

SAMPLING THEORY FOR GRAPH SIGNALS

Siheng Chen^{1,2}, Aliaksei Sandryhaila³, Jelena Kovačević^{1,2,4}

¹Dept. of ECE, ²Center for Bioimage Informatics, ³HP Vertica, ⁴Dept. of BME,
Carnegie Mellon University,
Pittsburgh, PA, USA

ABSTRACT

We propose a sampling theory for finite-dimensional vectors with a generalized bandwidth restriction, which follows the same paradigm of the classical sampling theory. We use this general result to derive a sampling theorem for bandlimited graph signals in the framework of discrete signal processing on graphs. By imposing a specific structure on the graph, graph signals reduce to finite discrete-time or discrete-space signals, effectively ensuring that the proposed sampling theory works for such signals. The proposed sampling theory is applicable to both directed and undirected graphs, the assumption of perfect recovery is easy both to check and to satisfy, and, under that assumption, perfect recovery is guaranteed without any probability constraints or any approximation.

Index Terms— Sampling theory, discrete signal processing on graphs

1. INTRODUCTION

The theory of time signal processing is the foundation of our discipline [1]. It consists of four closely related variants depending on the nature of time domain: infinite and finite discrete-time signals (sequences indexed by integers) and infinite and finite continuous-time signals (functions of a real variable). Each case has its own notion of filtering, spectrum, and Fourier transform. The connection between the discrete and continuous domains is through *sampling*, which produces a sequence from a function, and *interpolation*, which produces a function from a sequence. The ability to sample a function, manipulate the resulting sequence with a discrete-time system, and then interpolate to produce a function is the foundation of digital signal processing. Conversely, the ability to interpolate a sequence to create a function, manipulate the resulting function with a continuous-time system, and then sample to produce a sequence is the foundation of digital communications. As the bridge connecting sequences and functions, the classical sampling theory shows that a bandlimited function can be perfectly recovered from its sampled sequence if the sampling rate is high enough [2].

More generally, we can treat any decrease in dimension via a linear operator as sampling, and, conversely, any increase in dimension via a linear operator as interpolation [1]; formulating a sampling theory in this context means moving between the higher- and lower-dimensional spaces while ensuring perfect recovery. In this paper, we first consider the sampling and interpolation of finite-dimensional vectors. Following Chapter 5 from [1] and the same paradigm as the classical sampling theory, we consider sampling and interpolation of finite-dimensional vectors, and propose a sampling theory for *bandlimited* finite-dimensional vectors. We next apply this result to discrete signal processing on graphs.

Discrete signal processing on graphs is a theoretical framework that generalizes classical discrete signal processing from regular domains, such as lines and rectangular lattices, to irregular structures that are commonly described by graphs [3, 4]; it provides the standard signal processing concepts to graphs, including graph signal, graph filtering, and graph Fourier transform domain. When finite-dimensional vectors represent graph signals, we can use the proposed sampling theory to perfectly recover those graph signals that are *bandlimited*, that is, they have limited support in the graph Fourier transform domain. By imposing a specific structure on the graph, graph signals reduce to finite discrete-time or discrete-space signals, effectively ensuring that the proposed sampling theory works for such signals.

Previous work on sampling for graph signals applies to undirected graphs only [5, 6, 7]. The main contributions of this paper are thus: (1) we formulate a sampling theory for bandlimited finite-dimensional vectors that follows the same paradigm of the classical sampling theory; (2) the proposed sampling theory is applicable to signals supported on either directed or undirected graphs; (3) the assumption in the proposed sampling theory is easy to check and easy to satisfy; (4) perfect recovery is guaranteed without any probability constraints or any approximation as in compressed sensing [8].

2. DISCRETE SIGNAL PROCESSING ON GRAPHS

In this section, we briefly review relevant concepts of discrete signal processing on graphs; a thorough introduction can be found in [3, 4]. It is a theoretical framework that generalizes classical discrete signal processing from regular domains,

The authors gratefully acknowledge support from the NSF through awards 1130616 and 1017278, and CMU Carnegie Institute of Technology Infrastructure Award.

such as lines and rectangular lattices, to irregular structures that are commonly described by graphs. Among its applications are signal compression, prediction and classification, semi-supervised learning and data recovery [9, 10, 11].

Graph Shift. Discrete signal processing on graphs studies signals with complex, irregular structure represented by a graph $G = (\mathcal{V}, A)$, where $\mathcal{V} = \{v_0, \dots, v_{N-1}\}$ is the set of nodes and $A \in \mathbb{C}^{N \times N}$ is a *graph shift* (weighted adjacency matrix). It represents the connections of the graph G , which can be either directed or undirected (note that the standard graph Laplacian matrix can only represent undirected graphs [5, 12, 6, 7]). The n th signal coefficient corresponds to node v_n , and the edge weight $A_{n,m}$ between nodes v_n and v_m is a quantitative expression of the underlying relation between the n th and the m th signal coefficients, such as a similarity, a dependency, or a communication pattern.

Graph Signal. Given the graph representation $G = (\mathcal{V}, A)$, a *graph signal* is defined as the map on the graph nodes that assigns the signal coefficient $x_n \in \mathbb{C}$ to the node v_n . Once the node order is fixed, the graph signal can be written as a vector

$$\mathbf{x} = [x_0, x_1, \dots, x_{N-1}]^T \in \mathbb{C}^N. \quad (1)$$

Graph Fourier Transform. In general, a Fourier transform corresponds to the expansion of a signal using basis functions that are invariant to filtering; this basis is the eigenbasis of the graph shift A (or, if the complete eigenbasis does not exist, the Jordan eigenbasis of A).

For simplicity, assume that A has a complete eigenbasis and the spectral decomposition of A is [1, 13]

$$A = V \Lambda V^{-1}, \quad (2)$$

where the eigenvectors of A form the columns of matrix V , and $\Lambda \in \mathbb{C}^{N \times N}$ is the diagonal matrix of corresponding eigenvalues $\lambda_0, \dots, \lambda_{N-1}$ of A and $\lambda_0 > \lambda_1 > \dots > \lambda_{N-1}$. These eigenvalues represent frequencies on the graph, with λ_0 the lowest and λ_{N-1} the highest frequency, specified by the descending order of the eigenvalues.

Definition 1. The *graph Fourier transform* of $\mathbf{x} \in \mathbb{C}^N$ and the *inverse graph Fourier transform* are, respectively,

$$\hat{\mathbf{x}} = V^{-1} \mathbf{x}, \quad \mathbf{x} = V \hat{\mathbf{x}}. \quad (3)$$

The vector $\hat{\mathbf{x}}$ in (3) represents the signal's expansion in the eigenvector basis and describes the frequency content of the graph signal $\mathbf{x}$. The inverse graph Fourier transform reconstructs the graph signal from its frequency content by combining graph frequency components weighted by the coefficients of the signal's graph Fourier transform.

3. SAMPLING THEORY FOR FINITE-DIMENSIONAL VECTORS

In [1], the authors present sampling and interpolation as a means of moving between higher/lower-dimensional spaces;

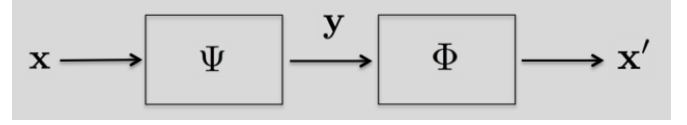


Fig. 1: Sampling followed by interpolation.

they do so for infinite as well as periodic functions, as well as infinite sequences and finite-dimensional vectors. For this last class, however, they do not consider a counterpart of the classical sampling theory; we do this here.

Sampling & Interpolation. Suppose that we want to sample a finite-dimensional vector $\mathbf{x} \in \mathbb{C}^N$ to obtain a shorter finite-dimensional vector $\mathbf{y} \in \mathbb{C}^M$ describing $\mathbf{x}$ ($M < N$); we then interpolate $\mathbf{y}$ to get $\mathbf{x}' \in \mathbb{C}^N$, which recovers $\mathbf{x}$ either exactly or approximately. The sampling operator Ψ is a linear mapping from $\mathbb{C}^N$ to $\mathbb{C}^M$ and the interpolation operator Φ is a linear mapping from $\mathbb{C}^M$ to $\mathbb{C}^N$ (see Figure 1),

$$\begin{aligned} \text{sampling :} \quad & \mathbf{y} = \Psi \mathbf{x} \in \mathbb{C}^M, \\ \text{interpolation :} \quad & \mathbf{x}' = \Phi \mathbf{y} = \Phi \Psi \mathbf{x} \in \mathbb{C}^N. \end{aligned}$$

Perfect recovery happens for all $\mathbf{x}$ when $\Phi \Psi$ is the identity matrix; it is not possible in general because $\text{rank}(\Phi \Psi) \leq M < N$.

Bandlimited Finite-Dimensional Vectors. We consider bandlimited finite-dimensional vectors here, where perfect recovery is possible.

Definition 2. Let $F \in \mathbb{C}^{N \times N}$ be an invertible linear operator. The *F-transform* of $\mathbf{x} \in \mathbb{C}^N$ and the *inverse F-transform* are, respectively,

$$\hat{\mathbf{x}} = F \mathbf{x}, \quad \mathbf{x} = F^{-1} \hat{\mathbf{x}}.$$

Note that in the definition, we did not specify any structure on F ; one can treat it as a generalized version of an N -point discrete Fourier transform.

Definition 3. A finite-dimensional vector $\mathbf{x}$ is called *bandlimited* under the F -transform when there exists $K \in \{0, 1, \dots, N-1\}$ such that its F -transform $\hat{\mathbf{x}}$ satisfies

$$\hat{x}_i = 0 \quad \text{for all } i \geq K.$$

The smallest such K is called the *bandwidth* of $\mathbf{x}$ under the F -transform. A vector that is not bandlimited is called a *full-band vector*.

Definition 4. The set of vectors in $\mathbb{C}^N$ with bandwidth of at most K under the F -transform is a closed subspace denoted $\text{BL}_K(F)$.

The definition describes a bandlimited subspace under the F -transform. Since we do not specify the structure of F , we can permute its rows to choose any band. For example, if we specify F to be the discrete Fourier transform matrix, the bandlimited subspace contains the lowpass finite discrete-time signals.

Sampling Theory. Let $\mathbf{x} \in \text{BL}_K(\mathbf{F})$; sample M coefficients in $\mathbf{x}$ to produce $\mathbf{x}_{\mathcal{M}}$, where $\mathcal{M}$ denotes the sequence of *measured* indices, $\mathcal{M} \subset \{0, 1, \dots, N-1\}$ and $|\mathcal{M}| = M$. The sampling operator is

$$\Psi_{i,j} = \begin{cases} 1, & j = \mathcal{M}_i; \\ 0, & \text{otherwise.} \end{cases} \quad (4)$$

We aim to recover $\mathbf{x}$ from $\mathbf{x}_{\mathcal{M}}$ perfectly.

Theorem 1. Let $\Psi \in \mathbb{C}^{M \times N}$ be the sampling operator (4). The interpolation operator $\Phi \in \mathbb{C}^{N \times M}$ that satisfies: (1) Φ spans the space $\text{BL}_K(\mathbf{F})$; (2) $\Phi\Psi$ is a projection operator, achieves perfect recovery,

$$\mathbf{x} = \Phi\Psi\mathbf{x}, \text{ for all } \mathbf{x} \in \text{BL}_K(\mathbf{F}).$$

Because of space constraints, we only outline the proof (it follows the proof of Theorem 5.2 in [1]). Since $\Phi\Psi$ is a projection operator, $\Phi\Psi\mathbf{x}$ is an approximation of $\mathbf{x}$ in the space of $\text{BL}_K(\mathbf{F})$. We achieve perfect recovery when $\mathbf{x}$ is in the space of $\text{BL}_K(\mathbf{F})$.

Theorem 2. Let $\mathbf{F}_{(K)}^{-1}$ be the first K columns of $\mathbf{F}^{-1}$ and let the sampling operator Ψ satisfy

$$\text{rank}(\Psi \mathbf{F}_{(K)}^{-1}) = K. \quad (5)$$

The interpolation operator $\Phi = \mathbf{F}_{(K)}^{-1} \mathbf{U}$, with $\mathbf{U} \Psi \mathbf{F}_{(K)}^{-1}$ a $K \times K$ identity matrix, achieves perfect recovery,

$$\mathbf{x} = \Phi\Psi\mathbf{x}, \text{ for all } \mathbf{x} \in \text{BL}_K(\mathbf{F}).$$

Proof. We only need to show that Φ satisfies the two properties from Theorem 1.

Since $\text{rank}(\Psi \mathbf{F}_{(K)}^{-1}) = K$ and $\text{rank}(\mathbf{U} \Psi \mathbf{F}_{(K)}^{-1}) = K$, the rank of $\mathbf{U}$ is K , that is $\mathbf{U}$ spans $\mathbb{C}^K$. We then have that the interpolation operator $\Phi = \mathbf{F}_{(K)}^{-1} \mathbf{U}$ spans $\text{BL}_K(\mathbf{F})$, satisfying the first property.

To prove that $\mathbf{P} = \Phi\Psi$ is a projection operator, we must prove it is idempotent,

$$\mathbf{P}^2 = \Phi\Psi\Phi\Psi = \mathbf{F}_{(K)}^{-1} \mathbf{U} \Psi \mathbf{F}_{(K)}^{-1} \mathbf{U} \Psi \stackrel{(a)}{=} \mathbf{F}_{(K)}^{-1} \mathbf{U} \Psi = \mathbf{P},$$

where (a) follows from (5), satisfying the second property. $\square$

From Theorem 2, we see that an arbitrary sampling operator may not leads to perfect recovery even for bandlimited vectors. The sampling operator should select at least one set of K linearly-independent rows in $\mathbf{F}_{(K)}^{-1}$. To find linearly-independent rows in a matrix, fast algorithms exist, such as QR decomposition; see [14, 1].

The sample size M should be no smaller than the bandwidth K . When they are equal, $M = K$, because of (5), $\mathbf{U}$ is the inverse of $\Psi \mathbf{F}_{(K)}^{-1}$; when $M > K$, $\mathbf{U}$ is a pseudo-inverse of $\Psi \mathbf{F}_{(K)}^{-1}$.

In some cases, the length of a vector is much larger than its bandwidth, $N \gg K$; the vector is then K -sparse under $\mathbf{F}$ -transform. If $\text{rank}(\mathbf{F}_{(K)}^{-1}) = K$, Theorem 2 shows that the vector can be recovered using only K measurements by choosing the proper sampling operator.

4. SAMPLING THEORY FOR GRAPH SIGNALS

In Section 3, we showed a general sampling theory for finite-dimensional vectors and that perfect recovery is possible for finite-dimensional vectors bandlimited under some $\mathbf{F}$ -transform. In this section, we specify finite-dimensional vectors to be graph signals, the $\mathbf{F}$ -transform to be the graph Fourier transform, and propose a sampling theory for bandlimited graph signals.

We thus aim to sample a graph signal $\mathbf{x}$ to obtain a measured part $\mathbf{x}_{\mathcal{M}}$; after that, we interpolate $\mathbf{x}_{\mathcal{M}}$ to recover $\mathbf{x}$. We choose the sampling operator be Ψ as in (4); then

$$\begin{aligned} \text{sampling :} \quad & \mathbf{x}_{\mathcal{M}} = \Psi\mathbf{x} \in \mathbb{C}^M, \\ \text{interpolation :} \quad & \mathbf{x} = \Phi\mathbf{x}_{\mathcal{M}} = \Phi\Psi\mathbf{x} \in \mathbb{C}^N; \end{aligned}$$

see Figure 2. From Section 3, we know that the perfect recovery is not possible in general; we thus focus on bandlimited graph signals. The following two definitions are graph counterparts of Definitions 3 and 4 for finite-dimensional vectors. Because we have access to a graph shift, graph Fourier transform is naturally given, which implies that the notion of bandlimitedness is an intuitive one (with respect to the graph Fourier transform as opposed to a general $\mathbf{F}$ -transform).

Definition 5. A graph signal is called *bandlimited* when there exists $K \in \{0, 1, \dots, N-1\}$ such that its graph Fourier transform $\hat{\mathbf{x}}$ satisfies

$$\hat{x}_i = 0 \quad \text{for all } i \geq K.$$

The smallest such K is called the *bandwidth* of $\mathbf{x}$. A graph signal that is not bandlimited is called a *full-band graph signal*.

Definition 6. The set of graph signals in $\mathbb{C}^N$ with bandwidth of at most K is a closed subspace denoted $\text{BL}_K(\mathbf{V}^{-1})$, with $\mathbf{V}^{-1}$ as in (2).

We apply now Theorem 2 to obtain the following result.

Theorem 3. Let $\mathbf{V}_{(K)}$ be the first K columns of $\mathbf{V}$ and let the sampling operator Ψ satisfy

$$\text{rank}(\Psi \mathbf{V}_{(K)}) = K.$$

The interpolation operator $\Phi = \mathbf{V}_{(K)} \mathbf{U}$, with $\mathbf{U} \Psi \mathbf{V}_{(K)}$ a $K \times K$ identity matrix, achieves perfect recovery,

$$\mathbf{x} = \Phi\Psi\mathbf{x}, \text{ for any } \mathbf{x} \in \text{BL}_K(\mathbf{V}^{-1}).$$

To prove the theorem, simply substitute the graph Fourier transform for the $\mathbf{F}$ -transform in Theorem 2.

Since $\mathbf{V}$ is invertible, the column vectors in $\mathbf{V}$ are linearly independent and $\text{rank}(\mathbf{V}_{(K)}) = K$ always holds. In other words, at least one set of K linearly-independent rows in $\mathbf{V}_{(K)}$ always exists. Since the graph shift $\mathbf{A}$ is given, one can find such set without any graph signals. Once such set is ready, Theorem 3 guarantees perfect recovery of bandlimited graph

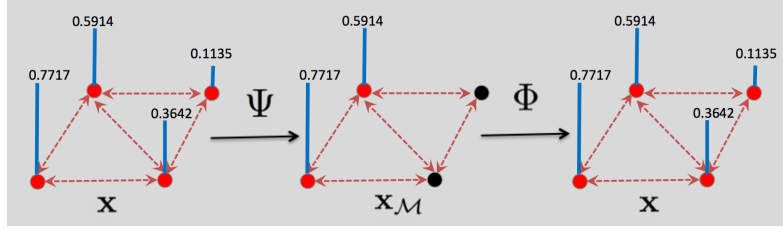


Fig. 2: Sampling followed by interpolation. The arrows indicate different edge weights for two nodes.

signals without any probability constraints or any approximation as in compressed sensing [8].

If we have multiple choices of K linearly-independent rows, which one is best? When $M = K$, U is a basis that spans $\mathbb{C}^K$; we thus check the condition for the Riesz basis. For each feasible Ψ , we compute the inverse of $\Psi V_{(K)}$ to obtain U ; the best choice comes from the tightest stability constants of U [1]. When $M > K$, U is a frame that spans $\mathbb{C}^K$; we thus check the condition for the frame. For each feasible Ψ , we compute the pseudo-inverse of $\Psi V_{(K)}$ to obtain U ; the best choice comes from the tightest frame bounds of U [1]. The reason is that we want U to span the space well, thus, the recovery is more robust and stable.

5. TOY EXAMPLE

We consider a four-node directed graph with graph shift

$$A = \begin{bmatrix} 0 & 0.5 & 0.25 & 0.25 \\ 0.667 & 0 & 0.333 & 0 \\ 0.333 & 0 & 0.333 & 0.333 \\ 0.5 & 0 & 0.5 & 0 \end{bmatrix}.$$

The corresponding inverse graph Fourier transform matrix is

$$V = \begin{bmatrix} 0.5 & 0.1827 & 0.4523 & 0.4997 \\ 0.5 & 0.5434 & -0.1508 & -0.5976 \\ 0.5 & -0.2717 & -0.7538 & 0.2988 \\ 0.5 & -0.7730 & 0.4523 & -0.5513 \end{bmatrix}.$$

We generate a bandlimited graph signal $\mathbf{x} \in \text{BL}_2(V^{-1})$ as

$$\mathbf{x} = [0.5914 \quad 0.7717 \quad 0.3642 \quad 0.1135]^T.$$

We can check the first two columns of V to see that all sets of two rows are independent. According to the sampling theorem, we can recover $\mathbf{x}$ perfectly by sampling any two of its coefficients; for example, sample the first two. Then, $\mathcal{M} = \{1, 2\}$, $\mathbf{x}_{\mathcal{M}} = [0.5914 \quad 0.7717]^T$, and the sampling operator

$$\Psi = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}.$$

We recover $\mathbf{x}$ by using the following interpolation operator (see Figure 2)

$$\Phi = V_{(2)}(\Psi V_{(2)})^{-1} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 2.2601 & -1.2601 \\ 3.6502 & -2.6502 \end{bmatrix}.$$

6. SAMPLING THEORY FOR FINITE DISCRETE-TIME & DISCRETE-SPACE SIGNALS

We now show that by imposing structure on the graph, graph signals reduce to finite discrete-time or discrete-space signals.

We get finite discrete-time signals by specifying the graph shift to be the cyclic permutation matrix, whose eigenvectors then form the N -point discrete Fourier transform matrix [1, 4]. We see that Definitions 5, 6 and Theorem 3 are immediately applicable to finite discrete-time signals. With this definition of the discrete Fourier transform matrix, the highest frequency is in the middle of the spectrum (although this is just a matter of ordering). From Definitions 3 and 5, we can actually permute the rows in the discrete Fourier transform matrix to choose any frequency band.

Similarly, we get discrete-space signals by defining the graph shift appropriately; the corresponding graph Fourier transform is exactly the N -point discrete cosine transform.

7. CONCLUSIONS

In this paper, we proposed a sampling theory for bandlimited finite-dimensional vectors, which follows the same paradigm of classical sampling theory. We then applied this general result to discrete signal processing on graphs and showed the sampling theory for bandlimited graph signals. By imposing a specific structure on the graph, graph signals reduce to finite discrete-time or discrete-space signals, effectively ensuring that the proposed sampling theory works for such signals. The proposed sampling theory is applicable to both directed and undirected graphs, the assumption of perfect recovery is easy both to check and to satisfy, and, under that assumption, perfect recovery is guaranteed without any probability constraints or any approximation.

8. REFERENCES

- [1] M. Vetterli, J. Kovačević, and V. K. Goyal, *Foundations of Signal Processing*, Cambridge University Press, 2014, <http://www.fourierandwavelets.org/>.
- [2] M. Unser, “Sampling – 50 years after Shannon,” *Proc. IEEE*, vol. 88, no. 4, pp. 569–587, Apr. 2000.
- [3] A. Sandryhaila and J. M. F. Moura, “Discrete signal processing on graphs,” *IEEE Trans. Signal Process.*, vol. 61, no. 7, pp. 1644–1656, 2013.
- [4] A. Sandryhaila and J. M. F. Moura, “Discrete signal processing on graphs: Frequency analysis,” *IEEE Trans. Signal Process.*, vol. 62, no. 12, pp. 3042–3054, 2014.
- [5] I. Z. Pesenson, “Sampling in paley-wiener spaces on combinatorial graphs,” *Trans. Amer. Math. Soc.*, vol. 360, no. 10, pp. 5603–5627, 2008.
- [6] A. Anis, A. Gadde, and A. Ortega, “Towards a sampling theorem for signals on arbitrary graphs,” in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, 2014, pp. 3864 – 3868.
- [7] X. Wang, P. Liu, and Y. Gu, “Local-set-based graph signal reconstruction,” *IEEE Trans. Signal Process.*, 2014, Submitted.
- [8] E. J. Candès, “Compressive sampling,” in *Int. Congr. Mathematicians*, Madrid, Spain, 2006.
- [9] S. Chen, F. Cerda, P. Rizzo, J. Bielak, J. H. Garrett, and J. Kovačević, “Semi-supervised multiresolution classification using adaptive graph filtering with application to indirect bridge structural health monitoring,” *IEEE Trans. Signal Process.*, vol. 62, no. 11, June 2014.
- [10] S. Chen, A. Sandryhaila, G. Lederman, Z. Wang, J. M. F. Moura, P. Rizzo, J. Bielak, J. H. Garrett, and J. Kovačević, “Signal inpainting on graphs via total variation minimization,” in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, Florence, May 2014, pp. 8267 – 8271.
- [11] S. Chen, A. Sandryhaila, J. M. F. Moura, and J. Kovačević, “Signal recovery on graphs,” *IEEE Trans. Signal Process.*, 2014, Submitted.
- [12] 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, 2013.
- [13] J. Kovačević, V. K. Goyal, and M. Vetterli, *Fourier and Wavelet Signal Processing*, Cambridge University Press, Cambridge, 2014, <http://www.fourierandwavelets.org/>.
- [14] R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge University Press, Cambridge, 1985.