2024 Asia Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC)

Generalized Graph Signal Sampling under Subspace Priors by Difference-of-Convex Minimization

Keitaro Yamashita*, Kazuki Naganuma*, and Shunsuke Ono*
* Tokyo Institute of Technology, Japan
E-mail: yamashita@mdi.c.titech.ac.jp; naganuma.k.aa@m.titech.ac.jp; ono@c.titech.ac.jp

***Abstract*—This paper proposes an effective approach for sampling graph signals under the subspace prior. Unlike conventional methods that assume bandlimited signals, our method, based on generalized sampling theory, designs a sampling operator suitable for general graph signals beyond bandlimitedness. We formulate a feasibility problem for designing the sampling operator for graph signals under the subspace prior, and then transform it into a difference-of-convex (DC) minimization problem. The DC problem involves a tight relaxation of the intractable rank constraint in the original problem via the nuclear norm. To solve this DC problem, we present an algorithm based on the general double-proximal gradient DC algorithm (GDPGDC), which ensures convergence to a critical point of our DC problem. The effectiveness of our approach is validated by sampling and recovery experiments on various graph signal models.**

I. Introduction

Graph Signal Processing (GSP) is a novel framework for analyzing data represented as signals located at the vertices of a graph. Currently, GSP is a thriving area of research within the signal processing field, covering a wide range of studies from theoretical approaches to practical applications. For a comprehensive overview of GSP, see, for example, [1]–[3].

In the realm of GSP, one of the fundamental technologies is *graph signal sampling* [4], [5]. A significant difference from traditional sampling methods is that graph signal sampling does not have a regular sampling pattern. Consequently, extensive research has been devoted to extending the principles of Shannon-Nyquist sampling theory to accommodate graph signals. Most of the existing sampling methods are based on ensuring bandlimitedness with respect to graph frequencies, as shown in previous works such as [4], [6]–[14]. However, there is a wide variety of graph signals *not* assuming bandlimitedness, which can also be perfectly recovered by extending the *generalized sampling* theory [15], [16].

In generalized sampling, there are two main sampling approaches, one in the *graph frequency domain* and the other in the *graph vertex domain*. For the graph frequency domain, the authors of [17] pioneered the extension of generalized sampling theory to sample and reconstruct graph signals beyond bandlimitedness, which led to branches in the graph frequency domain, one focusing on smoothness and subspace priors [18], and the other on the stochastic prior [19]. While approahes in the graph frequency domain are elegant, they have limitations, such as requiring graph Fourier transforms, so approaches in the vertex domain have also been explored.

For sampling approaches in the graph vertex domain, a vertex-wise sampling method under arbitrary signal priors (subspace, smoothness, and stochastic) is proposed in [20]. However, since the method is based on a greedy algorithm, it may select biased sampling positions and is sensitive to noise. On the other hand, the sampling method proposed in [21], while targeting only the subspace prior, takes an approach of directly designing a sampling operator through convex optimization. This method allows for more efficient and robust sampling because it can generate samples by mixing multiple signal values. However, in order to transform the sampling operator design problem into a convex optimization problem, this method significantly relaxes a certain invertibility constraint involving the sampling operator to be designed. This relaxation is undesirable from the perspective of generalized sampling theory.

A natural question arises: *Can the flexible sampling operator design problem with the constraint be reduced to a tractable optimization problem without significant relaxation?* In this paper, we propose a novel sampling operator design method for graph signals under the subspace prior by utilizing difference-of-convex (DC) optimization techniques, which we aim to design *flexible* sampling operator that mixes the signal values of multiple vertices to create a sampled signal.[1] The key contributions of our study are summarized as follows:

- Formulate a feasibility problem for designing the flexible sampling operator for graph signals under the subspace prior, incorporating a rank constraint derived from the generalized sampling theory.
- Transform the original problem into a tractable DC problem via a tight relaxation of the rank constraint using the nuclear norm.
- Develop an efficient algorithm based on the *general double-proximal gradient DC algorithm* (GDPGDC) [24] to solve the DC problem with guaranteed convergence to a critical point.

Finally, we perform sampling and recovery experiments on different types of graph signals to validate the effectiveness of our method.

[1] In [22], DC-based graph signal sampling for the smoothness prior is also discussed. DC optimization is also attractive to other signal processing applications [23].

II. Preliminaries

A. Notation and Definitions

Bold lowercase and uppercase letters represent vectors and matrices, respectively. We use x_i and X_{ij} to represent the i-th element of a vector $\mathbf{x}$ and the element in the i-th row and j-th column of a matrix $\mathbf{X}$, respectively. The ℓ_2 norm of a vector $\mathbf{x}$ is denoted by $\|\mathbf{x}\|_2 := \sqrt{\sum_i x_i^2}$. The transpose of a matrix $\mathbf{X}$ is denoted as $\mathbf{X}^\top$. The inverse and the pseudo-inverse of a matrix $\mathbf{X}$ are denoted by $\mathbf{X}^{-1}$ and $\mathbf{X}^\dagger$, respectively. The inner product of two matrices $\mathbf{X}$ and $\mathbf{Y}$ is denoted by $\langle \mathbf{X}, \mathbf{Y} \rangle := \sum_i \sum_j X_{ij} Y_{ij}$. The i-th singular value of $\mathbf{X}$ is denoted by $\sigma_i(\mathbf{X})$. The ℓ_1 norm, the Frobenius norm, and the nuclear norm of $\mathbf{X}$ are denoted by $\|\mathbf{X}\|_1 := \sum_i \sum_j |X_{ij}|$, $\|\mathbf{X}\|_F := \sqrt{\langle \mathbf{X}, \mathbf{X} \rangle}$, and $\|\mathbf{X}\|_* = \sum_i \sigma_i(\mathbf{X})$, respectively. A diagonal matrix with $\cdot$ as its principal diagonal is denoted by $\operatorname{diag}(\cdot)$. The indicator function of a set $\mathcal{C}$ is denoted by $\iota_{\mathcal{C}}(\mathbf{X})$ that is defined such that $\iota_{\mathcal{C}}(\mathbf{X}) = 0$ when $\mathbf{X} \in \mathcal{C}$, and $\iota_{\mathcal{C}}(\mathbf{X}) = \infty$ otherwise. The extended real line $\overline{\mathbb{R}}$ is defined as $\overline{\mathbb{R}} := \mathbb{R} \cup \{+\infty, -\infty\}$. An identity matrix is denoted by $\mathbf{I}$, and $\mathbf{I}_N$ represents $N \times N$ identity matrix.

B. Generalized Sampling of Graph Signals

In this section, we briefly outline the generalized sampling theory [5], [16], [18] for graph setting on which our method is based. We consider a weighted undirected graph $\mathcal{G} = \{\mathcal{V}, \mathcal{E}\}$, where $\mathcal{V}$ and $\mathcal{E}$ denote a set of vertices and a set of edges between the vertices, respectively. The number of vertices is denoted as $N = |\mathcal{V}|$. We define an adjacency matrix $\mathbf{E} \in \mathbb{R}^{N \times N}$, where E_{ij} is the weight of the edge between the i-th and j-th vertices. The degree matrix $\mathbf{D} \in \mathbb{R}^{N \times N}$ is a diagonal matrix, where the i-th diagonal element $D_{ii} := \sum_j E_{ij}$ represents the sum of weights connected to vertex i.

We define a graph Laplacian as $\mathbf{L} := \mathbf{D} - \mathbf{E}$ as a graph variation operator for clarity and specificity. Since $\mathbf{L}$ is a real symmetric matrix, it always admits an eigendecomposition $\mathbf{L} = \mathbf{U} \mathbf{\Lambda} \mathbf{U}^\top$, where $\mathbf{U} = [\mathbf{u}_1 \ldots \mathbf{u}_N]$ forms a unitary matrix containing the eigenvectors $\mathbf{u}_i$, and $\mathbf{\Lambda} = \operatorname{diag}(\lambda_1, \ldots, \lambda_N)$ comprises the eigenvalues λ_i. We denote $\mathbf{U}$ and λ_i as the graph Fourier basis and the graph frequency, respectively, and the graph frequency is smaller as λ_i is smaller.

There are two approaches for sampling and recovering: sampling and recovering in the vertex domain [4], [7] and those in the frequency domain [10]. We describe the sampling approach in the vertex domain as our proposal in this paper focuses in it.

Fig. 1 shows the outline of sampling in the vertex domain. Let $\mathbf{x} \in \mathcal{X} \subseteq \mathbb{R}^N$, $\mathbf{c} \in \mathbb{R}^M (M \leq N)$, and $\tilde{\mathbf{x}} \in \mathcal{X}$ be an original graph signal, the sampled signal, and the recovered graph signal, respectively. The graph signal $\mathbf{x}$ undergoes sampling by a sampling operator $\mathbf{S}^\top \in \mathbb{R}^{M \times N}$, i.e., $\mathbf{c} := \mathbf{S}^\top \mathbf{x}$. Subsequently, the sampled signal $\mathbf{c}$ is filtered with a correction operator $\mathbf{H}$ to reduce any errors or distortions introduced during the sampling and recovering process. Following this, it is further filtered by a reconstruction operator $\mathbf{W}$ to map the sampled and corrected signal back onto the original graph. Hence, the recovered signal $\tilde{\mathbf{x}}$ is represented as follows:

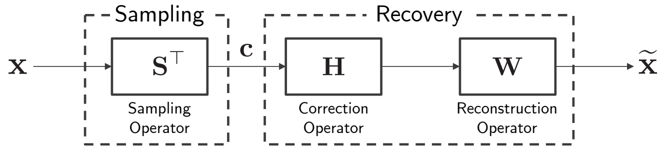


Fig. 1: Framework for generalized sampling of graph signals. Here, $\mathbf{x}$, $\mathbf{c}$, and $\tilde{\mathbf{x}}$ represent the original, sampled, and recovered graph signals, respectively.

$$\tilde{\mathbf{x}} = \mathbf{W}\mathbf{H}\mathbf{c} = \mathbf{W}\mathbf{H}\mathbf{S}^\top \mathbf{x}. \tag{1}$$

The reconstruction operator $\mathbf{W}$ may be constrained, i.e., it may be predefined for some reason, such as computational or hardware constraints. We call the cases where $\mathbf{W}$ is constrained as *predefined case* and the cases where $\mathbf{W}$ is not constrained as *unconstrained case*. The recovery problem entails finding the optimal $\mathbf{H}$ (and $\mathbf{W}$ if not predefined) based on the assumed priors of graph signals. This framework encompasses various situations involving sampling and recovery, including bandlimited graph signals(refer to [5] for more details).

C. Graph Signal Sampling under Subspace Prior

In this paper, we consider sampling and recovering graph signals under the subspace prior [5] for the unconstrained case. We provide the description of the subspace prior and the desigining method for the best $\mathbf{H}$ and $\mathbf{W}$ under the prior from the established results in this subsection.

Under the the subspace prior, we suppose that a graph signal $\mathbf{x} \in \mathbb{R}^N$ is characterized by a linear model as follows:

$$\mathbf{x} := \mathbf{A}\mathbf{d}, \tag{2}$$

where $\mathbf{A} \in \mathbb{R}^{N \times K}$ $(K \leq N)$ is a known generator matrix and $\mathbf{d} \in \mathbb{R}^K$ are expansion coefficients. For simplisity, we consider cases that $M \leq K$ in this paper.

The operators $\mathbf{H}$ and $\mathbf{W}$ are designed based on well-established strategies; the least-squares (LS) strategy and the minimax (MX) strategy. The LS strategy aims to find the recovered signal $\tilde{\mathbf{x}}$ that minimizes the ℓ_2 norm of the difference between the recovered signal $\tilde{\mathbf{x}}$ and the sampled signal $\mathbf{c}$:

$$\tilde{\mathbf{x}}_{\mathrm{LS}} = \underset{\tilde{\mathbf{x}} \in \mathcal{X},\ \mathbf{S}^\top \mathbf{x} = \mathbf{c}}{\operatorname{argmin}} \|\mathbf{S}^\top \tilde{\mathbf{x}} - \mathbf{c}\|_2^2, \tag{3}$$

The MX strategy attempts to directly control the recovery error $\|\tilde{\mathbf{x}} - \mathbf{x}\|_2$ for minimizing the error for the worst feasible signal:

$$\tilde{\mathbf{x}}_{\mathrm{MX}} = \underset{\tilde{\mathbf{x}} \in \mathcal{X}}{\operatorname{argmin}} \max_{\mathbf{S}^\top \mathbf{x} = \mathbf{c}} \|\tilde{\mathbf{x}} - \mathbf{x}\|_2^2. \tag{4}$$

The solutions of the LS and MX strategy are both given by

$$\tilde{\mathbf{x}} = \mathbf{A}(\mathbf{S}^\top \mathbf{A})^\dagger \mathbf{S}^\top \mathbf{x}, \tag{5}$$

followed by the correction operator $\mathbf{H}$ and the reconstruction operator $\mathbf{W}$ as

$$\mathbf{H} = (\mathbf{S}^\top \mathbf{A})^\dagger, \quad \mathbf{W} = \mathbf{A}. \tag{6}$$

Here are three examples of graph signal models assuming the subspace prior:

- Bandlimited (BL) signal [5] is one of the most studied signal models, which are characterized as:

$$\mathbf{x} = \sum_{i=0}^{K-1} d_i \mathbf{u}^{(i)} = \mathbf{U}_{\mathcal{BL}}\mathbf{d}, \tag{7}$$

where $\mathbf{U}_{\mathcal{BL}} \in \mathbb{R}^{N\times K}$ is the submatrix of $\mathbf{U}$ whose rows are extracted within $\mathcal{BL} = \{1, \dots, K\}$. In this case,

$$\mathbf{W}_{\mathrm{BL}} = \mathbf{A}_{\mathrm{BL}} = \mathbf{U}_{\mathcal{BL}}. \tag{8}$$

- Periodic graph spectrum (PGS) signal [18] assumes the periodicity of the graph spectrum as follows:

$$\mathbf{x} = \mathbf{U}A(\mathbf{\Lambda})\mathbf{D}_{\mathrm{samp}}^{\top}\mathbf{d}, \tag{9}$$

where $A(\mathbf{\Lambda}) := \mathrm{diag}(A(\lambda_0), \dots, A(\lambda_{N-1}))$ represents a graph spectral response of the generator, and $\mathbf{D}_{\mathrm{samp}} = [\mathbf{I}_K \ \mathbf{I}_K \ \dots] \in \mathbb{R}^{K\times N}$ is the matrix for the graph Fourier transform domain upsamling. In this case,

$$\mathbf{W}_{\mathrm{PGS}} = \mathbf{A}_{\mathrm{PGS}} = \mathbf{U}A(\mathbf{\Lambda})\mathbf{D}_{\mathrm{samp}}^{\top}. \tag{10}$$

- Piecewise constant (PWC) signal [25] is characterized by constant values in separated vertex regions and are defined as follows with the number of pieces K:

$$\mathbf{x} = \sum_{i=1}^{K} d_i \boldsymbol{\tau}^{(i)} = [\boldsymbol{\tau}^{(1)} \ \dots \ \boldsymbol{\tau}^{(K)}]\mathbf{d}, \tag{11}$$

where $\tau^{(i)}$ for any $i = 1, \dots, K$ is defined as $\tau_n^{(j)} = 1$ when the node j is in the i-th piece and $\tau_j^{(i)} = 0$ otherwise for any $j = 1, \dots, N$. In this case,

$$\mathbf{W}_{\mathrm{PWC}} = \mathbf{A}_{\mathrm{PWC}} = [\boldsymbol{\tau}^{(1)} \ \dots \ \boldsymbol{\tau}^{(K)}]. \tag{12}$$

D. General Double-Proximal Gradient Difference-of-Convex Algorithm

After discussing the recovery of graph signals using the subspace prior based on generalized sampling theory, we now turn our attention to an algorithm designed to solve *difference-of-convex* (DC) minimization problems, which form the basis of the optimization problems we will develop. In this section, we introduce the *General Double-Proximal Gradient Difference-of-Convex* (GDPGDC) algorithm, a flexible tool for solving DC minimization problems.

The GDPGDC [24] can solve DC minimization problems in the form of

$$\min_{\mathbf{X}} f_1(\mathbf{X}) + f_2(\mathbf{X}) - h(\mathbf{Z}) \quad \text{s.t. } \mathbf{Z} = \mathbf{B}\mathbf{X}, \tag{13}$$

where $f_1 : \mathbb{R}^{n\times m} \to \mathbb{R}$ is a differentiable convex function with $1/\beta$-Lipschitz continuous gradient for some $\beta > 0$, $f_2 : \mathbb{R}^{n\times m} \to \overline{\mathbb{R}}$ and $h : \mathbb{R}^{k\times m} \to \overline{\mathbb{R}}$ are proper lower-semi-continuous convex functions, and $\mathbf{B} : \mathbb{R}^{k\times n}$ is a matrix.

We introduce the *proximity operator* of a proper lower-semi-continuous convex function f with a parameter $\gamma > 0$ as follows:

$$\begin{aligned}&\mathrm{prox}_{\gamma f} : \mathbb{R}^{n\times m} \to \mathbb{R}^{n\times m} : \\ &\mathbf{Y} \mapsto \operatorname*{argmin}_{\mathbf{X}} f(\mathbf{X}) + \frac{1}{2\gamma}\|\mathbf{Y} - \mathbf{X}\|_F^2.\end{aligned} \tag{14}$$

Then, Prob. (13) can be solved by GDPGDC by the following procedures: for $\gamma_1 > 0$ and $\gamma_2 > 0$, iterate

$$\left\lfloor \begin{array}{l} \mathbf{X}^{(t+1)} \leftarrow \mathrm{prox}_{\gamma_1 f_2}(\mathbf{X}^{(t)} - \gamma_1(\nabla f_1(\mathbf{X}^{(t)}) - \mathbf{B}^*\mathbf{Z}^{(t)})); \\ \mathbf{Z}^{(t+1)} \leftarrow \mathrm{prox}_{\gamma_2 h^*}(\mathbf{Z}^{(t)} + \gamma_2 \mathbf{B}\mathbf{X}^{(t+1)}); \\ t \leftarrow t+1; \end{array} \right. . \tag{15}$$

Here, the *Fenchel–Rockafellar conjugate function* of h is defined as

$$h^*(\mathbf{X}) := \max_{\mathbf{Y}} \langle \mathbf{X}, \mathbf{Y} \rangle - h(\mathbf{Y}). \tag{16}$$

Thanks to Moreau's Identity[26, Theorem 14.3(ii)], the proximity operator of h^* is calculated with a parameter $\gamma_2 > 0$ as follows:

$$\mathrm{prox}_{\gamma_2 h^*}(\mathbf{X}) = \mathbf{X} - \gamma_2 \mathrm{prox}_{\frac{1}{\gamma_2}h}\left(\frac{1}{\gamma_2}\mathbf{X}\right). \tag{17}$$

We summarize the theoretical results for the convergence of GDPGDC as follows:

Theorem 1 ([24, Proposition 4] Convergence of the sequence generated by GDPGDC)**.** *Let* $\inf\{f_1(\mathbf{X}) + f_2(\mathbf{X}) - h(\mathbf{B}\mathbf{X}) \mid \mathbf{X} \in \mathcal{H}\} > -\infty$*, where* $\mathcal{H}$ *is a Hilbelt space, and let* $0 < \gamma_1 < 2\beta$ *and* $0 < \gamma_2 < +\infty$ *be satisfied. If* $\{\mathbf{X}^{(t)}\}_{t\in\mathbb{N}}$ *and* $\{\mathbf{Z}^{(t)}\}_{t\in\mathbb{N}}$*, generated by Algorithm* (15)*, are bounded, then every cluster point of* $\{\mathbf{X}^{(t)}\}_{t\in\mathbb{N}}$ *is a critical point of Prob.* (13)*.*

III. Proposed Method

The generalized sampling theory works under the assumption of a predefined sampling operator $\mathbf{S}^{\top}$. This assumption, which may be determined by hardware for example, does not provide any direction for the specific design of $\mathbf{S}$. In contrast, when dealing with graph signal sampling, the main challenge revolves around the strategic construction of $\mathbf{S}$. To address this, we formulate a feasibility problem to design the flexible sampling operator for graph signals under the subspace prior, which is then reformulated and relaxed into a DC problem. Finally, we develop an effective GDPGDC-based algorithm to solve this problem.

A. Problem Formulation

Suppose $\mathbf{W}$ and $\mathbf{H}$ are defined as shown in Eq. (6), our current task is to devise an appropr iate $\mathbf{S}$ that is consistent with the sampling and recovery process described in Eq. (1). To achieve the best possible recovery, a promising strategy is to *find* $\mathbf{S}$ *for which the correction operator* $\mathbf{H} = (\mathbf{S}^{\top}\mathbf{A})^{\dagger}$ *has the full column rank, i.e.,* $\mathbf{S}^{\top}\mathbf{A}$ *has the full row rank.*

The question then arises how to design such $\mathbf{S}$. To address this, we first formulate the sampling operator design problem as the following feasibility problem:

$$\text{find } \mathbf{S} \text{ s.t. } \begin{cases} \|\mathbf{S}\|_F \leq \varepsilon, \\ \mathbf{S}^\top \mathbf{A} \text{ has full column rank.} \end{cases} \tag{18}$$

The first constraint serves to control the absolute values of the elements of the sampling operator, ensuring stable sampling within a radius of Frobenius norm $\varepsilon > 0$. As $\mathbf{S}^\top \mathbf{A}$ is $M \times K$ matrix, the second constraint should be

$$\text{rank}(\mathbf{S}^\top \mathbf{A}) = \text{rank}(\mathbf{A}^\top \mathbf{S}) = M. \tag{19}$$

Then, Prob. (18) can be rewritten as follows:

$$\text{find } \mathbf{S} \text{ s.t. } \begin{cases} \mathbf{S} \in \mathcal{B}_{F,\varepsilon} := \{\mathbf{X} \mid \|\mathbf{X}\|_F \leq \varepsilon\}, \\ \text{rank}(\mathbf{A}^\top \mathbf{S}) = M. \end{cases} \tag{20}$$

Prob. (20) is still a challenge because of the rank constraint. To overcome this, as it is known that the nuclear norm of a matrix is the tightest convex envelope of the rank function [27], we relax the problem into a constrained convex *maximization* problem as follows:

$$\max_{\mathbf{S}} \|\mathbf{A}^\top \mathbf{S}\|_* \text{ s.t. } \mathbf{S} \in \mathcal{B}_{F,\varepsilon}. \tag{21}$$

By introducing the indicator function of $\mathcal{B}_{F,\varepsilon}$, i.e. $\iota_{\mathcal{B}_{F,\varepsilon}}$, and reversing maximization to minimization, we can reformulate Prob. (21) as depicted as follows:

$$\min_{\mathbf{S}} \iota_{\mathcal{B}_{F,\varepsilon}}(\mathbf{S}) - \|\mathbf{A}^\top \mathbf{S}\|_*. \tag{22}$$

Here, $\iota_{\mathcal{B}_{F,\varepsilon}}$ is a proper lower-semicontinuous convex function as $\mathcal{B}_{F,\varepsilon}$ is a nonempty closed convex set, and $\|\mathbf{A}^\top \mathbf{S}\|_*$ is also a proper lower-semicontinuous convex function. Thus, this problem is regarded as the minimization of the difference between two convex functions and is reduced to Prob. (13).

B. Optimization

The algorithmic procedure for solving Prob. (22) is summarized in Algorithm 1, where $f_1(\mathbf{X}) = 0$, $f_2(\mathbf{X}) = \iota_{\mathcal{B}_{F,\varepsilon}}(\mathbf{X})$, $h(\mathbf{Z}) = \|\mathbf{Z}\|_*$, $\mathbf{X} = \mathbf{S}$, and $\mathbf{B} = \mathbf{A}^\top$ in (13).

In what follows, we derive specific computations of each step of the algorithm. Since $\mathcal{B}_{F,\varepsilon}$ in the step 2 of the Algorithm 1 is a nonempty closed convex set, the proximity operator of its indicator function is equal to the metric projection[2] onto $\mathcal{B}_{F,\varepsilon}$, i.e.,

$$\text{prox}_{\iota_{\mathcal{B}_{F,\varepsilon}}}(\mathbf{X}) = \mathcal{P}_{\mathcal{B}_{F,\varepsilon}}(\mathbf{X}) = \begin{cases} \mathbf{X}, & \text{if } \mathbf{X} \in \mathcal{B}_{F,\varepsilon}; \\ \frac{\varepsilon \mathbf{X}}{\|\mathbf{X}\|_F}, & \text{otherwise.} \end{cases} \tag{23}$$

The proximity operator of h^* in the step 3 in Algorithm 1 is calculated by (17), and the proximity operator for the nuclear norm $\|\cdot\|_*$ is calculated by

$$\text{prox}_{\gamma\|\cdot\|_*}(\mathbf{X}) = \mathbf{U}_\mathbf{X} \mathcal{S}_\gamma(\mathbf{\Sigma}_\mathbf{X}) \mathbf{V}_\mathbf{X}^\top, \tag{24}$$

where $\mathbf{X} = \mathbf{U}_\mathbf{X} \mathbf{\Sigma}_\mathbf{X} \mathbf{V}_\mathbf{X}^\top$ is the SVD of $\mathbf{X}$, and $\mathcal{S}_\gamma(\cdot)$ is the soft-thresholding operator applied to the singluar values, defined as:

$$\mathcal{S}_\gamma(\sigma_i) = \max(\sigma_i - \gamma, 0), \tag{25}$$

and $\mathcal{S}_\gamma(\mathbf{\Sigma}_\mathbf{X})$ is a diagonal matrix with entries $\mathcal{S}_\gamma(\sigma_i(\mathbf{X}))$, i.e., $\mathcal{S}_\gamma(\mathbf{\Sigma}_\mathbf{X}) = \text{diag}(\mathcal{S}_\gamma(\sigma_i(\mathbf{X})))$. With (17) and (24), the step 3 in Algorithm 1 can be computed.

[2]The proximity operator of the indicator function of a nonempty closed convex set C is identical to the metric projection onto C, denoted by $\mathcal{P}_C$ [28].

Algorithm 1 Proposed sampling operator design algorithm

Input: $\mathbf{S}^{(0)}, \mathbf{Z}^{(0)}, \varepsilon > 0, \gamma_1 > 0, \gamma_2 > 0$
1: **while** until a stopping criterion is satisfied **do**
2: $\quad \mathbf{S}^{(t+1)} \leftarrow \mathcal{P}_{B_{F,\varepsilon}}\left(\mathbf{S}^{(t)} + \gamma_1 \mathbf{A}\mathbf{Z}^{(t)}\right)$ by (23);
3: $\quad \mathbf{Z}^{(t+1)} \leftarrow \text{prox}_{\gamma_2 h^*}\left(\mathbf{Z}^{(t)} + \gamma_2 \mathbf{A}^\top \mathbf{S}^{(t+1)}\right)$
by (17) and (24);
4: $\quad t \leftarrow t + 1$;
5: **end while**
Output: $\mathbf{S}^{(t)}$

Remark 1 (Convergence of Prob. (22))**.** Notice that $\mathcal{P}_{\mathcal{B}_{F,\varepsilon}}$ always returns a variable whose Frobenius norm is less than ε. This indicates that the sequence $\{\mathbf{S}^{(t)}\}_{t\in\mathbb{N}}$ generated by Algorithm 1 is bounded. Therefore, from Theorem 1, $\{\mathbf{S}^{(t)}\}_{t\in\mathbb{N}}$ is guaranteed to converge to a critical point of Prob. (22).

IV. Experiments

We demonstrate the effectiveness of our method through sampling and recovery experiments on several types of graph signals. All experiments were conducted using MATLAB (R2024a) on a Windows 11 computer with Intel Core i9-12900 3.19-GHz processor, 32 GB of RAM, and NVIDIA GeForce RTX 3090. Our method is compared with the following graph signal sampling methods: NLPD [4], SP [7], AVM [14], SASB [21], and GSSS [20]. The NLPD, SP, and AVM are methods for bandlimited graph signals, while SASB and GSSS are applicable to graph signals under the subspace prior.

A. Setup

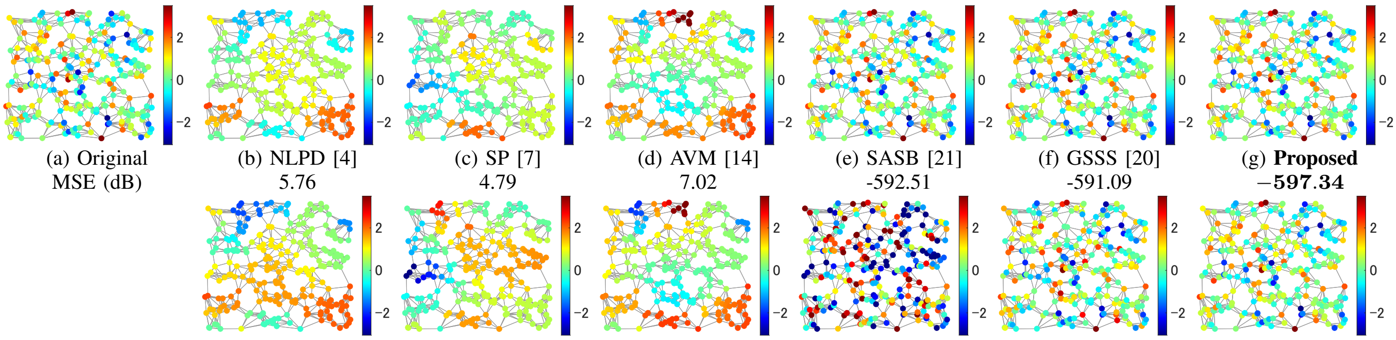
We constructed random sensor graphs consisting of $N = 256$ vertices by using GSPBox [29]. The size of the sampled signal was set to $M = 16$. We have generated the following types of graph signals:

- Bandlimited (BL) signals (7) with $K = 16$;
- Periodic graph spectrum (PGS) signals (9) with following [18] that the graph spectral response of the generator $\mathbf{A}$ was set to $A(\lambda_i) := \exp(-1.5\lambda_i/\lambda_{\max})$ for any $i = 1, \ldots, K$, where λ_i and $\lambda_{\max}$ are the i-th graph frequency and the largest graph frequency, respectively, and $K = 16$;
- Piecewise constant (PWC) signals (11) with $K = 16$.

For all types above, we set the expansion coefficients $\mathbf{d}$ as their element $d_i \sim \mathcal{N}(1, 1)$ for all $i = 1, \ldots, K$, and we used the generator matrix $\mathbf{A}$ and the reconstruction matrix $\mathbf{W}$ as presented in (8), (10), and (12) for each type of signals. We also experimented with noisy sampled signal $\mathbf{y} := \mathbf{c} + \boldsymbol{\eta}$, where the noise $\boldsymbol{\eta} \in \mathbb{R}^M$ is generated as a white Gaussian noise with its variance $\sigma^2 = 0.1$, for all types of signals. In this case, we recovered from $\mathbf{y}$, i.e., $\tilde{\mathbf{x}} = \mathbf{W}\mathbf{H}\mathbf{y} = \mathbf{W}\mathbf{H}(\mathbf{S}^\top \mathbf{x} + \boldsymbol{\eta})$.

TABLE I: Average MSEs in Decibel of the Recoveries for 20 Independent Runs.

Signal Model	Method					
	NLPD [4]	SP [7]	AVM [14]	SASB [21]	GSSS [20]	**Proposed**
BL	-607.38	-51.80	-44.44	-316.55	-605.38	$\mathbf{-607.58}$
BL + noise	-13.14	-15.03	-2.61	13.44	-12.73	$\mathbf{-46.74}$
PGS	3.53	2.30	24.69	-578.67	-586.66	$\mathbf{-591.99}$
PGS + noise	5.29	3.29	23.91	18.17	-12.39	$\mathbf{-46.48}$
PWC	-5.67	0.06	6.82	-590.25	-591.17	$\mathbf{-591.73}$
PWC + noise	-2.91	1.76	14.30	-11.08	-12.85	$\mathbf{-46.22}$



(a) Original MSE (dB) | (b) NLPD [4] 5.76 | (c) SP [7] 4.79 | (d) AVM [14] 7.02 | (e) SASB [21] -592.51 | (f) GSSS [20] -591.09 | (g) **Proposed** $\mathbf{-597.34}$

(h) NLPD [4] 8.07 | (i) SP [7] 7.39 | (j) AVM [14] 7.68 | (k) SASB [21] 15.79 | (l) GSSS [20] -14.29 | (m) **Proposed** $\mathbf{-48.15}$

Fig. 2: An example of PGS signal and recovered graph signals on a random sensor graph ($N = 256, M = 16$). Fig. (a) shows the original PGS signal. Fig. (b)-(g) show the recovered signals from the sampled signal without noise under each method. Fig. (h)-(m) show the recovered signals from the noisy sampled signal under each method. The color of each vertex indicates the magnitude of the signal value.

The parameter ε in Prob. (22) was set to $\sqrt{NM}/8$, and γ_1 and γ_2 in our algorithm were set to $\gamma_1 = \gamma_2 = 0.001$. We defined $\mathbf{S}^{(0)}$ as a matrix with random Gaussian entries and the stopping criterion as $\|\mathbf{S}^{(t+1)} - \mathbf{S}^{(t)}\|_F / \|\mathbf{S}^{(t)}\|_F \leq 10^{-5}$. The existing methods used parameters as described in their papers.

For the quantitative evaluations, we used the mean squared error (MSE): $\mathrm{MSE} = \|\tilde{\mathbf{x}} - \mathbf{x}\|_2^2 / N$.

B. Results and Discussion

Table I presents the averaged MSEs in decibel obtained from 20 independent runs. The results in the table are expressed in decibels, with lower numbers indicating better results.

Notably, the average MSEs for NLPD, SP, and AVM assuming bandlimitedness, were lower compared to that of our method, especially for the non-bandlimited signals, PGS and PWC signals. Since SASB and GSSS have the ability to sample graph signals assuming the subspace prior, their results were better than those of the three methods, especially for the PGS and PWC signals without the noise. However, our method outperformed SASB and GSSS in all cases. In particular, it can be seen that the recovery results from noisy sampled signals were particularly superior. This difference can be attributed to the limitation of GSSS, which restricts its strategy to local sampling, as opposed to our method, which allows the integration of non-local signal values into the sampling. In other words, the samples generated by our method tend to preserve more information of the original signal. The comparison with SASB also shows that our method was able to relax the formulation more appropriately.

Fig. 2 visualizes an example of PGS signal and recovered graph signals on a random sensor graph. The distribution of colors on the graphs shown in the figure also visually confirms that our method can recover the graph signal closer to the original signal than the other methods.

V. Conclusion

In this paper, we addressed the challenge of designing a sampling operator for graph signals under the subspace prior beyond bandlimitedness. Our approach involved formulating a feasibility problem based on generalized sampling theory to determine the optimal sampling operator. To deal with the rank constraint, we reformulated the problem in a DC form and developed an algorithm based on GDPGDC that guarantees convergence to a critical point.

In contrast to prevailing methods that select individual vertices for sampling, our method introduced the novel concept of combining non-local signal values at multiple vertices to generate a sampled signal. This innovation gave our approach a degree of flexibility and efficiency not found in traditional techniques. To validate our approach, we performed sampling and recovery experiments on various graph signal models. The results confirmed that our method consistently outperformed existing methods in terms of recovery accuracy.

Acknowledgment

This work was supported in part by JST PRESTO under Grant JPMJPR21C4, JST AdCORP under Grant JPMJKB2307 and JST ACT-X Grant JPMJAX23CJ, in part by JSPS KAKENHI under Grant 22H03610, 22H00512, 23H01415 and 23K17461, and in part by Grant-in-Aid for JSPS Fellows under Grant 23KJ0912.

References

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

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

[3] A. Ortega, P. Frossard, J. Kovačević, J. M. F. Moura, and P. Vandergheynst, "Graph signal processing: Overview, challenges, and applications," *Proceedings of the IEEE*, vol. 106, no. 5, pp. 808–828, 2018.

[4] 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.

[5] Y. Tanaka, Y. C. Eldar, A. Ortega, and G. Cheung, "Sampling signals on graphs: From theory to applications," *IEEE Signal Process. Mag.*, vol. 37, no. 6, pp. 14–30, 2020.

[6] 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, 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. Valsesia, G. Fracastoro, and E. Magli, "Sampling of graph signals via randomized local aggregations," *IEEE Trans. Signal Inf. Process. Netw.*, vol. 5, no. 2, pp. 348–359, 2018.

[10] Y. Tanaka, "Spectral domain sampling of graph signals," *IEEE Trans. Signal Process.*, vol. 66, no. 14, pp. 3752–3767, 2018.

[11] G. Puy, N. Tremblay, R. Gribonval, and P. Vandergheynst, "Random sampling of bandlimited signals on graphs," *Appl. Comput. Harmon. Anal.*, vol. 44, no. 2, pp. 446–475, 2018.

[12] A. Sakiyama, Y. Tanaka, T. Tanaka, and A. Ortega, "Eigendecomposition-free sampling set selection for graph signals," *IEEE Trans. Signal Process.*, vol. 67, no. 10, pp. 2679–2692, 2019.

[13] Y. Bai, F. Wang, G. Cheung, Y. Nakatsukasa, and W. Gao, "Fast graph sampling set selection using gershgorin disc alignment," *IEEE Trans. Signal Process.*, vol. 68, pp. 2419–2434, 2020.

[14] A. Jayawant and A. Ortega, "Practical graph signal sampling with log-linear size scaling," *Signal Process.*, vol. 194, p. 108 436, 2022.

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

[16] Y. C. Eldar, *Sampling theory: Beyond bandlimited systems*. Cambridge University Press, 2015.

[17] S. P. Chepuri, Y. C. Eldar, and G. Leus, "Graph sampling with and without input priors," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, 2018, pp. 4564–4568.

[18] Y. Tanaka and Y. C. Eldar, "Generalized sampling on graphs with subspace and smoothness priors," *IEEE Trans. Signal Process.*, vol. 68, pp. 2272–2286, 2020.

[19] J. Hara, Y. Tanaka, and Y. C. Eldar, "Graph signal sampling under stochastic priors," *IEEE Trans. Signal Process.*, vol. 71, pp. 1421–1434, 2023.

[20] J. Hara and Y. Tanaka, "Sampling set selection for graph signals under arbitrary signal priors," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, 2022, pp. 5732–5736.

[21] J. Hara, K. Yamada, S. Ono, and Y. Tanaka, "Design of graph signal sampling matrices for arbitrary signal subspaces," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, 2021, pp. 5275–5279.

[22] S. Ono, N. Naganuma, and K. Yamashita, *Graph signal sampling under smoothness priors: A difference-of-convex approach*, *arXiv:2306.14634*, [Online]. Available: https://arxiv.org/abs/2306.14634, 2023. arXiv: 2306.14634 [eess.IV]. [Online]. Available: https://arxiv.org/abs/2306.14634.

[23] K. Sato, K. Naganuma, and S. Ono, "Enhancing hyperspectral anomaly detection by difference-of-convex sparse anomaly modeling," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, IEEE, 2024, pp. 9921–9925.

[24] S. Banert and R. I. Boţ, "A general double-proximal gradient algorithm for dc programming," *Mathematical programming*, vol. 178, no. 1-2, pp. 301–326, 2019.

[25] S. Chen, R. Varma, A. Singh, and J. Kovačević, "Representations of piecewise smooth signals on graphs," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, IEEE, 2016, pp. 6370–6374.

[26] H. H. Bauschke and P. L. Combettes, *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*. 2nd ed. Cham: Springer International Publishing AG, 2017.

[27] M. Fazel, "Matrix rank minimization with applications," Ph.D. dissertation, PhD thesis, Stanford University, 2002.

[28] H. H. Bauschke and P. L. Combettes, *Convex analysis and monotone operator theory in Hilbert spaces*. Springer, 2011.

[29] N. Perraudin, J. Paratte, D. Shuman, *et al.*, "GSPBOX: A toolbox for signal processing on graphs," *arXiv:1408.5781*, 2014, [Online]. Available: https://arxiv.org/abs/1408.5781.