Hodge Laplacians on Graphs

Topological Data Analysis Spring 2025

Dahlen Elstran and Stephen Hu

May 12, 2025

Background Information

“Cohomology on a Bumper Sticker”

Given two matrices $A \in \mathbb{R}^{m\times n}$ and $B \in \mathbb{R}^{n\times p}$ satisfying the property that

$$AB = 0,$$

a property equivalent to

$$\operatorname{im}(B) \subset \ker(A),$$

the *cohomology group* with respect to A and B is the quotient vector space

$$\ker(A)/\operatorname{im}(B).$$

The harmonic representative

- The cohomology class $x + \mathrm{im}(B) \in \ker(A)/\mathrm{im}(B)$ is a coset of vectors:

$$x + \mathrm{im}(B) := \{x + y \in \mathbb{R}^n : y \in \mathrm{im}(B)\}$$

  for some $x \in \ker(A)$.
- Instead of working with equivalence classes (numerically weird), we can choose a unique representative x_H where x_H is orthogonal to every vector in $\mathrm{im}(B)$:
  - Since $\mathrm{im}(B)^\perp = \ker(B^T)$, we require $x_H \in \ker(A) \cap \ker(B^T)$
- From the first isomorphism theorem:

$$\ker(A)/\mathrm{im}(B) \cong \ker(A) \cap \ker(B^T)$$

- We can redefine the cohomology group as the subspace $\ker(A) \cap \ker(B^T)$ of $\mathbb{R}^n$

Why "harmonic"?

- An element in $\ker(A) \cap \ker(B^T)$ is called "harmonic" because:
  - For $AB = 0$, the *Hodge Laplacian* is defined as:

$$A^T A + BB^T \in \mathbb{R}^{n \times n}$$

  - We can show that:

$$\ker(A^T A + BB^T) = \ker(A) \cap \ker(B^T)$$

  - Thus the harmonic representative x_H satisfies the *Laplace equation*:

$$(A^T A + BB^T)x = 0$$

- Thus, we have: $\ker(A)/\mathrm{im}(B) \cong \ker(A^T A + BB^T)$

Version for Differential Forms

Note that...

The Hodge decomposition is an orthogonal decomposition of the space of differential k-forms on a closed Riemannian manifold.

$$\begin{aligned}\Omega^k(M) &= \mathrm{im}(d_{k-1}) \oplus \mathcal{H}^k(M) \oplus \mathrm{im}(d_k^*) \\ &= \mathcal{H}^k(M) \oplus d\Omega^{k-1}(M) \oplus d^*\Omega^{k+1}(M)\end{aligned}$$

where:

- $\Omega^k(M)$: Space of k-forms on manifold M
- $\mathrm{im}(d_{k-1})$: Exact forms
- $\mathcal{H}^k(M)$: Harmonic k-forms ($\Delta_k\omega = 0$)
- $\mathrm{im}(d_k^*)$: Coexact forms

Hodge decomposition (Lim's Edition)

- The *Hodge decomposition* provides an orthogonal direct sum:

$$\mathbb{R}^n = \operatorname{im}(A^T) \oplus \ker(A^T A + BB^T) \oplus \operatorname{im}(B)$$

- In other words, any $x \in \mathbb{R}^n$ can be uniquely decomposed as:

$$x = A^T w + x_H + Bv, \quad \langle A^T w, x_H \rangle = \langle x_H, Bv \rangle = \langle A^T w, Bv \rangle = 0$$

  for some $v \in \mathbb{R}^p$ and $w \in \mathbb{R}^m$
- Recall from linear algebra that

$$\mathbb{R}^n = \ker(A) \oplus \operatorname{im}(A^T),$$

  so combining this with the Hodge decomposition we get:

$$\mathbb{R}^n = \overbrace{\operatorname{im}(A^T) \oplus \underbrace{\ker(A^T A + BB^T)}_{}}^{\ker(B^T)}\!\!\underbrace{\vphantom{)}\oplus \operatorname{im}(B)}_{\ker(A)}.$$

- This is another illustration that $\ker(A^T A + BB^T) \cong \ker(A) \cap \ker(B^T)$.

Vector Space of k-Chains

Take a simplicial complex X.

Vector Space of k-Chains

The finite-dimensional vector space over $\mathbb{R}$ where:

- the basis elements are the oriented k-simplices $\{s_1^k, s_2^k, \ldots, s_{n_k}^k\}$
- Has dimension $n_k = |X^k|$, the number of k-simplices in the simplicial complex

Hence a k-**chain** $c_k \in C_k$ is a formal linear combination of oriented k-simplices:

$$c_k = \sum_{i=1}^{n_k} \gamma_i s_i^k, \quad \text{where } \gamma_i \in \mathbb{R}$$

One can also consider alternating functions $f : C_k \to \mathbb{R}$ (the cochains C^k, or "discrete differential forms.") Since everything is finite dimensional it doesn't matter too much.

Boundary map

- Recall $\partial_k : C_k \to C_{k-1}$ defined as:

$$\partial_k([i_0, \dots, i_k]) = \sum_{j=0}^{k} (-1)^j [i_0, \dots, i_{j-1}, i_{j+1}, \dots, i_k].$$

  Denote the matrix representation of ∂_k as B_k.
- The coboundary map $\partial^k : C^{k-1} \to C^k$ is the adjoint of the boundary map.
  - Hence its matrix representation is $\boldsymbol{B}_k^T$.

Hodge Laplacians

Definition

From the sequences of boundary maps, one can define a hierarchy of Laplacian operators.

k**th combinatorial Hodge Laplacian:**

$$\boldsymbol{L}_k = \boldsymbol{B}_k^T \boldsymbol{B}_k + \boldsymbol{B}_{k+1} \boldsymbol{B}_{k+1}^T$$

Special cases:

- Standard combinatorial graph Laplacian: $\boldsymbol{L}_0 = \boldsymbol{B}_1 \boldsymbol{B}_1^T$ (as $\boldsymbol{B}_0 = 0$)
- $\boldsymbol{L}_1$: Hodge 1-Laplacian, primary focus of paper

Random Walks and the Normalized Hodge Laplacian

Recall that the graph Laplacian $\boldsymbol{L}_0 := \boldsymbol{D} - \boldsymbol{A}$.
A (standard, unbiased) random walk on a graph with adjacency matrix A follows:

$$p_{t+1} = \boldsymbol{A}\boldsymbol{D}^{-1} p_t = (\boldsymbol{I} - \boldsymbol{L}_0 \boldsymbol{D}^{-1}) p_t,$$

where p_t is the probability distribution over nodes at time t, and $\mathcal{L}_0 = \boldsymbol{L}_0 \boldsymbol{D}^{-1}$ is the random walk Laplacian.

Note the following:

- The transition matrix is closely related to the normalized Hodge Laplacian.
- Harmonic functions of $\mathcal{L}_0$ reflect graph connectivity.
- The graph's topology determines the random walk dynamics.

Random walk on edges

- This fails when we do a random walk on the edges! Orientation makes things hard.
- What about the dual graph?
  - Laplacian of the dual graph isn't linked to original Laplacian.
- Solution? Lift to a new simplicial complex, let $\boldsymbol{L}_1$ act on it, and project back down. (lift, apply, project)

Normalized Hodge 1-Laplacian $\mathcal{L}_1$

Definition

Consider an simplicial complex X whose boundary operators can be represented by the matrices $\boldsymbol{B}_1$ and $\boldsymbol{B}_2$. The normalized Hodge 1-Laplacian matrix is then defined by

$$\mathcal{L}_1 = \boldsymbol{D}_2 \boldsymbol{B}_1^T \boldsymbol{D}_1^{-1} \boldsymbol{B}_1 + \boldsymbol{B}_2 \boldsymbol{D}_3 \boldsymbol{B}_2^T \boldsymbol{D}_2^{-1}, \tag{1}$$

where $\boldsymbol{D}_2$ is the diagonal matrix of (adjusted) degrees of each edge,

$$\boldsymbol{D}_2 = \max(\mathrm{diag}(|\boldsymbol{B}_2|), \boldsymbol{I}) \Leftrightarrow (\boldsymbol{D}_2)_{[i,j],[i,j]} = \max\{\deg([i,j]), 1\}, \tag{2}$$

$\boldsymbol{D}_1 = 2 \cdot \mathrm{diag}(|\boldsymbol{B}_1| \boldsymbol{D}_2 \mathbf{1})$ is a diagonal matrix of weighted degrees of the nodes (with the weight of an edge equal to the maximum of 1 and the number of cofaces of the edge), and $\boldsymbol{D}_3 = \frac{1}{3}\boldsymbol{I}$.

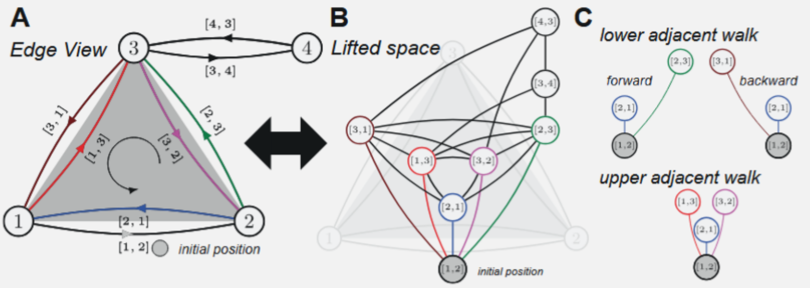

Fig. 3.1 Illustration of lifted simplicial complex. (A) *We may think of the lifted complex as an augmented complex in which each original edge is represented in both possible orientations.* (B) *Alternatively, we may interpret each oriented edge* $[i, j]$ *in the original complex as giving rise to two states* $[i, j]$ *and* $[j, i]$ *on a graph with* $2n_1$ *vertices.* (C) *Starting from* $[1, 2]$, *there are lower adjacent connections "forward" and "backward," as well as upper adjacent connections (see text).*

Stochastic matrix

This normalized version can be seen as a random walk on the lifted simplicial complex!

Stochastic lifting of the normalized Hodge 1-Laplacian

The matrix $-\mathcal{L}_1/2$ has a stochastic lifting, i.e., there exists a column stochastic matrix $\tilde{\boldsymbol{P}}$ such that $-\frac{1}{2}\mathcal{L}_1 \boldsymbol{V}^T = V^T \tilde{\boldsymbol{P}}$. Specifically, $\tilde{\boldsymbol{P}} := \frac{1}{2}\boldsymbol{P}_{\text{lower}} + \frac{1}{2}\boldsymbol{P}_{\text{upper}}$, where $\boldsymbol{P}_{\text{lower}}$ is the transition matrix of a random walk determined by the lower-adjacent connections and $\boldsymbol{P}_{\text{upper}}$ is the transition matrix of a random walk determined by the upper-adjacent connections.

Definitions

Two k-simplices in an simplicial complex X are **upper adjacent** if they are both faces of the same $(k+1)$-simplex and are **lower adjacent** if both have a common face.

Computing the Hodge Decomposition

Normalized Decomposition

For an edge flow (i.e. a 1-cochain) $\boldsymbol{e} \in \mathbb{R}^m$:

$$\boldsymbol{e} = \boldsymbol{g} \oplus \boldsymbol{r} \oplus \boldsymbol{h}$$

where:

- $\boldsymbol{g} = \boldsymbol{D}_2^{1/2} \boldsymbol{B}_1^T \boldsymbol{p}$ (gradient flow)
- $\boldsymbol{r} = \boldsymbol{D}_2^{-1/2} \boldsymbol{B}_2 \boldsymbol{w}$ (curl flow)
- $\boldsymbol{C}_1^T \boldsymbol{h} = \boldsymbol{0}$ (harmonic flow constraint)

Computation via Least Squares

Solve two least squares problems:

$$\min_{\boldsymbol{p}} \|\boldsymbol{D}_2^{1/2} \boldsymbol{B}_1^T \boldsymbol{p} - \boldsymbol{e}\|_2 \quad \text{and} \quad \min_{\boldsymbol{w}} \|\boldsymbol{D}_2^{-1/2} \boldsymbol{B}_2 \boldsymbol{w} - \boldsymbol{e}\|_2$$

Computation via Least Squares

Practical Implementation

Using the residual error vectors:

$$\boldsymbol{e}_p = \boldsymbol{D}_2^{1/2}\boldsymbol{B}_1^T\boldsymbol{p}^* - \boldsymbol{e}$$

$$\boldsymbol{e}_w = \boldsymbol{D}_2^{-1/2}\boldsymbol{B}_2\boldsymbol{w}^* - \boldsymbol{e}$$

Final decomposition:

$$\boldsymbol{g} = \boldsymbol{e}_p + \boldsymbol{e} = \boldsymbol{D}_2^{1/2}\boldsymbol{B}_1^T\boldsymbol{p}^*$$

$$\boldsymbol{r} = \boldsymbol{e}_w + \boldsymbol{e} = \boldsymbol{D}_2^{-1/2}\boldsymbol{B}_2\boldsymbol{w}^*$$

$$\boldsymbol{h} = \boldsymbol{e} - \boldsymbol{g} - \boldsymbol{r}$$

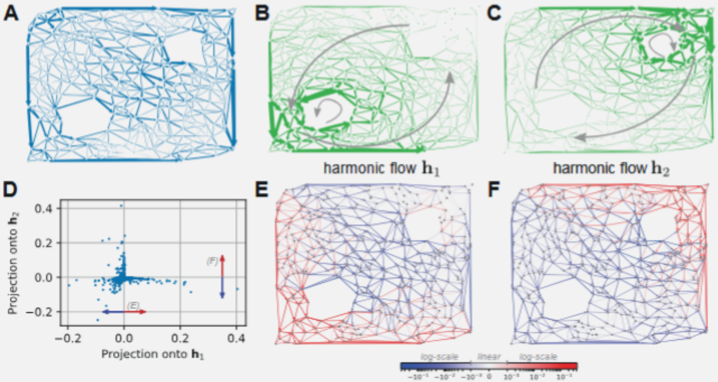

Fig. 5.1 Embedding of edge flows. (A) *A flow on an SC with two "holes," constructed as described in the text. Arrows indicate the direction of the flow on each edge; the magnitude is proportional to the width of the edge.* (B)–(C) *Harmonic functions* $\boldsymbol{h}_1$ *and* $\boldsymbol{h}_2$ *of the symmetric normalized Hodge 1-Laplacian* $\mathcal{L}_1^s$ *of the underlying SC. Edge directions correspond to the orientation induced by each harmonic function. Gray arrows indicate how the harmonic flows encircle the two holes.* (D) *Projection of each edge flow* $f_{[i,j]}$ *(depicted in Figure* 5.1A*) onto the harmonic functions.* (E)–(F) *Projection of the edge flow onto the harmonic functions* $\boldsymbol{h}_1$ (E) *and* $\boldsymbol{h}_2$ (F)*. Red indicates a positive projection, blue a negative projection. The arrow direction is the same as in* (A).

Applications

Dataset

- Compiled by the National Oceanic and Atmospheric Administration, the data we used modeled the movement of floating ocean buoys, referred to as drifters. These drifters are tracked with GPS to measure how surface currents move across the ocean over time.
- The dataset was millions of lines long, with a large amount of data. For the sake of efficiency, like the paper we followed, we proceeded with our analysis on a small portion of the data.

Generating Nodes and Edges

- Once the data is read, it must be sorted into nodes
- Each node represented a unique latitude and longitude point
- Each edge represented one drifter's movements from one node to another
- From here, we can construct a directed graph.

Building the Boundary Matrix and Computing the Hodge Laplacian

- We construct the boundary matrix like usual, with zeroes everywhere other than 1's (-1's) to indicate a positive (negative) edge between two nodes
- Next, the node degree matrix was found. This is a diagonal matrix where each entry (i, i) is the degree of node i.
- Finally, the Hodge Laplacian, and it's normalized version, could be calculated from the equations given previously.

Harmonic Eigenvectors and Flow Fields

- The smallest 5 eigenvalues, and their respecting eigenvectors, were found from the normalized Hodge Laplacian
- This forms an orthonormal basis for the harmonic subspace
- We can then visualize the Harmonic Flow Fields.

Hodge Decomposition

- We can also use the eigenvectors we found to compute the gradient flow, curl flow, and harmonic flow constraint
- From here, we can compute the Hodge Decomposition
- Finally, we plot the Hodge Decomposition components.

The End