

# Optimal Fractional Fourier Filtering for Graph Signals

Cuneyd Ozturk , *Student Member, IEEE*, Haldun M. Ozaktas, *Fellow, IEEE*, Sinan Gezici , *Senior Member, IEEE*, and Aykut Koç , *Senior Member, IEEE*

**Abstract**—Graph signal processing has recently received considerable attention. Several concepts, tools, and applications in signal processing such as filtering, transforming, and sampling have been extended to graph signal processing. One such extension is the optimal filtering problem. The minimum mean-squared error estimate of an original graph signal can be obtained from its distorted and noisy version. However, the best separation of signal and noise, and thus the least error, is not always achieved in the ordinary Fourier domain, but rather a fractional Fourier domain. In this work, the optimal filtering problem for graph signals is extended to fractional Fourier domains, and theoretical analysis and solution of the proposed problem are provided along with computational cost considerations. Numerical results are presented to illustrate the benefits of filtering in fractional Fourier domains.

**Index Terms**—Fractional Fourier transform, graph signal processing (GSP), optimal filtering, Wiener filter, graph Fourier transform (GFT), signal processing on graphs, graphs.

## I. INTRODUCTION

DISCRETE signal processing on graphs, or graph signal processing (GSP) is concerned with analyzing data which reside on irregular and/or complex structures [1]–[14]. In applications including social, neural, and sensor networks, data can be modeled on the vertices of a weighted graph [2], [4]. While conventional signal processing formulations may be inadequate to model and analyze such data, GSP may make possible the extraction of intrinsic complex relationships in such irregular and complex data [2], [10].

Many standard concepts in signal processing have been extended to graph signals [9], [10], [12]. Specifically, filtering [1]–[3], [7], [9], [10], [13], [15], frequency analysis [5], [13], [16], sampling [6], [17], interpolation [18], Fourier transformation [19], multiscale decompositions and approaches [20]–[23], and processing of stationary signals and processes [24], [25]

Manuscript received September 12, 2020; revised January 27, 2021; accepted May 6, 2021. Date of publication May 19, 2021; date of current version May 28, 2021. The associate editor coordinating the review of this manuscript and approving it for publication was S. Das. The work of Haldun M. Ozaktas was supported in part by the Turkish Academy of Sciences. (*Corresponding author: Aykut Koç.*)

Cuneyd Ozturk, Haldun M. Ozaktas, and Sinan Gezici are with the Department of Electrical and Electronics Engineering, Bilkent University, Ankara TR-06800, Turkey (e-mail: cuneyd@ee.bilkent.edu.tr; haldun@ee.bilkent.edu.tr; gezici@ee.bilkent.edu.tr).

Aykut Koç is with the Department of Electrical and Electronics Engineering, Bilkent University, Ankara TR-06800, Turkey and also with the UMRAM, Bilkent University, Ankara TR-06800, Turkey (e-mail: aykut.koc@bilkent.edu.tr).

Digital Object Identifier 10.1109/TSP.2021.3079804

have been considered for graph signals. Graph based extensions to areas such as machine learning and brain signal analysis have also been studied [26], [27].

There exists two main approaches for adapting standard signal processing concepts to GSP. One is based on the graph Laplacian matrix [2], [12], [14]. Although the graph Laplacian matrix is generally restricted to be symmetric and positive semi-definite, there exist attempts for generalizations [28]. However, it is generally not possible to deploy the graph Laplacian matrix approach for directed graphs or negative edge weights [6]. The other approach is based on the graph shift operator, which is also known as the graph adjacency matrix [1], [3]–[11], [13]. The graph adjacency matrix is not restricted to be either symmetric or positive semi-definite and therefore has broader applicability. In this work, we adopt the latter approach in its most general form to cover both directed/undirected and weighted/unweighted graphs.

A graph signal is defined as a mapping between the set of vertices of a graph and a set of complex numbers. The definition of graph filters is given in [1]–[5], [7], [9], [10], [13] based on the shift operator for graph signals. As a concept that is the counterpart of linear time-invariant (LTI) systems in standard signal processing, linear shift-invariant (LSI) filters are defined in [1], [2], [7], [9], [10], [13]. In [7], the optimal filtering problem on graphs is discussed. Given a noisy version of the actual graph signal, the filter leading to the minimum mean-squared error (MSE) is found under the assumptions that the filter is LSI and the noise is a white Gaussian process. This leads to the classical Wiener-Hopf equations. Adaptive versions of graph filters, which combine multiple graph filters, for semi-supervised classification, are discussed in [29], [30].

The concept of the graph spectral domain, which is the counterpart of the conventional frequency domain, is introduced through the graph Fourier transform (GFT) [1]–[11], [13], [14]. In the literature, there exists two main GFT definitions, which use the spectral structure of the graph Laplacian matrix or the graph adjacency matrix. The first is built on *spectral graph theory* and uses the graph Laplacian. This approach defines the GFT as a change of basis to the basis of the eigenvectors of the graph Laplacian [2], [13], [14]. In the second approach, which is built on *algebraic signal processing*, the GFT is introduced based on the graph adjacency matrix [1], [3]–[11], [13]. This approach defines the GFT as a change of basis into the eigenvectors of the adjacency matrix. Since the GFT is defined based on the graph adjacency matrix, which is analogous to the discrete shift

matrix, it reduces to the discrete-time Fourier transform matrix for circulant adjacency matrices, [4], [5], [7], [9], [10]. The graph adjacency matrix based GFT has also been extended to the graph fractional Fourier transform (GFRT) in [8], [11], [19].

An important issue in frequency analyses of graph signals is to develop an ordering among graph frequencies. The GFT matrix is used to define graph frequencies in [2], [4], [5], [7], [9], [10]. One approach for developing an ordering among graph frequencies is to use total variation based methods [5], [10]. In [5], it is proven that the ordering based on the total variation is unique for graph adjacency matrices with real spectra (all eigenvalues are real), but may not be unique for graphs with complex spectra (some eigenvalues are not real).

The  $a$ th order fractional Fourier transform (FRT) is the  $a$ th operator power of the ordinary Fourier transform (FT). It generalizes the FT to fractional orders  $a$ , with  $a = 0$  corresponding to the identity operation and  $a = 1$  to the ordinary FT [31]–[38]. As an operator power, it satisfies index additivity: The  $a$ th transform of the  $b$ th transform is the  $(a + b)$ th transform. The ordinary FT transforms a signal from the time or space domain to the temporal or spatial frequency domain. The FRT, on the other hand, transforms the signal to a fractional Fourier domain, which can be viewed as intermediate domains. If one considers the time-frequency or space-frequency plane (such as that in which the Wigner distribution lives), where the horizontal axis is time or space and the vertical axis is frequency, the  $a$ th order fractional Fourier domain is an axis making angle  $a\pi/2$  with the horizontal axis. Thus, the FRT leads to a generalization of the concept of the *frequency domain*. Further details can be found in [34], [37], [39]. The FRT has many applications in several areas of signal and image processing [40]–[47], including time/space-frequency representations, image and video processing, signal reconstruction, pattern recognition, radar signal processing, and beamforming [36], [39], [48]–[62]. FRT has also been widely used in wave and beam propagation, diffraction, optics, and optical signal processing, among several other applications [31], [63], [64].

Our main motivation can be stated from two alternative perspectives. First, we generalize the optimal Wiener filtering problem on graphs to fractional Fourier domains. This is a generalization of the corresponding problem in ordinary Fourier domains, and has the potential to offer improved performance in the sense of lower errors, thanks to the additional degree of freedom in choosing the fractional Fourier transform order. The ordinary (non-fractional) version of the optimal filtering problem has been addressed in [7] for the graph vertex-domain under the assumption that the filter is shift-invariant. The problem considered in [7] corresponds to a special case of the problem proposed in this manuscript when the fractional order  $a$  is equal to one. In this work, it is shown that the minimum MSE is not always attained at  $a = 0$  or  $a = 1$ ; hence, employing the fractional Fourier transform can offer improved performance. In other words, we propose a generalized version of the problem in [7] and provide its analytical solution for any order  $a$ . Our development is valid for all kinds of weighted/unweighted and directed/undirected graph structures, and for all graph signals that can be defined on them. Neither on the graph structure nor

on the graph signals, did we make any restrictive assumptions. Therefore, the proposed GFRT filtering framework can be applied under a wide variety of arbitrary conditions.

The second perspective is that we generalize the optimal fractional Fourier domain filtering problem to graph signal processing. The optimal fractional Fourier domain filtering problem has earlier been solved in a conventional signal processing context [40], [65]. Here we focus on the graph signal processing version of the problem. In any event, since in our generalized method, both the time (space) and frequency domains are included as special cases, consideration of fractional domains will never degrade the filtering performance and will most likely improve it.

The rest of the manuscript is organized as follows: In Section II, we present preliminaries for both GSP and FRT. Section III introduces the proposed optimal filtering approach in the fractional Fourier domain for graph signals and provides theoretical derivations. Section IV discusses computational cost issues. Numerical examples are given in Section V, followed by the concluding remarks in Section VI.

## II. BACKGROUND INFORMATION

In this section, some preliminaries for graph signals, graph filters, GFT, fractional FT (FRT) and graph fractional Fourier transform (GFRT) are provided.

### A. Graph Signals

Let  $\mathcal{G} = \{\mathcal{V}, \mathbf{A}\}$  be a graph where  $\mathcal{V} = \{v_0, v_1, \dots, v_{N-1}\}$  is the set of vertices and  $\mathbf{A}$  is the adjacency matrix of the graph  $\mathcal{G}$ . The adjacency matrix  $\mathbf{A} \in \mathbb{C}^{N \times N}$  stands for the relationships between the vertices. In other words, if there is no relation between the vertices  $v_i$  and  $v_j$ ,  $\mathbf{A}(i, j) = 0$ , otherwise  $\mathbf{A}(i, j) \neq 0$ , where  $\mathbf{A}(i, j)$  is the  $(i, j)$ th element of  $\mathbf{A}$ . The graph signal  $\mathbf{x} = [\mathbf{x}[0] \ \mathbf{x}[1] \ \dots \ \mathbf{x}[N-1]]^\top \in \mathbb{C}^N$  is a mapping from the set of vertices,  $\mathcal{V}$  to  $\mathbb{C}^N$ , i.e., each vertex is mapped to a complex number [7], [9], [10], [13] as follows:

$$\begin{aligned} \mathbf{x} : \mathcal{V} &\rightarrow \mathbb{C}^N \\ v_n &\rightarrow \mathbf{x}[n] \end{aligned}$$

### B. Graph Filters

A graph filter can be defined as any matrix  $\mathbf{H} \in \mathbb{C}^{N \times N}$ . The filtering is the process of multiplying a matrix  $\mathbf{H}$  with the graph signal  $\mathbf{x}$ , which is a linear operation. Before going into details, we define the graph shift operator and the shift-invariance property of a graph filter. In discrete signal processing, the basic building block for filters is given by the time shift, which is denoted by  $z^{-1}$  in the  $z$ -transform domain [3].

Consider the adjacency matrix  $\mathbf{A}$  of the simple graph given by Fig. 1, which can be written as

$$\mathbf{A}(i, j) = \begin{cases} 1, & \text{if } i - j \equiv 1 \pmod{N} \\ 0, & \text{otherwise} \end{cases}.$$

Note that  $\mathbf{A}\mathbf{x} = [\mathbf{x}[N-1] \ \mathbf{x}[0] \ \dots \ \mathbf{x}[N-2]]^\top$ , which is a circularly shifted version of the original signal  $\mathbf{x}$ . In other words,

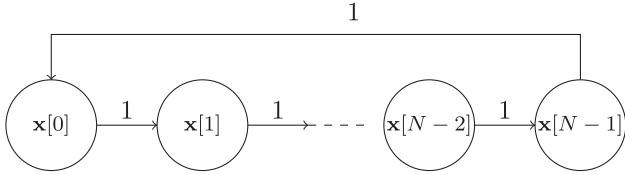


Fig. 1. Graph representation of periodic data.

the result of the unit time shift operation for the signal  $\mathbf{x}$  can be written as  $\mathbf{Ax}$ . In order for a graph filter  $\mathbf{H}$  to be considered LSI, it is required that  $\mathbf{AHx} = \mathbf{HAx}$  for any  $\mathbf{x} \in \mathbb{C}^N$ , where  $\mathbf{A}$  is the adjacency matrix. In other words, linear shift-invariance means that filtering and shifting commute with each other for any signal  $\mathbf{x}$ . From this definition, it follows that  $\mathbf{AH} = \mathbf{HA}$  when  $\mathbf{H}$  is an LSI filter.

In [1], it is proven that any LSI filter  $\mathbf{H}$  can be represented as a polynomial of the adjacency matrix  $\mathbf{A}$ , if the characteristic polynomial and the minimal polynomial of  $\mathbf{A}$  are the same. That is, there exist  $L \in \mathbb{N}$  and coefficients  $\{h_k\}_{k=0}^{L-1}$  such that

$$\mathbf{H} = \sum_{k=0}^{L-1} h_k \mathbf{A}^k \quad (1)$$

In the preceding discussion, instead of the adjacency matrix  $\mathbf{A}$ , one can substitute any alternative definition of a shift operator on graphs. For example,  $\frac{1}{\lambda_{\max}(\mathbf{A})} \mathbf{A}$  is the *normalized graph shift matrix* introduced in [5], where  $\lambda_{\max}(\mathbf{A})$  denotes the largest eigenvalue of  $\mathbf{A}$ .

If the adjacency matrix  $\mathbf{A}$  is assumed to be diagonalizable and its eigendecomposition is given by  $\mathbf{A} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^{-1}$ , a new graph shift operator is introduced in [7] as  $\mathbf{A}_\phi = \mathbf{V}\mathbf{\Lambda}_\phi\mathbf{V}^{-1}$ , where  $\mathbf{\Lambda}_\phi = \text{diag}(e^{j\phi_0}, e^{j\phi_1}, \dots, e^{j\phi_{N-1}})$  and  $\phi_k$  is an arbitrary phase in  $[0, 2\pi]$  such that  $\phi_k \neq \phi_l$  for  $k \neq l$ . The reason why such a shift operator is introduced is due to its energy preserving property, which will be clear in the next subsection.

If the considered graph shift operator is  $\mathbf{A}_\phi$ , the definition of an LSI filter is modified by replacing  $\mathbf{A}_\phi$  with  $\mathbf{A}$ . In other words, if  $\mathbf{H}$  is an LSI filter under  $\mathbf{A}_\phi$ , then  $\mathbf{H}$  should commute with  $\mathbf{A}_\phi$ . Moreover, by Theorem 1 in [1], if  $\mathbf{H}$  is an LSI filter under  $\mathbf{A}_\phi$ , then it should be a polynomial of  $\mathbf{A}_\phi$ , since all the eigenvalues of  $\mathbf{A}_\phi$  are different.

It is evident that  $\mathbf{A}$  commutes with  $\mathbf{A}_\phi$ , hence  $\mathbf{A}$  is an LSI filter under  $\mathbf{A}_\phi$  and  $\mathbf{A}_\phi$  is an LSI filter under  $\mathbf{A}$ . Since all the eigenvalues of  $\mathbf{A}_\phi$  are different, Theorem 1 in [1] implies that  $\mathbf{A}$  can be expressed as a polynomial of  $\mathbf{A}_\phi$ .

### C. Graph Fourier Transform (GFT)

It is known that under some conditions, an LSI filter for a graph signal can be represented as in (1), and that a matrix and its powers share the same set of eigenvectors. Therefore, the LSI filter  $\mathbf{H}$  shares a set of eigenvectors with the adjacency matrix  $\mathbf{A}$ . We can write the Jordan decomposition of  $\mathbf{A}$  as  $\mathbf{A} = \mathbf{V}\mathbf{J}_A\mathbf{V}^{-1}$ , where  $\mathbf{J}_A$  is the Jordan block form matrix, and  $\mathbf{V} = [\mathbf{v}_0 \ \mathbf{v}_1 \ \dots \ \mathbf{v}_{N-1}]$  contains the generalized eigenvectors of  $\mathbf{A}$  in its columns. Hence, the columns of  $\mathbf{V}$  provide a basis for the eigenvectors of an LSI filter, which is the graph signal

processing counterpart of the well-known result stating that complex exponentials are the eigenfunctions of (conventional) LTI systems.

In conventional signal processing, as complex exponentials form an orthonormal basis, one can express any signal in terms of the complex exponential basis, and the expansion coefficients constitute the Fourier representation of the signal. Similarly, for the graph signal processing case, one can write any signal  $\mathbf{x}$  in terms of the columns  $\mathbf{v}_i$  of  $\mathbf{V}$  as follows:

$$\begin{aligned} \mathbf{x} &= \sum_{i=0}^{N-1} \hat{x}_i \mathbf{v}_i \\ &= \mathbf{V}[\hat{x}_0 \ \hat{x}_1 \ \dots \ \hat{x}_{N-1}]^T = \mathbf{V}\hat{\mathbf{x}}. \end{aligned}$$

If  $\{\lambda_i\}_{i=0}^{N-1}$  denotes the set of eigenvalues of  $\mathbf{A}$ , we can view  $\lambda_i$ 's as frequencies and  $\hat{x}_i$ 's as the Fourier coefficients of  $\mathbf{x}$  corresponding to  $\lambda_i$ 's. Therefore,  $\hat{\mathbf{x}}$  can be defined as the graph Fourier transform of  $\mathbf{x}$ , which is given by

$$\hat{\mathbf{x}} \triangleq \mathcal{F}\mathbf{x} = \mathbf{V}^{-1}\mathbf{x}.$$

However, for an arbitrary graph, one should note that the GFT may not be unitary; i.e., Parseval's theorem does not hold unless  $\mathbf{V}^{-1} = \mathbf{V}^H$ , where  $(\cdot)^H$  denotes the conjugate transpose of its argument. From [7], it is known that  $\|\mathbf{V}^{-1}\mathbf{x}\|_2^2$  is bounded as

$$\alpha \|\mathbf{x}\|_2^2 \leq \|\mathbf{V}^{-1}\mathbf{x}\|_2^2 \leq \beta \|\mathbf{x}\|_2^2$$

with  $\alpha = 1/\|\mathbf{V}^{-1}\|_2^2$  and  $\beta = \|\mathbf{V}^{-1}\|_2^2$ , where  $\|\cdot\|_2$  and  $\|\cdot\|$  denote the  $l_2$  norm and the spectral norm, respectively. Moreover, in [7], it is proven that  $\|\mathcal{F}\mathbf{x}\|_2 = \|\mathcal{F}(\mathbf{A}_\phi^k \mathbf{x})\|_2$  for any integer  $k \geq 0$ . That is, the energy of the graph signal  $\mathbf{x}$  in the Fourier domain remains the same for any amount of shift  $k$ . This property is not satisfied if neither  $\mathbf{A}$  nor  $\frac{1}{\lambda_{\max}(\mathbf{A})} \mathbf{A}$  is used as the graph shift operator as the magnitudes of the eigenvalues of  $\mathbf{A}$  and  $\frac{1}{\lambda_{\max}(\mathbf{A})} \mathbf{A}$  do not have to be equal to 1.

### D. Fractional Fourier Transform (FRT)

Fractional Fourier transform (FRT) is a generalization to the Fourier transform (FT). The  $a$ th order FRT  $\mathcal{F}^a$  of a function  $f(u)$  is defined for  $0 < |a| < 2$  as follows [34]:

$$\begin{aligned} \mathcal{F}^a f(u) &= \int_{-\infty}^{\infty} K_a(u, u') f(u') du', \\ K_a(u, u') &= A_\theta \exp [i\pi(u^2 \cot \theta - 2uu' \csc \theta + u'^2 \cot \theta)], \\ A_\theta &= \frac{\exp(-i\pi \text{sgn}(\sin \theta)/4 + i\theta/2)}{|\sin \theta|^{1/2}} \text{ and } \theta = \frac{a\pi}{2}. \end{aligned}$$

When  $a = 1$ , FRT reduces to FT. More details can be found in [34].

### E. Graph Fractional Fourier Transform (GFRT)

The definition of FRT for graph signals is introduced in [8], [11]. Let the Jordan decomposition of the GFT matrix  $\mathbf{V}^{-1}$  be as follows:

$$\mathbf{V}^{-1} = \mathbf{P}\mathbf{J}\mathbf{P}^{-1}.$$

Then, the definition of the  $a$ th order GFRT matrix is given by

$$\mathbf{F}^a \triangleq \mathbf{P}\mathbf{J}^a\mathbf{P}^{-1}.$$

The  $a$ th power of  $\mathbf{J}$  can be computed as follows: For the  $j$ th Jordan block, the  $a$ th power of the  $k$ th super-diagonal is  $\prod_{i=1}^k \frac{a+1-i}{i} \lambda_j^{a-k}$ . The details of this computation can be found in [8]. Then, the definition of the  $a$ th order GFRT is given by

$$\hat{\mathbf{x}} \triangleq \mathcal{F}^a x = \mathbf{F}^a \mathbf{x}.$$

Then, the inverse GFRT is specified by the matrix  $\mathbf{F}^{-a}$ . Clearly, when  $a = 0$  and  $a = 1$ , the GFRT reduces to  $\mathbf{I}$  (the identity matrix) and  $\mathbf{F}$  (the non-fractional transform), respectively. It is also easy to verify that the index additivity property holds, i.e.,  $\mathbf{F}^a \mathbf{F}^b = \mathbf{F}^{a+b}$  for any  $a, b$ .

Thus, we have established the fractional Fourier transform for graph signals that will be employed in the following section.

### III. OPTIMAL FILTERING IN FRACTIONAL FOURIER DOMAINS FOR GRAPH SIGNALS

To formulate the proposed problem, the following signal observation model is used:

$$\mathbf{y} = \mathbf{G}\mathbf{x} + \mathbf{n} \quad (2)$$

where  $\mathbf{G}$  is a known matrix in  $\mathbb{C}^{N \times N}$  and  $\mathbf{x}$  is a stochastic graph signal defined over  $\mathcal{G} = \{\mathcal{V}, \mathbf{A}\}$  with given parameters  $\mathbb{E}\{\mathbf{x}\}$ ,  $\mathbb{E}\{\mathbf{x}^H \mathbf{x}\}$ , and  $\mathbb{E}\{\mathbf{x}\mathbf{x}^H\}$ . Moreover,  $\mathbf{n}$  is the additive noise term with the following statistical parameters assumed to be known:  $\mathbb{E}\{\mathbf{n}\}$ ,  $\mathbb{E}\{\mathbf{n}\mathbf{n}^H\}$ ,  $\mathbb{E}\{\mathbf{n}^H \mathbf{n}\}$ ,  $\mathbb{E}\{\mathbf{x}\mathbf{n}^H\}$ ,  $\mathbb{E}\{\mathbf{x}^H \mathbf{n}\}$ .

Before presenting the proposed problem, we first start with the optimal Wiener filtering problem for graph signal processing. If  $\mathbf{x}$  is a graph signal to be estimated and  $\mathbf{y}$  is the observed graph signal, the optimal Wiener filtering problem for graph signal processing can be formulated as

$$\min_{\mathbf{R}} \mathbb{E}\{\|\mathbf{R}\mathbf{y} - \mathbf{x}\|_2^2\} \quad (3)$$

This problem is exactly in the same form as the classical Wiener filtering problem. Assuming  $\mathbb{E}\{\mathbf{y}\mathbf{y}^H\}$  is invertible, the solution of (3) is given by [66]:

$$\arg \min_{\mathbf{R} \in \mathbb{C}^{N \times N}} \mathbb{E}\{\|\mathbf{R}\mathbf{y} - \mathbf{x}\|_2^2\} = \mathbb{E}\{\mathbf{x}\mathbf{y}^H\} \left( \mathbb{E}\{\mathbf{y}\mathbf{y}^H\} \right)^{-1} \quad (4)$$

In this manuscript, our aim is to find the optimal filter in the fractional Fourier domain for graph signals, which minimizes the MSE while recovering the original signal  $\mathbf{x}$ . This is the graph signal processing counterpart of the problem discussed in [40].

The aim is to recover  $\mathbf{x}$  in (2) with the minimum mean error in the  $l_2$  sense (i.e., with the minimum MSE). If the reconstructed version of  $\mathbf{x}$  is  $\tilde{\mathbf{x}}$ , which is a function of  $\mathbf{y}$ , the corresponding MSE is given by  $\mathbb{E}\{\|\tilde{\mathbf{x}} - \mathbf{x}\|_2^2\}$ . We consider filtering in the  $a$ th order fractional Fourier domain, so that the estimated signal  $\tilde{\mathbf{x}}$  is expressed as

$$\tilde{\mathbf{x}} = \mathbf{F}^{-a} \mathbf{H} \mathbf{F}^a \mathbf{y}. \quad (5)$$

In other words, the observed signal  $\mathbf{y}$  is fractional Fourier transformed, multiplicatively filtered with a ‘‘transfer function,’’ and then inverse fractional Fourier transformed to the original

domain. Therefore, our aim is to design the matrix  $\mathbf{H}$  to solve the following problem:

$$\min_{\mathbf{H}} \mathbb{E}\{\|\mathbf{F}^{-a} \mathbf{H} \mathbf{F}^a \mathbf{y} - \mathbf{x}\|_2^2\} \quad (6)$$

for any  $a$ . Then, we will search over possible values of  $a$  in order to find  $a^*$  that yields the minimum MSE.

The following lemma focuses on the unconstrained version of the problem in (6).

*Lemma 1:* If there is no restriction on  $\mathbf{H}$  in (6), the problem yields the same MSE for any  $a$ .

*Proof:* Consider any real numbers  $a, b$  such that  $a \neq b$  and define

$$\mathbf{H}^{(a)} = \arg \min_{\mathbf{H}} \mathbb{E}\{\|\mathbf{F}^{-a} \mathbf{H} \mathbf{F}^a \mathbf{y} - \mathbf{x}\|_2^2\} \quad (7)$$

$$\mathbf{H}^{(b)} = \arg \min_{\mathbf{H}} \mathbb{E}\{\|\mathbf{F}^{-b} \mathbf{H} \mathbf{F}^b \mathbf{y} - \mathbf{x}\|_2^2\} \quad (8)$$

Note that for any choice of  $\mathbf{H}$ , we know by (7) that

$$\mathbb{E}\{\|\mathbf{F}^{-a} \mathbf{H}^{(a)} \mathbf{F}^a \mathbf{y} - \mathbf{x}\|_2^2\} \leq \mathbb{E}\{\|\mathbf{F}^{-a} \mathbf{H} \mathbf{F}^a \mathbf{y} - \mathbf{x}\|_2^2\}. \quad (9)$$

Then, taking  $\mathbf{H} = \mathbf{F}^a \mathbf{F}^{-b} \mathbf{H}^{(b)} \mathbf{F}^b \mathbf{F}^{-a}$ , we obtain via (9) that

$$\mathbb{E}\{\|\mathbf{F}^{-a} \mathbf{H}^{(a)} \mathbf{F}^a \mathbf{y} - \mathbf{x}\|_2^2\} \leq \mathbb{E}\{\|\mathbf{F}^{-b} \mathbf{H}^{(b)} \mathbf{F}^b \mathbf{y} - \mathbf{x}\|_2^2\}. \quad (10)$$

Similarly, we can easily show that

$$\mathbb{E}\{\|\mathbf{F}^{-b} \mathbf{H}^{(b)} \mathbf{F}^b \mathbf{y} - \mathbf{x}\|_2^2\} \leq \mathbb{E}\{\|\mathbf{F}^{-a} \mathbf{H}^{(a)} \mathbf{F}^a \mathbf{y} - \mathbf{x}\|_2^2\}. \quad (11)$$

From (7), (8), (10), and (11), it can be concluded that the same minimum value is achieved for any  $a$  and  $b$ . Hence, the statement in the lemma is proven. ■

The result in Lemma 1 is quite intuitive since we can view  $\mathbf{F}^{-a}$  and  $\mathbf{F}^{-b}$  as similarity transformation matrices.

In [40], optimal fractional Fourier domain filters for both continuous-time and discrete-time signals are derived. This amounts to first taking the fractional Fourier transform of the given signal, applying a multiplicative filter with a certain ‘‘transfer function’’ or ‘‘fractional frequency response’’ in the fractional Fourier domain, and finally inverse fractional Fourier transforming to return to the original signal domain. In a discrete-time formulation, application of a multiplicative filter or ‘‘transfer function’’ can be modeled as multiplication by a diagonal matrix. Hence, we will force  $\mathbf{H}$  in (6) to be diagonal. To that aim, we define a set as  $\mathcal{D} \triangleq \{\mathbf{H} \mid \mathbf{H} \in \mathbb{C}^{N \times N} \text{ and } \mathbf{H} \text{ is diagonal}\}$ . Then, our aim is to solve the following optimization problem:

$$\min_{\mathbf{H}} \mathbb{E}\{\|\mathbf{F}^{-a} \mathbf{H} \mathbf{F}^a \mathbf{y} - \mathbf{x}\|_2^2\} \quad (12a)$$

$$\text{subject to } \mathbf{H} \in \mathcal{D} \quad (12b)$$

In [7], (3) is solved under the assumptions that  $\mathbf{A}$  is diagonalizable and  $\mathbf{R}$  is shift-invariant, i.e.,  $\mathbf{R}$  is the polynomial of the shift operator. More specifically, the following problem is considered in [7]:

$$\min_{\mathbf{H}} \mathbb{E}\{\|\mathbf{H}\mathbf{y} - \mathbf{x}\|_2^2\} \quad (13a)$$

$$\text{subject to } \mathbf{H} \text{ is LSI and } \mathbf{A} \text{ is diagonalizable.} \quad (13b)$$

*Proposition 1:* The optimization problems given by (12) and (13) are equivalent to each other for  $a = 1$  and diagonalizable adjacency matrix  $\mathbf{A}$ .

*Proof:* It is known that when  $\mathbf{H}$  is shift-invariant, it can be represented as a polynomial of the shift operator  $\mathbf{A}_\phi$ . Hence, there exist  $L \in \mathbb{N}$  and  $\{a_k\}_{k=0}^{L-1} \in \mathbb{R}$  such that

$$\mathbf{H} = \sum_{k=0}^{L-1} a_k \mathbf{A}_\phi^k$$

Since  $\mathbf{A} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^{-1}$ , we can write  $\mathbf{A}_\phi^k = \mathbf{V}\mathbf{\Lambda}_\phi^k\mathbf{V}^{-1}$ , where  $\mathbf{V}$  contains the eigenvectors of  $\mathbf{A}$ . Therefore, one can say that any LSI filter  $\mathbf{H}$  can be represented as

$$\mathbf{H} = \mathbf{V} \underbrace{\left( \sum_{k=0}^{L-1} a_k \mathbf{\Lambda}_\phi^k \right)}_{\text{diagonal}} \mathbf{V}^{-1}$$

Since  $\mathbf{V}^{-1} = \mathbf{F}^1$ , (13) can be transformed into the following optimization problem:

$$\min_{\mathbf{Q}} \mathbb{E}\{\|\mathbf{F}^{-1}\mathbf{Q}\mathbf{F}\mathbf{y} - \mathbf{x}\|_2^2\} \quad (14a)$$

$$\text{subject to} \quad \mathbf{Q} \in \hat{\mathcal{D}}$$

where  $\hat{\mathcal{D}} = \{\mathbf{Q} \mid \exists L \in \mathbb{N}, \{a_k\}_{k=0}^{L-1} \in \mathbb{R} \text{ such that } \mathbf{Q} = \sum_{k=0}^{L-1} a_k \mathbf{\Lambda}_\phi^k\}$ .

If we can show that  $\hat{\mathcal{D}} = \mathcal{D}$ , then the proof is completed. Since any  $\mathbf{Q} \in \hat{\mathcal{D}}$  is diagonal, it is clear that  $\hat{\mathcal{D}} \subseteq \mathcal{D}$ . Hence, the aim becomes showing that  $\mathcal{D} \subseteq \hat{\mathcal{D}}$ . To that aim, take any arbitrary diagonal matrix  $\mathbf{X} = \text{diag}(x_0, x_1, \dots, x_{N-1}) \in \mathcal{D}$  and consider the following matrix equation:

$$\underbrace{\begin{bmatrix} 1 & e^{j\phi_0} & \dots & e^{j(N-1)\phi_0} \\ 1 & e^{j\phi_1} & \dots & e^{j(N-1)\phi_1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & e^{j\phi_{N-1}} & \dots & e^{j(N-1)\phi_{N-1}} \end{bmatrix}}_{\mathbf{Y}} \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_{N-1} \end{bmatrix} = \begin{bmatrix} x_0 \\ x_1 \\ \vdots \\ x_{N-1} \end{bmatrix} \quad (15)$$

Another way of writing (15) is that

$$\sum_{k=0}^{N-1} a_k \mathbf{\Lambda}_\phi^k = \mathbf{X}$$

Note that as  $\mathbf{Y}$  is a Vandermonde matrix, (15) is solvable. Since (15) is solvable, one can claim that  $\mathbf{X} \in \hat{\mathcal{D}}$ . Therefore, we must have  $\mathcal{D} \subseteq \hat{\mathcal{D}}$ . Therefore,  $\hat{\mathcal{D}} = \mathcal{D}$  and the proof is completed. ■

*Remark 1:* By Proposition 1, one can infer that the proposed problem in (12) is the generalized version of (13). In other words, by solving (12) we can achieve lower MSE values than (13) for some  $0 < a < 1$ .

*Remark 2:* As a consequence of Proposition 1, we conclude that for graph signal processing, filtering via an LSI filter corresponds to multiplication in the vertex-frequency domain. This result matches with the classical signal processing theory since filtering with an LTI filter corresponds to a multiplicative filter in the Fourier domain for both continuous-time and discrete-time signals.

Instead of working with matrices in  $\mathcal{D}$ , we can transform the proposed problem in (12) into another problem where the

parameter of interest is simply a vector in  $\mathbb{C}^N$  as the number of nonzero elements of  $\mathbf{H}$  is at most  $N$ . To that end, let  $\mathbf{F}^{-a} = [\mathbf{w}_1 \mathbf{w}_2 \dots \mathbf{w}_N]$ ,  $(\mathbf{F}^a)^\top = [\tilde{\mathbf{w}}_1 \tilde{\mathbf{w}}_2 \dots \tilde{\mathbf{w}}_N]$ , and  $\mathbf{W}_i = \mathbf{w}_i \tilde{\mathbf{w}}_i^\top$  for  $i = 1, 2, \dots, N$ . Then, for  $\mathbf{H} = \text{diag}(h_1, h_2, \dots, h_N)$ , we can write

$$\mathbf{F}^{-a} \mathbf{H} \mathbf{F}^a = \sum_{i=1}^N h_i \mathbf{W}_i.$$

That is, the objective is a function of  $\mathbf{h} = [h_1 \ h_2 \ \dots \ h_N]^\top$ , and (12) can be transformed into the following problem:

$$\min_{\mathbf{h}} \mathbb{E} \left\{ \left\| \sum_{i=1}^N h_i \mathbf{W}_i \mathbf{y} - \mathbf{x} \right\|_2^2 \right\} \quad (16)$$

where we do not impose any restrictions on  $\mathbf{h}$ .

*Proposition 2:* (16) is a convex optimization problem.

*Proof:* Since we do not have any restriction on  $\mathbf{h}$ , the domain of the problem is  $\mathbb{C}^N$ , which is a convex set. Therefore, we only need to prove that the objective function in (16) is a convex function of  $\mathbf{h}$ . The convexity follows from the linearity of the expectation operation and the convexity of the norm. For the sake of completeness we provide the proof in Appendix A. ■

As a consequence of Proposition 2, by simply taking the derivative of the objective function in (16), we can find the optimal solution. The following proposition characterizes solution of the optimal filtering problem in (16).

*Proposition 3:* For the problem in (16), the optimal filter coefficients  $\mathbf{h}^{(\text{opt})} = [h_1^{\text{opt}} \ h_2^{\text{opt}} \ \dots \ h_N^{\text{opt}}]^\top$  are obtained from the following matrix equation:

$$\mathbf{T} \mathbf{h}^{(\text{opt})} = \mathbf{q} \quad (17)$$

where  $\mathbf{q} = \mathbb{E}\{\mathbf{S}^\text{H} \mathbf{x}\}$ ,  $\mathbf{T} = \mathbb{E}\{\mathbf{S}^\text{H} \mathbf{S}\}$ ,  $\mathbf{S} = [\mathbf{S}_1 \ \mathbf{S}_2 \ \dots \ \mathbf{S}_N]$  and  $\mathbf{S}_k = \mathbf{W}_k \mathbf{y}$  for any  $1 \leq k \leq N$ .

*Proof:* Let  $f(\mathbf{h}) \triangleq \mathbb{E}\{\|\sum_{i=1}^N h_i \mathbf{W}_i \mathbf{y} - \mathbf{x}\|_2^2\}$ . Thus,

$$\begin{aligned} f(\mathbf{h}) &= \mathbb{E} \left\{ \left( \sum_{i=1}^N h_i^* \mathbf{y}^\text{H} \mathbf{W}_i^\text{H} - \mathbf{x}^\text{H} \right) \left( \sum_{j=1}^N h_j \mathbf{W}_j \mathbf{y} - \mathbf{x} \right) \right\} \\ &= \sum_{i=1}^N \sum_{j=1}^N h_i^* h_j \mathbb{E} \left\{ \mathbf{y}^\text{H} \mathbf{W}_i^\text{H} \mathbf{W}_j \mathbf{y} \right\} - \sum_{j=1}^N h_j \mathbb{E} \left\{ \mathbf{x}^\text{H} \mathbf{W}_j \mathbf{y} \right\} \\ &\quad - \sum_{i=1}^N h_i^* \mathbb{E} \left\{ \mathbf{y}^\text{H} \mathbf{W}_i^\text{H} \mathbf{x} \right\} + \mathbb{E} \left\{ \mathbf{x}^\text{H} \mathbf{x} \right\}. \end{aligned}$$

By taking the derivative of  $f(\mathbf{h})$  with respect to the real and imaginary parts of  $\mathbf{h}$ , we can obtain the optimal solution. To that aim, for any  $1 \leq k \leq N$ , let  $h_k$  be given by  $h_k = h_{k,r} + j h_{k,im}$  where  $h_{k,r}$  and  $h_{k,im}$  denote the real and imaginary parts of  $h_k$ , respectively. Then the following equations must hold:

$$\begin{aligned} \frac{\partial f(\mathbf{h})}{\partial h_{k,r}} \Big|_{\mathbf{h}=\mathbf{h}^{\text{opt}}} &= \sum_{j=1}^N \left( h_j^{\text{opt}} \mathbb{E} \left\{ \mathbf{y}^\text{H} \mathbf{W}_k^\text{H} \mathbf{W}_j \mathbf{y} \right\} \right. \\ &\quad \left. + (h_j^{\text{opt}})^* \mathbb{E} \left\{ \mathbf{y}^\text{H} \mathbf{W}_j^\text{H} \mathbf{W}_k \mathbf{y} \right\} \right) - \mathbb{E} \left\{ \mathbf{x}^\text{H} \mathbf{W}_k \mathbf{y} + \mathbf{y}^\text{H} \mathbf{W}_k^\text{H} \mathbf{x} \right\} = 0. \end{aligned} \quad (18)$$

$$\begin{aligned} \frac{\partial f(\mathbf{h})}{\partial h_{k,im}} \Big|_{\mathbf{h}=\mathbf{h}^{\text{opt}}} &= \sum_{j=1}^N \left( -h_j^{\text{opt}} \mathbb{E} \left\{ \mathbf{y}^H \mathbf{W}_k^H \mathbf{W}_j \mathbf{y} \right\} \right. \\ &+ \left. (h_j^{\text{opt}})^* \mathbb{E} \left\{ \mathbf{y}^H \mathbf{W}_j^H \mathbf{W}_k \mathbf{y} \right\} \right) - \mathbb{E} \left\{ \mathbf{x}^H \mathbf{W}_k \mathbf{y} - \mathbf{y}^H \mathbf{W}_k^H \mathbf{x} \right\} = 0. \end{aligned} \quad (19)$$

By combining (18) and (19), for any  $1 \leq k \leq N$ , we obtain

$$\sum_{j=1}^N h_j^{\text{opt}} \mathbb{E} \left\{ \mathbf{y}^H \mathbf{W}_k^H \mathbf{W}_j \mathbf{y} \right\} = \mathbb{E} \left\{ \mathbf{y}^H \mathbf{W}_k^H \mathbf{x} \right\} \quad (20)$$

Then, for any  $1 \leq k \leq N$ , (20) can be rewritten as

$$\sum_{j=1}^N h_j^{\text{opt}} \mathbb{E} \left\{ \mathbf{S}_k^H \mathbf{S}_j \right\} = \mathbb{E} \left\{ \mathbf{S}_k^H \mathbf{x} \right\}. \quad (21)$$

which is exactly the statement in the proposition.  $\blacksquare$

*Remark 3:* One should note that via some algebraic manipulation, the  $(i, j)$ th entry of  $\mathbf{T}$  and the  $i$ th entry of  $\mathbf{q}$  can be written as

$$\begin{aligned} \mathbf{T}(i, j) &= \text{Tr} \left\{ \mathbf{G}^H \mathbf{W}_i^H \mathbf{W}_j \mathbf{G} \mathbb{E} \{ \mathbf{x} \mathbf{x}^H \} \right\} + \text{Tr} \left\{ \mathbf{G}^H \mathbf{W}_i^H \mathbf{W}_j \mathbb{E} \{ \mathbf{n} \mathbf{x}^H \} \right\} \\ &+ \text{Tr} \left\{ \mathbf{W}_i^H \mathbf{W}_j \mathbf{G} \mathbb{E} \{ \mathbf{x} \mathbf{n}^H \} \right\} + \text{Tr} \left\{ \mathbf{W}_i^H \mathbf{W}_j \mathbb{E} \{ \mathbf{n} \mathbf{n}^H \} \right\}, \\ \mathbf{q}(i) &= \text{Tr} \left\{ \mathbf{G}^H \mathbf{W}_i \mathbb{E} \{ \mathbf{x} \mathbf{x}^H \} \right\} + \text{Tr} \left\{ \mathbf{W}_i^H \mathbb{E} \{ \mathbf{x} \mathbf{n}^H \} \right\}. \end{aligned}$$

Since  $\mathbf{G}$ ,  $\mathbb{E} \{ \mathbf{x} \mathbf{x}^H \}$ ,  $\mathbb{E} \{ \mathbf{n} \mathbf{x}^H \}$ ,  $\mathbb{E} \{ \mathbf{x} \mathbf{n}^H \}$ ,  $\mathbb{E} \{ \mathbf{n} \mathbf{n}^H \}$ ,  $\{ \mathbf{W}_i \}_{i=1}^N$  are all known parameters, each entry of  $\mathbf{T}$  and  $\mathbf{q}$  is completely known. Entries of  $\mathbf{q}$  are not affected by the auto-correlation of the noise terms, whereas the entries of  $\mathbf{T}$  are affected by all the system parameters. The optimal filter coefficients  $\mathbf{h}^{\text{opt}}$  can be derived from the linear equation given by (17) based on  $\mathbf{G}$ ,  $\{ \mathbf{W}_i \}_{i=1}^N$ , and the statistical parameters related to  $\mathbf{x}$  and  $\mathbf{n}$ .

Proposition 3 provides an analytical solution to the problem of optimal filtering in fractional Fourier domains for graph signals. It should also be noted that, in the analysis above, we do not make any assumptions on the adjacency matrix  $\mathbf{A}$ . In other words, our analysis is valid for any kind of graph structures including directed/undirected or weighted/unweighted graphs.

#### IV. COMPUTATIONAL COST

The detailed analysis of the computational cost of solving (12), in other words, determining the optimal filter coefficients  $\mathbf{h}^{\text{opt}}$  is given in Appendix B. It is shown that for any GFRT order, the number of required operations is  $O(N^4)$ . For the GFRT order, the interval  $[0, 1]$  is discretized with  $N_A$  different elements. For any GFRT order, the corresponding optimal filter coefficients are found and among  $N_A$  different vectors of filter coefficients, the one which yields the lowest MSE is chosen. Thus, the resulting computational complexity is  $O(N^4 N_A)$ . However, these operations have to be performed only once for a particular graph structure, since the filter depends on the graph and signal statistics, not on the particular realization of the signal. Hence, in the following, we concentrate on the computational

cost of filtering in the fractional Fourier domain which should not be confused with the cost of computation of optimal filter coefficients.

The computational cost of filtering in the  $a$ th order fractional Fourier domain as given in (5) can be analyzed as follows. In the conventional version, fractional Fourier domain filtering corresponds to two instances of computing the discrete fractional Fourier transform (DFRT) and one multiplication with a diagonal matrix. The diagonal matrix multiplication in the center can be performed in  $O(N)$ , where  $N$  is the number of samples in the signal. The discrete definition of the FRT and its fast computation are well-studied in the literature and the following references can be listed to name a few, [35], [67]–[74]. A detailed recent review on this issue can also be found in [75]. There are also algorithms for efficient computation of FRT in  $O(N \log N)$  time [67], [76]. Therefore, the conventional fractional domain filtering can be implemented in  $O(N \log N)$  time.

When we generalize to graph signal processing, the computational cost becomes  $O(N^2)$ , since an efficient and fast  $O(N \log N)$  time algorithm for computation of the GFT, let alone the fractional GFT, does not exist. Development of such algorithms is an important future research direction [77]–[79]. In [77], the GFT matrix is factorized into sparse matrices using Flexible Approximate Multi-layer Sparse Transforms (FA  $\mu$  ST), in which each matrix has at most  $2N$  non-zero elements. It is shown that for some setups the complexity will reduce to  $O(N^\alpha)$  for some  $\alpha \in (1, 2)$ . Furthermore, in [80], fast GFT is proposed when the graph is bipartite or satisfies certain topological symmetries. To the best of our knowledge, there are not any attempts for developing a fast algorithm for GFRT.

The computational complexity of the unconstrained Wiener filtering problem for the graph signal processing given in (3) is  $O(N^2)$ . Even though the computational complexity of (12) is the same as (3), if a fast algorithm for GFRT is developed, the cost will be significantly reduced. It is important to note that in conventional (non-graph) signal processing, the complexity of fractional Fourier transformation and fractional Fourier domain filtering is the same as ordinary Fourier transformation and ordinary Fourier domain filtering. The advantages of working in fractional Fourier domains comes at no cost. Thus, the additional cost here does not come from working in fractional domains, but from working on graphs. If a fast algorithm is found for Fourier transformation on graphs, it is likely fractional Fourier transformation will not introduce additional costs.

#### V. NUMERICAL RESULTS

In this section, numerical examples are provided to illustrate the theoretical results. We provide two sets of numerical examples. The first set utilizes graph signals with a physical interpretation that is motivated from sensor networks. The second set presents results for a totally arbitrary graph signal, demonstrating that our results are general and not restricted to a particular choice of signal. In the simulations, we assume that the noise in (2) is uncorrelated with the graph signal  $\mathbf{x}$ ,  $\mathbb{E} \{ \mathbf{n} \} = \mathbf{0}$ , and  $\mathbb{E} \{ \mathbf{n} \mathbf{n}^H \} = \sigma^2 \mathbf{I}$  for some  $\sigma > 0$ , where  $\mathbf{I}$  is the

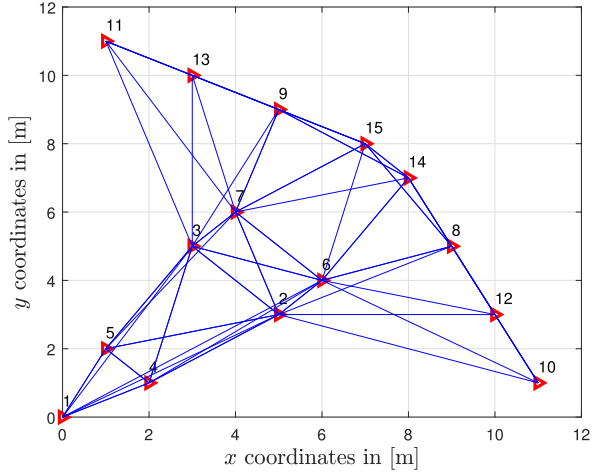


Fig. 2. Sensor network with 15 nodes (undirected and unweighted graph).

identity matrix in  $\mathbb{R}^{N \times N}$ . In addition, for the order of the GFRT, 11 different values  $a = \{\frac{k}{10} \}_{k=0}^{10}$  are used.

For the graph signal  $\mathbf{x} = [x[0] \ x[1] \ \dots \ x[N-1]]^T \in \mathbb{C}^N$ , we assume that its auto-correlation matrix  $\mathbb{E}\{\mathbf{x}\mathbf{x}^H\}$  is given by  $\frac{1}{\lambda_{\max}(\mathbf{C})}\mathbf{C}$ , where  $\lambda_{\max}(\mathbf{C})$  is the largest eigenvalue of  $\mathbf{C} \in \mathbb{R}^{N \times N}$  and for any  $1 \leq i, j \leq N$ ,

$$\mathbf{C}_{ij} = \begin{cases} 2, & \text{if } i = j \\ 1, & \text{if there is a connection between } i \text{ and } j \\ 0, & \text{otherwise.} \end{cases}$$

Division of the matrix  $\mathbf{C}$  by  $\lambda_{\max}(\mathbf{C})$  is just for normalization purposes. By this construction, any two signals corresponding to different nodes are correlated only if the nodes are connected. Moreover, the average power of the graph signal can be calculated as follows  $\mathbb{E}\{\mathbf{x}^H \mathbf{x}\} = \sum_{i=1}^N \mathbb{E}\{|x[i]|^2\} = 2N/\lambda_{\max}(\mathbf{C})$ .

In the simulations, we assume that the entries of  $\mathbf{G}$  are independent and identically distributed (i.i.d.) standard Gaussian random variables; i.e.,  $\mathcal{N}(0, 1)$ . Accordingly,  $N^2$  realizations are generated in a matrix form with dimensions  $N \times N$  by using MATLAB (the seed is equal to 1).

In the following first four subsections, examples are provided in four different setups. First, we consider a sensor network with an undirected and unweighted graph adjacency matrix. Then, we use an exponentially distance based weighted adjacency matrix for the same setup. Third, we consider a sensor network with a directed and unweighted graph adjacency matrix. Then, we work with a directed and weighted graph adjacency matrix.

The graph signal model above is dependent on the graph structure. We also consider an arbitrary graph signal  $\mathbf{x}$  which does not necessarily have any physical interpretation related to the underlying graph structure. Simulation results for an undirected and unweighted graph adjacency matrix with an arbitrary graph signal  $\mathbf{x}$  are presented in the last subsection.

#### A. Example With Undirected and Unweighted Graph

Fig. 2 illustrates the deployment of 15 sensor nodes in a two dimensional area. The sensors are located at  $\{(0, 0), [5, 3], [3, 5],$

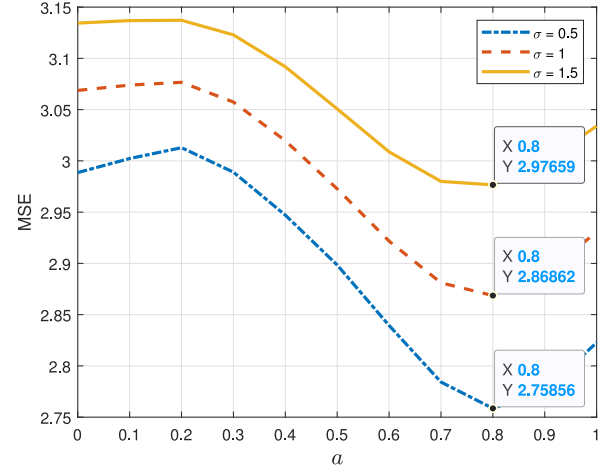


Fig. 3. MSE versus  $a$  for the undirected and unweighted graph for different levels of noise.

$[2, 1], [1, 2], [6, 4], [4, 6], [9, 5], [5, 9], [11, 1], [1, 11], [10, 3], [3, 10], [8, 7], [7, 8]$  meters. Each node is connected to 5 of its nearest neighbors. We assume that the graph is undirected. If node  $i$  and node  $j$  has a connection,  $\mathbf{A}_{ij} = \mathbf{A}_{ji} = 1$ , otherwise  $\mathbf{A}_{ij} = \mathbf{A}_{ji} = 0$ , where  $\mathbf{A}$  is the graph-adjacency matrix of the sensor network given in Fig. 2. Since  $\mathbf{A}$  is undirected, it is diagonalizable and its eigenvalues are real-valued.

For three different  $\sigma$  values, the resulting MSE versus the order of the GFRT,  $a$ , are plotted in Fig. 3. From Fig. 3, it can be observed that the minimum MSE is not attained at  $a = 0$  or  $a = 1$ ; hence, the optimal filtering in the ordinary (non-fractional) space or frequency domains would be suboptimal. The optimal value of the fractional order is  $a = 0.8$  for all values of  $\sigma$ , showing that filtering in fractional domains facilitates a reduction in the MSE.

#### B. Example With Undirected and Weighted Graph

In this example, we use the same setup as in the undirected and unweighted case except that the adjacency matrix is different. We again use the 5-nearest graph structure, but this time, an exponential distance based weighted adjacency matrix is used [5]. In particular,  $\mathbf{A}$  is given by

$$\mathbf{A}_{ij} = \frac{e^{-d_{i,j}^2}}{\sqrt{\sum_{k \in \mathcal{N}_i} e^{-d_{i,k}^2} \sum_{l \in \mathcal{N}_j} e^{-d_{j,l}^2}}} \quad (22)$$

where  $\mathcal{N}_i$  denotes the set of nodes which are the 5 nearest neighbors of node  $i$ , and  $d_{i,j}$  is the distance between node  $i$  and node  $j$ . Based on (22), when the positions of the graph nodes are explicitly known, we are able to construct the adjacency matrix  $\mathbf{A}$ . From (22), one can easily note that  $\mathbf{A}$  is symmetric, i.e.,  $\mathbf{A}_{ij} = \mathbf{A}_{ji}$ , hence diagonalizable. (A detailed discussion of why  $\mathbf{A}$  is constructed as in (22) is beyond the scope of this paper. Using inverse exponents of squared distances is a popular choice to construct weighted adjacency matrices. The construction of representation graphs is an important research question and there

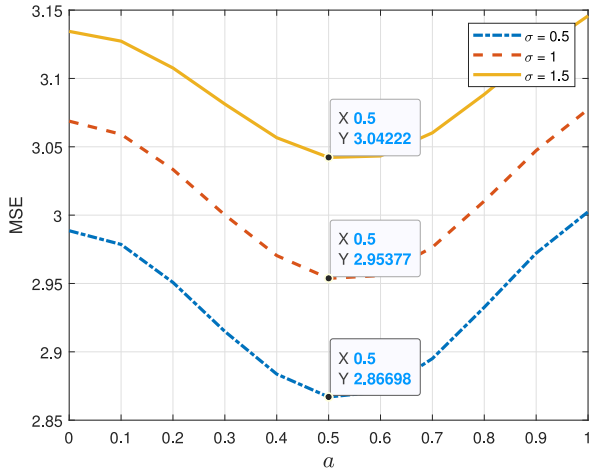


Fig. 4. MSE versus  $a$  for the undirected and weighted graph for different levels of noise.

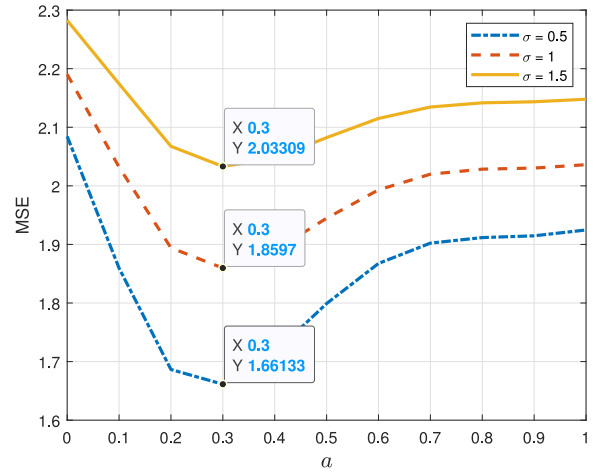


Fig. 6. MSE versus  $a$  for the directed and unweighted graph for different levels of noise.

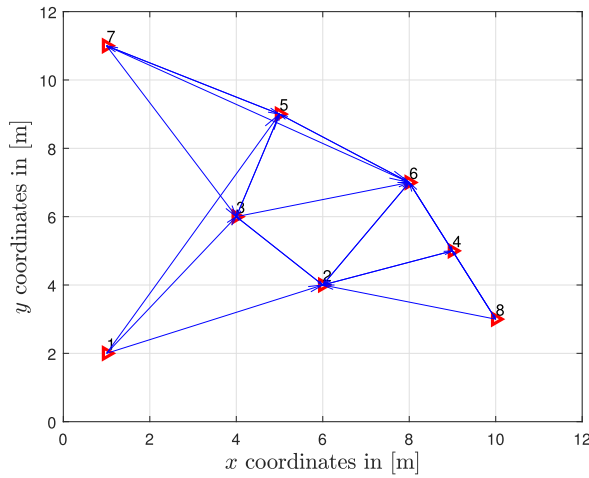


Fig. 5. Sensor network with 8 nodes (directed and unweighted graph).

are various other graph construction methods [2], [21], [81], [82].)

For various  $\sigma$  values, the MSE versus the order of the GFRT  $a$  are plotted in Fig. 4. It is concluded that the minimum MSE is attained for  $a = 0.5$  for all values of  $\sigma$ . In other words, we observe the benefit of filtering in fractional domains.

C. Example With Directed and Unweighted Graph

Fig. 5 represents the deployment of 8 sensor nodes in a two dimensional area. The sensors are located at  $\{[1, 2], [6, 4], [4, 6], [9, 5], [5, 9], [8, 7], [1, 11], [10, 3]\}$  meters and the connections between the nodes are as illustrated in Fig. 5. Each node is connected to 3 of its nearest neighbors. In this case, the graph is directed. If there is an edge from node  $j$  to node  $i$ ,  $\mathbf{A}_{ij} = 1$ , otherwise  $\mathbf{A}_{ij} = 0$ , where  $\mathbf{A}$  is the graph-adjacency matrix of the sensor network in Fig. 5. Since  $\mathbf{A}$  is directed, it is not necessarily diagonalizable and we need to work with its Jordan canonical form.

In Fig. 6, the MSE is plotted versus the order of the GFRT,  $a$ , for three different  $\sigma$  values. From Fig. 6, it can be observed

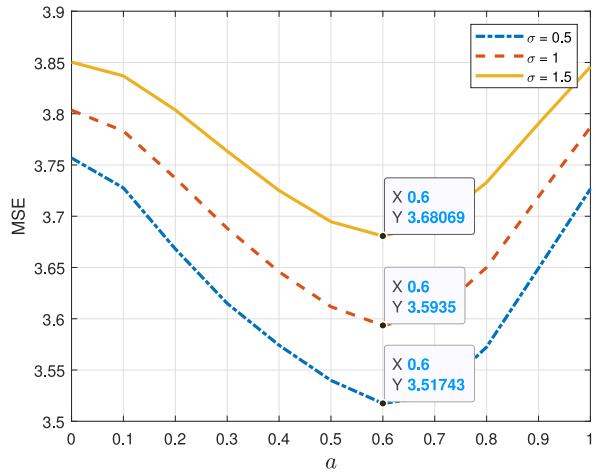


Fig. 7. MSE versus  $a$  for the directed and weighted graph for different levels of noise.

that the minimum MSE is attained at  $a = 0.3$  for all values of  $\sigma$ , not at  $a = 0$  or  $a = 1$ . Once again, we conclude that filtering in fractional domains leads to lower errors.

D. Example With Directed and Weighted Graph

In this example, we use the same setup as in the directed and unweighted case except that the adjacency matrix is weighted and each node is connected to 4 of its nearest neighbor nodes. We use the exponentially distance based weighted graph adjacency matrix, but this time we assume that if there is no edge from node  $j$  to node  $i$ , then  $\mathbf{A}_{ij} = 0$ .

In Fig. 7, the MSE is plotted versus the order of the GFRT,  $a$ , for three different  $\sigma$  values. As in the previous cases, we have again the same observation that the minimum MSE is not attained at  $a = 0$  or  $a = 1$ . For all  $\sigma$  values, the minimum MSE is attained for  $a = 0.6$ .



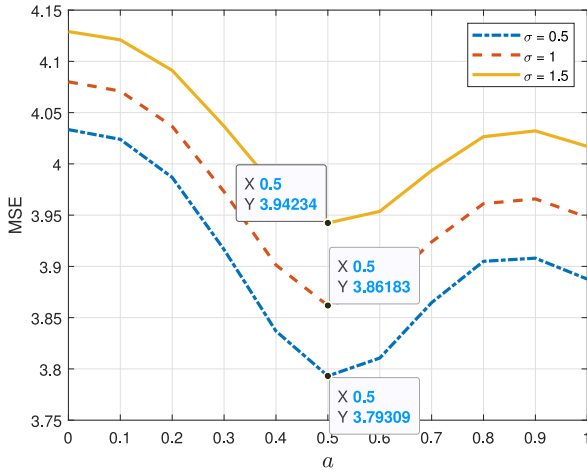


Fig. 8. MSE versus  $a$  for the undirected and unweighted graph for different levels of noise with an arbitrary graph signal.

### E. Example With Undirected and Unweighted Graph With an Arbitrary Graph Signal

To define an arbitrary graph signal  $\mathbf{x}$ , we generate a random matrix  $\mathbf{B}$  such that its entries are i.i.d. standard Gaussian random variables; i.e.,  $\mathcal{N}(0, 1)$ . Accordingly,  $N^2$  realizations are generated in a matrix form with dimensions  $N \times N$  by using MATLAB (the seed is equal to 2). Then,  $\mathbb{E}\{\mathbf{x}\mathbf{x}^H\}$  is set to  $\frac{1}{\lambda_{\max}(\mathbf{B}\mathbf{B}^H)}\mathbf{B}\mathbf{B}^H$ . One should note that,  $\frac{1}{\lambda_{\max}(\mathbf{B}\mathbf{B}^H)}\mathbf{B}\mathbf{B}^H$  is a valid auto-correlation matrix, as it is Hermitian and positive semi-definite matrix. For the simulations, we use the undirected and unweighted graph setup employed previously. In other words, we use the same graph adjacency matrix, the same matrix  $\mathbf{G}$  and the same statistical assumptions about the noise as in the undirected and unweighted graph case.

In Fig. 8, the MSE is plotted versus the order of the GFRT,  $a$ , for three different  $\sigma$  values. As in the previous cases, we have again the same observation that the minimum MSE is not attained at  $a = 0$  or  $a = 1$ . For all  $\sigma$  values, the minimum MSE is attained for  $a = 0.5$ . Hence, benefit of filtering in fractional domains is observed for the arbitrary graph signal scenario, as well.

## VI. CONCLUSION

The optimal fractional Fourier domain filtering problem has been formulated in a discrete graph signal processing context. It has been shown that the proposed problem generalizes the approach in [7], which is a special case when the order of the fractional Fourier transform,  $a$ , is equal to one. The problem has been solved analytically for any kind of graph structures including weighted/unweighted or directed/undirected graphs.

Numerical examples show that the minimum value of the MSE can occur at non-integer values of the fractional Fourier transform order  $a$ . This means that the optimal values do not necessarily correspond to the ordinary time/space or frequency domains but to fractional domains which are not accessible with conventional methods. Since fractional domain filtering does not

bring any additional complexity, the improvements come at no additional cost.

Many operations and problems in signal analysis and processing have been generalized in terms of the fractional Fourier transform. Our results suggest that similar generalizations can be carried over into graph signal processing with beneficial results.

## APPENDIX A

*Proof of Proposition 2:* To prove the convexity of the objective function in (16) with respect to  $\mathbf{h}$ , we take  $\mathbf{z} = [z_1 \ z_2 \ \dots \ z_N]^T$ ,  $\mathbf{w} = [w_1 \ w_2 \ \dots \ w_N]^T \in \mathbb{C}^N$ , and  $\alpha \in [0, 1]$ . If we define  $\mathbf{c} \triangleq \sum_{i=1}^N z_i \mathbf{W}_i \mathbf{y} - \mathbf{x}$  and  $\mathbf{d} \triangleq \sum_{i=1}^N w_i \mathbf{W}_i \mathbf{y} - \mathbf{x}$ , then, it is sufficient to prove the following inequality:

$$\|\alpha \mathbf{c} + (1 - \alpha) \mathbf{d}\|_2^2 \leq \alpha \|\mathbf{c}\|_2^2 + (1 - \alpha) \|\mathbf{d}\|_2^2 \quad (23)$$

Therefore, to prove (23), via some algebraic manipulations, we observe that it is sufficient to show that

$$\mathbf{c}^H \mathbf{d} + \mathbf{d}^H \mathbf{c} \leq \|\mathbf{c}\|_2^2 + \|\mathbf{d}\|_2^2. \quad (24)$$

Using the Cauchy-Schwarz inequality, we can write

$$|\mathbf{c}^H \mathbf{d}| = |\mathbf{d}^H \mathbf{c}| \leq \|\mathbf{c}\|_2 \|\mathbf{d}\|_2. \quad (25)$$

Hence, by using (25), it can be verified that

$$\underbrace{\mathbf{c}^H \mathbf{d} + \mathbf{d}^H \mathbf{c}}_{\mathbf{c}^H \mathbf{d} + \mathbf{d}^H \mathbf{c} \in \mathbb{R}} \leq |\mathbf{c}^H \mathbf{d} + \mathbf{d}^H \mathbf{c}| \leq |\mathbf{c}^H \mathbf{d}| + |\mathbf{d}^H \mathbf{c}| \leq 2\|\mathbf{c}\|_2 \|\mathbf{d}\|_2 \quad (26)$$

Thus, to prove (24), it is sufficient to show that

$$2\|\mathbf{c}\|_2 \|\mathbf{d}\|_2 \leq \|\mathbf{c}\|_2^2 + \|\mathbf{d}\|_2^2 \quad (27)$$

This clearly holds since (27) is true if and only if  $(\|\mathbf{c}\|_2 - \|\mathbf{d}\|_2)^2 \geq 0$ . ■

## APPENDIX B

Computational complexity of determining optimal filter coefficients for any GFRT order  $a$  is given by  $O(N^4)$ .

*Proof:* By Proposition 3, it is known that optimal filter coefficients  $\mathbf{h}^{\text{opt}}$  should satisfy  $\mathbf{T}\mathbf{h}^{\text{opt}} = \mathbf{q}$ . As a consequence of Remark 3, in order to compute each entry of  $\mathbf{T}$  and  $\mathbf{q}$ , we need  $O(N^2)$  operations. Hence, computational complexities of constructing  $\mathbf{T}$  and  $\mathbf{q}$  are  $O(N^4)$  and  $O(N^3)$ , respectively. Since the complexity of the Gauss-Jordan elimination is  $O(N^3)$ , the total complexity is  $O(N^4)$ . ■

## REFERENCES

- [1] A. Sandryhaila and J. M. F. Moura, "Discrete signal processing on graphs," *IEEE Trans. Signal Process.*, vol. 61, no. 7, pp. 1644–1656, Apr. 2013.
- [2] D. I. Shuman, S. K. Narang, P. Frossard, A. Ortega, and P. Vandergheynst, "The emerging field of signal processing on graphs: Extending high-dimensional data analysis to networks and other irregular domains," *IEEE Signal Process. Mag.*, vol. 30, no. 3, pp. 83–98, May 2013.
- [3] A. Sandryhaila and J. M. F. Moura, "Discrete signal processing on graphs: Graph filters," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, May 2013, pp. 6163–6166.
- [4] A. Sandryhaila and J. M. F. Moura, "Big data analysis with signal processing on graphs: Representation and processing of massive data sets with irregular structure," *IEEE Signal Process. Mag.*, vol. 31, no. 5, pp. 80–90, Sep. 2014.

- [5] 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, Jun. 2014.
- [6] S. Chen, R. Varma, A. Sandryhaila, and J. Kovačević, "Discrete signal processing on graphs: Sampling theory," *IEEE Trans. Signal Process.*, vol. 63, no. 24, pp. 6510–6523, Dec. 2015.
- [7] A. Gavili and X. Zhang, "On the shift operator, graph frequency, and optimal filtering in graph signal processing," *IEEE Trans. Signal Process.*, vol. 65, no. 23, pp. 6303–6318, Dec. 2017.
- [8] Y. Wang, B. Li, and Q. Cheng, "The fractional Fourier transform on graphs," in *Proc. Asia-Pacific Signal Inf. Process. Assoc. Annu. Summit Conf.*, Dec. 2017, pp. 105–110.
- [9] G. Ribeiro and J. Lima, "Graph signal processing in a nutshell," *J. Commun. Inf. Syst.*, vol. 33, pp. 219–233, Jan. 2018.
- [10] A. Ortega, P. Frossard, J. Kovačević, J. M. F. Moura, and P. Vandergheynst, "Graph signal processing: Overview, challenges, and applications," *Proc. IEEE*, vol. 106, no. 5, pp. 808–828, May 2018.
- [11] Y. Wang and B. Li, "The fractional Fourier transform on graphs: Sampling and recovery," in *Proc. 14th IEEE Int. Conf. Signal Process.*, 2018, pp. 1103–1108.
- [12] L. Stanković, D. P. Mandić, M. Daković, M. Brajović, B. S. Dees, and T. Constantinides, "Graph signal processing—Part I: Graphs, graph spectra, and spectral clustering," 2019, [arXiv:1907.03467](https://arxiv.org/abs/1907.03467).
- [13] L. Stanković, D. Mandić, M. Daković, M. Brajović, B. Scalzo Dees, and A. Constantinides, "Graph signal processing—Part II: Processing and analyzing signals on graphs," 2019, [arXiv:1909.10325](https://arxiv.org/abs/1909.10325).
- [14] L. Stanković *et al.*, "Vertex-frequency graph signal processing: A comprehensive review," *Digit. Signal Process.*, vol. 107, 2020, Art. no. 102802.
- [15] F. Hua, C. Richard, J. Chen, H. Wang, P. Borgnat, and P. Gonçalves, "Learning combination of graph filters for graph signal modeling," *IEEE Signal Process. Lett.*, vol. 26, no. 12, pp. 1912–1916, Dec. 2019.
- [16] F. Grassi, A. Loukas, N. Perraudin, and B. Ricaud, "A time-vertex signal processing framework: Scalable processing and meaningful representations for time-series on graphs," *IEEE Trans. Signal Process.*, vol. 66, no. 3, pp. 817–829, Feb. 2018.
- [17] 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.
- [18] S. K. Narang, A. Gadde, and A. Ortega, "Signal processing techniques for interpolation in graph structured data," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, 2013, pp. 5445–5449.
- [19] A. Sandryhaila and J. M. F. Moura, "Discrete signal processing on graphs: Graph Fourier transform," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, 2013, pp. 6167–6170.
- [20] X. Zheng, Y. Y. Tang, J. Pan, and J. Zhou, "Adaptive multiscale decomposition of graph signals," *IEEE Signal Process. Lett.*, vol. 23, no. 10, pp. 1389–1393, Oct. 2016.
- [21] R. R. Coifman and M. Maggioni, "Diffusion wavelets," *Appl. Comput. Harmon. Anal.*, vol. 21, no. 1, pp. 53–94, 2006.
- [22] D. K. Hammond, P. Vandergheynst, and R. Gribonval, "Wavelets on graphs via spectral graph theory," *Appl. Comput. Harmon. Anal.*, vol. 30, no. 2, pp. 129–150, 2011.
- [23] S. K. Narang and A. Ortega, "Perfect reconstruction two-channel wavelet filter banks for graph structured data," *IEEE Trans. Signal Process.*, vol. 60, no. 6, pp. 2786–2799, Jun. 2012.
- [24] N. Perraudin and P. Vandergheynst, "Stationary signal processing on graphs," *IEEE Trans. Signal Process.*, vol. 65, no. 13, pp. 3462–3477, Jul. 2017.
- [25] A. G. Marques, S. Segarra, G. Leus, and A. Ribeiro, "Stationary graph processes and spectral estimation," *IEEE Trans. Signal Process.*, vol. 65, no. 22, pp. 5911–5926, Nov. 2017.
- [26] L. Stanković *et al.*, "Graph signal processing - Part III: Machine learning on graphs, from graph topology to applications," 2020, [arXiv:2001.00426](https://arxiv.org/abs/2001.00426).
- [27] W. Huang, L. Goldsberry, N. F. Wymbs, S. T. Grafton, D. S. Bassett, and A. Ribeiro, "Graph frequency analysis of brain signals," *IEEE J. Sel. Topics Signal Process.*, vol. 10, no. 7, pp. 1189–1203, Oct. 2016.
- [28] R. Singh, A. Chakraborty, and B. S. Manoj, "Graph Fourier transform based on directed laplacian," in *Proc. Int. Conf. Signal Process. Commun.*, 2016, pp. 1–5.
- [29] S. Chen, A. Sandryhaila, J. M. Moura, and J. Kovačević, "Adaptive graph filtering: Multiresolution classification on graphs," in *Proc. IEEE Global Conf. Signal Inf. Process.*, 2013, pp. 427–430.
- [30] 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, pp. 2879–2893, Jun. 2014.
- [31] H. M. Ozaktas and D. Mendlovic, "Fourier transforms of fractional order and their optical interpretation," *Opt. Commun.*, vol. 101, no. 3–4, pp. 163–169, 1993.
- [32] D. Mendlovic and H. M. Ozaktas, "Fractional Fourier transforms and their optical implementation: I," *J. Opt. Soc. Amer. A*, vol. 10, pp. 1875–1881, 1993.
- [33] H. M. Ozaktas and D. Mendlovic, "Fractional Fourier transforms and their optical implementation: II," *J. Opt. Soc. Amer. A*, vol. 10, pp. 2522–2531, 1993.
- [34] H. M. Ozaktas, Z. Zalevsky, and M. A. Kutay, *The Fractional Fourier Transform With Applications in Optics and Signal Processing*. New York, NY, USA: Wiley, 2001.
- [35] S. C. Pei and J. J. Ding, "Closed-form discrete fractional and affine Fourier transforms," *IEEE Trans. Signal Process.*, vol. 48, no. 5, pp. 1338–1353, May 2000.
- [36] L. B. Almeida, "The fractional Fourier transform and time-frequency representations," *IEEE Trans. Signal Process.*, vol. 42, no. 11, pp. 3084–3091, Nov. 1994.
- [37] H. M. Ozaktas, M. A. Kutay, and C. Candan, "Transforms and applications handbook," in *Fractional Fourier Transform*. New York, NY, USA: CRC Press, Boca Raton, 2010, pp. 14–1–14–28.
- [38] A. I. Zayed and A. G. Garcia, "New sampling formulae for the fractional Fourier transform," *Signal Process.*, vol. 77, no. 1, pp. 111–114, 1999.
- [39] A. W. Lohmann, "Image rotation, wigner rotation, and the fractional Fourier transform," *J. Opt. Soc. Amer. A*, vol. 10, no. 10, pp. 2181–2186, 1993.
- [40] M. A. Kutay, H. M. Ozaktas, O. Arikan, and L. Onural, "Optimal filtering in fractional Fourier domains," *IEEE Trans. Signal Process.*, vol. 45, no. 5, pp. 1129–1143, May 1997.
- [41] A. I. Zayed, "A convolution and product theorem for the fractional Fourier transform," *IEEE Signal Process. Lett.*, vol. 5, no. 4, pp. 101–103, Apr. 1998.
- [42] H. M. Ozaktas, B. Barshan, D. Mendlovic, and L. Onural, "Convolution, filtering, and multiplexing in fractional Fourier domains and their relation to chirp and wavelet transforms," *J. Opt. Soc. Amer. A*, vol. 11, no. 2, pp. 547–559, 1994.
- [43] E. Sejdici, I. Djurović, and L. Stanković, "Fractional Fourier transform as a signal processing tool: An overview of recent developments," *Signal Process.*, vol. 91, no. 6, pp. 1351–1369, 2011.
- [44] S. C. Pei and J. J. Ding, "Relations between gabor transforms and fractional Fourier transforms and their applications for signal processing," *IEEE Trans. Signal Process.*, vol. 55, no. 10, pp. 4839–4850, Oct. 2007.
- [45] R. Tao, B. Z. Li, and Y. Wang, "Spectral analysis and reconstruction for periodic nonuniformly sampled signals in fractional Fourier domain," *IEEE Trans. Signal Process.*, vol. 55, no. 7, pp. 3541–3547, Jul. 2007.
- [46] P. Mohindru, R. Khanna, and S. S. Bhatia, "New tuning model for rectangular windowed FIR filter using fractional Fourier transform," *Signal, Image Video Process.*, vol. 9, no. 4, pp. 761–767, May 2015.
- [47] P. K. Anh, L. P. Castro, P. T. Thao, and N. M. Tuan, "New sampling theorem and multiplicative filtering in the FRFT domain," *Signal, Image Video Process.*, vol. 13, no. 5, pp. 951–958, Jul. 2019.
- [48] S. C. Pei and J. J. Ding, "Relations between the fractional operations and time-frequency distributions, and their applications," *IEEE Trans. Signal Process.*, vol. 49, no. 8, pp. 1638–1655, Aug. 2001.
- [49] S. C. Pei and J. J. Ding, "Relations between the fractional operations and the wigner distribution, ambiguity function," *IEEE Trans. Signal Process.*, vol. 49, no. 8, pp. 209–219, Oct. 2001.
- [50] J. Shi, M. Han, and N. Zhang, "Uncertainty principles for discrete signals associated with the fractional Fourier and linear canonical transforms," *Signal, Image Video Process.*, vol. 10, no. 8, pp. 1519–1525, 2016.
- [51] A. Koç, B. Bartan, E. Gundogdu, T. Çukur, and H. M. Ozaktas, "Sparse representation of two- and three-dimensional images with fractional Fourier, hartley, linear canonical, and haar wavelet transforms," *Expert Syst. Appl.*, vol. 77, pp. 247–255, 2017.
- [52] X. Zhang *et al.*, "Optimal design of orders of DFrFTs for sparse representations," *IET Signal Process.*, vol. 12, no. 8, pp. 1023–1033, 2018.
- [53] K. K. Sharma and M. Sharma, "Image fusion based on image decomposition using self-fractional Fourier functions," *Signal, Image Video Process.*, vol. 8, no. 7, pp. 1335–1344, Oct. 2014.

- [54] N. Jindal and K. Singh, "Image and video processing using discrete fractional transforms," *Signal, Image Video Process.*, vol. 8, no. 8, pp. 1543–1553, Nov. 2014.
- [55] N.-L. Tian, X.-Z. Zhang, B. W.-K. Ling, and Z.-J. Yang, "Two-dimensional discrete fractional Fourier transform-based content removal algorithm," *Signal, Image Video Process.*, vol. 10, no. 7, pp. 1311–1318, Oct. 2016.
- [56] L. Stanković, T. Alieva, and M. J. Bastiaans, "Time - frequency signal analysis based on the windowed fractional Fourier transform," *Signal Process.*, vol. 83, no. 11, pp. 2459–2468, 2003.
- [57] T. Alieva, M. J. Bastiaans, and L. Stanković, "Signal reconstruction from two close fractional Fourier power spectra," *IEEE Trans. Signal Process.*, vol. 51, no. 1, pp. 112–123, Jan. 2003.
- [58] D. Mendlovic, Z. Zalevsky, and H. M. Ozaktas, "Optical pattern recognition," in *The Applications of the Fractional Fourier Transform to Optical Pattern Recognition*. T. S. Yu and S. Jutamulia, Eds. Cambridge, UK: Cambridge Univ. Press, pp. 89–125, 1998.
- [59] R. Jacob, T. Thomas, and A. Unnikrishnan, "Applications of fractional Fourier transform in sonar signal processing," *IETE J. Res.*, vol. 55, no. 1, pp. 16–27, 2009.
- [60] Z. Zhao, R. Tao, G. Li, and Y. Wang, "Fractional sparse energy representation method for ISAR imaging," *IET Radar, Sonar Navigation*, vol. 12, no. 9, pp. 988–997, 2018.
- [61] M. I. Ahmad, M. U. Sardar, and I. Ahmad, "Blind beamforming using fractional Fourier transform domain cyclostationarity," *Signal, Image Video Process.*, vol. 12, no. 2, pp. 379–383, Feb. 2018.
- [62] M. I. Ahmad, "Optimum FrFT domain cyclostationarity based adaptive beamforming," *Signal, Image Video Process.*, vol. 13, no. 3, pp. 551–556, Apr. 2019.
- [63] H. M. Ozaktas and D. Mendlovic, "Fractional Fourier optics," *J. Opt. Soc. Amer. A*, vol. 12, pp. 743–751, 1995.
- [64] H. M. Ozaktas and M. F. Erden, "Relationships among ray optical, gaussian beam, and fractional Fourier transform descriptions of first-order optical systems," *Opt. Commun.*, vol. 143, pp. 75–86, 1997.
- [65] Z. Zalevsky and D. Mendlovic, "Fractional Wiener filter," *Appl. Opt.*, vol. 35, no. 20, pp. 3930–3936, Jul. 1996.
- [66] J. Gutiérrez-Gutiérrez, A. Podhorski, I. Iglesias, and J. Del Ser, "A note on the Wiener filter for vector random processes," in *Mobile Lightweight Wireless Systems*. J. Del Ser, E. A. Jorswieck, J. Miguez, M. Matinmikko, D. P. Palomar, S. Salcedo-Sanz, and S. Gil-Lopez, Eds., Berlin, Heidelberg: Springer, 2012, pp. 30–36.
- [67] H. M. Ozaktas, O. Arıkan, M. A. Kutay, and G. Bozdağı, "Digital computation of the fractional Fourier transform," *IEEE Trans. Signal Process.*, vol. 44, no. 9, pp. 2141–2150, Sep. 1996.
- [68] N. M. Atakishiyev, L. E. Vicent, and K. B. Wolf, "Continuous vs. discrete fractional Fourier transforms," *J. Comput. Appl. Math.*, vol. 107, no. 1, pp. 73–95, 1999.
- [69] L. Barker, C. Candan, T. Hakioglu, M. A. Kutay, and H. M. Ozaktas, "The discrete harmonic oscillator, harper's equation, and the discrete fractional Fourier transform," *J. Phys. A: Math. Gen.*, vol. 33, no. 11, 2000, Art. no. 2209.
- [70] C. Candan, M. A. Kutay, and H. M. Ozaktas, "The discrete fractional Fourier transform," *IEEE Trans. Signal Process.*, vol. 48, no. 5, pp. 1329–1337, May 2000.
- [71] S. C. Pei and M. H. Yeh, "Improved discrete fractional Fourier transform," *Opt. Lett.*, vol. 22, no. 14, pp. 1047–1049, 1997.
- [72] S. C. Pei, W. Hsue, and J. Ding, "Discrete fractional Fourier transform based on new nearly tridiagonal commuting matrices," *IEEE Trans. Signal Process.*, vol. 54, no. 10, pp. 3815–3828, Oct. 2006.
- [73] S. C. Pei, M. H. Yeh, and C. C. Tseng, "Discrete fractional Fourier transform based on orthogonal projections," *IEEE Trans. Signal Process.*, vol. 47, no. 5, pp. 1335–1348, May 1999.
- [74] A. Koç, "Operator theory-based discrete fractional Fourier transform," *Signal, Image Video Process.*, vol. 13, no. 7, pp. 1461–1468, 2019.
- [75] X. Su, R. Tao, and X. Kang, "Analysis and comparison of discrete fractional Fourier transforms," *Signal Process.*, vol. 160, pp. 284–298, 2019.
- [76] A. Koç, F. S. Oktem, H. M. Ozaktas, and M. A. Kutay, "Linear canonical transforms: Theory and applications," in *Fast Algorithms for Digital Computation of Linear Canonical Transforms*, J. J. Healy, M. A. Kutay, H. M. Ozaktas, J. T. Sheridan, Eds. New York, NY, USA: Springer, 2016, pp. 293–327.
- [77] L. L. Magoarou and R. Gribonval, "Are there approximate fast Fourier transforms on graphs?," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, 2016, pp. 4811–4815.
- [78] L. Le Magoarou, R. Gribonval, and N. Tremblay, "Approximate fast graph Fourier transforms via multilayer sparse approximations," *IEEE Trans. Signal Inf. Process. Netw.*, vol. 4, no. 2, pp. 407–420, Jun. 2018.
- [79] J. Domingos and J. M. F. Moura, "Graph Fourier transform: A stable approximation," *IEEE Trans. Signal Process.*, vol. 68, pp. 4422–4437, 2020.
- [80] K. Lu and A. Ortega, "Fast graph Fourier transforms based on graph symmetry and bipartition," *IEEE Trans. Signal Process.*, vol. 67, no. 18, pp. 4855–4869, Sep. 2019.
- [81] M. Belkin and P. Niyogi, "Laplacian eigenmaps for dimensionality reduction and data representation," *Neural Comput.*, vol. 15, no. 6, pp. 1373–1396, 2003.
- [82] C. Guestrin, P. Bodik, R. Thibaux, M. Paskin, and S. Madden, "Distributed regression: An efficient framework for modeling sensor network data," in *Proc. 3rd Int. Symp. Inf. Process. Sensor Netw.*, 2004, pp. 1–10.