Parametric Dictionary Learning for Topological Signal Representation

Claudio Battiloro[1,2], Paolo Di Lorenzo[1], Alejandro Ribeiro[2]

[1]Department of Information Engineering, Electronics, and Telecommunications, Sapienza University, Rome, Italy.
[2]Electrical and Systems Engineering Department, University of Pennsylvania, Philadelphia, USA.
Emails: {claudio.battiloro,paolo.dilorenzo}@uniroma1.it, aribeiro@seas.upenn.edu

***Abstract*—The aim of this paper is to introduce a novel dictionary learning algorithm for sparse representation of signals defined over regular cell complexes. Leveraging tools from Hodge theory, we inject the underlying topology in the dictionary structure by parametrizing it as a concatenation of sub-dictionaries that are polynomial of Hodge Laplacians. The learning problem is cast as the joint optimization of the topological dictionary coefficients and the sparse signal representation, which is efficiently solved via an iterative alternating algorithm. Numerical results on synthetic data show the effectiveness of the proposed procedure in learning sparse representations of topological signals.**

I. Introduction

In the last few years, many processing and learning techniques have been proposed for signals defined over irregular, not necessarily metric, domains. As an example, graph signal processing (GSP) introduced several methods for analyze and process signals defined over the vertices of a graph. The definition of various graph (shift) operators lead to a variety of filters on graphs and graph Fourier transforms [1]. The main feature of these processing tools is that they intrinsically depend on the connectivity of the graph, which is encoded into the structure of the adopted graph shift operator. However, despite their overwhelming popularity, graph representations can only take into account *pairwise* relationships among data. In many situations, the interactions cannot be reduced to simple pairwise relationships, making graphs an inefficient object to model them [2], [3]. Prototypical examples are biological networks, in which *multi-way* interactions among complex substances (such as genes, proteins, or metabolites) [2] happen, or brain networks, in which groups of neurons typically activate at the same time [3]. These applications require processing tools that go beyond GSP, thus leading to the emergent field of topological signal processing (TSP) [4].

Motivation and Related Works. Pioneering works in TSP showed the benefits obtained by processing signals defined over simplicial or cell complexes, here referred as *topological signals*, which are specific examples of hyper-graphs with a rigorous algebraic description [4], [5]. Then, a series of papers have given important contributions to TSP, spanning from the introduction of FIR filters for simplicial and cell complex signals [6], [7], to the definition of a generalized Laplacian for embedding simplicial complexes into traditional graphs [8]. TSP also encouraged the development of deep neural architectures able to learn from data defined over topological spaces, see, e.g., [9]–[15].

A fundamental problem in signal processing is *sparse signal representation* [16], whose aim is designing overcomplete dictionaries of atoms that can represent signals as linear combinations of only a few atoms in the dictionary. Two main approaches have been proven successful for Euclidean and graph signals: i) analitycal dictionaries, i.e. structured dictionaries based on mathematical modelling that can be designed starting from the given domain and assuming a certain signal class (e.g. Fourier transforms, wavelets, curvelets, etc.); ii) *learnable dictionaries*, i.e. unstructured dictionaries that are learned from a set of training signals [17], [18]. An important tradeoff has to be taken into account in the choice of a specific approach: analytic dictionaries are usually faster to implement but are not resilient to model mismatching, while learnable dictionaries are usually robust to different signal classes but have larger complexity due to the required training phase. Regarding analytical approaches for topological dictionaries, generalizing what has been done for graph signals, a natural basis for signal representation is given by the topological Fourier modes [19]. Then, since Fourier modes are generally non-sparse and thus inefficient for representing localized signals, the work in [20] proposed a family of wavelets for simplicial signals, respecting the Hodge decomposition, whereas the work in [21] introduced topological Slepians, i.e., a class of signals that are maximally concentrated on the topological domain and perfectly localized on the spectral domain. However, to the best of our knowledge, strategies to learn dictionaries for sparse topological signal representation are missing in the current literature.

Contributions. In this work, hinging on formal arguments from algebraic topology [22], we generalize the work in [23] for graphs to regular cell complexes, thus introducing a novel class of learnable dictionaries for topological signals that are computational efficient and incorporate the topological structure. We exploit Hodge theory for regular cell complexes to inject the underlying structure into the dictionary, and we build the dictionaries to be a concatenation of sub-dictionaries parametrized as polynomials of the Hodge Laplacians. We propose two different parametrizations of topological dictionaries that lead to the same (non-convex) problem formulation. An efficient iterative alternating procedure then jointly learns the topological dictionary coefficients and the sparse signal

representation. Finally, we illustrate the advantages of the proposed topological dictionary learning strategies on a sparse signal representation task.

II. Background

In this section, we review some basics of topological signal processing over cell complexes.

Definition 1 (Regular Cell Complex). A *regular cell complex* is a topological space $\mathcal{X}$ together with a partition $\{\mathcal{X}_\sigma\}_{\sigma\in\mathcal{P}_\mathcal{X}}$ of subspaces $\mathcal{X}_\sigma$ of $\mathcal{X}$ called *cells*, where $\mathcal{P}_\mathcal{X}$ is the indexing set of $\mathcal{X}$, such that [22]:

1) For each $c \in \mathcal{X}$, every sufficient small neighborhood of c intersects finitely many $\mathcal{X}_\sigma$;
2) For all τ,σ we have that $\mathcal{X}_\tau \cap \overline{\mathcal{X}}_\sigma \neq \varnothing$ iff $\mathcal{X}_\tau \subseteq \overline{\mathcal{X}}_\sigma$, where $\overline{\mathcal{X}}_\sigma$ is the closure of the cell;
3) Every $\mathcal{X}_\sigma$ is homeomorphic to $\mathbb{R}^k$ for some k;
4) For every $\sigma \in \mathcal{P}_\mathcal{X}$ there is a homeomorphism ϕ of a closed ball in $\mathbb{R}^k$ to $\overline{\mathcal{X}}_\sigma$ such that the restriction of ϕ to the interior of the ball is a homeomorphism onto $\mathcal{X}_\sigma$.

Condition 2 implies that the indexing set $\mathcal{P}_\mathcal{X}$ has a poset structure, given by $\tau \leq \sigma$ iff $\mathcal{X}_\tau \subseteq \overline{\mathcal{X}_\sigma}$, and we say that τ *bounds* σ. This is known as the *face poset* of $\mathcal{X}$. The regularity condition 4 implies that all of the topological information about $\mathcal{X}$ is encoded in the poset structure of $\mathcal{P}_\mathcal{X}$. Then, a regular cell complex can be identified with its face poset. For this reason, from now on we will indicate the cell $\mathcal{X}_\sigma$ with its corresponding face poset element σ. The dimension $\dim(\sigma)$ of a cell σ is k, we call it a $k-$cell and denote it with σ^k to make this explicit. Regular cell complexes can be described via an incidence relation (boundary relation) with a reflexive and transitive closure that is consistent with the partial order introduced in Definition 1.

Definition 2 (Boundary Relation). We have the boundary relation $\sigma \prec \tau$ iff $\dim(\sigma) \leq \dim(\tau)$ and there is no cell δ such that $\sigma \leq \delta \leq \tau$.

In other words, Definition 2 states that the boundary of a cell σ^k of dimension k is the set of all cells of dimension less than k bounding σ^k. The dimension or order of a cell complex is the largest dimension of any of its cells and we denote an order K regular cell complex with $\mathcal{X}^K$. A graph is a particular case of a cell complex of order 1, containing only cells of order 0 (nodes) and 1 (edges). An example of a cell complex of order 2 is a graph with order 2 cells being some of its induced cycles that we refer to as *polygons*. In general, there is little interest with dimensions above two.

Let us denote the set of k-cells in $\mathcal{X}^K$ as $\mathcal{D}_k := \{\sigma_i^k : \sigma_i^k \in \mathcal{X}_K\}$, with $|\mathcal{D}_k| = N_k$.

Definition 3 (Topological Signals). A k-topological signal $\mathbf{s}_k$ over a regular cell complex $\mathcal{X}^K$ is defined as a collection of mappings from the set of all k-cells contained in the complex to real numbers:

$$\mathbf{s}_k = [s_k(\sigma_1^k), \ldots, s_k(\sigma_i^k), \ldots, s_k(\sigma_{N_k}^k)] \in \mathbb{R}^{N_k}, \quad (1)$$

where $s_k : \mathcal{D}_k \to \mathbb{R}$.

It is possible to give a combinatorial description of regular cell complexes. To do so, it is essential to introduce an orientation of the cells. Orienting cells is not mathematically trivial but, in the end, it is only a "bookkeeping matter" [7]; for this reasons, here we assume that a reference orientation of the complex is given, detailed explanations can be found in [7], [24]. At this point, there are two ways in which two cells can be considered to be adjacent: lower and upper adjacent. Two $k-$cells are lower adjacent if they share a common face of order $k-1$ and upper adjacent if both are faces of a cell of order $k+1$. The structure of a oriented regular cell complex of order K is then fully captured by the set of its incidence (or boundary) matrices $\mathbf{B}_k \in \mathbb{R}^{N_{k-1}\times N_k}$, $k = 1, \ldots, K$, with entries $B_k(i,j) = 0$ if σ_i^{k-1} is not a face of σ_j^k, and $B_k(i,j) = 1$ (or -1), if σ_i^{k-1} is a face of σ_j^k and its orientation is coherent (or not) with the orientation of σ_j^k. From the incidence information, we build the Hodge (or combinatorial) Laplacian matrices of order $k = 0, \ldots, K$ as [25]:

$$\mathbf{L}_0 = \mathbf{B}_1\mathbf{B}_1^T, \quad (2)$$

$$\mathbf{L}_k = \underbrace{\mathbf{B}_k^T\mathbf{B}_k}_{\mathbf{L}_k^{(d)}} + \underbrace{\mathbf{B}_{k+1}\mathbf{B}_{k+1}^T}_{\mathbf{L}_k^{(u)}}, \quad k = 1, \ldots, K-1, \quad (3)$$

$$\mathbf{L}_K = \mathbf{B}_K^T\mathbf{B}_K. \quad (4)$$

All Laplacian matrices of intermediate orders, i.e., $k = 1, \ldots, K-1$, contain two terms: The first term $\mathbf{L}_k^{(d)}$, also known as lower Laplacian, encodes the lower connectivity among k-order cells; the second term $\mathbf{L}_k^{(u)}$, also known as upper Laplacian, encodes the upper connectivity among k-order cells. For example, two edges are lower adjacent if they share a common vertex, whereas they are upper adjacent if they are faces of a common polygon. High order Laplacians admit the following Hodge decomposition [26].

Proposition 1 (Hodge Decomposition). The k-topological signal space $\mathcal{C}^k(\mathbb{R})$ can be decomposed as [22]:

$$\mathcal{C}^k(\mathbb{R}) = \mathbf{im}\big(\mathbf{B}_k^T\big) \bigoplus \mathbf{im}\big(\mathbf{B}_{k+1}\big) \bigoplus \mathbf{ker}\big(\mathbf{L}_k\big), \quad (5)$$

where $\bigoplus$ is the direct sum of vector spaces, and $\mathbf{ker}(\cdot)$ and $\mathbf{im}(\cdot)$ are the kernel and image spaces of a matrix, respectively.

Therefore, any topological signal $\mathbf{s}_k \in \mathcal{C}^k(\mathbb{R})$ admits the following orthogonal decomposition:

$$\mathbf{s}_k = \mathbf{B}_k^T\,\mathbf{s}_{k-1} + \mathbf{B}_{k+1}\,\mathbf{s}_{k+1} + \bar{\mathbf{s}}_k. \quad (6)$$

An interesting interpretation of (6) about the decomposition of edge flow signals $\mathbf{s}_1$ can be found in [4], [27].

Topological signals of various orders can be represented over the bases of the eigenvectors of the corresponding Hodge Laplacian matrices. Hence, using the eigendecomposition $\mathbf{L}_k = \mathbf{U}_k\mathbf{\Lambda}_k\mathbf{U}_k^T$, the Cell Complex Fourier Transform (CFT) of order k is defined as the projection of a k-order signal $\mathbf{s}_k$ onto the eigenvectors of $\mathbf{L}_k$ [27]:

$$\widehat{\mathbf{s}}_k \triangleq \mathbf{U}_k^T\,\mathbf{x}_k. \quad (7)$$

We refer to the eigenvalue domain of the CFT as the frequency domain (or spectrum). A consequence of the Hodge decomposition in (6) is that the eigenvectors belonging to $\mathbf{im}\big(\mathbf{L}_k^{(d)}\big)$ are orthogonal to those belonging to $\mathbf{im}\big(\mathbf{L}_k^{(u)}\big)$,

for all $k = 1, \dots, K-1$. Therefore, the eigenvectors (and eigenvalues) of $\mathbf{L}_k$ are given by the union of the eigenvectors (and eigenvalues) of $\mathbf{L}_k^{(u)}$ that we collect in a set $\mathcal{F}_k^{(u)}$, the eigenvectors (and eigenvalues) of $\mathbf{L}_k^{(d)}$ that we collect in a set $\mathcal{F}_k^{(d)}$, and the kernel of $\mathbf{L}_k$.

III. Cell Complex Convolutional Filters

In this section, we briefly review the cell complex FIR filters introduced in [27] and [7], that represent the fundamental block of the topological parametric dictionaries we will introduce in the next section.

Definition 4 (Cell Complex FIR Filters). Given the k-th Hodge Laplacian $\mathbf{L}_k$ as in (3), a cell complex FIR filter acting on $k-$topological signals is defined as a polynomial of the Laplacian as:

$$\mathbf{S}_k = \sum_{i=1}^{J} h_i^{(u)} \left(\mathbf{L}_k^{(u)}\right)^i + \sum_{i=1}^{J} h_i^{(d)} \left(\mathbf{L}_k^{(d)}\right)^i + h\mathbf{I}, \tag{8}$$

where J is a positive integer and $h_i^{(u)}$, $h_i^{(d)}$, $h \in \mathbb{R}$. Assigning two different sets of coefficients to the lower and upper Laplacians gives more flexibility. Setting $h_i^{(u)} = h_i^{(d)}$, $i = 1, \dots, J$ leads to filters of the form:

$$\mathbf{S}_k = \sum_{i=0}^{J} h_i \mathbf{L}_k^i, \tag{9}$$

which is the FIR filter proposed [7]. However, the filter in (9) cannot differentiate between the two types of adjacencies, resulting in a limited expressive power. In the sequel we focus on the case $k = 1$, neglecting the order subscript k for the sake of exposition (e.g., we indicate $\mathbf{L}_1$ with $\mathbf{L}$, $\mathbf{s}_1$ with $\mathbf{s}$, $\mathbf{S}_1$ with $\mathbf{S}$, $\mathcal{F}_1^{(u)}$ with $\mathcal{F}^{(u)}$, $\mathcal{F}_1^{(d)}$ with $\mathcal{F}^{(d)}$, and N_1 with N).

IV. Topological Dictionary Learning

Generalizing the approach proposed in [23] for parametric graph dictionary learning, we build a novel class of overcomplete topological dictionaries of the following form:

$$\mathbf{D} = [\mathbf{S}_1, ..., \mathbf{S}_P] \in \mathbb{R}^{N \times PN}, \tag{10}$$

where each $\mathbf{S}_p$, $p = 1, \dots, P$ is defined as in (8) and has a different set of coefficients. We collect the coefficients of the p-th sub-dictionary in a vector $\mathbf{h}_p \in \mathbb{R}^{2J+1}$ and the overall coefficients in a vector $\mathbf{h} = [\mathbf{h}_1, \dots, \mathbf{h}_P] \in \mathbb{R}^{(2J+1)P}$. The chosen polynomial structure gives localization guarantees. In particular, the $v-$th atom of the $p-$th sub-dictionary will have a component localized on the J-hop lower neighborhood of the $v-$th cell, and a component localized on the J-hop upper neighborhood of the $v-$th cell. Moreover, from (7) and (8), it is easy to see that the $j-th$ atom of the $p-$th sub-dictionary (the j-th column of $\mathbf{S}_p$) has the following spectrum:

$$\widehat{\mathbf{s}}_{p,j}(n) = \Bigg(\sum_{i=1}^{J} h_{p,i}^{(u)} \lambda_n^i \, \mathbb{I}(\lambda_n \in \mathcal{F}^{(u)}) + \sum_{i=1}^{J} h_{p,i}^{(d)} \lambda_n^i \, \mathbb{I}(\lambda_n \in \mathcal{F}^{(d)}) + h_p \Bigg) \mathbf{u}_n(j), \tag{11}$$

where $\mathbb{I}(\cdot)$ is the indicator function, λ_n is the $n-th$ eigenvalue of $\mathbf{L}$, and $\mathbf{u}_n(j)$ is the j-th component of the n-th column of $\mathbf{U}$ (the n-th eigenvector of $\mathbf{L}$). We refer to the coefficient in brackets in (11) as kernel and we denote it with $\widehat{\mathbf{s}}_p(\cdot)$ since, fixed a frequency, it is the same for all the atoms belonging to the same sub-dictionary. For this reason, to control the atoms frequency behavior, we can impose constraints on the kernels $\widehat{\mathbf{s}}_p(\cdot)$ [23]. We first impose that the spectra are non-negative and bounded in each sub-dictionary, i.e.:

$$0 \leq \widehat{\mathbf{s}}_p(n) \leq d, \quad p = 1, \dots, P, \; n = 1, \dots, N, \tag{12}$$

where $d > 0$. Equivalently, this constraint is requiring that each sub-dictionary has to be positive semi-definite with maximum eigenvalue bounded by d, i.e.,

$$\mathbf{0} \preccurlyeq \mathbf{S}_p \preccurlyeq d\mathbf{I}, \quad p = 1, \dots, P. \tag{13}$$

Moreover, we need to ensure that the whole spectrum is covered by the dictionary $\mathbf{D}$. Thus, we impose the constraint:

$$d - \epsilon \leq \sum_{p=1}^{P} \widehat{\mathbf{s}}_p(n) \leq d + \epsilon, \quad n = 1, \dots, N, \tag{14}$$

where ϵ is a small positive constant. This constraint is equivalent to requiring that that the sum of all of the sub-dictionaries $\mathbf{D}_p$ has the minimum eigenvalue lower-bounded by $d - \epsilon$ and the maximum eigenvalue upper-bounded by $d + \epsilon$, i.e.,

$$(d - \epsilon)\mathbf{I} \preccurlyeq \sum_{p=1}^{P} \mathbf{S}_p \preccurlyeq (d + \epsilon)\mathbf{I}. \tag{15}$$

We are now in the condition of formulating the dictionary learning problem. Given a training set of M order k topological signals $\mathbf{Y} = [\mathbf{y}_1, \dots, \mathbf{y}_M] \in \mathbb{R}^{N \times M}$, we aim to learn a dictionary as in (10), which can represent the training signals as a sparse linear combination of the atoms. Clearly, this is equivalent to learn the filters coefficients $\mathbf{h}$. Thus, we formulate the dictionary learning problem as the joint optimization of the dictionary coefficients and the sparse signal representation in the following way:

$$\begin{aligned} (\mathbf{h}^*, \mathbf{X}^*) = \arg\min_{\mathbf{h}, \mathbf{X}} \; & \|\mathbf{Y} - \mathbf{D}\mathbf{X}\|_F^2 + \gamma \|\mathbf{h}\|_2^2 \\ \text{subject to:} \\ & a) \; \|\mathbf{x}_i\|_0 \leq K_0, \; i = 1, \dots, M \\ & b) \; \mathbf{0} \preccurlyeq \mathbf{S}_p \preccurlyeq d\mathbf{I}, \; p = 1, \dots, P \\ & c) \; (d - \epsilon)\mathbf{I} \preccurlyeq \sum_{p=1}^{P} \mathbf{S}_p \preccurlyeq (d + \epsilon)\mathbf{I} \\ & d) \; \mathbf{S}_p \text{ as in (8)}, \; p = 1, \dots, P, \end{aligned} \tag{16}$$

where $\mathbf{x}_i$ is the $i-$th column of $\mathbf{X} \in \mathbb{R}^{PN \times M}$, i.e. the sparse signal representation of the $i-$th training signal, $\|\cdot\|_F$ is the Frobenius norm, $\|\cdot\|_0$ is the l_0 norm (i.e., the number of non-zero components of its vector argument), K_0 is an integer between 1 and N, and $\gamma > 0$. The l_2 penalty on the coefficients $\mathbf{h}$ is useful to ensure numerical stability and mitigate overfitting. Problem (16) is clearly non-convex,

Algorithm 1 : Topological Dictionary Learning

Inputs:

$\mathbf{Y} \in \mathbb{R}^{N \times M}$: Training signals.

$\mathbf{X}[0] \in \mathbb{R}^{PN \times M}$: Initialization of sparse representation

T_{max}: number of iterations (or stopping criterion)

Outputs:

$\mathbf{h}^*, \mathbf{X}^*$: Learned dictionary and sparse representations

function DICTIONARY LEARNING (**Inputs**)

  for $t \in [1, T_{max}]$ **do**

    Dictionary Update Step. Set:

$$\mathbf{h}[t] = \arg\min_{\mathbf{h}} \|\mathbf{Y} - \mathbf{D}\mathbf{X}[t-1]\|_F^2 + \gamma \|\mathbf{h}\|_2^2$$

subject to:

$$b)\ \mathbf{0} \preccurlyeq \mathbf{S}_p \preccurlyeq d\mathbf{I},\ p = 1, \dots, P$$

$$c)\ (d-\epsilon)\mathbf{I} \preccurlyeq \sum_{p=1}^{P} \mathbf{S}_p \preccurlyeq (d+\epsilon)\mathbf{I} \tag{17}$$

$$d)\ \mathbf{S}_p \text{ as in (8)},\ p = 1, \dots, P$$

    Sparse Coding Step. Set:

$$\mathbf{X}[t] = \arg\min_{\mathbf{X}} \|\mathbf{Y} - \mathbf{D}[t]\mathbf{X}\|_F^2 + \gamma \|\mathbf{h}\|_2^2$$

subject to:

$$a)\ \|\mathbf{x}_i\|_0 \leq K_0,\ i = 1, \dots, M \tag{18}$$

  return :

    $\mathbf{h}^* = \mathbf{h}[T_{max}]$

    $\mathbf{X}^* = \mathbf{X}[T_{max}]$

however it reduces to a semidefinite program (SDP) if the signal representation $\mathbf{X}$ is hold fixed, and to a sparse coding problem if the dictionary $\mathbf{D}$ is hold fixed. This fact suggests to approximate its solution with an iterative alternating procedure, which we list in Algorithm 1. As it is, Problem (17) is an SDP that can be solved with off the shelf numerical tools for convex optimization with a polynomial complexity. However, it is possibile to show that Problem (17) can be recast as a quadratic program, which can be solved even more efficiently. Instead, Problem (18) is still NP-hard, but its solution can be approximated via the orthogonal matching pursuit (OMP) algorithm [28], i.e. a versatile tool for sparse coding problems having linear complexity in PN [29].

V. NUMERICAL RESULTS

We asses the performance of the proposed dictionary learning algorithm on a sparse edge flow signal representation task. In particular, we consider a simplicial complex, i.e. a regular cell complex $\mathcal{X}^2$ including nodes (40), edges (100) and triangles (62) [4], and we process edge flow signals ($k = 1$ from (1)) defined on it. We generate 3 different sets of synthetic training and test (denoted with $\widetilde{Y} = [\widetilde{\mathbf{y}}_1, \dots, \widetilde{\mathbf{y}}_{M_T}] \in \mathbb{R}^{N \times M_T}$) signals consisting of localized patterns on the complex, with each training set computed using a different generating dictionary. In particular we generate a training set and

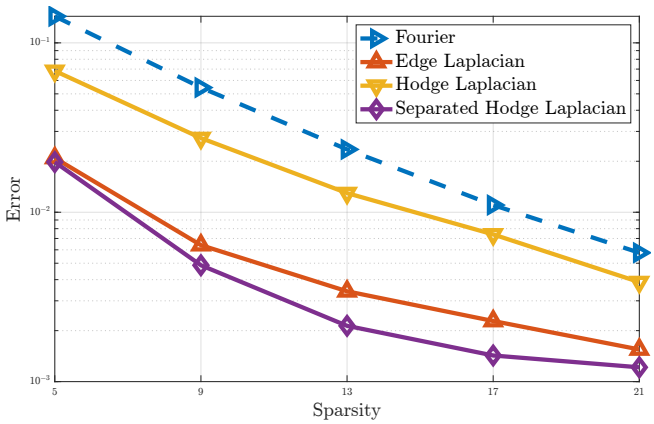

Fig. 1. NMSE versus sparsity, for different parametrization strategies. True dictionary generated from edge Laplacian.

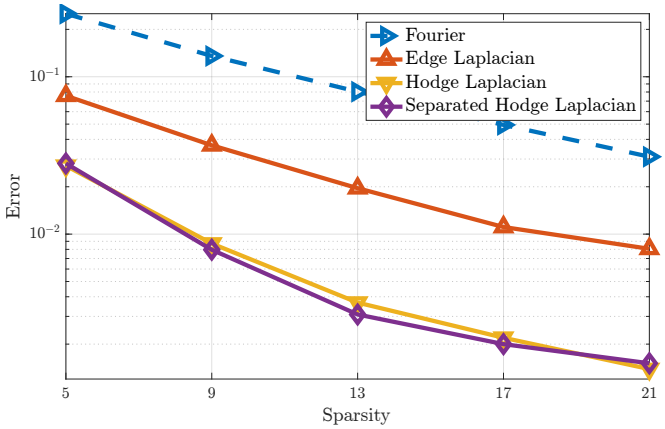

Fig. 2. NMSE versus sparsity, for different parametrization strategies. True dictionary generated from Hodge Laplacian.

a test set using: i) sub-dictionaries as in (9) (referred to as Hodge Laplacian); ii) a parametrization with polynomials of the Edge Laplacian [30], corresponding to (9) with only $\mathbf{L}^{(d)}$ in place of $\mathbf{L}$ (referred to as Edge Laplacian); iii) the proposed sub-dictionaries as in (8) (referred to as Separated Hodge Laplacian). The Edge Laplacian parametrization represents a fair comparison, because it is the naivest generalization to cell complexes of the work [23] for graphs, since it does not exploit Hodge theory. Per each dictionary type, we concatenate 3 subdictionaries, each of them being an order 3 polynomial ($P = 3$, $J = 3$); training and test sets are generated by linearly combining random atoms (at maximum 20) from the dictionaries with random coefficients; hyperparameters d and ϵ are chosen in agreement with the underlying generating dictionaries. In Figs. 1-2-3, we illustrate the behavior of the test NMSE $= \frac{1}{M_T} \sum_{m=1}^{M_T} \frac{\|\widetilde{\mathbf{y}}_m - \mathbf{D}\widetilde{\mathbf{x}}_m\|_2^2}{\|\widetilde{y}_m\|_2^2}$ (where $\mathbf{D}$ is the learned dictionary and $\widetilde{\mathbf{x}}_m$ is the sparse signal representation of the $m-$th test signal obtained via OMP) versus ℓ_0-norm of the sparse signal representations (i.e., K_0). We compare the aforementioned parametrizations per each type of generating dictionary (Fig. 1 for case i), Fig. 2 for case ii), Fig. 3 for

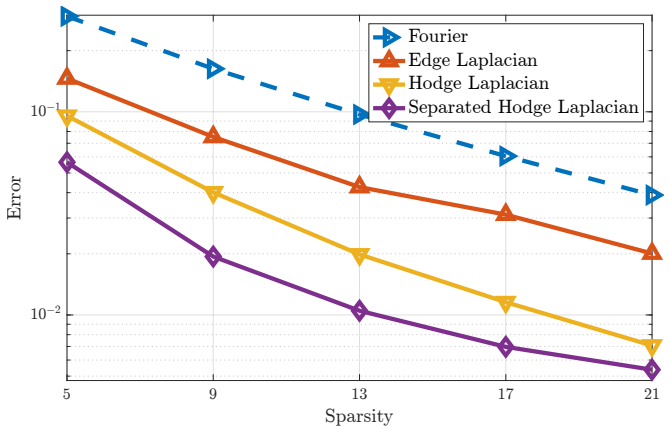

Fig. 3. NMSE versus sparsity, for different parametrization strategies. True dictionary generated from Separated Hodge Laplacian.

case iii)), and the results are averaged on 20 realizations of the dictionaries (thus, of the training and test sets). We also report the Fourier baseline to better enhance the advantages obtained by the dictionary learning strategies. As the reader can notice from Figs. 1-2-3, the parametrization corresponding to the true underlying generating dictionary always perform well but, notably, the proposed Separated Hodge Laplacian parametrization is able to mitigate the model mismatching in all the cases, showing the best performance and generalization capabilities with respect to all the considered strategies.

VI. Conclusions

In this work, we introduced a novel topological dictionary learning algorithm for sparse representation of signals defined over regular cell complexes. The proposed strategy is efficient thanks to a light parametrization of the dictionary structure driven by Hodge theory and topological signal filtering. We assessed its effectiveness and robustness on a synthetic sparse representation task. We plan to extend this work by refining the procedure to make it even more efficient, generalize it to exploit the interplay among multiple signals orders, provide theoretical guarantees in terms of frame bounds, and extensively test it on real world applications.

References

[1] Elvin Isufi, Fernando Gama, David I Shuman, and Santiago Segarra, "Graph filters for signal processing and machine learning on graphs," *arXiv preprint arXiv:2211.08854*, 2022.

[2] R. Lambiotte, M. Rosvall, and I. Scholtes, "From networks to optimal higher-order models of complex systems," *Nature physics*, vol. 15, no. 4, pp. 313–320, 2019.

[3] C. Giusti, R. Ghrist, and D. S. Bassett, "Two's company, three (or more) is a simplex," *Journal of computational neuroscience*, vol. 41, no. 1, pp. 1–14, 2016.

[4] S. Barbarossa and S. Sardellitti, "Topological signal processing over simplicial complexes," *IEEE Trans. on Signal Processing*, vol. 68, pp. 2992–3007, 2020.

[5] M. T. Schaub, Y. Zhu, J.B. Seby, T. M. Roddenberry, and S. Segarra, "Signal processing on higher-order networks: Livin'on the edge... and beyond," *Signal Processing*, vol. 187, pp. 108149, 2021.

[6] Maosheng Yang, Elvin Isufi, Michael T. Schaub, and Geert Leus, "Simplicial convolutional filters," 2022.

[7] T. Mitchell Roddenberry, Michael T. Schaub, and Mustafa Hajij, "Signal processing on cell complexes," in *ICASSP 2022 - 2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2022, pp. 8852–8856.

[8] Feng Ji, Giacomo Kahn, and Wee Peng Tay, "Signal processing on simplicial complexes with vertex signals," *IEEE Access*, vol. 10, pp. 41889–41901, 2022.

[9] C. Bodnar, F. Frasca, Y. Guang Wang, N. Otter, G. Montufar, P. Liò, and M. M. Bronstein, "Weisfeiler and Lehman go topological: Message passing simplicial networks," in *ICLR 2021 Workshop on Geometrical and Topological Representation Learning*, 2021.

[10] S. Ebli, M. Defferrard, and G. Spreemann, "Simplicial neural networks," in *NeurIPS 2020, Workshop on Topological Data Analysis and Beyond*.

[11] T. M. Roddenberry and S. Segarra, "Hodgenet: Graph neural networks for edge data," *Computing Research Repository (CoRR)*, vol. abs/1912.02354, 2019.

[12] L. Giusti, C. Battiloro, P. Di Lorenzo, S. Sardellitti, and S. Barbarossa, "Simplicial attention neural networks," arXiv:2203.07485, 2022.

[13] Ruochen Yang, Frederic Sala, and Paul Bogdan, "Efficient representation learning for higher-order data with simplicial complexes," in *Proceedings of the First Learning on Graphs Conference*, Bastian Rieck and Razvan Pascanu, Eds. 09–12 Dec 2022, vol. 198 of *Proceedings of Machine Learning Research*, pp. 13:1–13:21, PMLR.

[14] Claudio Battiloro, Zhiyang Wang, Hans Riess, Paolo Di Lorenzo, and Alejandro Ribeiro, "Tangent bundle filters and neural networks: from manifolds to cellular sheaves and back," Proc. of ICASSP 2023.

[15] Mustafa Hajij, Kyle Istvan, and Ghada Zamzmi, "Cell complex neural networks," in *Advances in Neural Information Processing Systems Workshop on TDA & Beyond*, 2020.

[16] O. Rioul and M. Vetterli, "Wavelets and signal processing," *IEEE Signal Proc. Mag.*, vol. 8, no. 4, pp. 14–38, 1991.

[17] Ivana Tošić and Pascal Frossard, "Dictionary learning," *IEEE Signal Processing Magazine*, vol. 28, no. 2, pp. 27–38, 2011.

[18] Pedro A Forero, Ketan Rajawat, and Georgios B Giannakis, "Prediction of partially observed dynamical processes over networks via dictionary learning," *IEEE transactions on signal processing*, vol. 62, no. 13, pp. 3305–3320, 2014.

[19] M. T. Schaub, A. R. Benson, P. Horn, G. Lippner, and A. Jadbabaie, "Random walks on simplicial complexes and the normalized hodge 1-laplacian," *SIAM Review*, vol. 62, no. 2, pp. 353–391, Jan 2020.

[20] T. Mitchell Roddenberry, Florian Frantzen, Michael T. Schaub, and Santiago Segarra, "Hodgelets: Localized spectral representations of flows on simplicial complexes," in *IEEE International Conference on Acoustics, Speech and Signal Processing*, 2022, pp. 5922–5926.

[21] Claudio Battiloro, Paolo Di Lorenzo, and Sergio Barbarossa, "Topological slepians: Maximally localized representations of signals over simplicial complexes," *Proc. of ICASSP 2023, arXiv:2210.14758*.

[22] L. J. Grady and J. R. Polimeni, *Discrete calculus: Applied analysis on graphs for computational science*, vol. 3, Springer, 2010.

[23] Dorina Thanou, David I Shuman, and Pascal Frossard, "Learning parametric dictionaries for signals on graphs," *IEEE Transactions on Signal Processing*, vol. 62, no. 15, pp. 3849–3862, 2014.

[24] Stefania Sardellitti, Sergio Barbarossa, and Lucia Testa, "Topological signal processing over cell complexes," in *2021 55th Asilomar Conference on Signals, Systems, and Computers*, 2021, pp. 1558–1562.

[25] T. E. Goldberg, "Combinatorial Laplacians of simplicial complexes," *Senior Thesis, Bard College*, 2002.

[26] L. Lim, "Hodge Laplacians on graphs," *Siam Review*, vol. 62, no. 3, pp. 685–715, 2020.

[27] Stefania Sardellitti and Sergio Barbarossa, "Topological signal representation and processing over cell complexes," 2022.

[28] Joel A. Tropp and Anna C. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *IEEE Trans. on Inf. Theory*, vol. 53, no. 12, pp. 4655–4666, 2007.

[29] Bob L Sturm and Mads Græsbøll Christensen, "Comparison of orthogonal matching pursuit implementations," in *2012 Proceedings of the 20th European Signal Processing Conference (EUSIPCO)*. IEEE, 2012, pp. 220–224.

[30] Daniel Zelazo, Amirreza Rahmani, and Mehran Mesbahi, "Agreement via the edge laplacian," in *2007 46th IEEE Conference on Decision and Control*, 2007, pp. 2309–2314.