpoint.

Frozen inference uses exact Euler multipliers $(1 + \tau\eta)^m$, with common rescaling for numerical stability, and records the lower and upper Rayleigh quotients defined in the main text. Four-disk curves use the first test field through depth 65,536; ocean curves extend to 1,024. Figure 4 shows means and pointwise interquartile ranges over test occurrences, grouped by the generator-predicted regime: 600 adjacency-like, 520 Laplacian-gradient and 80 Laplacian-curl occurrences. Figure 2 instead shows each checkpoint's mean over its 20 test inputs. These are frozen-flow diagnostics, not retraining at each depth. The predicted asymptotic limit is conditional on initial excitation of the leading eigenspace.