

# Graph Signal Sampling Under Stochastic Priors

Junya Hara, *Student Member, IEEE*, Yuichi Tanaka, *Senior Member, IEEE*, and Yonina C. Eldar, *Fellow, IEEE*

**Abstract**—We propose a generalized sampling framework for stochastic graph signals. Stochastic graph signals are characterized by graph wide sense stationarity (GWSS) which is an extension of wide sense stationarity (WSS) for standard time-domain signals. In this paper, graph signals are assumed to satisfy the GWSS conditions and we study their sampling as well as recovery procedures. In generalized sampling, a correction filter is inserted between sampling and reconstruction operators to compensate for non-ideal measurements. We propose a design method for the correction filters to reduce the mean-squared error (MSE) between original and reconstructed graph signals. We derive the correction filters for two cases: The reconstruction filter is arbitrarily chosen or predefined. The proposed framework allows for arbitrary sampling methods, i.e., sampling in the vertex or graph frequency domain. We also show that the graph spectral response of the resulting correction filter parallels that for generalized sampling for WSS signals if sampling is performed in the graph frequency domain. Furthermore, we reveal the theoretical connection between the proposed and existing correction filters. The effectiveness of our approach is validated via experiments by comparing its MSE with existing approaches.

**Index Terms**—Generalized sampling theory, stochastic prior, graph wide sense stationarity, graph Wiener filter.

## I. INTRODUCTION

Graph signal processing (GSP) is a developing field in signal processing [2], [3]. Applications of GSP are extensive, including learning of graphs [4]–[8], restoration of graph signals [9]–[12], image/point cloud processing [13], [14], and graph neural networks [15]. Recent interest in GSP is to extend classical signal processing theory to the graph setting [16]–[20]. One of the main differences between standard signal processing and GSP is that GSP systems are not shift-invariant (SI) in general. This leads to the challenge that GSP systems may have different definitions in the vertex and spectral (graph frequency) domains: They do not coincide in general. Such examples include sampling [21]–[29], translation [30]–[34], and filtering [35].

In this paper, we focus on graph signal sampling. Sampling theory for graph signals has been widely studied in GSP [21]–[28], [36], [37]. There are many promising applications of graph signal sampling, such as sensor placement [27], filter bank designs [38]–[41], traffic monitoring [42], and semi-supervised learning [43], [44]. Most works focus on building parallels of the Shannon-Nyquist theorem [45], [46] and its generalizations [47], [48] to the graph setting. Therefore, sampling of bandlimited graph signals has been widely studied [22], [23], [26], [28]. Other graph signal subspaces are

also useful for practical applications: For example, piecewise smooth graph signals and periodic graph spectrum signals have been considered [49]–[52].

Since sampling of deterministic signals has been well studied in the literature of sampling in Hilbert space [47], [48], [53], [54], we can immediately derive its GSP counterpart. In contrast, sampling of *random graph signals* has not been considered so far because existing (generalized) sampling methods for graph signals have mainly focused on deterministic signal models [17], [21]–[28], [55].

In this paper, we consider a graph signal sampling framework for random graph signals. Random graph signals are modeled by graph wide sense stationarity (GWSS), which is a counterpart of wide sense stationarity (WSS) of standard signals. WSS is characterized by the statistical moments, i.e., mean and covariance, that are invariant to shift. As previously mentioned, graph signals do not lie in SI spaces in general. Hence, existing definitions of GWSS are based on the spectral characteristics of signals [32], [56], [57], which are extensions of the power spectral density (PSD) of the standard WSS. However, existing definitions of GWSS require additional assumptions that are not necessary for WSS. Therefore, we first define GWSS as a natural extension of WSS based on signal modulation.

Subsequently, we develop a generalized sampling framework to best recover GWSS signals. Our framework parallels generalized sampling of SI signals [58]. It consists of sampling, correction, and reconstruction transforms. The correction transform is inserted between the sampling and reconstruction transforms to compensate for non-ideal measurements. We derive the correction transform that minimizes the mean-squared error (MSE). Our framework can be applied to any sampling method that is linear. In other words, both vertex and graph frequency domain sampling methods [22], [23], [26]–[29] are applicable without changing the framework.

Existing works of (generalized) sampling theory on graphs [17], [55] have been studied only for deterministic signal models. In contrast, our generalized sampling is designed for random graph signals as described in Table I. Interestingly, our solution parallels that in the SI setting [58] when sampling is performed in the graph frequency domain. Moreover, we reveal that the existing signal recovery under different priors [17] is special cases of our proposed recovery. Experiments for synthetic signals validate that our proposed recovery is effective for stochastic graph signals with sampling in both vertex and spectral domains.

The remainder of this paper is organized as follows. Section II reviews generalized sampling for time-domain signals. For WSS signals, we introduce the standard Wiener filter. In Section III, generalized sampling framework for stochastic graph signals is introduced. In addition, GWSS is defined as

Preliminary results of this work was presented in [1].

JH and YT are supported in part by JST PRESTO (JPMJPR1656) and JSPS KAKENHI (19K22864). YE has received funding from European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation program (grants no. 278025,677909,681514).

TABLE I  
COMPARISON OF GENERALIZED GRAPH SIGNAL SAMPLING.

Signal priors	SI spaces	Graph subspaces
Subspace	[47]	[17], [55]
Smoothness	[47]	[17], [55]
Stochastic	[58]	<b>This work</b>

an extension of WSS. Section IV derives graph Wiener filters for recovery of a GWSS process based on the minimization of the MSE between the original and reconstructed graph signals. In the special case of graph frequency domain sampling, the graph frequency response parallels that of sampling in SI spaces. In Section V, we discuss the relationship among the proposed Wiener filter and existing generalized graph signal sampling. Signal recovery experiments for synthetic and real-world signals are demonstrated in Section VI. Section VII concludes the paper.

*Notation:* We consider a weighted undirected graph  $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ , where  $\mathcal{V}$  and  $\mathcal{E}$  represent the sets of vertices and edges, respectively. The number of vertices is  $N = |\mathcal{V}|$  unless otherwise specified. The adjacency matrix of  $\mathcal{G}$  is denoted by  $\mathbf{A}$  where its  $(m, n)$ -element  $a_{mn} \geq 0$  is the edge weight between the  $m$ th and  $n$ th vertices;  $a_{mn} = 0$  for unconnected vertices. The degree matrix  $\mathbf{D}$  is defined as  $\mathbf{D} = \text{diag}(d_0, d_1, \dots, d_{N-1})$ , where  $d_m = \sum_n a_{mn}$  is the  $m$ th diagonal element. We use a graph Laplacian  $\mathbf{L} := \mathbf{D} - \mathbf{A}$  as a graph variation operator. A graph signal  $\mathbf{x} \in \mathbb{C}^N$  is defined as a mapping from the vertex set to the set of complex numbers, i.e.,  $\mathbf{x} : \mathcal{V} \rightarrow \mathbb{C}$ .

The graph Fourier transform (GFT) of  $\mathbf{x}$  is defined as

$$\hat{x}(\lambda_i) = \langle \mathbf{u}_i, \mathbf{x} \rangle = \sum_{n=0}^{N-1} u_i[n] x[n], \quad (1)$$

where  $\mathbf{u}_i$  is the  $i$ th column of a unitary matrix  $\mathbf{U}$  and it is obtained by the eigenvalue decomposition of the graph Laplacian  $\mathbf{L} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^*$  with the eigenvalue matrix  $\mathbf{\Lambda} = \text{diag}(\lambda_0, \lambda_1, \dots, \lambda_{N-1})$ . We refer to  $\lambda_i$  as a *graph frequency*. We let  $[\cdot]_{n,k}$  and  $(\cdot)_n$  denote the  $(n, k)$ -element in the matrix and  $n$ -th element in the vector, respectively,  $(*)$ ,  $\langle \cdot, \cdot \rangle$  and  $\text{tr}(\cdot)$  denote the convolution, the inner product between two vectors and trace of a matrix, respectively, and  $(\circ)$  represents the Hadamard (elementwise) product.

For time-domain signals,  $x(t)$  and  $x[n]$  denote a continuous-time signal and discrete-time signal, respectively. The continuous-time Fourier transform (CTFT) of a signal  $x(t) \in L_2$  is denoted by  $X(\omega)$  and the discrete-time Fourier transform (DTFT) of a sequence  $x[n] \in \ell_2$  is denoted by  $X(e^{j\omega})$ .

## II. GENERALIZED SAMPLING IN SI SPACES

Generalized sampling for stochastic standard signals is reviewed in this section [58]. Detailed derivations and discussions can be found in [47], [58]. The Wiener filter introduced in this section parallels that in our graph signal sampling framework.

### A. Sampling and recovery framework for time-domain signals

We first review standard generalized sampling for time-domain signals [47]. The framework is depicted in Fig. 1.

Let  $x(t)$  be a continuous-time signal and  $\eta[n]$  be stationary noise. We consider samples  $c[n]$  at  $t = nT$  of a filtered signal of  $x(t)$ ,  $c(t) = x(t) * s(-t)$ , where  $s(-t)$  denotes a sampling filter. The DTFT of the samples,  $C(e^{j\omega})$ , can be written as

$$C(e^{j\omega}) = \frac{1}{T} \sum_{k \in \mathbb{Z}} S^* \left( \frac{\omega - 2\pi k}{T} \right) X \left( \frac{\omega - 2\pi k}{T} \right), \quad (2)$$

where  $S(\omega)$  is the CTFT of  $s(t) \in L_2$ . The measured samples are corrupted by noise and are given by

$$y[n] = \langle s(t - nT), x(t) \rangle + \eta[n]. \quad (3)$$

The reconstructed signal is constrained to lie in a SI space  $\mathcal{W}$ , spanned by the shifts of a reconstruction kernel  $w(t) \in L_2$ , i.e.,

$$\tilde{x}(t) = \sum_{n \in \mathbb{Z}} d[n] w(t - nT), \quad (4)$$

where  $d[n] \in \ell_2$  are unknown expansion coefficients, which yield the best possible recovery  $\tilde{x}(t)$ . To obtain an appropriate  $d[n]$ , we apply a digital correction filter  $h[n]$  to  $y[n]$ , i.e.,  $d[n] = (h * y)[n]$ . The CTFT of  $\tilde{x}(t)$ ,  $\tilde{X}(\omega)$ , can be expressed as

$$\tilde{X}(\omega) = D(e^{j\omega T}) W(\omega), \quad (5)$$

where  $D(e^{j\omega})$  is the DTFT of  $d[n]$  and  $W(\omega)$  is the CTFT of  $w(t)$ .

In practice, sampling and reconstruction filters, i.e.,  $s(t)$  and  $w(t)$ , may be predefined prior to sampling based on technical requirements. For generalized sampling, instead, we assume that we can design the correction filter  $h[n]$  freely. The best correction filter  $h[n]$  is designed such that the input signal  $x(t)$  and the reconstructed signal  $\tilde{x}(t)$  are close enough in some metric. This is a widely-accepted sampling framework for time-domain signals and we follow it here for the graph setting [58], [59].

### B. Wiener filter

Next, suppose that  $x(t)$  is a zero-mean WSS process with known power spectral density (PSD)  $\Gamma_x(\omega)$  and  $\eta[n]$  is a zero-mean WSS noise process with known PSD  $\Gamma_\eta(e^{j\omega})$ , independent of  $x(t)$ . Detailed definitions of WSS are given in Appendix A.

The optimal correction filter proposed in [58] is obtained by solving the following problem:

$$\min_{h[n]} \mathbb{E} \left\{ |\tilde{x}(t) - P_{\mathcal{W}} x(t)|^2 \right\}, \quad (6)$$

where  $P_{\mathcal{W}}$  is an orthogonal projection onto the reconstruction subspace  $\mathcal{W}$ . Since the recovered signal  $\tilde{x}(t)$  is constrained by (4) to lie in  $\mathcal{W}$ , we want it to best approximate the orthogonal projection of  $x(t)$  onto that same space  $\mathcal{W}$  [58]. The frequency response of the solution of (6) is given by [58].

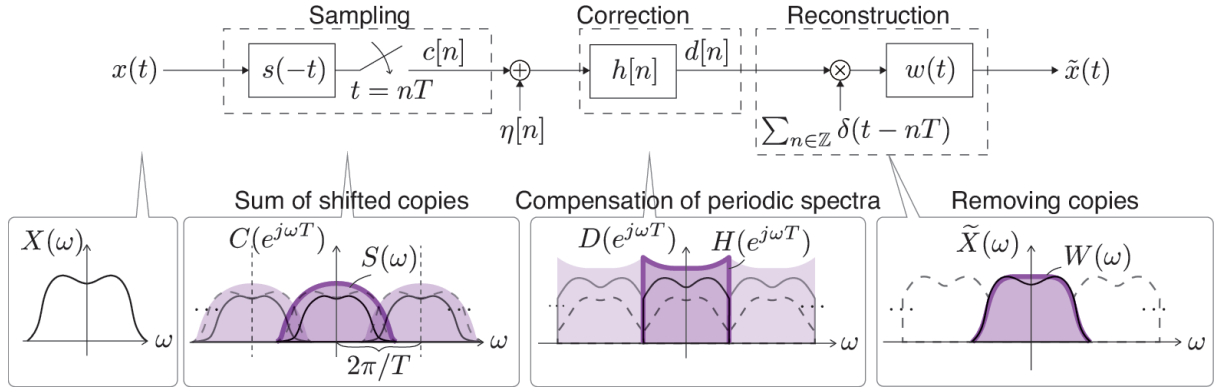


Fig. 1. Generalized sampling framework in SI spaces and its counterpart in the Fourier domain. Dashed-lined and solid-lined areas in the bottom represent input and filtered spectra, respectively. Colored areas represent spectral responses of each filter.

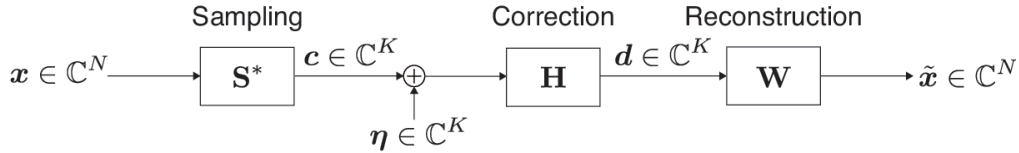


Fig. 2. Generalized sampling framework for graph signals.

$$H_{\text{PRE}}(e^{j\omega}) = \frac{\sum_{k \in \mathbb{Z}} \Gamma_x\left(\frac{\omega}{T} + \frac{2\pi k}{T}\right) S\left(\frac{\omega}{T} + \frac{2\pi k}{T}\right) W^*\left(\frac{\omega}{T} + \frac{2\pi k}{T}\right)}{R_W(e^{j\omega}) \left( \Gamma_\eta(e^{j\omega}) + \sum_{k \in \mathbb{Z}} \Gamma_x\left(\frac{\omega}{T} + \frac{2\pi k}{T}\right) |S\left(\frac{\omega}{T} + \frac{2\pi k}{T}\right)|^2 \right)}, \quad (7)$$

where  $R_W(e^{j\omega}) = \sum_{k \in \mathbb{Z}} |W\left(\frac{\omega}{T} + \frac{2\pi k}{T}\right)|^2$ .

When the reconstruction filter is unconstrained, we can optimize the correction filter jointly with the reconstruction filter. The optimal filter in the unconstrained setting is a special case of (7), where  $w(t)$  is optimally chosen. In order to obtain the unconstrained solution, we thus solve (6) with respect to  $w(t)$  by substituting (7) for (6) [47]. The resulting optimal correction filter in the unconstrained case is given by

$$H_{\text{UNC}}(e^{j\omega}) = \frac{1}{\Gamma_\eta(e^{j\omega}) + \sum_{k \in \mathbb{Z}} \Gamma_x\left(\frac{\omega}{T} + \frac{2\pi k}{T}\right) |S\left(\frac{\omega}{T} + \frac{2\pi k}{T}\right)|^2}, \quad (8)$$

with the reconstruction filter  $W(\omega) = \Gamma_x(\omega)S(\omega)$ .

In Section III, we will introduce the sampling framework for GWSS, which is a counterpart to what we describe in this section.

### III. SAMPLING OF GRAPH SIGNALS AND GRAPH WIDE SENSE STATIONARITY

In this section, we introduce the sampling and recovery framework used throughout the paper. Then, two representative sampling methods for graph signals are reviewed. We also formally define GWSS as an analog of WSS.

#### A. Sampling and recovery framework for graph signals

We begin by reviewing generalized graph signal sampling [1], [17], illustrated in Fig. 2. The main differences among

various methods stem from the sampling domain and signal generation model.

Let  $c \in \mathbb{C}^K$  ( $K \leq N$ ) be the sampled graph signal. When  $\mathbf{S}^*$  is the sampling transformation,  $c = \mathbf{S}^*x$ , we would like to recover  $x$  from the noisy samples  $y = c + \eta$  using a reconstruction transformation  $\mathbf{W} \in \mathbb{C}^{N \times K}$ . The recovered signal is then

$$\tilde{x} = \mathbf{W}\mathbf{H}y = \mathbf{W}\mathbf{H}(c + \eta) = \mathbf{W}\mathbf{H}(\mathbf{S}^*x + \eta), \quad (9)$$

where  $\mathbf{H} \in \mathbb{C}^{K \times K}$  is some transformation that compensates for non-ideal measurements. It is an analog of the generalized sampling framework introduced in Section II-A: We use three filters  $\mathbf{S}$ ,  $\mathbf{H}$ , and  $\mathbf{W}$ . In the following, we seek the best  $\mathbf{H}$  so that  $\tilde{x}$  is close to  $x$  in some sense under appropriate signal assumptions [58], [59].

#### B. Sampling Methods

Here, we introduce representative graph signal sampling operators in the vertex and graph frequency domains.

1) *Sampling in Vertex Domain*: Vertex domain sampling is expressed as the selection of samples on a vertex subset. Therefore, it corresponds to nonuniform sampling in the time domain.

Sampling in the vertex domain is defined as follows:

**Definition 1** (Sampling in the vertex domain). *Let  $x \in \mathbb{C}^N$  be the original graph signal and  $\mathbf{G} \in \mathbb{C}^{N \times N}$  be an arbitrary graph filter. In addition, let  $\mathbf{I}_{\mathcal{M}} \in \{0, 1\}^{K \times N}$  be a submatrix of the identity matrix  $\mathbf{I}_N$  extracting  $K = |\mathcal{M}|$  rows corresponding to the sampling set  $\mathcal{M}$ . The sampled graph signal  $c \in \mathbb{C}^K$  is given as follows:*

$$c = \mathbf{I}_{\mathcal{M}}\mathbf{G}x. \quad (10)$$

The sampling matrix is then expressed as  $\mathbf{S}^* = \mathbf{I}_{\mathcal{M}}\mathbf{G}$ . In contrast to time domain signals,  $\mathbf{I}_{\mathcal{M}}$  may depend on the graph

since local connectivity of the graph is often irregular. This implies that there may exist a “best” vertex set for graph signal sampling described in [21].

2) *Sampling in the GFT Domain*: When sampling SI signals, the spectrum folding phenomenon occurs [47]. Graph frequency domain sampling utilizes the behavior in the graph spectrum.

Formally, graph frequency domain sampling, which is a counterpart of (2), is defined as follows:

**Definition 2** (Sampling in the graph frequency domain [29]). *Let  $\hat{x} \in \mathbb{C}^N$  be the original signal in the graph frequency domain, i.e.,  $\hat{x} = \mathbf{U}^* \mathbf{x}$ , and let  $S(\lambda_i)$  be an arbitrary sampling filter defined in the graph frequency domain. For any sampling ratio  $M \in \mathbb{Z}^1$ , the sampled graph signal in the graph frequency domain is given by  $\hat{c} \in \mathbb{C}^K$ , where  $K = N/M$ , and*

$$\hat{c}(\lambda_i) = \sum_{l=0}^{M-1} S(\lambda_{i+Kl}) \hat{x}(\lambda_{i+Kl}). \quad (11)$$

In matrix form, the sampled graph signal is represented as  $\hat{c} = \mathbf{D}_{\text{samp}} S(\mathbf{\Lambda}) \hat{x}$ , where  $\mathbf{D}_{\text{samp}} = [\mathbf{I}_K \ \mathbf{I}_K \ \dots] \in \mathbb{C}^{K \times N}$ . Then, we define the sampling matrix  $\mathbf{S}^*$  as [17]

$$\mathbf{S}^* = \mathbf{U}_{\text{reduced}} \mathbf{D}_{\text{samp}} S(\mathbf{\Lambda}) \mathbf{U}^*, \quad (12)$$

where  $\mathbf{U}_{\text{reduced}} \in \mathbb{C}^{K \times K}$  is an arbitrary unitary matrix, which may correspond to the GFT for a reduced-size graph. Theoretically, we can use any unitary matrix for  $\mathbf{U}_{\text{reduced}}$ . One choice which has been studied in multiscale transforms for graph signals is graph reduction [60].

Here, we define reconstruction in the graph frequency domain, which is the counterpart of (5), as follows:

$$[\mathbf{U}^* \tilde{\mathbf{x}}](\lambda_i) = \hat{d}(\lambda_{i \bmod K}) W(\lambda_i), \quad (13)$$

where  $\hat{d} \in \mathbb{C}^K$  is a vector composed of expansion coefficients and  $W(\lambda_i)$  is the reconstruction filter kernel defined in the graph frequency domain. Correspondingly, the reconstruction matrix is represented as

$$\mathbf{W} = \mathbf{U} \mathbf{W}(\mathbf{\Lambda}) \mathbf{D}_{\text{samp}}^* \mathbf{U}_{\text{reduced}}^*. \quad (14)$$

The reconstruction in (13) is performed by replicating the original spectrum in the same manner as standard signal processing [29].

### C. Graph Wide Sense Stationarity

Several definitions of GWSS have been proposed [32], [56], [61]. A definition of GWSS based on the WSS by shift (see Collorary 1 in Appendix A) is shown in Appendix B-A.

In this paper, we define GWSS based on the WSS by modulation (see Collorary 2 in Appendix A) as follows:

**Definition 3** (Graph Wide Sense Stationary by Modulation (GWSS<sub>M</sub>)). *Let  $\mathbf{x}$  be a graph signal on a graph  $\mathcal{G}^2$ . Then,  $\mathbf{x}$  is a graph wide sense stationary process if and only if the following two conditions are satisfied for all  $m$ :*

<sup>1</sup> $M$  is assumed to be a divisor of  $N$  for simplicity.

<sup>2</sup>We suppose that  $\mathcal{G}$  is connected for simplicity. Nevertheless, we can extend our definition of GWSS to the case where  $\mathcal{G}$  is not connected.

$$1) \ \mathbb{E}[(\mathbf{x} \star \delta_n)_m] = \mu_x = \text{const}, \quad (15)$$

$$2) \ \mathbb{E}[(\mathbf{x} \star \delta_n - \mu_x \mathbf{1})_m (\mathbf{x} \star \delta_k - \mu_x \mathbf{1})_m^*] = [\mathbf{\Gamma}_x]_{n,k}. \quad (16)$$

The operator  $\cdot \star \delta_k$  is defined as follows:

$$\mathbf{x} \star \delta_k := \mathbf{M} \text{diag}(u_0[k], u_1[k], \dots) \mathbf{U}^* \mathbf{x}, \quad (17)$$

where  $\mathbf{M} = [e^{j2\pi \cdot 0/N} \mathbf{1}, e^{j2\pi \cdot 1/N} \mathbf{1}, \dots]$ .

Note that  $\cdot \star \delta_k$  corresponds to the standard sinusoidal modulation, since  $(\mathbf{x} \star \delta_k)_n = (\mathbf{U} \exp(j\Omega) \mathbf{U}^* \mathbf{x})_k$  where  $\Omega = \text{diag}(0, 2\pi/N, \dots, 2\pi(N-1)/N)$ . Therefore, we refer to  $\cdot \star \delta_k$  as a modulation operator on a graph. Derivation and motivation of Definition 3 are given in Appendix B-B.

One of the important properties of Definition 3 is that  $\mathbf{\Gamma}_x$  is diagonalizable by  $\mathbf{U}$  (cf. Appendix B). We refer to  $\hat{\mathbf{\Gamma}}_x(\mathbf{\Lambda}) := \mathbf{U}^* \mathbf{\Gamma}_x \mathbf{U}$  as the graph PSD. This property is preferred in GSP since it parallels the Wiener-Khinchin relation in the time domain [62] and it is advantageous for exploiting spectral tools in GSP.

Recall that the autocovariance function is invariant to shift in the standard WSS [62]. Since a graph operator is often used as a shift operator in the graph setting, the invariance with shift for GWSS can be translated to

$$\mathbf{\Gamma}_x \mathbf{L} = \mathbf{L} \mathbf{\Gamma}_x. \quad (18)$$

This property is also used in [63], [64]. Note that  $\mathbf{\Gamma}_x$  is diagonalizable by  $\mathbf{U}$  if and only if (18) is satisfied [65]. In fact, the alternative definitions of GWSS require additional assumptions to satisfy (18). For example, [32] requires that all eigenvalues of  $\mathbf{L}$  are distinct. However, we sometimes encounter graphs whose graph operator has an eigenvalue with multiplicity greater than one [56]. Moreover, [56], [61] require that the covariance  $\mathbf{\Gamma}_x$  is expressed by a polynomial in  $\mathbf{L}$ , while this assumption may not be true in general, including the multiple eigenvalue case. On the other hand, our GWSS definition does not require such assumptions, but it still satisfies (18). As a result, Definition 3 can be regarded as an extension of the existing GWSS. The relationship among the GWSS definitions is discussed in detail in Appendix B-C. Note that our generalized sampling is applicable under different definitions of GWSS with slight modifications.

## IV. GRAPH WIENER FILTER: RECOVERY FOR STOCHASTIC GRAPH SIGNALS

We now consider the design of the correction filter. It is optimized such that the reconstructed graph signal is close to the original one in the mean squared sense. We show that the resulting Wiener filter parallels that in the SI setting when graph spectral sampling (Definition 2) is performed.

### A. Graph Wiener Filter

Suppose that  $\mathbf{x}$  is a zero-mean GWSS process with a known covariance  $\mathbf{\Gamma}_x$  and  $\boldsymbol{\eta}$  is a zero-mean GWSS noise process with a known covariance  $\mathbf{\Gamma}_\eta$ , independent of each other. For simplicity, we suppose that columns of  $\mathbf{W}$  and  $\mathbf{S}$  satisfy the Riesz condition [47], i.e.,  $\mathbf{W}^* \mathbf{W}$  and  $\mathbf{S}^* \mathbf{S}$  are invertible.

We now formulate the design the correction filter. Unlike the formulation (6) in the time domain, we can directly minimize

the expectation of the normed error in the graph setting because the subspace of graph signals is finite-dimensional.

We consider minimizing the MSE between the original and reconstructed graph signals by solving

$$\min_{\mathbf{H}} \mathbb{E}[\|\tilde{\mathbf{x}} - \mathbf{x}\|^2]. \quad (19)$$

The optimal correction filter is obtained in the following Theorem.

**Theorem 1.** *Suppose that the input signal  $\mathbf{x}$  and noise  $\mathbf{w}$  are zero-mean GWSS processes with covariances  $\Gamma_x$  and  $\Gamma_w$ , respectively, which are uncorrelated with each other. Then, the solution of (19) is given by*

$$\mathbf{H}_{\text{PRE}} = (\mathbf{W}^* \mathbf{W})^{-1} \mathbf{W}^* \Gamma_x \mathbf{S} (\mathbf{S}^* \Gamma_x \mathbf{S} + \Gamma_\eta)^{-1}. \quad (20)$$

*Proof.* The MSE  $\varepsilon_{\text{MSE}} = \mathbb{E}[\|\tilde{\mathbf{x}} - \mathbf{x}\|^2]$  is given by

$$\begin{aligned} \varepsilon_{\text{MSE}} &= \mathbb{E}[\text{tr}(\mathbf{WHS}^* \mathbf{x} \mathbf{x}^* \mathbf{SH}^* \mathbf{W}^*)] \\ &\quad + \mathbb{E}[\text{tr}(-\mathbf{x} \mathbf{x}^* \mathbf{SH}^* \mathbf{W}^* - \mathbf{x} \mathbf{x}^* \mathbf{WHS}^* + \mathbf{x} \mathbf{x}^*)] \\ &\quad + \mathbb{E}[\text{tr}(\mathbf{WH}\eta\eta^* \mathbf{H}^* \mathbf{W}^* - \mathbf{x} \eta^* \mathbf{H}^* \mathbf{W}^*)] \\ &\quad + \mathbb{E}[\text{tr}(-\mathbf{WH}\eta \mathbf{x}^* + \eta \eta^*)] \\ &= \text{tr}(\mathbf{WHS}^* \Gamma_x \mathbf{SH}^* \mathbf{W}^* - \Gamma_x \mathbf{SH}^* \mathbf{W}^*) \\ &\quad + \text{tr}(-\Gamma_x \mathbf{WHS}^* + \Gamma_x) \\ &\quad + \text{tr}(\mathbf{WH}\Gamma_\eta \mathbf{H}^* \mathbf{W}^* + \Gamma_\eta) \\ &= \text{tr}(\mathbf{S}^* \Gamma_x \mathbf{SH}^* \mathbf{W}^* \mathbf{WH} - 2\mathbf{S}^* \Gamma_x \mathbf{WH} + \Gamma_x) \\ &\quad + \text{tr}(\Gamma_\eta \mathbf{H}^* \mathbf{W}^* \mathbf{WH} + \Gamma_\eta), \end{aligned} \quad (21)$$

where the second equality holds since expectation and trace are interchangeable and the third equality follows by the cyclic property of trace.

The infinitesimal perturbation of  $\varepsilon_{\text{MSE}}$  with respect to  $\mathbf{H}$  results in

$$\begin{aligned} d\varepsilon_{\text{MSE}} &= \text{tr}(2\mathbf{S}^* \Gamma_x \mathbf{SH}^* \mathbf{W}^* \mathbf{W} d\mathbf{H}) \\ &\quad + \text{tr}(-2\mathbf{S}^* \Gamma_x \mathbf{W} d\mathbf{H} + 2\Gamma_\eta \mathbf{H}^* \mathbf{W}^* \mathbf{W} d\mathbf{H}). \end{aligned} \quad (22)$$

Here,  $d\varepsilon_{\text{MSE}}$  is defined by

$$d\varepsilon_{\text{MSE}} = \sum_{n,m} \left[ \frac{\partial \varepsilon_{\text{MSE}}}{\partial \mathbf{H}} \circ d\mathbf{H} \right]_{n,m} = \text{tr} \left( \left( \frac{\partial \varepsilon_{\text{MSE}}}{\partial \mathbf{H}} \right)^* d\mathbf{H} \right). \quad (23)$$

As a result, the derivative of  $\varepsilon_{\text{MSE}}$  in (21) with respect to  $\mathbf{H}$  results in

$$\frac{\partial \varepsilon_{\text{MSE}}}{\partial \mathbf{H}} = \mathbf{W}^* \mathbf{WH} (\mathbf{S}^* \Gamma_x \mathbf{S} + \Gamma_\eta) - \mathbf{W}^* \Gamma_x \mathbf{S}. \quad (24)$$

Setting (24) to zero, we obtain the optimal  $\mathbf{H}$  of (20).  $\square$

In Theorem 1, we consider the predefined case. Next, we move on to the unconstrained case: The reconstruction filter  $\mathbf{W}$  can also be freely chosen along with  $\mathbf{H}$ . The optimal solution is obtained by solving

$$\min_{\mathbf{W}, \mathbf{H}} \mathbb{E}[\|\tilde{\mathbf{x}} - \mathbf{x}\|^2]. \quad (25)$$

The optimal correction filter is given in the following Theorem.

**Theorem 2.** *Suppose that the input signal  $\mathbf{x}$  and noise  $\mathbf{w}$  are zero-mean GWSS processes with covariances  $\Gamma_x$  and  $\Gamma_w$ ,*

*respectively, which are uncorrelated with each other. Then, the solution of (25) is given by*

$$\mathbf{H}_{\text{UNC}} = (\mathbf{S}^* \Gamma_x \mathbf{S} + \Gamma_\eta)^{-1}, \quad \mathbf{W}_{\text{UNC}} = \Gamma_x \mathbf{S}. \quad (26)$$

*Proof.* Since (20) is optimal for an arbitrary  $\mathbf{W}$ , we plug it into (21) and have

$$\begin{aligned} \varepsilon_{\text{MSE}} &= \text{tr}(\mathbf{P}_{\mathcal{W}} \mathbf{ZS}^* \Gamma_x \mathbf{SZ}^* \mathbf{P}_{\mathcal{W}} - 2\Gamma_x \mathbf{SZ}^* \mathbf{P}_{\mathcal{W}} + \Gamma_x) \\ &\quad + \text{tr}(\mathbf{P}_{\mathcal{W}} \mathbf{Z}\Gamma_\eta \mathbf{Z}^* \mathbf{P}_{\mathcal{W}} + \Gamma_w), \end{aligned} \quad (27)$$

where  $\mathbf{P}_{\mathcal{W}} = \mathbf{W}(\mathbf{W}^* \mathbf{W})^{-1} \mathbf{W}^*$  and  $\mathbf{Z} = \Gamma_x \mathbf{S} (\mathbf{S}^* \Gamma_x \mathbf{S} + \Gamma_\eta)^{-1}$ . In the same manner as (24), the derivative of  $\varepsilon_{\text{MSE}}$  with respect to  $\mathbf{W}$  is derived from (27) as

$$\begin{aligned} \frac{\partial \varepsilon_{\text{MSE}}}{\partial \mathbf{W}} &= 2\mathbf{P}_{\mathcal{W}} \mathbf{ZS}^* \Gamma_x \mathbf{SZ}^* \mathbf{W} (\mathbf{W}^* \mathbf{W})^{-1} \\ &\quad - 2\Gamma_x \mathbf{SZ}^* \mathbf{W} (\mathbf{W}^* \mathbf{W})^{-1} \\ &\quad + 2\mathbf{P}_{\mathcal{W}} \mathbf{Z}\Gamma_\eta \mathbf{Z}^* \mathbf{W} (\mathbf{W}^* \mathbf{W})^{-1}. \end{aligned} \quad (28)$$

By setting (28) to zero, we have

$$(\mathbf{P}_{\mathcal{W}} \Gamma_x \mathbf{S} - \Gamma_x \mathbf{S}) \mathbf{Z}^* \mathbf{W} (\mathbf{W}^* \mathbf{W})^{-1} = \mathbf{0}. \quad (29)$$

Therefore,  $\mathbf{P}_{\mathcal{W}}$  is necessary to be the identity mapping over  $\mathcal{R}(\Gamma_x \mathbf{S})$ , leading to  $\mathcal{W} = \mathcal{R}(\Gamma_x \mathbf{S})$  where  $\mathcal{R}(\cdot)$  is the range space of a matrix. Graph signal recovery with (20) can be expressed as

$$\mathbf{WH} = \mathbf{P}_{\mathcal{W}} \Gamma_x \mathbf{S} (\mathbf{S}^* \Gamma_x \mathbf{S} + \Gamma_\eta)^{-1}. \quad (30)$$

Since  $\mathcal{W} = \mathcal{R}(\Gamma_x \mathbf{S})$ , we can choose  $\mathbf{W} = \Gamma_x \mathbf{S}$  and obtain  $\mathbf{H} = (\mathbf{S}^* \Gamma_x \mathbf{S} + \Gamma_\eta)^{-1}$ . This completes the proof.  $\square$

So far, we considered a general solution for stochastic recovery. By choosing graph frequency domain sampling, we can show that the resulting recovery parallels that in SI space, studied in [47], [58].

### B. Special cases for GFT domain sampling

Suppose that the graph signal and noise conform to zero-mean GWSS processes with PSD  $\hat{\Gamma}_x(\lambda)$  and  $\hat{\Gamma}_\eta(\lambda)$ , respectively. We assume that  $\mathbf{S}^*$  and  $\mathbf{W}$  are defined in the graph frequency domain by (12) and (14), respectively. In this setting, (20) is diagonalizable by  $\mathbf{U}_{\text{reduced}}$ , i.e., it has a graph frequency response

$$H_{\text{PRE}}(\lambda_i) = \frac{\sum_l \Gamma_x(\lambda_{i+Kl}) W^*(\lambda_{i+Kl}) S(\lambda_{i+Kl})}{R_{\mathcal{W}}(\lambda_i) (\sum_l \hat{\Gamma}_x(\lambda_{i+Kl}) |S(\lambda_{i+Kl})|^2 + \hat{\Gamma}_\eta(\lambda_i))}, \quad (31)$$

where

$$R_{\mathcal{W}}(\lambda_i) := \sum_{l=0}^{M-1} W^*(\lambda_{i+Kl}) W(\lambda_{i+Kl}). \quad (32)$$

Similarly, the graph spectral responses of (26) are given by

$$H_{\text{UNC}}(\lambda_i) = \frac{1}{\sum_l \Gamma_x(\lambda_{i+Kl}) |S(\lambda_{i+Kl})|^2 + \Gamma_\eta(\lambda_i)}, \quad (33)$$

with reconstruction filter  $W(\lambda_i) = \Gamma_x(\lambda_i) S(\lambda_i)$ . The correction filters (31) and (33) inherit frequency responses of the Wiener filter in standard sampling (7) and (8), respectively.

In the following, we discuss the relationship among the proposed graph Wiener filter and existing generalized graph signal sampling.

## V. RELATIONSHIP TO OTHER PRIORS

In this section, we consider the relationship between the proposed graph Wiener filter and graph signal recovery under subspace and smoothness priors [17]. For simplicity, we only consider the predefined case. The unconstrained case can be derived in a similar fashion.

### A. Subspace prior

Under the subspace prior, we assume the following graph signal model [17].

$$\mathbf{x} = \mathbf{A}\mathbf{d}, \quad (34)$$

where  $\mathbf{A} \in \mathbb{C}^{N \times K}$  is a known generator and  $\mathbf{d} \in \mathbb{C}^K$  is expansion coefficients.

We consider a relationship between the subspace prior and the stochastic one. Suppose that  $\mathbf{d}$  is a random vector and  $\Sigma_d = \mathbb{E}[\mathbf{d}\mathbf{d}^*]$  is the covariance of  $\mathbf{d}$ . The covariance of  $\mathbf{x}$  is then written as

$$\begin{aligned} \Sigma_x &= \mathbb{E}[\mathbf{A}\mathbf{d}\mathbf{d}^*\mathbf{A}^*] \\ &= \mathbf{A}\mathbb{E}[\mathbf{d}\mathbf{d}^*]\mathbf{A}^* \\ &= \mathbf{A}\Sigma_d\mathbf{A}^*. \end{aligned} \quad (35)$$

By substituting (35) to (20), we have the following correction filter:

$$\mathbf{H}_{\text{MX,SUB}} = (\mathbf{W}^*\mathbf{W})^{-1}\mathbf{W}^*\mathbf{A}\Sigma_d\mathbf{A}^*\mathbf{S}(\mathbf{S}^*\mathbf{A}\Sigma_d\mathbf{A}^*\mathbf{S} + \Gamma_\eta)^{-1}. \quad (36)$$

For subspace prior, if  $\mathbf{S}^*\mathbf{A}$  is invertible and there is no noise, (36) reduces to

$$\begin{aligned} \mathbf{H}_{\text{MX,SUB}} &= (\mathbf{W}^*\mathbf{W})^{-1}\mathbf{W}^*\mathbf{A}\Sigma_d\mathbf{A}^*\mathbf{S}(\mathbf{S}^*\mathbf{A}\Sigma_d\mathbf{A}^*\mathbf{S})^{-1} \\ &= (\mathbf{W}^*\mathbf{W})^{-1}\mathbf{W}^*\mathbf{A}\Sigma_d\mathbf{A}^*\mathbf{S}(\mathbf{A}^*\mathbf{S})^{-1}\Sigma_d^{-1}(\mathbf{S}^*\mathbf{A})^{-1} \\ &= (\mathbf{W}^*\mathbf{W})^{-1}\mathbf{W}^*\mathbf{A}(\mathbf{S}^*\mathbf{A})^{-1}, \end{aligned} \quad (37)$$

where we assume that  $\Sigma_d$  is invertible. As observed,  $\mathbf{H}_{\text{MX,SUB}}$  does not depend on  $\Sigma_d$ . Interestingly, the solution in (37) coincides with the minimax solution for signal recovery under subspace prior [17]. As a result, we can view the minimax recovery under the subspace prior (37) as a special case of (20) with random expansion coefficients and known  $\mathbf{A}$ .

### B. Smoothness prior

We consider smoothness prior. Under the smoothness prior, graph signals satisfy the following condition [17]:

$$\|\mathbf{V}\mathbf{x}\|^2 \leq \rho^2, \quad (38)$$

where  $\mathbf{V} = \mathbf{U}\mathbf{V}(\Lambda)\mathbf{U}^*$  is a smoothness measuring function, i.e., graph high-pass filter, and  $\rho > 0$  is a constant. We assume that  $\mathbf{V}$  is bounded, i.e.,  $\mathbf{V}^*\mathbf{V}$  is invertible.

To derive the solution of (19) subject to the constraint in (38), the problem may be written as minimization of the Lagrangian

$$L(\mathbf{H}, \xi) = \mathbb{E}[\|\tilde{\mathbf{x}} - \mathbf{x}\|^2] + \xi(\mathbb{E}[\|\mathbf{V}\mathbf{x}\|^2] - \rho^2), \quad (39)$$

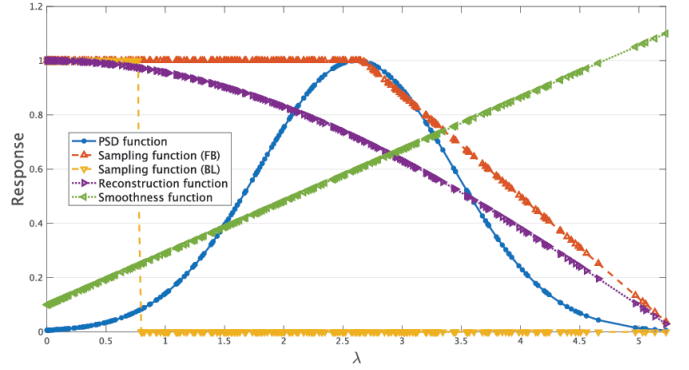


Fig. 3. Graph frequency responses of several functions used for experiments.

where  $\xi$  is a constant. The optimal solution of (39) satisfies the following identity [66]:

$$\lambda(\mathbb{E}[\|\mathbf{V}\mathbf{x}\|^2] - \rho^2) = \xi(\|\mathbf{V}\Sigma_x\mathbf{V}^*\|_F^2 - \rho^2) = 0. \quad (40)$$

Note that the second equality of (40) is satisfied with either  $\xi = 0$  or  $\|\mathbf{V}\Sigma_x\mathbf{V}^*\|_F^2 = \rho^2$ . Therefore, the nontrivial solution of (39) can be obtained with  $\|\mathbf{V}\Sigma_x\mathbf{V}^*\|_F^2 = \rho^2$ . Since  $\|\mathbf{V}(\mathbf{V}^*\mathbf{V})^{-1}\mathbf{V}^*\|_F^2 = \|\mathbf{I}\|_F^2 = N$ ,  $\Sigma_x$  can be written as

$$\Sigma_x = \frac{\rho^2}{N}(\mathbf{V}^*\mathbf{V})^{-1}. \quad (41)$$

Since the derivative of  $L(\mathbf{H}, \xi)$  with respect to  $\mathbf{H}$  coincides with (21), we immediately obtain the following solution:

$$\mathbf{H}_{\text{MX,SMO}} = (\mathbf{W}^*\mathbf{W})^{-1}\mathbf{W}^*(\mathbf{V}^*\mathbf{V})^{-1}\mathbf{S}(\mathbf{S}^*(\mathbf{V}^*\mathbf{V})^{-1}\mathbf{S})^{-1}, \quad (42)$$

where we also assume the noiseless case ( $\Gamma_\eta = \mathbf{0}$  in (21)) as in the subspace prior.

In fact, (42) coincides with the minimax solution under the smoothness prior [17]. Similar to (37), we can view minimax recovery under the smoothness prior (42) as a special case of (20).

## VI. SIGNAL RECOVERY EXPERIMENTS

In this section, we validate the effectiveness of the proposed method via signal recovery experiments. We demonstrate that the approach described in Section IV reduces the reconstruction error in comparison with existing recovery techniques under the stochastic setting.

### A. Synthetic Graph Signals

In this subsection, we perform signal recovery experiments for synthetic graph signals.

1) *Sampling and Recovery Setting*: We perform signal recovery experiments of synthetic graph signals. The recovery framework is illustrated in Fig. 2.

Here, we consider the following three graphs with  $N = 256$ :

- Random sensor graph<sup>3</sup>;
- Erdős-Rényi (ER) graph with edge connection probability  $p = 0.3$ ; and

<sup>3</sup>Random sensor graphs are implemented by  $k$  nearest neighbor graphs, whose vertices are randomly distributed in 2-D space  $[0, 1] \times [0, 1]$  [71].

TABLE II

AVERAGE MSEs OF THE RECONSTRUCTED SIGNALS FOR THE SYNTHETIC DATA ON 20 INDEPENDENT REALIZATIONS (IN DECIBELS). WE ALSO SHOW THE STANDARD DEVIATIONS ALONG WITH THE MSEs. ER, BL, FB, VS AND SS DENOTE AN ERDŐS-RÉNYI GRAPH, BANDLIMITED, FULL-BAND, VERTEX DOMAIN SAMPLING AND SPECTRAL DOMAIN SAMPLING, RESPECTIVELY. UNC, PRE, LS AND MX DENOTE UNCONSTRAINED, PREDEFINED, LEAST-SQUARES AND MINIMAX CRITERIA, RESPECTIVELY.

Prior			Stochastic		Smoothness [17]			BL	MKV V	MKVD	NLPD	GSOD	NLPI		
Reconst./ Criterion			UNC	PRE	UNC	PRE	LS	PRE	MX	[29]	[67]	[68]	[23]	[69]	[70]
Sensor	Noisy	BL	VS	-8.87±1.18	-8.86±1.19	443±14.4	208±11.5	442±14.4	-8.69±1.16	136±23.9	68.9±30.4	119±23.6	18.9±0.10	-6.79±1.15	-
		SS	-8.88±1.99	-8.88±1.99	-8.13±1.99	-6.99±1.98	-8.63±1.99	-8.19±1.99	-	-	-	-	-	-	-
		FB	VS	-11.9±1.15	-10.7±1.15	-4.73±1.44	-9.82±1.21	-9.59±1.21	-	143±18.8	203±22.7	162±17.8	18.9±0.11	-2.10±2.22	-
		SS	-14.1±1.99	-10.9±1.99	-5.05±1.99	-10.8±1.99	-10.6±1.99	-	-	-	-	-	-	-	-
	Clean	BL	VS	-8.99±1.20	-8.90±1.19	126±12.5	44.2±8.94	125±12.6	-8.90±1.19	97.0±20.4	-5.93±1.02	-6.94±1.22	18.7±0.09	-8.91±1.19	-
		SS	-8.99±1.99	-8.95±1.99	-8.99±1.99	-8.88±1.99	-8.95±1.99	-8.98±1.99	-	-	-	-	-	-	-
		FB	VS	-13.5±1.35	-11.5±1.23	-6.66±1.38	-11.4±1.23	-11.4±1.23	-	132±19.7	176±27.6	152±17.1	18.8±0.11	-4.01±2.13	-
		SS	-14.7±1.99	-11.2±1.99	-5.51±1.99	-11.1±1.99	-11.0±1.99	-	-	-	-	-	-	-	-
ER	Noisy	BL	VS	-18.6±1.40	-16.6±1.44	209±19.2	198±18.5	209±19.2	-17.3±1.43	-14.8±1.46	30.1±22.8	142±12.2	18.0±0.32	-16.5±1.40	-
		SS	-24.2±1.99	-18.2±1.99	-24.2±1.99	-24.3±1.99	-12.6±1.99	-18.2±1.99	-	-	-	-	-	-	-
		FB	VS	-18.8±1.53	-17.2±1.46	-14.3±1.50	-14.1±1.43	-14.1±1.43	-	-14.8±1.46	53.1±14.1	147±11.8	18.1±0.31	-16.8±1.42	-
		SS	-25.9±2.00	-19.5±2.00	-21.3±1.99	-19.5±1.99	-19.5±1.99	-	-	-	-	-	-	-	-
	Clean	BL	VS	-26.7±1.64	-17.5±1.51	62.4±15.9	66.1±14.0	62.1±15.9	-17.6±1.47	-14.8±1.46	-4.95±22.8	137±12.7	17.8±0.33	-16.9±1.46	-
		SS	-26.7±2.00	-18.7±1.99	-26.7±2.00	-13.9±2.00	-18.7±1.99	-26.5±2.00	-	-	-	-	-	-	-
		FB	VS	-25.2±1.91	-18.8±1.61	-19.9±1.63	-18.8±1.61	-18.8±1.61	-	-14.8±1.46	14.4±20.2	141±12.2	18.0±0.32	-17.7±1.54	-
		SS	-28.2±2.01	-20.1±2.00	-22.1±2.00	-20.1±2.00	-20.1±2.00	-	-	-	-	-	-	-	-
2D grid	Noisy	BL	VS	-8.25±1.16	-8.24±1.16	311±18.1	204±13.8	310±18.1	-8.11±1.15	-8.07±1.15	-2.85±17.1	132±23.5	18.9±0.09	-5.96±1.14	-
		SS	-8.31±1.99	-8.24±1.99	-7.57±1.99	-6.13±1.98	-7.85±1.99	-7.67±1.99	-	-	-	-	-	-	-
		FB	VS	-12.4±1.15	-10.6±1.17	-5.58±1.35	-10.5±1.16	-10.3±1.15	-	-8.12±1.15	66.2±18.7	156±18.4	18.9±0.10	-2.34±1.88	-
		SS	-13.7±1.99	-10.7±1.99	-5.75±1.99	-10.6±1.99	-10.4±1.99	-	-	-	-	-	-	-	-
	Clean	BL	VS	-8.51±1.18	-8.33±1.18	72.0±17.3	36.0±10.4	71.2±17.2	-8.32±1.17	-8.23±1.16	-6.38±1.08	-6.41±1.62	18.7±0.11	-8.32±1.17	-
		SS	-8.51±1.99	-8.39±1.99	-8.51±1.99	-8.25±1.99	-8.39±1.99	-8.48±1.99	-	-	-	-	-	-	-
		FB	VS	-14.0±1.29	-11.3±1.24	-6.45±1.37	-11.2±1.24	-11.2±1.24	-	-8.13±1.15	61.3±20.6	153±18.3	18.9±0.11	-3.29±1.82	-
		SS	-14.5±2.00	-11.0±1.99	-6.06±1.99	-10.9±1.99	-10.8±1.99	-	-	-	-	-	-	-	-

- 2D grid graph.

We use the following functions throughout the experiments.

- PSD function:

$$\hat{\Gamma}_x(\lambda_i) = \exp\left(-\left\{\frac{2\lambda_i - \lambda_{\max}}{\sqrt{\lambda_{\max}}}\right\}^2\right). \quad (43)$$

- Sampling function #1 (for full-band sampling):

$$S(\lambda_i) = \begin{cases} 1 & \lambda_{\max}/\lambda_i \leq 2 \\ 2 - 2\lambda_i/(\lambda_{\max}) & \text{otherwise.} \end{cases} \quad (44)$$

- Sampling function #2 (for bandlimited sampling):

$$S(\lambda_i) = \begin{cases} 1 & i \in [0, K - 1], \\ 0 & \text{otherwise.} \end{cases} \quad (45)$$

- Reconstruction function (for constrained recovery):

$$W(\lambda_i) = \cos\left(\frac{\pi}{2} \cdot \frac{\lambda_i}{\lambda_{\max}}\right). \quad (46)$$

- Smoothness function (for recovery with smoothness prior by [17]):

$$V(\lambda_i) = \frac{\lambda_i}{\lambda_{\max}} + \varepsilon \quad (47)$$

where we set to  $\varepsilon = 0.1$ .

Their spectral responses are shown in Fig. 3.

A stochastic graph signal is generated by  $\mathbf{x} \sim \mathcal{N}(\mathbf{0}, \mathbf{\Gamma}_x)$ , where  $\mathbf{\Gamma}_x$  satisfies the GWSS conditions described in Section III, i.e.,  $\mathbf{\Gamma}_x = \mathbf{U}\hat{\mathbf{\Gamma}}_x(\mathbf{\Lambda})\mathbf{U}^*$ .

As mentioned in Appendix B, existing definitions of GWSS coincide when  $\mathbf{\Gamma}_x$  is diagonalizable by  $\mathbf{U}$ . Therefore, this

setting simultaneously satisfies Definition 3 and the other existing definitions of GWSS [32], [56], [61].

Under the presence of noise, we suppose that noise conforms to i.i.d. zero-mean Gaussian distribution with  $\sigma^2 = 0.3$ , i.e.,  $\eta \sim \mathcal{N}(0, 0.3)$  and thus  $\Gamma_\eta(\lambda_i) = 0.3$  for all  $i$ . Graph signals are sampled with sampling ratio  $N/K = 4$ .

We perform experiments for two cases. 1) The sampling function  $S(\lambda_i)$  is full-band as in (44). 2)  $S(\lambda_i)$  is bandlimited as in (45). In addition, two sampling domains, i.e., samplings in vertex and graph frequency domains, are considered. For vertex domain sampling, sampled vertices are selected randomly. For the predefined recovery, we use the given kernel  $W(\lambda_i)$  in (46).

We calculate the average MSE in 20 independent runs and compare with existing graph signal interpolation methods, MKVV [67], MKVD [68], NLPD [23], GSOD [69] and NLPI [70], and recovery methods under smoothness priors proposed in [17]. Smoothness is measured by the energy of high frequency graph spectra with the measuring function  $V(\lambda_i)$ . We also show the result of the bandlimited reconstruction [29], i.e., reconstruction with the sampling filter #2 without correction.

Since there has been no prior work on graph signal sampling with stochastic prior, we also compare with graph signal recovery with smoothness prior [17] as a benchmark.

2) *Results*: Table II summarizes the average MSEs with the standard deviations (in decibels). Examples of recovered signals via full-band sampling are visualized in Fig. 4.

From Table II, it is observed that the proposed methods consistently present the best MSEs in most cases. It is regardless of the sampling domain. Sometimes, the MSEs for

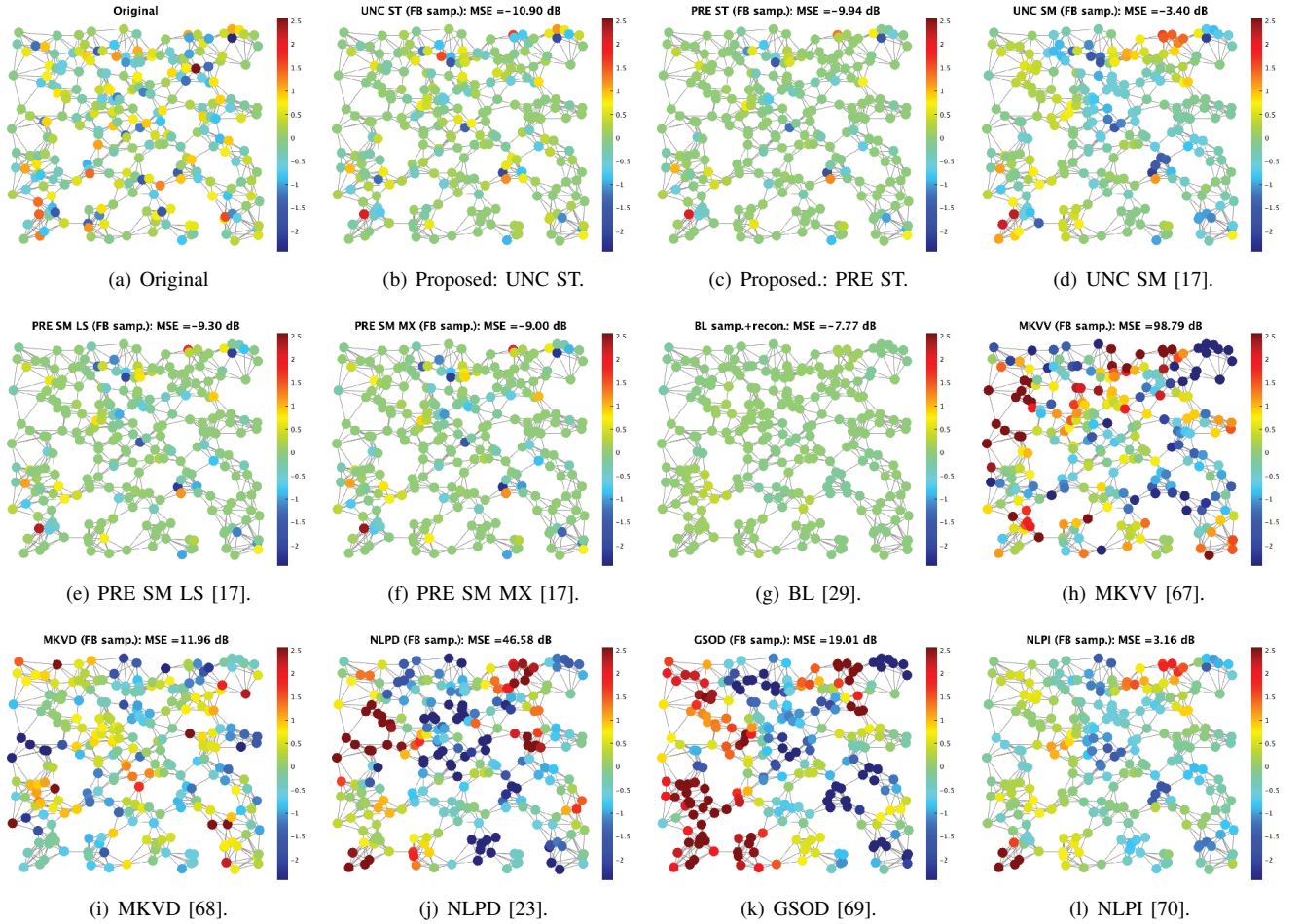


Fig. 4. Signal recovery experiments for noisy graph signals on a random sensor graph with  $N = 256$ . Sampling is performed by (44) in the vertex domain. UNC and PRE denote the unconstrained and predefined solutions. SM and ST denote recovery with smoothness and stochastic priors. LS and MX denote least-squares and minimax criteria.

the bandlimited sampling is close to the proposed method, especially for the ER graph. This is because of the eigenvalue distribution: The ER graph may produce eigenvalues sparsely distributed in low graph frequencies, i.e.,  $\lambda < \lambda_{\max}/2$  where we consider  $\mathbf{L}$  as a graph operator, and many eigenvalues exist in high graph frequencies ( $\lambda \geq \lambda_{\max}/2$ ). This implies that the cutoff frequency at the  $K$ th eigenvalue can be relatively high, resulting in wide bandwidth. This also results in the fact that the bandlimited sampling filter in (45) may pass the spectra in the middle graph frequencies ( $\lambda \in [\lambda_{\max}/4, 3\lambda_{\max}/4]$ ); This behavior improves the reconstruction MSEs. Note that the proposed method presents stable recovery performances for all the graph and signal models considered.

The recovery with smoothness prior increases MSEs especially for vertex domain sampling because the PSD function  $\hat{\Gamma}_x(\lambda_i)$  does not represent smooth signals in this setting. For the noiseless cases, the proposed methods demonstrate MSE improvements by 3–5 dB on average compared to the noisy case. The unconstrained solution of the proposed stochastic recovery results in the best performance for almost all cases.

## B. Real-world Data

We also perform signal recovery experiments on real-world data.

1) *Sampling and Recovery Setting:* We use the global sea surface temperature dataset [72]. It records snapshots of sea surface temperatures for every month from 2004 to 2021. They are spatially sampled at intersections of 1-degree latitude-longitude grids. For the experiment, we used data for 24 months from 2018 to 2020. For this dataset, we randomly sample  $N = 291$  intersections and they are regarded as vertices. We then construct a 5-nearest neighbor graph based on Euclidean distance. Edge weights are set to be unweighted. We then remove the edges crossing land areas. The number of samples is set to  $K = 73$ .

For each time instance, we estimate the covariance from the data in the previous year. We use the method proposed in [56]. The estimated covariance is used for the proposed graph Wiener filter. We calculate the average MSEs of 24 months for samplings in vertex and graph frequency domains, and compare with the existing methods in the previous subsection. For vertex domain sampling, we perform the experiments on 20 random sampling patterns of vertices for calculating the average MSE.



TABLE III

AVERAGE MSEs OF THE RECONSTRUCTED SIGNALS FOR THE SEA SURFACE TEMPERATURE DATA (IN DECIBELS). NOTATIONS ARE THE SAME AS TABLE II.

Prior		Stochastic		Smoothness [17]			BL	MKVV	MKVD	NLPD	GSOD	NLPI
Reconst./	Criterion	UNC	PRE	UNC	PRE LS	PRE MX	[29]	[67]	[68]	[23]	[69]	[70]
BL	VS	$20.6 \pm 4.05$	$54.4 \pm 3.17$	$232 \pm 15.5$	$154 \pm 8.71$	$232 \pm 15.5$	$56.0 \pm 3.14$	$60.9 \pm 3.13$	$209 \pm 15.8$	$206 \pm 12.3$	$61.0 \pm 3.13$	$827 \pm 7.05$
	SS	$14.1 \pm 1.16$	$55.1 \pm 0.15$	$14.1 \pm 1.16$	$63.3 \pm 0.14$	$551 \pm 0.15$	$14.2 \pm 1.15$	-	-	-	-	-
FB	VS	$28.3 \pm 4.34$	$54.4 \pm 3.17$	$32.0 \pm 3.93$	$54.5 \pm 3.17$	$54.4 \pm 3.17$	-	$60.9 \pm 3.13$	$210 \pm 15.8$	$206 \pm 12.0$	$61.0 \pm 3.13$	$828 \pm 6.88$
	SS	$19.5 \pm 1.11$	$55.1 \pm 0.14$	$22.0 \pm 0.94$	$55.1 \pm 0.14$	$55.1 \pm 0.14$	-	-	-	-	-	-

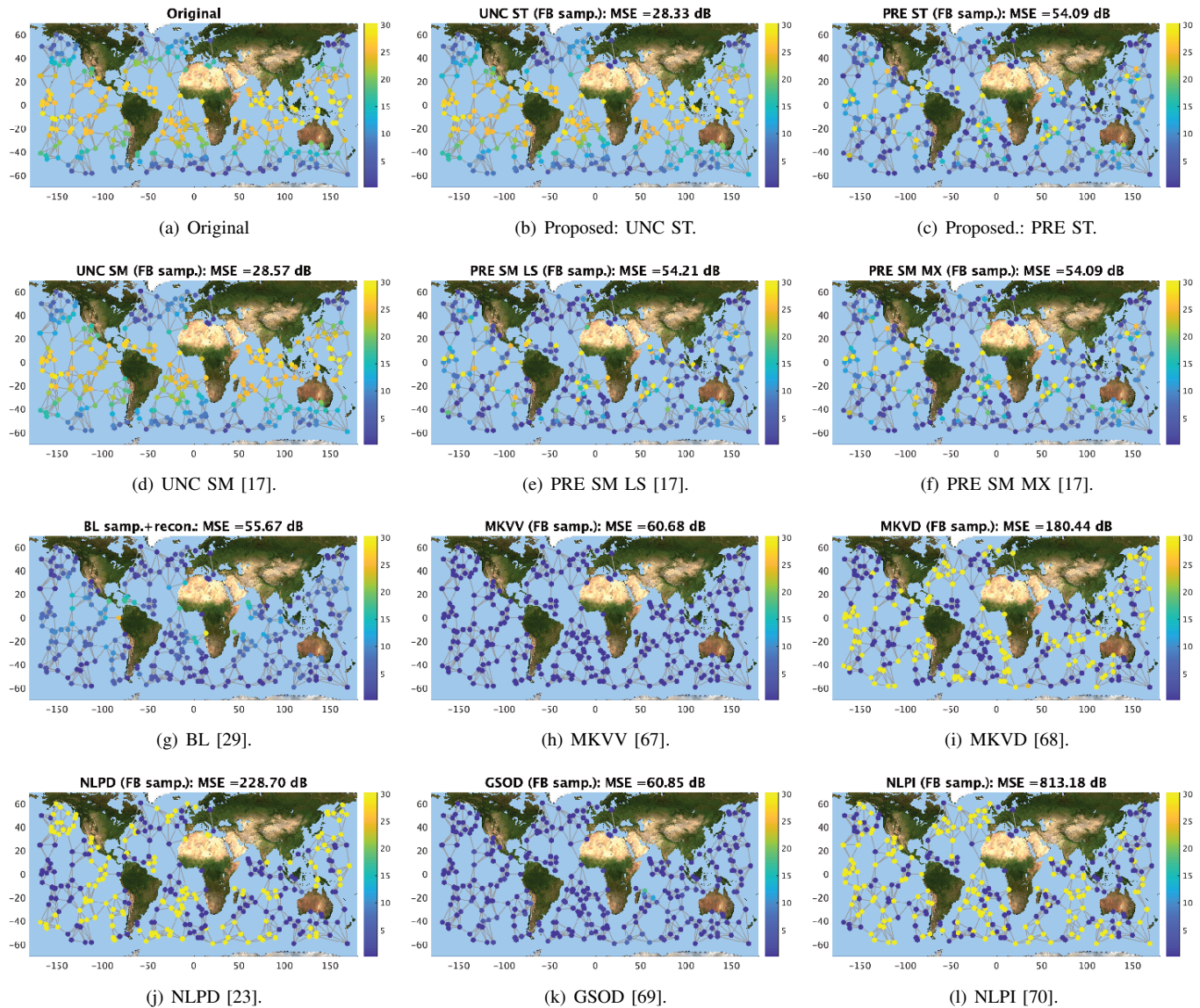


Fig. 5. Signal recovery experiments for the sea surface temperature data on a 5-nearest neighbor graph. Sampling is performed by (44) in the vertex domain. Abbreviations are the same as those in Fig. 4.

2) *Results:* Table III summarizes the average MSEs with the standard deviations (in decibels). Recovered graph signals are visualized in Fig. 5.

From Table III, we can observe that the proposed method outperforms existing methods for almost all cases, as expected to the experiment on synthetic data. This is also validated in Fig. 5.

## VII. CONCLUSION

Generalized graph signal sampling with stochastic priors is proposed in this paper. We derive the graph Wiener filter for unconstrained and predefined reconstruction, based on minimization of the expectation of the squared norm error between the original and reconstructed graph signals. We show that, when sampling is performed in the graph frequency domain, spectral responses of the graph Wiener filter parallel

those in generalized sampling for WSS signals. We also reveal a theoretical relationship between the proposed graph Wiener filter and existing signal recovery under different priors. In the recovery experiments, we validate the MSE improvement of the proposed methods for various graphs and two sampling domains.

#### APPENDIX A WIDE SENSE STATIONARITY

Stationarity of time-domain signals is reviewed here since this is useful to understand the connection between WSS and GWSS.

We consider a continuous-time signal  $x(t)$ . Its stationarity is defined as follows:

**Definition 4** (Wide sense stationary for time-domain signals). *Let  $x(t)$  and  $\gamma_x(t)$  be a stochastic signal in the time domain and its autocovariance function, respectively. The signal  $x(t)$  is a wide sense stationary process if and only if the following two conditions are satisfied:*

- 1)  $\mathbb{E}[x(t)] = \mu_x = \text{const}$ , (48)
- 2)  $\mathbb{E}[(x(t) - \mu_x)(x(\tau) - \mu_x)^*] = \gamma_x(t - \tau)$ . (49)

A WSS process can be characterized in the Fourier domain by Wiener-Khinchin theorem [73]. If  $x(t)$  is a WSS process, then its power spectral density (PSD) function coincides with the CTFT of the autocovariance function of  $x(t)$ ,  $\gamma_x(t) \in L_1$ , i.e.,

$$\Gamma_x(\omega) = \int_{-\infty}^{\infty} \gamma_x(t) e^{-j\omega t} dt. \quad (50)$$

Not surprisingly, WSS has several equivalent expressions because of the correspondence between time-shift and frequency-modulation, i.e.,  $x(t - \tau) \leftrightarrow e^{-j\tau\omega} X(\omega)$ .

Here, we present another two expressions of WSS to show the connection with GWSSs.

**Corollary 1** (WSS by Shift). *Let  $T_{t_0}\{\cdot\}$  be the shift operator that delays the signal by  $t_0$ , i.e.,  $T_{t_0}\{x(t)\} = x(t - t_0)$ . Definition 4 can be written as follows:*

- 1)  $\mathbb{E}[T_{t_0}\{x(t)\}] = \mu_x = \text{const}$ , (51)

- 2)  $\mathbb{E}[(T_{t_0}\{x(t)\} - \mu_x)(T_{t_0}\{x(\tau)\} - \mu_x)^*] = \gamma_x(t - \tau)$ . (52)

The equivalence of conditions (51) and (48) is easy to verify. Since the autocovariance only depends on the time difference, (52) is identical to (49).

Noting that

$$x(\tau) = x(t) * \delta(t - \tau), \quad (53)$$

leads to the following corollary.

**Corollary 2** (WSS by Modulation). *Definition 4 can be expressed equivalently by using  $\delta(t)$  as follows:*

- 1)  $\mathbb{E}[x(t_0) * \delta(t_0 - \tau)] = \mu_x = \text{const}$ , (54)

- 2)  $\mathbb{E}[(x(t_0) * \delta(t_0 - t) - \mu_x)(x(t_0) * \delta(t_0 - \tau) - \mu_x)^*] = \gamma_x(t - \tau)$ . (55)

In the paper, we utilize these expressions of WSS for formally defining GWSS.

#### APPENDIX B GRAPH WIDE SENSE STATIONARITY

In this Appendix, we describe and compare some definitions of GWSS, including ours. These definitions differ in whether Corollary 1 or 2 in Appendix A is used for the baseline. They coincide for time-domain signals, but this is not the case for graph signals, leading to slightly different definitions of GWSS. The definitions are mainly divided according to whether the covariance is diagonalizable by the GFT basis  $\mathbf{U}$ . These differences are summarized in Table IV.

##### A. GWSS followed by Corollary 1

First, we show the definition of GWSS [32] as a counterpart of Corollary 1. In the literature of GSP, the graph Laplacian  $\mathbf{L}$  is often referred to as a counterpart of the time-shift operator (translation operator) [24], [61]. However,  $\mathbf{L}$  changes the signal energy, i.e.,  $\|\mathbf{L}\mathbf{x}\| \neq \|\mathbf{x}\|$ , while time shift does not, i.e.,  $\|T_\tau\{x[n]\}\| = \|x[n]\|$ .

The first definition of GWSS introduced here is based on a graph-translation operator which preserves the signal energy.

**Definition 5** (Graph Wide Sense Stationary by Translation (GWSS<sub>T</sub>) [74]). *Let  $\mathbf{x}$  be a graph signal on a graph  $\mathcal{G}$ . Suppose that a graph-translation operator  $\mathbf{T}_G$  is defined by*

$$\mathbf{T}_G := \exp\left(j\pi\sqrt{\frac{\mathbf{L}}{\rho_G}}\right), \quad (56)$$

where  $\rho_G = \max_{m \in \mathcal{V}} \sqrt{2d_m(d_m + \bar{d}_m)}$ ,  $\bar{d}_m = \frac{\sum_{n=0}^{N-1} a_{mn}d_n}{d_m}$  and  $\rho_G \geq \lambda_{\max}$  [75]. Then,  $\mathbf{x}$  is a graph wide sense stationary process by translation, if and only if it the following two conditions are satisfied,

- 1)  $\mathbb{E}[(\mathbf{T}_G\mathbf{x})_n] = \mu_x = \text{const}$ , (57)

- 2)  $\mathbb{E}[(\mathbf{T}_G\mathbf{x} - \mu_x\mathbf{1})_n(\mathbf{T}_G\mathbf{x} - \mu_x\mathbf{1})_k^*] = [\Gamma_x]_{n,k}$ . (58)

The condition (57) implies that  $\mathbb{E}[\mathbf{x}]$  is also constant. To see this, note that (57) can be expressed as

$$\mathbf{T}_G\mathbb{E}[\mathbf{x}] = \mathbf{U} \exp(j\pi\sqrt{\Lambda/\rho_G})\mathbf{U}^*\mathbb{E}[\mathbf{x}]. \quad (59)$$

Since  $\lambda_0 = 0$  and  $\mathbf{u}_0 = \mathbf{1}$  are always satisfied, i.e.,  $\mathbf{L}\mathbf{1} = 0 \cdot \mathbf{1}$ , (59) holds if and only if  $\mathbb{E}[\mathbf{x}] = \mu_x\mathbf{u}_0 = \mu_x\mathbf{1}$ . It also implies that  $\mathbf{T}_G\mathbf{1} = \mathbf{1}$ .

When  $\{\lambda_i\}_{i=0, \dots, N-1}$  are distinct, the condition (58) implies that  $\Gamma_x$  is diagonalizable by  $\mathbf{U}$  because (58) can be expressed as

$$\begin{aligned} \Gamma_x &= \mathbf{T}_G\mathbb{E}[(\mathbf{x} - \mu_x\mathbf{1})(\mathbf{x} - \mu_x\mathbf{1})^*]\mathbf{T}_G^* \\ &= \mathbf{U} \exp(j\pi\sqrt{\Lambda/\rho_G})\mathbf{U}^*\Gamma_x\mathbf{U} \exp(-j\pi\sqrt{\Lambda/\rho_G})\mathbf{U}^* \\ &= \mathbf{U}(\hat{\Gamma}_x \circ \Theta)\mathbf{U}^*, \end{aligned} \quad (60)$$

where  $\hat{\Gamma}_x = \mathbf{U}^*\Gamma_x\mathbf{U}$  and  $[\Theta]_{i,l} = \exp\{j\pi(\sqrt{\lambda_i/\rho_G} - \sqrt{\lambda_l/\rho_G})\}$ . The third equality follows by the relationship  $\text{diag}(\mathbf{a})\mathbf{X}\text{diag}(\mathbf{b}^*) = \mathbf{X} \circ \mathbf{a}\mathbf{b}^*$ . Since  $[\Theta]_{i,l} = 1$  for  $\lambda_i = \lambda_l$  and  $[\Theta]_{i,l} \neq 1$  otherwise, the equality holds if and only if  $[\hat{\Gamma}_x]_{i,l} = 0$  for  $\lambda_i \neq \lambda_l$ . If the eigenvalues are distinct, then this condition is equivalent that  $\Gamma_x$  is diagonalizable by  $\mathbf{U}$ .

TABLE IV  
COMPARISON AMONG WSS AND GWSSs. OPE AND COV DENOTE OPERATOR AND COVARIANCE, RESPECTIVELY.

OPE	WSS		Mean	WSS	
	Shift	Modulation		Shift	Modulation
	$T_{t_0}\{x(t)\} = x(t-t_0)$	$x(t_0) * \delta_t = x(t)$		$\mathbb{E}[T_{t_0}\{x(t)\}] = \mu_x$	$\mathbb{E}[x(t_0) * \delta_\tau] = \mu_x$
	GWSS			GWSS	
GWSS <sub>T</sub>	GWSS <sub>M</sub>	GWSS <sub>T</sub>	GWSS <sub>M</sub>		
$\mathbf{T}_G \mathbf{x} = \mathbf{U} \exp(j\mathbf{\Pi}) \mathbf{U}^* \mathbf{x}$	$\mathbf{x} * \delta_n = \mathbf{M} \text{diag}(\mathbf{U}^* \delta_n) \mathbf{U}^* \mathbf{x}$	$\mathbb{E}[(\mathbf{T}_G \mathbf{x})_n] = \mu_x$	$\mathbb{E}[(\mathbf{x} * \delta_n)_m] = \mu_x$		
COV	WSS		PSD	WSS	
	Shift	Modulation		Shift	Modulation
	$\mathbb{E}[T_{t_0}\{\bar{x}(t)\} T_{t_0}\{\bar{x}(\tau)\}^*] = \gamma_x(t-\tau)$	$\mathbb{E}[(\bar{x}(t_0) * \delta_t)(\bar{x}(t_0) * \delta_\tau)^*] = \gamma_x(t-\tau)$		$\Gamma_x(\omega) = \int_{-\infty}^{\infty} \gamma_x(t) e^{-j\omega t} dt$ always	
	GWSS			GWSS	
GWSS <sub>T</sub>	GWSS <sub>M</sub>	GWSS <sub>T</sub>	GWSS <sub>M</sub>		
$\mathbb{E}[(\mathbf{T}_G \bar{\mathbf{x}})_n (\mathbf{T}_G \bar{\mathbf{x}})_k^*] = [\mathbf{\Gamma}_x]_{n,k}$	$\mathbb{E}[(\bar{\mathbf{x}} * \delta_n)_m (\bar{\mathbf{x}} * \delta_k)_m^*] = [\mathbf{\Gamma}_x]_{n,k}$	$\hat{\mathbf{\Gamma}}_x(\mathbf{\Lambda}) = \mathbf{U}^* \mathbf{\Gamma}_x \mathbf{U}$ if $\mathbf{\Lambda}$ is distinct	$\hat{\mathbf{\Gamma}}_x(\mathbf{\Lambda}) = \mathbf{U}^* \mathbf{\Gamma}_x \mathbf{U}$ always		

We denote by  $\delta_t = \delta(t_0 - t)$ ,  $\bar{x}(t) = x(t) - \mu_x$ ,  $\bar{\mathbf{x}} = \mathbf{x} - \eta_x \mathbf{1}$ , and  $\mathbf{\Pi} = \pi \sqrt{\Lambda} / \rho_G$ .

### B. GWSS followed by Corollary 2

Next, we study the properties of Definition 3 for GWSS used in our generalized sampling, as a counterpart of Corollary 2.

In Definition 3, the condition in (15) is identical to

$$\mathbb{E}[\mathbf{x} * \delta_n] = \mathbf{M} \text{diag}(u_0[n], u_1[n], \dots) \mathbf{U}^* \mathbb{E}[\mathbf{x}]. \quad (61)$$

Since  $\mathbf{M} \text{diag}(1, 0, \dots, 0) = [\mathbf{1} \ \mathbf{0} \ \dots \ \mathbf{0}]$ , the equality holds if and only if  $\mathbb{E}[\mathbf{x}] = \mu_x \mathbf{u}_0 = \mu_x \mathbf{1}$ . Therefore,  $\cdot * \delta_n$  does not change the mean of graph signals. The condition in (16) implies the covariance  $\mathbf{\Gamma}_x$  has to be diagonalizable by  $\mathbf{U}$ . We show this fact in the following lemma.

**Lemma 1.** *Let  $\mathbf{x}$  and  $\mathbf{\Gamma}_x$  be a stochastic signal on  $\mathcal{G}$  and its covariance matrix, respectively, by Definition 3. Then,  $\mathbf{\Gamma}_x$  is diagonalizable by  $\mathbf{U}$ .*

*Proof.* The LHS of (16) is expressed as

$$\begin{aligned} & \mathbb{E}[(\mathbf{x} * \delta_n - \mu_x \mathbf{1})_m (\mathbf{x} * \delta_k - \mu_x \mathbf{1})_m^*] \\ &= [\mathbf{M} \text{diag}(u_0[n], u_1[n], \dots) \hat{\mathbf{\Gamma}}_x \text{diag}(u_0^*[k], u_1^*[k], \dots) \mathbf{M}^*]_{m,m} \\ &= \delta_m^* \mathbf{M} \text{diag}(\mathbf{U}^* \delta_n) \hat{\mathbf{\Gamma}}_x \text{diag}(\mathbf{U}^* \delta_k) \mathbf{M}^* \delta_m \\ &= \delta_n^* \mathbf{U} \text{diag}(\mathbf{M}^* \delta_m) \hat{\mathbf{\Gamma}}_x \text{diag}(\mathbf{M}^* \delta_m) \mathbf{U}^* \delta_k \\ &= [\mathbf{U}(\hat{\mathbf{\Gamma}}_x \circ \mathbf{M}^* \delta_m \delta_m^* \mathbf{M}) \mathbf{U}^*]_{n,k} \\ &= [\mathbf{U}(\hat{\mathbf{\Gamma}}_x \circ \mathbf{\Xi}) \mathbf{U}^*]_{n,k}, \end{aligned} \quad (62)$$

where  $\hat{\mathbf{\Gamma}}_x = \mathbb{E}[\mathbf{U}^*(\mathbf{x} - \mu_x \mathbf{1})(\mathbf{x} - \mu_x \mathbf{1})^* \mathbf{U}]$  and  $[\mathbf{\Xi}]_{i,l} = [\mathbf{M}^* \delta_m \delta_m^* \mathbf{M}]_{i,l} = \exp(-j2\pi(i-l)/N)$ . Since  $[\mathbf{\Xi}]_{i,l} = 1$  for  $i = l$  and  $[\mathbf{\Xi}]_{i,l} \neq 1$  otherwise, the condition satisfying (62), i.e.,  $\hat{\mathbf{\Gamma}}_x = \hat{\mathbf{\Gamma}}_x \circ \mathbf{\Xi}$ , holds if and only if  $\hat{\mathbf{\Gamma}}_x$  is diagonal. Therefore,  $\mathbf{\Gamma}_x$  is diagonalizable by  $\mathbf{U}$ , which completes the proof.  $\square$

### C. Relationship among GWSS definitions

We now discuss the relationship among some representative definitions of GWSS from the viewpoint of the PSD. These results are summarized in Table V.

Existing definitions of GWSS are defined as counterparts of the classical WSS definitions in Corollaries 1 and 2. In terms of diagonalizability, GWSS<sub>M</sub> (the counterpart of Corollary 1) is stricter than GWSS<sub>T</sub> (the counterpart of Corollary 2). This

TABLE V  
DIAGONALIZABILITY OF THE COVARIANCE FOR GWSS DEFINITIONS. IF ALL EIGENVALUES IN  $\mathbf{\Lambda}$  ARE DISTINCT, THEY COINCIDE WITH EACH OTHER.

$\mathbf{\Gamma}_x$ is not necessarily diagonalizable by $\mathbf{U}$	GWSS <sub>T</sub> [32]
$\mathbf{\Gamma}_x$ is diagonalizable by $\mathbf{U}$	GWSS <sub>M</sub>
$\mathbf{\Gamma}_x$ is a polynomial in $\mathbf{L}$	[56], [61]

is because the covariance is always diagonalizable by  $\mathbf{U}$  in GWSS<sub>M</sub>, while that is not the case with GWSS<sub>T</sub> in general.

Next, we compare  $\Theta$  in (60) with  $\Xi$  in (62). In (62), off-diagonal entries in  $\Xi$  are not equal to 1. In contrast, those in  $\Theta$  can take the value 1 in the case  $\lambda_i = \lambda_l$  for  $i \neq l$ , i.e., eigenvalues with multiplicity greater than 1. Therefore, GWSS<sub>T</sub> allows the existence of non-zero off-diagonal elements of  $\hat{\mathbf{\Gamma}}_x$  for some graphs having repeated eigenvalues, while GWSS<sub>M</sub> always yields the diagonal  $\hat{\mathbf{\Gamma}}_x$  regardless of graphs. In fact, GWSS<sub>M</sub> and GWSS<sub>T</sub> coincide with each other if all eigenvalues of  $\mathbf{L}$  are distinct (cf. Lemma 1). Therefore, GWSS<sub>T</sub> effectively assumes distinct eigenvalues of  $\mathbf{L}$  in [32].

It is often assumed that  $\mathbf{\Gamma}_x$  is a polynomial in  $\mathbf{L}$  [56], [61]<sup>4</sup>. This is sufficient for the diagonalizability of  $\mathbf{\Gamma}_x$ . It is noteworthy that the polynomial assumption is equivalent to  $\hat{\mathbf{\Gamma}}_x(\lambda_i) = \hat{\mathbf{\Gamma}}_x(\lambda_l)$  for all  $\lambda_i = \lambda_l$  [56]. This is a special case of GWSS<sub>M</sub>, since GWSS<sub>M</sub> allows for  $\hat{\mathbf{\Gamma}}_x(\lambda_i) \neq \hat{\mathbf{\Gamma}}_x(\lambda_l)$  for any  $\lambda_i = \lambda_l$ . As a result, GWSS<sub>M</sub> is a good compromise between applicability and a rigorous relationship to the WSS definition. Therefore, we use GWSS<sub>M</sub> as our GWSS definition.

### REFERENCES

- [1] J. Hara, Y. Tanaka, and Y. C. Eldar, "Generalized graph spectral sampling with stochastic priors," in *Proc. 2020 IEEE Int. Conf. Acoust. Speech Signal Process. (ICASSP)*, pp. 5680–5684, May 2020.

<sup>4</sup>In this paper, we simplify a counterpart of time-shift by means of  $\mathbf{L}$  for simplicity. Nevertheless, we can easily extend the proposed GWSS for other graph variation operators, including  $\mathbf{A}$  and  $\mathcal{L} := \mathbf{I} - \mathbf{D}^{-1/2} \mathbf{A} \mathbf{D}^{-1/2}$ .

- [2] 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, pp. 83–98, May 2013.
- [3] A. Ortega, P. Frossard, J. Kovačević, J. M. F. Moura, and P. Vandergheynst, "Graph Signal Processing: Overview, Challenges and Applications," *arXiv:1712.00468*, Mar. 2018.
- [4] X. Dong, D. Thanou, P. Frossard, and P. Vandergheynst, "Learning Laplacian Matrix in Smooth Graph Signal Representations," *IEEE Trans. Signal Process.*, vol. 64, pp. 6160–6173, Dec. 2016. Conference Name: IEEE Transactions on Signal Processing.
- [5] V. Kalofolias, "How to learn a graph from smooth signals," p. 10.
- [6] H. E. Egilmez, E. Pavez, and A. Ortega, "Graph Learning from Filtered Signals: Graph System and Diffusion Kernel Identification," *arXiv:1803.02553 [cs, stat]*, Mar. 2018.
- [7] K. Yamada, Y. Tanaka, and A. Ortega, "Time-Varying Graph Learning with Constraints on Graph Temporal Variation," *arXiv:2001.03346 [cs, eess]*, Jan. 2020. *arXiv: 2001.03346*.
- [8] X. Dong, D. Thanou, M. Rabbat, and P. Frossard, "Learning Graphs From Data: A Signal Representation Perspective," *IEEE Signal Processing Magazine*, vol. 36, pp. 44–63, May 2019. Conference Name: IEEE Signal Processing Magazine.
- [9] S. Ono, I. Yamada, and I. Kumazawa, "Total generalized variation for graph signals," in *2015 IEEE Int. Conf. Acoust. Speech Signal Process. (ICASSP)*, (South Brisbane, Queensland, Australia), pp. 5456–5460, IEEE, Apr. 2015.
- [10] M. Onuki, S. Ono, M. Yamagishi, and Y. Tanaka, "Graph Signal Denoising via Trilateral Filter on Graph Spectral Domain," *IEEE Transactions on Signal and Information Processing over Networks*, vol. 2, pp. 137–148, June 2016. Conference Name: IEEE Transactions on Signal and Information Processing over Networks.
- [11] M. Nagahama, K. Yamada, Y. Tanaka, S. H. Chan, and Y. C. Eldar, "Graph Signal Restoration Using Nested Deep Algorithm Unrolling," *arXiv:2106.15910 [cs, eess]*, June 2021. *arXiv: 2106.15910*.
- [12] S. Chen, M. Li, and Y. Zhang, "Sampling and Recovery of Graph Signals based on Graph Neural Networks," *arXiv:2011.01412 [cs, eess]*, Nov. 2020. *arXiv: 2011.01412*.
- [13] G. Cheung, E. Magli, Y. Tanaka, and M. Ng, "Graph Spectral Image Processing," *arXiv:1801.04749 [eess]*, Jan. 2018.
- [14] Y. Wang, Y. Sun, Z. Liu, S. E. Sarma, M. M. Bronstein, and J. M. Solomon, "Dynamic Graph CNN for Learning on Point Clouds," *ACM Trans. Graph.*, vol. 38, pp. 1–12, Nov. 2019.
- [15] M. M. Bronstein, J. Bruna, Y. LeCun, A. Szlam, and P. Vandergheynst, "Geometric Deep Learning: Going beyond Euclidean data," *IEEE Signal Process. Mag.*, vol. 34, pp. 18–42, July 2017.
- [16] A. Agaskar and Y. M. Lu, "A spectral graph uncertainty principle," *IEEE Trans. Inf. Theory*, vol. 59, pp. 4338–4356, July 2013.
- [17] 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.
- [18] 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, pp. 80–90, Sept. 2014.
- [19] S. K. Narang and A. Ortega, "Perfect Reconstruction Two-Channel Wavelet Filter Banks for Graph Structured Data," *IEEE Transactions on Signal Processing*, vol. 60, pp. 2786–2799, June 2012.
- [20] D. I. Shuman, M. J. Faraji, and P. Vandergheynst, "A Multiscale Pyramid Transform for Graph Signals," *IEEE Transactions on Signal Processing*, vol. 64, pp. 2119–2134, Apr. 2016.
- [21] Y. Tanaka, Y. C. Eldar, A. Ortega, and G. Cheung, "Sampling signals on Graphs: From Theory to Applications," *IEEE Signal Process. Mag.*, vol. 37, pp. 14–30, Nov. 2020.
- [22] 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, pp. 3775–3789, July 2016.
- [23] S. Chen, R. Varma, A. Sandryhaila, and J. Kovačević, "Discrete signal processing on graphs: sampling theory," *IEEE Trans. Signal Process.*, vol. 63, pp. 6510–6523, Dec. 2015.
- [24] A. G. Marques, S. Segarra, G. Leus, and A. Ribeiro, "Sampling of graph signals with successive local aggregations," *IEEE Trans. Signal Process.*, vol. 64, pp. 1832–1843, Apr. 2016.
- [25] G. Puy, N. Tremblay, R. Gribonval, and P. Vandergheynst, "Random sampling of bandlimited signals on graphs," *Applied and Computational Harmonic Analysis*, vol. 44, pp. 446–475, Mar. 2018.
- [26] M. Tsitsvero, S. Barbarossa, and P. D. Lorenzo, "Signals on graphs: Uncertainty principle and sampling," *IEEE Trans. Signal Process.*, vol. 64, pp. 4845–4860, Sept. 2016.
- [27] A. Sakiyama, Y. Tanaka, T. Tanaka, and A. Ortega, "Eigendecomposition-free sampling set selection for graph signals," *IEEE Trans. Signal Process.*, vol. 67, pp. 2679–2692, May 2019.
- [28] L. F. O. Chamon and A. Ribeiro, "Greedy Sampling of Graph Signals," *IEEE Transactions on Signal Processing*, vol. 66, pp. 34–47, Jan. 2018.
- [29] Y. Tanaka, "Spectral Domain Sampling of Graph Signals," *IEEE Trans. Signal Process.*, vol. 66, pp. 3752–3767, July 2018.
- [30] D. I. Shuman, B. Ricaud, and P. Vandergheynst, "Vertex-frequency analysis on graphs," *Applied and Computational Harmonic Analysis*, vol. 40, pp. 260–291, Mar. 2016.
- [31] A. Sandryhaila and J. M. F. Moura, "Discrete signal processing on graphs: Frequency analysis," *IEEE Trans. Signal Process.*, vol. 62, pp. 3042–3054, June 2014.
- [32] B. Girault, "Stationary graph signals using an isometric graph translation," in *2015 23rd European Signal Processing Conference (EUSIPCO)*, pp. 1516–1520, Aug. 2015.
- [33] B. S. Dees, L. Stankovic, M. Dakovic, A. G. Constantinides, and D. P. Mandic, "Unitary Shift Operators on a Graph," *arXiv:1909.05767 [cs, eess, math]*, Sept. 2019.
- [34] A. Gavili and X.-P. Zhang, "On the Shift Operator, Graph Frequency and Optimal Filtering in Graph Signal Processing," *arXiv:1511.03512 [cs, math]*, Sept. 2017.
- [35] J. Shi and J. M. F. Moura, "Graph Signal Processing: Modulation, Convolution, and Sampling," *arXiv:1912.06762 [eess]*, Dec. 2019. *arXiv: 1912.06762*.
- [36] I. Pesenson, "Sampling in Paley-Wiener spaces on combinatorial graphs," *Trans. Amer. Math. Soc.*, vol. 360, pp. 5603–5627, May 2008.
- [37] F. Wang, Y. Wang, and G. Cheung, "A-Optimal Sampling and Robust Reconstruction for Graph Signals via Truncated Neumann Series," *IEEE Signal Processing Letters*, vol. 25, pp. 680–684, May 2018.
- [38] D. K. Hammond, P. Vandergheynst, and R. Gribonval, "Wavelets on graphs via spectral graph theory," *Applied and Computational Harmonic Analysis*, vol. 30, pp. 129–150, Mar. 2011.
- [39] S. K. Narang and A. Ortega, "Compact Support Biorthogonal Wavelet Filterbanks for Arbitrary Undirected Graphs," *IEEE Transactions on Signal Processing*, vol. 61, pp. 4673–4685, Oct. 2013.
- [40] Y. Tanaka and A. Sakiyama, "\$M\S\$ -Channel Oversampled Graph Filter Banks," *IEEE Trans. Signal Process.*, vol. 62, pp. 3578–3590, July 2014.
- [41] A. Sakiyama, K. Watanabe, Y. Tanaka, and A. Ortega, "Two-channel critically sampled graph filter banks with spectral domain sampling," *IEEE Trans. Signal Process.*, vol. 67, pp. 1447–1460, Mar. 2019.
- [42] S. Chen, Y. Yang, C. Faloutsos, and J. Kovacevic, "Monitoring Manhattan's traffic at 5 intersections?," in *2016 IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, (Washington DC, DC, USA), pp. 1270–1274, IEEE, Dec. 2016.
- [43] A. Gadde, A. Anis, and A. Ortega, "Active semi-supervised learning using sampling theory for graph signals," in *Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining - KDD '14*, (New York, New York, USA), pp. 492–501, ACM Press, 2014.
- [44] A. Anis, A. E. Gamal, A. S. Avestimehr, and A. Ortega, "A Sampling Theory Perspective of Graph-Based Semi-Supervised Learning," *IEEE Trans. Inf. Theory*, vol. 65, pp. 2322–2342, Apr. 2019.
- [45] A. J. Jerri, "The Shannon sampling theorem—Its various extensions and applications: A tutorial review," *Proc. the IEEE*, vol. 65, pp. 1565–1596, Nov. 1977.
- [46] M. Unser, "Sampling—50 Years After Shannon," *Proc. IEEE*, vol. 88, no. 4, pp. 569–587, 2000.
- [47] Y. C. Eldar, *Sampling Theory: Beyond Bandlimited Systems*. Cambridge University Press, Apr. 2015.
- [48] Y. C. Eldar and T. Michaeli, "Beyond bandlimited sampling," *IEEE Signal Process. Mag.*, vol. 26, pp. 48–68, May 2009.
- [49] S. Li, Y. Jin, and D. I. Shuman, "Scalable \$M\S\$-Channel Critically Sampled Filter Banks for Graph Signals," *arXiv:1608.03171 [cs, math]*, Jan. 2019. *arXiv: 1608.03171*.
- [50] S. Chen, A. Singh, and J. Kovačević, "Multiresolution Representations for Piecewise-Smooth Signals on Graphs," *arXiv:1803.02944 [cs, eess]*, Mar. 2018. *arXiv: 1803.02944*.
- [51] R. Varma and J. Kovacevic, "Passive and Active Sampling for Piecewise-Smooth Graph Signals," in *2019 13th International conference on Sampling Theory and Applications (SampTA)*, (Bordeaux, France), pp. 1–5, IEEE, July 2019.

- [52] S. Chen, R. Varma, A. Singh, and J. Kovacevic, "Representations of piecewise smooth signals on graphs," in *2016 IEEE Int. Conf. Acoust. Speech Signal Process. (ICASSP)*, pp. 6370–6374, IEEE, Mar. 2016.
- [53] Y. C. Eldar and T. Werther, *Genral framework for consistent sampling in Hilbert spaces*, vol. 3. Int. J. Wavelets, Multiresolution Inf. Process., 3 ed., Sept. 2005.
- [54] O. Christensen and Y. C. Eldar, "Oblique dual frames and shift-invariant spaces," *Applied and Computational Harmonic Analysis*, vol. 17, pp. 48–68, July 2004.
- [55] S. P. Chepuri, Y. C. Eldar, and G. Leus, "Graph sampling with and without input priors," in *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 4564–4568, Apr. 2018.
- [56] N. Perraudin and P. Vandergheynst, "Stationary signal processing on graphs," *IEEE Trans. Signal Process.*, vol. 65, pp. 3462–3477, July 2017.
- [57] A. G. Marques, S. Segarra, G. Leus, and A. Ribeiro, "Stationary graph processes and spectral estimation," *IEEE Trans. Signal Process.*, vol. 65, pp. 5911–5926, Nov. 2017.
- [58] Y. C. Eldar and M. Unser, "Nonideal sampling and interpolation from noisy observations in shift-invariant spaces," *IEEE Trans. Signal Process.*, vol. 54, pp. 2636–2651, July 2006.
- [59] Y. C. Eldar and T. G. Dvorkind, "A minimum squared-error framework for generalized sampling," *IEEE Trans. Signal Process.*, vol. 54, pp. 2155–2167, June 2006.
- [60] A. Sakiyama and Y. Tanaka, "Graph reduction method using localization operator and its application to pyramid transform," in *2017 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC)*, (Kuala Lumpur), pp. 050–055, IEEE, Dec. 2017.
- [61] S. Segarra, A. G. Marques, G. Leus, and A. Ribeiro, "Stationary graph processes: Parametric power spectral estimation," in *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, (New Orleans, LA), pp. 4099–4103, IEEE, Mar. 2017.
- [62] A. Papoulis and S. U. Pillai, *Probability, Random Variables, and Stochastic Processes, Fourth Edition*. McGraw-Hill Higher Education (a division of McGraw-Hill Companies), 2002.
- [63] X. Jian and W. P. Tay, "Wide-Sense Stationarity in Generalized Graph Signal Processing," *arXiv:2112.01127 [eess]*, Dec. 2021.
- [64] A. Hasanzadeh, X. Liu, N. Duffield, and K. R. Narayanan, "Piecewise Stationary Modeling of Random Processes Over Graphs With an Application to Traffic Prediction," *arXiv:1711.06954 [eess]*, Sept. 2019.
- [65] E. M. Stein and R. Shakarchi, *Real analysis: measure theory, integration, and Hilbert spaces*. No. v. 3 in Princeton lectures in analysis, Princeton, N.J: Princeton University Press, 2005.
- [66] S. Boyd, S. P. Boyd, and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, Mar. 2004.
- [67] A. Heimowitz and Y. C. Eldar, "Smooth Graph Signal Interpolation for Big Data," *IEEE Transactions on Signal Processing*, June 2018.
- [68] S. K. Narang, A. Gadde, and A. Ortega, "Signal processing techniques for interpolation in graph structured data," in *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 5445–5449, May 2013.
- [69] 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, pp. 4363–4378, Aug. 2016.
- [70] S. K. Narang, A. Gadde, E. Sanou, and A. Ortega, "Localized iterative methods for interpolation in graph structured data," in *2013 IEEE Global Conference on Signal and Information Processing*, pp. 491–494, Dec. 2013.
- [71] 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:1408.5781*, Aug. 2014.
- [72] N. A. Rayner, D. E. Parker, E. B. Horton, C. K. Folland, L. V. Alexander, D. P. Rowell, E. C. Kent, and A. Kaplan, "Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century," *Journal of Geophysical Research: Atmospheres*, vol. 108, no. D14, 2003.
- [73] N. Wiener, "Generalized harmonic analysis," *Acta Mathematica*, vol. 55, pp. 117–258, 1930.
- [74] B. Girault, *Signal Processing on Graphs - Contributions to an Emerging Field*. PhD Thesis, Ecole normale superieure de lyon - ENS LYON, Dec. 2015.
- [75] K. C. Das, "Extremal graph characterization from the bounds of the spectral radius of weighted graphs," *Applied Mathematics and Computation*, vol. 217, pp. 7420–7426, May 2011.