SPACE-SHIFT SAMPLING OF GRAPH SIGNALS

Santiago Segarra[†]*, Antonio G. Marques*[*]*, Geert Leus*[‡]*, and Alejandro Ribeiro*[†]

[†]Dept. of ESE, University of Pennsylvania, Philadelphia, PA, USA
[*]Dept. of TSC, King Juan Carlos University, Madrid, Spain
[‡]Dept. of EEMCS, Delft University of Technology, Delft, Netherlands

ABSTRACT

A novel scheme for sampling graph signals is proposed. *Space-shift* sampling can be understood as a hybrid scheme that combines selection sampling – observing the signal values on a subset of nodes – and aggregation sampling – observing the signal values at a single node after successive aggregation of local data. Under the assumption of bandlimitedness, we state conditions and propose strategies for signal recovery in different settings. Being a more general procedure, space-shift sampling achieves smaller reconstruction errors than current schemes, as we illustrate through the reconstruction of the industrial activity in a graph of the U.S. economy.

Index Terms— Graph signal processing, Space-shift sampling, Bandlimited signal, Reconstruction.

1. INTRODUCTION

Sampling is one of the most studied problems in classical signal processing [1]. The rise of new areas 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 [2–4]. 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 in a more general graph domain. For the case of sampling, the fact of the nodes of a general graph not having an inherent order – as opposed to what happens in discrete time – raises questions about the advantages of uniform sampling and, consequently, opens the door to the development of new sampling schemes.

This paper investigates a novel scheme for the sampling and posterior recovery of signals defined in the nodes of a graph. The underlying assumption is that such signals admit a sparse representation in a frequency domain related to the structure of this graph. Most existing works have focused on observing the value of the signal at a subset of nodes to recover the signal in the entire graph [5–8]. By contrast, a scheme where the samples are taken at a single node that aggregates local information sequentially was recently proposed in [9–11]. Here, we present a more general sampling scheme that contains both mentioned paradigms as particular cases. More specifically, we consider samples taken at several nodes in the graph which, in turn, aggregate signals in their neighborhood.

Section 2 introduces notation and basic concepts from graph signal processing that will be used throughout the paper. Section 3 describes the problem of sampling graph signals, along with the two existing sampling schemes: selection and aggregation sampling. Section 4 presents the *space-shift sampling* scheme, which is the main contribution of this paper, and discusses conditions for recovery. A simple algorithm to recover the signal in the presence of noise and when the (sparse) frequency support is unknown is also proposed. Section 5 describes numerical experiments assessing the recovery success of space-shift sampling in a graph of the U.S. economy and confirming that the proposed scheme achieves smaller reconstruction errors than aggregation and selection sampling.

Work in this paper is supported by USA NSF CCF-1217963 and Spanish MINECO TEC2013-41604-R.

2. PRELIMINARIES

Let $\mathcal{G} = (\mathcal{N}, \mathcal{E})$ denote a directed graph. The set of nodes or vertices $\mathcal{N}$ 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, we define the adjacency matrix $\mathbf{A}$ as a sparse $N \times N$ matrix with non-zero elements A_{ij} if and only if $(j,i) \in \mathcal{E}$. The value of A_{ij} 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.

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 non-zero 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 non-zero entries of $\mathbf{S}$. Common choices for $\mathbf{S}$ are the adjacency [3, 12] and the Laplacian [2] matrices. An 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, so that there exists a $N \times N$ matrix $\mathbf{V}$ and a $N \times N$ diagonal matrix $\mathbf{\Lambda}$ that can be used to decompose the shift as $\mathbf{S} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^{-1}$.

3. SAMPLING OF BANDLIMITED GRAPH SIGNALS

Recovery of the graph signal $\mathbf{x}$ from a sampled version is possible under the assumption that $\mathbf{x}$ 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. To facilitate exposition, it will be assumed throughout that the active frequencies are the first K ones, i.e., $\mathbf{x}$ is K-bandlimited, which are associated with the largest eigenvalues [5, 13], so that $\widehat{\mathbf{x}} = [\widehat{x}_1, ..., \widehat{x}_K, 0, ..., 0]^T$. However, our results can be applied to any set of active frequencies $\mathcal{K}$ of size K provided that $\mathcal{K}$ is known. For convenience, denote by $\mathbf{e}_i$ the i-th $N \times 1$ canonical basis vector (all entries of $\mathbf{e}_i$ are zero except the i-th one, which is one). Further, define $\mathbf{E}_K := [\mathbf{e}_1, ..., \mathbf{e}_K]$, $\mathbf{V}_K := \mathbf{V}\mathbf{E}_K$ and $\widehat{\mathbf{x}}_K := [\widehat{x}_1, ..., \widehat{x}_K]^T$.

3.1. Selection sampling of bandlimited graph signals

Let $\mathbf{C}$ denote a fat $K \times N$ *selection* matrix whose elements satisfy: $C_{ij} \in \{0, 1\}$, $\sum_j C_{ij} = 1$ for all i, and $\sum_i C_{ij} \leq 1$ for all j. Under the selection sampling approach [5–8], sampling a graph signal amounts to setting $\bar{\mathbf{x}} = \mathbf{C}\mathbf{x}$. Since the $K \times N$ binary matrix $\mathbf{C}$ selects the observed nodes, 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 this, it is assumed that $\mathbf{x}$ 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 $\bar{\mathbf{x}} = \mathbf{C}\mathbf{x} = \mathbf{C}\mathbf{V}_K\widehat{\mathbf{x}}_K$. Hence, if matrix $\mathbf{C}\mathbf{V}_K$ is invertible, $\widehat{\mathbf{x}}_K$ can be recovered from $\bar{\mathbf{x}}$ and, thus, the original signal $\mathbf{x}$ is obtained as

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

Perfect signal reconstruction can be guaranteed by selecting a subset of K nodes such that the corresponding rows in $\mathbf{V}_K$ are linearly independent. In the classical domain of time-varying signals, $\mathbf{V}_K$ has a row-wise *Vandermonde* structure since it is a submatrix of a Fourier matrix, which implies that any subset of K rows is invertible. However, for an arbitrary graph this is not guaranteed and algorithms to select a specific subset that guarantees recovery are required [6].

3.2. Aggregation sampling of bandlimited graph signals

An alternative procedure is to pick a node, say i, and sample the values observed by this node as the shift operator $\mathbf{S}$ is applied sequentially. Formally, define the l-th shifted signal $\mathbf{y}^{(l)} := \mathbf{S}^l\mathbf{x}$ and the $N \times N$ matrix

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

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

Define the $N \times N$ Vandermonde matrix $\mathbf{\Psi}$ with entries $\Psi_{ij} = \lambda_j^{i-1}$, the vector $\boldsymbol{\upsilon}_i := \mathbf{V}^T\mathbf{e}_i$, and the $N \times K$ matrix $\mathbf{\Psi}_i = \mathbf{\Psi}\text{diag}(\boldsymbol{\upsilon}_i)\mathbf{E}_K$. If matrix $\mathbf{C}\mathbf{\Psi}_i$ is invertible, then $\mathbf{x}$ can be recovered from $\bar{\mathbf{y}}_i$ as [10]

$$\mathbf{x} = \mathbf{V}_K(\mathbf{C}\mathbf{\Psi}_i)^{-1}\bar{\mathbf{y}}_i. \tag{3}$$

While for the selection sampling described in Section 3.1 there is no straightforward way to check the invertibility of $\mathbf{C}\mathbf{V}_K$ (existing algorithms typically do that by inspection [6]), for the aggregation sampling described in (3) the invertibility of $\mathbf{C}\mathbf{\Psi}_i$ is guaranteed by the fulfillment of two simple conditions on the eigendecomposition of $\mathbf{S}$ [10].

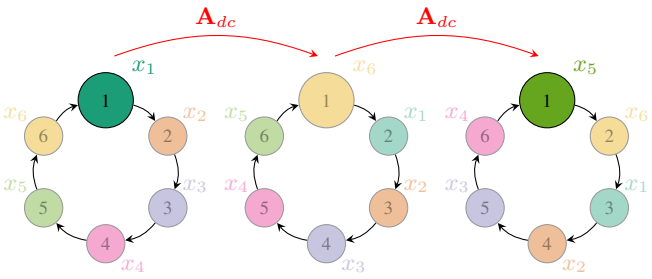


Fig. 1: Conventional sampling in the time domain as aggregation sampling in $\mathcal{G}_{dc}$. We apply the shift operator $\mathbf{S}$ successively and sample the resulting signal observed at a given node (here, node 1).

3.3. Aggregation and selection sampling of time signals

To illustrate the difference between selection and aggregation sampling, let us consider their application to a signal defined in the time domain. Classical time domain signals can be represented as graph signals defined on a directed cycle [2, 7]; see Fig. 1. With $\mathbf{A}_{dc}$ denoting the adjacency matrix of the directed cycle graph, define the shift operator $\mathbf{S} = \mathbf{A}_{dc}$ and the uniform selection matrix $\mathbf{C} = [\mathbf{e}_1, \mathbf{e}_{N/K+1}, \ldots, \mathbf{e}_{N-N/K+1}]^T$ where, we recall, $\mathbf{e}_i$ the i-th canonical basis vector. In selection sampling, the sampled signal is obtained as $\bar{\mathbf{x}} = \mathbf{C}\mathbf{x}$. By contrast, aggregation sampling applies $\mathbf{S} = \mathbf{A}_{dc}$ to $\mathbf{x}$ sequentially. Each of these applications amounts to rotating the signal clockwise. It follows that the aggregated signal $\mathbf{y}_1$ in (2) is given by $\mathbf{y}_1 = [x_1, x_N, x_{N-1}, \ldots, x_2]$, which upon multiplication by $\mathbf{C}$ results in a vector $\bar{\mathbf{y}}_1 = \mathbf{C}\mathbf{y}_1$ that contains the same elements that $\bar{\mathbf{x}}$ contains. Hence, when $\mathbf{S} = \mathbf{A}_{dc}$ both methods can be viewed as generalizations of conventional sampling. However, for more general topologies, the outputs generated are different. In selection sampling one moves through nodes to collect samples at points uniquely identified by $\mathbf{C}$, whereas in aggregation sampling the signal is moved through the graph while samples are collected at a fixed node.

4. SPACE-SHIFT SAMPLING

In selection sampling, a subset of the values in the graph or *space* domain are chosen, corresponding to the *first* column of $\mathbf{Y}$ [cf. (2)]. By contrast, in aggregation sampling, values obtained at a fixed node i are sampled when applying successive *shifts*, corresponding to the i-th row of $\mathbf{Y}$. In the more general case of *space-shift* sampling, we vary the sampling node *and* the number of shifts applied to the signal. Equivalently, we sample matrix $\mathbf{Y}$ without restricting ourselves to a particular row or column. To do so, we first define the vectorized version of $\mathbf{Y}$ as $\boldsymbol{\gamma} := \text{vec}(\mathbf{Y}^T)$. Furthermore, define $\bar{\boldsymbol{\upsilon}}_i := \mathbf{E}_K^T\boldsymbol{\upsilon}_i$ for all i and the $NK \times K$ matrix $\bar{\mathbf{\Upsilon}} =: [\text{diag}(\bar{\boldsymbol{\upsilon}}_1), \ldots, \text{diag}(\bar{\boldsymbol{\upsilon}}_N)]^T$.

Lemma 1 *The relation between the active frequency coefficients* $\widehat{\mathbf{x}}_K$ *and* $\boldsymbol{\gamma}$ *is given by*

$$\boldsymbol{\gamma} = \Big(\mathbf{I} \otimes (\mathbf{\Psi}\mathbf{E}_K)\Big)\bar{\mathbf{\Upsilon}}\widehat{\mathbf{x}}_K, \tag{4}$$

where $\otimes$ is the matrix Kronecker product.

Proof: From (3), we have $\mathbf{y}_i = \mathbf{\Psi}_i \widehat{\mathbf{x}}_K = \mathbf{\Psi}\mathbf{E}_K \text{diag}(\bar{\boldsymbol{\upsilon}}_i)\widehat{\mathbf{x}}_K$, for all i. The result follows from the definition of the Kronecker product $\otimes$ once we note that $\boldsymbol{\gamma} = [\mathbf{y}_1^T, \mathbf{y}_2^T, \ldots, \mathbf{y}_N^T]^T$. ■

To solve for $\widehat{\mathbf{x}}_K$ – and, hence, $\mathbf{x}$ – we must select at least $K < N$ linearly independent equations out of the N^2 stated in (4). Suppose that for a given node i we consider the problem of selecting K equations out of the N equations in positions $\{(i-1)N + n\}_{n=1}^N$, then space-shift sampling reduces to local aggregation sampling at node i. Similarly, if we restrict ourselves to select K equations out of the N equations in positions $\{1 + (n-1)N\}_{n=1}^N$, the problem reduces to selection sampling. In this sense, space-shift sampling is more general. To implement the selection of the K equations, we use a binary selection matrix $\mathbf{C}$ as done in previous sections but, in this case, the size of $\mathbf{C}$ is $K \times N^2$. If $\mathbf{C}(\mathbf{I} \otimes (\mathbf{\Psi}\mathbf{E}_K))\bar{\mathbf{\Upsilon}}$ is invertible, then $\mathbf{x}$ can be recovered from $\bar{\boldsymbol{\gamma}} := \mathbf{C}\boldsymbol{\gamma}$ as [cf. (4)]

$$\mathbf{x} = \mathbf{V}_K \Big(\mathbf{C}\big(\mathbf{I} \otimes (\mathbf{\Psi}\mathbf{E}_K)\big)\bar{\mathbf{\Upsilon}}\Big)^{-1} \bar{\boldsymbol{\gamma}}. \tag{5}$$

When $\mathbf{C}(\mathbf{I} \otimes (\mathbf{\Psi}\mathbf{E}_K))\bar{\mathbf{\Upsilon}}$ is not invertible, additional samples (rows of $\mathbf{C}$) are required. The aim is to have at least K linearly independent equations, so that the matrix $\mathbf{C}(\mathbf{I} \otimes (\mathbf{\Psi}\mathbf{E}_K))\bar{\mathbf{\Upsilon}}$ has full column rank and the original signal $\mathbf{x}$ can be recovered using the pseudoinverse. Beyond invertibility, taking additional samples improves the recovery performance in the presence of noise [9]. Although space limitations prevent us to present the details here, the structure in (5) can be used to design optimal sampling and recovery schemes that minimize the effects of the noise. Such optimal designs depend on the error metric to be optimized, with some of them leading to NP-hard formulations, hence calling for tractable approximations [14] that leverage the particularities of our graph setup. See, e.g., [11] for a simpler but related problem.

A sampling setup of particular interest is when the sampling schemes are implemented in a distributed manner using message passing. Suppose that the sampling is performed at node $i = 1$ with neighbors of $i = 1$ being $i = 2, \ldots, N_1 + 1$. To compute $y_i^{(l)}$, node i needs to have access to $y_j^{(l')}$ for all $j \in \mathcal{N}_i$ and $l' < l$. Suppose also that node $i = 1$ computes L_1 shifts, from $y_1^{(0)}$ up to $y_1^{(L_1)}$. This implies that node $i = 1$ has access to $L_1 + 1$ of its own samples and to L_1 samples of each of its N_1 neighbors, thus, node i can apply the space-shift reconstruction in (5). Note however that matrix $\mathbf{C}\left(\mathbf{I} \otimes (\mathbf{\Psi}\mathbf{E}_K)\right)\bar{\mathbf{\Upsilon}}$ is not full (row) rank. The reason is that all the samples obtained at node $i = 1$, except for the first one, are linear combinations of the samples at its neighbors. This implies that the number of frequencies that can be recovered is, at most, $1 + L_1 N_1$.

Alternative structured observation models can also be of interest. For example, one can consider setups where nodes from different parts of the graph take a few samples each and forward those samples to a central fusion center. In such a case, since the nodes gathering data need not be neighbors, the problem of some of the samples being a linear combination of the others will not necessarily be present.

Remark 1 Although the space (graph) and shift dimensions coincide for time-varying signals – Section 3.3 –, the fact of these two dimensions not being the same can be leveraged when designing algorithms for graph signals. For the particular case of sampling, taking samples in both dimensions simultaneously yields a more general scheme with additional degrees of freedom that can used to achieve better reconstruction performance; as demonstrated in Section 5.

4.1. Joint recovery and support identification

Thus far, it has been assumed that the frequency support of $\mathbf{x}$ corresponded to the K principal eigenvectors. A related but more challenging problem is to design the sampling and interpolation procedures when the frequency support $\mathcal{K}$ is not known, qualifying as sparse signal reconstruction [15–17]. Based on (4) and defining $\mathbf{\Upsilon} =: [\text{diag}(\boldsymbol{\upsilon}_1), \ldots, \text{diag}(\boldsymbol{\upsilon}_N)]^T$, we may reformulate the recovery problem as

$$\begin{aligned} \widehat{\mathbf{x}}^* := \arg\min_{\widehat{\mathbf{x}}} \quad & ||\widehat{\mathbf{x}}||_0 \\ \text{s.t.} \quad & \mathbf{C}\boldsymbol{\gamma} = \mathbf{C}\big(\mathbf{I} \otimes \mathbf{\Psi}\big)\mathbf{\Upsilon}\widehat{\mathbf{x}}, \end{aligned} \tag{6}$$

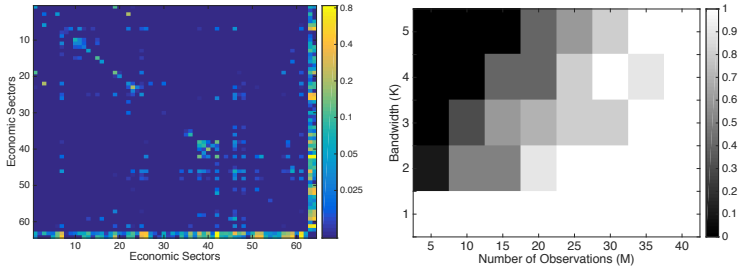
where the sparse support of $\widehat{\mathbf{x}}$ is unknown. For this case, there is no guarantee that the solution $\widehat{\mathbf{x}}^*$ coincides with the K-sparse representation of the observed signal. Nevertheless, identifiability is guaranteed whenever $\mathbf{C}(\mathbf{I} \otimes \mathbf{\Psi})\mathbf{\Upsilon}$ is full spark and contains at least $2K$ rows [15]. For specific forms of $\mathbf{C}$, the full-spark condition can be assessed by a quick inspection of $\{\lambda_i\}_{i=1}^N$ and $\mathbf{V}$; see [9, 10] for the case of aggregation sampling.

From a computational perspective, the presence of the 0-norm in (6) renders the optimization non-convex, thus challenging to solve. A straightforward way to convexify the 0-norm is to replace it with the 1-norm. Conditions under which this process is guaranteed to recover the frequency support can be found by analyzing the coherence and the restricted isometry property (RIP) of matrix $\mathbf{C}(\mathbf{I} \otimes \mathbf{\Psi})\mathbf{\Upsilon}$ [15, 18]. Unfortunately, determining the conditioning of all submatrices of a deterministic matrix (and, hence, the RIP) is challenging [19]. Finally, if the observations are noisy, the constraint in problem (6) can be replaced by $\|\mathbf{C}\boldsymbol{\gamma} - \mathbf{C}\big(\mathbf{I} \otimes \mathbf{\Psi}\big)\mathbf{\Upsilon}\widehat{\mathbf{x}}\|_2^2 < \epsilon$, where ϵ is a maximum level of expected (or tolerable) noise.

5. NUMERICAL EXPERIMENTS

The U.S. Department of Commerce publishes a yearly table of inputs and outputs organized by economic sectors [20]. More precisely, we have a set $\mathcal{N}$ of 62 industrial sectors and a 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 synthetic 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 obtain $\mathbf{S}$ from $\bar{\mathbf{S}}$ by setting to 0 all the values lower than 0.01; see Fig. 2(a).

Recovery and support identification. In order to assess recoverability when the frequency support of the signal is unknown (cf. Section 4.1) we generate synthetic signals of varying bandwidth $K \in \{1, 2, \ldots, 5\}$ on the economic network. Further, we consider different number of observations M, i.e. rows of $\mathbf{C}$, where $M \in \{5, 10, \ldots, 40\}$. The M observations are randomly chosen from a pool of 62×2 possibilities given by the original signal values and the values after one shift in the 62 real sectors (excluding AV and FU). We assume that K is known, but not the location of the K active frequencies. Thus, the recovery algorithm is the following. We first



(a) (b)

Sampling strategy				Error	
				Min.	Median
$[\mathbf{x}]_i$	$[\mathbf{x}]_j$	$[\mathbf{x}]_k$	$[\mathbf{x}]_l$	.0039	4.2
$[\mathbf{x}]_i$	$[\mathbf{Sx}]_i$	$[\mathbf{S^2x}]_i$	$[\mathbf{S^3x}]_i$	.0035	.019
$[\mathbf{Sx}]_i$	$[\mathbf{Sx}]_j$	$[\mathbf{Sx}]_k$	$[\mathbf{Sx}]_l$	.0035	.030
$[\mathbf{S^2x}]_i$	$[\mathbf{S^2x}]_j$	$[\mathbf{S^2x}]_k$	$[\mathbf{S^2x}]_l$	.0035	.0055
$[\mathbf{S^3x}]_i$	$[\mathbf{S^3x}]_j$	$[\mathbf{S^3x}]_k$	$[\mathbf{S^3x}]_l$	.0035	.0040
$[\mathbf{x}]_i$	$[\mathbf{Sx}]_i$	$[\mathbf{x}]_j$	$[\mathbf{Sx}]_j$	.0035	.039

(c)

Fig. 2: (a) Heat map of the graph-shift operator $\mathbf{S}$ of the economic network. The shift $\mathbf{S}$ is sparse across the real economic sectors (nodes from $n = 1$ to $n = 62$ are weakly connected), while the synthetic sectors AV and FU (nodes $n = 63$ and $n = 64$) are highly connected. (b) Successful rate of recovery as a function of the bandwidth K and the number of observations M of a bandlimited signal of unknown frequency support defined on the economic network. (c) Minimum and median reconstruction error for different sampling strategies. The first sampling strategy corresponds to selection sampling (cf. Section 3.1), the second one to aggregation sampling (cf. Section 3.2), and the remaining strategies correspond to more general space-shift sampling schemes (cf. Section 4).

solve a convex version of (6) where the 0-norm is replaced by the 1-norm. If the solution is K-sparse, we stop. Otherwise, we solve (6) again with the additional constraint that a randomly chosen non-zero element of the previous solution is set it to zero. We repeat this procedure, adding an additional constraint at each step, until we obtain a K-sparse solution or the problem becomes unfeasible. In Fig. 2(b) we present the success of the described procedure averaged over 10 realizations for every combination of K and M. The behavior is as expected. In the top-left corner of the figure, the recovery is null since the number of observations is low and the degrees of freedom (bandwidth) of the signal are relatively high. The opposite is true in the bottom-right corner while, for intermediate configurations, the success rate varies gradually.

Comparing sampling schemes. Associated with the economic graph, we consider the signal $\mathbf{x} \in \mathbb{R}^{64}$ that collects the total production – in trillion of dollars – of each sector during year 2011. Signal $\mathbf{x}$ is approximately bandlimited in $\mathbf{S}$ since the reconstructed signal $\mathbf{x}_4 = \mathbf{V}_4\widehat{\mathbf{x}}_4$ obtained by just keeping the first $K = 4$ frequency coefficients attains a reconstruction error of $\|\mathbf{x} - \mathbf{x}_4\|_2^2/\|\mathbf{x}\|_2^2 = .0035$.

The table in Fig. 2(c) lists the reconstruction errors attained by different sampling schemes when taking four samples. The first row shows the performance achieved by selection sampling (SS) strategies (cf. Section 3.1), i.e., by observing the value of $\mathbf{x}$ at 4 different nodes i, j, k, l. The second row shows the performance achieved by aggregation sampling (AS) strategies (cf. Section 3.2), i.e., we pick one node, say i, and take samples after successive applications of 0, 1, 2, and 3 graph-shifts $\mathbf{S}$. The remaining rows show the performance achieved by different configurations of our space-shift sampling (SSS) schemes. For every row, the table lists the minimum and median errors achieved by the particular strategy, which are computed across all possible n choose m configurations of sampling nodes, where $n = 62$ (we exclude the two synthetic sectors) and m depends on the sampling scheme. For instance, for AS we have that $m = 1$, since there is a single sampling node. Hence, the table shows that among the $\binom{n}{m} = 62$ single-node configurations, the smallest error is .0035 while the median error is .019.

First notice that the minimum error achieved by all the strategies that shift the original signal (AS and all SSS configurations) is the minimum one. This means that all those schemes are able to recover $\mathbf{x}_4$. This is not true for SS (first row), although the difference is not large. However, the median error achieved by SS is two orders of magnitude larger than that achieved by any of the other strategies. This problem disappears when more general SSS strategies are considered. For example, one can sample the value of the signal at 4 nodes after the application of 1, 2 or 3 graph shifts. The results (rows 3, 4 and 5) reveal that reduction in the median error after each graph shift application is conspicuous, especially when going from none to one application – the median error goes from 4.2 to 0.03. Another alternative is to sample both $\mathbf{x}$ and $\mathbf{Sx}$ in two different sectors (bottom row in the table). Note that with this sampling configuration, the two sectors are only required to obtain the aggregated activity of their one-hop neighbors.

The performance attained by a specific sampling scheme depends on factors like the operating conditions of the network, the structure of the graph, and the properties of the signal. As a general rule, when sampling an approximately bandlimited signal whose active frequencies are associated with large eigenvalues of $\mathbf{S}$ – which is a standard assumption in the literature [5–8] –, AS is expected to give rise to a better recovery. Successive applications of $\mathbf{S}$ amplify the active frequencies, entailing a better estimation of these frequencies and reducing the reconstruction error. By contrast, when the active frequencies are associated with small eigenvalues of $\mathbf{S}$, SS is preferred. SSS strategies are useful whenever some active frequencies are related to large eigenvalues and others are related to small eigenvalues. Moreover, SSS is also a suitable alternative when the magnitudes of the eigenvalues associated with the active frequencies are unknown.

6. CONCLUSIONS

A novel scheme for sampling graph signals that admit a sparse frequency representation was proposed. The scheme was based on the *local aggregation* of data at *several nodes* across the graph. We stated conditions for perfect reconstruction when the signal frequency support is known and proposed a recovery algorithm for the cases when it is unknown. Since existing sampling methods can be recovered as particular cases of the proposed space-shift sampling scheme, the latter exhibits a better recovery performance, as confirmed by numerical experiments.

7. REFERENCES

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

[2] D. Shuman, S. 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, Mar. 2013.

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

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

[5] 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.

[6] I. Shomorony and A. Avestimehr, "Sampling large data on graphs," *CoRR*, vol. abs/1411.3017, 2014. [Online]. Available: http://arxiv.org/abs/1411.3017

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

[8] X. Wang, P. Liu, and Y. Gu, "Local-set-based graph signal reconstruction," *arXiv preprint arXiv:1410.3944*, 2014.

[9] A. G. Marques, S. Segarra, G. Leus, and A. Ribeiro, "Sampling of graph signals with successive local aggregations," *arXiv preprint arXiv:1504.04687*, 2015.

[10] S. Segarra, A. G. Marques, G. Leus, and A. Ribeiro, "Sampling of graph signals: successive local aggregations at a single node," in *Asilomar Conf. on Signals, Systems, and Computers*, Pacific Grove, CA, Nov. 8-11, 2015.

[11] ——, "Aggregation sampling of graph signals in the presence of noise," in *Workshop on Comput. Adv. in Multi-Sensor Adaptive Process. (CAMSAP)*, Cancun, Mexico, Dec. 13-16, 2015.

[12] 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.

[13] 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.

[14] F. Pukelsheim, *Optimal Design of Experiments*. SIAM, 1993, vol. 50.

[15] D. L. Donoho and M. Elad, "Optimally sparse representation in general (nonorthogonal) dictionaries via ℓ^1 minimization," *Proc. Nat. Academy of Sciences*, vol. 100, no. 5, pp. 2197–2202, 2003.

[16] M. Elad, "Optimized projections for compressed sensing," *IEEE Trans. Signal Process.*, vol. 55, no. 12, pp. 5695–5702, Dec 2007.

[17] X. Zhu and M. Rabbat, "Graph spectral compressed sensing for sensor networks," in *IEEE Intl. Conf. Acoust., Speech and Signal Process. (ICASSP)*, March 2012, pp. 2865–2868.

[18] E. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. Inf. Theory*, vol. 52, no. 2, pp. 489–509, Feb 2006.

[19] B. Alexeev, J. Cahill, and D. G. Mixon, "Full spark frames," *J. of Fourier Analysis and Appl.*, vol. 18, no. 6, pp. 1167–1194, 2012.

[20] 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