Sampling of graph signals: successive local aggregations at a single node

Santiago Segarra, Antonio G. Marques, Geert Leus, and Alejandro Ribeiro

***Abstract*—A new scheme to sample bandlimited graph signals is proposed. The signals are defined in the nodes of a graph and admit a sparse representation in a frequency domain related to the structure of the graph, which is captured by the so-called graph-shift operator. Most of the existing works focused on using the value of the signal observed at a subset of nodes to recover the signal in the entire graph. Differently, the sampling scheme proposed here uses as input observations taken at a single node. The observations correspond to sequential applications of the graph-shift operator, which are linear combinations of the information gathered by the neighbors of the node. When the graph corresponds to a directed cycle, which is the support of time-varying signals, our method is equivalent to the classical sampling in the time domain. When the graph is more general, we show that the Vandermonde structure of the sampling matrix, which plays a critical role in guaranteeing recovery when sampling time-varying signals, is preserved.**

***Index Terms*—Graph signal processing, Sampling, Interpolation, Bandlimited graph signals.**

I. Introduction

The emergence of new fields of knowledge such as network science and big data calls for the extension of the results existing for classical time-varying signals to signals defined on graphs [1]–[3]. This not only entails modifying the algorithms currently available for time-varying signals, but also gaining intuition on what concepts are preserved (and lost) when a signal is defined, not in the classical time grid, but in a more general graph domain.

This paper investigates the sampling and posterior recovery of signals that are defined in the nodes of a graph. The underlying assumption is that such signals admit a sparse representation in a (frequency) domain which is related to the structure of the graph where these signals reside [4]. Most of the current efforts in this field have focused on using the value of the signal observed at a subset of nodes to recover the signal in the entire graph [4]–[6]. Our proposal in this paper is different. We present a new sampling method that accounts for the graph structure, can be run at a single node and only requires access to information of neighboring nodes. Moreover, we also show that the proposed method shares similarities with the classical sampling and interpolation of time-varying signals. If the graph corresponds to a directed cycle, which is the support of classical time-varying signals, our method is equivalent to classical sampling. When the graph is more general, the Vandermonde structure of the sampling matrix, which is critical to guarantee recovery in classical sampling [7], is preserved. Such a structure not only facilitates the interpolation process, but also helps to draw some connections between the proposed method and the sampling of time-varying signals. The conditions under which signal recovery is guaranteed are identified. These depend both on the structure of the graph and the particular node taking the observations. They also reveal that one way to understand bandlimited graph signals is to think of signals that can be well approximated by only observing the value of the signal at a small neighborhood. Simulations illustrating the application of our sampling scheme for both synthetic signals defined in random graphs and real-world graph signals close the paper.[1]

Work in this paper is supported by USA NSF CCF-1217963 and Spanish MINECO TEC2013-41604-R. S. Segarra and A. Ribeiro are with the Dept. of Electrical and Systems Eng., Univ. of Pennsylvania. A. G. Marques is with the Dept. of Signal Theory and Comms., King Juan Carlos Univ. G. Leus is with the Dept. of Electrical Eng., Math. and Computer Sci., Delft Univ. of Tech. Emails: {ssegarra,aribeiro}@seas.upenn.edu, antonio.garcia.marques@urjc.es, g.j.t.leus@tudelft.nl.

II. Sampling of graph signals

Let $\mathcal{G} = (\mathcal{N}, \mathcal{E})$ denote a directed graph. The set of nodes $\mathcal{N}$ has cardinality N, and the set of links $\mathcal{E}$ is such that $(i,j) \in \mathcal{E}$ if and only if node i is connected to node j. The set $\mathcal{N}_i : \{\, j \,|(j,i) \in \mathcal{E}\}$ contains all nodes with an incoming connection to i and is termed the incoming neighborhood of i. For any given graph, the adjacency matrix $\mathbf{A}$ is defined as a sparse $N \times N$ matrix with nonzero elements A_{ji} if and only if $(i,j) \in \mathcal{E}$. The value of A_{ji} captures the strength of the connection between i and j. The focus of this paper is not on analyzing $\mathcal{G}$, but a graph signal defined on the set of nodes $\mathcal{N}$. Such a signal can be represented as a vector $\mathbf{x} = [x_1, \ldots, x_N]^T \in \mathbb{R}^N$ where the i-th component represents the value of the signal at node i, or, equivalently, as a function $f : \mathcal{N} \to \mathbb{R}$, defined on the vertices of the graph.

The graph $\mathcal{G}$ is endowed with a *graph-shift operator* $\mathbf{S}$ defined as an $N \times N$ matrix whose entry (i,j), denoted as S_{ij}, can be nonzero only if $i = j$ or $(j,i) \in \mathcal{E}$. The sparsity pattern of the matrix $\mathbf{S}$ captures the local structure of $\mathcal{G}$, but we make no specific assumptions on the values of the nonzero entries of $\mathbf{S}$. Common choices for $\mathbf{S}$ are the adjacency matrix of the graph [2], [8] and its Laplacian [1]. The intuitive interpretation of $\mathbf{S}$ is that it represents a linear transformation that can be computed locally at the nodes of the graph. If $\mathbf{y} = [y_1, \ldots, y_N]^T$ is defined as $\mathbf{y} = \mathbf{S}\mathbf{x}$, then node i can compute y_i provided that it has access to the values of x_j at its incoming neighbors $j \in \mathcal{N}_i$. We assume henceforth that $\mathbf{S}$ is diagonalizable, i.e., that there exists a $N \times N$ matrix $\mathbf{V}$

[1] **Notation:** Boldface capital letters denote matrices and boldface lowercase letters column vectors. Generically, the entries of a matrix $\mathbf{X}$ and a vector $\mathbf{x}$ are denoted as X_{ij} and x_i; however, when contributing to avoid confusion, the alternative notation $[\mathbf{X}]_{ij}$ and $[\mathbf{x}]_i$ will be used. The notations ^{T} and ^{H} stand for transpose and transpose conjugate, respectively; $\text{diag}(\mathbf{x})$ is a diagonal matrix satisfying $[\text{diag}(\mathbf{x})]_{ii} = [\mathbf{x}]_i$; $\mathbf{e}_i$ is the i-th $N \times 1$ canonical vector (all entries of $\mathbf{e}_i$ are zero except the i-th one, which is one); $\mathbf{E}_K := [\mathbf{e}_1, ..., \mathbf{e}_K]$ is a tall matrix collecting the K first canonical vectors; and $\mathbf{0}$ is the all-zeros vector. The modulus obtained after dividing x by N is denoted as $\text{mod}_N(x)$.

and a $N \times N$ diagonal matrix $\mathbf{\Lambda}$ such that $\mathbf{S} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^{-1}$. In particular, the previous is true for normal matrices satisfying $\mathbf{S}\mathbf{S}^H = \mathbf{S}^H\mathbf{S}$. In that case we have that $\mathbf{V}$ is unitary, which implies $\mathbf{V}^{-1} = \mathbf{V}^H$.

A natural definition of sampling for a graph signal is to introduce a fat $K \times N$ selection matrix $\mathbf{C}$ and define

$$\bar{\mathbf{x}} = \mathbf{C}\mathbf{x}. \tag{1}$$

as the sampled signal [6]. If the matrix $\mathbf{C}$ is chosen as binary, i.e., with elements $C_{ij} \in \{0, 1\}$, and satisfying $\sum_j C_{ij} = 1$ for all i, then the signal $\bar{\mathbf{x}}$ is a selection of K out of the N elements of $\mathbf{x}$. In such a case, the ratio K/N is the sampling rate of the signal. Uniform sampling amounts to setting $\mathbf{C} = [\mathbf{e}_1, \mathbf{e}_{N/K+1}, \ldots, \mathbf{e}_{N-N/K+1}]^T$, while the selection of the first K elements of $\mathbf{x}$ is accomplished by setting $\mathbf{C} = \mathbf{E}_K^T := [\mathbf{e}_1, \ldots, \mathbf{e}_K]^T$. In general, it is not clear how to choose good selection matrices $\mathbf{C}$. This is in contrast to conventional sampling of signals in the time domain where uniform sampling is advantageous [7].

An equally valid, yet less intuitive, definition is to fix a node, say i, and consider the sampling of the signal seen by this node as the shift operator $\mathbf{S}$ is applied recursively. To describe this sampling methodology more clearly, define the l-th shifted signal $\mathbf{y}^{(l)} := \mathbf{S}^l\mathbf{x}$ and the $N \times N$ matrix

$$\mathbf{Y} := [\mathbf{y}^{(0)}, \mathbf{y}^{(1)}, \ldots, \mathbf{y}^{(N-1)}] = [\mathbf{x}, \mathbf{S}\mathbf{x}, \ldots, \mathbf{S}^{N-1}\mathbf{x}], \tag{2}$$

which groups the signal $\mathbf{x}$ and the result of the first $N-1$ applications of the shift operator. Associating the i-th row of $\mathbf{Y}$ with node i, we define the successively aggregated signal at i as $\mathbf{y}_i := (\mathbf{e}_i^T\mathbf{Y})^T = \mathbf{Y}^T\mathbf{e}_i$. Sampling is now reduced to the selection of K out of the N elements (rows) of $\mathbf{y}_i$, which we accomplish with a selection matrix $\mathbf{C}$ [cf. (1)]

$$\bar{\mathbf{y}}_i := \mathbf{C}\mathbf{y}_i = \mathbf{C}\left(\mathbf{Y}^T\mathbf{e}_i\right). \tag{3}$$

We say that the signal $\bar{\mathbf{y}}_i$ samples $\mathbf{x}$ with successive local aggregations. This nomenclature follows from the fact that $\mathbf{y}^{(l)}$ can be computed recursively as $\mathbf{y}^{(l)} := \mathbf{S}\mathbf{y}^{(l-1)}$ and that the i-th element of this vector can be computed using signals associated with itself and its incoming neighbors: $y_i^{(l)} = \sum_{j \in \mathcal{N}_i} S_{ij} y_j^{(l-1)}$. We can then think of the signal $\mathbf{y}_i$ as being computed locally at node i using successive variable exchanges with neighboring nodes. In fact, it is not difficult to see that $y_i^{(l)}$ is a linear combination of the values of x_j at nodes j whose distance (number of hops) from node i is less than or equal to l. Hence, the sampled signal $\bar{\mathbf{y}}_i$ in (3) is a selection of values that node i can determine locally.

Sampling in the time domain: To understand the difference between selection sampling [cf. (1)] and aggregation sampling [cf. (3)], it is instructive to consider their application to a signal defined in the time domain. Classical time domain signals can be represented as graph signals defined on top of a directed cycle graph [1], [6]. To do so define the directed cycle graph $\mathcal{G}_{dc}$ as one in which the edge set $\mathcal{E}_{dc} := \{(i, \mathrm{mod}_N(i)+1)\}_{i=1}^N$, connects node i to node $i+1$ for all nodes except N, which is connected to node 1. The leftmost graph in Fig. 1 is an example of a signal $\mathbf{x}$ defined on top of the directed cycle $\mathcal{G}_{dc}$ of length $N=6$. For this example, we consider selection sampling and aggregation sampling when using the adjacency matrix of $\mathcal{G}_{dc}$ as shift operator $\mathbf{S} = \mathbf{A}_{dc}$ and the uniform selection matrix $\mathbf{C} = [\mathbf{e}_1, \mathbf{e}_{N/K+1}, \ldots, \mathbf{e}_{N-N/K+1}]^T$. In selection sampling, we just multiply the graph signal $\mathbf{x}$ with the selection matrix $\mathbf{C}$ to obtain the sampled signal $\bar{\mathbf{x}} = \mathbf{C}\mathbf{x}$ as indicated by (1). If $K/N = 1/2$, in the example in Fig. 1 this amounts to selecting the values at the 3 odd nodes. In aggregation sampling, we consider subsequent applications of the shift matrix $\mathbf{S} = \mathbf{A}_{dc}$. Each of these shifts amounts to rotating the signal clockwise so that the element at node i moves to node $i+1$ for all $i < N$ and the element at node N moves to node 1. If we consider, e.g., node $i = 1$, the first shift moves signal x_N to this node so that $y_1^{(1)} = x_N$, the second shift moves signal x_{N-1} to this node so that $y_1^{(2)} = x_{N-1}$ and so on; see the 6 graphs in Fig. 1. It follows that the aggregated signal $\mathbf{y}_1$ in (2) is given by $\mathbf{y}_1 = [x_1, x_N, x_{N-1}, \ldots, x_2]$. This is just a shift of the original signal $\mathbf{x}$, which, upon multiplication by the selection matrix $\mathbf{C}$ as per (3) results in a vector $\bar{\mathbf{y}}_1 = \mathbf{C}\mathbf{y}_1$ that contains the same elements that $\bar{\mathbf{x}}$ contains.

For the cycle graph and shift operator $\mathbf{S} = \mathbf{A}_{dc}$ selection and aggregation sampling produce not only equivalent sampled signals but also reduce to conventional sampling. This is not a coincidence because both methods are designed as generalizations of conventional sampling. In general, selection sampling and aggregation sampling produce different outcomes. In selection sampling we move through nodes to collect samples at points uniquely identified by $\mathbf{C}$, whereas in aggregation sampling we move the signal through the graph while collecting samples at a fixed node. Observe that because aggregation sampling depends on the shift operator, it incorporates the structure of the graph into the sampling procedure. This is not true for selection sampling except for the choice of matrices $\mathbf{C}$ adapted to particular graphs.

III. Sampling of bandlimited graph signals

Recovery of the original signal from its sampled version is possible under the assumption that the original signal is bandlimited and therefore admits a sparse representation. The common practice when addressing the problem of sampling signals in graphs is to suppose that the graph-shift operator $\mathbf{S}$ plays a key *role in explaining* the signals of interest $\mathbf{x}$. More specifically, that $\mathbf{x}$ can be expressed as a linear combination of a *subset* of the columns of $\mathbf{V} = [\mathbf{v}_1, ..., \mathbf{v}_N]$, or, equivalently, that the vector $\widehat{\mathbf{x}} = \mathbf{V}^{-1}\mathbf{x}$ is sparse. In this context, vectors $\mathbf{v}_k$ are interpreted as the graph frequency basis and $\widehat{x}_k$ as the corresponding signal frequency coefficients. It will be assumed throughout that the active frequencies are the first K ones, associated with the eigenvalues of largest magnitude [4], [9]. Under this assumption, it holds that $\widehat{\mathbf{x}} = [\widehat{x}_1, ..., \widehat{x}_K, 0, ..., 0]^T$. However, the results presented in the paper can be applied to any set of active frequencies $\mathcal{K}$ of size K provided that $\mathcal{K}$ is known. For convenience, we define $\mathbf{V}_K := [\mathbf{v}_1, ..., \mathbf{v}_K]$ and $\widehat{\mathbf{x}}_K := [\widehat{x}_1, ..., \widehat{x}_K]^T$ so that we may write $\widehat{\mathbf{x}} = [\widehat{\mathbf{x}}_K^T, \ \mathbf{0}]^T$.

Notice that, for $\widehat{\mathbf{x}}$ to be sparse, it is reasonable to assume that $\mathbf{S}$ is involved in the generation of $\mathbf{x}$. When $\mathcal{G} = \mathcal{G}_{dc}$, it can be easily shown that setting the shift operator either to $\mathbf{S} = \mathbf{A}_{dc}$ or to $\mathbf{S} = \mathbf{L}_{dc} := \mathbf{I} - \mathbf{A}_{dc}$ gives rise to the Fourier

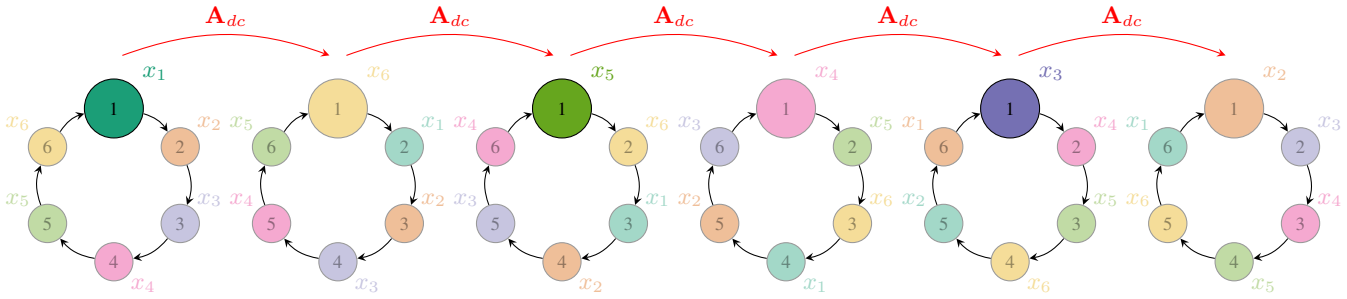


Fig. 1: Sampling in the time domain as aggregation sampling in a directed cycle graph. In aggregation sampling we utilize successive applications of a shift operator determined by the given graph and sample the resulting signal observed at a given node. Using the cycle adjacency matrix $\mathbf{A}_{dc}$ as shift operator results in the signal $\mathbf{x}$ rotating through the graph and the selection of elements of the aggregated signal reduces to conventional sampling. Aggregation sampling is, therefore, a generalization of conventional sampling to graph signals that utilizes the underlying graph structure in the construction of samples.

basis $\mathbf{F}$. More formally, that the right eigenvectors of $\mathbf{S}$ satisfy $\mathbf{V} = \mathbf{F}$, with $F_{ij} := N^{-1/2} e^{+\mathrm{j}\frac{2\pi}{N}(i-1)(j-1)}$ and $\mathrm{j} := \sqrt{-1}$. Selecting $\mathbf{S} = \mathbf{A}_{dc}$ has the additional advantage of satisfying $\Lambda_{ii} = e^{-\mathrm{j}\frac{2\pi}{N}(i-1)}$, i.e., the eigenvalues of the shift operator correspond to the classical discrete frequencies.

A. Selection sampling of bandlimited graph signals

Under the selection sampling approach [4]–[6], sampling a graph signal amounts to setting $\bar{\mathbf{x}} = \mathbf{C}\mathbf{x}$ [cf. (1)]. Since the $K \times N$ binary selection matrix $\mathbf{C}$ indexes the nodes that are observed, the issue then is how to design $\mathbf{C}$, i.e., which nodes to select, and how to recover the original signal $\mathbf{x}$ from its samples $\bar{\mathbf{x}}$. To answer these questions, it is assumed that the signal $\mathbf{x}$ is bandlimited, so that it can be expressed as a linear combination of the K principal eigenvectors in $\mathbf{V}$. The sampled signal $\bar{\mathbf{x}}$ is then $\bar{\mathbf{x}} = \mathbf{C}\mathbf{x} = \mathbf{C}\mathbf{V}_K\widehat{\mathbf{x}}_K$. Clearly, if the matrix $\mathbf{C}\mathbf{V}_K$ is *invertible*, then $\widehat{\mathbf{x}}_K$ can be recovered from $\bar{\mathbf{x}}$. Once $\widehat{\mathbf{x}}_K$ is known, the signal in the original domain is $\mathbf{x} = \mathbf{V}_K\widehat{\mathbf{x}}_K$. Combining the previous equations, we have

$$\mathbf{x} = \mathbf{V}_K\widehat{\mathbf{x}}_K = \mathbf{V}_K(\mathbf{C}\mathbf{V}_K)^{-1}\bar{\mathbf{x}}, \tag{4}$$

which reveals how the original signal can be interpolated from its samples. For (4) being true, $\mathbf{C}\mathbf{V}_K$ has to be invertible. Hence, the key for guaranteeing perfect reconstruction is to select a subset of nodes such that the corresponding rows in $\mathbf{V}_K$ are linearly independent. In the classical domain of time-varying signals, the (Fourier) basis has a Vandermonde structure, both row-wise and column-wise. This implies that any subset of K rows will give rise to a (row-wise) Vandermonde matrix and, hence, invertibility is guaranteed. However, for an arbitrary graph this is not guaranteed and algorithms to select subsets that ensure recovery are required [5].

B. Aggregation sampling of bandlimited graph signals

As explained in (3), under the aggregation approach the sampled signal is formed by observations of the shifted signals $\mathbf{y}^{(l)} = \mathbf{S}^l\mathbf{x}$ taken at a given node i. Under this second approach, the graph-shift operator $\mathbf{S}$ plays a key *role* not only in *explaining* and *recovering* $\mathbf{x}$, but also in *sampling* $\mathbf{x}$. Another reason to consider this scheme is that the entries of $\mathbf{y}^{(l)}$ can be found by sequentially exchanging information among neighbors. This implies that: a) for setups where graph vertices correspond to nodes of an actual network, the procedure can be implemented distributedly; and b) if recovery is feasible, the observations at a single node can be used to recover the signal in the entire graph.

Mimicking the approach in the previous section, we first analyze how the bandlimitedness of $\mathbf{x}$ is manifested on the sampled signal. Then, we identify under which conditions recovery is feasible and describe the corresponding interpolation algorithm. For the ease of exposition, the dependence of $\mathbf{y}_i$ on $\widehat{\mathbf{x}}$ is given in the form of a lemma.

Lemma 1: *Define the $N \times 1$ vector $\boldsymbol{\upsilon}_i := \mathbf{V}^T\mathbf{e}_i$, which collects the values of the frequency basis $\{\mathbf{v}_k\}_{k=1}^K$ at node i, and the $N \times N$ (column-wise) Vandermonde matrix*

$$\boldsymbol{\Psi} := \begin{pmatrix} 1 & \dots & 1 \\ \lambda_1 & \dots & \lambda_N \\ \vdots & & \vdots \\ \lambda_1^{N-1} & \dots & \lambda_N^{N-1} \end{pmatrix}. \tag{5}$$

Then, the shifted signal $\mathbf{y}_i$ can be expressed as

$$\mathbf{y}_i = \boldsymbol{\Psi}\mathrm{diag}(\boldsymbol{\upsilon}_i)\widehat{\mathbf{x}}. \tag{6}$$

Proof: Using the spectral decomposition of $\mathbf{S}$, signal $\mathbf{y}^{(l)}$ can be written as $\mathbf{y}^{(l)} = \mathbf{S}^l\mathbf{x} = (\mathbf{V}\boldsymbol{\Lambda}^l\mathbf{V}^{-1})\mathbf{x} = (\mathbf{V}\boldsymbol{\Lambda}^l)\widehat{\mathbf{x}}$. Based on the definitions of $\mathbf{y}_i$ and $\boldsymbol{\upsilon}_i$, it follows that $\mathbf{y}_i = \mathbf{Y}^T\mathbf{e}_i = (\mathbf{V}\mathbf{V}^{-1}\mathbf{Y})^T\mathbf{e}_i = (\mathbf{V}^{-1}\mathbf{Y})^T\mathbf{V}^T\mathbf{e}_i = (\mathbf{V}^{-1}\mathbf{Y})^T\boldsymbol{\upsilon}_i$. Since the l-th column of matrix $\mathbf{Y}$ is $\mathbf{y}^{(l-1)}$, this column can be written as $(\mathbf{V}\boldsymbol{\Lambda}^{l-1})\widehat{\mathbf{x}}$ and, therefore, the l-th column of $(\mathbf{V}^{-1}\mathbf{Y})$ can be written as $\boldsymbol{\Lambda}^{l-1}\widehat{\mathbf{x}}$ or, equivalently, as $\mathrm{diag}(\widehat{\mathbf{x}})[\lambda_1^{l-1}, ..., \lambda_N^{l-1}]^T$. Leveraging the fact that the vector containing the l-th power of the eigenvalues corresponds to the row $l+1$ of matrix $\boldsymbol{\Psi}$, the shifted signal $\mathbf{y}_i$ can be expressed as $\mathbf{y}_i = (\mathbf{V}^{-1}\mathbf{Y})^T\boldsymbol{\upsilon}_i = (\mathrm{diag}(\widehat{\mathbf{x}})\boldsymbol{\Psi}^T)^T\boldsymbol{\upsilon}_i = \boldsymbol{\Psi}\mathrm{diag}(\widehat{\mathbf{x}})\boldsymbol{\upsilon}_i = \boldsymbol{\Psi}\mathrm{diag}(\boldsymbol{\upsilon}_i)\widehat{\mathbf{x}}$, which is the claim in the lemma. ■

While in Sec. III-A the relationship between the sparse frequency coefficients $\widehat{\mathbf{x}}$ and the signal to be sampled was simply given by $\mathbf{x} = \mathbf{V}\widehat{\mathbf{x}}$, now it is given by $\mathbf{y}_i = \boldsymbol{\Psi}\mathrm{diag}(\boldsymbol{\upsilon}_i)\widehat{\mathbf{x}}$. Next, we use Lemma 1 to identify under which conditions recovery is feasible. To do this, let us define the $N \times K$ matrix $\boldsymbol{\Psi}_i = \boldsymbol{\Psi}\mathrm{diag}(\boldsymbol{\upsilon}_i)\mathbf{E}_K$. Then, the sampled signal $\bar{\mathbf{y}}_i$ is

$$\bar{\mathbf{y}}_i = \mathbf{C}\mathbf{y}_i = \mathbf{C}\boldsymbol{\Psi}\mathrm{diag}(\boldsymbol{\upsilon}_i)\widehat{\mathbf{x}} = \mathbf{C}\boldsymbol{\Psi}_i\widehat{\mathbf{x}}_K, \tag{7}$$

where $\mathbf{C}$ is the binary $K \times N$ selection matrix, and $\widehat{\mathbf{x}}_K$ the vector collecting the non-zero components of $\widehat{\mathbf{x}}$. To simplify exposition, for the time being we will assume that $\mathbf{C} = \mathbf{E}_K^T$,

i.e., that the observations correspond to the original signal and the first $K-1$ shifts. This assumption can be relaxed, as discussed in Remark 1.

If matrix $\mathbf{C}\mathbf{\Psi}_i$ is invertible, then $\widehat{\mathbf{x}}_K$ can be recovered from $\bar{\mathbf{y}}_i$ [cf. (7)] and, once $\widehat{\mathbf{x}}_K$ is known, $\mathbf{x}$ can be found as $\mathbf{x} = \mathbf{V}_K\widehat{\mathbf{x}}_K$. Combining the previous expressions, we have [cf. (4)]

$$\mathbf{x} = \mathbf{V}_K\widehat{\mathbf{x}}_K = \mathbf{V}_K(\mathbf{C}\mathbf{\Psi}_i)^{-1}\bar{\mathbf{y}}_i. \tag{8}$$

The equation shows how the original signal can be interpolated from its samples. As already stressed, for (8) to hold true, $\mathbf{C}\mathbf{\Psi}_i$ has to be invertible. Hence, the key for guaranteeing perfect signal reconstruction is to select samples such that the corresponding rows in $\mathbf{\Psi}_i$ are linearly independent. While for the selection sampling described in Sec. III-A there is no straightforward way to check the invertibility of $\mathbf{C}\mathbf{V}_K$ (existing algorithms typically do that by inspection [5]), for the aggregation sampling in (6)-(8), the invertibility of $\mathbf{C}\mathbf{\Psi}_i$ can be guaranteed if the two following conditions hold.

Proposition 1: *Let $\mathbf{x}$ and $\bar{\mathbf{y}}_i$ be, respectively, a bandlimited graph signal where at most K of the first frequency components are non-zero and the output of the sampling process as defined in (7) with $\mathbf{C} = \mathbf{E}_K^T$. Then, the N entries of signal $\mathbf{x}$ can be recovered from the K samples in $\bar{\mathbf{y}}_i$ if the two following conditions hold*

i) The first K eigenvalues of the graph-shift operator $\mathbf{S}$ are distinct; i.e., $\lambda_{k_1} \neq \lambda_{k_1}$ for all $k_1 \neq k_2$, $k_1 \leq K$ and $k_2 \leq K$.

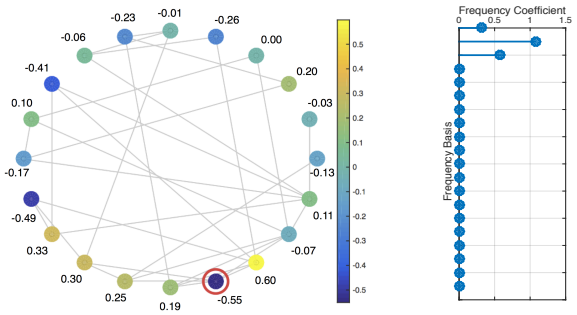
ii) The K first entries of $\boldsymbol{\upsilon}_i$ are non-zero.

Proof: To proposition is true because $\mathbf{C}\mathbf{\Psi}_i$ can be understood as the product of $\mathbf{C}\mathbf{\Psi}\mathbf{E}_K$ – invertible if *i)* holds true – and $\mathbf{E}_K^T\text{diag}(\boldsymbol{\upsilon}_i)\mathbf{E}_K$ – invertible if *ii)* holds true. ■

One of the implications of the proposition is that there is no need to compute or observe the entire vector $\mathbf{y}_i$, since its first K entries suffice to guarantee recovery. This readily implies that a K-bandlimited graph signal can be recovered by observing the values of the signal in a $K-1$ hop neighborhood.

The conditions in Prop. 1 are not difficult to check and they provide additional insights on the behavior of the sampling and interpolation procedure. Condition *i)* refers to the structure of the entire graph. It states that if a graph has two identical frequencies and the signal of interest is a linear combination of both of them, the sampling procedure will fail, regardless of the chosen node. Condition *ii)* refers to the specific node where the samples of the shifted signal are taken. It states that any node in the network can be used to sample the signal provided that $(\mathbf{e}_k^T\boldsymbol{\upsilon}_i) \neq 0$ for $k \leq K$; i.e., that the chosen node participates in the specific frequencies on which signal $\mathbf{x}$ is expressed. In practice, when noise is present, the recovery performance will depend on how strongly the particular node expresses the frequencies of interest. More precisely, since (8) requires inverting $\text{diag}(\mathbf{E}_K^T\boldsymbol{\upsilon}_i)$, nodes with low values of $\mathbf{e}_k^T\boldsymbol{\upsilon}_i$ will lead to reconstructions highly sensitive to noise.

Remark 1: The structure of the selection matrix $\mathbf{C}$ and, in particular, the fact that $\mathbf{C}\mathbf{\Psi}\mathbf{E}_K$ is a Vandermonde matrix are instrumental to guarantee the recovery of the original signal. Note that $\mathbf{C}\mathbf{\Psi}\mathbf{E}_K$ is Vandermonde not only when $\mathbf{C} = \mathbf{E}_K^T$, but also when $\mathbf{C} = [\mathbf{e}_1, \mathbf{e}_{1+N_0}, \ldots, \mathbf{e}_{1+(K-1)N_0}]^T$, provided that $1 \leq N_0 \leq N/K$ and $\lambda_{k_1}^{N_0} \neq \lambda_{k_2}^{N_0}$ for all $k_1 \neq k_2$, where $k_1 \leq K$ and $k_2 \leq K$. By setting $N_0 = N/K$, the counterpart of the classical time sampling theorem (which considers uniformly spaced samples) is recovered. Moreover, if $\lambda_k \neq 0$ for $k \leq K$, selection patterns of the form $\mathbf{C} = [\mathbf{e}_{n_0}, \mathbf{e}_{n_0+N_0}, \ldots, \mathbf{e}_{n_0+(K-1)N_0}]^T$ will also lead to invertible matrices. In this case, $\mathbf{C}\mathbf{\Psi}\mathbf{E}_K$ can be expressed as a product of a Vandermonde and a non-zero diagonal matrix.



(a) (b)

Fig. 2: A bandlimited graph signal. (a) The graph $\mathcal{G}$ and the graph signal $\mathbf{x}$ defined on the nodes of $\mathcal{G}$. The sampling node is circled in red. (b) Frequency components $\widehat{\mathbf{x}}$ of the graph signal $\mathbf{x}$. Given that there are three non-zero coefficients, the bandwidth of signal $\mathbf{x}$ is 3.

IV. Numerical experiments

We start by illustrating the perfect recovery of synthetic graph signals. Consider the 20-node undirected graph $\mathcal{G}$ depicted in Fig. 2a, which corresponds to a realization of a symmetric Erdős-Rényi graph with edge probability 0.20. With $\mathbf{A} = \mathbf{V}\mathbf{\Lambda}_A\mathbf{V}^H$ denoting the adjacency matrix of $\mathcal{G}$, three different graph-shift operators are considered: $\mathbf{S}_1 = \mathbf{A}$, $\mathbf{S}_2 = \mathbf{I} - \mathbf{A}$, and $\mathbf{S}_3 = 0.5\mathbf{A}^2$. Even though the support of $\mathbf{S}_3$ differs from that of $\mathbf{S}_1$ and $\mathbf{S}_2$, the graph-shift operator $\mathbf{S}_3$ still preserves the notion of locality as defined by a two-hop neighborhood. Note also that the three shift operators share the same set of eigenvectors $\mathbf{V}$, but they have a different set of eigenvalues. Let $\mathbf{x}$ be a graph signal supported on $\mathcal{G}$. This signal is represented in Fig. 2a where the value of the signal at every node is written explicitly next to each node and also coded by its color. Although seemingly random in the node domain, the structure of the signal $\mathbf{x}$ is highly determined by the graph. To illustrate this, Fig. 2b plots the frequency components $\widehat{\mathbf{x}}$ of signal $\mathbf{x}$, where the graph frequency basis is given by the columns of $\mathbf{V}$. The figure shows that $\mathbf{x}$ has a bandwidth of $K = 3$. Since $\mathbf{V}$ is the basis for $\mathbf{S}_1$, $\mathbf{S}_2$ and $\mathbf{S}_3$, the frequency representation $\widehat{\mathbf{x}}$ and the bandwidth K are the same for the three operators. Thus, the procedure in Sec. III-B will recover the whole signal using three aggregated samples, no matter which operator is chosen for the aggregation.

Suppose that we select node $i = 4$ as sampling node, which is circled in red in Fig. 2a. If the shift is $\mathbf{S}_1$, the 3 first observations taken by that node are $\mathbf{y}_4 = [-0.55, 1.27, -2.94]^T$. The first observation corresponds to the value of the signal at node 4, the second one to the aggregated signal at its neighbors and the third observation corresponds to a linear combination of the signal values within its two-hop neighborhood. Since $K = 3$, Prop. 1 guarantees recovery if: i) the 3 first eigenvalues of the shift operator are distinct and ii) the 3 first values of $\boldsymbol{\upsilon}_4$ are non-zero. It turns out that for $\mathbf{S}_1$ and node 4 these two conditions hold true and, hence, the interpolation in (8)

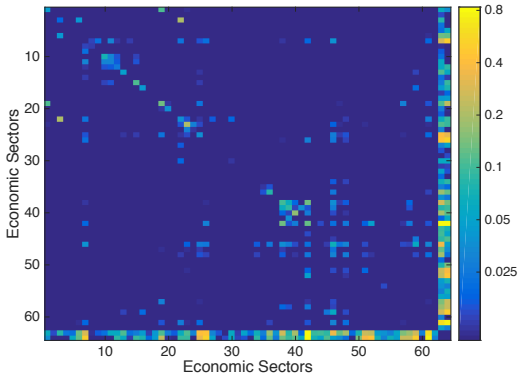


Fig. 3: Heat map of the graph-shift operator $\mathbf{S}$ of the economic network. It is sparse across the real economic sectors (from sector 1 to 62) while the artificial sectors AV and FU are highly connected.

yields the original signal in Fig. 2a. In fact, for the network at hand, these two conditions are satisfied for all nodes and shift operators considered. This implies that perfect reconstruction is achieved regardless of the node and shift operator – among the three presented – carrying out the aggregation. To better asses the conditions in Prop. 1, we build 10,000 different random connected graphs where the edge probability is chosen from the interval $[0.15, 0.25]$. We vary the number of nodes from 10 to 30 and the active frequencies from 1 to 5. For each random graph and signal defined on it, we test for perfect signal recovery on every node. The results show that in 99.89% of the cases the signal is successfully recovered.

We now present results for real-world graph signals corresponding to the exchange among the different sectors of the economy of the United States. The Bureau of Economic Analysis publishes a yearly table of inputs and outputs organized by economic sectors [10]. More precisely, we have a set $\mathcal{N}$ of 62 industrial sectors and a similarity function $U : \mathcal{N} \times \mathcal{N} \to \mathbb{R}_+$ where $U(i, i')$ represents how much of the production of sector i, expressed in trillions of dollars per year, was used as an input of sector i' on average during years 2008, 2009, and 2010. Moreover, for each sector we are given two economic markers: the added value (AV) generated and the level of production destined to the market of final users (FU). Thus, we define a graph on the set of $N = 64$ nodes comprising the original 62 sectors plus the two artificial ones (AV and FU) and an associated symmetric graph-shift operator $\bar{\mathbf{S}}$ defined as $\bar{\mathbf{S}}_{ij} = (U(i, j) + U(j, i))/2$. We then threshold $\bar{\mathbf{S}}$ in order to make it more sparse by setting to 0 all the values lower than 0.01 to obtain the shift operator $\mathbf{S} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^H$, which is normal given that it is symmetric; see Fig. 3. Consider now the signal $\mathbf{x} \in \mathbb{R}^{64}$ on the mentioned graph where $\mathbf{x}$ contains the total production – in trillion of dollars – of each sector (including AV and FU) during year 2011. Signal $\mathbf{x}$ is approximately bandlimited in $\mathbf{S}$ since most of the elements of $\widehat{\mathbf{x}} = \mathbf{V}^H\mathbf{x}$ are close to zero. In particular, the first 4 frequency coefficients contain 99.65% of the signal's energy. Hence, in what follows we interpret the graph signal $\mathbf{x}$ as a "noisy" realization of a signal of bandwidth $K = 4$.

Perfect recovery with $K = 4$ samples is not feasible in this case and the goal is to approximate $\mathbf{x}$. As anticipated in Sec. III-B, the quality of the reconstruction depends on the sampling node. The best reconstruction is achieved by the artificial sectors AV and FU, which are connected and closely related to every other sector. If those two sectors are excluded, the best reconstruction among real sectors is achieved by 'Insurance Carriers', which is also well connected to most sectors of the economy (cf. node 40 in Fig. 3). The worst is given by 'Publishing Industries' (node 34). This sector is poorly connected and the first 4 entries of $\boldsymbol{\upsilon}_{34}$ are very small (cf. discussion after Prop. 1). The median reconstruction error across nodes is 1.9%. An alternative approach is to implement selection sampling, i.e., to sample the signal $\mathbf{x}$ in 4 different sectors – excluding the artificial sectors AV and FU – and interpolate the whole signal from these 4 observations. Recall that reconstruction is not guaranteed for every subset of 4 nodes since we must have invertibility of $(\mathbf{C}\mathbf{V}_K)$ [cf. (4)]. Indeed, the *median* reconstruction error in this case is 420%. The differences in terms of the best recovery performance are less noticeable: 0.35% for the best node in aggregation sampling and 0.39% for the the best subset of four nodes in selection sampling. Equally interesting, if more general sampling strategies (like the ones described in Remark 1) are implemented, the median error can be reduced down to 0.40% while keeping the minimum error at 0.35%.

V. Conclusions

The problem of sampling bandlimited graph signals was investigated. A new scheme based on the aggregation of local information at a single node after successive applications of the graph-shift operator was proposed. This contrasts most existing works, which focus on observing the value of the signal at a subset of nodes. Our scheme was shown to be equivalent to classical sampling for directed cycle graphs whereas, for more general graphs, the Vandermonde structure of the sampling matrix was exploited to find the conditions for perfect reconstruction.

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 Process. Mag.*, vol. 30, no. 3, pp. 83–98, 2013.

[2] A. Sandryhaila and J. Moura, "Discrete signal processing on graphs," *IEEE Trans. Signal Process.*, vol. 61, no. 7, pp. 1644–1656, April 2013.

[3] X. Zhu and M. Rabbat, "Approximating signals supported on graphs," in *IEEE Intl. Conf. Acoust., Speech and Signal Process. (ICASSP)*, March 2012, pp. 3921–3924.

[4] A. Anis, A. Gadde, and A. Ortega, "Towards a sampling theorem for signals on arbitrary graphs," in *IEEE Intl. Conf. Acoust., Speech and Signal Process. (ICASSP)*, 2014, pp. 3864–3868.

[5] I. Shomorony and A. S. Avestimehr, "Sampling large data on graphs," *CoRR*, vol. abs/1411.3017, 2014.

[6] S. Chen, R. Varma, A. Sandryhaila, and J. Kovačević, "Discrete signal processing on graphs: Sampling theory," *arXiv preprint arXiv:1503.05432*, 2015.

[7] M. Unser, "Sampling-50 years after Shannon," *Proc. IEEE*, vol. 88, no. 4, pp. 569–587, 2000.

[8] A. Sandryhaila and J. Moura, "Discrete signal processing on graphs: Frequency analysis," *IEEE Trans. Signal Process.*, vol. 62, no. 12, pp. 3042–3054, June 2014.

[9] M. Rabbat and V. Gripon, "Towards a spectral characterization of signals supported on small-world networks," in *IEEE Intl. Conf. Acoust., Speech and Signal Process. (ICASSP)*, May 2014, pp. 4793–4797.

[10] Bureau of Economic Analysis, "Input-output accounts: the use of commodities by industries before redefinitions," *U.S. Dept. of Commerce*, 2011. [Online]. Available: http://www.bea.gov/iTable/index_industry.cfm