WIDE-SENSE STATIONARITY AND SPECTRAL ESTIMATION FOR GENERALIZED GRAPH SIGNAL

Xingchao Jian, and Wee Peng Tay, Senior Member, IEEE

School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore

ABSTRACT

We consider a probabilistic model for graph signal processing (GSP) in a generalized framework where each vertex of a graph is associated with an element from a Hilbert space. We introduce the notion of joint wide-sense stationarity in this generalized GSP (GGSP) framework, which allows us to characterize a random graph process as a combination of uncorrelated oscillation modes across both the vertex and Hilbert space domains. We also propose a method for joint power spectral density estimation in case of missing features. Experiment results corroborate the effectiveness of our estimation approach.

***Index Terms*—** Graph signal processing, Hilbert space, wide-sense stationarity, power spectral density.

1. INTRODUCTION

In the studies regarding biological network [1], sensor network [2, 3] and image processing [4, 5], it is beneficial to assign the data with a latent graph structure, so that the data of interest becomes a graph signal, i.e. a map from vertex set to real number ($\mathbb{R}$) or complex number ($\mathbb{C}$). Graph signal processing (GSP) techniques [6, 7] have been developed to tackle the problems such as sampling [8–10] and denoising [5] for such signals. A notable statistical model in GSP is graph wide-sense stationary (GWSS) signal [11–13], which is an analogy of wide-sense stationary (WSS) for time series. This model defines frequency to be the graph signal's variation along the edges of the graph, and assumes that the signal consists of uncorrelated oscillation modes on the graph.

In order to leverage the information contained in both time and vertex domain, the time-vertex model and its corresponding concept of WSS were introduced [14, 15]. The signal model is further generalized to a possibly infinite dimensional Hilbert space by [16], which is referred to as generalized GSP (GGSP) framework. It encompasses a variety of signals such as multichannel signal, discrete- and continuous-time graph signals, but the randomness is not considered in [16].

In this paper we introduce the statistical model for GGSP framework. We define the random graph process as a random element in a signal space, which is related to both graph structure and the Hilbert space where the vertex observation comes from. Given general shift operators on separate domains (i.e. vertex domain and Hilbert space domain), we define WSS for this framework via the Fourier basis induced by them, and elucidate the relations between WSS in different domains. We also provide a strategy for power spectral density (PSD) estimation which allows for parallelized computation. Numerical experiment demonstrates that this strategy can be utilized under a large proportion of missing features and large sample sizes.

This research is supported by the Singapore Ministry of Education Academic Research Fund Tier 2 grant MOE-T2EP20220-0002.

2. PRELIMINARIES

In this preliminary section, we briefly introduce the definition of a random element and its moments. They are, respectively, generalizations of the random vector and its moments in a Euclidean space to a Hilbert space. Readers are referred to [17–19] for more details. These concepts are the basic ingredients for the statistical model of random graph process in Section 3 and the notion of WSS in Section 4.

Consider a probability space $(\Omega, \mathcal{F}, \mu)$, where Ω is a sample space, $\mathcal{F}$ a σ-algebra and μ a probability measure, and a complex separable Hilbert space $\mathcal{H}$. A random element is defined as a measurable map $X : \Omega \mapsto \mathcal{H}$.

We assume that $\mathbb{E}\|X\|^2 < \infty$ for all random elements X throughout this paper. Under this assumption, X's mean element m_X and covariance operator $\mathbf{C}_X$ are well-defined as follows: for all $u, v \in \mathcal{H}$,

$$\langle m_X, u\rangle = \mathbb{E}[\langle X, u\rangle], \tag{1}$$

$$\langle \mathbf{C}_X(u), v\rangle = \mathbb{E}\Big[\overline{\langle X - m_X, u\rangle}\langle X - m_X, v\rangle\Big]. \tag{2}$$

It can be verified that the above definitions degenerate to the standard definitions of mean and covariance of random vectors in finite dimensional Hilbert spaces. To preclude the pathological cases (e.g., a uncorrelated complex random process' real and imaginary parts can be correlated), in this paper we only consider those random elements that are proper, i.e. satisfy the equivalent conditions in [20, Theorem 1].

3. GGSP STATISTICAL MODEL

In this section, we present the definition of random graph process in the GGSP framework. Let $\mathscr{G} = (\mathscr{V}, \mathscr{E})$ be an undirected (weighted) graph, where $\mathscr{V} = \{1, \dots, n\}$ is the vertex set, and $\mathscr{E} \subset \{(i,j) \in \mathscr{V} \times \mathscr{V} : i < j\}$ denotes the edge set. Suppose that each vertex of $\mathscr{G}$ is associated with an element from a separable Hilbert space $\mathcal{H}$. A generalized graph signal [16] has the form $x = (x_1, x_2, \dots, x_n)$ where each $x_v \in \mathcal{H}$, $v \in \mathscr{V}$. For example, if $\mathcal{H} = L^2[a,b]$ the space of square integrable functions on a bounded interval $[a,b]$, then at each vertex v of the graph $\mathscr{G}$, we have associated with it such a function $x_v(\cdot) \in L^2[a,b]$.

It is shown in [16] that the space of all generalized graph signals can be identified with the Hilbert space $\mathbb{C}^n \otimes \mathcal{H}$ via the isomorphism

$$x \cong \sum_{v=1}^{n} \delta_v \otimes x_v,$$

where $\{\delta_v : v = 1, \dots, n\}$ is the standard basis in $\mathbb{C}^n$. Suppose $\mathcal{H} = L^2(\Upsilon)$ for some set Υ. For $x \in \mathbb{C}^n \otimes \mathcal{H}$ and each $t \in \Upsilon$, we write

$$x(t) = (x_1(t), x_2(t), \cdots, x_n(t))^\intercal, \tag{3}$$

a vector-valued function.

Definition 1. *Consider a probability space $(\Omega, \mathcal{F}, \mu)$, where Ω is a sample space, $\mathcal{F}$ a σ-algebra and μ a probability measure. A* random graph process *X is a measurable map (i.e., a random element) from Ω to $\mathbb{C}^n \otimes \mathcal{H}$.*

For a random graph process X, if $\mathcal{H} = L^2(\Upsilon)$, we can define its pointwise mean $m_X(t) := \mathbb{E}[X(t)]$ and cross-covariance $\mathbf{K}_X(s,t) := \operatorname{cov}(X(s), X(t))$ for every $s, t \in \Upsilon$.

4. GENERALIZED JOINT WIDE-SENSE STATIONARITY

In this section, we develop the notion of WSS with respect to (w.r.t.) a shift operator for a random graph process. To this end, we need to define a shift operator on $\mathbb{C}^n \otimes \mathcal{H}$. Suppose we are given self-adjoint and compact operators $\mathbf{A}_\mathscr{G}$ on $\mathscr{G}$ and $\mathbf{A}_\mathcal{H}$ on $\mathcal{H}$. In this paper, we adopt $\mathbf{A}_\mathscr{G} \otimes \mathbf{A}_\mathcal{H}$ as the shift operator for $\mathbb{C}^n \otimes \mathcal{H}$ as in [16]. Suppose $\mathbf{A}_\mathscr{G}$ has eigenvalues $\lambda_1 \le \lambda_2 \cdots \le \lambda_n$ with corresponding eigenvectors $\{\phi_k\}_{k=1}^n$ and $\mathbf{A}_\mathcal{H}$ has eigenvalues $\nu_1 \le \nu_2 \le \cdots$ with corresponding eigenvectors $\{\psi_\tau\}_{\tau=1}^\infty$. Then, $\mathbf{A}_\mathscr{G} \otimes \mathbf{A}_\mathcal{H}$ has eigenvalues $\{\lambda_k \nu_\tau : k = 1, \dots, n, \tau \ge 1\}$ and eigenvectors $\{\phi_k \otimes \psi_\tau : k = 1, \dots, n, \tau \ge 1\}$. To simplify the exposition and to obtain a unique orthonormal eigenbasis (up to multiples of ± 1), similar to most of the GSP literature [12, 21], we make the following assumption throughout this paper.

Assumption 1. *The geometric multiplicity of each eigenvalue of $\mathbf{A}_\mathscr{G} \otimes \mathbf{A}_\mathcal{H}$ is one.*

The joint Fourier transform of a signal $x \in \mathbb{C}^n \otimes \mathcal{H}$ is then defined as its inner product with an eigenvector of the shift operator [16]: for $k = 1, \dots, n$ and $\tau \ge 1$,

$$\mathcal{F}_x(\phi_k \otimes \psi_\tau) = \langle x, \phi_k \otimes \psi_\tau \rangle. \tag{4}$$

A time domain scalar-valued stochastic process $x = (x_1, \dots, x_m)$ is said to be WSS if $\operatorname{cov}(x_i, x_j)$ only depends on $j - i$. In other words, $\operatorname{cov}(x)$ can be diagonalized by DFT matrix and its conjugate. A graph signal $x \in \mathbb{R}^n$ is defined to be GWSS if $\operatorname{cov}(x)$ can be diagonalized by graph Fourier transform matrix and its transpose [12]. As an analogy to these definitions, we define WSS on $\mathbb{C}^n \otimes \mathcal{H}$ as follows.

Definition 2. *A random graph process X is joint wide-sense stationary (JWSS) if*

$$\mathbf{C}_X \circ (\mathbf{A}_\mathscr{G} \otimes \mathbf{A}_\mathcal{H}) = (\mathbf{A}_\mathscr{G} \otimes \mathbf{A}_\mathcal{H}) \circ \mathbf{C}_X. \tag{5}$$

In other words, from Assumption 1 and [22, Chapter 4, Exercise 35 (a)], $\mathbf{C}_X$ and the shift operator $\mathbf{A}_\mathscr{G} \otimes \mathbf{A}_\mathcal{H}$ have the same complete orthonormal set of eigenvectors with

$$\mathbf{C}_X = \sum_{k,\tau} p_X(k,\tau) \mathbf{\Pi}_{\phi_k \otimes \psi_\tau}, \tag{6}$$

where $\{p_X(k,\tau) : k = 1, \dots, n, \tau \ge 1\}$ are the eigenvalues of $\mathbf{C}_X$, also known as the joint power spectral density (JPSD) *of X. If $\{(k,\tau) : p_X(k,\tau) > 0\}$ is a finite set, X is said to be* bandlimited*.*

According to [17, Theorem 7.2.6, Theorem 7.2.7], the Fourier coefficients of a JWSS signal X are uncorrelated. Like in [14], we can define WSS in separate domains, i.e. vertex domain and Hilbert space domain.

Definition 3. *Given a random graph process X with $\mathcal{H} = L^2(\Upsilon)$, we say that X is vertex wide-sense stationary (VWSS) if*

$$\mathbf{K}_X(t,t) \mathbf{A}_\mathscr{G} = \mathbf{A}_\mathscr{G} \mathbf{K}_X(t,t)$$

for every $t \in \Upsilon$.

Definition 4. *A random graph process X is said to be Hilbert space wide-sense stationary (HWSS) if*

$$\boldsymbol{\delta}_m^\mathcal{H} \mathbf{C}_X \boldsymbol{\delta}_m^\mathcal{H} \circ (\mathbf{I}_n \otimes \mathbf{A}_\mathcal{H}) = (\mathbf{I}_n \otimes \mathbf{A}_\mathcal{H}) \circ \boldsymbol{\delta}_m^\mathcal{H} \mathbf{C}_X \boldsymbol{\delta}_m^\mathcal{H} \tag{7}$$

for all $m = 1, \cdots, n$. Here, $\boldsymbol{\delta}_m^\mathcal{H} := \operatorname{diag}(\delta_m) \otimes \mathbf{I}_\mathcal{H}$ where δ_m is the m-th standard basis vector in $\mathbb{C}^n$ consisting of all zeros except a one at the m-th entry.

Definitions 3 and 4 can be regarded as traditional WSS definitions embedded in the GGSP framework. In the following theorem we show that JWSS implies WSS in both the vertex and Hilbert space domains. The proof is omitted due to the lack of space.

Theorem 1. *A JWSS random graph process X with $\mathcal{H} = L^2(\Upsilon)$ is both VWSS and HWSS.*

5. POWER SPECTRAL DENSITY ESTIMATION

In this section, we demonstrate that JWSS links the PSD values on different domains, i.e., the vertex domain, Hilbert space domain and joint domain. This relationship allows us to estimate the JPSD even if there are missing features on every sample. This situation has not been investigated to the best of our knowledge. To be specific, it only requires separate features' vertex domain PSD, therefore can be implemented in a parallelized way. In this section, we focus on the case where $\dim \mathcal{H} = d < \infty$.

Suppose each vertex observes a signal consisting of d features. Let $\mathbf{X} = (x_1, \dots, x_n)^\intercal \in \mathbb{C}^{n\times d}$ be a representation of a zero-mean JWSS random graph process of interest. The i-th row $x_i^\intercal$ of $\mathbf{X}$ represents the d-dimensional signal observed at the i-th vertex, and the j-th column of $\mathbf{X}$ represents the n-dimensional graph signal under the j-th feature. In this case, every x_i belongs to $L^2(\Upsilon)$ where $\Upsilon = \{1, 2, \cdots, d\}$. Let $\mathbf{\Psi} = (\psi_1, \dots, \psi_d) = (\tilde{\psi}_1, \dots, \tilde{\psi}_d)^\intercal$ and $\mathbf{\Phi} = (\phi_1, \dots, \phi_n) = (\tilde{\phi}_1, \dots, \tilde{\phi}_n)^\intercal$ be matrices whose columns are the eigenvectors of $\mathbf{A}_\mathcal{H}$ and $\mathbf{A}_\mathscr{G}$, respectively. Differently from the time-vertex framework [14, 23] which presumes $\mathbf{A}_\mathcal{H}$ as the translation operator, in GGSP framework $\mathbf{A}_\mathcal{H}$ can be inferred from the dataset, e.g. the empirical covariance of vertex signal.

Suppose the random graph process $\mathbf{X}$ is JWSS. The Fourier transform, in this case, can be written as

$$\mathcal{F}_\mathbf{X} = \mathbf{\Phi}^\intercal \mathbf{X} \bar{\mathbf{\Psi}}, \tag{8}$$

which in vectorized form $\operatorname{vec}(\mathcal{F}_\mathbf{X}) = \mathbf{\Psi}^* \otimes \mathbf{\Phi}^\intercal \operatorname{vec}(\mathbf{X})$ corresponds to (4). The entries of $\mathcal{F}_\mathbf{X}$ are uncorrelated according to previous discussion.

Suppose m independent and identically distributed (i.i.d.) copies of $\mathbf{X}$, $\{\mathbf{X}_i\}_{i=1}^m$, are observed. A natural way to estimate the PSD $p_X(k, \tau)$, for $k = 1, \dots, n$ and $\tau = 1, \dots, d$, is averaging the squared Fourier coefficients:

$$\hat{p}_X(k, \tau) = \frac{1}{m} \sum_{i=1}^{m} |\phi_k^\intercal \mathbf{X}_i \bar{\psi}_\tau|^2. \tag{9}$$

This method corresponds to the periodogram estimator of time signals and graph signals [12], and requires that there are no missing entries in each observed $\mathbf{X}_i$. In the following, we estimate $\hat{p}_X(k, \tau)$ from $\{\mathbf{X}_i\}_{i=1}^m$ with possible missing features in every sample.

Let $\mathbf{C}_X$ be the covariance of $\operatorname{vec}(\mathbf{X})$ and p_X be a vector that contains the JPSD values of $\mathbf{X}$, written in a block form so that $p_X = (r_{X,1}^\intercal, \dots, r_{X,d}^\intercal)^\intercal$, where every $r_{X,i} \in \mathbb{R}^n$ is a column vector

$$r_{X,i} = (p_X(1, i), p_X(2, i), \cdots, p_X(n, i))^\intercal.$$

By vectorizing both sides of (8) and computing their covariances, we obtain

$$\mathbf{C}_X = (\mathbf{\Psi} \otimes \mathbf{\Phi}) \operatorname{diag}(p_X)(\mathbf{\Psi}^* \otimes \mathbf{\Phi}^\intercal). \tag{10}$$

From the perspective of vertex signal, according to Theorem 1, every x_i is HWSS. We write $p^{(i)} \in \mathbb{R}^d$ to represent the Hilbert space domain PSD of x_i. Without loss of generality, we investigate the signal observed on the 1st vertex, i.e. x_1, and its Hilbert space domain PSD $p^{(1)}$. For ease of notation we write y to substitute x_1 in this section. Then the covariance matrix of y has the decomposition

$$\mathbf{C}_y = \mathbf{\Psi} \operatorname{diag}(p^{(1)}) \mathbf{\Psi}^*. \tag{11}$$

To investigate the relationship between $p^{(1)}$ and p_X, first notice that $y^\intercal = \delta_1^\intercal \mathbf{X}$, implying that $\mathbf{C}_y$ and $\mathbf{C}_X$ are linearly related. To see this, it suffices to vectorize both sides of $y^\intercal = \delta_1^\intercal \mathbf{X}$ and take the covariance:

$$\mathbf{C}_y = (\mathbf{I}_d \otimes \delta_1^\intercal) \mathbf{C}_X (\mathbf{I}_d \otimes \delta_1). \tag{12}$$

Now we have derived three linear relations (10) to (12): $\mathbf{C}_X \sim p_X$, $\mathbf{C}_y \sim p^{(1)}$ and $\mathbf{C}_y \sim \mathbf{C}_X$. Let $|a|^2$ denote the component-wise squared norm of a vector a, and p_i denotes the vertex domain PSD of the ith measurement x_i. The Hilbert space domain and vertex domain PSDs of the graph vertices are linearly related to the JPSD p_X as follows:

Theorem 2. *The PSD values of signals on all vertices can be linearly expressed by the JPSD p_X as*

$$\begin{pmatrix} p^{(1)} \\ p^{(2)} \\ \cdots \\ p^{(n)} \end{pmatrix} = \begin{pmatrix} \mathbf{I}_d \otimes |\tilde{\phi}_1^\intercal|^2 \\ \mathbf{I}_d \otimes |\tilde{\phi}_2^\intercal|^2 \\ \cdots \\ \mathbf{I}_d \otimes |\tilde{\phi}_n^\intercal|^2 \end{pmatrix} p_X. \tag{13}$$

The corresponding result also applies to vertex domain PSD values and JPSD:

$$\begin{pmatrix} p_1 \\ p_2 \\ \cdots \\ p_d \end{pmatrix} = \begin{pmatrix} |\tilde{\psi}_1^\intercal|^2 \otimes \mathbf{I}_n \\ |\tilde{\psi}_2^\intercal|^2 \otimes \mathbf{I}_n \\ \cdots \\ |\tilde{\psi}_d^\intercal|^2 \otimes \mathbf{I}_n \end{pmatrix} p_X. \tag{14}$$

Proof. Here we only need to prove (13), since (14) can be proved with the same approach.

By substituting (10) into (12) we have

$$\mathbf{C}_y = (\mathbf{I}_d \otimes \delta_1^\intercal)(\mathbf{\Psi} \otimes \mathbf{\Phi}) \operatorname{diag}(p_X)(\mathbf{\Psi}^* \otimes \mathbf{\Phi}^\intercal)(\mathbf{I}_d \otimes \delta_1).$$

To further exploit the linear relation, we vectorize both sides of it:

$$\begin{aligned} &\operatorname{vec}(\mathbf{C}_y) && (15)\\ &= (\mathbf{I}_d \otimes \delta_1^\intercal)(\bar{\mathbf{\Psi}} \otimes \mathbf{\Phi}) \odot (\mathbf{I}_d \otimes \delta_1^\intercal)(\mathbf{\Psi} \otimes \mathbf{\Phi}) p_X \\ &= ((\mathbf{I}_d \otimes \delta_1^\intercal) \otimes (\mathbf{I}_d \otimes \delta_1^\intercal))((\bar{\mathbf{\Psi}} \otimes \mathbf{\Phi}) \odot (\mathbf{\Psi} \otimes \mathbf{\Phi})) p_X, && (16) \end{aligned}$$

where $\odot$ represents the Khatri-Rao product. Here we have used the properties:

1. $\mathrm{vec}(\mathbf{A}\,\mathrm{diag}(p)\mathbf{B}) = (\mathbf{B}^\intercal \odot \mathbf{A})p$;

2. $(\mathbf{A}\otimes\mathbf{B})(\mathbf{C}\odot\mathbf{D}) = (\mathbf{AC}\odot\mathbf{BD})$,

for arbitrary matrices $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$, $\mathbf{D}$ and vector p.

To simplify $\mathrm{vec}(\mathbf{C}_y)$, let $\mathbf{U} = (\bar{\mathbf{\Psi}}\otimes\mathbf{\Phi})\odot(\mathbf{\Psi}\otimes\mathbf{\Phi})$. Since each column of $\mathbf{\Psi}\otimes\mathbf{\Phi}$ can be written as $\eta_{i,j} := \psi_i\otimes\phi_j$, we can write $\mathbf{U}$ in columns:

$$\mathbf{U} = (\bar{\eta}_{1,1}\otimes\eta_{1,1}, \bar{\eta}_{1,2}\otimes\eta_{1,2}, \cdots, \bar{\eta}_{N,d}\otimes\eta_{N,d})$$

For each column of $\mathbf{U}$,

$$\begin{aligned}&(\mathbf{I}_d\otimes\delta_1^\intercal)\otimes(\mathbf{I}_d\otimes\delta_1^\intercal)(\bar{\eta}_{i,j}\otimes\eta_{i,j})\\ &= (\mathbf{\Phi}_{1j})^2\bar{\psi}_i\otimes\psi_i.\end{aligned}$$

Note that $\{\bar{\psi}_i\otimes\psi_i : i = 1,\cdots,d\}$ form an orthonormal set in $\mathbb{C}^{d^2}$, hence are linearly independent. According to (16), this result implies that $\mathrm{vec}(\mathbf{C}_y) \in \mathrm{span}\{\bar{\psi}_i\otimes\psi_i : i = 1,\cdots,d\}$. To be specific, we have $\mathrm{vec}(\mathbf{C}_y) = \sum_{i=1}^d \alpha_i\bar{\psi}_i\otimes\psi_i$, in which $\alpha_i = |\tilde{\phi}_1^\intercal|^2 r_{X,i}$.

On the other hand, from (11) we know that $\mathrm{vec}(\mathbf{C}_y) = \bar{\mathbf{\Psi}}\odot\mathbf{\Psi}p^{(1)}$. Therefore $p^{(1)} = (\alpha_1,\cdots,\alpha_d)^\intercal$, i.e.,

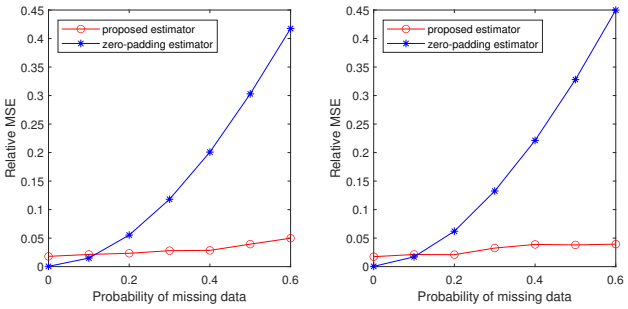
$$\begin{aligned}p^{(1)} &= \begin{pmatrix}|\tilde{\phi}_1^\intercal|^2 & & & \\ & |\tilde{\phi}_1^\intercal|^2 & & \\ & & \ddots & \\ & & & |\tilde{\phi}_1^\intercal|^2\end{pmatrix}\begin{pmatrix}r_{X,1}\\ r_{X,2}\\ \cdots\\ r_{X,d}\end{pmatrix}\\ &= (\mathbf{I}_d\otimes|\tilde{\phi}_1^\intercal|^2)p_X,\end{aligned}$$

which concludes the proof. □

It can be seen from (11) that each $p^{(i)}$ on the left-hand side (L.H.S.) of (13) can be computed by each vertex i separately and does not rely on the graph structure. Similarly, (14) can be utilized to estimate p_X via computing p_i separately on each feature (e.g. time step), rather than computing on the joint domain (e.g. time-vertex domain). Here we note that (13) cannot be utilized to estimate p_X if $|\mathbf{\Phi}|^2$ is not full-rank. For instance, when the graph Laplacian $\mathbf{L}$ is adopted as $\mathbf{A}_{\mathscr{G}}$, (13) contains at most $(nd-d)$ independent equations hence cannot be solved.

6. NUMERICAL EXPERIMENT

We test the quality of the JPSD estimator via solving (14). The underlying graph is generated by the Erdős–Rényi model with $n = 30$ vertices and connection probability 0.3 and then fixed throughout the experiment. Each vertex is assumed to observe a d-dimensional zero-mean random vector consisting of $d = 10$ features, hence the resulting sample set can be written as $\{\mathbf{X}_i : i = 1,\ldots,m\}$, where $\mathbf{X}_i = (X_{i,1},\ldots,X_{i,N})^\intercal \in \mathbb{R}^{N\times d}$. The orthonormal basis $\mathbf{\Psi}$ of $\mathcal{H} = \mathbb{R}^d$ is randomly generated and also fixed throughout



(a) JPSD in form (i) (b) JPSD in form (ii)

Fig. 1. Performance of JPSD estimators under different ground-truth JPSD forms. The sample size in this experiment is $m = 4000$.

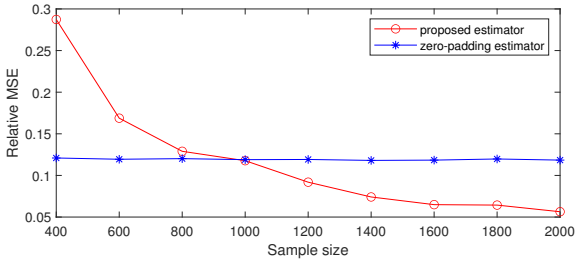


Fig. 2. Performance of JPSD estimators under varying sample sizes. The ground-truth JPSD in this experiment is (i).

the experiment. We consider two choices of ground-truth JPSD values: (i) $p_X(k,\tau) = |\sin(k + N(\tau-1))|$; (ii) $p_X(k,\tau) = \exp(-k\tau/10)$.

In this experiment, we randomly mask the observations and estimate $\mathbf{X}$'s JPSD from the partial observations. In particular, we assume that each feature measurement is unavailable in each sample $\mathbf{X}_i$ with probability p. We test two methods to recover JPSD: 1) by solving (14), and 2) by filling in the missing values with their corresponding mean values (i.e., 0) first, and then estimating the JPSD using the empirical mean of the squared Fourier coefficients (9). We take the relative mean-square error (MSE) given by

$$\text{Relative MSE} = \frac{\mathbb{E}\|\hat{p}_X - p_X\|^2}{\|p_X\|^2}$$

as the performance metric.

The results in Figs. 1 and 2 are averaged over 20 experiments. From Fig. 1 we observe that when the percentage of missing data is large, the proposed method via solving (14) outperforms the empirical estimation approach by (9), and this advantage does not rely on the ground-truth JPSD. From Fig. 2 it is clear that when a large sample set is available , the proposed method's result converges to the ground-truth JPSD, while the empirical approach cannot reach it due to the existence of bias.

7. REFERENCES

[1] Weiyu Huang, Leah Goldsberry, Nicholas F. Wymbs, Scott T. Grafton, Danielle S. Bassett, and Alejandro Ribeiro, "Graph frequency analysis of brain signals," *IEEE J. Sel. Topics Signal Process.*, vol. 10, no. 7, pp. 1189–1203, Aug. 2016.

[2] Chien-Cheng Tseng and Su-Ling Lee, "A graph Fourier transform based method for missing temperature data detection," in *2019 IEEE 8th Global Conference on Consumer Electronics (GCCE)*, Osaka, Japan, Oct. 2019.

[3] Théo Gnassounou, Pierre Humbert, and Laurent Oudre, "Adaptive subsampling of multidomain signals with product graphs," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing*, Toronto, Canada, jun 2021.

[4] Gene Cheung, Enrico Magli, Yuichi Tanaka, and Michael K. Ng, "Graph spectral image processing," *Proc. IEEE*, vol. 106, no. 5, pp. 907–930, Apr. 2018.

[5] Ali Can Yağan and Mehmet Tankut Özgen, "Spectral graph based vertex-frequency Wiener filtering for image and graph signal denoising," *IEEE Trans. Signal Inf. Process. Netw.*, vol. 6, pp. 226–240, Feb. 2020.

[6] David I Shuman, Sunil K. Narang, Pascal Frossard, Antonio Ortega, and Pierre 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, Apr. 2013.

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

[8] PaoloDi Lorenzo, Sergio Barbarossa, and Paolo Banelli, "Chapter 9 - sampling and recovery of graph signals," in *Cooperative and Graph Signal Processing*. Academic Press, 2018.

[9] Sijie Lin, Xuan Xie, Hui Feng, and Bo Hu, "Active sampling for approximately bandlimited graph signals," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing*, Brighton, UK, May 2019.

[10] Feng Ji, Pratibha, and Wee Peng Tay, "On folded graph signals," in *Proc. IEEE Global Conf. on Signal and Information Processing*, Ottawa, Canada, Nov. 2019.

[11] Benjamin Girault, "Stationary graph signals using an isometric graph translation," in *Proc. of European Signal Processing Conference (EUSIPCO)*, Nice, France, Aug./Sept. 2015.

[12] Antonio G. Marques, Santiago Segarra, Geert Leus, and Alejandro Ribeiro, "Stationary graph processes and spectral estimation," *IEEE Trans. Signal Process.*, vol. 65, no. 22, pp. 5911–5926, Aug. 2017.

[13] Nathanaël Perraudin and Pierre Vandergheynst, "Stationary signal processing on graphs," *IEEE Trans. Signal Process.*, vol. 65, no. 13, pp. 3462–3477, Apr. 2017.

[14] Nathanaël Perraudin, Andreas Loukas, Francesco Grassi, and Pierre Vandergheynst, "Towards stationary time-vertex signal processing," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing*, New Orleans, US, Mar. 2017.

[15] Andreas Loukas and Nathanaël Perraudin, "Stationary time-vertex signal processing," *EURASIP J. Adv. Signal Process*, vol. 2019, no. 1, pp. 1–19, Aug. 2019.

[16] Feng Ji and Wee Peng Tay, "A Hilbert space theory of generalized graph signal processing," *IEEE Trans. Signal Process.*, vol. 67, no. 24, pp. 6188–6203, Nov. 2019.

[17] Tailen Hsing and Randall Eubank, *Theoretical Foundations of Functional Data Analysis, With an Introduction to Linear Operators*, John Wiley & Sons, 2015.

[18] Charles R. Baker, "Joint measures and cross-covariance operators," *Transactions of the American Mathematical Society*, vol. 186, pp. 273–289, Dec. 1973.

[19] N. N. Vakhania, V. I. Tarieladze, and S. A. Chobanyan, *Probability distributions on Banach spaces*, Springer Science & Business Media, 1987.

[20] N. N. Vakhania and N. P. Kandelaki, "Random vectors with values in complex Hilbert spaces," *Theory Prob. and its Applications*, vol. 41, no. 1, pp. 116—-131, Feb. 1995.

[21] Aliaksei Sandryhaila and José M. F. Moura, "Discrete signal processing on graphs," *IEEE Trans. Signal Process.*, vol. 61, no. 7, pp. 1644–1656, Jan. 2013.

[22] Elias M. Stein and Rami Shakarchi, *Real Analysis*, Princeton, Princeton, US, 2005.

[23] Francesco Grassi, Andreas Loukas, Nathanaël Perraudin, and Benjamin Ricaud, "A time-vertex signal processing framework: Scalable processing and meaningful representations for time-series on graphs," *IEEE Trans. Signal Process.*, vol. 66, no. 3, pp. 817–829, Nov. 2018.