Graph Fourier Transform Enhancement through Envelope Extensions

Ali Bagheri Bardi, Taher Yazdanpanah, Miloš Daković, Ljubiša Stanković

Abstract—Many real-world networks are characterized by directionality; however, the absence of an appropriate Fourier basis hinders the effective implementation of graph signal processing techniques. Inspired by discrete signal processing, where embedding a line digraph into a cycle digraph facilitates the powerful Discrete Fourier Transform for signal analysis, addressing the structural complexities of general digraphs can help overcome the limitations of the Graph Fourier Transform (GFT) and unlock its potential.

The Discrete Fourier Transform (DFT) serves as a Graph Fourier Transform for both cycle graphs and Cayley digraphs on the finite cyclic groups $\mathbb{Z}_N$. We propose a systematic method to identify a class of such Cayley digraphs that can encompass a given directed graph. By embedding the directed graph into these Cayley digraphs and opting for envelope extensions that naturally support the Graph Fourier Transform, the GFT functionalities of these extensions can be harnessed for signal analysis.

Among the potential envelopes, optimal performance is achieved by selecting one that meets key properties. This envelope's structure closely aligns with the characteristics of the original digraph. The Graph Fourier Transform of this envelope is reliable in terms of numerical stability, and its columns approximately form an eigenbasis for the adjacency matrix associated with the original digraph.

It is shown that the envelope extensions possess a convolution product, with their GFT fulfilling the convolution theorem. Additionally, shift-invariant graph filters (systems) are described as the convolution operator, analogous to the classical case. This allows the utilization of systems for signal analysis.

Index Terms—Graph signal processing, Graph Fourier transform, Graph filters, Cayley digraphs

I. Introduction

Graph Signal Processing (GSP) is an expanding field focused on the analysis and manipulation of data residing on graphs. GSP finds diverse applications in areas such as social network analysis, sensor networks, biological networks, recommendation systems, image and video processing, brain network analysis, transportation networks, and communication systems. [1]–[17]. By leveraging the underlying graph structure, GSP enables more effective data analysis and signal processing, leading to improved performance and insights across these domains.

The principal challenge in Graph Signal Processing revolves around the development of methodologies to effectively analyze and process signals within the context of graph structures.

Ali Bagheri Bardi (bagheri@pgu.ac.ir) is with Persian Gulf University, Bushehr, Iran, Taher Yazdanpanah (yazdanpanah@pgu.ac.ir) is with Persian Gulf University, Bushehr, Iran, Miloš Daković (milos@ucg.ac.me) and Ljubiša Stanković (ljubisa@ucg.ac.me) are with University of Montenegro, Podgorica, Montenegro.

The core objective of this discourse is proposing an approach to explore and address this central challenge on directed graphs, with a primary emphasis on devising strategies that harness and exploit the inherent properties and characteristics of the underlying graph structure. This involves developing algorithms that can accommodate the directionality of edges, which are critical in many applications such as traffic flow analysis, social network dynamics.

Early methods in GSP often relied on spectral graph theory, which deals with the eigen-decomposition techniques from linear algebra [1], [18]–[24]. The key idea is to use the eigenvectors of graph Laplacians or adjacency matrices to define transformations and operations on graph signals. Spectral methods enable tasks such as filtering, clustering, and classification by exploiting the spectral properties of the graph.

Despite their theoretical appeal, spectral methods face significant limitations when applied to directed graphs. One major challenge stems from the non-diagonalizability of adjacency matrices in many directed graphs, which complicates their eigen value decomposition for spectral analysis. Unlike undirected graphs, the adjacency matrices of directed graphs frequently yield non trivial Jordan blocks in their Jordan normal form. In addition of numerical instability, this characteristic complicates the application of shift-invariant filters, a challenge well articulated in [3], [25]. The limitation is even more significant in large-scale graphs or in dynamic graph settings where the structure of the graph changes over time.

To ensure effective progression in our discourse, we establish a convention to guide the discussion. Directed graphs featuring diagonalizable adjacency matrices with distinct non-zero eigenvalues will be termed admissible digraphs.

The main contribution of this work is the establishment of a systematic approach to propose admissible options for a given directed graph, thereby enabling the application of the Graph Fourier Transform (GFT) for processing of graph signals.

This paper is organized as follows. To facilitate the illustration of the main idea and highlight the advantages, the Introduction section is supported and enhanced by three subsections. Part A provides a general definition and deals with the importance of admissibility, and Part B reviews Digital Signal Processing (DSP) concepts that motivated our outline. Finally, Part C includes a brief review of two related works most closely aligned with the current subject.

Section II focuses on the theory that supports our algorithm, offering an in-depth exploration of the underlying principles and mathematical foundations. In Section III, we apply our approach to a real-world dataset, detailing the step-by-step implementation process and highlighting key findings. Lastly,

an Appendix section is included to recap the essential concepts and terminology encountered throughout the paper, facilitating a comprehensive understanding of the subject matter.

LIST OF ACRONYMS AND NOTATIONS

Acronym/Notations	Definition
DFT (GFT)	Discrete (Graph) Fourier Transform
DSP (GSP)	Discrete (Graph) Signal Processing
L_{DepCol} (L_{DepRow})	Dependency column (row) lists
$\mathcal{C}$ ($\mathcal{R}$)	A dependency column (row) list
$\text{Cay}(\mathbb{Z}_{\mathbb{N}}, \Gamma)$	Cayley graph with connetion set Γ
$\mathbf{G}$	Directed Graph
$\mathbf{G}_e$	Addmisible extention of $\mathbf{G}$
$\mathfrak{F}$	DFT
$\mathbf{V}$	Graph Fourier Basis on $\mathbf{G}$
$\circledast$	convolution
$\star$	conventional Convolution

A. Graph Fourier Transform : Framework

Let $\mathbf{G}$ be a directed graph (digraph) with size N. A given signal $\mathbf{s}$ on $\mathbf{G}$ is represented as an N-dimensional vector in $\mathbb{R}^N$ (or $\mathbb{C}^N$). Theoretically, based on any (orthonormal) basis $\mathcal{B} = \{\mathbf{v}_0, \ldots, \mathbf{v}_{N-1}\}$, the processing of the signal $\mathbf{s}$ involves finding the scalars $y[k]$ to represent $\mathbf{s}$ in terms of the vectors in $\mathcal{B}$:

$$\mathbf{s} = y[0]\mathbf{v}_0 + \cdots + y[N-1]\mathbf{v}_{N-1} \qquad (1)$$

Let $\mathbf{V}$ be the matrix whose columns are the vectors in $\mathcal{B}$. Under these conditions, the representation of the signal $\mathbf{s}$ in terms of the basis vectors, as described in equation (1), can be equivalently expressed using matrix notation:

$$\begin{cases} \mathbf{y} = \mathbf{V}^{-1}\mathbf{s}, \\ \mathbf{y} = (y[0], \ldots, y[N-1])^T. \end{cases} \qquad (2)$$

When working with real datasets organized on directed graphs, various challenges can arise. To address these challenges, it is crucial to develop specific mathematical criteria that guide the analysis and processing of the data. If these criteria can be expressed in terms of an (orthonormal) basis $\mathcal{B}$, graph signal processing techniques can be effectively applied, provided the following key conditions are met:

1) Graph structure compatibility. The basis $\mathcal{B}$ serves as a set of principal graph signals, referred to as harmonics, which act as the primary components for representing signals on the graph. When processing a graph signal $\mathbf{s}$, one crucial step involves decomposing it into these principal components to analyze how each basis vector influences the signal's properties. This decomposition process underscores the importance of establishing a connection between the basis $\mathcal{B}$ and the underlying directed graph structure's characteristics.

2) Efficiency. Coefficients $y[k]$s denote the coordinates of a graph signal when expressed in terms of the principal components constituting the graph Fourier basis. Manipulating the initial graph signal generally involves two primary steps: altering the Fourier coefficients and reconstructing the modified graph signal.

The adjacency matrix functor serves as a mapping between the category of directed graphs and the square matrices, encapsulating the connections between the graph's vertices.

Given the adjacency matrix $\mathbf{A}$ associated with the digraph $\mathbf{G}$, the Jordan normal decomposition allows us to factorize $\mathbf{A}$ into the product of a Jordan matrix $\mathbf{J}$ and a similarity transformation matrix $\mathbf{V}$. This can be represented as:

$$\mathbf{A} = \mathbf{V}\mathbf{J}\mathbf{V}^{-1} \qquad (3)$$

Building upon this formula, we define the graph Fourier transform $\mathbf{F}$ associated with $\mathbf{G}$ as the inverse of the similarity transformation matrix $\mathbf{F} = \mathbf{V}^{-1}$. The columns of matrix $\mathbf{V}$ serve as the basis generalized eigenvectors of the adjacency matrix $\mathbf{A}$. These generalized eigenvectors constitute the graph Fourier basis $\mathcal{B}_{\mathbf{F}}$, providing a set of principal graph signals on the directed graph $\mathbf{G}$.

While the graph Fourier basis $\mathcal{B}_{\mathbf{F}}$, offers valuable insights by encapsulating the principal components that capture the essential characteristics of the digraph $\mathbf{G}$, there are two significant challenges that limit its widespread adoption in the realm of general graph signal processing:

(i) The first issue associated with employing the Jordan normal decomposition of $\mathbf{A}$ is its computational complexity. The decomposition process can be resource-intensive, particularly when dealing with large-scale graphs.

(ii) When the Jordan matrix is not a diagonal matrix, the utilization of *shift-invariant graph filters*, also known as *systems*, becomes impractical. These types of filters are known as essential tools for modifying the signals. To further elaborate on this item, we reiterate that

- A graph filter $\mathbf{H}$ is shift-invariant if $\mathbf{H}\mathbf{A} = \mathbf{A}\mathbf{H}$.
- For a given polynomial h, the linear transformation $h(\mathbf{A})$ is termed a system.

When the minimal and characteristic polynomials of $\mathbf{A}$ are identical, Theorem 1 of [3] provides a characterization for shift-invariant graph filters as follows:

Shift-invariant filters = $\{h(\mathbf{A}) : h$ is a polynomial$\}$

In the next section, the motivation for addressing both of the mentioned challenges is provided in detail.

B. Motivation : Graph Fourier Basis in classical DSP

In discrete signal processing, signals are initially analyzed on a line graph where the adjacency matrix corresponds to the backward shift matrix.

$$\mathbf{S} = \begin{bmatrix} 0 & \mathbf{I}_{N-1} \\ 0 & 0 \end{bmatrix} \qquad (4)$$

Here, $\mathbf{I}_{N-1}$ represents the identity matrix of size $N-1$ and $\mathbf{S}$ denotes the backward shift matrix of size N.

Backward shift matrices are nil-potent, meaning all its eigenvalues are zero. While this simplifies the Jordan decomposition, the resulting graph Fourier basis does not offer substantial practical utility in signal processing tasks.

When addressing one-dimensional digital signal processing involving periodic signals, line digraphs are replaced with cycle digraphs. This substitution makes the backward shift
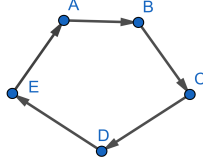
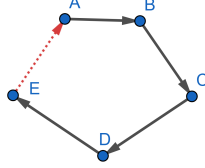Fig. 1. The directed cycle graph with 5 vertices.

Fig. 2. By adding the red dotted line, the directed line graph with 5 vertices is included in the directed cycle graph.

matrix circulant, which is a specific type of matrix with particular properties well-suited for handling periodic signals.

$$C = \begin{bmatrix} 0 & I_{N-1} \\ 1 & 0 \end{bmatrix} \qquad (5)$$

Here, C represents the circulant matrix of size N.

In this circular domain, the Fourier basis becomes the most organized and widely used basis for signal processing tasks. Central to this framework is the normalized discrete Fourier transform, denoted by $\mathfrak{F}$, defined as follows:

$$\mathfrak{F} = \frac{1}{\sqrt{N}} \left[\exp\left(-\frac{2\pi i k l}{N}\right) \right]_{k,l=0}^{N-1}. \qquad (6)$$

The discrete Fourier transform converts the convolution product into a pointwise operation, thereby streamlining the signal processing procedure by enabling convenient modification of the coefficients. Notably, the inverse DFT can be readily derived through its hermitian, which inherently underscores its numerical stability.

Any periodic signal can be conceptualized as a graph signal on the directed cycle graph, as depicted in Figure 1. The eigenvalue decomposition of the circulant matrix, which is the adjacency matrix, is derived using the discrete Fourier transform, thereby demonstrating the intrinsic connection between the DFT and the underlying graph structure upon which periodic signals are modeled.

$$\begin{cases} C = V D V^{-1} \\ V = \mathfrak{F}^{-1} \\ D = \operatorname{diag}\left(1, \exp\frac{2\pi i}{N}, \ldots, \exp\frac{2\pi i(N-1)}{N}\right) \end{cases} \qquad (7)$$

In discrete signal processing, the foundational approach involves embedding the directed line graph into a directed cycle graph (see Figure 2). This transformation is achieved by connecting the sink of the line graph to its source, effectively altering a zero entry in the forward shift matrix and resulting in a circulant matrix structure. This modification turns the initial linear signal representation into a periodic signal on a circular domain. By embedding the line graph into a cycle graph, DSP practitioners leverage the power of the discrete

Fourier transform for efficient signal processing.

The primary contribution of this work is the development of an approach applicable to general digraphs.

Definition 1. *Let G be a digraph and let A be the adjacency matrix associated with G with the Jordan normal decomposition $A = VJV^{-1}$.*
1) *The digraph G is non-singular if J is a diagonal matrix with all non-zero diagonal entries.*
2) *The non-singular digraph G is admissible if the diagonal entries of J are distinct.*

Section II-C establishes that admissible digraphs support a convolution product, denoted as $\circledast$. This convolution product enables the characterization of any shift-invariant filter as a convolution operator: for a given polynomial h, there exists a graph signal s such that for every signal x, the following relationship holds:

$$h(A)x = s \circledast x \qquad (8)$$

Furthermore, the graph Fourier transform $F = V^{-1}$ satisfies the convolution theorem, which is stated as follows:

$$F(x \circledast y) = F x \cdot F y \qquad (9)$$

In this equation, x and y represent graph signals, while $\cdot$ denotes the pointwise product between the transformed vectors Fx and Fy.

Therefore admissible digraphs satisfy the minimal conditions necessary to employ GFT techniques for graph signal processing within the broader category of digraphs. This observation raises a question:

For a given directed graph G, can we devise a systematic method to construct an (optimal) admissible directed graph G_e that includes G as a directed subgraph?

Identifying such an approach could have significant implications for the efficient processing of graph signals on G. This would allow us to exploit the intrinsic characteristics of admissible graph and capitalize on the potential of the GFT, which emerge from the eigenvalue decomposition of the adjacency matrix associated with G_e given in Definition 1.

The primary contribution of this research lies in addressing the aforementioned question by effectively embedding any directed graph into a Cayley graph based on a cyclic finite group $\mathbb{Z}_N$, with the inclusion of a minimal number of new edges. These types of Cayley graphs are recognized as enriched admissible directed graphs, as explained in Theorem 2. It outlines a method for obtaining the optimal graph Fourier transform with respect to numerical stability.

C. Related works

Within the existing body of literature, two fundamentally distinct methodologies have been pursued for the theoretical development of the Graph Fourier Transform in the domain of signal processing on directed graphs.

(I) One technique for formulating the GFT on a directed graph involves the application of strategic topological modifications to the graph structure. These modifications, which may include the addition of carefully selected edges or vertices, are designed to transform the graph's adjacency or Laplacian matrix into a diagonalizable form. The resulting modified

matrix then serves as the foundation for defining the GFT through its eigendecomposition.

The concept of incorporating new edges was first explored in [1], where a method was proposed to eliminate non-trivial Jordan blocks present in the adjacency matrix during each iteration. This process relies on the computation of the Jordan normal form to dismantle the blocks. The method proceeds iteratively to ensure the complete removal of all non-trivial Jordan blocks. However, a major challenge arises due to the potential formation of new Jordan blocks when one is eliminated, leading to a time-consuming and prolonged process that can become difficult to control.

In the study [18], an intriguing attempt is proposed to define a class of graph Fourier transforms where the columns can be regarded as an approximate Fourier basis on graph **G**. The aim is to identify some graph Fourier transforms **F** by solving the following optimization problem:

$$\begin{cases} \min_{\mathbf{F},\mathbf{T}} \| \mathbf{A}\,\mathbf{F} - \mathbf{F}\,\Lambda \| \\ \text{subject to} \quad \mathbf{A}\,\mathbf{F} = \mathbf{F}\,\mathbf{T} \\ \mathbf{T} \text{ is an uppertirangular matrix with zero diagonal entries} \\ \sigma_{\min}\,\mathbf{F} \geq \alpha \text{ and } 0 < \alpha \leq 1 \\ \Lambda \text{ represents the diagonal matrix of eigenvalues of } \mathbf{A}. \end{cases}$$
$$(10)$$

The proposed method has demonstrated its ability to ensure spectral compatibility with the corresponding adjacency matrix. However, its effectiveness hinges on the preservation of structural properties. To elaborate this, let **F** denote the matrix obtained from the optimization problem (10), and let $\mathfrak{G}_F$ represent the set of all directed graphs for which **F** functions as a GFT. The greater the disparity between the graph structures in $\mathfrak{G}_F$ and the original digraph **G**, the less reliable the results become. There is currently no evidence to suggest that this method can provide a satisfactory solution that ensures both spectral and structural compatibility.

(II) The second approach to introducing a GFT on directed graphs involves defining a total variation measure and identifying an orthonormal basis that minimizes this variation. Recent studies, such as those referenced in [20], [24], [26]–[28], explore this approach across various adaptations. A broader generalization of this method is presented in [26], offering a more unified framework that integrates related concepts.

The derivation of an orthonormal basis ensures that the corresponding GFT possesses desirable mathematical properties, such as energy conservation and straightforward signal reconstruction. However, two ambiguities arise regarding this method. Firstly, the non-uniqueness of variations leads to different GFTs. Secondly, a meaningful connection between the graph structure and the shift operator associated with the GFT derived from the orthonormal basis $\mathcal{B}$ remains elusive. This latter issue poses significant challenges in defining shift invariant filters (systems), which should be induced by an appropriate convolution on the directed graph. In Part B of Section III in [26], the authors discuss how this shortcoming leads to the non-uniqueness of shift-invariant filters.

II. THEORY

This section thoroughly presents the analytical evidence required to justify the methodologies employed in this study.
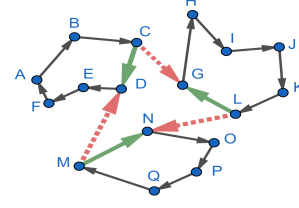

Fig. 3. Hamiltonian Cycle Creation

A. Admissible Extensions

The primary focus of this section is to highlight that for a directed graph **G** with N vertices, embedding it within a specific type of Cayley digraph $\mathrm{Cay}(\mathbb{Z}_N, \Gamma)$ can be achieved by introducing a minimal number of additional edges. This embedding ensures the existence of suitable admissible extensions of **G**, thereby enabling efficient graph signal processing.

The vertices of the Cayley digraph $\mathrm{Cay}(\mathbb{Z}_N, \Gamma)$ are given by the set $\mathbb{Z}_N = \{0, \ldots, N-1\}$, where the connection set Γ determines the edges. All necessary details are provided in the appendix. We need to revisit a key point concerning the Cayley graph structure $\mathrm{Cay}(\mathbb{Z}_N, \Gamma)$, which will serve as a foundational framework for our subsequent analysis. The proof of this key point is also provided in the appendix.

Theorem 2. *Suppose that $\mathbf{A}_\Gamma$ is the adjacancy matrix of the Cayley digraph $\mathrm{Cay}(\mathbb{Z}_N, \Gamma)$.*

$$\begin{cases} \mathbf{A}_\Gamma = \mathbf{V}\,\mathbf{D}_\Gamma\,\mathbf{V}^{-1} \qquad (\mathbf{V} = \mathfrak{F}^{-1}) \\ \mathbf{D}_\Gamma = \sum_{q=1}^k \mathbf{D}^{n_q} \\ \mathbf{D} = diag\Big(1, \exp\big(\tfrac{2\pi i}{N}\big), \ldots, \exp\big(\tfrac{2\pi i(N-1)}{N}\big)\Big) \end{cases}$$
$$(11)$$

Theorem 2 demonstrates that the discrete Fourier transform serves as a GFT for all types of these Cayley digraphs.

The following introduces the principal theorem that underpins the algorithm used to introduce the GFT for any digraph. To ensure a comprehensive understanding of this theorem, it is beneficial to adopt the following convention:

Dependency lists and pseudo-permutation. Given an $N \times N$ matrix **A** with a nullity index g, representing the dimension of its null space, we introduce the concepts of column and row dependency lists, denoted as L_{DepCol} and L_{DepRow}, respectively. A row dependency list $\mathcal{R}$ in L_{DepRow} consists of g rows of matrix **A**, chosen such that removing the rows indicated in $\mathcal{R}$ does not affect the rank of **A**. Similarly, a column dependency list $\mathcal{C}$ in L_{DepCol} comprises g columns of **A** with the same property. For a given pair $(\mathcal{R}, \mathcal{C})$, we define an $N \times N$ matrix **Q** with rank g and unit non-zero entries located exclusively at the intersections of $\mathcal{R}$ and $\mathcal{C}$ as a pseudo-permutation. Notably, there are $g!$ distinct pseudo-permutations corresponding to each pair $(\mathcal{R}, \mathcal{C})$.

Theorem 3. *Let **G** be a digraph with N vertices. There is a Cayley graph $Cay(\mathbb{Z}_N, \Gamma)$ that includes **G** as a directed subgraph.*

Proof. Let **A** denote the adjacency matrix associated with **G**. The proof is conducted in three steps.

Step1. Non-singularization: If $\mathbf{A}$ is singular, this step must be performed; otherwise, it can be skipped. Suppose that the nulity index of $\mathbf{A}$ is g. Let $\mathcal{C}$ (resp. $\mathcal{R}$) be g columns (resp. g rows) representing a column dependency list (row dependency list) of $\mathbf{A}$ and $\mathbf{Q}$ be a corresponding pseudo-permutation to the pair $(\mathcal{R},\mathcal{C})$. We define $\mathbf{A}_{\text{Inv}} = \mathbf{A} + \mathbf{Q}$.

By excluding the columns $\mathcal{C}$ and rows $\mathcal{R}$ from A, we obtain the $(N-g) \times (N-g)$ submatrix of $\mathbf{A}$, denoted as $\mathbf{A}_0$ that is definitely non-singular. Using the Laplace expansion to compute the determinant of $\mathbf{A}_{\text{Inv}}$ and starting from either the columns $\mathcal{C}$ or the rows $\mathcal{R}$, we directly observe that

$$|\det(\mathbf{A}_{\text{Inv}})| = |\det(\mathbf{A}_0)|$$

Let $\mathbf{G}_{\text{Inv}}$ be the digraph whose adjacency matrix is $\mathbf{A}_{\text{Inv}}$. It is indeed created by adding g new edges to $\mathbf{G}$, as specified by the pseudo-permutation $\mathbf{Q}$.

Step2. Hamiltonian Cycle Creation: We represent,

$$\mathbf{A}_{\text{Inv}} = \left[a_{ij}\right]_{i,j=0}^{N-1} \tag{12}$$

Employing the definition of the determinant of $\mathbf{A}_{\text{Inv}}$ as a sum over permutations,

$$|\det(\mathbf{A}_{\text{Inv}})| = \sum_{\sigma} \prod_{0 \le i \le N-1} a_{i,\sigma(i)} \tag{13}$$

where the sum is taken over all permutations σ on the set $\{0, \dots, N-1\}$. Because $\mathbf{A}_{\text{Inv}}$ is non-singular, there exists a permutation σ such that $\prod_{0 \le i \le N-1} a_{i,\sigma(i)}$ is non-zero. This σ corresponds to a collection of k disjoint cycles $\mathcal{O}_1, \dots, \mathcal{O}_k$ that cover all vertices of $\mathbf{G}_{\text{Inv}}$. As visualized in Figure 3 for $k = 3$, by adding at most k new edges, it is possible to connect the cycles $\mathcal{O}_1, \dots, \mathcal{O}_k$ to form a cycle $\mathcal{O}$ that covers all vertices of $\mathbf{G}_{\text{Inv}}$. First, select an arbitrary edge on each of these cycles (green ones). Then, using the new edges (red ones), connect the cycles $\mathcal{O}_m$ to form the cycle $\mathcal{O}$.

Step3. Constructing the Cayley Graph: The cycle digraph $\mathcal{O}$ is isomorphic to digraph $\mathbb{Z}_N$. Let E be the set of all edges present in $\mathbf{G}_{\text{Inv}}$, except those on the cycle $\mathcal{O}$. Note that E may be considered as subset of $\mathbb{Z}_N \times \mathbb{Z}_N$. We define:

$$\begin{cases} \Gamma_0 = \{|a - b| : (a,b) \in E\} \\ \Gamma = \Gamma_0 \cup \{1\} \end{cases}$$

It is evident that the initial digraph $\mathbf{G}$ is embedded into the Cayley digraph $\text{Cay}(\mathbb{Z}_N, \Gamma)$. $\qquad\square$

Remark 4. *Upon examining the proof of Theorem 3, it becomes evident that the assignment of specific non-zero weights to the additional edges does not affect the attainment of non-singularity. Instead, the existence of these edges plays a pivotal role in ensuring the adjacency matrix's non-singularity. This implies that, as long as the newly introduced edges are included in the adjacency matrix with any non-zero weight, the resultant matrix will inherently possess the property of non-singularity. During this study, by an admissible extension we mean the newly added edges are considered with no weight unless specifically pointed out.*

In light of Theorem 3, we find it suitable to present the following definition as a foundation for our continued discussion:

Definition 5. *Let $\mathbf{G}$ be a digraph. Any admissible (non-singular) digraph that contains $\mathbf{G}_e$ as a subgraph is referred to as an admissible extension (non-singular extension) of $\mathbf{G}$.*

Several considerations deserve attention with regard to Theorem 3.

(i) The challenge of embedding a digraph within Cayley graphs is of great interest in graph theory, a field concentrating on the analysis and manipulation of graph structures [29], [30]. Theorem 3 serves as a solution to this important problem, offering a straightforward tool for researchers in this area.

(ii) When the nullity index g of the adjacency matrix associated with the directed graph is greater than one, several non-isomorphic non-singular extensions will emerge. Let $\beta_{\text{Dep\_Col}}$ and $\beta_{\text{Dep\_Row}}$ denote the number of different dependency column and row lists, respectively, and define

$$\text{Total}_{\text{Inv}}(\mathbf{G}) = \beta_{\text{Dep\_Col}} \times \beta_{\text{Dep\_Row}} \times g! \tag{14}$$

Theorem 3 provides numerous $\text{Total}_{\text{Inv}}$ non-singular extensions of the digraph $\mathbf{G}$. This leads to the potential existence of many non-isomorphic admissible extensions.

(ii) Drawing from experience, it is widely observed that on any directed graph, the non-singular extension is typically just an admissible extension as well. However, in cases where an admissible extension does not immediately materialize after forming the non-singular extension, the corresponding Cayley graph extension identifies the positions where additional edges could be inserted to achieve the desired outcome. In these scenarios, the admissible extension is often achieved after just one or two more edges are added. These observations serve as a guide for achieving the admissible extension and provide theoretical support for the existence of admissible extensions for any arbitrary directed graph.

(iii) The methodology, as delineated in Theorem 3, comprises three distinct adjustments made to both the directed graph and its corresponding adjacency matrix through the introduction of new edges. Initially, the focus is on ensuring non-singularity, followed by the establishment of a Hamiltonian cycle in the subsequent step. Finally, the process progresses towards achieving diagonalizability.

The inclusion of new edges in the first phase precisely corresponds to the nullity of the adjacency matrix. After completing this step, all Jordan blocks associated with the eigenvalue 0 are eliminated.

B. Spectral components: Admissible Extensions

Let $\mathbf{G}_e$ be an admissible extension of $\mathbf{G}$ with the associated adjacency matrices $\mathbf{A}_e$ and $\mathbf{A}$. Let $\mathbf{Q}$ be a corresponded pseud-permutation with

$$\mathbf{A}_e = \mathbf{A} + \mathbf{Q} \tag{15}$$

Let $\mathbf{F} = \mathbf{V}^{-1}$ be a GFT on $\mathbf{G}_e$. The columns of $\mathbf{V}$ serve a basis compsed of right eigenvectors of $\mathbf{A}_e$. We recall that each eigenvalue of $\mathbf{A}_e$ possesses an algebraic multiplicity of one, indicating that the eigenvectors are unique up to scalar multiplication.

$$\begin{cases} \mathbf{A}_e \mathbf{V} = \mathbf{V} \mathbf{D} \\ \mathbf{D} = \text{Diag}(\lambda_0, \dots, \lambda_{N-1}), \\ \lambda_k \text{s are eigenvalues of } \mathbf{A}_e \end{cases} \tag{16}$$

Algorithm 1 Admissible Extension Algorithm

Input: Adjacency matrix $\mathbf{A}$ of directed graph $\mathbf{G}$
  $N \leftarrow \text{size}(\mathbf{A})$ ▷ Number of nodes in graph $\mathbf{G}$
  $r \leftarrow \text{rank}(\mathbf{A})$ ▷ Rank of matrix $\mathbf{A}$
  $g \leftarrow N - r$ ▷ Nullity of matrix $\mathbf{A}$
  Find all lists $\mathcal{R}$ of r linearly independent rows of $\mathbf{A}$
  Find all lists $\mathcal{C}$ of r linearly independent columns of $\mathbf{A}$
  For each pair $(\mathcal{R}, \mathcal{C})$, construct the $N \times N$ pseudo-permutation matrix $\mathbf{Q}$ using g rows not in $\mathcal{R}$ and g columns not in $\mathcal{C}$
  $\mathbf{A}_{\text{inv}} \leftarrow \mathbf{A} + \mathbf{Q}$ ▷ Modified adjacency matrix
  Verify all possible choices of $(\mathcal{R}, \mathcal{C})$ to ensure the admissibility of $\mathbf{G}_{\text{Inv}}$, the digraph associated with $\mathbf{A}_{\text{inv}}$
Output: A class of admissible digraphs that include $\mathbf{G}$ as a subgraph.

By leveraging (15), we can derive the equation

$$-(\mathbf{A}\,\mathbf{V} - \mathbf{V}\,\mathbf{D}) = \mathbf{Q}\,\mathbf{V} \tag{17}$$

This expression indicates that $\mathbf{F}$, the graph Fourier transform of $\mathbf{G}_e$, functions as an approximate graph Fourier transform for $\mathbf{G}$, with an error term represented by $\mathbf{Q}\,\mathbf{V}$. The non-zero rows of $\mathbf{Q}\,\mathbf{V}$ capture the influence of the edges required to achieve the admissibility of $\mathbf{G}_e$. This matrix helps determine how effectively the eigenvectors of $\mathbf{A}_e$ can serve as approximations for the eigenvectors of $\mathbf{A}$.

Let's consider the k-th column vector of $\mathbf{V}$, denoted as $\mathbf{w}_k$, which satisfies the equation

$$\mathbf{A}\,\mathbf{w}_k - \lambda_k\,\mathbf{w}_k = \mathbf{Q}\,\mathbf{w}_k \tag{18}$$

The vector $\mathbf{Q}\,\mathbf{w}_k$, whose nonzero components of $\mathbf{w}_k$ are just appeared in the row indices addressed by $\mathbf{Q}$, helps us evaluate how well $\mathbf{w}_k$ can serve as an approximate eigenvector for $\mathbf{A}$. To continue with the discussion, let us consider the following (δ, Δ) indices:

$$\begin{cases} \Delta_{\mathbf{V}} = \|\mathbf{Q}\,\mathbf{V}\| \\ \delta_{\mathbf{V}} = \max\{\|\mathbf{Q}\,\mathbf{w}_k\| : k = 0, \ldots, N-1\} \end{cases} \tag{19}$$

(i) Trivially, for any arbitrary small $\epsilon > 0$ and eigenvector $\mathbf{w}_k$, the scaled vector $\epsilon\,\mathbf{w}_k$ is also an eigenvector of $\mathbf{A}_e$ associated with the eigenvalue λ_k. We may write,

$$\begin{cases} \mathbf{A}_e\,\mathbf{V}_\epsilon = \mathbf{V}_\epsilon\,\mathbf{D} \\ \mathbf{V}_\epsilon = \epsilon\,\mathbf{V} \end{cases} \tag{20}$$

Indeed $\mathbf{V}_\epsilon$ corresponds some graph Fourier transform of the admissible digraph $\mathbf{G}_e$. The second condition in (20) clarifies that each column of $\mathbf{V}_\epsilon$ is a scaled version of the original one $\mathbf{V}$ by the factor ϵ. Theoretically, introducing $\mathbf{V}_\epsilon$ allows for assuming both indices δ_{V_ϵ} and $\Delta_{\mathbf{V}_\epsilon}$ to be sufficiently small. However, in practice, the choice of ϵ poses a trade-off with the numerical stability of $\mathbf{V}_\epsilon$. A smaller ϵ might satisfy the assumption of small indices but could also result in a reduced numerical stability when working with $\mathbf{V}_\epsilon$. This trade-off necessitates careful consideration in real-world applications, as both spectral and numerical stability are essential factors for accurate and reliable results. Striking an optimal balance between the two aspects is crucial to ensure the best

performance of the graph-based system under investigation.

(ii) In the context of a weighted admissible extension for $\mathbf{G}$, there is a relationship between the weights assigned to the newly added edges and the norm of the weighted pseudo-permutation matrix. Specifically, as the edge weights are chosen to be smaller, the norm of the weighted pseudo-permutation matrix also decreases. This direct impact is evident in the delta indices (19), which become smaller as well. Choosing excessively small weights might compromise the numerical stability of the graph Fourier transform, potentially leading to unreliable results.

If the primary objective is to obtain a numerically stable graph Fourier transform for a given graph $\mathbf{G}$ while minimizing the indices introduced in (19), it is advisable to explore the entire domain of weighted and unweighted admissible extensions. By examining various admissible extensions, one can identify the optimal graph that strikes the best balance between accurate eigenvector estimates and numerical stability, highlighting the benefits of this current approach.

Example 6. *Let $\mathbf{G}$ be the directed line graph of size N whose adjacency matrix is $\mathbf{S}$, the backward shift matrix. Connecting the sink to the source in this case, a unique rank-one pseudo-permutation $\mathbf{Q}$ exists, making $\mathbf{S}$ the circulant matrix $\mathbf{C}$.*

In the unweighted scenario, the graph Fourier transform is just $\mathfrak{F}$. The following values are simply formulated:

$$\begin{cases} \Delta_{\mathbf{V}} = 1 & \mathbf{V} = \mathfrak{F}^{-1} \\ \delta_{\mathbf{V}} = \frac{1}{\sqrt{N}} \\ k(\mathbf{V}) = 1 & \text{condition number} \end{cases} \tag{21}$$

Now suppose that the connection of the sink to the source is considered with the weight w and denoted as $\mathbf{C}_w$ with the adjacency matrix of the weighted admissible digraph.

$$\mathbf{C}_w = \begin{bmatrix} 0 & \mathbf{I}_{N-1} \\ w & 0 \end{bmatrix} \tag{22}$$

The characteristic polynomial of $\mathbf{C}_w$ is just $p(\lambda) = \lambda^N - w$. Thus, the eigenvalues of $\mathbf{C}_w$ are just the N-th roots of w and they are given as follows when w is a positive real number:

$$\lambda_k = w^{1/N} \exp\left(\frac{2\pi i k}{N}\right) \quad \text{for} \quad k = 0, 1, \ldots, N-1 \tag{23}$$

We define,

$$\mathbf{V}_w = \frac{1}{\sqrt{N}} \begin{bmatrix} 1 & 1 & \cdots & 1 \\ \lambda_0 & \lambda_1 & \cdots & \lambda_{N-1} \\ \vdots & \vdots & \cdots & \vdots \\ \lambda_0^{N-1} & \lambda_1^{N-1} & \cdots & \lambda_{N-1}^{N-1} \end{bmatrix} \tag{24}$$

One may directly check that the kth column of $\mathbf{V}_w$ is a right eigenvector of $\mathbf{C}_w$ corresponding to the eigenvalue λ_k. This yields:

$$\begin{cases} \mathbf{C}_w = \mathbf{V}_w\,\mathbf{D}_w\,\mathbf{V}_w^{-1} \\ \mathbf{D}_w = diag\,(\lambda_0, \ldots, \lambda_{N-1}) \\ \mathbf{V}_w = \mathbf{E}_w\,\mathfrak{F}^{-1} \\ \mathbf{E}_w = diag\,\left(1, w^{\frac{1}{N}}, \ldots, w^{\frac{N-1}{N}}\right). \end{cases} \tag{25}$$
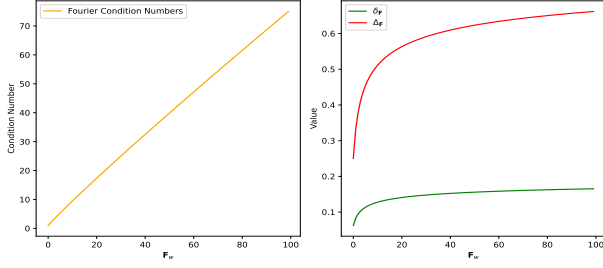
Fig. 4. The figure shows the (δ, Δ) indices and condition numbers of $\mathbf{F}_w$ for weights ranging from 1 to 0.01 with a step size of 0.01 when $N = 16$. The minimum and maximum condition numbers are reported as 1 and 74.98942.

Notably $\mathbf{V}_w$ is normalized when $w \leq 1$. Furthermore,

$$\begin{cases} \Delta_{\mathbf{V}_w} = w \\ \delta_{\mathbf{V}_w} = \dfrac{w}{\sqrt{N}} \\ k(\mathbf{V}_w) = \dfrac{1}{w^{\frac{N-1}{N}}} \end{cases} \tag{26}$$

In Figure 4, the condition numbers and (δ, Δ) indices of $\mathbf{F}_w$ are illustrated for weights ranging from 1 to 0.01 with a step size of 0.01, when $N = 16$.

C. Convolution Product: Admissible Digraphs

As mentioned in the introduction, to apply graph signal processing techniques to a digraph $\mathbf{G}$, it is essential for the graph to support a convolution product capable of accommodating all possible systems.

In the domain of digital signal processing, the cyclic group $\mathbb{Z}_N$ provides such a convolution product. For two discrete-time signals $x[n]$ and $y[n]$ with a period N, their (normalized) cyclic convolution is defined as follows:

$$(x * y)[n] = \frac{1}{\sqrt{N}} \sum_{m=0}^{N-1} x[m] y[(n-m)] \tag{27}$$

Two key points make the conventional convolution product $*$ efficient. First, the discrete Fourier transform $\mathfrak{F}$ converts it to the point-wise product:

$$\mathfrak{F}(\mathbf{x} * \mathbf{y}) = \mathfrak{F}\mathbf{x} \cdot \mathfrak{F}\mathbf{y} \tag{28}$$

Any system $\mathbf{H}$ is formulated by $\mathbf{H}\mathbf{x} = \mathbf{x} * \mathbf{s}$ for some unique signal $\mathbf{s}$.

The class of digraphs supporting such a convolution would be reliable for utilizing graph filters for signal analysis. The subsequent definition offers a mathematical expression for this desired convolution product:

Definition 7. *Let $\mathbf{G}$ be a digraph with the associated adjacency matrix $\mathbf{A}$ and GFT $\mathbf{F}$ on $\mathbf{G}$. We say that a binary operation $\circledast$ on the set of signals on $\mathbf{G}$ defines a convolution product if the following conditions are satisfied:*

1) *The convolution theorem holds: for given signals $\mathbf{x}$ and $\mathbf{y}$ on $\mathbf{G}$,*

$$\mathbf{F}(\mathbf{x} \circledast \mathbf{y}) = \mathbf{F}\mathbf{x} \cdot \mathbf{F}\mathbf{y} \tag{29}$$

where $\cdot$ denotes the pointwise product between the transformed signals $\mathbf{F}\mathbf{x}$ and $\mathbf{F}\mathbf{y}$.

2) *Any system is a convolution operator: for a given polynomial h, there exists a signal $\mathbf{s}$ on $\mathbf{G}$ such that for an arbitrary signal $\mathbf{x}$ on $\mathbf{G}$,*

$$h(\mathbf{A})\mathbf{x} = \mathbf{s} \circledast \mathbf{x} \tag{30}$$

Theorem 8. *Every admissible digraph $\mathbf{G}$ possesses a convolution product.*

Proof. Let $\mathbf{A}$ be the associative adjacency matrix of $\mathbf{G}$ with the eigenvalue decomposition $\mathbf{A} = \mathbf{F}^{-1} \mathbf{D} \mathbf{F}$.

Let $\mathbf{1}$ be the constant signal whose all components are 1. For any given polynomial h, we define

$$\mathbf{s}_h = \mathbf{F}^{-1} h(\mathbf{D}) \mathbf{1} \tag{31}$$

The admissibility property allows, by leveraging the Lagrange interpolation polynomials, any graph signal $\mathbf{x}$ on $\mathbf{G}$ to be uniquely represented by $\mathbf{x} = \mathbf{s}_{h_\mathbf{x}}$ for some polynomial $h_\mathbf{x}$. Utilizing this key point, we define:

$$\begin{cases} \circledast : \mathbb{C}^N \times \mathbb{C}^N \to \mathbb{C}^N \\ \mathbf{x} \circledast \mathbf{y} = \mathbf{s}_{h_\mathbf{x} h_\mathbf{y}} \end{cases} \tag{32}$$

Here $h_\mathbf{x} h_\mathbf{y}$ is just the product of polynomials $h_\mathbf{x}$ and $h_\mathbf{y}$. It is evident that $\mathbf{x} \circledast \mathbf{y} = \mathbf{y} \circledast \mathbf{x}$. Using (31), it can be verified that for a set of signals $\mathbf{x}_1, \ldots, \mathbf{x}_k$ and $\mathbf{y}_1, \ldots, \mathbf{y}_k$, and scalars $x_1, \ldots, x_k$ and $y_1, \ldots, y_k$, the following property holds:

$$\left(\sum_{m=1}^{k} a_m \mathbf{x}_m \right) \circledast \left(\sum_{n=1}^{k} b_n \mathbf{y}_n \right) = \sum_{m,n=1}^{k} a_m b_n \mathbf{x}_m \mathbf{y}_n \tag{33}$$

Now, we check $\circledast$ satisfies the convolution theorem.

$$\begin{aligned} \mathbf{F}(\mathbf{x} \circledast \mathbf{y}) = \mathbf{F}(\mathbf{s}_{h_\mathbf{x} h_\mathbf{y}}) &= \mathbf{F}(\mathbf{F}^{-1} h_\mathbf{x} h_\mathbf{y}(\mathbf{D}) \mathbf{1}) \\ &= h_\mathbf{x} h_\mathbf{y}(\mathbf{D})\mathbf{1} = \left(h_\mathbf{x}(\mathbf{D})\mathbf{1} \right) \cdot \left(h_\mathbf{y}(\mathbf{D})\mathbf{1} \right) \\ &= \left(\mathbf{F}\, \mathbf{s}_{h_\mathbf{x}} \right) \cdot \left(\mathbf{F}\, \mathbf{s}_{h_\mathbf{y}} \right) \\ &= \mathbf{F}\mathbf{x} \cdot \mathbf{F}\mathbf{y} \end{aligned}$$

Finally we check that the system $h(\mathbf{A})$ is a convolution operator.

$$\begin{aligned} h(\mathbf{A})\,\mathbf{x} = \mathbf{F}^{-1} h(\mathbf{D})\,\mathbf{F}\,\mathbf{x} &= \left(\mathbf{F}^{-1} h(\mathbf{D})\,\mathbf{F} \right) \mathbf{s}_{h_\mathbf{x}} \\ &= \mathbf{F}^{-1} h h_\mathbf{x}(\mathbf{D})\mathbf{1} = \mathbf{s}_{h h_\mathbf{x}} = \mathbf{s}_h \circledast \mathbf{x} \end{aligned}$$

$\square$

Remark 9. *With the notation of Theorem 8, let $\mathbf{d}_h$ be the signal obtained by the diagonal entries of the matrix $h(\mathbf{D})$, forming the set of its eigenvalues. As mentioned earlier, the columns of $\mathbf{V} = \mathbf{F}^{-1}$ constitute the eigenvectors of $h(\mathbf{D})$. One may directly see that,*

$$\mathbf{s}_h = \mathbf{V}\,\mathbf{d}_h \tag{34}$$

Example 10. *Consider the line digraph $\mathbf{G}$ with N vertices and its admissible extension, the cycle digraph $\mathbf{G}_w$, which is formed by connecting the sink to the source with a positive weight w. We compute the convolution product $\circledast$ on $\mathbf{G}_w$. As demonstrated in Example 6, $\mathbf{F}_w = \mathbf{V}_w^{-1}$ serves as the GFT on $\mathbf{G}_w$, where $\mathbf{V}_w$ is defined in (24). For given unit impulse*

TABLE II
MINIMUM AND MAXIMUM VALUES OF $(\delta_{\mathbf{V}}, \Delta_{\mathbf{V}})$ AND CONDITION
NUMBER FOR $w = 1$ AND $w = 0.01$

Weight		$\delta_{\mathbf{V}}$	$\Delta_{\mathbf{V}}$	Condition Number
$w = 1$	Minimum	0.0359	0.1707	39.011
	Maximum	0.1195	0.3165	$5.021e + 16$
$w = 0.01$	Minimum	0.00108	0.00395	427.899
	Maximum	0.00119	0.00457	$3.181e + 15$

signals $\delta[m]$ and $\delta[n]$, we have:

$$\delta[m] \circledast \delta[n] = \mathbf{F}_w^{-1} \left(\mathbf{F}_w \delta[m] \cdot \mathbf{F}_w \delta[n] \right)$$

$$= \mathbf{E}_w \mathfrak{F}^{-1} \left(\mathfrak{F} \mathbf{E}_w^{-1} \delta[m] \cdot \mathfrak{F} \mathbf{E}_w^{-1} \delta[n] \right)$$

$$= \mathbf{E}_w \mathfrak{F}^{-1} \left(\mathfrak{F} \left(\mathbf{E}_w^{-1} \delta[m] * \mathbf{E}_w^{-1} \delta[n] \right) \right)$$

$$= \mathbf{E}_w \left(\mathbf{E}_w^{-1} \delta[m] * \mathbf{E}_w^{-1} \delta[n] \right)$$

$$= \mathbf{E}_w \left(w^{-\frac{m}{N}} \delta[m] * w^{-\frac{n}{N}} \delta[n] \right)$$

$$= w^{-\frac{m+n}{N}} \mathbf{E}_w \left(\delta[m] * \delta[n] \right) = \sqrt{N} w^{-\frac{m+n}{N}} \mathbf{E}_w \delta[m+n]$$

$$= \sqrt{N} \delta[m+n] = \delta[m] * \delta[n]$$

where $m + n$ is taken modulo N.

Since any signal can be represented as a linear combination of unit impulse signals, the convolution product $\circledast$ on the directed graph $\mathbf{G}_w$ becomes the same as the conventional convolution product $*$.

To determine the optimal admissible extension for the line digraph with N vertices, we can combine the insights obtained from Examples 6 and 10. These examples demonstrate that the cyclic digraph $\mathbb{Z}_N$ is the most suitable admissible extension for the line digraph.

III. APPLICATION : FRIENDSHIP DIGRAPH

Upon examining the discourse leading to equation (14), it becomes apparent that the number of admissible extensions for a digraph with a singular adjacency matrix can potentially be quite substantial. This finding is further supported a by real-world data sets, given in subsequent section, which demonstrate that the actual number of admissible extensions far surpasses what one might anticipate at the outset. Recognizing the abundance of admissible extensions, identifying the optimal choice that caters to the targeted goal based on the specific data set under examination is highly beneficial.

A. Friendship Digraph in a High School in Illinois

The digraph in Figure 5 shows the friendship connections among boys at a small high school in Illinois, which has 70 students. We will denote this digraph as $\mathbf{G}_f$. Each boy was surveyed once in the fall of 1957 and again in the spring of 1958. The data set[1] consolidates the results from the initial survey. In this network, a node represents a boy, and an edge between two nodes indicates that the boy on the left selected the boy on the right as a friend. The adjacency matrix is of rank 66, signifying its singular nature. Consequently, we
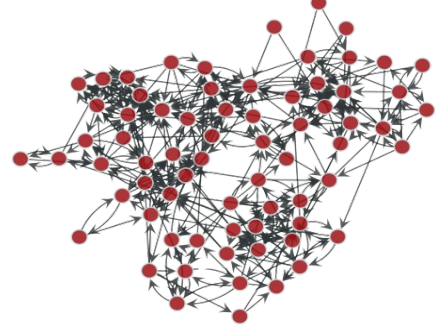
[1] https://networks.skewed.de/net/highschool

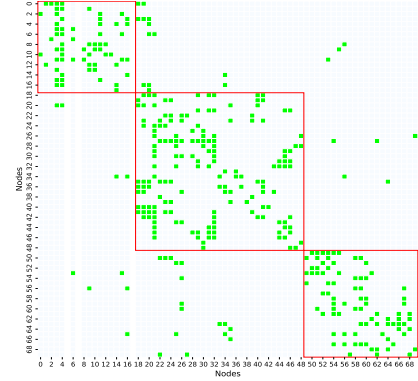Fig. 5. Friendship digraph $\mathbf{G}_f$ with 70 vertices.

Fig. 6. Adjacency matrix associated with Friendship digraph $\mathbf{G}_f$ clustered into three clusters using spectral clustering.

must firstly identify and incorporate 4 additional edges into the digraph to successfully dismantle all Jordan blocks linked to the eigenvalue 0. To enhance our visualization, the heat map of the adjacency matrix (Figure 6) is arranged into three distinct clusters using the spectral clustering algorithm. The digraph contains three sinks, which are highlighted in the heat map. Each sink must be incorporated into every column dependency list. There are only 5 column dependency lists in total, but there are 15495 different row dependency lists. Therefore, through this approach, the total number of ways to make the initial adjacency matrix non-singular is as follows:

$$\text{Total}_{\text{Inv}} = 15,495 \times 5 \times 4! = 1,859,400$$

By avoiding multiple edges, 596,713 different cases of admissible extensions are emerged.

Remark 11. *It is worth mentioning that while the adjacency matrix lacks a source, it does contain two duplicate rows. Interestingly, the exclusion of these duplicates does not impact the rank of the matrix. Additionally, there exist only 1555 row dependency lists, which encompass the two duplicate rows, referred to as D-Rows. The numerical findings suggest that, to attain the desired outcome, it is adequate to limit our consideration to instances where pseudo-permutations are constructed from row dependency lists derived from D-Rows. Within this context, we focus on a total of 7775 non-singular*

Indices	Kendall's Tau value	Condition number
$\mathbf{G}_{\text{Max-Tau}}$	0.92148	81062064.706
$\mathbf{G}_{\text{env-pr}}$	0.914854	69.14729

extensions, with 7243 of these cases constituting admissible extensions avoiding multiple edges. Henceforth referred to as the 7243-sample set in the remainder of this discussion.

1) Analyzing Eigenvector Approximation: We begin by examining the indices (δ, Δ) indices (see (21)) derived from the graph Fourier transforms on the adjacency matrices in the 7243-sample set, without considering any weights on the newly added edges. The datas presented in Table II indicate a high degree of spectral compatibility. However, there is a noticeable variation in the condition numbers of these transforms, which warrants further investigation.

The distribution of these indices is illustrated in Figures 7. A similar analysis is presented in Table II, which focus on the scenario where new edges are introduced with a fixed weight of $w = 0.01$.

The high condition numbers observed when the newly added edges are assigned a weight of 0.01, along with the small indices of $\delta_{\mathbf{V}}$ and $\Delta_{\mathbf{V}}$ when no weight is assigned, serve as compelling evidence in favor of implementing the 7243-sample set without considering any weight. This decision is based on the analysis of the provided data, which suggests that this approach yields.

2) Structure Compatibility: As the structural integrity of the original digraph is increasingly preserved upon the addition of new edges, the graph Fourier transform logically becomes a more appropriate choice for processing graph signals.

In the subsequent discourse, we propose a framework for determining suitable admissible extensions of the given digraph that prioritize maintaining the structural integrity between the original graph and its extended version.

Remark 12. *In the process of comparing social networks, numerous indices are frequently employed, including centrality measures such as PageRank, and additional metrics like the Global Clustering Coefficient, Network Motif Density, and Number of Core/Periphery nodes. By applying these metrics to the 7243-sample set, a filtering mechanism is established to identify the most appropriate admissible extensions that exhibit the highest structural integrity with the Friendship digraph. This enables the preservation of essential properties and characteristics while extending the network. Brief explanations of these metrics are provided in the Appendix section.*

Filter 1: PageRank-based Filtering. To initiate our analysis, we prioritize the PageRank criterion, a fundamental measure of node significance within networks. The Kendall's Tau index informs us about the rate of changes in the PageRanks of
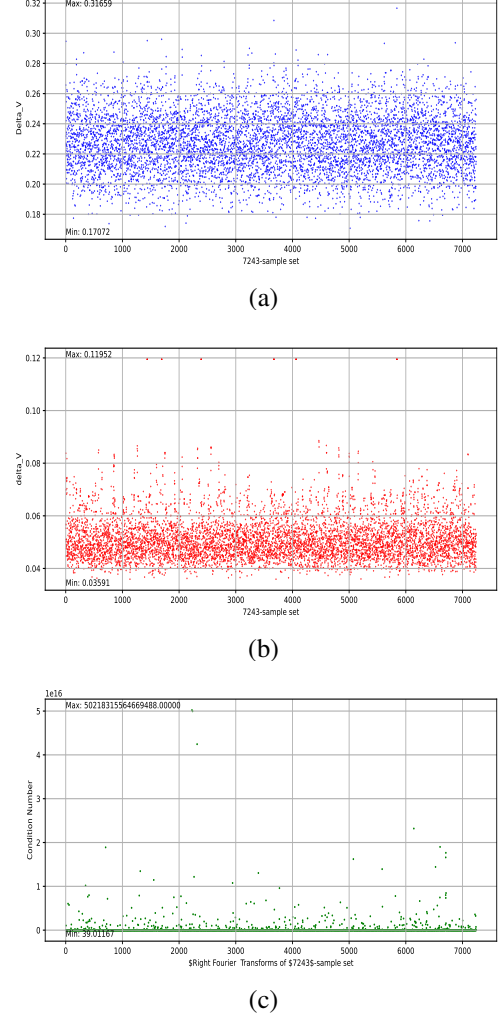
(a)

(b)

(c)

Fig. 7. Figures (a), (b), and (c) illustrate the distribution of (δ, Δ) indices and condition numbers of the GFTs obtained from the 7243-sample set.
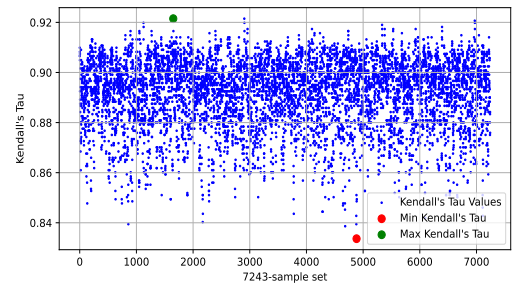
Fig. 8. The distribution of Kendall's Tau indices corresponding to the 7243-sample set due to reshuffling after adding new edges.
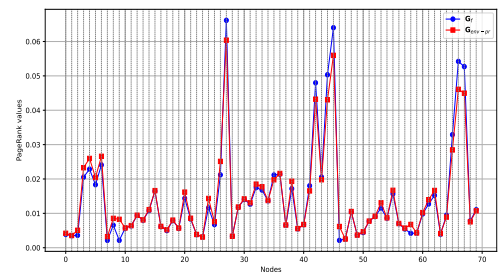
Fig. 9. Comparison of PageRanks for digraphs $\mathbf{G}_{\text{f}}$ and $\mathbf{G}_{\text{env-pr}}$.
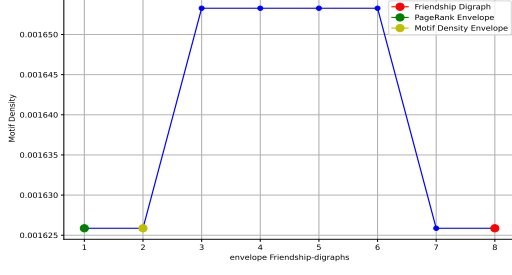
Fig. 10. The Motif density indices of $\mathbf{G}_\text{f}$ and envelope Friendship digraphs.

Fig. 11. Comparison of Core and Periphery Node Counts in $\mathbf{G}_\text{f}$ and Envelope Friendship Digraphs

nodes that occur due to reshuffling after adding new edges. The distribution of Kendall's Tau indices corresponding to the 7243-sample set is depicted in Figure 8.

We focus on those admissible digraphs that satisfy two specific conditions: a Kendall's Tau index of at least 0.91 and a condition number of 80 or less. These criteria ensure a balance between spectral and structural integrity in relation to the original Friendship digraph.

Out of the admissible digraphs in the 7243-sample set, 223 possess a Kendall's Tau index of 0.91 or higher, indicating strong concordance with the original graph's ordering. However, only 7 of these digraphs meet the condition number threshold of 80 or less. We refer to these 7 digraphs as *envelope Friendship-digraphs*.

Within the 7243-sample set, the digraph exhibiting the highest Kendall's Tau index is designated as $\mathbf{G}_\text{Max-Tau}$, while $\mathbf{G}_\text{env-pr}$ represents the digraph that achieves the maximum Kendall's Tau index among all envelope Friendship-digraphs. In Figure 9, PageRanks of the digraphs $\mathbf{G}_\text{f}$ and $\mathbf{G}_\text{env-pr}$ are compared together.

The findings in Table III underscore the importance of achieving a balance between condition number and structural compatibility when selecting the optimal graph Fourier transform for the Friendship digraph.

Filter 2: Motif Density Filtering. Motif Density assesses the frequency and prevalence of recurring significant patterns or motifs within a graph structure. As shown in Figure 10, a comparative analysis of the network motif density indices for the envelope Friendship-digraphs and the Friendship-digraph index reveals a striking similarity. In order to facilitate a more comprehensive analysis, an additional digraph, termed $\mathbf{G}_\text{env-motif}$, is introduced and highlighted in the plot using a yellow marker.

Filter 2: Core and Periphery counts Filtering. Core-Periphery counts evaluate the number of nodes that can be classified as either core or periphery within a graph structure. Core nodes are highly interconnected and central, while periphery nodes are less connected and occupy the outer layers of the network. In Figure 11, the Core and Periphery counts for the digraphs within the envelope Friendship-digraphs list are provided. The indices for the digraphs and notable digraphs such as $\mathbf{G}_\text{f}$, $\mathbf{G}_\text{env-pr}$, $\mathbf{G}_\text{env-motif}$, and a newly introduced $\mathbf{G}_\text{env-cor}$ are highlighted for emphasis.

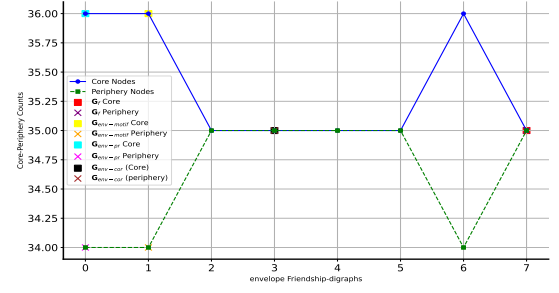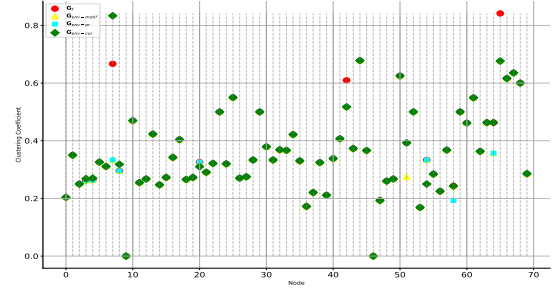*Local clustering coefficients.* The local clustering coefficient

Fig. 12. Local Clustering Coefficients in $\mathbf{G}_\text{f}$, $\mathbf{G}_\text{env-pr}$, $\mathbf{G}_\text{env-motif}$, and $\mathbf{G}_\text{env-cor}$.
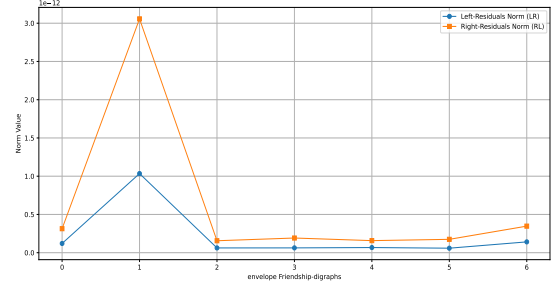

Fig. 13. This figure illustrates the operator norms of $(\mathbf{F}\,\mathbf{F}^{-1} - \mathbf{I})$ and $(\mathbf{F}^{-1}\,\mathbf{F} - \mathbf{I})$ for GFT of envelope Friendship digraphs.

of a node in a graph measures how close its neighbors are to being a complete subgraph (clique). It assesses the extent to which nodes tend to cluster together locally within their immediate neighborhoods. As demonstrated in Figure 12, a comparative analysis of the local clustering coefficients for the notable digraphs $\mathbf{G}_\text{f}$, $\mathbf{G}_\text{env-pr}$, $\mathbf{G}_\text{env-motif}$, and $\mathbf{G}_\text{env-cor}$ is conducted, offering valuable insights into their integrity.

Numerical stability : envelope Friendship-digraphs. To assess numerical stability, we compute the operator norms of $(\mathbf{F}\,\mathbf{F}^{-1} - \mathbf{I})$ and $(\mathbf{F}^{-1}\,\mathbf{F} - \mathbf{I})$, where $\mathbf{F}$ denotes the GFT of the envelope Friendship-digraphs. This computation evaluates the deviation between the product of inverse GFTs and the identity matrix, providing a thorough analysis of numerical stability. The results are illustrated in Figure 13.

B. Systems : Target envelopes

The methodology involves calculating three distinct graph Fourier transforms, referred to as *target envelopes*, specifically $\mathbf{G}_\text{env-pr}$, $\mathbf{G}_\text{env-cor}$, $\mathbf{G}_\text{env-motif}$, for the Friendship digraph $\mathbf{G}_\text{f}$. These target envelopes introduce the GFTs $\mathbf{F}_\text{pr}$, $\mathbf{F}_\text{cor}$, $\mathbf{F}_\text{motif}$

	(r,c)	(r,c)	(r,c)	(r,c)
$\mathbf{G}_{\text{env-pr}}$	(43,5)	(29,7)	(44,17)	(7,37)
$\mathbf{G}_{\text{env-motif}}$	(43,5)	(29,7)	(44,17)	(7,66)
$\mathbf{G}_{\text{env-cor}}$	(11,5)	(29,7)	(44,17)	(7,68)

for $\mathbf{G}_{\text{f}}$, respectively. Each of these transforms establishes a specific graph Fourier basis on $\mathbf{G}_{\text{f}}$. Our subsequent analysis focuses on checking and interpreting the systems induced on $\mathbf{G}_{\text{f}}$ by these target envelopes.

Let $\mathbf{A}_{\text{pr}}, \mathbf{A}_{\text{cor}}, \mathbf{A}_{\text{motif}}$ be the adjacency matrices associated with the three envelope targets. The list of newly added edges is addressed in Table IV. The heatmaps presented in Figure 14 provide a visual comparison of the spectral characteristics of the graphs for the three target envelopes. These heatmaps illustrate the magnitude differences between the eigenvectors of the adjacency matrices corresponding to the target envelope digraphs, offering a comprehensive representation of their spectral properties and allowing for an intuitive understanding of their similarities.
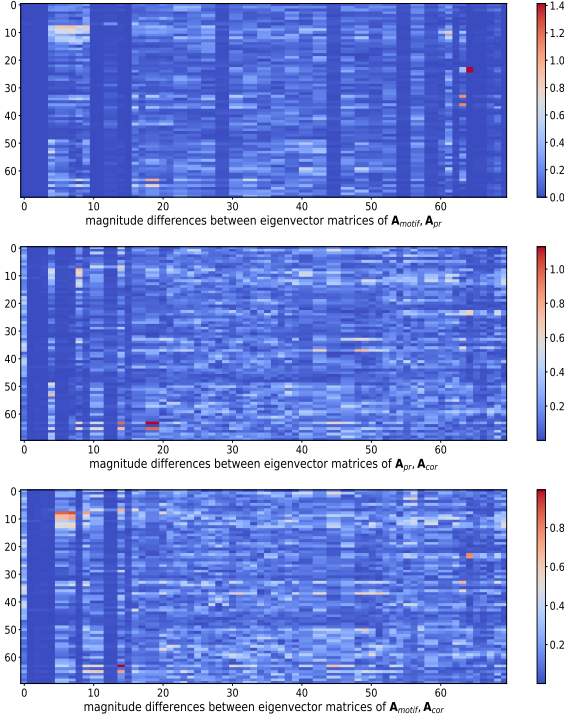

Fig. 14. The figures visualize the magnitude differences between columns for any given pair of eigenvector matrices of the adjacency matrices corresponding to the target envelope digraphs.

The following graph signals illustrate the systems induced by the target envelopes.

$$\begin{cases} \mathbf{s}_{\text{pr}} = \mathbf{F}_{\text{pr}}^{-1} \mathbf{D}_{\text{pr}} \mathbf{1} \\ \mathbf{s}_{\text{cor}} = \mathbf{F}_{\text{cor}}^{-1} \mathbf{D}_{\text{cor}} \mathbf{1} \\ \mathbf{s}_{\text{motif}} = \mathbf{F}_{\text{motif}}^{-1} \mathbf{D}_{\text{motif}} \mathbf{1} \end{cases} \quad (35)$$

where $\mathbf{D}_{\text{pr}}$, $\mathbf{D}_{\text{cor}}$, and $\mathbf{D}_{\text{motif}}$ are diagonal matrices containing the eigenvalues of $\mathbf{A}_{\text{pr}}$, $\mathbf{A}_{\text{cor}}$, and $\mathbf{A}_{\text{motif}}$. Figures 15 compare the real and imaginary parts of the signals defined in (35) .

Summary. In summary, the findings of this section are:

i) A filter mechanism was developed to identify appropriate GFTs for the digraph $\mathbf{G}_{\text{f}}$, focusing on achieving maximal spectral and structural compatibility. Consequently, three target envelopes were obtained, and it was observed that the resultant digraphs converged to similar characteristics as the filtering process progressed. The adjacency matrices corresponding to $\mathbf{G}_{\text{env-pr}}$ and $\mathbf{G}_{\text{env-motif}}$ were found to be identical, except for the addition of a single new edge with the same source.

ii) To analyze the impact of these envelopes on $\mathbf{G}_{\text{f}}$, comparing shift-invariant graph filters offers a suitable approach. The systems defined by them on $\mathbf{G}_{\text{f}}$ are characterized by their convolution operators, as established in Theorem 8.

As illustrated in Figure 15, the systems induced by these three convolution operators exhibit sufficiently distinct characteristics. This indicates that even with a certain level of spectral and structural compatibility, each GFT derived from the envelope digraphs corresponds to a unique case. Consequently, the different GFTs demonstrate distinct behaviors, thereby validating the hypothesis that these GFTs have unique impacts on signal processing within graph-based data frameworks.

C. Conclusion

When a real dataset is represented as a digraph, graph signal processing techniques might be employed for signal analysis. One significant contribution is formulating the challenges to obtain a well-structured graph Fourier basis, which is intricately linked to the nature of the graph structure. To achieve the desired target, modifying the graph structure to develop the GFT could be a viable solution.

Admissibility is a minimal property in a digraph that allows the utilization of shift-invariant graph filters for signal analysis, making them highly effective tools.

For a given digraph $\mathbf{G}$, a systematic approach is delineated to identify a class of sophisticated admissible Cayley digraphs $\text{Cay}(\mathbb{Z}_N, \Gamma)$. This ensures that $\mathbf{G}$ fits within various admissible envelopes, enabling the application of GSP techniques.

To achieve spectral and structural compatibility with $\mathbf{G}$, it is necessary to design an appropriate filtering mechanism. The initial problem considered on the real dataset dictates the selection of the optimal option among the admissible envelopes of $\mathbf{G}$ that emerge through the filtering process.

Any filtering mechanism will lead to a specific envelope digraph that is suitable only for that particular case. Consequently, choosing admissible extensions randomly may not yield reliable results.

IV. APPENDIX

A. Cayley digraphs $Cay(\mathbb{Z}_N, \Gamma)$

Let $\mathbb{Z}_N$ denote the cyclic group of integers modulo N, where N is a positive integer. The group $\mathbb{Z}_N$ consists of the elements $\{0, 1, 2, \ldots, N - 1\}$ under addition modulo N.

Given a subset $\Gamma \subseteq \mathbb{Z}_N$, the *Cayley digraph* $\text{Cay}(\mathbb{Z}_N, \Gamma)$ is a directed graph defined as follows:

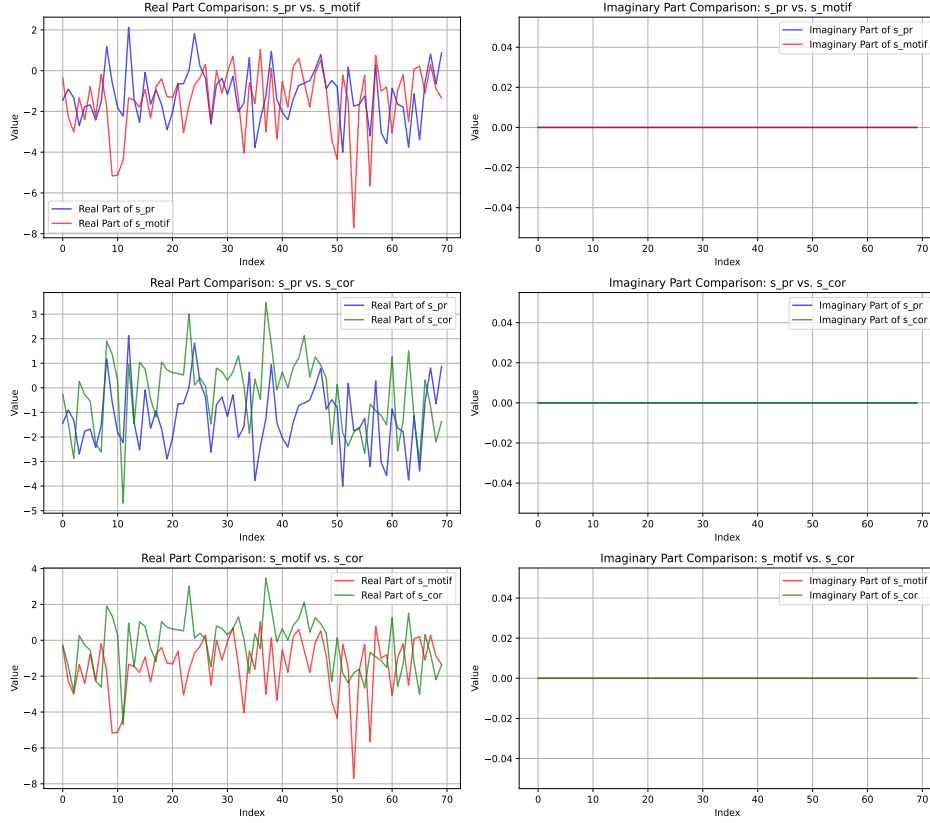- The vertex set of the digraph is the set of elements of $\mathbb{Z}_N$.

Fig. 15. The real and imaginary parts of signals defined in (35) are compared together in the figures.

- There is a directed edge from vertex m to vertex n if and only if $n - m \in \Gamma \pmod{N}$.

In other words, the adjacency relation in the Cayley digraph $\mathrm{Cay}(\mathbb{Z}_N, \Gamma)$ is given by $(m, n) \in \mathbb{Z}_N \times \mathbb{Z}_N$ iff $n \equiv m + k \pmod{N}$ for some $k \in \Gamma$.

The set Γ is often referred to as the *connection set* of the Cayley digraph. It is important to note that Γ should not contain the identity element (0 in the case of $\mathbb{Z}_N$) to avoid loops at each vertex unless loops are desired in the digraph.

Proof of Theorem 2. In the specific case $\Gamma = \{1\}$, the Cayley digraph $\mathrm{Cay}(\mathbb{Z}_N, \Gamma)$ simplifies to the directed cycle digraph with N vertices whose adjacency matrix is just the circulant matrix $\mathbf{C}$. It is known diagonalizable with the following eigenvalue decomposition.

$$\begin{cases} \mathbf{C} = \mathfrak{F}^{-1} \mathbf{D} \, \mathfrak{F} \\ \mathbf{D} = \mathrm{diag}\Big(1, \exp\left(-\frac{2\pi i}{N}\right), \ldots, \exp\left(-\frac{2\pi i(N-1)}{N}\right)\Big) \end{cases} \quad (36)$$

Let $\Gamma = \{n_1, \ldots, n_k\}$ be a connection set and $\mathbf{A}_\Gamma$ as the associative adjacency matrix for the Cayley graph $\mathrm{Cay}(\mathbb{Z}_N, \Gamma)$. Then one may directly check that,

$$\mathbf{A}_\Gamma = \sum_{q=1}^{k} \mathbf{C}^{n_q} = \sum_{q=1}^{k} (\mathfrak{F}^{-1} \mathbf{D} \, \mathfrak{F})^{n_q}$$

$$= \mathfrak{F}^{-1} \Big(\sum_{q=1}^{k} \mathbf{D}^{n_q} \Big) \mathfrak{F}$$

The matrix $\mathbf{D}_\Gamma = \sum_{q=1}^{k} \mathbf{D}^{n_q}$ is a diagonal matrix whose diagonal entries are precisely the eigenvalues of $\mathbf{A}_\Gamma$. Conse-

quently, $\mathbf{A}_\Gamma = \mathfrak{F}^{-1} \mathbf{D}_\Gamma \, \mathfrak{F}$ represents the eigenvalue decomposition of $\mathbf{A}_\Gamma$. This proves that the discrete Fourier matrix is the graph Fourier transform for all Cayley digraphs $\mathrm{Cay}(\mathbb{Z}_N, \Gamma)$.

B. A brief description for metrics

1) PageRank Centrality: The PageRank algorithm, initially created by Google for ranking web pages, is used in network analysis to measure node importance based on incoming link structure. Higher scores indicate more central nodes.

$$PR(i) = \frac{1-d}{N} + d \sum_{j \in M(i)} \frac{PR(j)}{L(j)}$$

- $PR(i)$ is the PageRank of node i.
- d is the damping factor, usually set to 0.85.
- N is the total number of nodes in the graph.
- $M(i)$ is the set of nodes that link to node i.
- $L(j)$ is the number of outbound links from node j.

The Kendall rank correlation coefficient (Kendall's tau) measures similarity between two rankings. It evaluates ordinal association by calculating differences in item pair orderings.

The formula for Kendall's tau is:

$$\tau = \frac{\big(\text{number of concordant pairs}\big) - \big(\text{number of discordant pairs}\big)}{\binom{n}{2}}$$

- τ is Kendall's tau coefficient,
- n is the number of paired observations,
- $\binom{n}{2}$ represents the number of pairs of items.

2) Motif Density: Motif Density quantifies the prevalence of recurring significant patterns or motifs within a graph. The formula for Motif Density can be expressed as:

$$\text{Motif Density}(M) = \frac{n_M}{N_G}$$

- n_M is the number of occurrences of motif M in G.
- N_G is the total number of potential occurrences of motif M in graph G, considering its size and configuration.

3) Core-Periphery Criterion:: Core-Periphery counts evaluate how nodes within a graph can be categorized into core and periphery groups based on their connectivity patterns.

- **Core Nodes**: These nodes have higher degrees of connectivity and play central roles within the network.
- **Periphery Nodes**: These nodes have fewer connections and are typically situated on the edges of the network.

The core-periphery structure score S_i for each node i is given:

$$S_i = \frac{k_i - \langle k \rangle}{k_i + \langle k \rangle}$$

- k_i is the degree (number of connections) of node i,
- $\langle k \rangle$ is the average degree of all nodes in the graph.

Nodes with $S_i > 0$ are classified as core nodes, while nodes with $S_i \leq 0$ are classified as periphery nodes.

4) Global Clustering Coefficient: The local clustering coefficient of a node in a graph measures the degree to which its neighbors are interconnected. It quantifies how close the node's neighbors are to forming a complete subgraph (clique), thereby indicating the local density of connections around the node. The local clustering coefficient C_i for a node i is given:

$$C_i = \frac{2 \times e_i}{k_i \times (k_i - 1)}$$

- k_i is the degree of node i.
- e_i is the number of edges between the neighbors of node i.

This formula computes the ratio of the actual number of edges between the neighbors of node i (denoted by e_i) to the maximum possible number of edges that could exist between them (given by $k_i \times (k_i - 1)/2$, which is the number of edges in a complete graph of k_i nodes).

REFERENCES

[1] B. Seifert and M. Püschel, "Digraph signal processing with generalized boundary conditions," *IEEE Transactions on Signal Processing*, vol. 69, pp. 1422–1437, 2021.

[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 processing magazine*, vol. 30, no. 3, pp. 83–98, 2013.

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

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

[5] 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 processing magazine*, vol. 31, no. 5, pp. 80–90, 2014.

[6] X. Dong, D. Thanou, L. Toni, M. Bronstein, and P. Frossard, "Graph signal processing for machine learning: A review and new perspectives," *IEEE Signal processing magazine*, vol. 37, no. 6, pp. 117–127, 2020.

[7] B.-N. S. O. GRAPH, "Graph signal processing: Foundations and emerging directions," *IEEE Signal Processing Magazine*, 2020.

[8] E. Isufi, F. Gama, D. I. Shuman, and S. Segarra, "Graph filters for signal processing and machine learning on graphs," *IEEE Transactions on Signal Processing*, 2024.

[9] R. Li, X. Yuan, M. Radfar, P. Marendy, W. Ni, T. J. O'Brien, and P. M. Casillas-Espinosa, "Graph signal processing, graph neural network and graph learning on biological data: a systematic review," *IEEE Reviews in Biomedical Engineering*, vol. 16, pp. 109–135, 2021.

[10] G. Leus, A. G. Marques, J. M. Moura, A. Ortega, and D. I. Shuman, "Graph signal processing: History, development, impact, and outlook," *IEEE Signal Processing Magazine*, vol. 40, no. 4, pp. 49–60, 2023.

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

[12] L. Stanković, D. Mandic, M. Daković, M. Brajović, B. Scalzo, S. Li, A. G. Constantinides *et al.*, "Data analytics on graphs part i: Graphs and spectra on graphs," *Foundations and Trends® in Machine Learning*, vol. 13, no. 1, pp. 1–157, 2020.

[13] ——, "Data analytics on graphs part ii: Signals og graphs," *Foundations and Trends® in Machine Learning*, vol. 13, no. 1, pp. 158–331, 2020.

[14] J. Shi and J. M. Moura, "Graph signal processing: Modulation, convolution, and sampling," *arXiv preprint arXiv:1912.06762*, 2019.

[15] W. Huang, T. A. Bolton, J. D. Medaglia, D. S. Bassett, A. Ribeiro, and D. Van De Ville, "A graph signal processing perspective on functional brain imaging," *Proceedings of the IEEE*, vol. 106, no. 5, pp. 868–885, 2018.

[16] M. Chen, Z. Wei, Z. Huang, B. Ding, and Y. Li, "Simple and deep graph convolutional networks," in *International conference on machine learning*. PMLR, 2020, pp. 1725–1735.

[17] L. Stanković, D. P. Mandic, M. Daković, I. Kisil, E. Sejdic, and A. G. Constantinides, "Understanding the basis of graph signal processing via an intuitive example-driven approach [lecture notes]," *IEEE Signal Processing Magazine*, vol. 36, no. 6, pp. 133–145, 2019.

[18] J. Domingos and J. M. Moura, "Graph fourier transform: A stable approximation," *IEEE Transactions on Signal Processing*, vol. 68, pp. 4422–4437, 2020.

[19] H. Sevi, G. Rilling, and P. Borgnat, "Harmonic analysis on directed graphs and applications: From fourier analysis to wavelets," *Applied and Computational Harmonic Analysis*, vol. 62, pp. 390–440, 2023.

[20] S. Sardellitti, S. Barbarossa, and P. Di Lorenzo, "On the graph fourier transform for directed graphs," *IEEE Journal of Selected Topics in Signal Processing*, vol. 11, no. 6, pp. 796–811, 2017.

[21] B. Seifert, C. Wendler, and M. Püschel, "Causal fourier analysis on directed acyclic graphs and posets," *IEEE Transactions on Signal Processing*, vol. 71, pp. 3805–3820, 2023.

[22] R. Shafipour, A. Khodabakhsh, G. Mateos, and E. Nikolova, "Digraph fourier transform via spectral dispersion minimization," in *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2018, pp. 6284–6288.

[23] S. Furutani, T. Shibahara, M. Akiyama, K. Hato, and M. Aida, "Graph signal processing for directed graphs based on the hermitian laplacian," in *Machine Learning and Knowledge Discovery in Databases: European Conference, ECML PKDD 2019, Würzburg, Germany, September 16–20, 2019, Proceedings, Part I*. Springer, 2020, pp. 447–463.

[24] L. Yang, A. Qi, C. Huang, and J. Huang, "Graph fourier transform based on 1 norm variation minimization," *Applied and Computational Harmonic Analysis*, vol. 52, pp. 348–365, 2021.

[25] A. Sandryhaila and J. M. Moura, "Discrete signal processing on graphs: Frequency analysis," *IEEE Transactions on signal processing*, vol. 62, no. 12, pp. 3042–3054, 2014.

[26] B. Girault, A. Ortega, and S. S. Narayanan, "Irregularity-aware graph fourier transforms," *IEEE Transactions on Signal Processing*, vol. 66, no. 21, pp. 5746–5761, 2018.

[27] R. Shafipour, A. Khodabakhsh, G. Mateos, and E. Nikolova, "A directed graph fourier transform with spread frequency components," *IEEE Transactions on signal processing*, vol. 67, no. 4, pp. 946–960, 2018.

[28] D. Wei and S. Yuan, "Hermitian random walk graph fourier transform for directed graphs and its applications," *Digital Signal Processing*, vol. 155, p. 104751, 2024.

[29] C. D. Godsil and W. Imrich, "Embedding graphs in cayley graphs," *Graphs and combinatorics*, vol. 3, no. 1, pp. 39–43, 1987.

[30] A. J. Przeździecki, "An almost full embedding of the category of graphs into the category of abelian groups," *Advances in Mathematics*, vol. 257, pp. 527–545, 2014.