Sampling and Reconstruction of Band-limited Graph Signals using Graph Syndromes

A Anil Kumar, N Narendra, M Girish Chandra and Kriti Kumar
TCS Research and Innovation, Bangalore, India.
Email:{achannaanil.kumar, n.narendra, m.gchandra, kriti.kumar}@tcs.com.

***Abstract*—The problem of sampling and reconstruction of band-limited graph signals is considered in this paper. A new sampling and reconstruction method based on the idea of error and erasure correction is proposed. We visualize the process of sampling as removal of nodes akin to introducing erasures, due to which the *graph syndromes* of a sampled signal gives rise to significant values, which otherwise would be minuscule for a band-limited signal. A reconstruction method by making use of these significant values in the graph syndromes is described and correspondingly the necessary and sufficient conditions for unique recovery and some key properties is provided. Additionally, this method allows for robust reconstruction i.e., reconstruction in the presence of few corrupted sampled nodes and a method based on weighted ℓ_1 - norm is described. Simulation results are provided to demonstrate the efficiency of the method which shows better mean squared error performance compared to existing methods.**

***Index terms*— Graph signal processing, Graph syndrome, error correction, Sampling and reconstruction, Robust reconstruction.**

I. Introduction

The limitation of the traditional discrete signal processing (DSP) in handling complex irregular structured data which are continuously being generated from various physical sources such as social networks, biological networks etc, [1], [2] has led to the growth of other alternative fields; a predominant field among them being the *graph signal processing* (GSP) [3]. It is well known that graph provides a natural representation of the data in many of the above mentioned areas and GSP broadly refers to processing of signals that reside on the vertices by taking into consideration the underlying graph topology. In this paper, we consider an important problem in this emerging area of GSP namely the sampling and reconstruction, and propose a new method based on the idea of error and erasure correction, which besides reconstruction, has the additional advantage of robust reconstruction i.e., reconstruction in the presence of outliers in the sampled signal.

In traditional DSP, which deals with discrete-time signals, it is well known that sampling and reconstruction is an important problem whose goal fundamentally is to recover high dimensional signal (i.e., reconstruction) from a low dimensional signal (i.e., from small subset of samples). Further, it is well established through the classical Nyquist-Shannon sampling theorem that band-limitedness is an essential key pre-requisite for sampling without significant information loss. Similarly, even in GSP, band-limitedness forms an important pre-requirement for sampling and analogous to traditional DSP, band-limitedness of graph signals in GSP is defined with the support of graph Fourier transform (GFT) [3]. Several new techniques over the recent past are being proposed to address the various aspects of sampling and reconstruction of band-limited graph signals. Sampling set selection i.e., choosing the appropriate subset of vertices, which facilitates *unique* and *stable* reconstruction is an important step and the works such as [4] - [6] (and references there-in) addressed this problem and described techniques for obtaining an appropriate choice of sampling set. While [4] proposed an approach for obtaining the sampling set for unique recovery, [5] proposed a greedy approach for stable reconstruction. Whereas these approaches require the graph Fourier basis, [6] described a computationally efficient method for determining the cut-off frequency and choosing a sampling set using only the graph variational operator such as graph Laplacian. Various methods for reconstruction are also proposed for example in [5] - [9]. While [7] - [9] described an approach based on appropriately designing the graph shift-invariant filters that are similar to traditional linear time-invariant filter, [5], [6] focused on consistent reconstruction[1] and used the frame-theoretic methods borrowed from [10]. All these reconstruction methods assume that the sampled signal is uncorrupted and reconstruction suffers in the presence of large outliers. To add robustness against these errors, [11] proposed an approach based on graph total variation minimization.

In this paper, we revisit this problem of sampling and reconstruction and provide an alternative method based on *graph syndromes*. We visualize the process of sampling as removal of a subset of nodes akin to introducing erasures in coding theory. The removal of nodes causes disturbance to the GFT spectrum. The graph syndrome which corresponds to the disturbance in the stop-band region is captured via the corresponding matrix referred as *graph parity check matrix*. The various relationships between the graph syndromes, the sampled signal and the reconstructed signal is outlined and based on which the necessary and sufficient condition for the unique reconstruction is provided. Furthermore, we show that the proposed reconstruction is consistent and is identical to the frame-theoretic method outlined in [6] for the same given sampling set. Now, any larger outliers on the sampled

[1]Consistent reconstruction, in brief refers to reconstruction without changing the observed variables. For more details, readers may refer to [10]

ISBN 978-90-827970-1-5 © EURASIP 2018

signal which is analogous to errors also causes disturbance and gets captured in the graph syndrome; a method based on weighted-ℓ_1-norm [12] is described for robust reconstruction. Simulation results are provided to corroborate the methods discussed in this paper and also to compare the performance with the existing schemes. The results show a similar sampling set selection compared to [5] and a better mean squared error (MSE) performance compared to [11] in case of robust reconstruction.

II. Preliminaries

In this section we provide the general graph notations and briefly describe the band-limitedness and existing sampling and reconstruction of band-limited graph signals. This is then followed by the definition of graph syndromes.

A. Graph signals

Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ denote a known, connected, undirected and weighted graph consisting of N nodes indexed by set $\mathcal{V} = \{1, 2, ..., N\}$ and connected by edges $\mathcal{E} = \{(i, j, w_{ij})\}, i, j \in \mathcal{V}$, where w_{ij} denotes the weight of the edge between i^{th} and the j^{th} node and $w_{ii} = 0$. The adjacency matrix $\mathbf{W}$ is an $N \times N$ matrix with $[\mathbf{W}]_{i,j} = w_{i,j}$. Due to the assumption of undirected graphs, $\mathbf{W} \in \mathbb{S}^N$. The degree of the i^{th} node can be defined as $d_i = \sum_{j=1}^{N} [\mathbf{W}]_{i,j}$, and the degree matrix $\mathbf{D} = \text{diag}(d_1, d_2, ..., d_N)$. The important key matrix, graph Laplacian is defined as $\mathbf{L} = \mathbf{D} - \mathbf{W}$. The graph shift operator $\mathbf{S} \in \mathbb{S}^N$, where $[\mathbf{S}]_{i,j}$ can be non-zero only if $i = j$ or if $(i, j) \in \mathcal{E}$. $\mathbf{S}$ captures the local structures of the graphs and choices for the graph shift operator are usually either the Laplacian matrix or the adjacency matrix [2], [3]. A graph signal is a scalar value assigned on each vertex or alternately, it is a function $f : \mathcal{V} \to \mathbb{R}$. Since $\mathbf{S}$ is a symmetric matrix, it admits the eigenvalue decomposition $\mathbf{S} = [\mathbf{u}_1, ..., \mathbf{u}_N]\text{diag}(\lambda_1, ..., \lambda_N)[\mathbf{u}_1, ..., \mathbf{u}_N]^H$. The eigenvectors and eigenvalues, $\mathbf{U} = \{\mathbf{u}_1, ..., \mathbf{u}_N\}$ and $\{\lambda_1, ..., \lambda_N\}$ provide a notion of frequency in the context of graphs [3]. Now, the GFT and the inverse GFT of the graph signal $\mathbf{f}$ can be defined as $\tilde{\mathbf{f}} = \mathbf{U}^H\mathbf{f}$ and $\mathbf{f} = \mathbf{U}\tilde{\mathbf{f}}$ respectively.

B. Band-limitedness and Sampling

1) Band-limited graph signal: A graph signal $\mathbf{f}$ is said to be ω-bandlimited if $\tilde{\mathbf{f}}_i = 0$ for all i with $|\lambda_i| > \omega$. Let $\mathcal{R} = \{1, 2, ..., r\}$ and the complimentary set $\mathcal{R}^{\mathcal{C}} = \mathcal{V} \setminus \mathcal{R}$, where r denotes the number of eigenvalues that are less than ω. Now, we can express the GFT matrix as $\mathbf{U} = [\mathbf{U}_{\mathcal{VR}}|\mathbf{U}_{\mathcal{VR}^c}]$, where $\mathbf{U}_{\mathcal{VR}}$ and $\mathbf{U}_{\mathcal{VR}^c}$ are matrices of dimension $N \times r$ and $N \times N-r$ respectively. Now, using $\mathbf{U}_{\mathcal{VR}}$ we can easily express the ω-band-limited signal $\mathbf{f}$ as [6]

$$\mathbf{f} = \sum_{i=1}^{r} \mathbf{u}_i\tilde{\mathbf{f}}_i = \mathbf{U}_{\mathcal{VR}}\tilde{\mathbf{f}}_{\mathcal{R}}. \tag{1}$$

The set $\{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_r\}$ spans a vector space which is referred as Paley-Wiener space [13] denoted by $PW_\omega(\mathcal{G})$ and consists of all ω-band-limited signals. In the following section, we briefly outline the sampling and reconstruction techniques of a graph signal $\mathbf{f} \in PW_\omega(\mathcal{G})$.

2) Sampling and reconstruction in $PW_\omega(\mathcal{G})$: A. Uniqueness of Sampling Set:
Let $\mathcal{S} \subseteq \mathcal{V}$ denote the sampling set with cardinality, i.e., $|\mathcal{S}| = d$, and the complement sampling set $\mathcal{S}^{\mathcal{C}} = \mathcal{V} \setminus \mathcal{S}$. The uniqueness set on $PW_\omega(\mathcal{G})$ is defined as follows [13].

Definition 2.1: Sampling set $\mathcal{S}$ is a uniqueness set for the space $PW_\omega(\mathcal{G})$ if and only if $\mathbf{f}^1_{\mathcal{S}} = \mathbf{f}^2_{\mathcal{S}}$ implies $\mathbf{f}^1 = \mathbf{f}^2$ for all $\mathbf{f}^1, \mathbf{f}^2 \in PW_\omega(\mathcal{G})$.
Here $\mathbf{f}_{\mathcal{S}}$ denotes the sampled signal i.e., it contains the elements corresponding to the sampling set $\mathcal{S}$. Let $L_2(\mathcal{S}^{\mathcal{C}})$ denote the space of all vectors in $\mathbb{R}^N$ that are zero at all places corresponding to $\mathcal{S}$. The following result provides the necessary and sufficient condition for uniqueness of sampling set $\mathcal{S}$ for any signal $\mathbf{f} \in PW_\omega(\mathcal{G})$ [4], [6].

Lemma 2.2: $\mathcal{S}$ is a uniqueness set for $PW_\omega(\mathcal{G})$ if and only if $PW_\omega(\mathcal{G}) \cap L_2(\mathcal{S}^{\mathcal{C}}) = \{\mathbf{0}\}$.

B. Reconstruction and Sampling set Selection:
Let $\mathbf{S}_d$ be a matrix whose columns are indicator functions for $\mathcal{S}$ and the sampling operator $\mathbf{S}_d^T : \mathbb{R}^N \to \mathbb{R}^d$ [6]. The sampled sub-vector $\mathbf{f}_{\mathcal{S}} = \mathbf{S}_d^T\mathbf{f}$. A well known method for consistent reconstruction is achieved based on the frame operator theory described in [10] by making use of the following expression [5], [6]

$$\hat{\mathbf{f}}_{PW} = \underbrace{\mathbf{U}_{\mathcal{VR}}(\mathbf{U}^H_{\mathcal{SR}}\mathbf{U}_{\mathcal{SR}})^{-1}\mathbf{U}^H_{\mathcal{SR}}}_{\mathbf{U}_{PW}}\mathbf{f}_{\mathcal{S}} \tag{2}$$

where the sub matrix $\mathbf{U}_{\mathcal{SR}} = \mathbf{S}_d^T\mathbf{U}_{\mathcal{VR}}$.

For optimal sampling set selection, various methods as stated in Section I have been proposed. Among them a key method is based on choosing appropriate rows of $\mathbf{U}_{\mathcal{VR}}$, since from the above expression it is quite evident that the performance depends upon $\mathbf{U}_{\mathcal{SR}}$. Hence, in the existing works such as [5], $\mathbf{U}_{\mathcal{VR}}$ is used for choosing the optimal sampling set $\mathcal{S}$ based on various objectives such as minimum reconstruction error etc. For a comprehensive summary of the objectives and the corresponding optimality criterion on $\mathbf{U}_{\mathcal{SR}}$, interested readers may refer to [6, Table II].

In the following section, we introduce the notion of graph syndrome which provide an alternative means for sampling and reconstruction by using the matrix $\mathbf{U}_{\mathcal{VR}^c}$. This new framework as shall be described later in Section III, in addition to reconstruction, it facilitates robust reconstruction i.e., reconstruction in the presence of outliers in the sampled signal.

C. Graph Syndrome

The graph syndrome of a graph signal $\mathbf{f}$ is defined as

$$\mathbf{m}_{\mathbf{f}} = \mathbf{U}^H_{\mathcal{VR}^c}\mathbf{f}. \tag{3}$$

Now, we have the following property on $(\mathbf{U}^H_{\mathcal{VR}^c})$

Proposition 2.3: Assume $\mathbf{f} \neq \mathbf{0}$, then $\mathbf{f} \in \mathcal{N}(\mathbf{U}^H_{\mathcal{VR}^c})$ (i.e., null space of $\mathbf{U}^H_{\mathcal{VR}^c}$) if and only if $\mathbf{f} \in PW_\omega(\mathcal{G})$.
The proof is easy and is clearly evident from the definitions of $PW_\omega(\mathcal{G})$ and GFT, and hence is not elaborated here. The above proposition implies that graph syndrome $\mathbf{m}_{\mathbf{f}} = \mathbf{0}$ for a perfect ω-band-limited signal. Now, one can easily draw the

ISBN 978-90-827970-1-5 © EURASIP 2018

parallels between the above definitions and the error correction codes [14]. It can easily be seen that the matrices $\mathbf{U}_{\mathcal{VR}}$ and $\mathbf{U}_{\mathcal{VR}^c}$ are analogous to generator and the parity check matrix respectively. Hence, in this context of graphs, we refer to $\mathbf{U}_{\mathcal{VR}}$ and $\mathbf{U}_{\mathcal{VR}^c}$ as the *graph generator matrix* and *graph parity check matrix* respectively and correspondingly $\mathbf{m}_\mathbf{f}$ as *graph syndrome* of the graph signal $\mathbf{f}$. The following section describes the process of sampling and reconstruction based on these graph syndromes.

III. Sampling and Reconstruction using Graph Syndromes

In this section we first describe the reconstruction of the sampled signal using the graph syndromes; followed by this the conditions for unique reconstruction and sampling set selection is discussed. Later it is extended for robust reconstruction.

A. Graph Syndromes based Reconstruction

The sampled signal $\mathbf{f}_\mathcal{S}$ can equivalently be expressed as

$$\mathbf{S}_d\mathbf{f}_\mathcal{S} = \mathbf{f} - \mathbf{S}_N^C\mathbf{f} \tag{4}$$

where $\mathbf{S}_N^C$ is a diagonal matrix of order $N \times N$ with diagonal entries of one's corresponding to $\mathcal{S}^\mathcal{C}$ and zero's corresponding to $\mathcal{S}$. Notice the difference between $\mathbf{S}_d$ and $\mathbf{S}_N$; unlike $\mathbf{S}_N$, $\mathbf{S}_d$ contains one's corresponding to $\mathcal{S}$ and zero's corresponding to $\mathcal{S}^\mathcal{C}$ and further the columns corresponding to all zeros are removed. Also, notice that the vector $\mathbf{S}_N^C\mathbf{f} \in L_2(\mathcal{S}^\mathcal{C})$ i.e., it contains zeros corresponding to sampling nodes in $\mathcal{S}$. Now using (3) the graph syndrome of the above equation can be estimated as

$$\mathbf{m}_{\mathbf{f}_\mathcal{S}} = \mathbf{U}_{\mathcal{VR}^c}^H\mathbf{S}_d\mathbf{f}_\mathcal{S}. \tag{5}$$

Substituting (4) in the above equation, we can express the graph syndrome as

$$\mathbf{m}_{\mathbf{f}_\mathcal{S}} = -\mathbf{U}_{\mathcal{VR}^c}^H\mathbf{S}_N^C\mathbf{f} \tag{6}$$

$$= -\mathbf{U}_{\mathcal{S}^c\mathcal{R}^c}^H\mathbf{f}_{\mathcal{S}^c} \tag{7}$$

for (6) Proposition 2.3 is used, while for (7) the structure of $\mathbf{S}_N^C$ is used. $\mathbf{U}_{\mathcal{S}^c\mathcal{R}^c}^H$ and $\mathbf{f}_{\mathcal{S}^c}$ contains the columns and the elements corresponding to the complementary sampling set $\mathcal{S}^\mathcal{C}$ respectively. Now, observe that $\mathbf{S}_N^C = \mathbf{S}_N^C\mathbf{S}_N^C$ and by substituting it in (6), the vector $\mathbf{S}_N^C\mathbf{f} = -(\mathbf{U}_{\mathcal{S}^c\mathcal{R}^c}^H)^\dagger\mathbf{m}_{\mathbf{f}_\mathcal{S}}$, where $(.)^\dagger$ denotes the Moore-Penrose Pseudoinverse. Again substituting for $\mathbf{m}_{\mathbf{f}_\mathcal{S}}$ from (5), $\mathbf{S}_N^C\mathbf{f} = -(\mathbf{U}_{\mathcal{S}^c\mathcal{R}^c}^H)^\dagger\mathbf{U}_{\mathcal{VR}^c}^H\mathbf{S}_d\mathbf{f}_\mathcal{S}$. Now, using this expression and (4), the reconstructed signal using graph syndrome approach $\hat{\mathbf{f}}_{GS}$ can be expressed as

$$\begin{aligned}\hat{\mathbf{f}}_{GS} &= \mathbf{S}_d\mathbf{f}_\mathcal{S} + \mathbf{S}_N^C\mathbf{f} = \mathbf{S}_d\mathbf{f}_\mathcal{S} - (\mathbf{U}_{\mathcal{S}^c\mathcal{R}^c}^H)^\dagger\mathbf{U}_{\mathcal{VR}^c}^H\mathbf{S}_d\mathbf{f}_\mathcal{S} \\ &= \underbrace{(\mathbf{I}_N - (\mathbf{U}_{\mathcal{S}^c\mathcal{R}^c}^H)^\dagger\mathbf{U}_{\mathcal{VR}^c}^H)\mathbf{S}_d}_{\mathbf{U}_{GS}}\mathbf{f}_\mathcal{S}.\end{aligned} \tag{8}$$

Notice the difference between (2) and (8); while (2) relates the reconstruction signal and the sampled signal using $\mathbf{U}_{\mathcal{VR}}$ and $\mathbf{U}_{\mathcal{SR}}$, (8) relates using $\mathbf{U}_{\mathcal{VR}^c}$ and $\mathbf{U}_{\mathcal{S}^c\mathcal{R}^c}$. Now, we have the following property on reconstruction with graph syndromes.

Proposition 3.1: The reconstruction obtained using the graph syndromes approach is consistent and is identical to (2) i.e., $\hat{\mathbf{f}}_{GS} = \hat{\mathbf{f}}_{PW}$.

We only provide an outline of the proof here. From (7) it is clear that only $\hat{\mathbf{f}}_{\mathcal{S}^c}$ is estimated using the graph syndromes approach and hence $\hat{\mathbf{f}}_\mathcal{S} = \mathbf{f}_\mathcal{S}$ implying consistent reconstruction. For the reconstruction equivalence, it can be shown that $\mathbf{U}_{GS} = \mathbf{U}_{PW}$ using matrix manipulations. The manipulation steps goes along the lines as provided in [15, Section V.C] and hence is not elaborated here. ■

In the following, we provide the necessary and sufficient conditions for unique reconstruction and briefly describe a method for sampling set selection with this graph syndromes approach.

1) Uniqueness of Reconstruction and Sampling set Selection: The following theorem provides the necessary and sufficient condition for the unique reconstruction with this approach.

Theorem 3.2: For any signal $\mathbf{f} \neq \mathbf{0}$ and $\mathbf{f} \in PW_\omega(\mathcal{G})$, the sampling set $\mathcal{S}$ is a uniqueness set if and only if $\mathbf{U}_{\mathcal{S}^c\mathcal{R}^c}^H$ is full rank.

Proof Assume $\mathcal{S}$ is a uniqueness set and recall that $\mathbf{S}_N^C\mathbf{f} \in L_2(\mathcal{S}^c)$. Thus from Lemma 2.2, $\mathbf{S}_N^C\mathbf{f} \notin PW_\omega(\mathcal{G})$ which further implies from Proposition 2.3 that $\mathbf{m}_{\mathbf{f}_\mathcal{S}} \neq \mathbf{0}$, and hence $\mathbf{U}_{\mathcal{S}^c\mathcal{R}^c}^H$ must be full rank (see (7)). Conversely, if $\mathbf{U}_{\mathcal{S}^c\mathcal{R}^c}^H$ is full rank then it can easily be observed that matrix $\mathbf{U}_{GS}$ is also full rank. Hence, using (8) it is easy to notice that for any $\mathbf{f}_\mathcal{S}^1 = \mathbf{f}_\mathcal{S}^2$, $\hat{\mathbf{f}}_{GS}^1 = \hat{\mathbf{f}}_{GS}^2$ which by definition 2.1 implies $\mathcal{S}$ is uniqueness set. ■

The above theorem leads the following observation in terms of the graph syndrome; for unique recovery the graph syndromes must satisfy the two conditions i) for any $\mathbf{f} \in PW_\omega(\mathcal{G})$, $\mathbf{m}_{\mathbf{f}_\mathcal{S}} \neq \mathbf{0}$, ii) for any $\mathbf{f}^1, \mathbf{f}^2 \in PW_\omega(\mathcal{G})$ and $\mathbf{f}^1 \neq \mathbf{f}^2$, $\mathbf{m}_{\mathbf{f}_\mathcal{S}^1} \neq \mathbf{m}_{\mathbf{f}_\mathcal{S}^2}$.

2) Sampling set selection and discussion: Similar to using graph generator matrix i.e., $\mathbf{U}_{\mathcal{VR}}$ for selection of best sampling set with certain objectives and optimality criterion on $\mathbf{U}_{\mathcal{SR}}$ as mentioned in Section II, the graph parity check matrix $\mathbf{U}_{\mathcal{VR}^c}^H$ can be used for choosing the sampling set. One can use identical objectives and optimality criterion on $\mathbf{U}_{\mathcal{S}^c\mathcal{R}^c}$ as listed in [6, Table II] to obtain the sampling set. Further, the same greedy algorithms as described in the existing works such as in [5] can be employed here for choosing the optimal sampling set. However, it is important to observe one key difference between using $\mathbf{U}_{\mathcal{SR}}$ and $\mathbf{U}_{\mathcal{S}^c\mathcal{R}^c}$ for sampling set selection; while using $\mathbf{U}_{\mathcal{SR}}$ one obtains the sampling set $\mathcal{S}$, by using $\mathbf{U}_{\mathcal{S}^c\mathcal{R}^c}$ we obtain the complementary sampling set $\mathcal{S}^\mathcal{C}$.

Now, for the same given objective, due to matrix equivalence (i.e., $\mathbf{U}_{GS} = \mathbf{U}_{PW}$), one may expect to obtain similar sampling set by employing either the graph generator matrix $\mathbf{U}_{\mathcal{VR}}$ or the graph parity check matrix $\mathbf{U}_{\mathcal{VR}^c}^H$. It is important to note that even though we have mentioned similar sampling set, they are not identical due to adoption of sub optimal greedy algorithm; as would be demonstrated in Section IV. Recall, from Section II that the dimensions of matrices $\mathbf{U}_{\mathcal{VR}}$ and $\mathbf{U}_{\mathcal{VR}^c}$ are $N \times r$ and $N \times (N - r)$ respectively. Thus, if

ISBN 978-90-827970-1-5

$r > N/2$ then it will be computationally advantageous to use $\mathbf{U}_{\mathcal{VR}^c}^H$ for sampling set selection, else one can use $\mathbf{U}_{\mathcal{VR}}$.

In addition to providing identical performance as suggested by Proposition 3.1 in the absence of outliers, the proposed graph syndrome based framework scores over the frame-theoretic approach by providing robust reconstruction in the presence of outliers.

B. Robust Graph Syndromes based Reconstruction

The above described reconstruction inherently trusts the sampled signal $\mathbf{f}_{\mathcal{S}}$ i.e., it assumes that all samples of $\mathbf{f}_{\mathcal{S}}$ are uncorrupted. However, when large errors are present, these disturbances influences or affects the reconstruction process. Robust reconstruction aims to minimize the influence of these disturbances by separating out these errors from the sampled signal. Let $\mathbf{e}$ denote the vector of size $d \times 1$ containing these large errors which we assume it to be sparse and $\mathcal{E}$ denote the set of erroneous locations with $|\mathcal{E}| = K$. The erroneous sampled signal can be expressed as

$$\mathbf{t}_{\mathcal{S}} = \mathbf{f}_{\mathcal{S}} + \mathbf{e}. \tag{9}$$

Using (5) the syndrome of $\mathbf{t}_{\mathcal{S}}$ can be expressed as

$$\begin{aligned} \mathbf{m}_{\mathbf{t}_{\mathcal{S}}} &= \mathbf{U}_{\mathcal{VR}^c}^H \mathbf{S}_d \mathbf{t}_{\mathcal{S}} = \mathbf{U}_{\mathcal{VR}^c}^H \mathbf{S}_d (\mathbf{f}_{\mathcal{S}} + \mathbf{e}) \\ &= \mathbf{m}_{\mathbf{f}_{\mathcal{S}}} + \mathbf{m}_{\mathbf{e}_{\mathcal{E}}} \end{aligned} \tag{10}$$

where the error syndrome $\mathbf{m}_{\mathbf{e}_{\mathcal{E}}} = -\mathbf{U}_{\mathcal{ER}^c}^H \mathbf{e}_{\mathcal{E}}$, $\mathbf{U}_{\mathcal{ER}^c}^H$ and $\mathbf{e}_{\mathcal{E}}$ denotes the columns and elements of $\mathbf{U}_{\mathcal{VR}^c}^H$ and $\mathbf{e}$ respectively corresponding to the error location set $\mathcal{E}$. Unlike $\mathcal{S}^{\mathcal{C}}$ which is known, the error location set $\mathcal{E}$ is unknown and hence the reconstruction described above is not directly applicable. Further, due to non-separability of the contribution to syndrome corresponding to sampling and errors, the reconstruction has to be jointly addressed. We borrow the idea of weighted-ℓ_1 -norm minimization [12] and propose to solve the following optimization problem for joint reconstruction:

$$\min_{\mathbf{x}} |\mathbf{B}\mathbf{x}|_{\ell_1} \text{ subject to } \mathbf{m}_{\mathbf{t}_{\mathcal{S}}} = \mathrm{U}_{\mathcal{VR}^c}^H \mathbf{x} \tag{11}$$

where $\mathbf{B}$ is a diagonal matrix of size $N \times N$ referred as weighting matrix whose diagonal elements corresponding to locations $\mathcal{S}$ contains one, while at locations corresponding to $\mathcal{S}^{\mathcal{C}}$ it has a very low value, say $\sigma \ll 1$. Suppose if the model is assumed to have some background noise then the equality constraint has to be replaced with $|\mathbf{m}_{\mathbf{t}_{\mathcal{S}}} - \mathrm{U}_{\mathcal{VR}^c}^H \mathbf{x}|_{\ell_2} \leq \epsilon$. By solving the above optimization problem we shall obtain the vector $\mathbf{x} = \mathbf{S}_N^{\mathcal{C}} \hat{\mathbf{f}} + \mathbf{S}_d \hat{\mathbf{e}}$. Since $\mathcal{S}^{\mathcal{C}} \cap \mathcal{E} = \{\mathbf{0}\}$, by keeping a threshold on $\mathbf{x}$, one can easily obtain the elements corresponding to $\hat{\mathbf{f}}_{\mathcal{S}^{\mathcal{C}}}$ and $\hat{\mathbf{e}}_{\mathcal{E}}$.

IV. Simulation Results

In this section we present the numerical results of the proposed method and compare it with the existing methods. For the experiments, we use a random sensor graph comprising of $N = 100$ nodes which is generated using the GSPBOX [16]. In all the experiments we generated a band-limited graph signal by assuming $r = 30$ and hence, the size of $\mathbf{U}_{\mathcal{VR}} = 100 \times 30$ and $\mathbf{U}_{\mathcal{VR}^c} = 100 \times 70$.

In the first simulation, we compare the sampling set obtained with the two approaches; using the graph generator matrix and graph parity check matrix i.e., with $\mathbf{U}_{\mathcal{VR}}$ and $\mathbf{U}_{\mathcal{VR}^c}$ respectively. In both the cases we keep the optimality criterion as maximizing the minimum eigenvalue of the sampled matrices $\mathbf{U}_{\mathcal{SR}}$ and $\mathbf{U}_{\mathcal{S}^c\mathcal{R}^c}$ respectively and use the greedy algorithm provided in [5, Algorithm 1] for choosing the sampling set. Recall from Section III-A2 that by using the graph parity check matrix $\mathbf{U}_{\mathcal{VR}^c}$, we obtain the set $\mathcal{S}^{\mathcal{C}}$ and subsequently, the sampling set is obtained using $\mathcal{V} \setminus \mathcal{S}^{\mathcal{C}}$. Fig.1 shows the graph topology of 100 nodes and the sampling nodes obtained with the two approaches. From the figure notice that except for few sampled node locations (in this case two nodes) the other sampling node locations obtained with both the approaches are identical. This small difference in sampling set is due to the usage of sub-optimal greedy algorithm. Further, it is important to notice that the two locations even though are different, they are in close vicinity. Although the results are not provided here, we observed similar behavior for other optimality criterion listed in [6, Table II]. Thus, as mentioned in Section III-A2 depending upon the size of r, one to can employ either the graph generator matrix or the graph parity check matrix for sampling set selection.

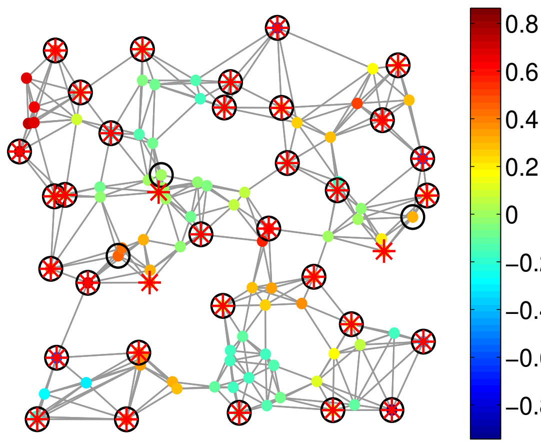

Fig. 1. Comparison of sampling nodes obtained with the graph generator matrix and graph parity check matrix approach. The sampled nodes obtained with the graph generator approach is depicted using black 'o', whereas with the graph parity check approach it is depicted with red '*'.

In the next simulation, we compare the reconstruction performance in the presence of outliers. $K = 5$ outliers are added to the sampled signal and Fig. 2 shows the MSE reconstruction performance as a function of the sampling ratio of the three approaches; the robust graph total variation regularization of [11], the graph syndrome based reconstruction (i.e., using just (8)) and the robust graph syndrome based reconstruction i.e., by solving (11) mentioned as RGTVR, GSR and RGSR respectively in the figure. For the sake of comparison, the reconstruction obtained with the Oracle method is also provided.

ISBN 978-90-827970-1-5 © EURASIP 2018

In oracle method, the reconstruction is achieved by assuming the error or the outliers locations to be known, which sort of provides a lower bound. Notice from the figure that as the number of sampling nodes increase, the MSE reduces corresponding to all three approaches, which is along the expected lines. Now, between the three methods, RGTVR and the proposed RGSR reconstructs by taking into consideration the outliers and hence shows better performance than GSR. In RGTVR, contrary to the proposed RGSR method, several performance sensitive hyper parameters are required to be chosen appropriately. In spite of choosing these parameters for the best performance by exhaustive search, one can see from the figure that RGSR performs better than tuned RGTVR and is closer to the oracle approach, thus demonstrating better robustness to outliers.

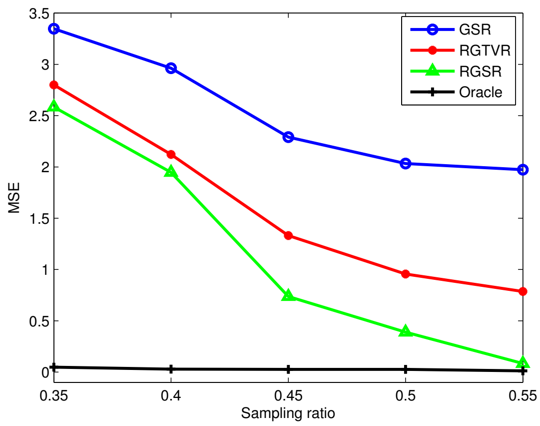

Fig. 2. MSE vs sampling ratio comparison of reconstruction in the presence of outliers.

V. Conclusion

By visualizing the sampling of a graph signal as removal of nodes, which in turn can be seen as introducing erasures, the paper suggests a new method for graph signal sampling and reconstruction. Linking the graph syndrome to the band-stop spectrum, which manifests as a disturbance in the sampling process, a reconstruction method has been put forward. Further, it is shown that the sampling set must be chosen such that the sub graph parity check matrix must be full rank for unique reconstruction. Additionally, this framework inherently allows for robust reconstruction by treating outliers as errors and a method based on weighted ℓ_1-norm minimization is arrived for this purpose. From the simulation results and the theory discussed in the paper, it can be concluded that the proposed graph syndrome based method has better robustness to outliers while reconstructing the graph signals.

References

[1] M. Newman, *Networks: an introduction*. Oxford university press, 2010.

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

[3] 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 Processing Magazine*, vol. 30, no. 3, pp. 83–98, 2013.

[4] H. Shomorony and A. S. Avestimehr, "Sampling large data on graphs," in *Signal and Information Processing (GlobalSIP), 2014 IEEE Global Conference on*. IEEE, 2014, pp. 933–936.

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

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

[7] S. Segarra, A. G. Marques, G. Leus, and A. Ribeiro, "Interpolation of graph signals using shift-invariant graph filters," in *Signal Processing Conference (EUSIPCO), 2015 23rd European*. IEEE, 2015, pp. 210–214.

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

[9] S. Segarra, A. G. Marques, G. Leus, and A. Ribeiro, "Reconstruction of graph signals through percolation from seeding nodes," *IEEE Transactions on Signal Processing*, vol. 64, no. 16, pp. 4363–4378, 2016.

[10] 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.

[11] S. Chen, A. Sandryhaila, J. M. Moura, and J. Kovačević, "Signal recovery on graphs: Variation minimization," *IEEE Transactions on Signal Processing*, vol. 63, no. 17, pp. 4609–4624, 2015.

[12] E. J. Candes, M. B. Wakin, and S. P. Boyd, "Enhancing sparsity by reweighted 1 minimization," *Journal of Fourier analysis and applications*, vol. 14, no. 5, pp. 877–905, 2008.

[13] I. Pesenson, "Sampling in paley-wiener spaces on combinatorial graphs," *Transactions of the American Mathematical Society*, vol. 360, no. 10, pp. 5603–5627, 2008.

[14] T. K. Moon, "Error correction coding," *Mathematical Methods and Algorithms. John Wiley and Son*, 2005.

[15] G. Rath and C. Guillemot, "Frame-theoretic analysis of dft codes with erasures," *IEEE Transactions on Signal Processing*, vol. 52, no. 2, pp. 447–460, 2004.

[16] N. Perraudin, J. Paratte, D. Shuman, L. Martin, V. Kalofolias, P. Vandergheynst, and D. K. Hammond, "Gspbox: A toolbox for signal processing on graphs," *arXiv preprint arXiv:1408.5781*, 2014.