SIGNAL PROCESSING ON PRODUCT SPACES

T. Mitchell Roddenberry[*], *Vincent P. Grande*[†], *Florian Frantzen*[†], *Michael T. Schaub*[†], *Santiago Segarra*[*]

[*]Department of Electrical and Computer Engineering, Rice University, USA
[†]Department of Computer Science, RWTH Aachen University, Germany

ABSTRACT

We establish a framework for signal processing on product spaces of simplicial and cellular complexes. For simplicity, we focus on the product of two complexes representing time and space, although our results generalize naturally to products of simplicial complexes of arbitrary dimension. Our framework leverages the structure of the eigenmodes of the Hodge Laplacian of the product space to jointly filter along time and space. To this end, we provide a decomposition theorem of the Hodge Laplacian of the product space, which highlights how the product structure induces a decomposition of each eigenmode into a spatial and temporal component. Finally, we apply our method to real world data, specifically for interpolating trajectories of buoys in the ocean from a limited set of observed trajectories.

1. INTRODUCTION

In recent years, graph signal processing (GSP) has proven itself to be an important tool for analyzing, filtering, denoising, and interpolating signals defined on abstract domains. Extending the tools of ordinary signal processing to graphs opened the door for a multitude of more combinatorially inspired applications [1, 2, 3, 4]. Recently, this shift has been extended further to signal processing on topological spaces using simplicial complexes (SCs) [5, 6, 7, 8, 9], cellular complexes (CCs) [10, 11, 12], or cellular sheaves [13]. This move to topological signal processing allows both for a more general geometry of the underlying domain to be captured, as well as for additional constraints on the filtering to be taken into account. Example applications where this topological viewpoint has been used include wireless traffic networks [7] or outlier detection on whales traveling in the Canadian Arctic Archipelago [14].

Sometimes the data considered varies not only in spatial dimensions, but also has a temporal dimension. An effective filter then should consider both of these dimensions. However, filtering on the simplicial complex will only take the spatial dimension into account, whereas filtering individually along each edge in time disregards the underlying topology and geometry of the problem. To solve this issue we propose to construct a new complex that captures both the spatial and temporal dimensions of the problem.

The idea of signal processing for product spaces has already been considered for product graphs for data defined on the nodes [15, 16, 17, 18]. However, processing of edge flow data is not possible within these product graph frameworks: while the product of two 0-dimensional vertices is again a $(0 + 0)$-dimensional vertex

and can be considered via standard graph products, the product of two (1-dimensional) edges is not an edge again but a 2-dimensional face. Hence, more sophisticated techniques and models are needed to model and process such data.

1.1. Contributions and Outline

We present a novel framework for signal processing on product spaces of simplicial complexes, which enables us to filter time-varying signals on SCs. By interpreting the time axis as an SC (in particular, as a graph), we construct a cellular complex that represents the spatially distributed, time-varying signal. It turns out that the Hodge Laplacian for the resulting product complex is a concise way to express a filter with the desired properties.

Section 2 gives an overview of the mathematical concepts of simplicial and cellular complexes, the Hodge Laplacian, and product complexes. Section 3 elaborates on the motivation and the theory for temporal flow interpolation. Specifically, Proposition 1 describes the Hodge Laplacian of a product space. Finally, we apply our method to real world ocean currents and show that interpolating along both time and space using the Hodge Laplacian of the product space outperforms interpolation solely along a single dimension.

2. BACKGROUND

In this section, we present an elementary overview of concepts used to process signals defined on higher order networks such as simplicial or cell complexes. For more details, [19, 20] give a background in algebraic topology and [10, 5] in topological signal processing.

Notation. For some $n \in \mathbb{N}$, let $[n]$ denote the set of integers $\{1, \ldots, n\}$. Matrices and vectors are denoted by letters $\mathbf{A}, \mathbf{x}$. We use $\cong$ to denote two isomorphic vector spaces, and $\oplus$ to denote the direct sum of vector spaces. Further, $\otimes$ denotes the tensor product.

Simplicial Complexes. A simplicial complex [19, 20] (SC) X consists of a finite set of points $\mathcal{V}$, and a set of nonempty subsets of $\mathcal{V}$ that is closed under taking nontrivial subsets. A *k-simplex* $\mathcal{S}^k$ is a subset of $\mathcal{V}$ with $k + 1$ points. Further, the *faces* of $\mathcal{S}^k$ are the subsets of $\mathcal{S}^k$ with cardinality k, that is, exactly one element of the simplex is omitted. If $\mathcal{S}^{k-1}$ is a face of $\mathcal{S}^k$, $\mathcal{S}^k$ is called a *coface* of $\mathcal{S}^{k-1}$. We denote the set of k-simplices and their number in X by X_k and N_k, respectively. We can ground the understanding of SCs as an extension of the well known concept of graphs: X_0 is the set of vertices, X_1 is the set of edges, X_2 is the set of filled in triangles (not all 3-cliques must be filled in triangles), and so on. The orientation of a k-simplex $\mathcal{S}^k$ is represented by an ordering of its $k + 1$ elements of $\mathcal{V}$. Two orientations are considered the same if they differ by an even permutation. Thus every simplex admits two possible orientations.

The structure of an SC can be encoded by *boundary operators* $\mathbf{B}_k$, which record the incidence relations between $(k-1)$-simplices and k-simplices according to the chosen reference orientations [19]. Rows of $\mathbf{B}_k$ are indexed by $(k-1)$-simplices and columns of $\mathbf{B}_k$ are indexed by k-simplices. Thus, the matrix $\mathbf{B}_1$ is the vertex-to-edge incidence matrix and $\mathbf{B}_2$ is the edge-to-triangle incidence matrix.

Simplicial Signals. A k-simplicial signal on top of an SC assigns a real number to each oriented k-simplex. We can think of this as a vector in $\mathbb{R}^{N_k}$. For example, $\mathbb{R}^{N_0}$ is the *node signal* space and $\mathbb{R}^{N_1}$ is the space of *edge flows*. We denote the k-simplicial signal space on the SC X by $\mathcal{C}^k(X)$. Denote by $\mathcal{C}^\bullet(X)$ the direct sum $\bigoplus_{k=0}^\infty \mathcal{C}^k(X)$ of the k-simplicial signal spaces. The sign of the signal indicates whether the signal is aligned with the reference orientation or not. More abstractly, a k-simplicial signal describes a skew-symmetric function on the set of oriented k-simplices. The k-th boundary operator $\mathbf{B}_k$ can be interpreted as a map from $\mathbb{R}^{N_k}$ to $\mathbb{R}^{N_{k-1}}$. Furthermore, denote by $\partial_X \colon \mathcal{C}^\bullet(X) \to \mathcal{C}^\bullet(X)$ the direct sum of the maps $\mathbf{B}_k$. Because of the way we defined $\mathbf{B}_k$ and ∂_X, applying ∂_X twice is trivial: $\partial_X \circ \partial_X = 0$.

Hodge Laplacian and Hodge Decomposition. Using the sequence of boundary operators $\mathbf{B}_k$, we can define the k-th *Hodge Laplacian* on the k-simplicial signal space $\mathbb{R}^{N_k}$ as $\mathbf{L}_k = \mathbf{B}_k^\top \mathbf{B}_k + \mathbf{B}_{k+1}\mathbf{B}_{k+1}^\top$ [21, 22]. The standard graph Laplacian corresponds to $\mathbf{L}_0 = \mathbf{B}_1 \mathbf{B}_1^\top$ ($\mathbf{B}_0 = 0$ by convention). The Hodge Laplacian gives rise to the *Hodge decomposition* [22, 7, 5, 23]:

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

We denote by $\Delta_X \colon \mathcal{C}^\bullet(X) \to \mathcal{C}^\bullet(X)$ the direct sum of the Hodge Laplacians $\mathbf{L}_k$.

Cell Complexes. Intuitively, *regular cell complexes* (CCs) generalize the notion of SCs [19, 20]. Rather than allowing only for simplices which are fixed-cardinality subsets of the vertex set as building blocks, we allow for the use of cells with more general shape. A cell complex X consists of cells of nonnegative dimension. The n-skeleton X_n of X consists of all cells in X of dimension at most n. Starting with 0-cells, a k-cell c_k is homeomorphic to the unit ball $B^k := \{\mathbf{x} \in \mathbb{R}^k : \|\mathbf{x}\| \le 1\}$. Its boundary $S^{k-1} := \{\mathbf{y} \in \mathbb{R}^k : \|\mathbf{y}\| = 1\}$ is glued to X_{k-1} along an attaching map $\alpha\colon S^{k-1} \to X_{k-1}$. We will call the cells contained in the image of the attaching map α the boundary of c_k. For example, a line is a 1-cell with its two endpoints as 0-dimensional boundary. Similarly, a polygon is a 2-cell and its boundary consists of the line segments defining it. Here, the difference between a simplicial and a cell complex becomes clear: While a 2-cell can be any arbitrary polygon, a 2-simplex must be a triangle.

Analogous to SCs, we equip a cell complex with a reference orientation for bookkeeping purposes. A cell inherits its orientation from its boundary. That is, by orienting the boundary of a cell, we obtain a corresponding orientation of the cell itself.

A CC can again be encoded by boundary operators $\mathbf{B}_k$. The rows of $\mathbf{B}_k$ are indexed by $(k-1)$-cells and columns of $\mathbf{B}_k$ are indexed by k-cells. See [19] for more details. If we forget the topological data of a cell complex and only record the incidence and boundary relation, we arrive at the notion of an *abstract cell complex*. As in the case of SCs, we can associate a k-cellular signal space $\mathcal{C}^k(X)$ to X. We define ∂_X analogously to the simplicial complex case.

3. TEMPORAL FLOW INTERPOLATION

In this section, we demonstrate how to use product spaces of SCs to process flows varying in time and space. To that end, we assume
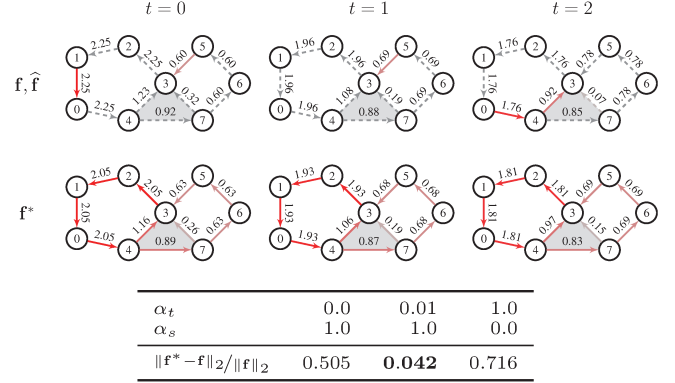

Fig. 1. **Illustration: Importance of considering both spatial and temporal components in sample-limited interpolation.** The first row illustrates the true flow $\mathbf{f}$ as it varies over time ($t \in \{1, 2, 3\}$). The solid edges indicate the observed flows in each snapshot while dashed edges are unobserved and their flow value should be interpolated. We fix $\lambda = 10^{-6}$. The second row shows $\alpha_t/\alpha_s = 0.01$.

that we are given a static simplicial complex and we observe a time-varying edge flow (*i.e.*, at each time an element of $\mathcal{C}^1(X)$) supported on this complex.

3.1. Example: Interpolation in Time and Space

Let $X = (\mathcal{V}, \mathcal{E}, \mathcal{T})$ be an SC consisting of vertices, edges, and filled in triangles, and assume that the set of edges has been endowed with an orientation. Suppose for some time-varying flow $f\colon [T]\times\mathcal{E} \to \mathbb{R}$, we observe flows on a subset of the edges within our spatial domain described by X. Denote the collection of observed edges by $\Omega \subseteq [T] \times \mathcal{E}$. For such a flow, write the vectorized form as $\mathbf{f}$, indexed so that $\mathbf{f}_t^e = f(t, e)$ for some time index $t \in [T]$ and oriented edge $e \in \mathcal{E}$. For convenience, we write the flow of an edge e at all time steps as $\mathbf{f}^e$, and the flow on the SC at a single time t as $\mathbf{f}_t$.

To interpolate such a partially observed flow, we proceed by setting up the following optimization problem akin to [24, 25], which enforces the flows to vary smoothly (jointly) across space and time:

$$\mathbf{f}^* = \arg\min_{\mathbf{f}} \Bigg(\mathrm{MSE}(\mathbf{f}|_\Omega, \widehat{\mathbf{f}}|_\Omega) + \alpha_s \sum_{t=0}^T \mathbf{f}_t^\top \mathbf{L}_s \mathbf{f}_t$$
$$+ \alpha_t \sum_{e\in\mathcal{E}} \mathbf{f}^{e\top} \mathbf{L}_t \mathbf{f}^e + \lambda \mathbf{f}^\top \mathbf{f} \Bigg). \tag{2}$$

Here $\widehat{\mathbf{f}}|_\Omega$ denotes the partially observed flow restricted to the observed domain Ω and MSE denotes the mean squared error. Further, $\mathbf{L}_s$ is the spatial Hodge Laplacian obtained by considering the SC X for each time index t independently. Moreover, the temporal Hodge Laplacian $\mathbf{L}_t$ is here simply the graph Laplacian of the path graph of length T, obtained by comparing each edge flow with the flow on the same edge at $t' = t \pm 1$. In this setting, "smooth" means "having a small quadratic form with respect to the Hodge Laplacian." The scalars α_s and α_t are regularization parameters, allowing us to enforce varying degrees of temporal smoothness relative to spatial smoothness. The ridge penalty λ is assumed to be small and is included only to force a unique solution.

As a concrete example, let us assume we observe only part of the edge flow at $T = 3$ different snapshots and want to infer an approximation of the remaining edge flows for all snapshots, as illustrated in Fig. 1. The true (only partially observed) flow is a smooth flow on the edges of an SC and changes gradually over time ($t \in \{1, 2, 3\}$).
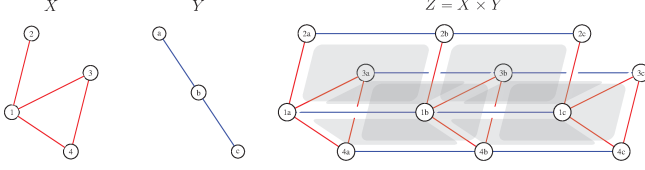
Fig. 2. **Cell complex yielded by the Cartesian product of two simplicial complexes.** For ease of illustration, we consider two graphs X, Y as the component spaces. Taking their Cartesian product, we obtain a cell complex Z. Using our notation, $\mathcal{C}^{0,0}(Z)$ is supported on the vertices of Z, $\mathcal{C}^{1,0}(Z)$ is supported on the red edges of Z, $\mathcal{C}^{0,1}(Z)$ is supported on the blue edges of Z, and $\mathcal{C}^{1,1}(Z)$ is supported on the gray faces of Z.

Our observation set now consists of the flow over 2 edges at $t = 1$, 1 edge at $t = 2$, and 2 edges at $t = 3$.

One can clearly see that many edges of the graph are never observed, rendering purely temporal smoothing unsuitable. Moreover, for any snapshot, the observed edges do not cover a sufficient subset of the graph, so that purely spatial interpolation at each point in time under the assumption of smooth flows is insufficient as well. This justifies the importance of interpolating in a way that requires joint smoothness in both the temporal and spatial domain. Indeed, without such smoothness assumptions, the interpolation problem is impossible, since there are 30 degrees of freedom, but only 5 known variables. To test this, we vary the regularization parameters α_s, α_t to explore the relationship between spatial and temporal smoothness, with results shown in the table of Fig. 1. Observe that the best interpolation performance is attained when both α_s and α_t are nonzero, with poor performance when either one is zero.

3.2. The Product Space Framework

In this subsection we will show how the above filtering procedure emerges naturally from considering the Hodge Laplacian of the product space. This not only provides a theoretical justification for considering the above optimization problem, but provides also the means to extend these ideas to signals supported on arbitrary products of complexes.

Product Complexes. Let X and Y be SCs. Observe that their Cartesian product $Z := X \times Y$ *cannot* easily be given the structure of a SC. Instead, we can view Z as a cell complex: The elements of Z are tuples of the form (σ_X, σ_Y), where $\sigma_X \in X$ and $\sigma_Y \in Y$. If σ_X is a k-simplex, and σ_Y is an ℓ-simplex, then we can view (σ_X, σ_Y) as a $(k + \ell)$-cell in Z. For instance, a tuple consisting of two vertices yields another vertex, and a tuple consisting of a vertex and an edge yields an edge. However, a tuple consisting of two edges does not yield a 2-simplex (triangle), but a 2-cell resembling a "filled in rectangle," as pictured in Fig. 2.

The faces of a k-cell $\sigma_Z = (\sigma_X, \sigma_Y)$ are the faces of the form (σ'_X, σ_Y) or (σ_X, σ'_Y), where σ'_X and σ'_Y are faces of σ_X or σ_Y respectively. One can easily see that the product $e \times f$ of two edges $e = [v_1, v_2]$ and $f = [u_1, u_2]$ has four faces corresponding to the cells (v_1, f), (v_2, f), (e, u_1), and (e, u_2). However, a 2-*simplex* is only allowed to have 3 faces. To maintain a simplicial structure, we would thus need to subdivide this cell into two triangles.

Signal Space of a Product Complex. Recall that there is a bijective correspondence between k-cells of Z and pairs (σ^i_X, σ^j_Y) of i- and j-cells of X and Y with $i + j = k$. Thus, we can decompose the k-cellular signal space of Z into a direct sum of tensor products of the i- and j-simplicial signal spaces of X and Y.

$$\mathcal{C}^k(Z) \cong \bigoplus_{i+j=k} \mathcal{C}^i(X) \otimes \mathcal{C}^j(Y). \tag{3}$$

We will denote the direct summands by $\mathcal{C}^{i,j}(Z) := \mathcal{C}^i(X) \otimes \mathcal{C}^j(Y)$. A correct definition of the boundary map of Z can be extracted from the boundary maps ∂_X and ∂_Y as

$$\begin{aligned} \partial_Z^{i,j} &: \mathcal{C}^{i,j}(Z) \to \mathcal{C}^{i-1,j}(Z) \oplus \mathcal{C}^{i,j-1}(Z), \\ \partial_Z^{i,j} &= \partial_X \otimes \mathrm{id}_Y + (-1)^i \mathrm{id}_X \otimes \partial_Y. \end{aligned} \tag{4}$$

This implicitly equips the cells of Z with an orientation inherited from the orientations of the simplices in X and Y. The multiplication with $(-1)^i$ is necessary for $\partial_Z \circ \partial_Z = 0$ to hold.

Hodge Laplacian of a Product Space. With these definitions in place, we can compute the Hodge Laplacian Δ_Z of the product space $Z := X \times Y$ as follows. Recall that we use X to denote the spatial simplicial complex, whereas the temporal simplicial complex Y consists of a path graph with a vertex for each time step t and an edge between each pair of successive time steps t and $t + 1$. For brevity, we will write ∂_X instead of $\partial_X \otimes \mathrm{id}_Y$ and similarly for ∂_Y, Δ_X, and Δ_Y. Denoting the restriction of the Hodge Laplacian Δ_Z to $\mathcal{C}^{i,j}(Z)$ by $\Delta_Z^{i,j}$ and using the commutativity of ∂_X with $\partial_Y^\top$, and $\partial_X^\top$ with ∂_Y, we obtain the following result.

Proposition 1. *The Hodge Laplacian Δ_Z restricted to $\mathcal{C}^{i,j}(Z)$ is given by*

$$\Delta_Z^{i,j} = \partial_Z \partial_Z^\top + \partial_Z^\top \partial_Z = \Delta_X + \Delta_Y, \tag{5}$$

where Δ_X and Δ_Y are the Hodge Laplacians of the spatial and temporal domain, constructed from ∂_X and ∂_Y, respectively.

Proof.

$$\begin{aligned} \Delta_Z^{i,j} &= \partial_Z \partial_Z^\top + \partial_Z^\top \partial_Z \\ &= \partial_X \partial_X^\top + (-1)^i \partial_X \partial_Y^\top + (-1)^{i+1} \partial_Y \partial_X^\top + \partial_Y \partial_Y^\top \\ &\quad + \partial_X^\top \partial_X + (-1)^i \partial_X^\top \partial_Y + (-1)^{i-1} \partial_Y^\top \partial_X + \partial_Y^\top \partial_Y \\ &= \Delta_X + \Delta_Y \qquad\qquad\qquad\qquad\qquad\qquad \square \end{aligned}$$

Because the above claim holds for arbitrary i and j, we see that the Hodge Laplacian acting on $\mathcal{C}^k(Z)$ simply sums the Hodge Laplacians Δ_X and Δ_Y acting on each constituent vector space of $\mathcal{C}^k(Z)$. Thus, it is reasonable to write the Hodge Laplacian on Z as being made up of restricted Laplacians of the form $\Delta_Z^{i,j} : \mathcal{C}^{i,j}(Z) \to \mathcal{C}^{i,j}(Z)$. This can be interpreted as an advantage of working with the product structure on a cell complex. When we are interested in evaluating the Hodge Laplacian only on the $(k + l)$-cells of Z which arose as products of k-simplices in X and l-simplices in Y, the signal values in $\mathcal{C}^{k,l}(Z)$ on the $(k + l)$-cells suffice as input. In contrast, in order to compute the Hodge Laplacian of an arbitrary subset of n-cells of Z, one needs the entire $(k + l)$-th signal space $\bigoplus_{i+j=k+l} \mathcal{C}^{i,j}(Z)$ of Z as input. This simplification yields a high degree of separability when considering the vector spaces $\mathcal{C}^{i,j}(Z)$, leading to a finer Hodge decomposition.

Filtering with the Product Space Hodge Laplacian. From our above discussion, we can conclude that the interpolation Eq. (2) for the spatiotemporal flows is simply governed by the Hodge Laplacian of the product space in disguise. To see this, let again X denote the spatial simplicial complex and Y the temporal simplicial complex. Furthermore, denote their product cell complex by $Z := X \times Y$ with

signal space $\mathcal{C}^\bullet(Z)$. Then the edge flow corresponds to the component $\mathcal{C}^{1,0}(Z)$ of the signal space. Using (5), we can split the Hodge Laplacian into Δ_X and Δ_Y. Then Δ_X is the spatial Hodge Laplacian $\mathbf{L}_s$ from (2) and Δ_Y the temporal Hodge Laplacian $\mathbf{L}_t$. We can weight the Hodge Laplacian of X by α_s to represent the spatial smoothness, and similarly, we can weight the Hodge Laplacian of Y by α_t to represent the temporal smoothness of the signal, so that $\Delta_Z^{1,0} = \alpha_s \Delta_X^1 + \alpha_t \Delta_Y^0$. This weighting of the Hodge Laplacian yields a quadratic form for temporal flows sufficient to express the interpolation program (2), for example.

4. REAL WORLD EXAMPLE: OCEAN DRIFTERS

In this section, we consider a dataset from the Global Drifter Program, which we restricted to the Caribbean[1]. The dataset consists of 1378 buoys floating in the ocean and the location of each buoy is logged every 6 hours, resulting in $532\,696$ location pings from the years 1992 to 2020.

The spatial simplicial complex is constructed following the setup considered in [26]: We place a hexagonal grid on the earth's surface, with the size of each hexagon corresponding to $0.3°$ (latitude). We associate a vertex with each hexagon and connect two vertices by an edge if their hexagons have a common face. Every triplet of hexagons that meet at a common point form a filled in triangle. Hexagons that cover landmasses are removed, which leads to some "holes" in the simplicial complex. The temporal domain corresponds to a path graph of length 29.

We discretize the observed trajectories according to our simplicial complex. A trajectory is represented by a vector $\mathbf{f}$ with entry $f_{[i,j]} = 1$ if the edge $[i, j]$ is traversed in the reference orientation, $f_{[i,j]} = -1$ if the edge is traversed in the opposite orientation, and 0 otherwise. The flow vector $\mathbf{f}_t$ of a particular year $t = 1, \ldots, T$ is the sum of the trajectory vectors of that year, where we split buoy trajectories that span multiple years into separate trajectories.

Finally, we divide the drifters into 80% training and 20% test trajectories and denote with $\widehat{\mathbf{f}}^{\mathrm{tr}}$ and $\widehat{\mathbf{f}}^{\mathrm{tst}}$ the training and test flows, respectively. We use the training trajectories to estimate the yearly ocean currents in the Caribbean and then check if these currents are compatible with the trajectories of the test drifters, e.g., if the buoy follows the ocean currents or moves perpendicular to them.

More specifically, from $\widehat{\mathbf{f}}^{\mathrm{tr}}$ we want to infer some suitable ocean currents $\mathbf{f}^*$ that describe the given trajectories best. Denote by $\mathbf{f}_{\widehat{\mathbf{f}}}$ the restriction of $\mathbf{f}$ to $\mathrm{supp}\,(\widehat{\mathbf{f}})$. Then this corresponds to minimizing the cosine similarity between $\widehat{\mathbf{f}}^{\mathrm{tr}}$ and $\mathbf{f}_{\widehat{\mathbf{f}}}$, i.e., we get the loss function

$$\mathcal{L}\left(\mathbf{f}, \widehat{\mathbf{f}}\right) = \frac{1}{2}\left(1 - \frac{\langle \mathbf{f}, \widehat{\mathbf{f}}\rangle}{\|\mathbf{f}_{\widehat{\mathbf{f}}}\|_2 \cdot \|\widehat{\mathbf{f}}\|_2}\right). \quad (6)$$

Notice that we normalize the cosine similarity such that 0 corresponds to perfect alignment and 1 to opposite flows.

We further want to infer the flow for hexagons without any training data, i.e., we want to use the resulting flow to infer the test trajectories. For that we add smoothness assumptions both in space and in time to the objective, i.e., we penalize if a hexagon's flow differs significantly from its neighbors' flows and if a hexagon's flow varies largely over time. The resulting convex optimization problem is as

Year 2003, $\alpha_t = 0.0$, $\alpha_s = 0.01$

Year 2003, $\alpha_t = 1.0$, $\alpha_s = 0.001$

| | Water | Land | Training | Test |

α_t	0.0	1.0	1.0	0.0
α_s	0.0	0.0	10^{-3}	10^{-2}
$\mathcal{L}(\mathbf{f}^*, \widehat{\mathbf{f}}^{\mathrm{tr}})$	**0.000**	**0.000**	**0.000**	**0.000**
$\mathcal{L}(\mathbf{f}^*, \widehat{\mathbf{f}}^{\mathrm{tst}})$	0.316	0.290	**0.264**	0.303

Fig. 3. **Quality of the inferred ocean currents with different smoothness hyperparameters.** Shown are the best cases without smoothness assumptions (first column), with pure temporal and pure spatial smoothness constraints (second and last column) and a combination of both (third). A loss of 0 corresponds to perfect alignment while 1 indicates opposite flows.

follows:

$$\mathbf{f}^* = \arg\min_{\mathbf{f}} \left(\mathcal{L}\left(\mathbf{f}, \widehat{\mathbf{f}}^{\mathrm{tr}}\right) + \frac{\frac{\alpha_s}{T}\sum_t \mathbf{f}_t^\top \mathbf{L}_s \mathbf{f}_t + \frac{\alpha_t}{|\mathcal{E}|}\sum_e \mathbf{f}^{e\top} \mathbf{L}_t \mathbf{f}^e}{\|\mathbf{f}\|_2^2}\right). \quad (7)$$

This amounts to minimizing $\mathcal{L}(\mathbf{f}, \widehat{\mathbf{f}}^{\mathrm{tr}}) + \langle \mathbf{f}, \Delta(\alpha_s, \alpha_t)\mathbf{f}\rangle/\|\mathbf{f}\|_2^2$, where $\Delta(\alpha_s, \alpha_t) = \alpha_s \Delta_X/T + \alpha_t \Delta_Y/|\mathcal{E}|$ denotes the weighted Hodge Laplacian. We solve this optimization problem for hyperparameters $\alpha_t, \alpha_s \in \{0\} \cup \{10^i : i = 0, \ldots, -5\}$ and report the losses in the table of Fig. 3. Almost all hyperparameters attain low loss on the training data, especially with those that give higher weight to spatial smoothing. For the test data, we observe that $\alpha_t = 1.0$ and $\alpha_s = 10^{-3}$ perform best. That is, enforcing more temporal smoothing than spatial smoothing is beneficial to infer unseen buoy trajectories. We assess that this is the case because ocean currents have fewer variations over the years as they have spatially.

5. DISCUSSION

We have shown that the Hodge Laplacian of product cell complexes provides a concise and effective framework for filtering spatiotemporal data. We showed the effectiveness of this product space framework by extrapolating ocean flows in the Caribbean from a limited set of data points varying in time and space. Further research could be made using more general product spaces. For instance, one could model a time complex where each time step is connected to the next time step and the time one year later. This approach could harvest periodicities on multiple scales in the data. This is connected to using a fiber bundle structure on the time space.

---

[1]Data available at https://www.aoml.noaa.gov/phod/gdp/data.php. We only consider entries in the rectangle spanned by 25°N 90°W (top left) and 10°N 55°W (bottom right). We ignore years before 1992 since there is no prior data in this area.
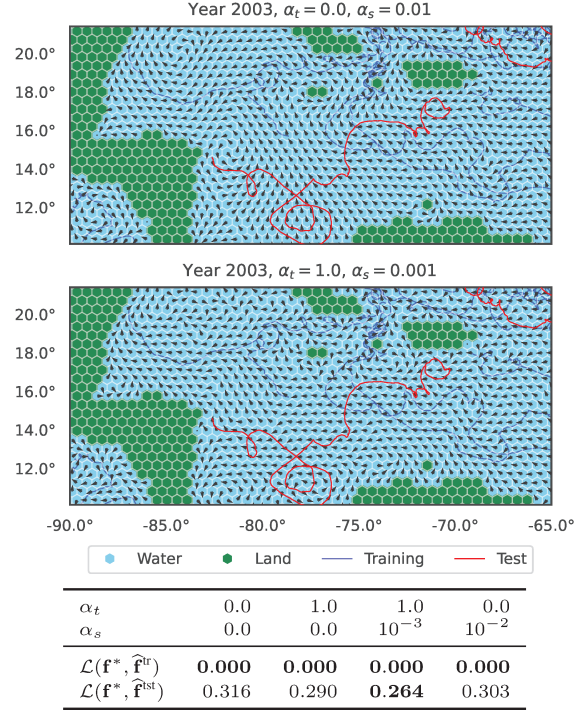
6. REFERENCES

[1] David I Shuman, Sunil K Narang, Pascal Frossard, Antonio Ortega, and Pierre Vandergheynst, "The emerging field of signal processing on graphs: Extending high-dimensional data analysis to networks and other irregular domains," *IEEE Signal Processing Magazine*, vol. 30, no. 3, pp. 83–98, Apr. 2013.

[2] Aliaksei Sandryhaila and José M. F. Moura, "Discrete signal processing on graphs," *IEEE Transactions on Signal Processing*, vol. 61, no. 7, pp. 1644–1656, 2013.

[3] Michael Scholkemper and Michael T. Schaub, "Blind extraction of equitable partitions from graph signals," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2022, pp. 5832–5836.

[4] Nathanaël Perraudin, Johan Paratte, David Shuman, Lionel Martin, Vassilis Kalofolias, Pierre Vandergheynst, and David K. Hammond, "Gspbox: A toolbox for signal processing on graphs," 2014.

[5] Michael T Schaub, Yu Zhu, Jean-Baptiste Seby, T Mitchell Roddenberry, and Santiago Segarra, "Signal processing on higher-order networks: Livin'on the edge... and beyond," *Signal Processing*, vol. 187, pp. 108149, 2021.

[6] Michael T. Schaub, Jean-Baptiste Seby, T. Mitchell Roddenberry, Yu Zhu, and Santiago Segarra, "Signal processing on simplicial complexes," in *Higher-Order Systems*, Federico Battiston and Giovanni Petri, Eds., pp. 301–328. Springer International Publishing, 2022.

[7] Sergio Barbarossa and Stefania Sardellitti, "Topological signal processing over simplicial complexes," *IEEE Transactions on Signal Processing*, vol. 68, pp. 2992–3007, 2020.

[8] Sergio Barbarossa, Stefania Sardellitti, and Elena Ceci, "Learning from signals defined over simplicial complexes," in *2018 IEEE Data Science Workshop (DSW)*. IEEE, 2018, pp. 51–55.

[9] Maosheng Yang, Elvin Isufi, Michael T Schaub, and Geert Leus, "Finite impulse response filters for simplicial complexes," *arXiv preprint*, 2021.

[10] T Mitchell Roddenberry, Michael T Schaub, and Mustafa Hajij, "Signal processing on cell complexes," *arXiv preprint arXiv:2110.05614*, 2021.

[11] Michael Robinson, *Topological signal processing*, vol. 81, Springer, 2014.

[12] Stefania Sardellitti, Sergio Barbarossa, and Lucia Testa, "Topological signal processing over cell complexes," in *Asilomar Conference on Signals, Systems, and Computers*. IEEE, 2021, pp. 1558–1562.

[13] Cristian Bodnar, Francesco Di Giovanni, Benjamin Paul Chamberlain, Pietro Liò, and Michael M. Bronstein, "Neural sheaf diffusion: A topological perspective on heterophily and oversmoothing in gnns," 2022.

[14] Florian Frantzen, Jean-Baptiste Seby, and Michael T. Schaub, "Outlier detection for trajectories via flow-embeddings," in *Asilomar Conference on Signals, Systems, and Computers*, 2021, pp. 1568–1572.

[15] Guillermo Ortiz-Jiménez, Mario Coutino, Sundeep Prabhakar Chepuri, and Geert Leus, "Sampling and reconstruction of signals on product graphs," in *IEEE Global Conference on Signal and Information Processing (GlobalSIP)*. IEEE, 2018, pp. 713–717.

[16] Aliaksei Sandryhaila and Jose MF Moura, "Big data analysis with signal processing on graphs: Representation and processing of massive data sets with irregular structure," *IEEE signal processing magazine*, vol. 31, no. 5, pp. 80–90, 2014.

[17] Jie Shi, Brandon Foggo, Xianghao Kong, Yuanbin Cheng, Nanpeng Yu, and Koji Yamashita, "Online event detection in synchrophasor data with graph signal processing," in *IEEE International conference on communications, control, and computing technologies for smart grids (SmartGridComm)*. IEEE, 2020, pp. 1–7.

[18] Francesco Grassi, Andreas Loukas, Nathanaël Perraudin, and Benjamin Ricaud, "A time-vertex signal processing framework: Scalable processing and meaningful representations for time-series on graphs," *IEEE Transactions on Signal Processing*, vol. 66, no. 3, pp. 817–829, 2017.

[19] Allen Hatcher, *Algebraic topology*, Cambridge University Press, 2002.

[20] Glen E. Bredon, *Topology and Geometry*, Graduate texts in mathematics. Springer-Verlag, New York, 1993.

[21] Danijela Horak and Jürgen Jost, "Spectra of combinatorial Laplace operators on simplicial complexes," *Advances in Mathematics*, vol. 244, pp. 303–336, 2013.

[22] Lek-Heng Lim, "Hodge Laplacians on graphs," *SIAM Review*, vol. 62, no. 3, pp. 685–715, 2020.

[23] T. Mitchell Roddenberry, Nicholas Glaze, and Santiago Segarra, "Principled simplicial neural networks for trajectory prediction," in *International Conference on Machine Learning*, 2021, vol. 139, pp. 9020–9029.

[24] Michael T Schaub and Santiago Segarra, "Flow smoothing and denoising: Graph signal processing in the edge-space," in *IEEE Global Conference on Signal and Information Processing*. IEEE, 2018, pp. 735–739.

[25] Junteng Jia, Michael T Schaub, Santiago Segarra, and Austin R Benson, "Graph-based semi-supervised & active learning for edge flows," in *ACM International Conference on Knowledge Discovery and Data Mining*, 2019, pp. 761–771.

[26] 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*, vol. 62, no. 2, pp. 353–391, 2020.