AN INTRODUCTION TO HYPERGRAPH SIGNAL PROCESSING

Sergio Barbarossa, Mikhail Tsitsvero

Sapienza University of Rome, DIET, Via Eudossiana 18, 00184 Rome, Italy
E-mail: sergio.barbarossa@uniroma1.it, tsitsvero@gmail.com

ABSTRACT

Developing tools to analyze signals defined over a graph is a research area that is attracting a significant amount of contributions because of its many applications. However, a graph representation does not capture the overall information about the data, as it implicitly takes into account only pairwise relations. The goal of this paper is to extend signal processing tools to signals defined over hypergraphs, which represent a formal framework to describe multi-way relations among the data. First, we suggest alternative ways to introduce a Fourier Transform (FT) for signals defined over hypergraphs and, in particular, for simplicial complexes. Then, building on the notion of Fourier Transform, we derive a sampling theorem aimed at identifying the minimum number of samples necessary to encode all information about band-limited hypergraph signals.

Index Terms— hypergraph signals, simplicial complex, sampling

1. INTRODUCTION

The study of signals defined over a graph is a field that is receiving a lot of attention because of its many potential applications [1]. Given a graph $G(\mathcal{V}, \mathcal{E})$ composed by a set $\mathcal{V}$ of vertices and a set $\mathcal{E}$ of edges, a graph signal is typically represented as a mapping from the vertex set $\mathcal{V}$ to the the space of reals $\mathbb{R}$. As an example, we may think of a biological network, where the vertices represent proteins, enzymes, etc, whereas the presence of an edge between two nodes implies that the corresponding substances take part in a chemical reaction. The value of the signal over each vertex represents the concentration of the corresponding substance. In this case, the signal is evolving through time as a result of the chemical reactions. Another example is image segmentation. In such a case, the graph is given by the pixels of an image and the signal represents the luminance of each pixel. The edges among pixels represent similarities, expressed in terms of neighborhood and close values of luminance. Building on such a graph, it is possible to achieve image segmentation through graph partitioning [3].

In spite of its potentials, the graph signal model suffers from some major limitations. First of all, a graph representation is only indicative of pairwise relations. In general, multi-way correspondences are more informative and should be taken into account [10]. As an example, the need for going beyond pairwise representations in computer vision was motivated in [4]. Talking about image segmentation, similarity in terms of texture involves the analysis of a set of pixels. In such a case, pairwise relations are clearly insufficient. In molecular interaction networks, multilateral chemical interactions are not compatible with graph edges; graph representations are possible but they may imply a loss of information that can yield wrong interpretations [5], [6]. Moreover, graph representations are particularly well suited to emphasize clustering behaviors, which is an important category in many applications. However, clustering is not the only way to establish features of a set of points. Further categorizations may shed more light on the data under analysis. In this broader perspective, algebraic topology can provide very useful tools [8], [9]. Finally, a further limitation is that, typically, a graph signal establishes a correspondence between vertices and real (complex) numbers. However, in many cases, it is more informative to associate a signal value not only to vertices, i.e. singleton sets, but to subset of vertices of size greater than one, like edges, triads, and so on. The aim of this paper is to adopt convenient signal processing tools, such as Fourier transform and sampling theory, to deal with signals defined over hypergraphs and, in particular, over simplicial and cell complexes.

A hypergraph $\mathcal{H}(\mathcal{V}, \mathcal{E})$ is identified by a set $\mathcal{V}$ of elements (equivalently, nodes or vertices) and by a family $\mathcal{E}$ of subsets of $\mathcal{V}$ [7]. Each element e of $\mathcal{E}$ is called a *hyperedge*. A *weighted* hypergraph $\mathcal{H}(\mathcal{V}, \mathcal{E}, \boldsymbol{w})$ is a hypergraph with a non-negative number $w(e)$ associated to each hyperedge and $\boldsymbol{w}$ is the vector collecting all weights. A hypergraph signal is defined as a mapping from the set of hyperedges $\mathcal{E}$ to real numbers $\mathbb{R}$. Given two vertices v_1 and v_k, there is a *hyperpath* between them if there exists an alternating sequence of distinct vertices and hyperedges $v_1, e_1, v_2, e_2, \ldots, e_{k-1}, v_k$, such that $\{v_i, v_{i+1}\} \in e_i \in \mathcal{E}$. A hypergraph is *connected* if there is a hyperpath between every pair of vertices.

In many applications, the hypergraph signal $\boldsymbol{x}$ admits a compact representation, i.e., it can be expressed as $\boldsymbol{x} = \mathcal{U}\boldsymbol{s}$, where $\boldsymbol{s}$ is either exactly or approximately sparse. For example, a signal could be a smooth function over clusters and it could be sufficiently localized around the "voids" of a hyper-

The work of S. Barbarossa has been supported by TROPIC Project, Nr. 318784.

graph. In such a case, the columns of $\mathcal{U}$ may represent clusters or some topological features and the only nonzero (or approximately nonzero) entries of $\boldsymbol{s}$ are the ones associated to these features.

Sampling theory, on the other hand, aims to recover a sparse or band-limited graph signal from a subset of values. Given a subset of hyperedges $\mathcal{S} \subseteq \mathcal{E}$, we define a vertex-limiting operator as a diagonal matrix $\mathbf{D}_{\mathcal{S}}$ such that

$$\mathbf{D}_{\mathcal{S}} = \mathrm{Diag}\{\mathbf{1}_{\mathcal{S}}\} \tag{1}$$

where $\mathbf{1}_{\mathcal{S}}$ is the set indicator vector, whose i-th entry is equal to one, if $i \in \mathcal{S}$, or zero otherwise. Similarly, given a unitary matrix $\mathcal{U}$ and a subset of indices $\mathcal{F} \subseteq \mathcal{E}^*$, we introduce the operator

$$\mathbf{B}_{\mathcal{F}} = \mathcal{U}\,\mathbf{\Sigma}_{\mathcal{F}}\,\mathcal{U}^*, \tag{2}$$

where $\mathbf{\Sigma}_{\mathcal{F}}$ is a diagonal matrix defined as $\mathbf{\Sigma}_{\mathcal{F}} = \mathrm{Diag}\{\mathbf{1}_{\mathcal{F}}\}$ and $(\cdot)^*$ is the symbol denoting conjugate transpose. The role of $\mathbf{B}_{\mathcal{F}}$ is to project a vector $\boldsymbol{x}$ onto the subspace spanned by the columns of $\mathcal{U}$ whose indices belong to $\mathcal{F}$. It is immediate to check that both operators $\mathbf{D}_{\mathcal{S}}$ and $\mathbf{B}_{\mathcal{F}}$ represent orthogonal projectors. A signal $\boldsymbol{x}$ is perfectly *band-limited* over a frequency set $\mathcal{F}$ if $\mathbf{B}_{\mathcal{F}}\boldsymbol{x} = \boldsymbol{x}$. In the following and for the sake of simplicity, we will drop the subscripts referring to the sets whenever this will not cause any ambiguity.

In the following, we start by providing basic definitions of simplicial and cell complexes and considering alternative representations of hypergraphs in Section 2. Then, in Section 3 we illustrate possible ways to define a Fourier transform for these structures. Building on the definition of Fourier transform, we derive a sampling theorem establishing the conditions for perfect reconstruction of hypergraph signals from its samples based on the properties of the operators $\mathbf{D}$ and $\mathbf{B}$.

2. CELL AND SIMPLICIAL COMPLEXES

Alternative representations of hypergraphs are available. In the following, we will mostly rely on two basic and most commonly used classes of hypergraphs - *simplicial complexes* and *cell complexes*.

We start by considering a notorious subclass of hypergraphs that is called cell complexes [10]. A cell complex consists of a collection of finite dimensional p-cells, where each p-cell σ^p is defined as a set of points homeomorphic to a closed unit p-ball

$$\mathcal{B}_p = \{\boldsymbol{x} \in \mathbb{R}^p : \|\boldsymbol{x}\| \leq 1\}. \tag{3}$$

The boundary of the p-cell σ^p in this case is a part of the p-cell that is homeomorphic to the boundary of the unit ball, i.e.,

$$\partial\mathcal{B}_p = \{\boldsymbol{x} \in \mathbb{R}^p : \|\boldsymbol{x}\| = 1\}, \tag{4}$$

and consists of the set of $(p-1)$-cells. From the intuition above, a p-cell may be described by an ordered set of vertices that comprise a convex p-polytope. The lowest order, i.e. 0-cells are usually called vertices, 1-cells - edges and 2-cells - faces. Next, we consider a subclass of cell complexes - simplicial complexes, where each p-cell consists exactly of $(p+1)$ vertices. In general, cell complexes provide a more effective representation, because they do not restrict each p-cell to have the same number of vertices, however simplicial complexes are simpler to handle.

An abstract simplicial complex K on the set of vertices $\mathcal{V}$ is a collection of subsets of $\mathcal{V}$ that is closed under inclusion [11]. Formally, if $\sigma \in K$ and $\tau \subset \sigma$, then also $\tau \in K$. The simplicial complex may be weighted, in which case to each element of K may be assigned some weight. An i-simplex σ of K is an element of K with cardinality $i+1$, so that a 0-simplex is a vertex, a 1-simplex is an edge, a 2-simplex is a triangle, and so on. The set of i-simplices of K is denoted by $S_i(K)$. Next we fix an order of the vertices in $\mathcal{V}$, and assume that the orientation of faces is given by the ordering of the vertices. There are two possibilities to orient a simplex σ: all the simplices given by even permutations of the initial ordering are called positively oriented and each such simplex we denote by $+\sigma$, on the other hand a simplex belonging to the class of odd permutations is negatively oriented and is denoted by $-\sigma$. A simplicial i-chain is a formal finite sum of i-simplices

$$\sum_{\sigma \in S_i(K)} \alpha_\sigma \sigma, \tag{5}$$

with coefficients $\alpha_\sigma \in \mathbb{R}$. The group of all i-chains is called i-chain group of K and is denoted by $C_i(K, \mathbb{R})$. The group of functions from $C_i(K, \mathbb{R})$ to $\mathbb{R}$, that is a dual of the i-chain group $C_i(K, \mathbb{R})$, is denoted by $C^i(K, \mathbb{R})$ and is called i-cochain group. In fact, our signal $f \in \bigcup_{i=0}^{P} C^i(K, \mathbb{R})$, for some $P \leq |\mathcal{V}|$. The co-boundary map $\delta_i : C^i(K, \mathbb{R}) \to C^{i+1}(K, \mathbb{R})$ is given by

$$\delta_i f(\sigma) = \sum_{j=0}^{i+1} (-1)^j f([v_0, \ldots, v_{j-1}, v_{j+1}, \ldots, v_{i+1}]), \tag{6}$$

where $-1 \leq i < \dim K$. The map δ_i is therefore evaluated on the simplex σ and equals to the sum of f over all faces of σ taking into account their orientation. It is easy to observe that $\delta_i \delta_{i-1} = 0$, which means that $B^i(K, \mathbb{R}) := \mathrm{im}\ \delta_{i-1} \subseteq Z^i(K, \mathbb{R}) := \ker \delta_i$. The elements of the group $B^i(K, \mathbb{R})$ are called coboundaries and the elements of $Z^i(K, \mathbb{R})$ are cocycles. The cohomology group is defined as a quotient group $H^i(K, \mathbb{R}) = Z^i(K, \mathbb{R})/B^i(K, \mathbb{R})$. Using the above definition of the coboundary mapping one can define combinatorial k-Laplacians as [21], [20]

$$\mathcal{L}_k^{down} := \delta_{k-1}\delta_{k-1}^*, \tag{7}$$

$$\mathcal{L}_k^{up} := \delta_k^*\delta_k, \tag{8}$$

$$\mathcal{L}_k := \mathcal{L}_k^{down} + \mathcal{L}_k^{up}. \tag{9}$$

All k-Laplacians are self-adjoint and positive semidefinite. The spectrum of $\mathcal{L}_k$ is uniquely defined by the spectra of

$\mathcal{L}_k^{down}$ and $\mathcal{L}_k^{up}$. Moreover spectra of $\mathcal{L}_k^{up}$ and $\mathcal{L}_k^{down}$ coincide up to the multiplicity of zero eigenvalue. Therefore we will consider the eigenvalues of $\mathcal{L}_k^{up}$. By the Rayleigh-Ritz theorem, the eigenvalues of $\mathcal{L}_k^{up}$ are given by the following iterative optimization problem [19]

$$\lambda_j^{up} = \min_{\substack{f \perp B^i(K,\mathbb{R}) \\ f \perp u_l, l<j}} \frac{\langle \mathcal{L}_i^{up} f, f \rangle}{\langle f, f \rangle} = \min_{\substack{f \perp B^i(K,\mathbb{R}) \\ f \perp u_l, l<j}} \frac{\|\delta_i f\|^2}{\|f\|^2}, \quad (10)$$

where $\{u_l\}_{l=1,\ldots,i-1}$ are the optimal vectors found during the previous iterations. To interpret the result provided by (10) let us consider the case when $i = 0$. In this case the operator δ_0 assigns to each edge the value of difference of the linked vertices, therefore $\|\delta_0 f\|$ may be interpreted as a total variation of the signal f defined over vertices. In the same manner we can interpret $\|\delta_k f\|$ as a higher order total variation of some signal $f \in C^k(K,\mathbb{R})$ which lives on k-simplices.

The cohomology group $H^i(K,\mathbb{R})$ is the dual of the corresponding homology group $H_i(K,\mathbb{R})$, which in turn characterizes k-dimensional cycles or voids appearing in K [11]. Eigenvectors belonging to the $\ker \mathcal{L}_k$ can be used to characterize such cycles due to the fact that the kernel of the Laplacian $\mathcal{L}_k$ is isomorphic to the k-th homology group, i.e. $\ker\mathcal{L}_k \cong H_k(K,\mathbb{R})$. Homology group $H_k(K,\mathbb{R})$ in turn is a vector space over $\mathbb{R}$. The dimension of the corresponding homology group is denoted by $\beta_k = \dim H_k(K,\mathbb{R})$. The numbers β_k are known as *Betti numbers* and each of them characterizes the presence of a k-dimensional cycle. For example β_0 is equal to the number of connected components in K, β_1 denotes the number of non-contractible loops in K and so on for higher dimensions. Therefore we could also expect that eigenvectors of $\mathcal{L}_k$ corresponding to the small eigenvalues will characterize subsets where k-dimensional cycles are up to appear, in pretty much the same manner as the first eigenvectors of $\mathcal{L}_0$ identify clusters in a graph, i.e. dense subsets of K separated by small cuts.

It is interesting to note that the eigenvectors $\{\boldsymbol{u}_i^0\}$ of $\mathcal{L}_0$ are linked to the eigenvectors $\{\boldsymbol{u}_i^1\}$ of $\mathcal{L}_1$. To demonstrate this let us consider some eigenvector $\boldsymbol{u}^0$ of $\mathcal{L}_0$, i.e.

$$\mathcal{L}_0 \boldsymbol{u}^0 = \delta_1^* \delta_1 \boldsymbol{u}^0 = \lambda^0 \boldsymbol{u}^0. \quad (11)$$

By applying a co-boundary operator δ_1 to $\boldsymbol{u}^0$, we can see that the resulting vector is an eigenvector of $\mathcal{L}_1$ corresponding to the same eigenvalue λ^0. In fact,

$$\begin{aligned} \mathcal{L}_1 \delta_1 \boldsymbol{u}^0 &= \delta_1 \delta_1^* \delta_1 \boldsymbol{u}^0 + \delta_2^* \delta_2 \delta_1 \boldsymbol{u}^0 \\ &= \delta_1 \delta_1^* \delta_1 \boldsymbol{u}^0 = \lambda^0 \delta_1 \boldsymbol{u}^0, \end{aligned} \quad (12)$$

where we used $\delta_2 \delta_1 = 0$ which follows directly from (6).

3. HYPERGRAPH TOTAL VARIATION

The Graph Fourier Transform (GFT) has been recently introduced as a fundamental tool to analyze graph signals [22], [23], [1], [25], [2] . Alternative definitions exist, based on the projection of the observed signal onto the space spanned by the eigenvectors of either the Laplacian matrix [1] or the adjacency matrix [2]. Indeed, it should come with no surprise that there is no unique way to introduce the GFT, as alternative ways capture different features of the signal under analysis or may shed different light on the processing tools. A formal way to identify the Fourier basis involves the minimization of the *total variation* of a graph signal. Alternative definitions of total variations for graph signals have been used in [1], [2] for undirected and directed graphs.

A unified approach to define the total variation comes from the observation that the cut size is a *submodular* function. First we remind the definition of a cut set and its size. Given a subset of vertices $\mathcal{S}$, together with its complement $\bar{\mathcal{S}}$, the *hyperedge boundary* $\partial\mathcal{S}$ is defined as the set of hyperedges which are cut, i.e. $\partial\mathcal{S} := \{e \in \mathcal{E} | e \cap \mathcal{S} \neq \emptyset, e \cap \bar{\mathcal{S}} \neq \emptyset\}$. The cut size between $\mathcal{S}$ and $\bar{\mathcal{S}}$ is then

$$\mathrm{cut}(\mathcal{S}, \bar{\mathcal{S}}) := \sum_{e \in \partial\mathcal{S}} w(e). \quad (13)$$

As a consequence, its Lovász extension, which is a function defined over the real domain, is a convex function [17]. Hence, the total variation can be defined as the Lovász extension of the cut function. As an example, for directed graphs, the total variation of a real vector $\boldsymbol{x}$ over a directed graph is

$$\mathrm{TV}(\boldsymbol{x}) := \sum_{i,j} a_{ij} (x_i - x_j)^+ \quad (14)$$

where a_{ij} is the (i,j) entry of the adjacency matrix and $(x)^+ := \max(0,x)$. The total variation of a signal defined over an undirected hypergraph is [15]:

$$\mathrm{TV}(\boldsymbol{x}) = \sum_{e \in \mathcal{E}} w(e) \max_{i,j \in \mathcal{E}} |x_i - x_j|. \quad (15)$$

The Fourier basis can then be sought as the set of orthonormal vectors that minimize the total variation, as given in (14) for graphs, or (15) for hypergraphs. In all cases, in spite of the convexity of the objective function to be minimized, the problem is non-convex because of the unit norm vector constraint. For this reason, in graphs the following alternative definition of total variation is typically used [1]:

$$\mathrm{TV}(\boldsymbol{x}) := \sum_{i,j} a_{ij} (x_i - x_j)^2 = \boldsymbol{x}^* \mathbf{L} \boldsymbol{x}, \quad (16)$$

built on the Laplacian matrix $\mathbf{L}$. The minimization of this quadratic form, normalized to the square norm of $\boldsymbol{x}$, leads to an eigenproblem whose solution is given by the eigenvectors of $\mathbf{L}$ associated to the smallest eigenvalues of $\mathbf{L}$. Because of its simplicity, this approach can then be extended to define the Fourier Transform for signals defined over simplicial complexes, due to the fact that they posses simple topological

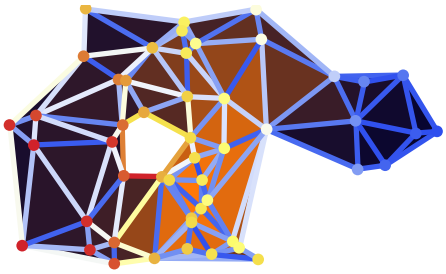

Fig. 1. Example of band-limited hypergraph signal defined over a simplicial complex of order $K = 2$.

structure. Let us denote by $\mathcal{L}_k = \mathcal{U}_k \Lambda_k \mathcal{U}_k^*$ the eigendecomposition of the k-th order Laplacian defined in (9). Every signal $\boldsymbol{x} = [\boldsymbol{x}_0, \dots, \boldsymbol{x}_K]^*$ defined over a simplicial complex of orders up to K, can be represented as

$$\boldsymbol{x} = \begin{pmatrix} \boldsymbol{x}_0 \\ \boldsymbol{x}_1 \\ \vdots \\ \boldsymbol{x}_K \end{pmatrix} = \begin{pmatrix} \mathcal{U}_0 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathcal{U}_1 & \cdots & \mathbf{0} \\ \cdots & \cdots & \cdots & \cdots \\ \mathbf{0} & \cdots & \cdots & \mathcal{U}_K \end{pmatrix} \begin{pmatrix} \boldsymbol{s}_0 \\ \boldsymbol{s}_1 \\ \vdots \\ \boldsymbol{s}_K \end{pmatrix} := \mathcal{U}\boldsymbol{s}.$$

The block $\boldsymbol{x}_0$ has dimension $|\mathcal{V}|$ and contains the values over all the vertices, the block $\boldsymbol{x}_1$ has dimension $|\mathcal{E}|$ and it contains the values of the signal over the edges, and so on. Typically, the vector $\boldsymbol{s}$ is sparse, because only a few eigenvectors are involved. The Fourier Transform of a signal defined over a simplicial complex can then be defined as

$$\hat{\boldsymbol{x}} = \mathcal{U}^* \boldsymbol{x} \tag{17}$$

with inverse Fourier transform $\boldsymbol{x} = \mathcal{U}\,\hat{\boldsymbol{x}}$. The main property of a FT defined in this way is that the signal is sparse when it varies smoothly either over the clusters defined over the vertices, or it tends to be localized over the edge cycles around the holes, or around voids enclosed by the faces of a simplicial complex. Therefore the sparsity pattern is informative about the topological properties which, in turn, are reflected in signal's structure. In Fig. 1, we show an example of simplicial complex of order $K = 2$ along with the structure of the eigenvectors corresponding to the second smallest eigenvalue of $\mathcal{L}_0$ and to the smallest eigenvalues of $\mathcal{L}_1$ and $\mathcal{L}_2$. The colors on the vertices represent the values of the second eigenvector of $\mathcal{L}_0$ and indeed emphasize two clusters. The colors on the edges encode absolute values of the first eigenvector of $\mathcal{L}_1$ and we can see that the energy is maximally concentrated around the hole. The absolute value of the entry of the first eigenvector of $\mathcal{L}_2$ is shown over triangles as different colors. The white region corresponds to the maximum amplitudes, thus emphasizing the region of the 2-simplices with higher degree of irregularities.

4. SAMPLING HYPERGRAPH SIGNALS

Let us study now the problem of sampling a hypergraph signal defined over a simplicial complex and derive the conditions for perfect reconstruction of a signal from its samples. For simplicity of notation, let us consider a signal defined over simplicial complex of order 2, i.e. vertices, edges and triangles and denote by $\mathcal{A} \equiv \mathcal{V} \cup \mathcal{E} \cup \mathcal{T}$ the ensemble of vertices, edges, and triangles where the signal is defined. By definition, the cardinality of $\mathcal{A}$ is $|\mathcal{A}| = |\mathcal{V}| + |\mathcal{E}| + |\mathcal{T}|$. Given a family of subsets $\mathcal{S} \subseteq \mathcal{A}$, we defined a set-limiting operator as a diagonal matrix $\mathbf{D}$ of size $|\mathcal{A}| \times |\mathcal{A}|$, whose diagonal entry is equal to 1 if the corresponding element (either vertex, edge or triangle) belongs to $\mathcal{S}$ or 0 otherwise. Correspondingly, Sampling a graph signal $\boldsymbol{x}$ over a set $\mathcal{S}$ of vertices gives rise to a vector $\boldsymbol{x}_s := \mathbf{D}\boldsymbol{x}$. The basic question with sampling is whether it is possible to reconstruct the overall signal from its samples and how. Generalizing our recent findings about graph signals in [18], we can prove the following theorem:

Theorem 1. *Given a band-limited signal $\boldsymbol{x} = \mathbf{B}\boldsymbol{x}$, it is possible to reconstruct $\boldsymbol{x}$ from its sampled version $\boldsymbol{x}_s = \mathbf{D}\boldsymbol{x}$ if and only if*

$$\|\mathbf{B}\overline{\mathbf{D}}\|_2 = \|\overline{\mathbf{D}}\mathbf{B}\|_2 < 1. \tag{18}$$

If condition (18) holds, the reconstruction formula is provided by the following theorem [18].

Theorem 2. *If condition (18) of the sampling theorem holds true, then any $\mathcal{F}$-band-limited signal $\boldsymbol{x}$ can be reconstructed from its sampled version $\boldsymbol{x}_\mathcal{S}$ by the following formula*

$$\boldsymbol{x} = \sum_{i=1}^{|\mathcal{F}|} \frac{1}{\sigma_i^2} \langle \boldsymbol{x}_\mathcal{S}, \boldsymbol{\psi}_i \rangle \boldsymbol{\psi}_i, \tag{19}$$

where $\{\boldsymbol{\psi}_i\}_{i=1..K}$ and $\{\sigma_i^2\}_{i=1..K}$ with $K = |\mathcal{F}|$, are the eigenvectors and eigenvalues of $\mathbf{BDB}$.

It is worth emphasizing that, even when condition (18) is met, the selection of the samples' location on each simplical complex is critical, as it directly affects the conditioning of the matrix to be inverted. Some of the strategies proposed for graphs in [18] can be extended to simplicial complexes, but the topic is clearly still open. Furthermore, in situations involving some kind of diffusion, the bandwidth over simplices of nearby orders can be correlated. This raises an interesting question on how to estimate the minimal bandwidth enabling a signal recovery under a maximum reconstruction error.

5. CONCLUSIONS

In this work we considered signals defined over hypergraphs. First we examined possible ways to define Fourier basis for hypergraph signals and for signals defined over cell or simplicial complexes in particular. Based on the definition of the Fourier transform we outlined some basic properties of basis vectors illustrating them with a numerical example. Finally, we approached the problem of sampling of hypergraph signals giving the condition and reconstruction formula.

6. REFERENCES

[1] D. I. Shuman, S. K. Narang, P. Frossard, A. Ortega, and P. Vandergheynst, "The emerging field of signal processing on graphs: Extending high-dimensional data analysis to networks and other irregular domains," *IEEE Signal Proc. Mag.*, vol. 30, no. 3, pp. 83-98, 2013.

[2] A. Sandryhaila and J. M. F. Moura, "Discrete signal processing on graphs: Frequency analysis," *IEEE Trans. on Signal Proc.*, vol. 62, pp. 3042-3054, 2014.

[3] J. Shi, J. Malik, "Normalized cuts and image segmentation," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Aug. 2000, pp. 888–905.

[4] Agarwal, S., J. Lim, L. Zelnik-Manor, P. Perona, D. Kriegman, and S. Belongie, "Beyond pairwise clustering," *IEEE Conf. on Computer Vision and Pattern Recognition*, 2005, pp. 838–845.

[5] S. Klamt, U.U. Haus, F. Theis, "Hypergraphs and cellular networks", *PLoS Computational Biology*, May 2009, pp. 1–6.

[6] Michoel, T. and B. Nachtergaele, "Alignment and integration of complex networks by hypergraph-based spectral clustering." *Physical Review E*, 2012.

[7] C. Berge, *Graphs and hypergraphs*, Vol. 7. Amsterdam: North-Holland publishing company, 1973.

[8] M. Robinson, *Topological Signal Processing*, Springer, 2014.

[9] H. Krim, A. Ben Hamza, *Geometric Methods in Signal and Image Analysis*, Cambridge University Press, 2015.

[10] L. J. Grady and J. Polimeni, *Discrete calculus: Applied analysis on graphs for computational science*, Springer Science & Business Media, 2010.

[11] J. R. Munkres, *Elements of Algebraic Topology.* Addison-Wesley Publ. Comp., 1984

[12] D. Zhou, J. Huang, and B. Schölkopf, "Learning with hypergraphs: Clustering, classification, and embedding", *Advances in Neural Information Processing Systems (NIPS 2007)*, pp. 1601–1608.

[13] S. Agarwal, K. Branson, and S. Belongie, "Higher order learning with graphs", *ICML 2006*, pp. 17-24.

[14] E. Ihler, D. Wagner, and F. Wagner, "Modeling hypergraphs by graphs with the same mincut properties", *Information Processing Letters*, pp. 171-175, 1993.

[15] M. Hein, S. Setzer, L. Jost, S. S. Rangapuram, "The Total Variation on Hypergraphs - Learning on Hypergraphs Revisited," *Advances in Neural Information Processing Systems (NIPS 2013)*, pp. 2427–2435.

[16] U. von Luxburg, "A tutorial on spectral clustering", *Statistics and Computing*, Dec. 2007, pp. 395–416.

[17] F. Bach, "Learning with Submodular Functions: A Convex Optimization Perspective", *Foundations and Trends in Machine Learning*, Nov. 2013.

[18] M. Tsitsvero, S. Barbarossa, P. Di Lorenzo, "Signals on Graph: Uncertainty Principle and Sampling", submitted to *IEEE Trans. on Signal Proc.*, July 2015; available at http://arxiv.org/abs/1507.08822; see also *Proc. of Asilomar Conf.*, Nov. 2015.

[19] A. Gundert and U. Wagner, "On Laplacians of random complexes", *Proceedings of the twenty-eighth annual symposium on Computational geometry*, ACM, 2012.

[20] D. Horak and J. Jürgen Jost, "Spectra of combinatorial Laplace operators on simplicial complexes", *Advances in Mathematics*, 244 (2013): 303-336.

[21] B. Eckmann, "Harmonische Funktionen und Randwertaufgaben in einem Komplex", *Comment. Math. Helv.* 17 (1945), 240255. MR 0013318 (7,138f)

[22] I. Z. Pesenson, "Sampling in Paley-Wiener spaces on combinatorial graphs," *Trans. of the American Mathematical Society*, vol. 360, no. 10, pp. 5603-5627, 2008.

[23] X. Zhu and M. Rabbat, "Approximating signals supported on graphs," in *IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, March 2012, pp. 3921-3924.

[24] A. Sandryhaila and J. M. F. Moura, "Discrete signal processing on graphs, *IEEE Trans. on Signal Proc.*, vol. 61, pp. 1644-1656, 2013.

[25] S.K. Narang, A. Gadde, and A. Ortega, "Signal processing techniques for interpolation in graph structured data," in *IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, May 2013, pp. 5445-5449.