Hilbert Transform, Analytic Signal, and Modulation Analysis for Graph Signal Processing

Arun Venkitaraman, Saikat Chatterjee, Peter Händel
School of Electrical Engineering and ACCESS Linnaeus Center
KTH Royal Institute of Technology, SE-100 44 Stockholm, Sweden
arunv@kth.se, sach@kth.se, ph@kth.se

Abstract—We propose Hilbert transform and analytic signal construction for signals over graphs. This is motivated by the popularity of Hilbert transform, analytic signal, and modulation analysis in conventional signal processing, and the observation that complementary insights are often obtained by viewing conventional signals in the graph setting. Our definitions of Hilbert transform and analytic signal use a conjugate-symmetry-like property exhibited by the graph Fourier transform (GFT), resulting in a 'one-sided' spectrum for the graph analytic signal. Using the graph analytic signal, we define amplitude, phase, and frequency modulations for a graph signal. Further, we use convex optimization to develop an alternative definition of envelope for a graph signal that explicitly enforces smoothness. We illustrate the proposed concepts by showing applications to synthesized and real-world signals. For example, we show that the graph Hilbert transform can indicate presence of anomalies and that graph analytic signal, and associated amplitude and frequency modulations reveal complementary information in speech signals.

Index Terms—Graph signal, analytic signal, Hilbert transform, demodulation, anomaly detection.

EDICS—ADEL-SIPOG.

I. INTRODUCTION

The analysis of data over networks or graphs poses unique challenges to the signal processing community, since data must be seen with due regard to the connections between various data points or nodes of a graph [1]–[3]. Given the wealth of techniques and models in conventional signal analysis, it is desirable to extend existing concepts to signals over graphs [4]–[7]. The collective efforts along this line of thought have led to the emergence of the notion of signal processing over graphs [2], [3], [8], [9]. In this paper, we generalize the concepts of Hilbert transform, analytic signal, and modulation analysis to signals over graphs. This is motivated by two observations. Firstly, Hilbert transform, analytic signal, and associated modulation analysis have been used extensively for one-dimensional (1D) and two-dimensional (2D) signals in various applications [10], [11]. By extending modulation analysis to graphs, we endeavour to provide similar tools for signals over graphs. Secondly, viewing of 1D /2D signals in a graph setting has been shown to give additional insight into the signals, leading to improved performance in tasks such as compression and denoising [3], [12].

A. Review of literature

Some of the early works in graph signal processing include windowed Fourier transforms [13], filterbanks [14]–[16], wavelet transforms and multiresolution representations for graphs [17]–[20]. A number of strategies for efficient sampling of signals over graphs have been proposed [21]–[25]. The notions of stationarity and power spectral density have also been considered extensively for signals over graphs [26]–[29]. A parametric dictionary learning approach for graph signals was proposed by Thanou et al. [30]. In [31], Shuman et al. generalized the notion of time-frequency analysis to the graph setting using windowed graph Fourier transforms. Shahid et al. proposed variants of principal component analysis for graph signals and developed scalable and efficient algorithms for recovery of low-rank matrices [32], [33]. Tremblay et al. proposed an efficient spectral clustering algorithm based on graph signal filtering [34]. In [35], Benzi et al. developed a song recommendation system based on non-negative matrix factorization and graph total variation. Shuman et al. proposed a multi-scale pyramid transform on graphs that generates a multiresolution of both the graph and the signal [36]. Segarra et al. proposed convex optimization based approaches for blind identification of graph filters [37], [38]. Recently, Chen et al. considered signal recovery on graphs based on total-variation minimization formulated as a convex optimization problem [39]. A sampling scheme for graph signals preserving the first-order difference of the graph signal was also proposed by Chen et al. [40]. Sakiyama et al. proposed spectral graph wavelets and filterbanks constructed as sum of sinusoids in spectral domain with low approximation error [41]. Deutsch et al. showed the application of spectral graph wavelets to manifold denoising [42]. A trilateral filter based denoising scheme was proposed by Onuki et al. [43]. Anis et al. showed that bandlimited graph signals can be efficiently sampled using graph signal proxies [44]. Multirate signal processing concepts including M-channel filter banks were also recently extended to the graph setting by Teke and Vaidyanathan [4], [5]. Kernel regression approaches for reconstruction of graph signals have also been proposed in the framework of reproducing kernel Hilbert spaces [45], [46].

The authors would like to acknowledge the support received from the Swedish Research Council.

The codes necessary for reproducing the material in this article may be found at http://www.ee.kth.se/reproducible.

arXiv:1611.05269v2 [cs.IT] 29 May 2017
B. Our contributions

In this paper, we propose definitions for the Hilbert transform and analytic signal for real signals over graphs\textsuperscript{1}. We show that a real graph signal with a real-valued adjacency matrix may be represented using fewer number of GFT coefficients than the signal length, akin to the ‘one-sided’ spectrum for 1D signals. We generalize the Hilbert transform and analytic signal construction \cite{cardoso2012convex} to graph signals by using the conjugate-symmetry-like property of the GFT basis. We show that the resulting graph Hilbert transform and graph analytic signal operations are linear and shift-invariant over graphs. We also show that graph Hilbert transform and graph analytic signal inherit properties such as isometry, phase-shifting, and orthogonality from their 1D counterparts. We illustrate experimentally that the graph Hilbert transform highlights anomalies in certain graphs of interest. We discuss how the graph Hilbert transform does not possess a Bedrosian-type property in general unlike its conventional counterpart. As a natural consequence of the graph Hilbert transform construction, we show that it is possible to define amplitude, phase, and frequency modulations for graph signals. We also develop an algorithm for unwrapping the phase of the graph analytic signal. We show that the proposed definitions contain the conventional counterparts as a special case. In order to obtain a graph signal envelope that is smooth over the graph, we also propose an alternative definition by formulating a convex optimization problem. We illustrate the proposed concepts with applications to synthesized and real-world signals. Our experiments show that graph Hilbert transform and convex graph envelope can reveal edge connections and presence of anomalies over the graph. We also demonstrate that viewing the speech signal as a graph signal brings improves speaker classification performance. We summarize the key similarities and differences between the conventional Hilbert transform/analytic signal and our graph Hilbert transform/graph analytic signal in Table I.

\begin{table}[h]
\centering
\begin{tabular}{|c|c|}
\hline
Hilbert transform/analytic signal & Graph Hilbert transform/graph analytic signal \\
\hline
Linear and shift invariant & Linear and graph-shift-invariant \\
\hline
Quadrature phase-shifting & Generalized quadrature phase-shifting \\
\hline
Changes with permutation of samples & Invariant to node permutations \\
\hline
Bedrosian property & Lacks Bedrosian property \\
\hline
One-sided spectrum & Near one-sided spectrum \\
\hline
Special case of graph Hilbert transform/graph analytic signal & Added insight into 1D signals \\
\hline
\end{tabular}
\caption{Key similarities and differences between conventional Hilbert transform/analytic signal and proposed graph Hilbert transform/analytic signal.}
\end{table}

II. Preliminaries

A. Graph signal processing

Let $x \in \mathbb{R}^\mathcal{V}$ be a real signal on the graph $G = (\mathcal{V}, \mathcal{A})$, where $\mathcal{V}$ and $\mathcal{A}$ denote the node set and the adjacency matrix, respectively. Then, the GFT of $x$ is defined as \cite{cardoso2012convex, cardoso2012two}:

\[ x \triangleq \left( \hat{x}(1), \hat{x}(2), \ldots, \hat{x}(i), \ldots, \hat{x}(N) \right)^\top = \mathbf{V}^{-1}x, \]

where $\mathbf{V}$ denotes the eigenvector matrix such that $\mathbf{A} = \mathbf{VJV}^{-1}$, and $\mathbf{J}$ the diagonal eigenvalue matrix $\mathbf{J} = \text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_N)$. An $N$-sample periodic 1D signal can be seen as a graph signal $x$ with the adjacency matrix $A = C \triangleq \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & 0 & \cdots & 0 \end{pmatrix}$, \hspace{1cm} (1)

and in this case, the GFT coincides with the discrete Fourier transform (DFT) \cite{ sugars2012quadrature, corrias2013graph}.

The smoothness of a graph signal may be measured in terms of the following mean-squared cost:

\[ \text{MS}_g(x) = \left\| \mathbf{x} - \mathbf{Ax} \right\|_2^2 \left/ \|\mathbf{x}\|_2^2 \right. \]

A graph signal with low $\text{MS}_g$ is smooth over the graph: connected nodes have signal values close to each other, stronger the edge closer the values. A unit shift of $x$ over the graph is defined as $\mathbf{A}x$, generalizing the notion of unit delay for the 1D graph.

A linear shift-invariant filter $\mathbf{H}$ on graph is defined as a polynomial $h(\cdot)$ of the adjacency matrix, such that:

\[ \mathbf{H} = \sum_{l=0}^{L} h_l \mathbf{A}^l = h(\mathbf{A}), \]

where $h_l \in \mathbb{R}$ and $L \leq N$. Such a filter also follows the convolution property such that the $k$th GFT coefficient of filtered graph signal $y = \mathbf{H}x$ is given by:

\[ \hat{y}(k) = \sum_{l=0}^{L} h_l \lambda_k^l \hat{x}(k) = h(\lambda_k) \hat{x}(k). \]

Graph filters are used in spectral analysis and processing of graph signals and have been applied in various applications \cite{boccaletti2014complex, zhou2014graph, cardoso2012convex, cardoso2012two, mukherjee2014graph, lasithas2014spectral}.

B. The conventional analytic signal

Let $\hat{x}(\omega)$ denote the DFT of the real 1D signal $x$ evaluated at frequency $\omega$. Then, the discrete analytic signal of $x$, denoted by $x_{a,c}$, has the following frequency-domain definition \cite{cardoso2012convex, cardoso2012two}:

\[ \hat{x}_{a,c}(\omega) = \begin{cases} 2\hat{x}(\omega), & \omega \in \left\{ \frac{2\pi}{N}, \cdots, \pi - \frac{2\pi}{N} \right\} \\ \hat{x}(\omega), & \omega \in \left\{ 0, \pi \right\} \\ 0, & \omega \in \left\{ \pi + \frac{2\pi}{N}, \cdots, \frac{2\pi(N-1)}{N} \right\} \end{cases}, \] \hspace{1cm} (2)

Taking the inverse DFT on both sides of (2), we get that $x_{a,c} = x + jx_{h,c}$, where $j = \sqrt{-1}$ and $x_{h,c}$ is known as the discrete Hilbert transform of $x$ \cite{cardoso2012convex}. The graph Hilbert transform has the following frequency-domain specification:

\[ \hat{x}_{a,c}(\omega) = \begin{cases} -j\hat{x}(\omega), & \omega \in \left\{ \frac{2\pi}{N}, \cdots, \pi - \frac{2\pi}{N} \right\} \\ \hat{x}(\omega), & \omega \in \left\{ 0, \pi \right\} \\ +j\hat{x}(\omega), & \omega \in \left\{ \pi + \frac{2\pi}{N}, \cdots, \frac{2\pi(N-1)}{N} \right\} \end{cases}. \] \hspace{1cm} (3)

\textsuperscript{1}Part of this work has appeared in the Proceedings of the Sampling Theory and Applications Conference, 2015 \cite{cardoso2012two}.
III. GRAPH ANALYTIC SIGNAL

We next define an analytic signal for signals over graphs. In our analysis, we make the following assumptions:

1. \( A \) is real, asymmetric, and diagonalizable with at least one conjugate-pair of eigenvalues.

2. The eigenvalues of \( A \) are sorted in the ascending order of their phase angle from 0 to \( 2\pi \) to form the diagonal matrix \( J \), and correspondingly the eigenvectors are sorted such that \( A = V J V^{-1} \). If multiple eigenvalues with same phase angle occur, we order them in the descending order of their magnitude.

We recall that the eigenvalues of a real-valued matrix, and the corresponding eigenvectors are either real-valued or occur in complex-conjugate pairs. Let \( K_1 \) and \( K_2 \) denote the number of real-valued positive and negative eigenvalues of \( A \), respectively, and \( K = K_1 + K_2 \). Let us define the sets:

\[
\begin{align*}
\Gamma_1 &= \{1, \cdots, K_1\} \text{ (positive real eigenvalues)}, \\
\Gamma_2 &= \{K_1 + 1, \cdots, K + \frac{N-K}{2}\} \text{ (eigenvalues with phase angle in } (0, \pi)), \\
\Gamma_3 &= \{K + \frac{N-K}{2} + 1, \cdots, \frac{N-K}{2}\} \text{ (negative real eigenvalues),} \\
\Gamma_4 &= \{\frac{N-K}{2} + 1, \cdots, N\} \text{ (eigenvalues with phase angle in } (\pi, 2\pi)),
\end{align*}
\]

and denote the vector spaces spanned by the corresponding eigenvectors by \( V_1, V_2, V_3, \) and \( V_4 \), respectively. For example, \( V_1 \) is the space spanned by the eigenvectors related to \( \Gamma_1 \). On ordering as per Assumption 2, we have that the \( i \)th eigenvector of \( A \), \( i \in \Gamma_2 \) is the complex conjugate of eigenvector indexed by the \((N-i+K_1+1)\)th entry of \( \Gamma_4 \):

\[
v_i = v^*_{(N-i+K_1+1)}, \quad i \in \Gamma_2.
\]

Since the \( i \)th GFT coefficient \( \hat{x}(i) \) is given by the inner product of \( x \) with \( v_i \), we have that

\[
\hat{x}(i) = \hat{\bar{x}}(N-i+K_1+1), \quad i \in \Gamma_2.
\]

(4)

For real-valued \( A \), \( N \) and \( K \) are always of the same parity (odd or even). In the case of 1D signals, (4) reduces to the conjugate-symmetry property of the DFT [52]. Equation (4) indicates that a real graph signal can be represented using \( \theta \) GFT coefficients, where \( \theta = |\Gamma_1| + |\Gamma_2| + |\Gamma_3| = (N+K)/2 \), and \( |\Gamma| \) denotes the cardinality of the set \( \Gamma \). For \( K \ll N \), \( \theta \approx N/2 \). We note that (4) holds only if \( x \) is real, which means that a graph signal which does not satisfy (4) is necessarily complex-valued. Motivated by (4) and conventional analytic signal construction, we next define the graph analytic signal and graph Hilbert transform.

Definition. We define the graph analytic signal of \( x \) as \( x_a = V \hat{x}_a \), where

\[
\hat{x}_a(i) = \begin{cases} 
2\hat{x}(i), & i \in \Gamma_2 \\
\hat{x}(i), & i \in \Gamma_1 \cup \Gamma_3 \\
0, & i \in \Gamma_4
\end{cases}.
\]

(5)

As a consequence of the ‘one-sidedness’ of the GFT spectrum, we have that \( x_a \) is complex and hence, is expressible as \( x_a = x_a^r + j x_a^i \). We define \( x_a^r \) as the graph Hilbert transform of \( x_a \).

Then, from the definition of the graph analytic signal, we have that

\[
\hat{x}_a(i) = \begin{cases} 
\hat{x}(i), & i \in \Gamma_2 \\
\hat{x}(i), & i \in \Gamma_1 \cup \Gamma_3 \\
0, & i \in \Gamma_4
\end{cases}.
\]

(6)

On setting \( A = C \), we observe that (5) reduces to the conventional Hilbert transform/analytic signal definitions given by (2) and (3), that is, \( x_a = x_{a,c} \) and \( x_h = x_{h,c} \) since the \( i \)th graph frequency is equal to \( e^{j\omega_i} \) where \( \omega_i \in \{0, \frac{2\pi}{N}, \cdots, \pi\} \). This corresponds to \( \Gamma_1 = \{1\} \), \( \Gamma_2 = \{2, \cdots, N/2-1\} \), \( \Gamma_3 = \{N/2\} \), and \( \Gamma_4 = \{N/2+1, \cdots, N-1\} \) for even \( N \). For odd \( N \), this corresponds to \( \Gamma_1 = \{1\} \), \( \Gamma_2 = \{2, \cdots, (N+1)/2\} \), \( \Gamma_3 = \{\} \), and \( \Gamma_4 = \{(N+1)/2+1, \cdots, N-1\} \).

As an illustration of the graph Hilbert transform/graph analytic signal construction, consider a graph with an adjacency matrix with eigenvalues distributed according to Figure 1(a). Let us consider a signal \( x \) having unit GFT magnitude for all the graph frequencies. Then, graph Hilbert transform of \( x \) has the GFT spectrum shown in Figure 1(c) since \( \Gamma_1 = \{1\} \), \( \Gamma_2 = \{2, 3, 4\} \), \( \Gamma_3 = \{5\} \), and \( \Gamma_4 = \{6, 7, 8\} \). The corresponding graph analytic signal has the GFT spectrum shown in Figure 1(d).

A. One-sided spectrum of the graph analytic signal

The exact number of nonzero values in the graph analytic signal depends on the structure of the graph and its adjacency matrix \( A \). In the case when all the eigenvalues of \( A \) are complex \((K = 0)\), the number of non-zero coefficients in \( x_a \) is exactly one half of the total resulting in a one-sided spectrum, that is, \( \theta = N/2 \). We list the \( \theta \) values for the 1D graph and random graphs Table II. For asymmetric matrices with entries drawn from independently and identically distributed
(IID) mean zero unit variance Gaussian distribution \( \mathcal{N}(0, 1) \), the fraction of real eigenvalues asymptotically tends to zero [54]. This was also shown to hold experimentally for matrices with independent and identically distributed entries from the uniform distribution over \([-1, 1]\): \( \mathcal{U}(-1, 1) \), and Bernoulli \((-1, 1)\) entries [54–56]. We note here that adjacency matrix with Bernoulli entries represents the Erdős–Rényi model for small-world graphs\(^2\). This implies that the corresponding graphs with adjacency matrices drawn from these distributions asymptotically have one-sided graph analytic signal spectrum. Since the Gaussian random matrix is a good approximation to general random matrices in terms of spectral properties, one can conclude that, on an average, asymmetric matrices have most of the eigenvalues as complex-valued. This in turn indicates that most directed graphs have graph analytic signal with approximately ’one-sided’ spectrum. Similarly, we can conclude that, on an average, small-world graphs\(^1\) have most of the eigenvalues as complex-valued. This in turn means that the corresponding graphs have graph analytic signal spectrum.

### B. Discussions on graph analytic signal and graph Hilbert transform

We next show that the graph Hilbert transform \( x_h \) of a real graph signal \( x \) is real. Since \( x_h = \mathbf{V} \hat{x} \), we have that

\[
\hat{x} = \sum_{i \in \Gamma_2} \mathbf{J}_h(i) v_i + \sum_{i \in \Gamma_4} \hat{x}(i) v_i \tag{6}
\]

where \( v_i \) denotes the \( i \)-th column of \( \mathbf{V} \), and \( \Im(a) \) denotes the imaginary part \( a \). The third equality in (6) follows because the eigenvectors indexed by \( \Gamma_2 \) and \( \Gamma_4 \) form complex conjugates. Thus, \( \hat{x}_h \) is purely imaginary which in turn means that \( x = \Re(x_h) \), where \( \Re(a) \) denotes the real part of \( a \). We express (5) as

\[
x_h = \mathbf{J}_h \hat{x}_h \quad \text{or} \quad x_h = \mathbf{V} \mathbf{J}_h \mathbf{V}^{-1} x, \tag{7}
\]

where \( \mathbf{J}_h \) is the diagonal matrix with \( i \)-th diagonal element:

\[
\mathbf{J}_h(i) = \begin{cases} -j, & i \in \Gamma_2 \\ 0, & i \in \Gamma_1 \cup \Gamma_3 \\ j, & i \in \Gamma_4 \end{cases} \tag{8}
\]

**Proposition 1.** The graph Hilbert transform is a linear shift-invariant graph filtering operation.

**Proof.** From (7), we have that \( x_h = \mathbf{V} \mathbf{J}_h \mathbf{V}^{-1} x = \mathbf{H} x \), where \( \mathbf{H} = \mathbf{V} \mathbf{J}_h \mathbf{V}^{-1} \). By definition, graph filter \( \mathbf{H} \) is linear and shift-invariant if for any graph filter of the form \( M = \sum_{i=0}^{N} m_i \mathbf{A}^i \), \( M \leq N \) we have \( \mathbf{H} \mathbf{M} \mathbf{x} = \mathbf{M} \mathbf{H} \mathbf{x} \), which in turn means that \( \mathbf{H} \) should be a polynomial of \( \mathbf{M} \), or equivalently, of \( \mathbf{A} \). Since \( \mathbf{A} = \mathbf{V} \mathbf{J} \mathbf{V}^{-1} \), we have that \( \mathbf{M} = \mathbf{V} m(\mathbf{J}) \mathbf{V}^{-1} \). Let \( y \) denote the output of filter \( \mathbf{M} \) for the input \( x \): \( y = \mathbf{M} x \). Then, we have that \( \hat{y}_h = \mathbf{V}^{-1} y = \mathbf{V}^{-1} \mathbf{m}(\mathbf{J}) \hat{x} \). Since \( \hat{x}_h = \mathbf{J}_h \hat{x} \), we get that

\[
\hat{y}_h = \mathbf{J}_h \hat{y}_h = \mathbf{J}_h m(\mathbf{J}) \hat{x} = m(\mathbf{J}) \hat{x}_h, \tag{9}
\]

where we use the commutativity of the diagonal matrices \( m(\mathbf{J}) \) and \( \mathbf{J}_h \). Taking inverse GFT on both sides of (9), we get that

\[
y_h = \mathbf{H} m(\mathbf{A}) x = m(\mathbf{A}) x_h = m(\mathbf{A}) \mathbf{H} x.
\]

This completes the proof. \( \square \)

**Table II.** \( \theta \) value for some graphs of interest

| Graph                        | Number of real eigenvalues | \( \theta \)          |
|------------------------------|----------------------------|-----------------------|
| ID graph, odd \( N \)        | 1                         | \( \frac{N+1}{2} \)   |
| ID graph, even \( N \)       | 2                         | \( \frac{N+2}{2} \)   |
| Entries drawn \( \mathcal{N}(0,1) \) or \( \mathcal{U}(-1,1) \) or Bernoulli \( \{0,1\} \) | \( \sqrt{\frac{2N}{\pi}} \) (asymptotic expected value) | \( \frac{N+1}{2} + \sqrt{\frac{N}{2\pi}} \) |

**Fig. 2.** 1D signal graph (a), (b) Eigenvalues of A, (c) Signal and its graph Hilbert transform computed using \( L = N \) and \( L = N/2 \), (d) GFT spectrum of graph Hilbert transform.

The graph Hilbert transform being a shift-invariant filter means that there exists a polynomial \( h(x) = \sum_{i=0}^{L} h_i x^i \) such that \( \mathbf{H} = h(\mathbf{A}) \), where \( h_i \)'s are obtained by solving the following system of equations:

\[
h_0 + h_1 \lambda_1 + \cdots + h_L \lambda_1^L = 0, \quad i \in \Gamma_1 \cup \Gamma_3 \tag{10}
\]

The solution of (10) obtained by setting \( \mathbf{A} = \mathbf{C} \) and \( L = N \) is the impulse response of the discrete Hilbert transform. In order to avoid ill-conditioning of (10), \( L \) is usually restricted to be much less than \( N \), thereby leading to approximations of the ideal graph Hilbert transform.

\( ^2 \)The Erdős–Rényi model is a popular model used for generating small-world random graphs [57].
In Figure 2, we show the graph Hilbert transform computed using (10) for various values of $L$ for the 1D signal graph. We observe from Figure 2(d) that as $L$ is decreased, the spectrum of the graph Hilbert transform differs from the ideal case. This is because the corresponding columns of each graph shift ($A^r$) are linearly independent and restricting the number of taps restricts the dimension of the signal space.

C. Some properties of graph analytic signal/graph Hilbert transform

Let $A$, $I$, and $H$ denote the graph analytic signal, identity, and graph Hilbert transform operators, respectively, such that $x_a = A\{x\}$, $x_h = H\{x\}$ and $A = I + jH$. For $x = \eta f + \eta^* f^*$ such that $f \in V_2 \cup V_4$ and $\eta \in \mathbb{C}$, we have the following properties:

1) **Graph-shift invariance:** $H\{\alpha A x\} = \alpha A H\{x\}$, $\alpha \in \mathbb{C}$.

2) **Superposition:** For $x_1, x_2 \in V_2 \cup V_4$ and $\alpha, \beta \in \mathbb{C}$, we have that $H\{\alpha x_1 + \beta x_2\} = \alpha H\{x_1\} + \beta H\{x_2\}$.

---

Proof: From (7), we have that for $x = \alpha x_1 + \beta x_2$

\[
H\{x\} = VJ_hV^{-1}\{x\} = VJ_h\{\alpha V^{-1}x_1 + \beta V^{-1}x_2\} = VJ_h\{\alpha x_1 + \beta x_2\} = \alpha H\{x_1\} + \beta H\{x_2\},
\]

where $\hat{x}_{1,h}$ and $\hat{x}_{2,h}$ denote the GFT of $H\{x_1\}$ and $H\{x_2\}$, respectively.

3) **Phase-shifting action**\(^3\): For $i \in \Gamma_2 \cup \Gamma_4$:

\[
H\{\Re\{v_i\}\} = \Im\{v_i\}, \text{ and } H\{\Im\{v_i\}\} = -\Re\{v_i\}. \quad (11)
\]

---

Proof: Consider $i \in \Gamma_2$. Using the property 2, we have

\[
H\{2\Re\{v_i\}\} = H\{v_i + v_i^*\} = (jv_i - jv_i^*) = 2\Im\{v_i\} \\
H\{2\Im\{v_i\}\} = H\{v_i - v_i^*\} = (jv_i + jv_i^*) = 2\Re\{v_i\}.
\]

The proof for $i \in \Gamma_4$ follows similarly.

Equation (11) generalizes the quadrature phase-shifting action of the discrete Hilbert transform $H_c$ on sinusoids, since:

\[
H_c\{\cos(\omega_i n)\} = H_c\{\Re\{v_i\}\}\{n\} = \sin(\omega_i n) = 3\{v_i\}\{n\}, \\
H_c\{\sin(\omega_i n)\} = H_c\{\Im\{v_i\}\}\{n\} = -\cos(\omega_i n) = -3\{v_i\}\{n\},
\]

noting when $A = C$, $v_i = e^{j(\omega_i n)}$, where $\omega_i = \frac{2\pi(i-1)}{N}$.

4) **Inverse:** $H^2 = -I$ or, $H^{-1} = -H$.

---

Proof: From (7), we have $\hat{x}_h = J_h\hat{x}$. Hence, the GFT of $H^2\{x\}$ is given by $J_h^2\hat{x}$. We have from (8) that

\[
J_h^2\{i\} = J_h(i)J_h(i) = \begin{cases} 1 & , i \in \Gamma_2 \cup \Gamma_4 \\
-1 & , i \in \Gamma_1 \cup \Gamma_3 \\
0 & , \text{else}
\end{cases}
\]

which shows that $J_h^2\hat{x} = -\hat{x}$ for $x \in V_2 \cup V_4$. In other words, $H^2\{x\} = -x$ which completes the proof.

\(^3\)For simplicity, we use the same operator notation to denote the corresponding operation for both the signal seen as a vector and as a function of the node. For example, $H\{\cos(\omega_n)\}$ denotes the operator action directly on the function $\cos(\omega_n)$ evaluated at the $i$th node, whereas $H\{x\}$ denotes the vector that comes out of applying the graph Hilbert transform operation on the signal $x$. In the case when $x(n) = \cos(\omega_n)$, we have $H\{x\}(n) = H\{\cos(\omega_n)\}$.

5) **Repeated operation:** $A^2\{x\} = (I + jH)^2\{x\} = 2A\{x\}$ and $H^4\{x\} = x$. (Follows from Property 4).

6) **Isometry:** Since $\forall i$, $|\hat{x}_h(i)| = |\hat{x}(i)|$, we have that $|\hat{x}|_p = |\hat{x}_h|_p$, where $|\hat{x}|_p = (\sum |\hat{x}(i)|^p)^{\frac{1}{p}}$ is the $l_p$ norm of $x$, $p \geq 1$, assuming $x \in l^p(C)$. In particular, $|x_n|^2 = |\hat{x}_n|^2 = \frac{1}{2}||\hat{x}_n||^2$. If $V$ is unitary, $\|x_n\|^2 = \|\hat{x}_n\|^2 = \frac{1}{2}||x_n||^2$.

7) **Preservation of orthogonality:** If $V$ is unitary and $x_1, x_2$ are orthogonal, $(x_1, x_2) = 0$, then $(\hat{x}_1, \hat{x}_2) = 0$.

---

Proof: Since $V^{-1}$ is unitary, $(\hat{x}_1, x_2) = (x_1, x_2)$. Then we have that

\[
(H\{x_1\}, H\{x_2\}) = (V^{-1}\hat{x}_1, V^{-1}\hat{x}_2) = (J_h\hat{x}_1, J_h\hat{x}_2) = (\hat{x}_1, \hat{x}_2) = (x_1, x_2) = 0.
\]

Since $A = I + jH$, properties 1 and 2 are also satisfied by $A$. Properties 1 to 7 do not hold if $x$ has contribution from the subspaces $V_1$ or $V_3$ as they form the null-space of the graph Hilbert transform operator, that is, $H\{x\} = 0$ for $x \in V_1 \cup V_3$. We note here that the `one-sidedness’ and other properties of the graph analytic signal/graph Hilbert transform of a real signal $x$ are decided entirely by the underlying graph or equivalently, its adjacency matrix $A$. A real signal $x$ may have a graph analytic signal $x_a$ with a larger number of nonzero GFT coefficients in one graph than in another graph.

D. Graph Hilbert transform and highlighting of singularities/anomalies

The conventional Hilbert transform has been shown to be useful for highlighting singularities in 1D/2D signals [58], [59]. This is a consequence of the functional form of impulse response of the Hilbert transform. Since the graph Hilbert transform generalizes the discrete Hilbert transform, our hypothesis is that the graph Hilbert transform also highlights singularities. We have already seen how the conventional 1D-Hilbert transform, as a special case of the graph Hilbert transform when $A = C$ highlights edges or anomalies (cf. Figure 2). We next consider a 40 x 40 2D signal or image signal. The image is a section of the coins image taken from the MATLAB library. Since there is no unique directed graph for an image signal, we define the graph as an extension from the 1D-setting, that is, we consider that $j$th pixel in $i$th row to be connected to the $(j+1)$th pixel in the same row and to the $j$th pixel in the $(i+1)$th row. The corresponding graph then has the adjacency matrix $A = C \otimes C$ as shown in Figure 3(a).

The image signal and its graph Hilbert transform (reshaped as an image) are shown in Figures 3(b) and 3(c), respectively. We observe that the graph Hilbert transform specialized to the 2D signal case exhibits edge highlighting behavior. We note here that connecting the pixels differently leads to alternative directed graphs, and we find in our experiments that the corresponding graph Hilbert transforms also highlight edges. However, all these cases are not reported here to avoid repetition.

We next consider a synthesized social network graph consisting of 10 communities with 6 member nodes each. The nodes within each community are strongly connected in addition to having inter-community edges. The intra-community
edge-weights are drawn from the uniform distribution over $[0, 1]$, and the inter-community edge-weights are drawn from uniform distribution over $[0, 0.5]$, and randomly placed across nodes from different communities (Note that the resulting graph is highly assymmetric). Graphs with real edge weights have been extensively employed in analyzing data occurring in many practical applications such as road traffic analysis, brain connectivity [3]. We consider the case of weighted random graphs to demonstrate the potential of our concepts to such application areas. We normalize $A$ have $|\lambda|_{\text{max}} = 1$. The nodes are labelled to correspond to the row index of the adjacency matrix. We consider two different cases, one with few inter-community edges ($1\%$ of the total number of possible edges in the graph) and the other with denser edges ($10\%$ percent of the total edges possible in the graph). For each case, we compute the graph Hilbert transform using (5) for the graph signal which is zero everywhere except at nodes 18 to 23 (which lie in communities 3 and 4) being active. By intuition, we expect all the nodes connected to these nodes which have value zero (thus making a singularity or anomaly) to be highlighted by the graph Hilbert transform. We observe from Figure 4(b) that this is indeed the case. The graph Hilbert transform takes large values at nodes 15 and 16 since they are strongly connected to node 18 (cf. Figures 4(a)-(b)). Similarly, presence of strong edge between nodes 18 to 50 results in node 50 being highlighted by the graph Hilbert transform. Similar arguments can be made for nodes 2, 3, and 55, all of which are highlighted by the graph Hilbert transform. We also note that the extent to which a node is highlighted also varies with the strength of the connecting edge. In the case of dense inter-community edges, we observe that the graph Hilbert transform highlights a large number of nodes since the nodes from 18 to 23 are connected to many nodes (cf. Figures 4(c)-(d)).

In Figure 5, we consider an unweighted community graph with very few inter community edges. The graph consists of 5 communities of 10 nodes each. Each 10-node community subgraph is randomly generated from the Erdős Rényi model with an edge probability $p = 0.5$. The communities are then connected with very few links also generated randomly. The resulting adjacency matrix is shown in Figure 5(a). We consider the signal to be all ones corresponding to nodes of community 3. In the present example, community 3 has only one incoming edge from community 4 (from node 38 to node 25), highlighted by the circle in Figure 5(a). We observe that the graph Hilbert transform highlights both nodes 38 and 26,
as expected.

Our experiments suggest that the graph Hilbert transform could be potentially used in anomaly/edge detection in graphs.

**E. On the graph Hilbert transform and the Bedrosian property**

We show a limitation of the graph Hilbert transform in this section. One of the important properties possessed by the conventional Hilbert transform (and its fractional versions [59]) is that it obeys the Bedrosian property, that is, if \( f \) and \( g \) are two signals with disjoint Fourier spectra (DFT or discrete-time Fourier transform, such that \( f \) is low-pass and \( g \) is high-pass, then we have that the discrete-Hilbert transform \( \mathcal{H}_c \) satisfies

\[
\mathcal{H}_c\{f(n)g(n)\} = f(n)\mathcal{H}_c\{g(n)\}. \quad \forall n
\]

As we show through experiments next, the graph Hilbert transform of a general graph does not possess the Bedrosian property. In our opinion, there are two important factors for this limitation. First, the point-wise product in the node domain does not correspond to a graph-frequency domain convolution (no definition for frequency domain convolution for graphs exists currently). Second, unlike the 1D case where the DFT basis is functionally related to the frequency (as the entries exists currently). Second, unlike the 1D case where the DFT column are given by complex exponentiation of the \( i \)th frequency to different powers), the GFT of a general graph usually is not related to its eigenvalues through analytical expressions. We demonstrate this by considering the jittered 1D signal modeled as a graph signal with

\[
\begin{pmatrix}
0 & w_1 & 0 & \cdots & 0 \\
0 & 0 & w_2 & \cdots & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
w_N & 0 & 0 & \cdots & 0
\end{pmatrix},
\]

where \( w_i \) denotes the normalized spacing between the \( i \)th and \((i + 1)\)th samples. For uniformly sampled 1D signal, \( w_i = 1 \) for all \( i \). Then from the Bedrosian property we have that

\[
\mathcal{H}_c\{\cos(\omega_i n) \cos(\omega_j n)\} = \cos(\omega_i n) \sin(\omega_j n), \quad \text{for } \omega_j > \omega_i
\]

which when expressed in terms of the corresponding DFT (GFT) vectors \( \mathbf{v}_i(n) = e^{j\omega_i n} \) becomes

\[
\mathcal{H}\{\Re(\mathbf{v}_i) \cdot \Re(\mathbf{v}_j)\} = \Re(\mathbf{v}_i) \cdot \Im(\mathbf{v}_j), \tag{12}
\]

where \( \Re \cdot \Im \) denote the vector obtained by component-wise products of \( \Re \) and \( \Im \). We test the validity of the Bedrosian property by computing the graph Hilbert transform of the signal \( \Re(\mathbf{v}_i) \cdot \Re(\mathbf{v}_j) \) where \( i \) and \( j \) correspond to the low and high frequency GFT basis vectors, respectively (Here we use the frequency-ordering proposed in [8] based on graph total-variation \(|x - Ax|_1\)). We compare the graph Hilbert transform of \( \Re(\mathbf{v}_i) \cdot \Re(\mathbf{v}_j) \) with \( \Re(\mathbf{v}_i) \cdot \Im(\mathbf{v}_j) \). We consider the low-jitter case \( w_i = 1 + d_i, d_i \) is drawn independently from the Gaussian distribution \( \mathcal{N}(0, 0.01) \). Repeating the experiment multiple times, we observe that the left and right hand sides of (12) almost never coincide.

In Figure 6, we consider a particular realization for \( x = \Re(\mathbf{v}_3) \cdot \Re(\mathbf{v}_{57}) \), where \( N = 100 \). We observe that the graph Hilbert transform does not coincide with either \( \Re(\mathbf{v}_3) \cdot \Im(\mathbf{v}_{57}) \) or \( \Im(\mathbf{v}_3) \cdot \Re(\mathbf{v}_{57}) \). This shows that the graph Hilbert transform does not possess a Bedrosian property even for graphs similar to the 1D graph.

**IV. THE GRAPH ANALYTIC SIGNAL AND MODULATION ANALYSIS**

The concept of analytic signal is used extensively in the demodulation of amplitude-modulated frequency-modulated signals [60]–[64]. Modulation analysis decomposes a signal into two components: one varying smoothly, capturing the average information in the signal (referred to as the AM), and the second, capturing the finer variations (referred to as the phase or frequency modulation (PM or FM)). Most demodulation techniques involve the construction of the analytic
signal, implicitly or explicitly. Motivated by 1D modulation definitions [60], [61], we next propose AM and PM for graph signals:

**Definition.** The AM $A_{x,v}$ and PM $\phi_{x,v}$ of a graph signal $x$ are defined as the magnitude and phase angle of the graph analytic signal, respectively:

$$A_{x,v}(i) = |x_a(i)|, \quad \forall i \in \{1, 2, \cdots, N\}$$

$$\phi_{x,v}(i) = \arg(x_a(i)),$$

where $\arg(\cdot)$ denotes the 4-quadrant arctangent function which takes values in the range $(-\pi, \pi]$.

For $A = C$, (13) reduces to 1D AM and PM definitions. This is because setting $A = C$ results in the graph analytic signal to coincide with the conventional analytic signal as we have discussed in Section III. This in turn implies that the amplitude and phase of the graph analytic signal also coincide with the conventional definitions. We hereafter refer to $A_{x,v}$ as the graph AM and $\phi_{x,v}$ as the graph PM. We next discuss computing unwrapped phase and frequency-modulation for the graph signal.

### A. Phase unwrapping and frequency modulation

The phase of complex signals with $\arg(\cdot)$ function returns phase values wrapped in the range $(-\pi, \pi]$. In practice, it is more convenient to work with unwrapped phase functions [65]. For the case of 1D signals, the phase-unwrapping is performed by keeping the causality in mind: unwrapping begins from the first sample, successively compensating for discontinuities in the phase [52], [66]. Phase-unwrapping algorithms for high-dimensional signals are also based on phase-discontinuity compensation though the exact strategy depends on the application [67], [68]. The unwrapping of the phase of the graph analytic signal poses additional challenges due to the connected nature of the signal. Since each node of a general graph may be connected to multiple nodes, it is desirable to unwrap the phase incorporating the connectivity information. We next propose a potential approach for unwrapping the phase of the graph analytic signal.

Let $A(i,j)$ denote the $(i,j)$th entry of $A$. Starting from node 1, we search for the node connected to 1 with the maximum edge-weight magnitude. Let us denote this node by $2'$. We find the next node $3'$ most strongly connected to $2'$, excluding node 1, and continue till all the nodes are numbered to obtain the sequence $\{1, 2', \cdots, N\}$, assuming that it is possible to traverse all the nodes in the graph. We construct the new phase sequence $\phi'_{x,v}(i) = \phi_{x,v}(i')$, to which we apply standard 1D phase-unwrapping to obtain $\phi''_{x,v}(i)$. Algorithm 1 describes the steps involved in the process. In the case when multiple nodes connected to the current node have equal edge-weights, we break the tie arbitrarily. Our motivation for Algorithm 1 is the observation that for signals that are smooth over the graph, nodes connected through strong edges take similar signal values [3], [8]. We note that our algorithm requires the graph to be weakly connected in addition to the assumptions considered in Section I. A directed graph is said to be weakly connected if the undirected underlying graph obtained by replacing all directed edges of the graph with undirected edges is a connected graph [69]. Algorithm 1 implicitly looks for a path connecting all the nodes in the network and is hence, similar to the travelling salesman problem for asymmetric graphs [70]. Our choice of ordering the nodes for unwrapping is by no means exhaustive and alternative definitions could be proposed.

We next define frequency modulation for graph signals.

**Definition (Frequency modulation).** The frequency modulation (FM) of a graph signal $x$ is defined as

$$\omega_{x,v} = \phi''_{x,v} - A_{x,v} \phi_{x,v},$$

where $\phi''_{x,v}$ denotes the unwrapped phase of the graph analytic signal.

The definition generalizes the backward-difference operator used to compute FM for 1D signals [52]. We assume that $A$ is normalized such that $|A|_{max} = 1$. In order to visualize the proposed graph AM and graph FM, we consider speech signal viewed as a graph signal using the linear prediction coefficients as proposed in [12]. For each speech frame, we construct $A$ by connecting every sample to its preceding $P$ samples with edge-weights equal to the corresponding $P$th-order linear prediction coefficients. We plot the obtained graph AM and graph FM in Figure 7. We also include the 1D AM and FM for comparison. We observe that the graph FM is smoother than 1D FM, and the graph AM and 1D AM nearly coincide.

### V. GRAPH AMPLITUDE IN A CONVEX FRAMEWORK

In defining an envelope, one of the desirable properties is smoothness over the graph and capturing the slow varying or average information of the signal. However, the graph AM does not necessarily satisfy these properties. The conventional 1D AM is also known to not satisfy these properties, and hence several alternative definitions for AM exist [60]. Given a graph signal $x$, we seek a graph envelope $\chi$, with low MS [$(\chi_c)$] while faithfully representing the average signal characteristics. By formulating the problem as a convex optimization problem,
we propose the following definition for the convex graph envelope $x_c$:

$$ x_c = \arg \min \| x_e - Ax_e \|^2 + \alpha \| x_e \|^2 \quad \text{s.t.} \quad |x| \leq |x_e|. \quad (14) $$

where the inequality $\leq$ is taken in the element-wise sense. The term $\alpha \| x_e \|^2$ ensures that the envelope remains bounded and closely follows the signal, particularly when the signal is close to zero. The optimization problem (14) is convex and has a global optimum. A similar approach has been used in demodulation of acoustic signals by Sell and Slaney [72].

We note here that unlike the graph AM proposed in Section IV, the convex graph envelope definition (14) is applicable to both directed and undirected graphs. In Figures 8 and 9, we show the comparison of convex graph envelope (with $\alpha = 0.1$) and the graph AM computed for a community graph. The adjacency matrix is constructed in the same manner as for the community graph example considered in Section III-D. The graph signal consists of two active communities with varying opinions within the community in addition to additive white Gaussian noise present at all nodes. We observe that convex graph envelope is a more intuitive envelope than the graph AM since the graph AM does not always follow the signal closely and takes large values even over nodes which are not active. This is expected from the discussion on anomaly highlighting property of the graph Hilbert transform in Section III-D: the graph Hilbert transform and hence, the graph AM highlights all the edges connected to the active nodes, which explains the graph AM taking large values even at nodes where the signal is close to zero. In contrast, the convex graph envelope gives a more realistic picture of the average activity of communities over the graph, which may be further used to decide on which communities are active. As expected, we obtain the lowest $MS_g$ with the convex graph envelope (cf. Table in Figure 9(b)).

We next return to the case of the jittered 1D signal described in Section III-E and compare the envelopes given by the graph AM and convex graph envelope. We take the graph signal as the real part of $10$th eigenvector of $A$ of size $N = 100$. The choice of the signal is only to ensure that the signal is related to the underlying graph and not arbitrary. We compute the graph AM, convex graph envelope (with $\alpha = 1$), and the AM obtained from the conventional analytic signal (which we denote as 1D AM). In our experiments, we find that both graph AM and convex graph envelope are suitable choices for the signal envelope as they fit the signal more closely, preserving the onset and tail decay characteristics, whereas the 1D AM tends to smear the onset and decay. In Figure 10, we show the results for a particular realization corresponding to the tenth eigenvector.

We also observe from Table III that the $MS_g$ is least for the convex graph envelope as expected.
Table IV

| Number of hidden neurons | Classifier 1 (DFT) | Classifier 2 (GFT) | Classifier 3 (DFT+GFT) |
|--------------------------|--------------------|--------------------|-------------------------|
| 1                        | 69.4               | 59.7               | 69.6                    |
| 5                        | 70.8               | 59.7               | 72.6                    |
| 10                       | 71.0               | 59.3               | 72.2                    |

**VI. EXPERIMENTS**

We next illustrate the applications of the proposed concepts on few real-world signal examples: for speaker classification and anomaly detection over graphs.

**A. Male-female voice classification**

We consider the speech signal as a signal over a graph. Our hypothesis is that viewing the speech signal as a graph signal provides additional information that could help improve the speaker recognition performance. In order to test our hypothesis, we construct a speech graph from learning set data consisting of speech samples from two speakers. We then compute the conventional AM and FM, and the graph AM and FM and use them features for classification. We use two-layer neural network classifiers trained from data distinct from test and learning data. Let \( X_1 = [x_{1,1}, \cdots, x_{1,n}] \) and \( X_2 = [x_{2,1}, \cdots, x_{2,n}] \) denote the speech sample matrices from two speakers S1 and S2 such that \( x_{i,j} \in \mathbb{R}^N \) denotes the \( j \)th frame of speech samples from \( i \)th speaker. The speech frames are taken from different sentences uttered by the speakers (one male and other female) from the CMU Arctic Database [73]. We choose a frame-length of \( N = 50 \) and total number of frames 4000 (where \( n = 2000 \)) such that \( X_t = [X_1, X_2] \). We compute the adjacency matrix by solving the following optimization problem:

\[
A^* = \arg \min_A \|X_t - AX_t\|_2^2 \\
\text{subject to } \text{diag}(A) = 0, A1 = 1, AT1 = 1. \tag{15}
\]

We use the constraints \( A1 = 1, A^T1 = 1 \) to avoid ill-conditioning in case of insufficient learning data. In Figure 11, we plot the adjacency matrix obtained from (15). We consider three different classifiers:

**Classifier 1**: The classifier uses magnitudes of the DFT coefficients of 1D AM and FM as feature vectors. The feature vector has length 100.

**Classifier 2**: The classifier uses the magnitudes of the GFT of the graph AM and FM as feature vectors, where the GFT is obtained from the eigen-decomposition of \( A^* \). The feature vector has length 100.

**Classifier 3**: The classifier uses the magnitudes of the DFT coefficients of 1D AM and FM, concatenated with magnitudes of the GFT of the graph AM and FM as feature vectors. The length of the feature vector is equal to 200.

The classifiers are trained using the features from training data \( X_tr \) and tested on \( X_test \), both data sets being different from \( X_t \) used in computing the adjacency matrix. The composite dataset \([X_tr, X_test]\) consists of 50000 samples of which 60\% is \( X_tr \) and the rest in \( X_test \). The classifier performance is computed for different number of sigmoidal neurons in the hidden layer. The performance is averaged over 50 runs where the training and test data are randomly partitioned. We observe from Table IV that Classifier 3 outperforms the other classifiers with a classification improvement of up to 2\% in comparison with the DFT-based classifier. We also observe that the performance of the neural classifier saturates after 5 hidden neurons. This shows that the proposed graph amplitude and frequency modulations improve speaker classification performance, and that viewing speech as a graph signal indeed provides complementary information.

**B. Anomaly detection over graphs**

We next consider the application of the convex graph envelope in anomaly detection over graphs. Anomalies may occur either at isolated nodes or over a connected subset of the network. Our hypothesis is that convex graph envelope can be used to indicate presence of anomalies in a connected network. We test our hypothesis by considering two different graphs. We consider first the undirected Minnesota road network graph with binary (one or zero) edge-weights. We generate realistic anomalies by randomly selecting a subset \( \Omega \) of nodes and setting the value of the graph signal \( x \) to 5 at those nodes. We further add white Gaussian noise \( v \) of unit variance to simulate normal traffic conditions. Thus, we have two situations \( H_0 \) (normal traffic) and \( H_1 \) (traffic anomalies) such that:

\[
H_1 : \quad x = 5.1\Omega + v \\
H_0 : \quad x = v,
\]

where \( \Omega \) denotes the indicator function for the subset \( \Omega \subset \mathcal{V} \) of nodes where the anomaly occurs, and \( v \) denotes a vector drawn from a correlated multivariate Gaussian distribution with unit mean. We propose the metric \( S_\alpha \) for anomaly detection:

\[
S_\alpha = \sum_{i=1}^{K} \tilde{x}(i)^2,
\]
where \( K = N/10 \), the GFT coefficients being frequency-ordered with the lowest corresponding to the index 1, motivated by the observation that significant energy of the signal is contained in the lower graph frequencies. Since \( N = 2642 \), we have \( K = N/10 = 264 \). Our motivation to define \( S_a \) metric comes from the recently proposed Graph Fourier Scan Statistic for undirected graphs [74]. We compute \( S_a \) for the signal under both \( H_0 \) and \( H_1 \) for various realizations. We also compute \( S_a \) for the convex graph envelope of the graph signal. The computed values of \( S_a \) for both the signal \( x \) and its convex graph envelope for 200 realizations are shown in Figure 12. The first 100 realizations correspond to \( H_0 \) and the remaining correspond to \( H_1 \). The red line indicates the mean value of \( S_a \) for the two hypotheses. We observe that \( S_a \) of the convex graph envelope shows better discrimination power in comparison with \( S_a \) of the signal. This is explained by noting that a graph signal with isolated anomalies results in a GFT spectrum that is nearly uniformly distributed over the graph frequencies. In contrast, the convex graph envelope by definition results in a signal that activates connected neighbors of the isolated anomaly giving rise to localized blocks with nodes that have similar value. This in turn results in the energy of the convex graph envelope to be concentrated in the low frequency region. We show an example realization in Figure 12 (c)-(d). In the case of anomalies that activate a connected region, that is, when \( \Omega \) contains connected subsets, the signal as well as its convex graph envelope have comparable smoothness and hence, one does not expect significant difference in the discriminative power of \( S_a \). Our experiments show that this is indeed the case. In Figure 12 (e)-(f), we show \( S_a \) for the signal and its convex graph envelope for anomalies that consist of connected sets with 5-hop neighbors active. We next repeat the experiment for the synthesized directed community graph described corresponding to Figure 8. The values of \( S_a \) computed for different realizations of \( H_0 \) and \( H_1 \) for the directed community graph is shown in Figures 13. We observe again that the \( S_a \) of the convex graph envelope exhibits discriminative power in detecting presence of anomalies in the graph, showing that the convex graph envelope is suitable for detection of anomalies over the graph.

**VII. DISCUSSIONS AND CONCLUSIONS**

We proposed definitions for the analytic signal and Hilbert transform of real graph signals over directed graphs. We showed that graph Hilbert transform and graph analytic signal are linear and shift-invariant over graphs, and that they inherit many properties such as shift-invariance, isometry, and phase-shifting from their conventional counterparts. We showed that for some graphs of interest the graph Hilbert transform highlights singularities/anomalies. We also demonstrated through numerical example that the graph Hilbert transform does not inherit the Bedrosian property in general. Using the graph analytic signal, we defined amplitude, phase, and frequency modulations for graph signals, and proposed an approach for unwrapping the phase of graph analytic signals. Motivated by the need to obtain an envelope minimizing a graph smoothness measure, we defined a graph envelope using convex optimization. We observed that the convex graph envelope is often a more intuitive choice for the envelope than that obtained through the graph analytic signal. In addition, the convex formulation is applicable to both directed and undirected graphs. In order to illustrate the proposed notions, we considered their application to synthesized and real-world signal examples. We observed in the context of speaker recognition that viewing the speech signal as a graph signal resulted in improved classification performance. This is because the graph signal model captures signal correlation across all samples in a speech frame, unlike the 1D graph which considers only the preceding sample. We also observed that both the graph Hilbert transform and the convex graph envelope exhibit the potential to detect anomalies over the graph. They are therefore of much relevance in analyzing scenarios such as malfunctioning in power grids, spread of
disease or epidemics, identifying activity sources in the brain, and traffic bottlenecks over transportation networks, and study of outliers in social network trends.

We note that applications chosen in this article serve the purpose of illustrating the proposed concepts and are by no means exhaustive. As with 1D modulation analysis, the utility varies across applications and can only be revealed by detailed analysis on various datasets. It is known that complex signal analysis often results in complementary information over real signal analysis, as in the case of analytic or complex wavelets [75]. It would be interesting to investigate if complex graph signals could be similarly used to arrive at unique insights into signal analysis over graphs.

REFERENCES

[1] M. E. J. Newman, Networks: An Introduction. Oxford University Press, 2010.
[2] A. Sandryhaila and J. M. F. Moura, “Discrete signal processing on graphs: Frequency analysis,” IEEE Trans. Signal Process., vol. 62, no. 12, pp. 3042–3054, 2014.
[3] D. I. Shuman, S. Narang, S. Press, A. Ortega, and P. Vanderheyst, “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, 2013.
[4] O. Teke and P. P. Vaidyanathan, “Extending classical multirate signal processing theory to graphs—part i: Fundamentals,” IEEE Trans. Signal Process., vol. 65, no. 2, pp. 409–422, Jan 2017.
[5] ——, “Extending classical multirate signal processing theory to graphs—part ii: M-channel filter banks,” IEEE Trans. Signal Process., vol. 65, no. 2, pp. 423–437, Jan 2017.
[6] R. R. Coifman and M. Maggioni, “Diffusion wavelets,” Proc. Adv. Neural Inform. Process. Syst., pp. 998–1006, 2013.
[7] M. Jansen, C. P. Nason, and B. W. Silverman, “Multiscale methods for data on graphs and irregular multidimensional situations,” J. Roy. Statist. Soc.: Series B, vol. 71, no. 1, pp. 97–125, 2009.
[8] A. Sandryhaila and J. M. F. Moura, “Discrete signal processing on graphs,” IEEE Trans. Signal Process., vol. 61, no. 7, pp. 1644–1656, 2013.
[9] ——, “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, 2014.
[10] L. Cohen, Time-frequency analysis: Theory and applications. Upper Saddle River, NJ, USA: Prentice-Hall, Inc., 1995.
[13] D. I. Shuman, B. Ricaud, and P. Vanderheyst, “A windowed graph Fourier transform,” IEEE Statist. Signal Process. Workshop (SSP), pp. 133–136, Aug 2012.
[14] S. K. Narang and A. Ortega, “Local two-channel critically sampled filter-banks on graphs,” Proc. IEEE Int. Conf. Image Process. (ICIP), pp. 333–336, 2010.
[15] ——, “Compact support biorthogonal wavelet filterbanks for arbitrary undirected graphs,” IEEE Trans. Signal Process., vol. 61, no. 19, pp. 4673–4685, 2013.
[16] ——, “Perfect reconstruction two-channel wavelet filter banks for graph structured data,” IEEE Trans. Signal Process., vol. 60, no. 6, pp. 2786–2799, 2012.
[17] R. R. Coifman and M. Maggioni, “Diffusion wavelets,” Appl. Comput. Harmonic Anal., vol. 21, no. 1, pp. 53–94, 2006.
[18] D. Ganesan, B. Greenstein, D. Estrin, J. Heidemann, and R. Govindan, “Multiresolution storage and search in sensor networks,” ACM Trans. Storage, vol. 1, no. 3, pp. 277–315, 2005.
[19] D. K. Hammond, P. Vanderheyst, and R. Grinov, “Wavelets on graphs via spectral graph theory,” Appl. Comput. Harmonic Anal., vol. 30, no. 2, pp. 129–150, 2011.
[20] R. Wagner, V. Delouille, and R. Baraniuk, “Distributed wavelet denoising for sensor networks,” Proc. 45th IEEE Conf. Decision Control, pp. 373–379, 2006.
[21] F. Gama, A. G. Marques, G. Mateos, and A. Ribeiro, “Rethinking sketching as sampling: Linear transforms of graph signals,” Proc. Asilomar Conf. Signals, Systems Comput., pp. 522–526, Nov 2016.
[22] S. Chen, R. Varma, A. Singh, and J. Kovačević, “Signal recovery on graphs: Fundamental limits of sampling strategies,” IEEE Trans. Signal Inform. Process. over Networks, vol. 2, no. 4, pp. 539–554, Dec 2016.
[23] A. G. Marques, S. Segarra, G. Leus, and A. Ribeiro, “Sampling of graph signals with successive local aggregations,” IEEE Trans. Signal Process., vol. 64, no. 7, pp. 1832–1843, April 2016.
[24] H. Q. Nguyen and M. N. Do, “Downsampling of signals on graphs via maximum spanning trees,” IEEE Trans. Signal Process., vol. 63, no. 1, pp. 182–191, Jan 2015.
[25] D. Tsiatsios, S. Barbashina, and P. D. Lorenzo, “Graphs on graphs: Uncertainty principle and sampling,” IEEE Trans. Signal Processing, vol. 64, no. 18, pp. 4845–4860, 2016.
[26] B. Girault, “Stationary graph signals using an isometric graph translation,” Proc. Eur. Signal Process. Conf. (EUSIPCO), pp. 1516–1520, Aug 2015.
[27] S. Segarra, A. G. Marques, G. Leus, and A. Ribeiro, “Stationary graph processes: Nonparametric spectral estimation,” Proc. IEEE Sensor Array Multichannel Signal Process. Workshop (SAM), pp. 1–5, July 2016.
[28] N. Perraudin and P. Vanderheyst, “Stationary signal processing on graphs,” CoRR, vol. abs/1601.02522, 2016. [Online]. Available: http://arxiv.org/abs/1601.02522
[29] P. G. B. Girault and E. Fleury, “Translation on graphs: An isometric shift operator,” Signal Process. Lett., vol. 22, no. 12, pp. 2416–2420, Dec. 2015.
[30] D. Thanou, D. I Shuman, and P. Frossard, “Learning parametric dictionaries for signals on graphs,” IEEE Trans. Signal Process., vol. 62, no. 15, pp. 3849–3862, 2014.
[31] D. I. Shuman, B. Ricaud, and P. Vanderheyst, “Vertex-frequency analysis on graphs,” Appl. Comput. Harmonic Anal., vol. 40, no. 2, pp. 260 – 291, 2016.
[32] N. Shahid, V. Kalofolias, X. Bresson, M. Bronstein, and P. Vanderheyst, “Robust principal component analysis on graphs,” IEEE Int. Conf. Comput. Vision (ICCV), pp. 2812 – 2820, 2015.
[33] N. Shahid, N. Perraudin, V. Kalofolias, G. Puy, and P. Vanderheyst, “Fast robust PCA on graphs,” IEEE J. Select. Topics Signal Process., vol. 10, no. 4, pp. 740–756, June 2016.
[34] N. Tremblay, G. Puy, P. Borgnat, R. Gribonval, and P. Vandergheynst, “Accelerated spectral clustering using graph filtering of random signals,” IEEE Int. Conf. Acoust. Speech Signal Process., 2016.
[35] K. Benzi, V. Kalofolias, X. Bresson, and P. Vandergheyst, “Song recommendation with non-negative matrix factorization and graph total variation,” Proc. IEEE Int. Conf. Acoust. Speech Signal Process., pp. 2439–2443, 2016.
[36] D. I. Shuman, M. J. Faraji, and P. Vanderheyst, “A multiscale pyramid transform for graph signals,” IEEE Trans. Signal Process., vol. 64, no. 8, pp. 2119–2134, April 2016.
[37] S. Segarra, A. G. Marques, G. Mateos, and A. Ribeiro, “Blind identification of graph filters with multiple sparse inputs,” IEEE Int. Conf. Acoust. Speech Signal Process. (ICASSP), pp. 4099–4103, March 2016.
[38] S. Segarra, G. Mateos, A. G. Marques, and A. Ribeiro, “Blind identification of graph filters,” IEEE Trans. Signal Process., vol. 65, no. 5, pp. 1146–1159, Mar. 2017.
[39] S. Chen, A. Sandryhaila, J. M. F. Moura, and J. Kovacevic, “Signal recovery on graphs: Variation minimization,” IEEE Trans. Signal Process., vol. 63, no. 17, pp. 4609–4624, Sept 2015.
[40] S. Chen, R. Varma, A. Sandryhaila, and J. Kovacevic, “Discrete signal processing on graphs: Sampling theory,” IEEE Trans. Signal Process., vol. 63, no. 4, pp. 6510–6523, Dec. 2015.
[41] A. Sakiyama, K. Watanabe, and Y. Tanaka, “Spectral graph wavelets and graph total variation,” Proc. IEEE Int. Conf. Image Process. (ICIP), pp. 3268–3271, 2015.
[42] F. Gama, A. G. Marques, G. Mateos, and A. Ribeiro, “Rethinking sketching as sampling: Linear transforms of graph signals,” Proc. Asilomar Conf. Signals, Systems Comput., pp. 522–526, Nov 2016.
13

[46] ——, “Estimating signals over graphs via multi-kernel learning,” in IEEE Statist. Signal Process. Workshop (SSP), June 2016, pp. 1–5.
[47] A. Venkitaraman, S. Chatterjee, and P. Handel, “On Hilbert transform of signals on graphs,” Proc. Sampling Theory Appl., 2015. [Online]. Available: http://w.american.edu/cas/sampta/papers/a13-venkitaraman.pdf
[48] D. Gabor, “Theory of communication,” J. Inst. Elec. Eng., vol. 93, 1946.
[49] R. M. Gray, “Toeplitz and circulant matrices: A review,” vol. 2, no. 3, pp. 155–239, 2006.
[50] A. Venkitaraman, S. Chatterjee, and P. Handel, “On Hilbert transform of signals on graphs,” Proc. Sampling Theory Appl., 2015. [Online]. Available: http://w.american.edu/cas/sampta/papers/a13-venkitaraman.pdf
[51] D. I. Shuman, P. Vandergheynst, and P. Frossard, “Chebyshev polynomial approximation for distributed signal processing,” in 2011 Int. Conf. Distributed Comput. Sensor Syst. Workshops (DCOSS), June 2011, pp. 1–8.
[52] A. V. Oppenheim and R. W. Schafer, Discrete-Time Signal Processing, 3rd ed. Upper Saddle River, NJ, USA: Prentice Hall Press, 2009.
[53] B. Gold, A. V. Oppenheim, and C. M. Rader, “Theory and implementation of the discrete Hilbert transform,” Symp. Comput. Process. Commun., Polytechnic Institute of Brooklyn, Tech. Rep., 1969.
[54] A. Edelman, E. Kostlan, and M. Shub, “How many eigenvalues of a random matrix are real?” J. Amer. Math. Soc., vol. 7, no. 1, 1994.
[55] V. L. Girko, “The strong circular law. Twenty years later. ii,” Random Oper. Stochastic Equations, vol. 12, no. 3, pp. 255–312, 2004.
[56] D. I. Shuman, P. Vandergheynst, and P. Frossard, “Chebyshev polynomial approximation for distributed signal processing,” in 2011 Int. Conf. Distributed Comput. Sensor Syst. Workshops (DCOSS), June 2011, pp. 1–8.
[57] A. V. Oppenheim and R. W. Schafer, Discrete-Time Signal Processing, 3rd ed. Upper Saddle River, NJ, USA: Prentice Hall Press, 2009.
[58] B. Gold, A. V. Oppenheim, and C. M. Rader, “Theory and implementation of the discrete Hilbert transform,” Symp. Comput. Process. Commun., Polytechnic Institute of Brooklyn, Tech. Rep., 1969.
[59] A. Edelman, E. Kostlan, and M. Shub, “How many eigenvalues of a random matrix are real?” J. Amer. Math. Soc., vol. 7, no. 1, 1994.
[60] V. L. Girko, “The strong circular law. Twenty years later. ii,” Random Oper. Stochastic Equations, vol. 12, no. 3, pp. 255–312, 2004.
[61] T. Tao, V. Vu, and M. Krishnapur, “Random matrices: Universality of esds and the circular law,” Ann. Probab., vol. 38, no. 5, pp. 2023–2065, 2010.
[62] R. Cohen and S. Havlin, Complex Networks: Structure, Robustness and Function. Cambridge University Press, 2010.
[63] A. W. Lohmann, D. Mendlovic, and Z. Zalevsky, “Fractional Hilbert transform,” Opt. Lett., vol. 24, no. 4, pp. 1532–1550, 1993.
[64] V. P. Stokes and P. Handel, “Comments on “On amplitude and frequency demodulation using energy operators,” IEEE Trans. Signal Process., vol. 41, no. 4, pp. 1532–1550, 1993.
[65] J. Kominek and T. Volgenant, “A technique to compute smooth amplitude, phase, and frequency modulations from the analytic signal,” IEEE Signal Process. Lett., vol. 19, no. 10, pp. 623–626, 2012.
[66] J. Tribot, “A new phase unwrapping algorithm,” IEEE Trans. Acoust., Speech, Signal Process., vol. 25, no. 2, pp. 170–177, 1977.
[67] W. Schwarzkopf, T. Milner, J. Ghosh, B. Evans, and A. Bovik, “Two-dimensional phase unwrapping using neural networks,” Proc. 4th IEEE Southwest Symp. Image Anal. Interpretation, pp. 274–277, 2000.
[68] R. Sivley and J. Havlicek, “Multidimensional phase unwrapping for consistent APF estimation,” Proc. IEEE Int. Conf. Image Process., vol. 2, no. 4, pp. 58–61, 2005.
[69] J. Bang-Jensen and G. Z. Gutin, Digraphs: Theory, Algorithms and Applications. Springer-Verlag London, 2009.
[70] R. Jonker and T. Volgenant, “Transforming asymmetric into symmetric traveling salesman problems,” Operations Research Lett., vol. 2, no. 4, pp. 161 – 163, 1983.
[71] North Texas vowel database. [Online]. Available: http://www.utdallas.edu/~assmann/KIDVOW1/North_Texas_vowel_database.html
[72] G. Sell and M. Slaney, “Solving demodulation as an optimization problem,” IEEE Trans. Audio Speech Language Process., vol. 18, no. 8, pp. 2051–2066, 2010.
[73] J. Kominek, A. W. Black, and V. Ver, “CMU Arctic databases for speech synthesis,” Tech. Rep., 2003.
[74] J. Sharpnack, A. Rinaldo, and A. Singh, “Detecting anomalous activity on networks with the graph Fourier scan statistic,” IEEE Trans. Signal Process., vol. 64, no. 2, pp. 364–379, 2016.
[75] S. Mallat, A Wavelet Tour of Signal Processing, Third Edition: The Sparse Way, 3rd ed. Academic Press, 2008.