

Article

A New Surrogating Algorithm by the Complex Graph Fourier Transform (CGFT)

Jordi Belda ¹, Luis Vergara ^{1,*}, Gonzalo Safont ¹, Addisson Salazar ¹ and Zuzanna Parcheta ²

¹ Institute of Telecommunications and Multimedia Applications, Universitat Politècnica de València, 46022 València, Spain

² Sciling SL, C/Camí a La Mar 75, 46022 València, Spain

* Correspondence: lvergara@dcom.upv.es; Tel.: +34-963-877-308

Received: 4 July 2019; Accepted: 2 August 2019; Published: 4 August 2019



Abstract: The essential step of surrogating algorithms is phase randomizing the Fourier transform while preserving the original spectrum amplitude before computing the inverse Fourier transform. In this paper, we propose a new method which considers the graph Fourier transform. In this manner, much more flexibility is gained to define properties of the original graph signal which are to be preserved in the surrogates. The complex case is considered to allow unconstrained phase randomization in the transformed domain, hence we define a Hermitian Laplacian matrix that models the graph topology, whose eigenvectors form the basis of a complex graph Fourier transform. We have shown that the Hermitian Laplacian matrix may have negative eigenvalues. We also show in the paper that preserving the graph spectrum amplitude implies several invariances that can be controlled by the selected Hermitian Laplacian matrix. The interest of surrogating graph signals has been illustrated in the context of scarcity of instances in classifier training.

Keywords: surrogates; graph Fourier transform; Hermitian Laplacian matrix

1. Introduction

1.1. Statement of the Problem and Related Works

Starting from an available original signal, synthetic signals can be generated which are surrogates of it. Essentially, surrogate means that the (Fourier) spectrum amplitude (we will use the term amplitude to define the magnitude or modulus of a complex number or a real number) of the original signal is retained while the phase is randomized. Surrogating methods have been considered in a variety of applications [1–7], where the synthetic signals are used to obtain a non-parametric estimate of the distribution of some statistic considered in a given test of hypotheses. In spite of its attractive approach, existing surrogating methods are too constrained due to the use of the Fourier transform (FT)—Both the original signal and its surrogates share the same autocorrelation function. This implies that pairwise interrelations are assumed to be the same for pairs of samples equally separated. While this can be appropriate in some cases, much more flexibility would be gained if specific pairwise relations between samples could be imposed in the surrogates. Thus, for example, we could enhance short-term properties by prioritizing interrelations between neighbor samples. We could also introduce application domain interrelation constraints given by an expert that must agree with the surrogates. In fact, we will show that FT methods are particular cases of the newly proposed method.

In this paper, we propose a new method for generating surrogates which exploit the concept of graph signal processing [8–10]. Graph signal processing (GSP) is an emerging area which extends the classical concept of signals defined on a uniform 1-D (time signal) or 2-D (images) grid, to more general grids. Thus it allows the ability to model interrelations between samples in a more general way

and benefits from consolidated concepts and methods emanating from signal processing and related areas. Notice that a graph signal may represent any arbitrary domain: Time samples, image pixels, a vector of features extracted from some preprocessing, the outputs of a sensor network, and so on.

In a previous related work [11], the authors propose a sign randomization of the graph Fourier transform (GFT). The GFT represents the graph signal in a domain expanded by the eigenvectors of the graph Laplacian matrix (other GFT are possible like the one proposed in [12], which considers the generalized eigenvectors of the adjacency matrix. This may lead to further extensions of the work presented here) [8,13]. The algorithm is applied to generate surrogates of EEG signals defined in the space domain, which preserve the smoothness of the original signals in the defined graph. The authors demonstrate the interest of the method in the detection of the effects of behavioral tasks in electroencephalography signals. The good results of this novel approach suggest continuing to working on it.

1.2. New Contributions and Paper Organization

So, in this paper, we go deeper into the concept in several aspects. First, the complex case is considered. Notice that the GFT of a real graph signal is also real, so phase randomization is actually limited to be a sign randomization as the authors do in [11]. However, working with complex signals allows unconstrained phase randomization. But a complex signal should be expressed as a linear combination of complex eigenvectors. Thus we have defined a complex Hermitian Laplacian matrix (Section 2.1), while most of the previous work on GSP assumes real Laplacian matrices. We derive properties of the Hermitian Laplacian, in particular, a generalization of the smoothness [8,14–16] relative to the conjugate complex weights. Then, we generalize the GFT to the complex case (Section 2.2) defining a Complex GFT (CGFT).

Besides, with the aim of gaining understanding and motivation of the proposed surrogating algorithm, we have made an analysis of what properties of the original signals are preserved in the surrogates. This is presented in Section 3 through the concept of graph spectrum amplitude (GSA) invariants. In particular, we demonstrate that the smoothness, the precision matrix [17,18], and the covariance matrix of graph wide-sense stationary (GWSS) stochastic signals, as recently defined in [19,20] are GSA invariants.

In Section 4.1 we present the new surrogating algorithm. First, existing algorithms based on the FT are briefly reviewed. We also show that these algorithms are equivalent to a particular definition of graph connectivity. The issue of how to select a Hermitian Laplacian matrix is considered in Section 4.2.

Finally, Section 5 presents some experiments to assess the improvements of the new surrogating method based on the CGFT in comparison with the one based on the FT. The application target is the training of a classifier with a scarcity of data. This may be considered as a generalization of the classical hypothesis testing problem where surrogates are routinely considered. Thus Section 4.1 is devoted to present experiments of hand gestures recognition. This is a difficult classification problem due to the large number of classes, large, and variant number of features.

1.3. Notation

Lower or upper case italics denote a scalar; lower case bold letters denote vectors with v_n denoting the n -th element of \mathbf{v} ; upper case bold letters denote matrix with $\mathbf{V}(n, m)$ or v_{nm} denoting the (m, n) -th element of \mathbf{V} and \mathbf{v}_n the n -th column; $(\cdot)^*$, $(\cdot)^T$, $(\cdot)^H$ denote conjugate, transpose, and conjugate transpose; $\text{Re}[\cdot] \equiv [\cdot]^R$, $\text{Im}[\cdot] \equiv [\cdot]^I$ denote real and imaginary parts; $|\cdot|$, $\|\cdot\|$ denote modulus and Euclidian norm; $\angle v$ denotes phase of the complex number v ; $\det \mathbf{V}$ denotes determinant; $\hat{\mathbf{v}}$ denotes a surrogate of \mathbf{v} ; $E[\cdot]$ denotes expectation; \mathbf{C}_{vv} , $\mathbf{Q}_{vv} = \mathbf{C}_{vv}^{-1}$ denote the covariance and the precision matrices of the random vector \mathbf{v} ; $\mathbf{v}^{(i)}$, $\mathbf{V}^{(i)}$ denote vector or matrix values corresponding to i -th iteration of an iterative algorithm. We will use j to indicate the complex number verifying $j^2 = -1$.

Finally, a list of acronyms is included in Appendix C, to facilitate tracking.

2. The Hermitian Laplacian Matrix

2.1. Definition and Properties

There are not many precedents of using complex Laplacian matrices. In [21] a Hermitian Laplacian matrix is proposed to include some possible oriented edges, but this leads to a very limited form of Hermitian Laplacian matrices—Its off-diagonal elements are 1 for non-oriented edges and $\pm j$ for the oriented edge. A more general Hermitian Laplacian matrix is proposed in [22] by adding a phase term to the off-diagonal elements of a real Laplacian matrix. Then some permutation invariants are found in the pattern recognition context. This extension implies that the real part of the corresponding adjacency matrix is not guaranteed to have positive entries. Here we propose a different Hermitian Laplacian matrix by adding an imaginary part to the (positive) off-diagonal elements of the adjacency matrix, while the n -th element of the diagonal matrix \mathbf{D} is kept as the sum of the real parts of the edges incident to node n . So, let us consider the weighted graph $G\{V, E, \mathbf{W}\}$, where V represents the set of N vertices or nodes of the graph, E represents the set of edges connecting the nodes, and \mathbf{W} is the adjacency matrix. The generic element $\mathbf{W}(n, m)$ is the weight corresponding to the edge connecting node m to node n . It is rather common in the literature to assume \mathbf{W} to be a real and symmetric matrix with non-negative elements $\mathbf{W}(n, m) \geq 0 \quad \forall n, m$. While this seems reasonable for real graph signals, it seems too constrained for complex graph signals. Thus we will consider that \mathbf{W} is, in general, a Hermitian matrix $\mathbf{W} = \mathbf{W}^H, \quad \mathbf{W}(n, m) = \mathbf{W}^*(m, n) \quad \forall n, m$. We will consider that the real part is semi-positive $\text{Re}[\mathbf{W}(n, m)] \geq 0 \quad \forall n, m$. This allows including the real case as a special case. Notice that in the real case all the edges are undirected because $\mathbf{W}(n, m) = \mathbf{W}(m, n) \quad \forall n, m$. However, in the complex case, the imaginary part of the complex weights changes the sign depending on the direction $\text{Im}[\mathbf{W}(n, m)] = -\text{Im}[\mathbf{W}(m, n)]$, so these edges are directed. This situation is sometimes known as mixed graphs [21].

We may define a Hermitian Laplacian matrix in the form $\mathbf{L} = \mathbf{D} - \mathbf{W}$ where \mathbf{D} is a diagonal matrix having the generic element $\mathbf{D}(n, n) = \sum_{m=1}^N \text{Re}[\mathbf{W}(n, m)]$. This is an obvious generalization of the real case where $\mathbf{D}(n, n) = \sum_{m=1}^N \mathbf{W}(n, m)$ and \mathbf{D} is termed the degree matrix. The foregoing definition of the Hermitian Laplacian matrix is appropriate for extending the concept of smoothness to complex graph signals. Let us define the graph signal $\mathbf{s} = [s_1, \dots, s_N]^T \in \mathbb{C}^N$, where the value s_n is associated with the n -th node of the graph, for $n = 1 \dots N$. Notice that \mathbf{s} may correspond to any arbitrary domain.

The graph smoothness can be defined as a simple generalization of the real case

$$S(\mathbf{s}) = \mathbf{s}^H \mathbf{L} \mathbf{s} \tag{1}$$

We show in Appendix A that

$$S(\mathbf{s}) = \mathbf{s}^H \mathbf{L} \mathbf{s} = \sum_{n=1}^{N-1} \sum_{m>n}^N w_{nm}^R |s_n - s_m|^2 + \sum_{n=1}^{N-1} \sum_{m>n}^N 2w_{nm}^I \text{Im}[s_n^* s_m] \tag{2}$$

where $w_{nm}^R \equiv \text{Re}[\mathbf{W}(n, m)]$ and $w_{nm}^I \equiv \text{Im}[\mathbf{W}(n, m)]$. Notice that if the Laplacian matrix and/or the graph signal are real, the second term in Equation (2) vanishes, and the smoothness becomes the classical weighted measure of the (squared) signal differences between every two samples. But in the most general case where both the Laplacian matrix and the graph signal may be complex, a new term appears. The latter can be expressed in the form:

$$\sum_{n=1}^{N-1} \sum_{m>n}^N 2w_{nm}^I \text{Im}[s_n^* s_m] = \sum_{n=1}^{N-1} \sum_{m>n}^N 2w_{nm}^I |s_n| |s_m| \sin(\angle s_m - \angle s_n) \tag{3}$$

where $s_n = |s_n| \exp(j\angle s_n)$. Hence, by using the defined Hermitian Laplace matrix, we are incorporating into the smoothness information about the phase difference of the value assigned to node m with respect to node n . Let us illustrate it with a simple example—A line graph, i.e., a graph in which node n is connected only with node $n + 1$ and node $n - 1$, except node 1 which is connected only with node 2, and node N which is connected only with node $N - 1$. Let us also assume that the weights are constants, and let us define $\mathbf{s}_r = [s_N, \dots, s_1]^T$, a reversed version of \mathbf{s} . It is straightforward to show that the first term of Equation (1) is the same for both the smoothness of \mathbf{s} and \mathbf{s}_r , while the second term changes the sign. Thus the smoothness depends on graph signal amplitude variability over the graph (as it happens in the real case) and on the sense of rotation when moving in the complex plane from one node to the other. This second term can be positive or negative, so the smoothness could eventually be negative for some graph signals. This would imply that the corresponding Hermitian Laplacian matrix could not be a positive semidefinite as a real Laplacian matrix is. This can be demonstrated using Sylvester’s criterion [23] which states that a necessary and sufficient condition for a Hermitian matrix to be positive definite is that all its leading principal minors (determinants of the leading submatrices) are positive. Let us call \mathbf{L}_2 the (2×2) leading submatrix of \mathbf{L} .

$$\mathbf{L}_2 = \begin{pmatrix} w_{12}^R + \sum_{m>2} w_{1m}^R & -w_{12}^R - jw_{12}^I \\ -w_{21}^R - jw_{21}^I & w_{21}^R + \sum_{m>2} w_{2m}^R \end{pmatrix} = \begin{pmatrix} w_{12}^R + \sum_{m>2} w_{1m}^R & -w_{12}^R - jw_{12}^I \\ -w_{12}^R + jw_{12}^I & w_{12}^R + \sum_{m>2} w_{2m}^R \end{pmatrix} \tag{4}$$

$$\det \mathbf{L}_2 = \left[w_{12}^R \sum_{m>2} (w_{1m}^R + w_{2m}^R) + \sum_{m>2} w_{1m}^R \sum_{m>2} w_{2m}^R \right] - (w_{12}^I)^2$$

We see that depending on the selection of the weights corresponding to the nodes 1 and 2, the leading principal minor of order 2 could be non-positive. The result, Equation (4), could be applied to any pair of nodes as reordering the nodes leads to a new Hermitian Laplacian matrix which is permutation similar to the original one [24], and the positive or non-positive semidefinite condition remains the same.

In consequence, given that the contributions to the second term in Equation (2) may be positive or negative, one may wonder about the actual meaning of increasing or decreasing the smoothness in the complex case. In the real case, this has an intuitive answer, as the smoothness is a measure of graph signal variability on the graph. However, in the complex case, the variability measure of the first term can eventually be compensated by negative contributions of the second term. Hence, to give meaning (even intuitive) to Equation (2) we must consider particular graph topologies (as we did in the above example of a line graph) or define some constraints to the graph signals whose smoothness is being compared. Thus, for example, Equation (2) is a measure of graph signal variability on a given graph, for the graph signal subset satisfying $\text{sign}[w_{nm}^I] = \text{sign}[\sin(\angle s_m - \angle s_n)]$ so that all the contributions of the second term in Equation (2) are positive. By doing so we guarantee that both the distances and the phase differences make positive contributions to the smoothness, and that zero smoothness is obtained only when $s_n = s_m \quad \forall nm$, as it is in the real case.

2.2. The Complex Graph Fourier Transform (CGFT)

Any Hermitian matrix can be diagonalized by a unitary matrix, then

$$\mathbf{L} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^H = \sum_{n=1}^N \lambda_n \mathbf{u}_n \mathbf{u}_n^H \tag{5}$$

where $\lambda_n \quad n = 1 \dots N$ are the eigenvalues of \mathbf{L} ($\mathbf{\Lambda}$ a diagonal matrix, $\mathbf{\Lambda}(n, n) = \lambda_n$) and $\mathbf{u}_n \quad n = 1 \dots N$ (columns of \mathbf{U}) are the corresponding orthonormal eigenvectors, $\mathbf{U}^H = \mathbf{U}^{-1}$ so \mathbf{U} is unitary. The eigenvalues are real due to the Hermitian property. The GFT is defined for the real case as the projection of the graph signal on the vector space expanded by a basis formed by the eigenvectors of the real Laplacian matrix. Then the notion of frequency is assimilated to the eigenvalue associated

with every eigenvector in much the same manner as every complex exponential is defined by a specific frequency when using the FT. We may extend this concept to the complex case by simply considering the eigenvectors of the Hermitian Laplacian, hence defining the direct and inverse CGFT as

$$\mathbf{r} = \text{CGFT}(\mathbf{s}) = \mathbf{U}^H \mathbf{s} \quad , \quad \mathbf{s} = \text{CGFT}^{-1}(\mathbf{r}) = \mathbf{U} \mathbf{r} \quad (6)$$

In the real case, the Laplacian matrix is positive semidefinite, then $\lambda_n \geq 0 \quad \forall n$, in fact, the minimum eigenvalue is 0 [25]. However, we have seen that in the complex case the Hermitian matrix \mathbf{L} could not be positive semidefinite, so at least one eigenvalue will be negative. Notice that the different modes (eigenvectors) of a Hermitian matrix can be ordered according to increasing smoothness because $\mathbf{u}_n^H \mathbf{L} \mathbf{u}_n = \lambda_n$ and that for any vector \mathbf{s} $\lambda_{\min} \leq \frac{\mathbf{s}^H \mathbf{L} \mathbf{s}}{\mathbf{s}^H \mathbf{s}} \leq \lambda_{\max}$. Hence we can still assimilate the concept of frequency to the eigenvalue, i.e., CGFT^{-1} in Equation (6) expresses the graph signal as a linear combination of modes having increasing smoothness (from λ_{\min} to λ_{\max}). However notice that, as explained at the end of Section 2.1, it is difficult to get a general interpretation of the meaning of smoothness. Thus the appearance of negative eigenvalues suggests that the sense of rotation when moving in the complex plane from one node to the other can be relevant for a complete characterization of a graph signal by the CGFT. This has been verified in the particular line graph example of the previous section. Another simple numerical example follows. Let us consider a graph of two nodes, where $w_{12} = 1 + j = w_{21}^*$. If we compute the two eigenvalues and two eigenvectors of the corresponding 2×2 Hermitian Laplacian we obtain $\lambda_1 = -0.41$ and $\lambda_2 = 2.41$, $\mathbf{u}_1 = [0.5 + 0.5j, 0.707]^T$, and $\mathbf{u}_2 = [-0.5 - 0.5j, 0.707]^T$. We see that a negative eigenvalue appears and that in the first eigenvector there is a clockwise rotation when moving in the complex plane from its first element to the second. However, in the second eigenvector, the rotation is counterclockwise. Similarly, we have shown at the end of Section 2.1 that the newly defined smoothness may keep the classical meaning of variability measure by imposing some constraints to the graph signals to be compared. Now, this measure incorporates distances as well as phase differences, which is consistent with the complex domain. Finally, notice that FT can be interpreted as a particular case of a CGFT (as it is of GFT) where the associated Hermitian Laplacian is a circulant matrix (see below, end of Section 4.1).

3. Graph Spectrum Amplitude Invariants

The key element of the surrogating algorithms is to keep the GSA of the original graph signal while the graph spectrum phases are randomized. So let us find GSA invariants before presenting the detailed algorithms in the next section. This helps to motivate the graph methods and to gain insights into the implications of surrogating graph signals. GSA invariants are those functions or statistics of the original graph signal which do not change if the GSA does not change. Given a signal \mathbf{s} , the set of signals having the same GSA as \mathbf{s} is

$$\hat{\mathbf{s}} = (\mathbf{U}^H)^{-1} \mathbf{\Phi} \mathbf{U}^H \mathbf{s} = \mathbf{U} \mathbf{\Phi} \mathbf{U}^H \mathbf{s} \quad (7)$$

where $\mathbf{\Phi}$ is the set of $N \times N$ diagonal matrices whose diagonal coefficients are unimodular, i.e., $\Phi(n, n) = \exp(j\phi_n) \quad n = 1 \dots N$. Notice that the spectral amplitudes of the transformed vector are retained, but a phase ϕ_n is added to the n -th graph spectrum phase, before the inverse transformation. Let us define the unitary matrix $\mathbf{P}^H = \mathbf{U} \mathbf{\Phi} \mathbf{U}^H = \mathbf{P}^{-1}$, from Equation (7) we can write

$$\hat{\mathbf{s}} \hat{\mathbf{s}}^H = \mathbf{U} \mathbf{\Phi} \mathbf{U}^H \mathbf{s} \mathbf{s}^H \mathbf{U} \mathbf{\Phi}^H \mathbf{U}^H = \mathbf{P}^{-1} \mathbf{s} \mathbf{s}^H \mathbf{P} \quad (8)$$

Then the sample correlation matrices of the original graph signal and its surrogates are unitarily similar [26]. Specht's criterion [27] establishes a necessary and sufficient condition for the unitary

similarity of two matrices in terms of some traces invariance. The application of this criterion to the sample correlation matrices leads to

$$\text{trace}\left(\left(\hat{\mathbf{s}}\hat{\mathbf{s}}^H\right)^k\right) = \text{trace}\left(\left(\mathbf{s}\mathbf{s}^H\right)^k\right), \quad (9)$$

for any integer k . Equation (9) defines invariants functions involving polynomials of the elements of \mathbf{s} and $\hat{\mathbf{s}}$. In particular, for $k = 1$ it implies that the energy of the graph signal is a GSA invariant

$$\sum_{n=1}^N |\hat{s}_n|^2 = \sum_{n=1}^N |s_n|^2, \quad (10)$$

and for $k = 2$ we can see that the Frobenius norm of the sample correlation matrix is also a GSA invariant

$$\sum_{n=1}^N \sum_{m=1}^N |\hat{s}_n \hat{s}_m^*|^2 = \sum_{n=1}^N \sum_{m=1}^N |s_n s_m^*|^2. \quad (11)$$

Notice that, so far, the value of the GSA invariant is the same for every possible Hermitian Laplacian matrix, i.e., they are implicit in the preservation of the GSA. Let us consider invariants which depend on the Hermitian Laplacian. First, notice that \mathbf{s} and $\hat{\mathbf{s}}$ are equally smooth on graph $G\{V, E, \mathbf{W}\}$,

$$\begin{aligned} S(\hat{\mathbf{s}}) &= \hat{\mathbf{s}}^H \hat{\mathbf{L}} \hat{\mathbf{s}} = \mathbf{s}^H \mathbf{U} \Phi^H \mathbf{U}^H \mathbf{L} \mathbf{U} \Phi \mathbf{U}^H \mathbf{s} \\ &= \mathbf{s}^H \mathbf{U} \Phi^H \mathbf{U}^H \mathbf{U} \Lambda \mathbf{U}^H \mathbf{U} \Phi \mathbf{U}^H \mathbf{s} \\ &= \mathbf{s}^H \mathbf{L} \mathbf{s} = S(\mathbf{s}), \end{aligned} \quad (12)$$

so the smoothness on the graph defined by the selected Hermitian Laplacian is a GSA invariant.

Let us now adopt a stochastic perspective. We assume that \mathbf{s} is a zero mean random vector having a known precision matrix $\mathbf{Q}_{ss} = \mathbf{C}_{ss}^{-1}$, where $\mathbf{C}_{ss} = E[\mathbf{s}\mathbf{s}^H]$ is the covariance matrix. Taking expectations and inverting in Equation (8) we conclude that the precision matrices are also unitarily similar

$$\mathbf{Q}_{\hat{s}\hat{s}} = \mathbf{P}^{-1} \mathbf{Q}_{ss} \mathbf{P}. \quad (13)$$

The elements of \mathbf{Q}_{ss} properly normalized are the partial correlations [28–31] defining linear conditional dependences between every two elements of \mathbf{s} . This motivates the extended use of \mathbf{Q}_{ss} as the Laplacian matrix. So, let us assume that we select $\mathbf{L} = \mathbf{Q}_{ss}$, then, considering Equation (5), and the definition of \mathbf{P} , we get:

$$\mathbf{Q}_{\hat{s}\hat{s}} = \mathbf{U} \Phi \mathbf{U}^H \mathbf{Q}_{ss} \mathbf{U} \Phi^H \mathbf{U}^H = \mathbf{U} \Phi \mathbf{U}^H \mathbf{U} \Lambda \mathbf{U}^H \mathbf{U} \Phi^H \mathbf{U}^H = \mathbf{Q}_{ss}. \quad (14)$$

Hence the precision matrix is invariant to the GSA when used as the Hermitian Laplacian. Notice that the interpretation of the smoothness in Equation (2) is still valid in this case. This is demonstrated in [17] for the real Laplacian and can be straightforwardly extended to the Hermitian Laplacian as the complex component only affects the nondiagonal elements.

Finally, some recent works [19,20] have proposed equivalent extensions of the concept of stationarity to graph signals by considering definitions of graph signal translation. Then, in [19,20] the authors demonstrate that a stochastic graph signal is GWSS if the covariance matrix is jointly diagonalizable with the Laplacian matrix. So let us assume that the pairwise interrelations are defined such that \mathbf{s} is a realization of a (zero mean) random graph signal which is GWSS with respect to the Laplacian used to generate the surrogates, i.e., that the graph signal is invariant to some defined operator of translation, then $\mathbf{C}_{ss} = \mathbf{U} \mathbf{\Gamma}_{ss} \mathbf{U}^H$, where $\mathbf{\Gamma}_{ss}$ is a diagonal matrix, hence

$$C_{\hat{s}\hat{s}} = \mathbf{U}\Phi\mathbf{U}^H\mathbf{U}\Gamma_{ss}\mathbf{U}^H\mathbf{U}\Phi^H\mathbf{U}^H = C_{ss}. \tag{15}$$

Therefore, the covariance matrix is a GSA invariant for GWSS signals, which is consistent with the well-known invariance of the autocovariance function with respect to the power spectral density of (conventional) WSS signals. Notice that the preserved covariance matrix is not necessarily Toeplitz as it is implicitly assumed when preserving the autocorrelation function with classical surrogating methods. Table 1 summarizes the GSA invariants demonstrated in this section.

Table 1. Graph Spectrum Amplitude (GSA) invariants.

Invariant	Equation
Signal energy	$\sum_{n=1}^N s_n ^2 = \sum_{n=1}^N \hat{s}_n ^2$
Frobenious norm of the sample correlation matrix	$\sum_{n=1}^N \sum_{m=1}^N s_n s_m^* ^2 = \sum_{n=1}^N \sum_{m=1}^N \hat{s}_n \hat{s}_m^* ^2$
Traces of the sample correlation matrix powers	$trace\left((\mathbf{s}\mathbf{s}^H)^k\right) = trace\left(\left(\hat{\mathbf{s}}\hat{\mathbf{s}}^H\right)^k\right)$
Smoothness	$\mathbf{s}^H \mathbf{L} \mathbf{s} = \hat{\mathbf{s}}^H \hat{\mathbf{L}} \hat{\mathbf{s}}$
Precision matrix	If $\mathbf{L} = \mathbf{Q}_{ss}$ then $\mathbf{Q}_{\hat{s}\hat{s}} = \mathbf{Q}_{ss}$
Graph Wide-Sense Stationarity	If $C_{ss} = \mathbf{U}\Gamma_{ss}\mathbf{U}^H$ then $C_{\hat{s}\hat{s}} = C_{ss}$

4. Surrogating Algorithms

4.1. Iterative Surrogate Graph Signal Algorithm

One of the first proposals of a surrogating algorithm was the amplitude adjusted Fourier transform (AAFT), [32]. The FT of the original (real) signal is computed, then the samples obtained after spectrum phase randomization followed by inverse Fourier transform are corrected to comply with the empirical distribution of the original samples. Random phases are obtained by sampling a uniform distribution between $-\pi$ and π , constrained to keep the antisymmetry property of the spectral phases of real signals. Corrections consist of replacing every sample of the surrogate so obtained, by the value of the original sample having the same rank order, thus recovering the empirical distribution of the original signals. Hence, the surrogate is a shuffled version of the original signal.

An improved version which iteratively adjusts the spectrum amplitude and corrects the sample empirical distribution was proposed—Iterative AAFT (IAAFT) [1,33]. Minor improvements or variations of the AAFT have been proposed in [34–36]. More recently, extensions to the multivariate case have been considered where the added problem is to keep the correlation among the different surrogate channels [37–40].

So far, these surrogating algorithms have been applied only to real signals in the time domain. We present in this section an iterative algorithm for surrogating complex graph signals by means of the CGFT (ICGFT). Everything starts with the original graph signal $\mathbf{s} = [s_1, \dots, s_N]^T \in C^N$. We compute an initial uncorrected surrogate $\hat{\mathbf{s}}^{(0)}$ by randomizing the phases of the CGFT of \mathbf{s} (Equation (16a)) while the GSA is maintained. Random phases are obtained by sampling a uniform distribution between $-\pi$ and π (no antisymmetry constraint is imposed as complex signals are considered). Then $\hat{\mathbf{s}}^{(0)}$ is corrected. This correction can be made to match the empirical distribution of the original graph signal [1,33], but other corrections could be better suited to the specific application context (see Section 5), so we will represent the correction by a generic function $C(\cdot)$. Thus we get the initial corrected surrogate $\mathbf{s}^{(0)}$ (Equation (16b)) as

$$\hat{\mathbf{s}}^{(0)} = \mathbf{U}\Phi\mathbf{U}^H\mathbf{s}, \tag{16a}$$

$$\mathbf{s}^{(0)} = C\left(\hat{\mathbf{s}}^{(0)}\right). \quad (16b)$$

Notice that in general correction $C(\cdot)$ will modify the GSA, so we proceed iteratively, as in IAAFT, computing the CGFT of the corrected surrogate, recovering the original GSA and generating a new uncorrected surrogate $\hat{\mathbf{s}}^{(i)}$ Equation (17a). The latter is again corrected to match the empirical distribution of the original graph signal Equation (17b), and thus successively:

$$\hat{\mathbf{s}}^{(i)} = \mathbf{U}\mathbf{A}^{(i-1)}\mathbf{U}^H\mathbf{s}^{(i-1)}, \quad (17a)$$

$$\mathbf{s}^{(i)} = C\left(\hat{\mathbf{s}}^{(i)}\right) \quad i = 1 \dots I. \quad (17b)$$

Defining $\mathbf{r}^{(i-1)} = \mathbf{U}^H\mathbf{s}^{(i-1)}$ then $\mathbf{A}^{(i-1)}$ is a diagonal matrix having elements $\mathbf{A}^{(i-1)}(n,n) = |r_n|/|r_n^{(i-1)}|$. This iterates until convergence, i.e., until no corrections are required. Algorithm 1 below is a pseudocode description of ICGFT.

Algorithm 1: Iterative surrogating complex graph signals by means of the CGFT (ICGFT).

- 1: **Input:** Graph signal \mathbf{s} , Hermitian Laplacian \mathbf{L} , number of surrogates K , maximum number of iterations I , convergence threshold ε
 - 2: **Compute eigenvector matrix \mathbf{U}**
 - 3: **for** $k = 1, 2, \dots, K$ **do:**
 - 4: Generate graph spectrum random phases Φ_k
 - 5: Compute initial surrogate, Equation (16a,b) $\mathbf{s}_k^{(0)} = C(\mathbf{U}\Phi_k\mathbf{U}^H\mathbf{s})$
 - 6: **for** $i = 1, 2, \dots, I$ **do:**
 - 7: Compute surrogate, Equation (17a,b) $\mathbf{s}^{(i)} = C(\mathbf{U}\mathbf{A}^{(i-1)}\mathbf{U}^H\mathbf{s}^{(i-1)})$ where $\mathbf{A}^{(i-1)} = \text{diag}\left(|r_n|/|r_n^{(i-1)}|\right)$
 - 8: If $\|\mathbf{s}^{(i)} - \mathbf{s}^{(i-1)}\| < \varepsilon$ then $\mathbf{s}_k = \mathbf{s}^{(i)}$ and go to 11
 - 9: **end for**
 - 10: $\mathbf{s}_k = \mathbf{s}^{(I)}$
 - 11: **end for**
 - 12: **Output:** $\mathbf{s}_1 \dots \mathbf{s}_K$
-

Unfortunately, there is not a clear understanding or theoretical analysis of convergence in all the different surrogating algorithms proposed so far. However, it has been experimentally verified that convergence is reached after a few iterations in most cases. Notice that measuring how close $\mathbf{s}^{(i-1)}$ is to $\mathbf{s}^{(i)}$ in line 8 of Algorithm 1 (i.e., fitting the threshold ε) is arbitrary, with no control of the computational time required to generate surrogates. That is why a maximum number I of iterations is allowed. The value ε may be experimentally adjusted in every particular application. In any case, notice that, ultimately, the performance indicators of the underlying application will assess the validity of the generated surrogates rather than a specific criterion to fit ε . Thus, for example, in all the experiments of Section 5, we have finally set a constant number of 10 iterations, i.e., $\varepsilon = 0 \quad I = 10$, after verifying that there was not a substantial improvement in the performance indicator (probability of error) with an (experimentally) optimized selection of ε .

4.2. Selecting the Hermitian Laplacian

A key aspect of the ICGFT algorithm is selecting an appropriate Laplacian (or adjacency) matrix. This is a classical problem in graph signal processing. One option could be to estimate \mathbf{Q}_{SS} from the set of available original graph signals (see for example [17,18,28–31,41,42]), and then using it as the Hermitian Laplacian. In such a case, we have seen in Section 3 that the precision matrix thus estimated is preserved in the surrogates. Other alternatives make a direct estimation of the Laplacian matrix

by minimizing cost functions, which include some regularization constraints [15,16]. However, in a scenario of scarcity, it would be difficult to obtain stable estimates of the Hermitian Laplacian matrix by all these methods.

A different option is to define connections from a general or specific desired connectivity that should be preserved in the surrogates. For example, a general option could be to connect every node only with its neighbors (this is done in [11] in the space domain), so that local properties of the graph signal can be retained in the surrogates. More specific connections can be established depending on the context and the application domain.

A question remains regarding the sign of w_{nm}^I . Remember that $w_{nm}^R = w_{mn}^R \geq 0$, but $w_{nm}^I = -w_{mn}^I$, for example, we may connect two samples because they are neighbors, but this can be made with both a positive or negative value of the imaginary part w_{nm}^I . As indicated in Equation (2), a different sign selection of w_{nm}^I leads to a different smoothness that eventually could be negative. We propose to select the signs in such a way that $S(\mathbf{s})$ is maximized, i.e., by fitting $\text{sign}[w_{nm}^I] = \text{sign}[\sin(\angle s_m - \angle s_n)]$. We show in Appendix B, that the spectrum bandwidth $B = \lambda_{\max} - \lambda_{\min}$ of the Hermitian Laplacian matrix is lower bounded by

$$B \geq \max\left(l_{\max}, \frac{S(\mathbf{s})}{\mathbf{s}^H \mathbf{s}}\right) - l_{\min}, \quad (18)$$

where l_{\max} and l_{\min} are respectively the maximum and minimum of the main diagonal elements of the Hermitian Laplacian matrix (which depends only on w_{nm}^R). Then, given w_{nm}^R , w_{nm}^I , and \mathbf{s} , the lower bound is maximized when $S(\mathbf{s})$ is the maximum. As the key step of surrogating is phase randomization in the transformed domain, it is appropriate to have available the maximum possible bandwidth.

Finally, notice that FT (more precisely its numerical implementation by the discrete Fourier transform (DFT)) can be interpreted as a particular form of selecting the Hermitian Laplacian matrix. This is because it is well-known that the DFT basis vectors are eigenvectors of circulant matrices [43]. So, implicitly, the use of the DFT imposes constraints in the elements of the adjacency matrix. In particular, pairwise connectivity is the same for nodes located at the same “distance” in the node numbering, which implies that stationarity (from the conventional concept of translation) is assumed. Moreover, the required cyclic symmetry of the coefficients imposes additional constraints which may be nonsense in a general context. Also, notice that the eigenvectors of a circulant Hermitian Laplacian matrix do not depend on the specific value of the weights. Hence, much more flexibility can be obtained by defining specific connectivity in concordance with the desired properties of the original graph signal that are to be preserved in the surrogates.

5. Experiments

Many applications of signal processing are characterized by a scarcity of signals during the development phase. This, in turn, produces a shortage of instances (vectors of features extracted from the preprocessed signals) for automatic training of classifiers. For example, in automatic non-destructive testing of materials, a classifier is to be trained for every class of defect. This requires a collection of a large number of signals including all the possible defect types, which can be impractical or even impossible [44,45]. Another totally different example is credit card fraud detection [46], where the number of legitimate operations is an order of magnitudes larger than the fraudulent operations. This also happens in multi-class problems where some classes may have a reduced training set [47]. A solution to alleviate scarcity is to generate synthetic replicas using traditional methods (see for example [48] and references therein). But traditional methods are model-oriented, e.g., the model fits some first-order probability density and correlation properties. Model-oriented methods require the definition of an appropriate model and the estimation of the model parameters. But in a scenario of scarcity, both things are rather limited. In contrast, surrogating may be considered a signal-oriented method of synthesis, it has actually already been proposed as a general method for synthesizing multivariate data [38]. Hence, in this experimental section, we consider the application of surrogates to increase the data set for classifier learning. Moreover, notice that to some extent, this is a natural

generalization of the use of surrogates in hypothesis testing. In this latter case, the null hypothesis is that the instance under test satisfies a given model (for example a linear model followed by a scalar nonlinearity). But, ultimately, some empirical properties of the original instance (autocorrelation, amplitude distribution, etc) are assumed to represent the model and are accordingly imposed to the surrogates. The same principle can be applied to generate instance replicas for deficient classes in classifier training: The replicas (surrogates) must keep some relevant properties of the original instances. In particular, considering that the instances are graph signals, the surrogates preserve the GSA invariants deduced in Section 3, which relates to preserving the pairwise interrelations defined by the corresponding Hermitian Laplacian matrix.

5.1. Automatic Hand Gesture Recognition

In this section, we consider the use of the surrogating methods in a real data case study—Automatic hand gesture recognition. This has been selected due to the complexity of the problem—A large number of classes, with a large and variant number of features. Thus, results should be hopefully extrapolated to simpler cases of automatic classification. Hand gesture recognition is very useful in everyday life for computer–human interaction in applications such as navigating in virtual environments or recognizing hand poses [49]. Originally, hands were tracked by camera images to determine their pose and their 2-D trajectory, but the introduction of depth sensors like Kinect or Leap Motion allowed the ability to obtain 3-D information to recognize the gestures [50]. These applications typically use hidden Markov models (HMM) [51] or support vector machine [52] based classifiers. We have used a public dataset (available at <https://github.com/Sasanita/spanish-sign-language-db>) composed by 91 approximations of the gestures of Spanish sign language acquired by a Leap Motion sensor. This dataset contains 40 instances per sign. Each instance is a matrix having 42 columns. Every column is a real signal corresponding to the temporal evolution of one observed variable (feature). The whole 42 observed variables describe the trajectory of the fingers and hands and their inclination according to X, Y, and Z axes of a three-dimensional system. The sampling frequency is 30 Hz for all instances, however, the observation time may vary between different instances because the hand sign can be produced at different velocities. Thus, the column dimension of the instance matrices varies from 7 (0.23 s) to 101 (3.36 s) with an average value of 26 (0.86 s). In all cases, the classifier was based on a hidden Markov model (HMM) having seven states. The conditioned probability of the observations to the hidden states was considered a Gaussian mixture of two components. The Viterbi algorithm was used for estimating the most probable sequence of states. Actually, the Hidden Markov Model Toolkit was used (details can be found in [53] and references therein). In all cases the testing of the classifier once trained, was made with original instances not included in the training sets to avoid overfitting.

With the aim of clarifying the experiments, we have defined an Enlarged Training Set (ETS)—a set formed by only a reduced number of original instances enlarged by a varying number of surrogates obtained from them. We also define an Original Training Set (OTS) which is only formed by original instances. We have considered four different types of training sets:

- OTS, the original training set formed by only original instances.
- ETS-ICGFT, an enlarged training set formed by only a reduced number of original instances enlarged by a varying number of surrogates obtained from them by the ICGFT method (Algorithm 1).
- ETS-IAAFT, an enlarged training set formed by only a reduced number of original instances (real part) enlarged by a varying number of surrogates obtained from them by the IAAFT method.
- ETS-*n*IGFT, an enlarged training set formed by only a reduced number of the original instances (real part) enlarged by a varying number of surrogates obtained from them by the method proposed in [11], which uses the (real) GFT. This method does not iterate as no corrections are made in the original graph domain, so we call it non-Iterative GFT (*n*IGFT).

Several details are relevant about the implementation. Thus, in IAAFT the discrete Fourier transform (DFT) was computed having the same size as the transformed column so that the original domain dimension can be preserved once the inverse DFT is computed. On the other hand, ICGFT is conceived to work with complex graph signals, so we first compute the analytic signal corresponding to each one of the 42 columns of the instance matrix (this is a resource that we can always resort to in applying ICGFT to real instances, thus benefiting from the unconstrained phase randomization of the CGFT in contrast with the limited sign randomization of the GFT). The analytic signal is a complex signal, whose real part is the original real signal and whose imaginary part is the Hilbert transform of the original real signal. It can be computed by using a Hilbert transform filter or by Fourier transforming the original real signal, retaining only the spectral content corresponding to the positive frequencies and inverse Fourier transforming. Hence, the analytic signal keeps the information of the original real signal (whose Fourier transform is symmetric), while its complex nature allows the application of ICGFT. Then the graph signal is built by assigning every complex component of the corresponding instance column to the node of a graph. Regarding the Laplacian matrix required in Algorithm 1, it was computed from an adjacency matrix defined as

$$\mathbf{W}(n, m) = \begin{cases} 1 \pm j & |n - m| = 1 \\ 0 & \text{otherwise} \end{cases} = \mathbf{W}^*(n, m), \quad (19)$$

where the signs of the imaginary part have been selected as indicated in Section 4.2 (the real part of Equation (19) is retained to implement *n*IGFT). This is made in an effort to replicate the local behavior of the original graph signals (neighbors are connected), but many other options are possible. In particular, we could try to replicate some special properties of the original instances which are known to be relevant for the final goal of classification. This normally should be made with the collaboration of an application domain specialist. Moreover, notice that the selection of normalized weights of magnitude equal 1 in Equation (19) implicitly indicates that no preference is given to the connection of some neighbor samples with respect to other neighbor samples. Actually, the magnitude of the weights connecting neighbor samples is an additional degree of freedom that could be exploited to better replicate in the surrogates some possible nonstationarity of the original signals (e.g., some neighbor samples are much more correlated than others). This requires a detailed analysis of the original signals dynamics, and a non-obvious procedure to deduce the weight magnitudes for all the graph connections. We have preferred instead to keep the simplicity and the general applicability of Equation (19), and to recover the signal dynamics by an “ad hoc” correction step as explained in the following.

In hand gesture recognition, the dynamics of the signals must be captured by the classifier. This is done by the HMM by first determining the successive states through which the signals are passing. Recovering the empirical distribution of the original graph signals may seriously affect the signal dynamics, so we have considered in this case a different type of correction, namely, recovering the original instantaneous frequency, which is a largely accepted measure of the dynamic of complex signals [54]. In such a way we take advantage of working in the complex domain and the “waveform” evolution of the original graph signals and the surrogates will be much more comparable. This specific correction can be expressed in the form

$$\mathbf{s}^{(i)} = \mathbf{C}^{(i)} \hat{\mathbf{s}}^{(i)} \quad i = 1 \dots I, \quad (20)$$

where $\mathbf{C}^{(i)}$ is a diagonal matrix having elements $\mathbf{C}^{(i)}(n, n) = \exp\left(j\left(\angle s_n - \angle \hat{s}_n^{(i)}\right)\right)$.

Two experiments have been conceived to illustrate the interest of using surrogates in two representative scenarios. The first one will be that of overall scarcity: We will assume that only a reduced number of original instances are available in every class, then the instance set will be enlarged by surrogates to improve the quality of the training. In the second scenario, we will consider the

imbalanced situation in which there is a significantly smaller number of available instances in some classes than in others.

5.2. Experiment 1. Enlarging the Original Training Set: Overall Scarcity

In this experiment, we consider that only a reduced number of instances are available for training in every class, specifically 10. Then this reduced set is enlarged by adding surrogate instances of it to form an ETS. Figure 1 shows the mean error rate achieved by the different surrogating methods as a function of the ETS size, from 10 (only the original 10 training instances per class) to 30 (10 original training instances plus 20 surrogates per class). The mean error rate was evaluated over 10 original instances, which are different from the ones available from training. The error rates of four partitions were averaged. In every partition, the surrogates were generated from a reduced set of 10 original instances to reproduce the practical case of scarcity. We have also included in Figure 1, for comparison, the curves corresponding to an OTS having the same size as the corresponding ETS.

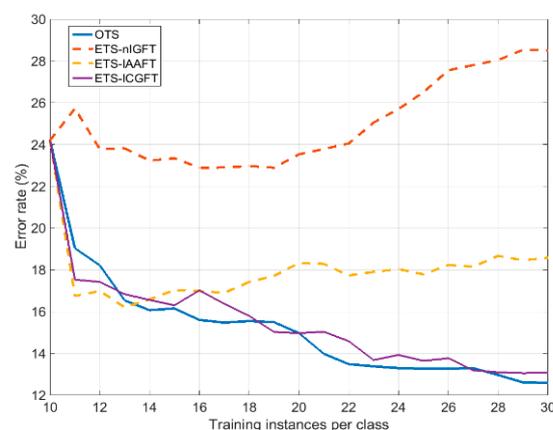


Figure 1. The classification error rate for increasing the number of training instances.

We can see in Figure 1 that training with ETS-ICGFT yields similar results to training with OTS, so that in case of scarcity we can add surrogates keeping the training quality. This is not the same when training with ETS-IAAFT, because, after an initial decrease, the error rate increases with the number of added surrogates. Something similar happens by training with ETS-nIGFT, but with much larger error rates. This suggests that ICGFT is able to implement more realistic surrogates and a better oversampling of the instance space. Table 2 compares the error rates corresponding to the use of an ETS of the maximum size (10 original + 20 surrogates) for every class, obtained from different surrogating methods. We also include the error rates corresponding to OTS of size 10 and OTS of size 30.

Table 2. Error rates for different original and enlarged training sets.

	OTS10	ETS-nIGFT	ETS-IAAFT	ETS-ICGFT	OTS30
Error rate (%)	24.1	28.5	18.58	13.07	12.6

5.3. Experiment 2. Enlarging the Original Training Set: Imbalanced Classes

In this experiment, we assume that the number of original training instances is different for every class. Thus, we will consider that 30 original training instances are available for half the classes, while only a number $N \leq 30$ is available for the other half (the same for all of them). Figure 2 shows the results corresponding to this experiment: Mean error rate (4 partitions) as a function of the imbalance number. This latter is defined as the difference $30 - N$, where N varies from 30 (no imbalance) to 10 (largest imbalance). In Figure 2 we have represented the curve corresponding to training with just the original (imbalanced) training set (OTS), as expected the error rate (evaluated again on a set of

10 original instances different from the ones available from training) significantly increases with the imbalance number. Then we show in Figure 2 the error rates obtained by enlarging the deficient class training sets with $30 - N$ surrogate instances, so that all classes have 30 instances for training.

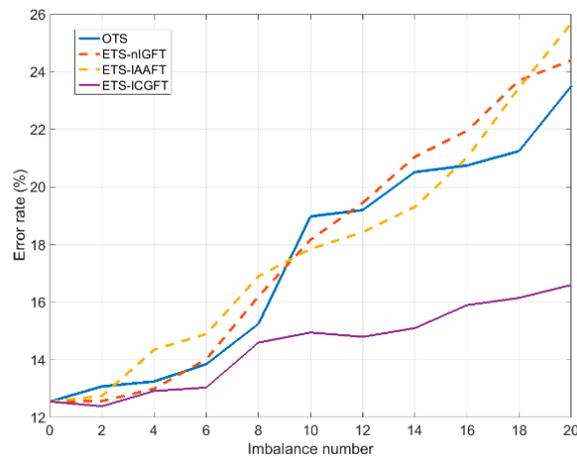


Figure 2. The classification error rate for increasing the imbalance number.

We can see in Figure 2 that only ETS-ICGFT is able to significantly reduce the error corresponding to training with OTS, thus alleviating the problem of imbalanced training. Table 3 shows the error rate averaged over the imbalance number domain for the OTS and for the ETS obtained from the different surrogating methods.

Table 3. Error rate averaged over the imbalance number for the original and enlarged training sets.

	OTS	ETS-nIGFT	ETS-IAAFT	ETS-ICGFT
Error rate (%)	17.47	17.91	17.92	14.82

6. Conclusions

We have presented a new algorithm for generating surrogates. By considering that instances of any arbitrary domain are defined on the vertices of a graph, much more flexibility can be obtained by the new algorithms in comparison with traditional surrogating methods based on the FT. This flexibility emanates from the definition of the graph topology by the Laplacian matrix. We have considered the general case of surrogating complex graph signals, so a Hermitian Laplacian matrix and an associated CGFT have been defined. Some theoretical analysis and interpretations regarding these new definitions have been made accordingly. This may have a general interest in the new area of GSP. Moreover, by means of the concept of GSA invariants, we have deduced significant properties of the original graph signals that are preserved into the surrogates.

Experiments on automatic hand gesture recognition have confirmed the efficacy of enlarging the original training set with surrogates and that the best results correspond to the use of ICGFT.

Author Contributions: The work reported here was developed in collaboration among all authors. All authors have contributed to the preparation of the manuscript, and have approved it. Conceptualization, L.V. and A.S.; data curation, J.B., G.S. and Z.P.; formal analysis, J.B. and L.V.; funding acquisition, L.V. and A.S.; investigation, J.B., G.S. and A.S.; software, J.B. and G.S.; supervision, L.V. and A.S.; writing—Original draft, J.B. and L.V.; writing—Review and editing, J.B. and L.V.

Funding: This research was funded by the Spanish Administration and the European Union under grant TEC2017-84743-P.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

Appendix A. Derivation of Equation (2)

Let us separate the Hermitian adjacency matrix into its real and imaginary parts: $\mathbf{W} = \mathbf{W}_R + j\mathbf{W}_I$. It is clear that \mathbf{W}_R is symmetric with non-negative real elements and \mathbf{W}_I is antisymmetric with real elements, more specifically

$$\begin{aligned} \mathbf{W}_R(n, m) &= \text{Re}[\mathbf{W}(n, m)] = \text{Re}[\mathbf{W}(m, n)] \geq 0, \\ \mathbf{W}_I(n, m) &= \text{Im}[\mathbf{W}(n, m)] = -\text{Im}[\mathbf{W}(m, n)]. \end{aligned} \tag{A1}$$

To ease the notation we define $\text{Re}[\mathbf{W}(n, m)] \equiv w_{nm}^R$ $\text{Im}[\mathbf{W}(n, m)] \equiv w_{nm}^I$. Let us compute the Hermitian Laplacian quadratic form:

$$\begin{aligned} S(\mathbf{s}) &= \mathbf{s}^H \mathbf{L} \mathbf{s} = \mathbf{s}^H (\mathbf{D} - \mathbf{W}) \mathbf{s} \\ &= \mathbf{s}^H (\mathbf{D} - \mathbf{W}_R) \mathbf{s} - j \mathbf{s}^H \mathbf{W}_I \mathbf{s} \end{aligned} \tag{A2}$$

But

$$\begin{aligned} &\mathbf{s}^H (\mathbf{D} - \mathbf{W}_R) \mathbf{s} \\ &= \sum_{n=1}^N \sum_{m=1}^N w_{nm}^R (|s_n|^2 - s_n^* s_m) \\ &= \sum_{n=1}^N \sum_{m=1}^N w_{nm}^R |s_n - s_m|^2 - \sum_{n=1}^N \sum_{m=1}^N w_{nm}^R (|s_m|^2 - s_n s_m^*) \\ &= \sum_{n=1}^N \sum_{m=1}^N w_{nm}^R |s_n - s_m|^2 - \sum_{n=1}^N \sum_{m=1}^N w_{nm}^R (|s_m|^2 - s_m^* s_n) \\ &= \sum_{n=1}^N \sum_{m=1}^N w_{nm}^R |s_n - s_m|^2 - (\mathbf{s}^H (\mathbf{D} - \mathbf{W}_R) \mathbf{s}). \end{aligned} \tag{A3}$$

where we have made use of the symmetry of \mathbf{W}_R , i.e., $w_{nm}^R = w_{mn}^R$. From Equation (A3) we conclude that

$$\begin{aligned} &\mathbf{s}^H (\mathbf{D} - \mathbf{W}_R) \mathbf{s} \\ &= \frac{1}{2} \sum_{n=1}^N \sum_{m=1}^N w_{nm}^R |s_n - s_m|^2 \\ &= \sum_{n=1}^{N-1} \sum_{m>n}^N w_{nm}^R |s_n - s_m|^2. \end{aligned} \tag{A4}$$

On the other hand

$$\begin{aligned} \mathbf{s}^H \mathbf{W}_I \mathbf{s} &= \sum_{n=1}^N \sum_{m=1}^N w_{nm}^I s_n^* s_m \\ &= \sum_{n=1}^{N-1} \sum_{m>n}^N (w_{nm}^I s_n^* s_m + w_{mn}^I s_m^* s_n) \\ &= \sum_{n=1}^{N-1} \sum_{m>n}^N w_{nm}^I (s_n^* s_m - s_m^* s_n) \\ &= \sum_{n=1}^{N-1} \sum_{m>n}^N w_{nm}^I j 2 \text{Im} [s_n^* s_m]. \end{aligned} \tag{A5}$$

In the first equality of Equation (A5) we have taken into account that $w_{nn}^I = 0 \ \forall n$, and in the second equality that $w_{nm}^I = -w_{mn}^I \ \forall n \neq m$. So finally we obtain:

$$\begin{aligned} S(\mathbf{s}) &= \mathbf{s}^H \mathbf{L} \mathbf{s} = \\ &= \sum_{n=1}^{N-1} \sum_{m>n}^N w_{nm}^R |s_n - s_m|^2 + \sum_{n=1}^{N-1} \sum_{m>n}^N 2w_{nm}^I \text{Im} [s_n^* s_m]. \end{aligned} \tag{A6}$$

Appendix B. Derivation of Equation (18)

The Schur–Horn theorem [55] applied to the Hermitian Laplacian states that

$$\sum_{n=1}^M (\lambda_n - l_n) \geq 0 \quad M \leq N. \quad (\text{A7})$$

where λ_n and l_n $n = 1 \dots N$ are, respectively, the eigenvalues and the main diagonal elements of the Hermitian Laplacian sorted in non-increasing order. Considering in Equation (A7) $M = 1$ we obtain

$$\lambda_1 \geq l_1, \quad (\text{A8})$$

which can be written $\lambda_{\max} \geq l_{\max}$. On the other hand, for $M = N$ Equation (A7) holds with equality, then we can write:

$$\lambda_N + \sum_{n=1}^{N-1} \lambda_n = l_N + \sum_{n=1}^{N-1} l_n. \quad (\text{A9})$$

Hence $\sum_{n=1}^{N-1} (\lambda_n - l_n) = l_N - \lambda_N \geq 0$ and therefore $\lambda_N \leq l_N$, which can be written $\lambda_{\min} \leq l_{\min}$.

We know that for any vector \mathbf{s}

$$\lambda_{\max} \geq \frac{S(\mathbf{s})}{\mathbf{s}^H \mathbf{s}}. \quad (\text{A10})$$

Hence, given a vector \mathbf{s} we may write a lower bound for the spectrum bandwidth $B = \lambda_{\max} - \lambda_{\min}$

$$B \geq \max\left(l_{\max}, \frac{S(\mathbf{s})}{\mathbf{s}^H \mathbf{s}}\right) - l_{\min}. \quad (\text{A11})$$

Appendix C. List of Acronyms

FT	Fourier Transform
GSP	Graph Signal Processing
GFT	Graph Fourier Transform
CGFT	Complex Graph Fourier Transform
GSA	Graph Spectrum Amplitude
GWSS	Graph Wide-Sense Stationary
AAFT	Amplitude Adjusted Fourier Transform surrogating algorithm
IAAFT	Iterative AAFT surrogating algorithm
ICGFT	Iterative CGFT surrogating algorithm
nIGFT	non-Iterative GFT surrogating algorithm
DFT	Discrete Fourier Transform
HMM	Hidden Markov Models
OTS	Original Training Set
ETS	Enlarged Training Set
GMM	Gaussian Mixture Model

References

- Schreiber, T.; Schmitz, A. Surrogate time series. *Phys. D Nonlinear Phenom.* **2000**, *142*, 346–382. [[CrossRef](#)]
- Maiwald, T.; Mammen, E.; Nandi, S.; Timmer, J. Surrogate data—A qualitative and quantitative analysis. In *Mathematical Methods in Signal Processing and Digital Image Analysis*; Dahlhaus, R., Kurths, J., Maass, P., Timmer, J., Eds.; Springer: New York, NY, USA, 2008; pp. 41–74.
- Miralles, R.; Vergara, L.; Salazar, A.; Igual, J. Blind Detection of Nonlinearities in Ultrasonic Grain Noise. *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **2008**, *55*, 637–647. [[CrossRef](#)] [[PubMed](#)]

4. Mandic, D.P.; Chen, M.; Gautama, T.; Van Hulle, M.M.; Constantinides, A. On the characterization of the deterministic/stochastic and linear/nonlinear nature of time series. *Proc. R. Soc. A* **2008**, *464*, 1141–1160. [[CrossRef](#)]
5. Araújo, R.; Small, M.; Fernandes, R. Testing for Linear and Nonlinear Gaussian Processes in Nonstationary Time Series. *Int. J. Bifurc. Chaos* **2015**, *25*, 1550013.
6. Jelfs, B.; Vayanos, P.; Javidi, S.; Goh, V.S.L.; Mandic, D.P. Collaborative adaptive filters for online knowledge extraction and information fusion. In *Signal Processing Techniques for Knowledge Extraction and Information Fusion*; Mandic, D., Golz, M., Kuh, A., Obradovic, D., Tanaka, T., Eds.; Springer: Boston, MA, USA, 2007; pp. 3–21.
7. Borgnat, P.; Flandrin, P.; Honeine, P.; Richard, C.; Xiao, J. Testing stationarity with surrogates: A time-frequency approach. *IEEE Trans. Signal Process.* **2010**, *58*, 3459–3470. [[CrossRef](#)]
8. Shuman, D.I.; Narang, S.K.; Frossard, P.; Ortega, A.; Vandergheynst, P. The emerging field of signal processing on graphs: Extending high-dimensional data analysis to networks and other irregular domains. *IEEE Signal Process. Mag.* **2013**, *30*, 83–98. [[CrossRef](#)]
9. Sandryhaila, A.; Moura, J.M.F. Discrete signal processing on graphs. *IEEE Trans. Signal Process.* **2013**, *61*, 1644–1656. [[CrossRef](#)]
10. Sandryhaila, A.; Moura, J.M.F. Big Data Analysis with signal processing on graphs. *IEEE Signal Process. Mag.* **2014**, *31*, 80–90. [[CrossRef](#)]
11. Pirondini, E.; Vybornova, A.; Coscia, M.; Van De Ville, D. A Spectral Method for Generating Surrogate Graph Signals. *IEEE Signal Process. Lett.* **2016**, *13*, 1275–1278. [[CrossRef](#)]
12. Sandryhaila, A.; Moura, J.M.F. Discrete signal processing on graphs: Frequency analysis. *IEEE Trans. Signal Process.* **2014**, *62*, 3042–3054. [[CrossRef](#)]
13. Shuman, D.I.; Ricaud, B.; Vandergheynst, P. Vertex-frequency analysis on graphs. *Appl. Comput. Harmonic Anal.* **2016**, *40*, 260–291. [[CrossRef](#)]
14. Spielman, D. Spectral graph theory. In *Combinatorial Scientific Computing*; Naumann, U., Schnek, O., Eds.; Chapman and Hall/CRC Press: Boca Raton, FL, USA, 2012; Chapter 16; pp. 1–23.
15. Hu, C.; Cheng, L.; Sepulcre, J.; Fakhri, G.E.; Lu, Y.M.; Li, Q. A graph theoretical regression model for brain connectivity learning of Alzheimer’s disease. In Proceedings of the IEEE 10th International Symposium on Biomedical Imaging, California, CA, USA, 7–11 April 2013; pp. 616–619.
16. Dong, X.; Thanou, D.; Frossard, P.; Vandergheynst, P. Learning Laplacian matrix in smooth graph signal representations. *IEEE Trans. Signal Process.* **2016**, *64*, 6160–6173. [[CrossRef](#)]
17. Zhang, C.; Florencio, D.; Chou, P.A. *Graph Signal Processing-A Probabilistic Framework*; Technical Report MSR-TR-2015-31; Microsoft Res.: Redmond, WA, USA, 2015.
18. Pávez, E.; Ortega, A. Generalized precision matrix estimation for graph signal processing. In Proceedings of the IEEE International Conference on Acoustics, Speech Signal Process (ICASSP), Sanghai, China, 20–25 March 2016; pp. 6350–6354.
19. Girault, B. Stationary graph signals using an isometric graph translation. In Proceedings of the 23rd European Signal Processing Conference (EUSIPCO), Nice, France, 31 August–4 September 2015; pp. 1516–1520.
20. Perraudin, N.; Vandergheynst, P. Stationary Signal processing on graphs. *IEEE Trans. Signal Process.* **2017**, *65*, 3462–3477. [[CrossRef](#)]
21. Yu, G.; Qu, H. Hermitian Laplacian Matrix and positive of mixed graphs. *Appl. Math. Comput.* **2015**, *269*, 70–76. [[CrossRef](#)]
22. Wilson, R.C.; Hancock, E.R. Spectral Analysis of Complex Laplacian Matrices. In *Structural, Syntactic and Statistical Pattern Recognition LNCS*; Fred, A., Caelli, T.M., Duin, R.P.W., Campilho, A., de Ridder, D., Eds.; Springer: Berlin/Heidelberg, Germany, 2004; pp. 57–65.
23. Gilbert, G.T. Positive definite matrices and Sylvester’s criterion. *Am. Math. Mon.* **1991**, *98*, 44–46. [[CrossRef](#)]
24. Merris, R. Laplacian matrices of a graph: A survey. *Linear Algebra Appl.* **1994**, *197*, 143–176. [[CrossRef](#)]
25. Zhang, X.D. The Laplacian Eigenvalues of Graphs: A Survey. In *Linear Algebra Research Advances*; Ling, G.D., Ed.; Nova Science Publishers Inc.: New York, NY, USA, 2007; pp. 201–228.
26. Shapiro, H. A survey of canonical forms and invariants for unitary similarity. *Linear Algebra Appl.* **1991**, *147*, 101–167. [[CrossRef](#)]
27. Futorny, V.; Horn, R.A.; Sergeichuk, V.V. Spetch’s criterion for systems of linear mapping. *Linear Algebra Appl.* **2017**, *519*, 278–295. [[CrossRef](#)]

28. Mazumder, R.; Hastie, T. The graphical lasso: New insights and alternatives. *Electr. J. Stat.* **2012**, *6*, 2125–2149. [[CrossRef](#)]
29. Baba, K.; Shibata, R.; Sibuya, M. Partial correlation and conditional correlation as measures of conditional independence. *Austr. New Zeal. J. Stat.* **2004**, *46*, 657–664. [[CrossRef](#)]
30. Chen, X.; Xu, M.; Wu, W.B. Covariance and precision matrix estimation for high-dimensional time series. *Ann. Stat.* **2013**, *41*, 2994–3021. [[CrossRef](#)]
31. Öllerer, V.; Croux, C. Robust high-dimensional precision matrix estimation. In *Modern Multivariate and Robust Methods*; Nordhausen, K., Taskinen, S., Eds.; Springer: New York, NY, USA, 2015; pp. 329–354.
32. Theiler, J.; Eubank, S.; Longtin, A.; Galdrikian, B.; Farmer, J.D. Testing for nonlinearity in time series: The method of surrogate data. *Phys. D* **1992**, *58*, 77–94. [[CrossRef](#)]
33. Schreiber, T.; Schmitz, A. Improved surrogate data for nonlinearity tests. *Phys. Rev. Lett.* **1996**, *77*, 635–638. [[CrossRef](#)] [[PubMed](#)]
34. Mammen, E.; Nandi, S.; Maiwald, T.; Timmer, J. Effect of Jump Discontinuity for Phase-Randomized Surrogate Data Testing. *Int. J. Bifurc. Chaos* **2009**, *19*, 403–408. [[CrossRef](#)]
35. Lucio, J.H.; Valdes, R.; Rodriguez, L.R. Improvements to surrogate data methods for nonstationary time series. *Phys. Rev. E* **2012**, *85*, 056202. [[CrossRef](#)] [[PubMed](#)]
36. Schreiber, T. Constrained randomization of time series data. *Phys. Rev. Lett.* **1998**, *80*, 2105–2108. [[CrossRef](#)]
37. Prichard, D.; Theiler, J. Generating surrogate data for time series with several simultaneously measured variables. *Phys. Rev. Lett.* **1994**, *73*, 951–954. [[CrossRef](#)] [[PubMed](#)]
38. Borgnat, P.; Abry, P.; Flandrin, P. Using surrogates and optimal transport for synthesis of stationary multivariate series with prescribed covariance function and non-Gaussian joint distribution. In Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Kyoto, Japan, 25–30 March 2012; pp. 3729–3732.
39. Salazar, A.; Safont, G.; Vergara, L. Surrogate Techniques for Testing Fraud Detection Algorithms in Credit Card Operations. In Proceedings of the IEEE International Carnahan Conference on Security Technology (ICCSR), Rome, Italy, 13–16 September 2014.
40. Hirata, Y.; Suzuki, H.; Aihara, K. Wind Modelling and its Possible Application to Control of Wind Farms. In *Signal Processing Techniques for Knowledge Extration and Information Fusion*; Mandic, D., Golz, M., Kuh, A., Obradovic, D., Tanaka, T., Eds.; Springer: New York, NY, USA, 2008; pp. 23–35.
41. Belda, J.; Vergara, L.; Salazar, A.; Safont, G. Estimating the Laplacian matrix of Gaussian mixtures for signal processing on graphs. *Signal Process.* **2018**, *148*, 241–249. [[CrossRef](#)]
42. Belda, J.; Vergara, L.; Safont, G.; Salazar, A. Computing the Partial Correlation of ICA Models for Non-Gaussian Graph Signal Processing. *Entropy* **2019**, *21*, 22. [[CrossRef](#)]
43. Gray, R.M. *Toeplitz and Circulant Matrices: A Review*; Information System Laboratory, Stanford University: Stanford, CA, USA, 1971.
44. Liao, T.W. Classification of weld flaws with imbalanced class data. *Expert Syst. Appl.* **2008**, *35*, 1041–1052. [[CrossRef](#)]
45. Song, S.J.; Shin, Y.G. Eddy current flaw characterization in tubes by neural networks and finite element modeling. *NDT&E Int.* **2000**, *33*, 233–243.
46. Bhattacharyya, S.; Jha, S.; Tharakunnel, K.; Westland, J.C. Data mining for credit card fraud: A comparative study. *Decis. Support Syst.* **2011**, *50*, 602–613. [[CrossRef](#)]
47. Kumar, M.N.A.; Sheshadri, H.S. On the classification of imbalanced datasets. *Int. J. Comput. Appl.* **2012**, *44*, 1–7.
48. Bardet, J.; Lang, G.; Oppenheim, G.; Philippe, A.; Taqqu, M. Generators of Long-Range Dependent Processes: A Survey. In *Long-Range Dependence: Theory and Applications*; Doukhan, P., Oppenheim, G., Taqqu, M., Eds.; Birkhauser: Boston, MA, USA, 2003; pp. 579–623.
49. Mitra, S.; Acharya, T. Gesture Recognition: A Survey. *IEEE Trans. Syst. Man Cybern.* **2007**, *37*, 311–324. [[CrossRef](#)]
50. Marin, G.; Dominio, F.; Zanuttigh, P. Hand Gesture Recognition with Leap Motion and Kinect Devices. In Proceedings of the IEEE International Conference on Image Processing (ICIP), Paris, France, 27–30 October 2014; pp. 1565–1569.

51. Moni, M.A.; Shawkat, A.B.M. HMM based hand gesture recognition: A review on techniques and approaches. In Proceedings of the IEEE International Conference on Computer Science and Information Technology (ICCSIT), Beijing, China, 8–12 August 2009; pp. 433–437.
52. Dardas, N.H.; Georganas, N.D. Real-Time Hand Gesture Detection and Recognition Using Bag-of-Features and Support Vector Machine Techniques. *IEEE Trans. Instrum. Meas.* **2011**, *60*, 3592–3606. [[CrossRef](#)]
53. Parcheta, Z.; Martínez-Hinarejos, C.D. Sign language gesture recognition using HMM. In *Pattern Recognition and Image Analysis: 8th Iberian Conference, Proceedings of the Iberian Conference on Pattern Recognition and Image Analysis (IbPRIA), Faro, Portugal, 20–23 June 2017*; Springer: Cham, Switzerland, 2017; pp. 419–426.
54. Boashash, B. Estimating and interpreting the instantaneous frequency of a signal—part 1: Fundamentals. *Proc. IEEE* **1992**, *80*, 520–538. [[CrossRef](#)]
55. Horn, A. Doubly Stochastic Matrices and the Diagonal of a Rotation Matrix. *Am. J. Math.* **1954**, *76*, 620–630. [[CrossRef](#)]



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).