GRAPH SAMPLING WITH AND WITHOUT INPUT PRIORS

Sundeep Prabhakar Chepuri[†], *Yonina C. Eldar*[‡], *and Geert Leus*[†]

[†]Delft University of Technology (TU Delft), The Netherlands
[‡]Technion - Israel Institute of Technology, Israel

ABSTRACT

In this paper the focus is on sampling and reconstruction of signals supported on nodes of arbitrary graphs or arbitrary signals that may be represented using graphs, where we extend concepts from generalized sampling theory to the graph setting. To recover such signals from a given set of samples, we develop algorithms that incorporate prior knowledge on the original signal when available such as smoothness or subspace priors related to the underlying graph. For reconstructing arbitrary signals, we constrain the reconstruction to the graph, and provide a consistent reconstruction method, in which both the reconstructed signal and the input yield exactly the same measurements. Given a set of graph frequency domain samples, the sampling and interpolation operations may be efficiently implemented using linear shift-invariant graph filters.

Index Terms— Graph sampling, graph signal processing, consistent reconstruction, subspace prior, frequency domain sampling.

1. INTRODUCTION

Graph signal processing is an emerging research area that is recently gaining a lot interest as it can be used to process datasets supported on irregular domains. Such complex datasets often appear in modern data analysis, e.g., brain networks [1], transportation networks [2], socio-economic networks [3], to list a few. Similar to how filtering, sampling, and compression play a fundamental role in traditional signal processing, these operations also form the basic building blocks of graph signal processing [4, 5].

Sampling and recovery of signals on graphs have been studied in the context of graph signal processing under the assumption that the graph signal is bandlimited; see [6–12]. The sampling functions that have been mostly considered in [6, 7] are node selection operators, which are typically designed using sparse sensing techniques [13–15]. Instead of observing a subset of nodes, aggregation sampling is suggested in [16], in which the graph signal is observed at a single node and the observations correspond to linear combinations of the information gathered by the neighbors of that node. When only the second-order statistics of the graph signals are to be recovered from the samples, the bandlimited assumption may be relaxed [17].

In this paper, the aim is to extend some of the concepts from *generalized sampling theory* [18–21] to the graph setting. A signal $\boldsymbol{x} \in \mathbb{C}^N$ is sampled by taking inner products with a set of vectors that span a subspace $\mathcal{S}$ in $\mathbb{C}^N$, which is referred to as the *sampling space*.

To reconstruct $\boldsymbol{x}$ from these samples we use knowledge on the original signal that it lies on a graph and translate this fact into a prior for sampling, such as smoothness or that the signal is known to lie in a subspace. To begin with, we focus on a least squares recovery method that recovers $\boldsymbol{x}$ from the acquired samples under the assumption that $\boldsymbol{x}$ lies in a subspace $\mathcal{A}$ (e.g., a subspace of the graph Laplacian). The subspace $\mathcal{A}$ is referred to as the *input space*. For this case, we show that existing works (e.g., [6, 7, 16]) specialize to the generalized sampling formulation used in this paper. In many cases, a subspace prior (i.e., information about $\mathcal{A}$) might not be available. Instead we might know that the signal is smooth with respect to the underlying graph, where the extent of smoothness is quantified by the graph Laplacian quadratic form. In this setting, the focus will be to develop a linear estimator that reconstructs a smooth signal from the given samples.

To recover arbitrary input signals, not necessarily supported on the graph, we will seek a reconstruction, call it $\hat{\boldsymbol{x}}$, which is *consistent* in the sense that $\hat{\boldsymbol{x}}$ might not be equal to $\boldsymbol{x}$, but both $\boldsymbol{x}$ and $\hat{\boldsymbol{x}}$ produce the same measurements. To do so, we force the reconstruction to be linear combinations of a set of vectors that span a subspace $\mathcal{R}$ (e.g., a subspace of a graph Laplacian). The subspace $\mathcal{R}$ is referred to as the *reconstruction space*. If the input signal also lies in $\mathcal{R}$ (e.g., $\mathcal{A} \subseteq \mathcal{R}$), then the reconstruction will be exact, thus including the least squares recovery with the subspace prior as a special case. The concept of consistent recovery is reasonable when we do not have prior information about the underlying signal and we are interested in a data representation that takes into account the hidden geometric structure as it allows us to consider different reconstruction subspaces $\mathcal{R}$ that fits for our needs while being consistent.

The contributions of this paper are (a) to extend the generalized sampling theory with and without input priors (e.g., subspace or smoothness) to the graph setting and (b) the interpretation of the sampling and reconstruction operations as linear shift-invariant graph filters, for which we introduce a *sparse sampler* in the graph spectral domain.

2. GRAPH SIGNALS

Consider an undirected graph $\mathcal{G} = \{\mathcal{V}, \mathcal{E}\}$, which consists of a finite set of vertices $\mathcal{V}$ with $|\mathcal{V}| = N$ and a set of edges $\mathcal{E}$. If there is an edge connecting vertices i and j, then $(i, j) \in \mathcal{E}$. A signal or function $x : \mathcal{V} \rightarrow \mathbb{C}$ defined on the vertices of the graph can be collected in a length-N vector $\boldsymbol{x} = [x_1, x_2, \dots, x_N]^T$, where the nth element of $\boldsymbol{x}$ represents the function value at the nth vertex in $\mathcal{V}$. Since $\boldsymbol{x}$ resides on the graph, we refer to the function $\boldsymbol{x}$ as a *graph signal*.

2.1. Graph filtering

Let us introduce an operator $\boldsymbol{G} \in \mathbb{C}^{N \times N}$, where the (i, j)th entry of $\boldsymbol{G}$ denoted by $g_{i,j}$ can only be nonzero if $(i, j) \in \mathcal{E}$ or $i = j$. The pattern of $\boldsymbol{G}$ captures the structure of the graph and for a graph

This work is part of the ASPIRE project (project 14926 within the STW OTP program), which is financed by the Netherlands Organization for Scientific Research (NWO).

signal $\boldsymbol{x}$, the signal $\boldsymbol{Gx}$ denotes the unit shifted version of $\boldsymbol{x}$. Thus the operator $\boldsymbol{G}$ is commonly referred to as the *graph-shift* operator [22]. Typical choices for $\boldsymbol{G}$ include the adjacency matrix [22], the graph Laplacian [4], or their respective variants. For undirected graphs, $\boldsymbol{G}$ is Hermitian, and thus it admits the following eigenvalue decomposition

$$\boldsymbol{G} = \boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{U}^H = [\boldsymbol{u}_1, \cdots, \boldsymbol{u}_N]\,\mathrm{diag}[\lambda_1, \cdots, \lambda_N]\,[\boldsymbol{u}_1, \cdots, \boldsymbol{u}_N]^H,$$

where the eigenvectors $\{\boldsymbol{u}_n\}_{n=1}^N$ and the eigenvalues $\{\lambda_n\}_{n=1}^N$ of $\boldsymbol{G}$ provide the notion of frequency in the graph setting [4,5]. Specifically, $\{\boldsymbol{u}_n\}_{n=1}^N$ forms an orthonormal Fourier-like basis for graph signals with the graph frequencies denoted by $\{\lambda_n\}_{n=1}^N$. The *graph Fourier transform* of $\boldsymbol{x}$, denoted by $\boldsymbol{x}_f$, is given by

$$\boldsymbol{x}_f = \boldsymbol{U}^H\boldsymbol{x} \Leftrightarrow \boldsymbol{x} = \boldsymbol{U}\boldsymbol{x}_f. \qquad (1)$$

The frequency content of graph signals may be modified using *linear shift-invariant graph filters* [22], denoted by $\boldsymbol{H} \in \mathbb{C}^{N\times N}$. Such a shift-invariant graph filter can be expressed as a polynomial in $\boldsymbol{G}$ [22]:

$$\boldsymbol{H} = \sum_{l=0}^{L-1} h_l \boldsymbol{G}^l = \boldsymbol{U}\boldsymbol{H}_f\boldsymbol{U}^H, \qquad (2)$$

where the filter $\boldsymbol{H}$ is of degree $L-1$ with filter coefficients $\{h_l\}_{l=0}^{L-1}$ and the diagonal matrix $\boldsymbol{H}_f = \sum_{l=0}^{L-1} h_l \boldsymbol{\Lambda}^l$ is the frequency response of the graph filter. Here, $L \leq N$ as $N-1$ is the maximum degree of any polynomial of $\boldsymbol{G}$ [22].

2.2. Representation subspace

The class of bandlimited graph signals may be viewed as signals that lie in a subspace. If the signal is bandlimited, then its frequency domain will be sparse. Suppose $\boldsymbol{x}_f$ is sparse with *known* support, and also assume, without loss of generality, that the first P entries are nonzero. Then the bandlimited signal $\boldsymbol{x}$ can be expressed as a linear combination of the first few modes as

$$\boldsymbol{x} = \sum_{i=1}^{P} \boldsymbol{u}_i x_{f,i} = \boldsymbol{U}_P \boldsymbol{x}_{f,P}. \qquad (3)$$

The above model for the graph signal places $\boldsymbol{x}$ in a P-dimensional linear subspace $\mathcal{A}$ known as the *input space* and spanned by the columns of $\boldsymbol{U}_P = [\boldsymbol{u}_1, \cdots, \boldsymbol{u}_P]$, i.e., $\mathcal{A} = \mathrm{range}(\boldsymbol{U}_P)$.

2.3. Laplacian smoothness

Smoothness of a graph signal depends on the underlying graph topology. Typically, the Laplacian quadratic form given by $\boldsymbol{x}^T\boldsymbol{G}\boldsymbol{x}$, where $\boldsymbol{G}$ is the graph Laplacian can be used to quantify the extent of smoothness of $\boldsymbol{x}$ with respect to the underlying graph [4]. We will say that the signal $\boldsymbol{x}$ is smooth with respect to the graph $\boldsymbol{G}$ for low values of $\boldsymbol{x}^T\boldsymbol{G}\boldsymbol{x}$.

3. THE SAMPLING PROBLEM

Suppose we are given samples of $\boldsymbol{x} \in \mathbb{C}^N$ obtained by taking inner products of $\boldsymbol{x}$ with a set of K sampling functions $\{\boldsymbol{s}_k, 1 \leq k \leq K\}$ as

$$y_k = \boldsymbol{s}_k^H \boldsymbol{x}, \quad k = 1, 2, \ldots, K.$$

The sampling functions $\{\boldsymbol{s}_k, 1 \leq k \leq K\}$ span an M-dimensional linear subspace $\mathcal{S}$ in $\mathbb{C}^N$, which is referred to as the *sampling space*. If $K = M$, then the measurements are *nonredundant*, whereas when $K > M$, the measurements are *redundant*. Collecting these samples in a K-dimensional vector $\boldsymbol{y} = [y_1, y_2, \ldots, y_K]^T$ and defining an operator $\boldsymbol{S} \in \mathbb{C}^{N\times K}$ that maps $\boldsymbol{x}$ into $\boldsymbol{y}$, the sampling scheme can be equivalently expressed as

$$\boldsymbol{y} = \boldsymbol{S}^H\boldsymbol{x}. \qquad (4)$$

Given $\boldsymbol{y}$ and $\boldsymbol{S}$, the sampling problem boils down to finding a reconstruction $\hat{\boldsymbol{x}}$ that is close to $\boldsymbol{x}$ is some sense. We wish to address this problem by incorporating our knowledge about the input signal when available. When no input priors are available, we seek an approximation that is consistent in the sense that it produces exactly the same measurements as the input to the sampling function.

4. RECOVERY WITH INPUT PRIORS

In this section, the focus will be on developing reconstruction methods that take into account the available input prior on the signal. We will first discuss the minimum squared error (i.e., least squares) recovery technique from [6, 7, 16], in which the signal is known to lie in a subspace of the graph Laplacian. Then, we will consider a recovery method, in which we assume that the signal is smooth on the known graph. Using simple examples, we show that the sampling and reconstruction operations have an elegant graph filter interpretation.

4.1. Subspace prior

Let us assume that the graph signal $\boldsymbol{x}$ lies in a *known* P-dimensional subspace $\mathcal{A}$ spanned by the columns of an $N \times P$ matrix $\boldsymbol{A}$, but its precise location is unknown. For example, $\mathcal{A}$ could be the subspace of bandlimited graph signals as detailed in Section 2.2.

Since $\boldsymbol{x} \in \mathcal{A}$, it can be decomposed as $\boldsymbol{x} = \boldsymbol{A}\boldsymbol{d}$ for some nonzero length-P vector $\boldsymbol{d}$. Now, we may find an estimate of $\boldsymbol{d}$ by minimizing the squared norm of the residual:

$$\underset{\boldsymbol{d}}{\text{minimize}}\ \|\boldsymbol{y} - \boldsymbol{S}^H\boldsymbol{A}\boldsymbol{d}\|_2^2. \qquad (5)$$

Suppose $K \geq P$ and the matrix $\boldsymbol{S}^H\boldsymbol{A}$ has full column rank, i.e., $\mathcal{A} \cap \mathcal{S}^\perp = \{0\}$. Then the unique reconstruction minimizing (5) is $\hat{\boldsymbol{d}} = (\boldsymbol{S}^H\boldsymbol{A})^\dagger\boldsymbol{y}$, and $\hat{\boldsymbol{x}} = \boldsymbol{A}\hat{\boldsymbol{d}} = \boldsymbol{A}(\boldsymbol{S}^H\boldsymbol{A})^\dagger\boldsymbol{y}$. Substituting (4) we have

$$\hat{\boldsymbol{x}} = \boldsymbol{A}(\boldsymbol{A}^H\boldsymbol{S}\boldsymbol{S}^H\boldsymbol{A})^{-1}\boldsymbol{A}^H\boldsymbol{S}\boldsymbol{S}^H\boldsymbol{x} = \boldsymbol{E}_{\mathcal{A}\mathcal{S}^\perp}\boldsymbol{x}. \qquad (6)$$

In other words, the reconstruction of $\boldsymbol{x}$ is given by its *oblique projection* onto the range $\mathcal{A}$ and along the null space $\mathcal{S}^\perp$, denoted by $\boldsymbol{E}_{\mathcal{A}\mathcal{S}^\perp}$. By assumption, since $\boldsymbol{x} \in \mathcal{A}$, $\hat{\boldsymbol{x}} = \boldsymbol{E}_{\mathcal{A}\mathcal{S}^\perp}\boldsymbol{x} = \boldsymbol{x}$. In the special case, in which $\mathcal{A} = \mathcal{S}$, the oblique projector $\boldsymbol{E}_{\mathcal{A}\mathcal{S}^\perp}$ simplifies to an orthogonal projector $\boldsymbol{P}_{\mathcal{S}} = \boldsymbol{S}(\boldsymbol{S}\boldsymbol{S}^H)^{-1}\boldsymbol{S}^H$.

This approach specializes to the recovery methods proposed in [6,7,16]. For the vertex-domain sample selection approach [6,7], $\boldsymbol{S}^H$ reduces to a selection matrix and the input prior is $\mathcal{A} = \mathrm{range}(\boldsymbol{U}_P)$, for which $\boldsymbol{d}$ corresponds to $\boldsymbol{x}_{f,P}$. In the aggregation sampling [16] approach based on observations gathered at a single node, $\boldsymbol{S}^H$ again corresponds to a selection matrix and the input prior specializes to $\mathcal{A} = \mathrm{range}(\boldsymbol{V}_P)$, where $\boldsymbol{V}_P$ is an $N \times P$ Vandermonde matrix with the (i,j)th entry λ_i^{j-1} (recall that λ_i is the ith eigenvalue of the graph-shift $\boldsymbol{G}$). It should be noted that in the vertex-domain and aggregation sampling approaches the sampling function is not related to the graph, but the reconstruction function depends on the graph.

We next show that the oblique projector $\boldsymbol{E}_{\mathcal{AS}^\perp}$ can be expressed as a graph filter by using a specific sampling function which depends on the graph. Suppose $\boldsymbol{A} = \boldsymbol{U}_P$ with $\boldsymbol{A}\boldsymbol{A}^H = \boldsymbol{H}_P$ being a graph filter. Also, let us choose a sampling function of the form $\boldsymbol{S}^H = \boldsymbol{\Phi}\boldsymbol{U}^H$, where $\boldsymbol{\Phi} = [\boldsymbol{e}_1^T, \dots, \boldsymbol{e}_K^T]^T \in \{0,1\}^{K\times N}$ is the selection matrix with $\boldsymbol{e}_i$ being the ith column of the identity matrix so that $K = M$. This sampling function essentially collects, without loss of generality, the first K *contiguous* frequencies. In fact, $\boldsymbol{S}\boldsymbol{S}^H = \boldsymbol{H}_{\text{samp}}$ is a shift-invariant graph filter with frequency response $\boldsymbol{H}_{f,\text{samp}} = \boldsymbol{\Phi}^T\boldsymbol{\Phi}$.

With this sampling function, it is easy to see that the term $\boldsymbol{A}(\boldsymbol{A}^H\boldsymbol{S}\boldsymbol{S}^H\boldsymbol{A})^{-1}\boldsymbol{A}^H$ in (6) is a graph filter of the form $\boldsymbol{H}_{\text{interp}} = \boldsymbol{U}_P\boldsymbol{H}_{f,\text{interp}}\boldsymbol{U}_P^H$, where the diagonal matrix $\boldsymbol{H}_{f,\text{interp}}$ is related to the portion of the sampling filter that overlaps with $\mathcal{A}$ and is given by $\boldsymbol{H}_{f,\text{interp}}^{-1} = \boldsymbol{U}_P^H\boldsymbol{H}_{\text{samp}}\boldsymbol{U}_P$. The oblique projector is then simply given by $\boldsymbol{E}_{\mathcal{AS}^\perp} = \boldsymbol{H}_{\text{interp}}\boldsymbol{H}_{\text{samp}}$. This means that

$$\hat{\boldsymbol{x}} = \boldsymbol{H}_{\text{interp}}\boldsymbol{H}_{\text{samp}}\boldsymbol{x}.$$

4.2. Smoothness prior

We now shift our focus to smoothness priors (cf. Section 2.3), where we assume that the signal $\boldsymbol{x}$ is smooth with respect to the underlying graph. With only the smoothness prior, we cannot achieve a perfect reconstruction, in general. In such cases, we can only obtain a consistent reconstruction. Therefore, we seek a reconstruction $\hat{\boldsymbol{x}}$ that is consistent in the sense that $\hat{\boldsymbol{x}}$ produces exactly the same measurements as $\boldsymbol{x}$ when "reapplied" as an input to the sampling function. In other words, the reconstruction $\hat{\boldsymbol{x}}$ should satisfy:

$$\boldsymbol{S}^H\hat{\boldsymbol{x}} = \boldsymbol{S}^H\boldsymbol{x}. \tag{7}$$

The reconstruction problem that takes into account the smoothness prior then becomes

$$\underset{\boldsymbol{x}}{\text{minimize}} \;\; \frac{1}{2}\boldsymbol{x}^H\boldsymbol{G}\boldsymbol{x} \quad \text{subject to} \quad \boldsymbol{S}^H\boldsymbol{x} = \boldsymbol{y} \tag{8}$$

which is an equality constrained quadratic program. The graph Laplacian matrix $\boldsymbol{G}$ is positive semidefinite.

The solution to the above problem can be computed in closed form using Lagrange multipliers. The Lagrangian of (8) is given by

$$J(\boldsymbol{x},\boldsymbol{\lambda}) = \boldsymbol{x}^H\boldsymbol{G}\boldsymbol{x} + \boldsymbol{\lambda}^T(\boldsymbol{S}^H\boldsymbol{x} - \boldsymbol{y})$$

where $\boldsymbol{\lambda} \in \mathbb{R}^K$ is a Lagrange multiplier corresponding to the measurement equation. Setting the derivative of $J(\boldsymbol{x},\boldsymbol{\lambda})$ with respect to $\boldsymbol{x}$ and $\boldsymbol{\lambda}$ to zero, we get

$$\frac{\partial J(\boldsymbol{x},\boldsymbol{\lambda})}{\partial \boldsymbol{x}} = \boldsymbol{G}\boldsymbol{x} - \boldsymbol{S}\boldsymbol{\lambda} = \boldsymbol{0}; \quad \frac{\partial J(\boldsymbol{x},\boldsymbol{\lambda})}{\partial \boldsymbol{\lambda}} = \boldsymbol{S}^H\boldsymbol{x} - \boldsymbol{y} = \boldsymbol{0}.$$

To handle cases with a singular $\boldsymbol{G}$, the above system is rewritten as

$$\begin{bmatrix} \boldsymbol{G} + \boldsymbol{S}\boldsymbol{S}^H & \boldsymbol{S} \\ \boldsymbol{S}^H & \boldsymbol{0} \end{bmatrix} \begin{bmatrix} \boldsymbol{x} \\ \boldsymbol{\lambda} \end{bmatrix} = \begin{bmatrix} \boldsymbol{S}\boldsymbol{y} \\ \boldsymbol{y} \end{bmatrix}. \tag{9}$$

Suppose the coefficient matrix in (9) is nonsingular, i.e., $\boldsymbol{G} + \boldsymbol{S}\boldsymbol{S}^H$ is positive definite, which requires $\mathcal{G}^\perp \cap \mathcal{S}^\perp = \{0\}$. (The nullspace of $\boldsymbol{G}$ is denoted by $\mathcal{G}^\perp$.) Then the solution to (9) is

$$\hat{\boldsymbol{x}} = \tilde{\boldsymbol{A}}(\boldsymbol{S}^H\tilde{\boldsymbol{A}})^{-1}\boldsymbol{y} = \tilde{\boldsymbol{A}}(\boldsymbol{S}^H\tilde{\boldsymbol{A}})^{-1}\boldsymbol{S}^H\boldsymbol{x}, \tag{10}$$

where $\tilde{\boldsymbol{A}} = (\boldsymbol{G} + \boldsymbol{S}\boldsymbol{S}^H)^{-1}\boldsymbol{S}$. In general, $\hat{\boldsymbol{x}} \neq \boldsymbol{x}$, except for the special case, in which $\boldsymbol{x}$ lies in $\text{range}(\tilde{\boldsymbol{A}})$.

Consider the sampling function $\boldsymbol{S}^H = \boldsymbol{\Phi}\boldsymbol{U}^H$ that was used before. Using the graph-shift operator $\boldsymbol{G} = \boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{U}^H$, we can express the term $\boldsymbol{S}(\boldsymbol{S}^H\tilde{\boldsymbol{A}})^{-1}\boldsymbol{S}^H$ in (10) as a graph filter $\tilde{\boldsymbol{H}}_{\text{samp}} = \boldsymbol{U}\tilde{\boldsymbol{H}}_{f,\text{samp}}\boldsymbol{U}^H$ with frequency response

$$\tilde{\boldsymbol{H}}_{f,\text{samp}} = \boldsymbol{\Phi}^T[\boldsymbol{\Phi}(\boldsymbol{\Lambda} + \boldsymbol{\Phi}^T\boldsymbol{\Phi})^{-1}\boldsymbol{\Phi}^T]^{-1}\boldsymbol{\Phi} \quad \text{(diagonal)}.$$

Thus, the reconstruction may be expressed as

$$\hat{\boldsymbol{x}} = \tilde{\boldsymbol{H}}_{\text{interp}}\tilde{\boldsymbol{H}}_{\text{samp}}\boldsymbol{x},$$

where $\tilde{\boldsymbol{H}}_{\text{interp}} = \boldsymbol{U}(\boldsymbol{\Lambda} + \boldsymbol{\Phi}^T\boldsymbol{\Phi})^{-1}\boldsymbol{U}^H$.

5. CONSISTENT RECOVERY WITHOUT INPUT PRIORS

In this section, we provide a reconstruction method that allows arbitrary inputs, which are not necessarily defined on a graph. In order to incorporate the geometric structure of the data that might be used to further process the data, we restrict ourselves to reconstruction spaces that are related to a graph. Since no input priors are available, we cannot recover the input exactly, in general. Nonetheless, we can obtain a consistent estimate [cf. (7)] that fits our needs.

The reconstructed signal $\hat{\boldsymbol{x}}$ is constrained to lie in an M-dimensional linear subspace $\mathcal{R}$ known as the *reconstruction space* and spanned by the columns of an $N \times M$ matrix $\boldsymbol{R}$. Recall that the sampling space is also of dimension M. Since $\hat{\boldsymbol{x}} \in \mathcal{R}$, it can be decomposed as

$$\hat{\boldsymbol{x}} = \boldsymbol{R}\boldsymbol{c}; \quad \boldsymbol{R} \in \mathbb{C}^{N\times M}, \boldsymbol{c} \in \mathbb{C}^M. \tag{11}$$

Now define an operator $\boldsymbol{M} \in \mathbb{C}^{M\times K}$ that maps $\boldsymbol{y}$ to $\boldsymbol{c}$ as

$$\boldsymbol{M}\boldsymbol{y} = \boldsymbol{c}. \tag{12}$$

Then the sampling and reconstruction schemes in (4) and (11), respectively, can be summarized as

$$\hat{\boldsymbol{x}} = \boldsymbol{R}\boldsymbol{M}\boldsymbol{y} = \boldsymbol{R}\boldsymbol{M}\boldsymbol{S}^H\boldsymbol{x}. \tag{13}$$

Given $\boldsymbol{S}$ and $\boldsymbol{R}$, finding a consistent reconstruction of *any* $\boldsymbol{x} \in \mathbb{C}^N$ amounts to finding the correction transform $\boldsymbol{M}$, which may be obtained by solving

$$\text{find} \quad \boldsymbol{M} \quad \text{such that} \quad \boldsymbol{S}^H\boldsymbol{R}\boldsymbol{M}\boldsymbol{S}^H\boldsymbol{x} = \boldsymbol{S}^H\boldsymbol{x}.$$

Suppose the sampling and the reconstruction spaces (which are of the same dimension) satisfy the condition $\mathcal{R} \cap \mathcal{S}^\perp = \{0\}$. Then $\hat{\boldsymbol{x}}$ will be a unique consistent reconstruction of $\boldsymbol{x}$ if and only if $\hat{\boldsymbol{x}}$ and $\boldsymbol{x}$ have the same orthogonal projection onto $\mathcal{S}$ [23]. Specifically, a consistent reconstruction of $\boldsymbol{x}$ is given by its *oblique projection* onto the range $\mathcal{R}$ and *along* the null space $\mathcal{S}^\perp$, denoted by $\boldsymbol{E}_{\mathcal{RS}^\perp}$ with $\boldsymbol{E}_{\mathcal{RS}^\perp} = \boldsymbol{R}(\boldsymbol{S}^H\boldsymbol{R})^\dagger\boldsymbol{S}^H$. That is,

$$\hat{\boldsymbol{x}} = \boldsymbol{E}_{\mathcal{RS}^\perp}\boldsymbol{x} = \boldsymbol{R}(\boldsymbol{S}^H\boldsymbol{R})^\dagger\boldsymbol{S}^H\boldsymbol{x} = \boldsymbol{R}(\boldsymbol{S}^H\boldsymbol{R})^\dagger\boldsymbol{y} \tag{14}$$

with $\boldsymbol{M} = (\boldsymbol{S}^H\boldsymbol{R})^\dagger$. Since $\hat{\boldsymbol{x}}$ always lies in $\mathcal{R}$, if $\boldsymbol{x}$ is not in $\mathcal{R}$, then $\hat{\boldsymbol{x}} \neq \boldsymbol{x}$. Nonetheless, as a special case, when $\boldsymbol{x} \in \mathcal{R}$, consistent reconstruction (in fact, also a perfect reconstruction) reduces to the solution derived with the subspace prior in Section 4.1.

We show that the oblique projector $\boldsymbol{E}_{\mathcal{RS}^\perp}$ can be interpreted as a graph filter the special case when $\mathcal{S}$ is equal to $\mathcal{R}$. Consider the sampling function $\boldsymbol{S}^H = \boldsymbol{\Phi}\boldsymbol{U}^H$ with $M = K$ that was introduced before. Suppose we use a reconstruction function of the form $\boldsymbol{R} = \boldsymbol{U}\boldsymbol{\Phi}^H$ so that $\boldsymbol{R}\boldsymbol{R}^H$ is also a shift-invariant graph filter (in fact, the sampling and the reconstruction functions are the same in this example). Then the oblique projector $\boldsymbol{R}(\boldsymbol{S}^H\boldsymbol{R})^\dagger\boldsymbol{S}^H$ simplifies to $\boldsymbol{R}(\boldsymbol{S}^H\boldsymbol{R})^{-1}\boldsymbol{S}^H$, which is a graph filter with frequency response $\boldsymbol{\Phi}^H\boldsymbol{\Phi}$.

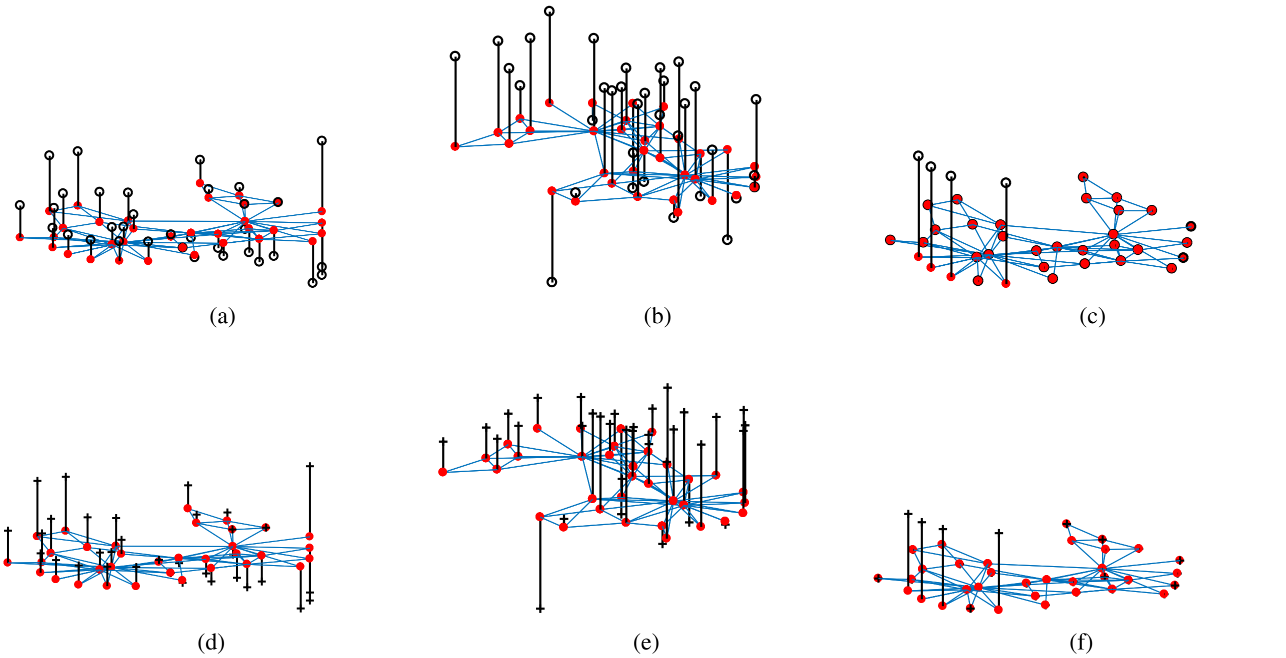

Fig. 1: Zachary's Karate club network with $N = 34$. Top row consists of the true signals, while the bottom row consists of the reconstructed signals. (a) and (d) Least squares recovery with the subspace prior. True graph signal in $\mathcal{A} = \mathrm{range}(\boldsymbol{U}_{\mathrm{P}})$ with $P = 5$, and $K = M = 5$. (b) and (e) Consistent recovery with the smoothness prior. (c) and (f) Consistent recovery forced to $\mathcal{R}$, where $\mathcal{R}$ is the space of graph signals x_n such that $x_n = 0$ for $n \in \mathcal{M} \subset \mathcal{N}$.

6. NUMERICAL EXPERIMENTS

In this section, we perform numerical experiments[1] to test the developed reconstruction methods. For experiments, we consider the Zachary's Karate club network [24]. This graph consists of 34 nodes representing members of the club.

To begin with, we focus on reconstruction with the subspace prior. We generate a signal that lies in the subspace $\mathcal{A} = \mathrm{range}(\boldsymbol{U}_P)$ of dimension $P = 5$, where $\boldsymbol{U}_P = [\boldsymbol{u}_1, \cdots, \boldsymbol{u}_P]$. This means that, the input is a bandlimited graph signal. The true signal can be seen in Fig. 1a. We observe $K = 5$ graph domain observations, i.e., we sample nodes $\{1, 2, \cdots, 5\}$ using $\boldsymbol{S}^H$ with $\boldsymbol{S}^H$ being the first 5 rows of the identity matrix of size 34×34. Since we precisely known the input subspace, the reconstruction is exact as can be seen in Fig. 1d.

Next, we consider reconstruction with the smoothness prior, for which we consider a smooth graph signal (generated based on the coordinates of the nodes); see the true signal in Fig. 1b. We observe $K = 15$ nodes, i.e., we sample nodes $\{20, 21, \cdots, 34\}$ such that $\boldsymbol{G} + \boldsymbol{S}\boldsymbol{S}^H$ is positive definite with $\boldsymbol{S}^H$ being the last 15 rows of the identity matrix of size 34×34. Although we do not explicitly use any information about the input subspace in this setting, the reconstruction is similar (not exact) to that of the true signal and it is consistent [cf. (7)].

Finally, we consider an example of recovering a graph-domain limited signal (i.e., a sparse graph signal) using bandlimited sampling. The true signal shown in Fig. 1c has four non-zero components at vertices $\{15, 16, 19, 21\}$ and is not bandlimited. We force the recovery to $\mathcal{R}$, where $\mathcal{R}$ is the space of graph signals x_n such that $x_n = 0$ for $n \in \mathcal{M} \subset \mathcal{N}$. In this example, we use $\mathcal{M} = \{14, 15, \cdots, 21\}$ and for $\boldsymbol{R}$ we use a binary matrix composed of the columns of the identity matrix of size $N \times N$ defined by the set $\mathcal{M}$. We gather $K = 7$ frequency domain observations such that $\boldsymbol{S}^H\boldsymbol{R}$ is full rank. Since $\mathcal{R}$ contains the input subspace, we can see in Fig. 1f that the reconstruction is exact. In contrast, if $\mathcal{R}$ does not contain the input subspace, the reconstruction will only be consistent, but not exact.

[1]Software to reproduce results of this paper may be downloaded from http://cas.et.tudelft.nl/~sundeep/sw/GsampConf.zip

7. CONCLUDING REMARKS

We discussed sampling and recovery of signals supported on graphs in this paper. It is possible to unambiguously recover the graph signal from its samples when we restrict ourselves to a subclass of graph signals, namely, signals that lie in a subspace related to the graph. To allow arbitrary signals or signals that are smooth with respect to the graph, we discussed consistent sampling and recovery, in which the input and the reconstruction yield exactly the same measurements. By choosing a sampling function, which essentially performs sample selection in the graph spectral domain, we can interpret the sampling and recovery operations as graph filters. Although we have provided an interpretation of the graph sampling and reconstruction operations as graph filtering operations, it will be interesting to study, as future work, the implementation details (e.g., filter order, polynomial fitting method) of the discussed graph filters.

8. REFERENCES

[1] E. Bullmore and O. Sporns, "Complex brain networks: graph theoretical analysis of structural and functional systems," *Nature Reviews Neuroscience*, vol. 10, no. 3, pp. 186–198, 2009.

[2] R. Guimera, S. Mossa, A. Turtschi, and L. N. Amaral, "The worldwide air transportation network: Anomalous centrality, community structure, and cities' global roles," *Proc. of the National Acad. of Sciences*, vol. 102, no. 22, pp. 7794–7799, 2005.

[3] M. O. Jackson, *Social and economic networks*. Princeton university press, 2010.

[4] 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.

[5] A. Sandryhaila and J. M. Moura, "Big data analysis with signal processing on graphs: Representation and processing of massive data sets with irregular structure," *IEEE Signal Process. Mag.*, vol. 31, no. 5, pp. 80–90, 2014.

[6] S. Chen, R. Varma, A. Sandryhaila, and J. Kovačević, "Discrete signal processing on graphs: Sampling theory," *IEEE Trans. Signal Process.*, vol. 63, no. 24, pp. 6510–6523, 2015.

[7] A. Anis, A. Gadde, and A. Ortega, "Efficient sampling set selection for bandlimited graph signals using graph spectral proxies," *IEEE Trans. Signal Process.*, vol. 64, no. 14, pp. 3775–3789, 2016.

[8] M. Tsitsvero, S. Barbarossa, and P. Di Lorenzo, "Signals on graphs: Uncertainty principle and sampling," *IEEE Trans. Signal Process.*, vol. 64, no. 18, pp. 4845–4860, 2016.

[9] D. Romero, M. Ma, and G. B. Giannakis, "Kernel-based reconstruction of graph signals," *IEEE Trans. Signal Process.*, vol. 65, no. 3, pp. 764–778, 2017.

[10] X. Zhu and M. Rabbat, "Approximating signals supported on graphs," in *Acoustics, Speech and Signal Processing (ICASSP), 2012 IEEE International Conference on*. IEEE, 2012, pp. 3921–3924.

[11] J. Leskovec and C. Faloutsos, "Sampling from large graphs," in *Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 2006, pp. 631–636.

[12] L. F. Chamon and A. Ribeiro, "Greedy sampling of graph signals," *arXiv preprint arXiv:1704.01223*, 2017.

[13] S. P. Chepuri, G. Leus *et al.*, "Sparse sensing for statistical inference," *Foundations and Trends® in Signal Processing*, vol. 9, no. 3–4, pp. 233–368, 2016.

[14] S. Joshi and S. Boyd, "Sensor selection via convex optimization," *IEEE Trans. Signal Process.*, vol. 57, no. 2, pp. 451–462, Feb. 2009.

[15] A. Krause, *Optimizing sensing: Theory and applications*, ser. Ph.D. dissertation, School of Comput. Sci. Carnegie Mellon Univ., Pittsburgh, PA, United States, 2008.

[16] A. G. Marques, S. Segarra, G. Leus, and A. Ribeiro, "Sampling of graph signals with successive local aggregations," *IEEE Trans. Signal Process.*, vol. 64, no. 7, pp. 1832–1843, 2016.

[17] S. P. Chepuri and G. Leus, "Graph sampling for covariance estimation," *arXiv preprint arXiv:1704.07661*, 2017.

[18] A. Hirabayashi and M. Unser, "Consistent sampling and signal recovery," *IEEE Trans. Signal Process.*, vol. 55, no. 8, pp. 4104–4115, 2007.

[19] Y. C. Eldar and T. Michaeli, "Beyond bandlimited sampling," *IEEE Signal Process. Mag.*, vol. 26, no. 3, pp. 48–68, 2009.

[20] T. Michaeli and Y. C. Eldar, *Optimization techniques in modern sampling theory*. Cambridge, UK: Cambridge Univ. Press, 2010.

[21] Y. C. Eldar, *Sampling Theory: Beyond Bandlimited Systems*. Cambridge University Press, 2015.

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

[23] Y. C. Eldar, "Sampling with arbitrary sampling and reconstruction spaces and oblique dual frame vectors," *Journal of Fourier Analysis and Applications*, vol. 9, no. 1, pp. 77–96, 2003.

[24] W. W. Zachary, "An information flow model for conflict and fission in small groups," *Journal of anthropological research*, vol. 33, no. 4, pp. 452–473, 1977.