

## PERSISTENT SHEAF LAPLACIANS

XIAOQI WEI<sup>✉1</sup> AND GUO-WEI WEI<sup>✉1,2,3</sup>

<sup>1</sup>Department of Mathematics, Michigan State University,  
MI 48824, USA

<sup>2</sup>Department of Electrical and Computer Engineering, Michigan State University,  
MI 48824, USA

<sup>3</sup>Department of Biochemistry and Molecular Biology, Michigan State University,  
MI 48824, USA

(Communicated by Vasileios Maroulas)

**ABSTRACT.** Recently, various types of topological Laplacians have been studied from the perspective of data analysis. The spectral theory of these Laplacians has significantly extended the scope of algebraic topology and data analysis. Inspired by the theory of persistent Laplacians and cellular sheaves, this work develops the theory of persistent sheaf Laplacians for cellular sheaves and describes how to construct sheaves for a point cloud where each point is associated with a quantity that can be devised to embed physical properties. The spectra of persistent sheaf Laplacians encode both geometrical and non-geometrical information of the given point cloud. The theory of persistent sheaf Laplacians provides an elegant method for fusing different types of data and has significant potential for future development.

**1. Introduction.** Recent years have witnessed a dramatic growth of research interest in topological data analysis (TDA) [19, 40], driven by its success in science and technology, particularly in computational biology and computer-aided drug design, one of the most challenging scientific fronts of the 21<sup>st</sup> century [27, 28]. The main workhorse of TDA is persistent homology [6, 12, 45], a new branch of algebraic topology that is able to capture multiscale topological information of data. Topological deep learning, coined in 2017, integrates persistent homology and artificial intelligence (AI) [4]. It has had enormous success in data science [23, 26, 37].

Persistent homology has many limitations. For example, it cannot capture the shape evolution of data that does not change topologically. Evolutionary de Rham-Hodge theory, or persistent Hodge Laplacians, was proposed in 2019 to overcome this limitation [8]. The de Rham-Hodge theory connects algebraic topology with differentiable geometry, partial differential equations, harmonic analysis, and algebraic geometry [11, 44]. Persistent Hodge Laplacians defined on a family of differentiable manifolds offer a powerful multiscale spectral analysis of volumetric data [8].

---

2020 *Mathematics Subject Classification.* Primary: 62R40; Secondary: 92B99.

*Key words and phrases.* Persistent Laplacian, persistent sheaf Laplacian, persistent spectral graph, combinatorial graph, Hodge Laplacian, algebraic topology, data fusion.

This work was supported in part by NIH grants R01GM126189, R01AI164266, and R01AI146210, NSF grants DMS-2052983, DMS-1761320, and IIS-1900473, NASA grant 80NSSC21M0023, MSU Foundation, Bristol-Myers Squibb 65109, and Pfizer.

<sup>✉</sup>Corresponding author: Xiaoqi Wei.

Additionally, persistent spectral graph theory, also called persistent combinatorial Laplacians (or simply persistent Laplacians), was also introduced in 2019 to overcome the drawback of persistent homology [38]. Combinatorial Laplacians enable us to perform spectral analysis for simplicial complexes constructed from a point cloud [9, 15, 20]. Persistent combinatorial Laplacians have stimulated much interest in the past few years [21, 22, 39] and has had success in various applications [7, 24, 29]. Both persistent Hodge Laplacians and persistent combinatorial Laplacians are persistent topological Laplacians, an emerging research topic in TDA. However, these PTLs cannot deal with labeled data, or the local non-geometric properties at individual data points.

A natural extension of the aforementioned algebraic topology, differential geometry, and graph theory approaches for data science is the sheaf theory. A sheaf is a mathematical tool for systematically tracking data (or states), or extracting spatial-temporal patterns, or organizing data (or states), or fusing data (or states) according to certain physical/mathematical rules or structures. It defines a relationship for adjacent data points lying in a topological space, which may be abstracted from point clouds, manifolds, or spatial-temporal data. Such relationships become particularly valuable when they are devised to embed certain physical laws governing the underlying data. A particular class of sheaves, namely cellular sheaves, has attracted much attention in the past decade for its application potentials. The notion of a *cellular sheaf* was first introduced by Shepard in his Ph.D. thesis in 1985 [35]. In the past decade, researchers have revived and expanded the cellular sheaf theory with applications in science and engineering [10, 17, 32]. Tools from sheaf theory have also been developed to study persistence modules [2, 3, 10, 16]. Roughly speaking, a cellular sheaf consists of a simplicial complex, an assignment of vector spaces for each simplex, and a definition of linear maps for each face relation, satisfying certain rules so that it gives rise to a *sheaf cochain complex*. Therefore, one can define sheaf cohomology, which reflects the property of the sheaf. Spectral sheaf theory extends spectral graph theory to cellular sheaves, leading to sheaf Laplacians [17].

The aim of this paper is to introduce persistent sheaf Laplacians (PSL) as an extension of persistent Laplacians to the setting of cellular sheaves. This work is also motivated by the need to elegantly fuse geometric and non-geometric information as in the persistent cohomology [5]. In this approach, the geometrical shape can be successfully captured by persistent homology and particularly, by persistent Laplacians, while local non-geometric properties, such as atomic biophysical properties of a molecule, requires additional treatment. It is beneficial to have mathematical tools that can distinguish between data that have very similar geometrical shapes but carrying different non-geometrical information. From a mathematical perspective, the extension of persistent Laplacians is natural, since the key theorem in the theory of persistent Laplacians is indeed true for more general (co)chain complexes. The rest of this paper is organized as follows. In Section 2, we introduce the basics of cellular sheaf theory and discuss how to define a sheaf on a labeled simplicial complex (i.e., a simplicial complex with a quantity associated to each vertex). In Section 3, we define the  $q$ -th persistent sheaf Laplacian for a pair of complexes  $X, Y$ . In Section 4 we demonstrate the spectra of persistent sheaf Laplacians for several examples of applications.

**2. Preliminaries.** In this section, we give a very concise introduction to cellular sheaves and persistent Laplacians. We assume that the reader is familiar with simplicial homology. We follow the standard notational convention to distinguish between chain complexes and cochain complexes. Most researchers define cellular sheaves on regular cell complexes [10, 13, 17, 32]. For the sake of simplicity, we only discuss cellular sheaves on simplicial complexes.

### 2.1. Cellular sheaves.

**Definition 2.1.** A cellular sheaf  $\mathcal{S}$  on a simplicial complex  $X$  consists of the following data:

- (1) a simplicial complex  $X$ , where the face relation that  $\sigma$  is a face of  $\tau$  is denoted by  $\sigma \leq \tau$ , and
- (2) an assignment to each simplex  $\sigma$  of  $X$  a (finite dimensional) vector space  $\mathcal{S}(\sigma)$  and to each face relation  $\sigma \leq \tau$  a linear morphism of vector spaces denoted by  $\mathcal{S}_{\sigma \leq \tau}$  or  $\mathcal{S}(\sigma \leq \tau) : \mathcal{S}(\sigma) \rightarrow \mathcal{S}(\tau)$ , satisfying the rule

$$\rho \leq \sigma \leq \tau \Rightarrow \mathcal{S}_{\rho \leq \tau} = \mathcal{S}_{\sigma \leq \tau} \circ \mathcal{S}_{\rho \leq \sigma}$$

and  $\mathcal{S}_{\sigma \leq \sigma} = \text{id}$  is the identity map.

The vector space  $\mathcal{S}(\sigma)$  is referred to as the stalk of  $\mathcal{S}$  over  $\sigma$ , and the linear morphism  $\mathcal{S}_{\sigma \leq \tau}$  is referred to as the restriction map of the face relation  $\sigma \leq \tau$ .

**Example 2.1.** Let  $X$  be a finite simplicial complex. We attach to every simplex of  $X$  a fixed vector space  $V$  and let every restriction map be the identity map. This sheaf is referred to as the constant sheaf  $\underline{V}$  on  $X$ .

**Definition 2.2.** Suppose that  $f : X \rightarrow Y$  is a simplicial map [25] and  $\mathcal{S}$  is a cellular sheaf on  $Y$ . The pullback sheaf  $f^*\mathcal{S}$  on  $X$  is given by

$$(f^*\mathcal{S})(\sigma) = \mathcal{S}(f(\sigma)),$$

and for any face relation  $\sigma \leq \tau$  of  $X$ ,

$$(f^*\mathcal{S})_{\sigma \leq \tau} = \mathcal{S}_{f(\sigma) \leq f(\tau)}.$$

**Example 2.2.** Suppose that  $X$  is a subcomplex of  $Y$  and  $\mathcal{S}$  is a sheaf on  $Y$ . We can define a sheaf  $\mathcal{T}$  on  $X$  using the data of  $Y$ . For  $\sigma \in X$ , let  $\mathcal{T}(\sigma) = \mathcal{S}(\sigma)$ . For any face relation  $\sigma \leq \tau$  in  $X$ , let  $\mathcal{T}(\sigma \leq \tau) = \mathcal{S}(\sigma \leq \tau)$ . The sheaf  $\mathcal{T}$  is a pullback of  $\mathcal{S}$ .

**Definition 2.3.** Suppose that  $\mathcal{S}$  is a sheaf. A global section  $s$  of  $\mathcal{S}$  is an assignment to each simplex  $\sigma$  an element  $s_\sigma \in \mathcal{S}(\sigma)$  such that  $\mathcal{S}_{\sigma \leq \tau}(s_\sigma) = s_\tau$  for any face relation  $\sigma \leq \tau$ . The set of global sections is denoted by  $\Gamma(X; \mathcal{S})$ .

**Example 2.3.** This example is related to quantum physics. The general form of the Schrödinger equation for an isolated quantum system is

$$i\hbar \frac{\partial}{\partial t} |\psi(t)\rangle = H |\psi(t)\rangle$$

where  $t$  is time,  $|\psi(t)\rangle = |\psi(\mathbf{r}, t)\rangle$  is the state vector of the quantum system that belongs to a Hilbert space  $\mathcal{H}$ ,  $\mathbf{r}$  is the position vector, and  $H$  is the Hamiltonian consisting of kinetic energy and potential energy operators. In the position space representation, the kinetic energy operator is given by the Laplacian. In the Schrödinger representation,  $H$  is independent of time ( $\frac{\partial H}{\partial t} = 0$ ), and the time evolution of  $|\psi(t)\rangle$  is given by

$$|\psi(t)\rangle = e^{-iH(t-t_0)/\hbar} |\psi(t_0)\rangle,$$

where  $e^{-iH(t-t_0)/\hbar}$  is known as the time-evolution operator. We can define a cellular sheaf  $\mathcal{S}$  as follows. First  $\mathbb{R}$  is seen as a simplicial complex such that its vertices are integers and its edges are intervals  $[n, n+1]$  where  $n \in \mathbb{Z}$ . We attach  $\mathcal{H}$  to each simplex. Restriction maps  $\mathcal{S}_{n \leq [n, n+1]}$  and  $\mathcal{S}_{n+1 \leq [n, n+1]}$  are defined by linear maps  $e^{\frac{-iH}{2\hbar}}$  and  $e^{\frac{iH}{2\hbar}}$ . The assignment  $n \rightarrow |\psi(n)\rangle$  for all  $n$  is a global section of sheaf  $\mathcal{S}$ . Similarly, cellular sheaves can be defined for many other linear (partial) differential equations. Note that in general there is no need to work with  $\mathbb{Z}$ .

**2.2. Sheaf cohomology.** For a sheaf  $\mathcal{S}$  on a finite simplicial complex  $X$ , we can construct the sheaf cochain complex of  $\mathcal{S}$  as follows. Let the  $q$ -th cochain group  $C^q(X; \mathcal{S})$  be the direct sum of  $\mathcal{S}(\sigma)$  over all  $q$ -simplices  $\sigma$ . To define the coboundary map  $d$ , we need a signed incidence relation [10].

**Definition 2.4.** A signed incidence relation is an assignment to every face relation  $\sigma \leq \tau$  an integer  $[\sigma : \tau]$  satisfying the following conditions:

- (1) if  $\dim \tau - \dim \sigma > 1$ , then  $[\sigma : \tau] = 0$ ; and
- (2) if  $\gamma \leq \tau$  and  $\dim \tau - \dim \gamma = 2$ , the sum  $\sum_{\sigma} [\gamma : \sigma][\sigma : \tau] = 0$ .

If a signed incidence relation is given, we then define the coboundary map  $d_q : C^q(X; \mathcal{S}) \rightarrow C^{q+1}(X; \mathcal{S})$  by

$$d^q|_{\mathcal{S}(\sigma)} = \sum_{\sigma \leq \tau} [\sigma : \tau] \mathcal{S}_{\sigma \leq \tau}.$$

Since  $d^q$  is a linear morphism, its action on each stalk  $\mathcal{S}(\sigma)$  determines itself. We can verify that  $d^q d^{q-1} = 0$  [10, Lemma 6.2.2], so there is the sheaf cochain complex

$$0 \longrightarrow C^0(X; \mathcal{S}) \xrightarrow{d} C^1(X; \mathcal{S}) \xrightarrow{d} C^2(X; \mathcal{S}) \xrightarrow{d} \cdots.$$

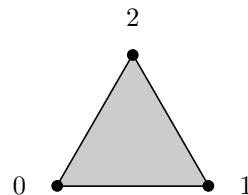
The  $q$ -th sheaf cohomology group  $H^q(X; \mathcal{S})$  is defined by  $\ker d^q / \text{im } d^{q-1}$ .

A natural signed incidence relation exists for every oriented simplicial complex. Recall that the orientation of a simplex is determined by the ordering of its vertices. For an oriented simplex  $\tau = [v_0, v_1, \dots, v_n]$  and its oriented face  $\sigma = [v_0, \dots, \hat{v}_i, \dots, v_n]$ , we let  $[\sigma : \tau] = (-1)^i$ . If  $\sigma$  or  $\tau$  is oriented alternatively, we let  $[\sigma : \tau] = (-1)^{i+1}$ . This signed incidence relation is used throughout this paper. In practice we can orient a simplicial complex by a global ordering of vertices. We remind the reader that we do not need orientation information to define a sheaf.

**Example 2.4.** We examine the sheaf cochain complex of the constant sheaf  $\mathbb{R}$  on a finite simplicial complex  $X$ . The dimension of the  $q$ -th cochain group  $C^q(X; \mathcal{S})$  is equal to the number of  $q$ -simplices. We use simplices to distinguish different stalks  $\mathbb{R}$  over different simplices. The  $q$ -th sheaf cochain group has the canonical basis  $\{\sigma \mid \dim \sigma = q\}$ , and

$$d^q \sigma = \sum_{\sigma \leq \tau} [\sigma : \tau] \tau.$$

Suppose  $X$  is  $\{0, 1, 2, 01, 02, 12, 012\}$ ,



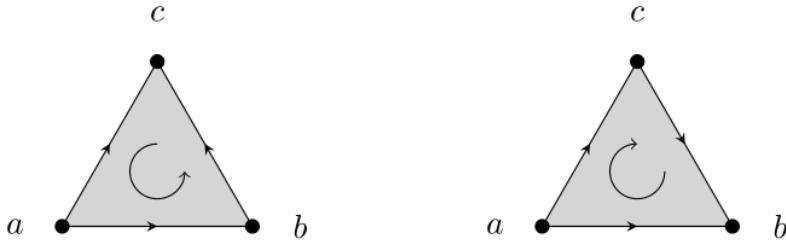


FIGURE 1. Different orderings determine different orientations. On the left, the ordering is  $a, b, c$ , and the oriented simplices are  $[a, b], [b, c], [a, c], [a, b, c]$ . On the right, the ordering is  $a, c, b$ , and the oriented simplices are  $[a, b], [c, b], [a, c], [a, c, b]$ . In practice, one would label vertices by natural numbers, and orient simplices such that labels are increasing. For example, if we label  $a, b, c$  by  $0, 1, 2$ , then we will get the left orientation. For the sake of simplicity, we often denote a vertex by its index, so  $v_i$  is written as  $i$ ; for edges, we just write  $ij$  instead of  $[v_i, v_j]$ , and adopt similar notations for higher dimensional simplices.

then the matrix representation of  $d^0$  is

$$\begin{matrix} 0 & 1 & 2 \\ 01 & \begin{pmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \\ -1 & 0 & 1 \end{pmatrix} \\ 12 & \\ 02 & \end{matrix}.$$

and the matrix representation of  $d^1$  is

$$\begin{matrix} 01 & 02 & 12 \\ 012 & \begin{pmatrix} 1 & -1 & 1 \end{pmatrix} \end{matrix}.$$

In fact for any finite simplicial complex  $X$ , the sheaf cochain complex of  $\mathbb{R}$  coincides with the dual of the simplicial chain complex of  $X$  with coefficient ring  $\mathbb{R}$ .

The following fact is well-known and we omit the proof.

**Proposition 2.5.** [13, Lemma 9.5]  $H^0(X; \mathcal{S}) = \Gamma(X; \mathcal{S})$ .

**2.3. Cellular sheaves on a labeled simplicial complex.** Suppose that there is a 1-dimensional simplicial complex (i.e., graph)  $X$  where each vertex  $v_i$  is associated with a quantity  $q_i \in \mathbb{R}$ . Denote the edge connecting  $v_i$  and  $v_j$  by  $e_{ij}$ . We can define a sheaf  $\mathcal{S}$  on  $X$  with  $q_i$  such that each stalk is  $\mathbb{R}$ , and for the face relation  $v_i \leq e_{ij}$ , the morphism  $\mathcal{S}_{v_i \leq e_{ij}}$  is the multiplication by  $q_j/r_{ij}$  where  $r_{ij}$  is the length of  $e_{ij}$ . This sheaf is inspired by molecular biology and chemistry. Given a molecule, each atom and its atomic partial charge can be seen as a vertex  $v_i$  and the associated quantity  $q_i$ , respectively. Conceptually, one may build a 1-dimensional Rips or Alpha complex from the point cloud of vertices or just take chemical bonds as edges. The assignment  $q_i \rightarrow v_i$  and  $q_i q_j / r_{ij} \rightarrow e_{ij}$  is a global section, since  $\mathcal{S}_{v_i \leq e_{ij}}(q_i) = \mathcal{S}_{v_j \leq e_{ij}}(q_j) = q_i q_j / r_{ij}$ . The quantity  $q_i q_j / r_{ij}$  is the potential energy.

The above sheaf can be generalized to high dimensional simplicial complexes (cf. [42]). Suppose we have the following data: a simplicial complex  $X$ , a set

of nonzero  $q_i \in \mathbb{R}$  associated to each vertex  $v_i$ , and a nowhere zero function  $F : X \rightarrow \mathbb{R}$ . We can define a sheaf where each stalk is  $\mathbb{R}$ , and for the face relation  $[v_0, \dots, v_n] \leq [v_0, \dots, v_n, v_{n+1}, \dots, v_m]$ , the linear morphism  $\mathcal{S}([v_0, \dots, v_n] \leq [v_0, \dots, v_n, v_{n+1}, \dots, v_m])$  is the scalar multiplication by

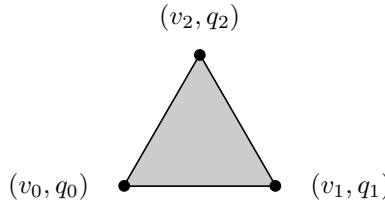
$$\frac{F([v_0, \dots, v_n])q_{n+1} \cdots q_m}{F([v_0, \dots, v_n, v_{n+1}, \dots, v_m])}.$$

This is indeed a sheaf since if we have  $[v_0, \dots, v_n] \leq [v_0, \dots, v_m] \leq [v_0, \dots, v_l]$ , then

$$\frac{F([v_0, \dots, v_m])q_{m+1} \cdots q_l}{F([v_0, \dots, v_l])} \frac{F([v_0, \dots, v_n])q_{n+1} \cdots q_m}{F([v_0, \dots, v_m])} = \frac{F([v_0, \dots, v_n])q_{n+1} \cdots q_l}{F([v_0, \dots, v_l])}.$$

The assignment  $q_{i_0} \cdots q_{i_n} / F([v_{i_0}, \dots, v_{i_n}]) \rightarrow [v_{i_0}, \dots, v_{i_n}]$  is a nontrivial global section.

**Example 2.5.** Consider the oriented simplicial complex  $\{v_0, v_1, v_2, v_0v_1, v_1v_2, v_0v_2, v_0v_1v_2\}$



where  $q_i \in \mathbb{R}$  is associated to  $v_i$ . Let  $r_{01}, r_{12}, r_{02}$  be the lengths of  $e_{01}, e_{12}, e_{02}$ . We can define the above sheaf on this complex where  $F$  maps every vertex to 1, every edge  $e_{ij}$  to its length  $r_{ij}$ , and the 2-simplex  $[v_0, v_1, v_2]$  to  $r_{01}r_{12}r_{02}$ . The matrix representation of  $d^0$  is

$$\begin{matrix} & v_0 & v_1 & v_2 \\ v_0v_1 & \begin{pmatrix} -q_1/r_{01} & q_0/r_{01} & 0 \\ 0 & -q_2/r_{12} & q_1/r_{12} \\ -q_2/r_{02} & 0 & q_0/r_{02} \end{pmatrix}, \\ v_1v_2 & & & \\ v_0v_2 & & & \end{matrix}$$

and the matrix representation of  $d^1$  is

$$\begin{matrix} & v_0v_1 & v_1v_2 & v_0v_2 \\ v_0v_1v_2 & \begin{pmatrix} \frac{q_2}{r_{02}r_{12}} & \frac{q_0}{r_{01}r_{02}} & \frac{-q_1}{r_{01}r_{12}} \end{pmatrix}. \end{matrix}$$

Note that many alternative sheaf constructions are available by appropriate choices of  $F$ . For example,  $F$  may map a 2-cell to the sum of its edge lengths. In practical applications, a good choice of  $F$  can embed physical information into spectral representation.

**2.4. Sheaf Laplacians.** If cochain groups of a cochain complex

$$\dots \xrightarrow{d^{q-2}} A^{q-1} \xrightarrow{d^{q-1}} A^q \xrightarrow{d^q} A^{q+1} \xrightarrow{d^{q+1}} \dots$$

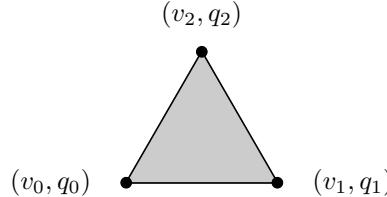
are all finite dimensional inner product spaces, the  $q$ -th combinatorial Laplacian  $\Delta_q : A^q \rightarrow A^q$  is defined by

$$\Delta_q = (d^q)^* d^q + d^{q-1} (d^{q-1})^*,$$

where  $(d^q)^*$  is the adjoint of  $d^q$ , and it is well-known that the kernel of  $\Delta_q$  is isomorphic to the  $q$ -th cohomology group  $H^q$ . Hansen and Ghrist [17] applied

this construction to sheaf cochain complexes and the resulting new combinatorial Laplacian is referred to as the sheaf Laplacian. If every stalk of a sheaf  $\mathcal{S}$  is a finite dimensional inner product space, we can equip an inner product structure on every  $C^q(X; \mathcal{S})$  such that  $\mathcal{S}(\sigma)$  and  $\mathcal{S}(\sigma')$  are orthogonal if  $\sigma \neq \sigma'$ .

**Example 2.6.** We consider the  $F$  defined in the same way as in Example 2.5 to evaluate the spectra of sheaf Laplacians. Consider the 2-simplex



whose edges are all of length 1. The matrix representation of  $\Delta_0$  is

$$\begin{pmatrix} q_1^2 + q_2^2 & -q_0 q_1 & -q_0 q_2 \\ -q_0 q_1 & q_0^2 + q_2^2 & -q_1 q_2 \\ -q_0 q_2 & -q_1 q_2 & q_0^2 + q_1^2 \end{pmatrix}$$

and its eigenvalues are  $\{q_0^2 + q_1^2 + q_2^2, q_0^2 + q_1^2 + q_2^2, 0\}$  and the corresponding eigenvectors are  $(-q_1/q_0, 1, 0)^T$ ,  $(-q_2/q_0, 0, 1)^T$ , and  $(q_0/q_2, q_1/q_2, 1)^T$ . Moreover, the matrix representation of  $\Delta_1$  is

$$\begin{pmatrix} q_0^2 + q_1^2 + q_2^2 & 0 & 0 \\ 0 & q_0^2 + q_1^2 + q_2^2 & 0 \\ 0 & 0 & q_0^2 + q_1^2 + q_2^2 \end{pmatrix}$$

and its only eigenvalue is  $q_0^2 + q_1^2 + q_2^2$ .

This example shows that the eigenvalues of  $\Delta_0$  and  $\Delta_1$  are dependent on the amplitude of  $q_i$ , which allows the embedding of non-geometric information in practical applications. However, they are not sensitive to the sign of  $q_i$ . Therefore, a (persistent) sheaf Dirac theory as an extension of recent Dirac formulation or quantum persistent homology [1] may enable us to further eliminate the sign degeneracy.

Furthermore, the eigenvectors of  $\Delta_0$  depend on the signs of  $q_i$ . As such, it is advantageous to use sheaf eigenvectors to embed physical properties of a molecule and utilize them in machine learning.

**3. Persistent sheaf Laplacians.** We review persistent combinatorial Laplacians before introducing persistent sheaf Laplacians.

**3.1. Persistent Laplacians.** Given two chain complexes  $(V, d)$  and  $(W, d)$ , if  $V_q$ , as an inner product space, is a subspace of the finite dimensional inner product space  $W_q$  for all  $q$ , and the boundary operator of  $(V, d)$  inherits from  $(W, d)$ , then we have the following commutative diagram

$$\begin{array}{ccccccc} \cdots & \xrightarrow{d_{q+2}} & V_{q+1} & \xrightarrow{d_{q+1}} & V_q & \xrightarrow{d_q} & V_{q-1} \xrightarrow{d_{q-1}} \cdots \\ & & \downarrow & & \downarrow & & \downarrow \\ \cdots & \xrightarrow{d_{q+2}} & W_{q+1} & \xrightarrow{d_{q+1}} & W_q & \xrightarrow{d_q} & W_{q-1} \xrightarrow{d_{q-1}} \cdots \end{array}$$

where dashed hooked arrows represent the inclusion map  $\iota : V \hookrightarrow W$ . If we use superscripts  $V$  and  $W$  to distinguish among maps and subspaces of  $(V, d)$  and  $(W, d)$ , then the  $q$ -th persistent homology group  $H_q(V, W)$  is defined by

$$\frac{Z_q^V}{B_q^W \cap Z_q^V} \cong \iota^\bullet(H_q(V)),$$

whose dimension is called the  $q$ -th persistent Betti numbers of the pair  $V, W$ . Let  $\Theta_{q+1}^{V,W} = \{x \in W_{q+1} \mid d_{q+1}^W x \in V_q\}$ . In other words,  $\Theta_{q+1}^{V,W}$  consists of chains of  $W_{q+1}$  whose boundaries are in  $V_q$  (actually in  $Z_q^V$ , since  $d^2 = 0$ ), and we have  $\text{im } \Theta_{q+1}^{V,W} = B_q^W \cap Z_q^V$ . If we denote the adjoint of the inclusion  $V \hookrightarrow W$  by  $\pi$  (which is a projection), the  $q$ -th persistent Laplacian  $\Delta_q^{V,W}$  [21, 22, 38] is defined by

$$\Delta_q^{V,W} = (d_q^V)^* d_q^V + \pi d_{q+1}^W |_{\Theta_{q+1}^{V,W}} (\pi d_{q+1}^W |_{\Theta_{q+1}^{V,W}})^*.$$

Since we know that  $\ker \Delta_q^{V,W} \cong H_q(V, W)$  [21, 22, 38], the theory of persistent Laplacian offers an alternative method of computing persistent Betti numbers. Besides that, non-zero eigenvalues and eigenvectors of a persistent Laplacian contain extra information that can be utilized.

**3.2. Persistent sheaf cohomology and persistent sheaf Laplacians.** Persistent sheaf cohomology is known to many researchers [31, 33, 43]. Given two oriented simplicial complexes  $X, Y$ , if  $X \subset Y$  and the orientation of  $X$  is identical to  $Y$ , we can let sheaf  $\mathcal{F}$  on  $X$  to be the pullback of the sheaf  $\mathcal{G}$  on  $Y$ , then we have the following commutative diagram

$$\begin{array}{ccccccc} \cdots & \xrightarrow{d} & C^{q-1}(X; \mathcal{F}) & \xrightarrow{d} & C^q(X; \mathcal{F}) & \xrightarrow{d} & C^{q+1}(X; \mathcal{F}) \xrightarrow{d} \cdots \\ & & \uparrow \pi & & \uparrow \pi & & \uparrow \pi \\ \cdots & \xrightarrow{d} & C^{q-1}(Y; \mathcal{G}) & \xrightarrow{d} & C^q(Y; \mathcal{G}) & \xrightarrow{d} & C^{q+1}(Y; \mathcal{G}) \xrightarrow{d} \cdots \end{array}$$

where  $\pi : C^q(Y; \mathcal{G}) \rightarrow C^q(X; \mathcal{F})$  is a projection map such that  $\pi|_{\mathcal{G}(\sigma)}$  is the identity map if  $\sigma \in X$ , and  $\pi|_{\mathcal{G}(\sigma)} = 0$  otherwise. Since  $\pi$  is a cochain map, it induces a map  $\pi^\bullet$  between sheaf cohomology groups of  $\mathcal{F}$  and  $\mathcal{G}$ , and the  $q$ -th persistent sheaf group is defined by

$$\pi^\bullet(H^q(Y; \mathcal{G}))$$

whose dimension is the  $q$ -th persistent sheaf Betti number. Our goal here is to extend the notion of persistent Laplacian to the setting of cellular sheaves. We will assume that each stalk of  $\mathcal{G}$  is a finite dimensional inner product space, and regard any cochain group as a direct sum of inner product spaces. Note that the adjoint of  $\pi$  is the inclusion map  $\iota : C^q(X; \mathcal{F}) \hookrightarrow C^q(Y; \mathcal{G})$ . We can dualize the above diagram (i.e., reverse all arrows) and define the persistent sheaf Laplacian by the persistent Laplacian of the dualized diagram. More specifically, we have the

following diagram

$$\begin{array}{ccccc}
 C^{q-1}(X; \mathcal{F}) & \xrightleftharpoons[d_X^{q-1}]{} & C^q(X; \mathcal{F}) & \xrightleftharpoons[d_X^q]{} & \Theta_{X,Y}^{q+1} \\
 (d_X^{q-1})^* & & \downarrow (d_{X,Y}^q)^* & \searrow d_{X,Y}^q & \\
 & & \Theta_{X,Y}^{q+1} & & \\
 & \downarrow & & \searrow & \\
 C^q(Y; \mathcal{G}) & \xrightleftharpoons[d_Y^q]{} & C^{q+1}(Y; \mathcal{G}) & \xrightleftharpoons[(d_Y^q)^*]{} & 
 \end{array}$$

(note that an inner product space is self-dual) where  $\Theta_{X,Y}^{q+1} = \{x \in C^{q+1}(Y; \mathcal{G}) \mid (d_Y^q)^*(x) \in C^q(X; \mathcal{F})\}$  and  $d_{X,Y}^q$  is the adjoint of  $\pi(d_Y^q)^*|_{\Theta_{X,Y}^{q+1}} : \Theta_{X,Y}^{q+1} \rightarrow C^q(X; \mathcal{F})$ . We define the  $q$ -th persistent sheaf Laplacian  $\Delta_q^{X,Y}$  by

$$\Delta_q^{X,Y} = (d_{X,Y}^q)^* d_{X,Y}^q + d_X^{q-1} (d_X^{q-1})^*.$$

The nullity of  $\Delta_q^{X,Y}$  is equal to the  $q$ -th persistent Betti number of the dualized diagram (since we dualize everything, the two cochain complexes become chain complexes). By the universal coefficient theorem for cohomology,  $\pi^\bullet$  and  $\iota^\bullet$  have the same rank (here  $\bullet$  means the induced map between cohomology or homology groups). So the  $q$ -th persistent Betti number of the dualized diagram is equal to the  $q$ -th persistent sheaf Betti number. In other words, we have

$$\ker \Delta_q^{X,Y} \cong \pi^\bullet(H^q(Y; \mathcal{G})).$$

The matrix representation of a persistent sheaf Laplacian can be calculated in the same way as a persistent Laplacian.

**Proposition 3.1** (The matrix representation of a persistent sheaf Laplacian). *Choose a basis  $\{v_1, \dots, v_n\}$  of  $\Theta_{X,Y}^{q+1}$  and denote its inner product matrix by  $P$ . Denote the matrix representation of  $(d_{X,Y}^q)^*$  with respect to the canonical basis of  $C^q(X; \mathcal{F})$  and  $\{v_1, \dots, v_n\}$  as  $D_{X,Y}^{*,q}$  and the matrix representation of  $d_X^{q-1}$  with respect to the canonical bases of  $C^{q-1}(X; \mathcal{F})$  and  $C^q(X; \mathcal{F})$  as  $D_X^{q-1}$ . Then, the matrix representation of  $\Delta_q^{X,Y}$  is*

$$D_X^{q-1} (D_X^{q-1})^T + D_{X,Y}^{*,q} P^{-1} (D_{X,Y}^{*,q})^T.$$

*Proof.* We only have to determine the matrix representation  $X$  of  $d_{X,Y}^q$ . Take two vectors  $v \in \Theta_{X,Y}^{q+1}, w \in C^q(X; \mathcal{F})$ . We abuse the notation a bit and use  $v, w$  to denote their coordinates in the form of column vector as well. We have

$$\langle (d_{X,Y}^q)^* v, w \rangle_{C^q(X; \mathcal{F})} = \langle v, d_{X,Y}^q w \rangle_{\Theta_{X,Y}^{q+1}} \quad (1)$$

$$(D_{X,Y}^{*,q} v)^T w = v^T P X w \quad (2)$$

$$v^T (D_{X,Y}^{*,q})^T w = v^T P X w. \quad (3)$$

As  $v, w$  are arbitrarily taken, we conclude that  $X = P^{-1} (D_{X,Y}^{*,q})^T$ .  $\square$

**Proposition 3.2.** *The spectrum of the  $q$ -th persistent sheaf Laplacian does not depend on the orientation of  $Y$ .*

*Proof.* Fixing a choice of orientation for each simplex of  $Y$ , it suffices to show that the spectrum of  $\Delta_q^{X,Y}$  is unchanged if the orientation of one simplex of  $Y$  is altered. We first fix some notations. Suppose we change the orientation of a simplex  $\sigma$ , then every morphism defined with respect to this new orientation will have a bar. We also sometimes denote  $\Theta_{X,Y}^q$  by  $\Theta$ . We define a linear map  $I_{\sigma,-}$  such that  $I_{\sigma,-}|_{\mathcal{G}(\sigma)} = -I$  and  $I_{\sigma,-}|_{\mathcal{G}(\sigma')} = I$  if  $\sigma' \neq \sigma$ . The adjoint of  $I_{\sigma,-}$  is itself. Depending on context, the domain of  $I_{\sigma,-}$  will be understood as  $C^{q-1}(X; \mathcal{F})$ ,  $C^q(Y; \mathcal{G})$  or  $C^{q+1}(Y; \mathcal{G})$ . The proof is divided into cases.

**Case I.** If  $\sigma \in C^{q-1}(X; \mathcal{F})$ , then  $\overline{d_Y^q} = d_Y^q$  and  $\overline{d_X^{q-1}} = d_X^{q-1} I_{\sigma,-}$ . So

$$\overline{d_X^{q-1}}(d_X^{q-1})^* = d_X^{q-1} I_{\sigma,-} I_{\sigma,-} (d_X^{q-1})^* = d_X^{q-1} (d_X^{q-1})^*$$

**Case II.** If  $\sigma \in C^q(Y; \mathcal{G})$ , then  $\overline{d_X^{q-1}} = \pi I_{\sigma,-}|_{C^q(X; \mathcal{F})} d_X^{q-1}$  and  $\overline{d_Y^q} = d_Y^q I_{\sigma,-}$ . So  $(\overline{d_Y^q})^* = I_{\sigma,-}(d_Y^q)^*$ . As

$$(\overline{d_Y^q})^* \Theta_{X,Y}^q = I_{\sigma,-}(d_Y^q)^* \Theta_{X,Y}^q \subset I_{\sigma,-} C^q(X; \mathcal{F}) = C^q(X; \mathcal{F}),$$

we see that  $\Theta_{X,Y}^q \subset \overline{\Theta_{X,Y}^q}$ . Similarly

$$(d_Y^q)^* \overline{\Theta_{X,Y}^q} = I_{\sigma,-} I_{\sigma,-} (d_Y^q)^* \overline{\Theta_{X,Y}^q} = I_{\sigma,-} (\overline{d_Y^q})^* \overline{\Theta_{X,Y}^q} \subset I_{\sigma,-} C^q(X; \mathcal{F}) = C^q(X; \mathcal{F}),$$

we see that  $\Theta_{X,Y}^q \supset \overline{\Theta_{X,Y}^q}$ , so  $\Theta_{X,Y}^q = \overline{\Theta_{X,Y}^q}$ .

Then

$$(\pi \overline{d_Y^q})^*|_{\Theta_{X,Y}^q} = \pi I_{\sigma,-}|_{C^q(X; \mathcal{F})} (d_Y^q)^*|_{\Theta_{X,Y}^q},$$

and

$$((\pi \overline{d_Y^q})^*|_{\Theta_{X,Y}^q})^* = ((d_Y^q)^*|_{\Theta_{X,Y}^q})^* \pi I_{\sigma,-}|_{C^q(X; \mathcal{F})}.$$

So

$$\begin{aligned} \overline{\Delta_q^{X,Y}} &= \overline{d_X^{q-1}}(d_X^{q-1})^* + (\overline{d_Y^q})^*|_{\Theta_{X,Y}^q} ((\overline{d_Y^q})^*|_{\Theta_{X,Y}^q})^* \\ &= \pi I_{\sigma,-}|_{C^q(X; \mathcal{F})} \Delta_q^{X,Y} \pi I_{\sigma,-}|_{C^q(X; \mathcal{F})}. \end{aligned}$$

**Case III.** If  $\sigma \in C^{q+1}(Y; \mathcal{G})$ , then  $\overline{d_X^{q-1}} = d_X^{q-1}$  and  $\overline{d_Y^q} = I_{\sigma,-} d_Y^q$ . So  $(\overline{d_Y^q})^* = (d_Y^q)^* I_{\sigma,-}$ . As  $C^q(X; \mathcal{F}) \supset (\overline{d_Y^q})^* \overline{\Theta} = (d_Y^q)^* I_{\sigma,-} \overline{\Theta}$ , we see that  $I_{\sigma,-} \overline{\Theta} \subset \Theta$ . Similarly  $C^q(X; \mathcal{F}) \supset (d_Y^q)^* \Theta = (\overline{d_Y^q})^* I_{\sigma,-} \Theta$ , implying that  $I_{\sigma,-} \Theta \subset \overline{\Theta}$ . So  $I_{\sigma,-} \Theta = \overline{\Theta}$ . Denote by  $I_{\sigma,-}^{\overline{\Theta}} : \overline{\Theta} \rightarrow \Theta$  the restriction of  $I_{\sigma,-}$  on  $\overline{\Theta}$ . We have  $\pi(\overline{d_Y^q})^*|_{\overline{\Theta}} = \pi(d_Y^q)^* I_{\sigma,-}|_{\overline{\Theta}} = \pi(d_Y^q)^*|_{\Theta} I_{\sigma,-}^{\overline{\Theta}}|_{\Theta}$ . Then

$$\begin{aligned} \overline{\Delta_q^{X,Y}} &= \overline{d_X^{q-1}}(d_X^{q-1})^* + \pi(\overline{d_Y^q})^*|_{\overline{\Theta}} (\pi(\overline{d_Y^q})^*|_{\overline{\Theta}})^* \\ &= d_X^{q-1}(d_X^{q-1})^* + \pi(d_Y^q)^*|_{\Theta} I_{\sigma,-}^{\overline{\Theta}} (I_{\sigma,-}^{\overline{\Theta}})^* (\pi(d_Y^q)^*|_{\Theta})^* \\ &= d_X^{q-1}(d_X^{q-1})^* + \pi(d_Y^q)^*|_{\Theta} (\pi(d_Y^q)^*|_{\Theta})^* \\ &= \Delta_q^{X,Y}. \end{aligned}$$

□

**Example 3.1.** Consider the 1-dimensional simplicial complex  $Y$



and the constant sheaf  $\underline{\mathbb{R}}$  over  $Y$ . We compute  $\Delta_0^{X,Y}$  when  $X = \{1, 2\}$ . The matrix representation of  $(d_Y^0)^*$  is

$$\begin{array}{ccc} 13 & 34 & 42 \\ 1 & \begin{pmatrix} -1 & 0 & 0 \\ 0 & 0 & 1 \\ 1 & -1 & 0 \\ 0 & 1 & -1 \end{pmatrix} \\ 2 & & \\ 3 & & \\ 4 & & \end{array}.$$

According to the definition of  $\Theta_{X,Y}^0$ , we want to find all elements of  $C^1(Y; \underline{\mathbb{R}})$  that are sent to  $C^0(X; \underline{\mathbb{R}}) = \text{span}\{1, 2\}$  by  $(d_Y^0)^*$ . After a few steps of column Gauss elimination, we get a new matrix representation of  $(d_Y^0)^*$

$$\begin{array}{ccc} 13 & 13 + 34 & 13 + 34 + 42 \\ 1 & \begin{pmatrix} -1 & -1 & -1 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \\ 2 & & \\ 3 & & \\ 4 & & \end{array}.$$

From this representation, we see that for any vector  $v = a13 + b(13 + 34) + c(13 + 34 + 42)$  with  $a, b$ , and  $c$  being coefficients,  $(d_Y^0)^*v \in \text{span}\{1, 2\}$  if and only if  $a$  and  $b$  are both zero. In other words,  $\Theta_{X,Y}^0 = \text{span}\{13 + 34 + 42\}$ ,  $P = 3$ , and the matrix representation of  $(d_{X,Y}^0)^*$  is

$$\begin{array}{c} 13 + 34 + 42 \\ 1 \begin{pmatrix} -1 \\ 1 \end{pmatrix} \\ 2 \end{array}.$$

Then, the matrix representation of  $\Delta_0^{X,Y}$  is

$$\begin{pmatrix} 1/3 & -1/3 \\ -1/3 & 1/3 \end{pmatrix}$$

and its spectrum is  $\{0, 2/3\}$ .

**4. Experiments.** Given a labeled point cloud  $P$  (i.e., a point cloud with a nonzero quantity  $q_i$  associated with each point  $v_i$ ), we can build a Rips or Alpha filtration out of it and construct a sheaf  $\mathcal{S}_t$  on each  $X_t$  consistently as described in section 2.3 provided a suitable global  $F : 2^P \rightarrow \mathbb{R}$  is chosen. If we take  $F_t$  to be the restriction of  $F : 2^P \rightarrow \mathbb{R}$  on  $X_t$  and construct  $\mathcal{S}_t$ , then we ensure that the stalks of  $\sigma$  and  $\tau$  and the restriction map between  $\sigma$  and  $\tau$  remain the same for any  $\mathcal{S}_t$  containing  $\sigma \leq \tau$ . Since  $\mathcal{S}_t$  is the pullback of  $\mathcal{S}_{t+p}$  for any  $t$  and  $p$ , we get a persistent module of sheaf cochain complexes and can compute the spectra of persistent sheaf Laplacians. In this way, we calculate the spectra of persistent sheaf Laplacians for a few examples of point clouds in this section. Some examples are the vertices of simple geometrical shapes and some are the atomic coordinates of molecules. We assign the quantities  $q_i$  to simple geometrical shapes, and take the partial charges as  $q_i$  for molecules. An Alpha complex filtration  $\{X_r\}$  is built for each labeled point cloud, parametrized by radius  $r$ . We choose  $F$  such that  $F$  maps every vertex  $v_i$  to 1, every edge  $v_i v_j$  to the length of itself  $r_{ij}$ , and every 2-cell  $v_i v_j v_k$  to the product of lengths of its edges  $r_{ij} r_{ik} r_{jk}$ . The spectra of  $\Delta_d^{X_r, X_{r+p}}$  for  $d = 0, 1$  and selected  $r, p$

will be calculated. The radius  $r$  will be a multiple of 0.01 or 0.01 Å. The spectrum can provide information for various analyses, but for simplicity here we only plot the minimal nonzero eigenvalue and the nullity against the radius  $r$ . The minimal nonzero eigenvalue of the persistent sheaf Laplacian  $\Delta_d^{X_r, X_{r+p}}$  is denoted by  $\lambda_d^{r,p}$  and the  $d$ -th persistent sheaf Betti number of the pair  $(X_r, X_{r+p})$  (i.e., the nullity of  $\Delta_d^{X_r, X_{r+p}}$ ) is denoted by  $\beta_d^{r,p}$ .

**4.1. Simple shapes.** The examples of simple shapes are the vertex sets of a 2-dimensional square or a 2-dimensional trapezoid as shown in Figure 2 with different choices of local property  $q_i$ . More specifically, we consider four labeled point clouds and two point clouds in  $\mathbb{R}^2$ . The two point clouds are  $\{(0, 0), (1, 0), (1, 1), (0, 1)\}$  and  $\{(0, 0), (1, 0), (3/4, \sqrt{15}/4), (1/4, \sqrt{15}/4)\}$ . For  $\{(0, 0), (1, 0), (1, 1), (0, 1)\}$  we assign  $q = \pm 1$  to  $(0, 1)$  and  $q = 1$  to the rest. For  $\{(0, 0), (1, 0), (3/4, \sqrt{15}/4), (1/4, \sqrt{15}/4)\}$  we assign  $q = \pm 1$  to  $(1/4, \sqrt{15}/4)$  and  $q = 1$  to the rest. For the two point clouds we construct the constant sheaf and compute persistent sheaf Laplacians, whose spectra coincide with persistent Laplacians. For the four labeled point clouds we construct filtrations of sheaves in the way described earlier in this section. The results are shown in Figures 3 and 4. The first thing we can infer is that, the minimal nonzero eigenvalue and the nullity usually change significantly when the topology of  $X_r$  changes (when  $p$  is nonzero, both of them change at  $r - p$  rather than  $r$ ). When we consider labeled point clouds instead of point clouds, sometimes the nullity show less change but the minimal nonzero eigenvalue gives more change (compare (a) and (b) with (c) and (d) in Figure 3 and Figure 4).

If we compare the results of the square and the trapezoid, we see more changes in the results of the trapezoid. This is due to the fact that the filtration constructed out of the vertex set of trapezoid contains more different complexes. We also observe that the change of signs of  $q_i$  does not affect the minimal nonzero eigenvalue and the nullity, though the eigenvectors of Laplacians will be different.



FIGURE 2. A square and a trapezoid. Coordinates of vertices are shown.

**4.2. Complex molecules.** Next we study the molecule CB8 [34] shown in Figure 5. We associate each atom with the corresponding partial charge (obtained using [30]). The results for CB8 are shown in Figures 6 and 7. Due to complexity of the molecule, it is very difficult to explain the spectral details of the system. However, this information can be very useful for machine learning analysis.

Finally, to demonstrate our method for practical problems, we study a small protein called bacteriocin AS-48 (PDB ID: 1E68) [14]. We select the model 1 of AS-48 and compute the pqr file by PDB2PQR with the Amber force field [18]. For the sake of faster computation, we only use the coordinates of carbon atoms as the point cloud. Results are shown in Figures 9 and 10.

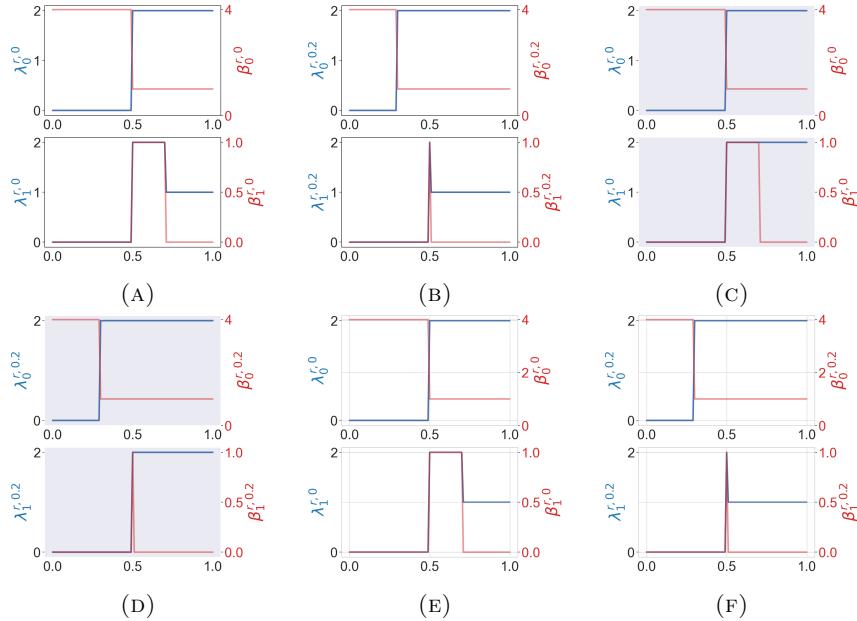


FIGURE 3. The results of persistent sheaf Laplacians for the square. We consider pairs  $(X_r, X_r)$  or  $(X_r, X_{r+0.2})$  in a filtration. The results of the labeled point cloud  $\{((0,0), 1), ((1,0), 1), ((1,1), 1), ((0,1), 1)\}$  are shown in (a) and (b). The results of the point cloud  $\{(0,0), (1,0), (1,1), (0,1)\}$  are shown in (c) and (d). The results of the labeled point cloud  $\{((0,0), 1), ((1,0), 1), ((1,1), 1), ((0,1), -1)\}$  are shown in (e) and (f).

**5. Concluding remarks.** The success of persistent homology in data science lies in its integration of multiscale analysis and topological abstraction in terms of topological invariants [12, 45]. However, this approach has a significant drawback, i.e., its inability to detect homotopic shape evolution of data that involves no topological changes. This limitation was addressed by persistent topological Laplacians, including persistent spectral graph, also called persistent Laplacians [38], and evolutionary de Rahm Hodge theory, also called persistent Hodge Laplacians [8]. The multiscale geometric analysis in these new approaches has dramatically empowered classical spectral theory. The harmonic spectra of persistent Laplacians fully retain the topological invariants of persistent homology, while their non-harmonic spectra capture the homotopic shape evolution of data that cannot be obtained from persistent homology. Nonetheless, persistent Laplacians cannot handle non-geometric information of the data, such as the atomic property of a molecule and the node property of a network. This work proposes persistent sheaf Laplacians (PSL) to overcome this limitation. The theory of PSLs introduces multiscale analysis to cellular sheaf Laplacians to further extend their application potential via a filtration. The proposed PSLs have been demonstrated on various point cloud data with localized labels.

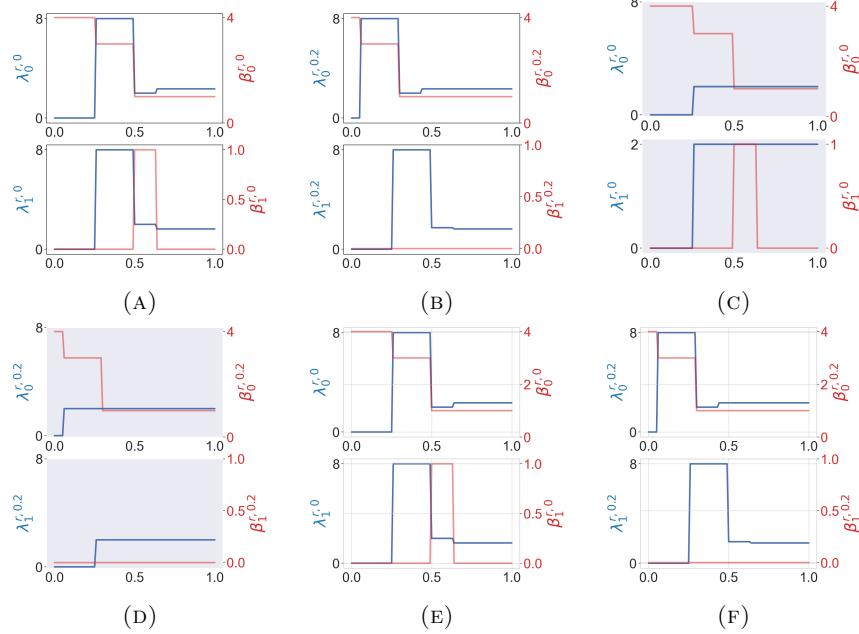


FIGURE 4. The results of persistent sheaf Laplacians for the trapezoid. We consider pairs  $(X_r, X_r)$  or  $(X_r, X_{r+0.2})$  in a filtration. The results of the labeled point cloud  $\{((0, 0), 1), ((1, 0), 1), ((3/4, \sqrt{15})/4, 1), ((1/4, \sqrt{15}/4), 1)\}$  are shown in (a) and (b). The results of the point cloud  $\{((0, 0), (1, 0), (3/4, \sqrt{15}/4), (1/4, \sqrt{15}/4))\}$  are shown in (c) and (d). The results of the labeled point cloud  $\{((0, 0), 1), ((1, 0), 1), ((3/4, \sqrt{15}/4), 1), ((1/4, \sqrt{15}/4), -1)\}$  are shown in (e) and (f).

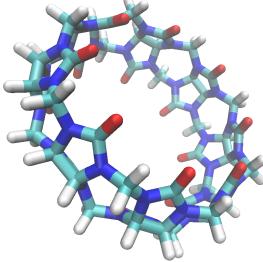


FIGURE 5. The structure of CB8.

Potential future development is widely open. First, a fast implementation of PSLs and a good understanding of the eigenvalues and eigenspaces of PSLs will dramatically boost the applications of PSLs to complex and big data. Second, the incorporation of foundational laws of physics, principles of chemistry, and rules of life into PSLs to study various types of data will be exciting. Third, the application

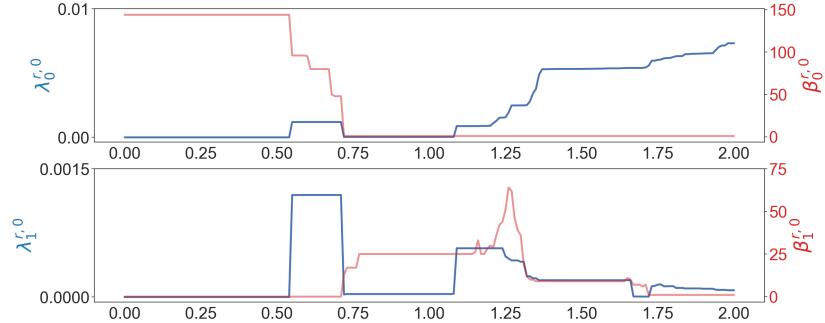
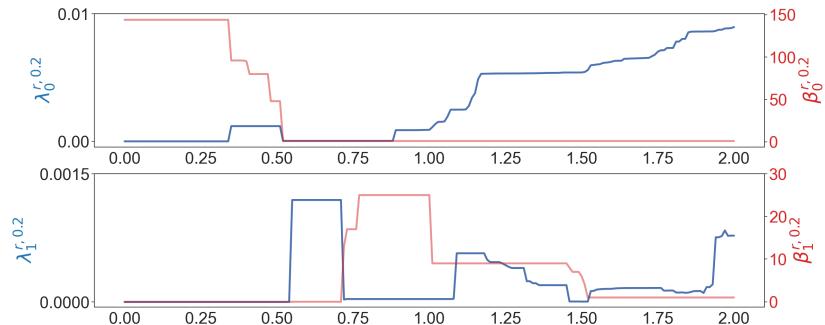
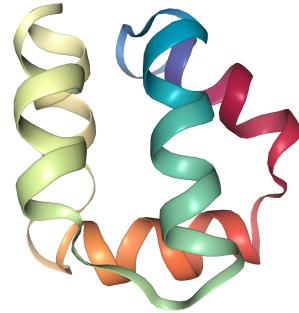
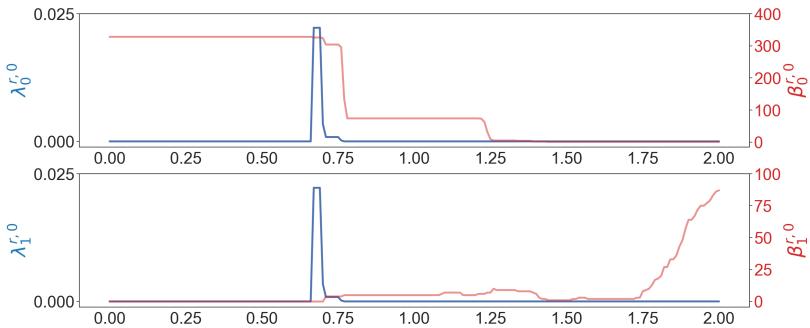
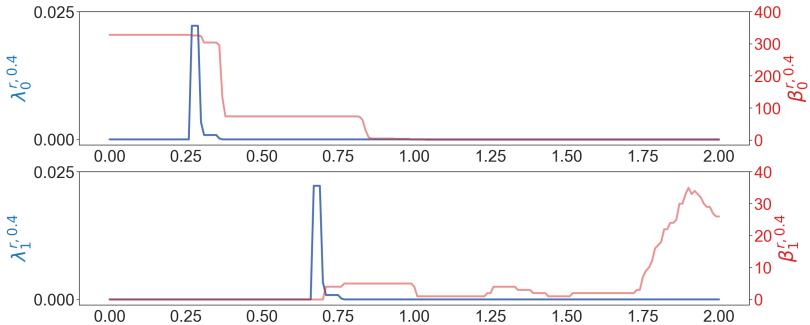
FIGURE 6. The results of the CB8 when  $p = 0$ .FIGURE 7. The results of the CB8 when  $p = 0.2$ .

FIGURE 8. Illustration of the structure of bacteriocin AS-48.

of PSLs for data fusion, i.e., integrating multiple data sources to produce more consistent, accurate, and useful information than that concatenates individual data sources, is another direction. Fourth, we need to explore various specific types

FIGURE 9. The results of AS-48 when  $p = 0$ .FIGURE 10. Results of AS-48 when  $p = 0.4$ .

of sheaf Laplacians for each specific problem and test their power in applications. Fifth, a persistent sheaf Dirac theory will extend the Dirac formulation of persistent homology [1, 36, 41] and provide a potentially more powerful approach for data science. A Dirac operator can be roughly seen as the square root of combinatorial Laplacians, and recently the extension of Dirac operators to the setting of persistent homology has been proposed since it offers a new perspective in topological data analysis. Wee et al. [41] demonstrated that persistent Dirac model discriminates the 9 types of organic-inorganic halide perovskites clearly. Finally, in the spirit of evolutionary de Rahm-Hodge theory [8], one can develop evolutionary sheaf Dirac on manifolds for volumetric data. However, like persistent Hodge Laplacians, this approach can be computationally demanding to implement.

**Code availability.** The code for calculating persistent sheaf Laplacians can be found at [https://github.com/weixiaoqimath/persistent\\_sheaf\\_Laplacians](https://github.com/weixiaoqimath/persistent_sheaf_Laplacians).

## REFERENCES

- [1] B. Ameneviro, V. Maroulas and G. Siopsis, [Quantum persistent homology](#), *Journal of Applied and Computational Topology*, (2024), 1-20.
- [2] N. Berkouk and G. Ginot, [A derived isometry theorem for constructible sheaves on  \$\mathbb{R}\$](#) , 2021.

- [3] P. Bubenik and N. Milićević, **Homological algebra for persistence modules**, *Foundations of Computational Mathematics*, **21** (2021), 1233-1278.
- [4] Z. Cang and G.-W. Wei, **Topologynet: Topology based deep convolutional and multi-task neural networks for biomolecular property predictions**, *PLoS Computational Biology*, **13** (2017), e1005690.
- [5] Z. Cang and G.-W. Wei, **Persistent cohomology for data with multicomponent heterogeneous information**, *SIAM Journal on Mathematics of Data Science*, **2** (2020), 396-418.
- [6] A. Cerri, B. Di Fabio, M. Ferri, P. Frosini and C. Landi, **Betti numbers in multidimensional persistent homology are stable functions**, *Mathematical Methods in the Applied Sciences*, **36** (2013), 1543-1557.
- [7] J. Chen, Y. Qiu, R. Wang and G.-W. Wei, **Persistent Laplacian projected Omicron BA. 4 and BA. 5 to become new dominating variants**, *Computers in Biology and Medicine*, **151** (2022), 106262.
- [8] J. Chen, R. Zhao, Y. Tong and G.-W. Wei, **Evolutionary de Rham-Hodge method**, *Discrete Contin. Dyn. Syst. Ser. B*, **26** (2021), 3785-3821.
- [9] F. R. Chung, *Spectral Graph Theory*, volume 92, American Mathematical Soc., 1997.
- [10] J. Curry, *Sheaves, Cosheaves and Applications*, PhD thesis, University of Pennsylvania, 2014.
- [11] J. Dodziuk, **de Rham-Hodge theory for  $l^2$ -cohomology of infinite coverings**, *Topology*, **16** (1977), 157-165.
- [12] H. Edelsbrunner and J. Harer, **Persistent homology-a survey**, *Contemporary Mathematics*, **453** (2008), 257-282.
- [13] R. Ghrist, *Elementary Applied Topology*, Createspace, 1 edition, 2014.
- [14] C. González, G. M. Langdon, M. Bruix, A. Gálvez, E. Valdivia, M. Maqueda and M. Rico, **Bacteriocin AS-48, a microbial cyclic polypeptide structurally and functionally related to mammalian NK-lysin**, *Proceedings of the National Academy of Sciences*, **97** (2000), 11221-11226.
- [15] R. Grone, R. Merris and V. Sunder, **The Laplacian spectrum of a graph**, *SIAM Journal on Matrix Analysis and Applications*, **11** (1990), 218-238.
- [16] H. Hang and W. Mio, **Correspondence modules and persistence sheaves: A unifying perspective on one-parameter persistent homology**, *Jpn. J. Ind. Appl. Math.*, **40** (2023), 41-93.
- [17] J. Hansen and R. Ghrist, **Toward a spectral theory of cellular sheaves**, *Journal of Applied and Computational Topology*, **3** (2019), 315-358.
- [18] E. Jurrus, D. Engel, K. Star, K. Monson, J. Brandi, L. E. Felberg, D. H. Brookes, L. Wilson, J. Chen, K. Liles, M. Chun, P. Li, D. W. Gohara, T. Dolinsky, R. Konecny, D. R. Koes, J. E. Nielsen, T. Head-Gordon, W. Geng, R. Krasny, G.-W. Wei, M. J. Holst, J. A. McCammon and N. A. Baker, **Improvements to the APBS biomolecular solvation software suite**, *Protein Science: A Publication of the Protein Society*, **27** (2018), 112-128. 28836357[pmid].
- [19] T. Kaczynski, K. Mischaikow and M. Mrozek, *Computational Homology*, Appl. Math. Sci., 157, Springer-Verlag, New York, 2004.
- [20] L.-H. Lim, **Hodge Laplacians on graphs**, *Siam Review*, **62** (2020), 685-715.
- [21] J. Liu, J. Li and J. Wu, **The algebraic stability for persistent Laplacians**, arXiv preprint, 2023. [arXiv:2302.03902](https://arxiv.org/abs/2302.03902).
- [22] F. Mémoli, Z. Wan and Y. Wang, **Persistent Laplacians: Properties, algorithms and implications**, *SIAM Journal on Mathematics of Data Science*, **4** (2022), 858-884.
- [23] Z. Meng, D. V. Anand, Y. Lu, J. Wu and K. Xia, **Weighted persistent homology for biomolecular data analysis**, *Scientific Reports*, **10** (2020), 2079.
- [24] Z. Meng and K. Xia, **Persistent spectral-based machine learning (PerSpect ML) for protein-ligand binding affinity prediction**, *Science Advances*, **7** (2021), eabc5329.
- [25] J. R. Munkres, *Elements of Algebraic Topology*, Addison-Wesley Publishing Company, Inc., 1984.
- [26] D. D. Nguyen, Z. Cang and G.-W. Wei, **A review of mathematical representations of biomolecular data**, *Physical Chemistry Chemical Physics*, **22** (2020), 4343-4367.
- [27] D. D. Nguyen, Z. Cang, K. Wu, M. Wang, Y. Cao and G.-W. Wei, **Mathematical deep learning for pose and binding affinity prediction and ranking in D3R grand challenges**, *Journal of Computer-Aided Molecular Design*, **33** (2019), 71-82.
- [28] D. D. Nguyen, K. Gao, M. Wang and G.-W. Wei, **MathDL: Mathematical deep learning for D3R grand challenge 4**, *Journal of Computer-Aided Molecular Design*, **34** (2020), 131-147.
- [29] Y. Qiu and G.-W. Wei, **Persistent spectral theory-guided protein engineering**, *Nature Computational Science*, **3** (2023), 149-163.

- [30] T. Raček, O. Schindler, D. Toušek, V. Horský, K. Berka, J. Koča and R. Svobodová, [Atomic charge calculator II: Web-based tool for the calculation of partial atomic charges](#), *Nucleic Acids Research*, **48** (2020), W591-W596.
- [31] M. Robinson, How do we deal with noisy data?
- [32] M. Robinson, *Topological Signal Processing*, Springer, Berlin, Heidelberg, 1 edition, 2014.
- [33] F. Russold, Persistent sheaf cohomology, arXiv preprint, 2022. [arXiv:2204.13446](https://arxiv.org/abs/2204.13446).
- [34] SampLchallenges, the mol2 file of CB8, [https://github.com/sampLchallenges/SAMPL6/blob/master/host\\_guest/CB8AndGuests/CB8.mol2](https://github.com/sampLchallenges/SAMPL6/blob/master/host_guest/CB8AndGuests/CB8.mol2).
- [35] A. D. Shepard, *A Cellular Description of the Derived Category of a Stratified Space*, PhD Thesis, Brown University, 1985.
- [36] F. Suwayyid and G.-W. Wei, [Persistent Dirac of path and hypergraph](#), *Foundations of Data Science*, **6** (2024), 124-153.
- [37] J. Townsend, C. P. Micucci, J. H. Hymel, V. Maroulas and K. D. Vogiatzis, [Representation of molecular structures with persistent homology for machine learning applications in chemistry](#), *Nature Communications*, **11** (2020), 3230.
- [38] R. Wang, D. D. Nguyen and G.-W. Wei, [Persistent spectral graph](#), *International Journal for Numerical Methods in Biomedical Engineering*, **36** (2020), e3376.
- [39] R. Wang, R. Zhao, E. Ribando-Gros, J. Chen, Y. Tong and G.-W. Wei, [HERMES: Persistent spectral graph software](#), *Foundations of Data Science*, **3** (2021), 67-97.
- [40] L. Wasserman, [Topological data analysis](#), *Annual Review of Statistics and its Application*, **5** (2018), 501-532.
- [41] J. Wee, G. Bianconi and K. Xia, [Persistent Dirac for molecular representation](#), *Scientific Reports*, **13** (2023), Article number: 11183.
- [42] C. Wu, S. Ren, J. Wu and K. Xia, Weighted (co)homology and weighted Laplacian, arXiv preprint, 2018. [arXiv:1804.06990](https://arxiv.org/abs/1804.06990).
- [43] K. Yegnesh, Persistence and sheaves, arXiv preprint, 2016. [arXiv:1612.03522](https://arxiv.org/abs/1612.03522).
- [44] R. Zhao, M. Wang, J. Chen, Y. Tong and G.-W. Wei, [The de Rham–Hodge analysis and modeling of biomolecules](#), *Bulletin of Mathematical Biology*, **82** (2020), Paper No. 108, 38 pp.
- [45] A. Zomorodian and G. Carlsson, [Computing persistent homology](#), *Discrete & Computational Geometry*, **33** (2005), 249-274.

Received November 2023; revised March 2024; early access July 2024.

Copyright of Science, Education & Innovations in the Context of Modern Problems is the property of International Meetings & Conferences Research Association (IMCRA) and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.