Uncertainty Principle on Graphs

Bastien Pasdeloup, Vincent Gripon, Réda Alami, and Michael G. Rabbat

1 Introduction

In classical signal processing, signals under study are generally defined on very regular domains, such as a path graph for temporal signals like audio, a two-dimensional lattice for spatial signals like images, and a three-dimensional lattice for spatio-temporal signals like video. Frequency analysis of such objects is performed thanks to the Fourier transform operator, which projects the signal under study into a basis of sines, providing a convenient dual representation for it. In the case of more complex signals, such as images or videos, the idea remains the same, except that the sines are now two or three-dimensional.

When the underlying domain on which signals are studied becomes irregular — *e.g.*, a sensor network, a social network, an affinity graph, etc. —, defining an adapted Fourier transform is not as natural. However, providing a frequency representation for signals evolving over such domains is of real interest, and has applications such as anomaly detection [11], brain activity comprehension [12], or topology analysis [4].

In order to study signals over such complex topologies, *graph signal processing* (GSP) has emerged as a field proposing to extend classical signal processing tools to irregular domains modeled as graphs. The field developed from the observation that the basis of eigenvectors associated with the Laplacian matrix of a ring graph are exactly sines, with increasing frequency as the associated eigenvalues increase [19].

Bastien Pasdeloup
EPFL, Route Cantonale, 1015 Lausanne, Switzerland, e-mail: bastien.pasdeloup@epfl.ch

Vincent Gripon
IMT Atlantique, Technopole Brest-Iroise, France, e-mail: vincent.gripon@imt-atlantique.net

Réda Alami
Orange Labs, e-mail: reda.alami@imt-atlantique.net

Michael Rabbat
McGill University, Montréal, Canada, e-mail: michael.rabbat@mcgill.ca

The Fourier transform operator is therefore directly given by this particular matrix, which is strongly tied with the graph modeling the support of information. Interestingly, this correspondence between the eigenvectors of the Laplacian matrix and the Fourier modes also holds for arbitrary graphs. As a matter of fact, the eigenvectors associated with the lowest eigenvalues vary smoothly on the graph, while more important variations appear as the associated eigenvalues increase. Using this so-called Graph Fourier Transform (GFT), numerous tools have been successfully developed, allowing operations such as filtering [19], wavelet decomposition of signals on graphs [9], translation, modulation, etc. (see [19, 14] for overviews of such tools).

Existence of a Fourier transform for signals on graphs also raises the question of uncertainty. In classical signal processing, it is established that a signal cannot be simultaneously localized both in time and in the frequency domain. In this chapter, we present a corresponding uncertainty principle for signals on graphs. We show that the extent to which a signal can be localized in both the graph vertex domain and in the graph spectral domain is tied to the graph topology, and that different graphs have different locality properties.

2 Definitions

In the rest of this chapter, we adopt the following notation. Sets are denoted in calligraphic letters (*e.g.*, $\mathcal{V}$, $\mathcal{E}$). Constants and scalar variables are written in italic font, respectively using capital and lower-case letters (*e.g.*, constant N, and scalar variables i, j). Matrices and vectors are denoted in bold, respectively with capital and lower-case letters, with entries having subscripted indices (*e.g.*, matrix $\mathbf{A}$ and vector $\mathbf{v}$ with entries $\mathbf{A}_{ij}$, $\mathbf{v}_i$).

2.1 Graphs and matrices

Let us consider a graph $\mathcal{G} = \langle \mathcal{V}, \mathcal{E} \rangle$, where $\mathcal{V}$ is the finite set of vertices and $\mathcal{E}$ is the set of edges. Graphs are mathematical models that are useful to describe relations between objects (vertices). In the context of graph signal processing, it often makes sense to consider edges to be pairs of distinct, unordered vertices, meaning that an edge conveys just enough information to denote whether two vertices are connected or not. As such, there are at most $\binom{N}{2}$ edges in a graph containing N vertices. An example of a graph is depicted in Figure 1.

In order to ease readability, let us consider vertices to be indexed from 1 to N: $\mathcal{V} = \{1, 2, \ldots, N\}$.

Any graph can be conveniently described by its adjacency matrix. The adjacency matrix $\mathbf{A}$ of a graph $\mathcal{G} = \langle \mathcal{V}, \mathcal{E} \rangle$ with N vertices is an N-by-N matrix with entries defined as follows:

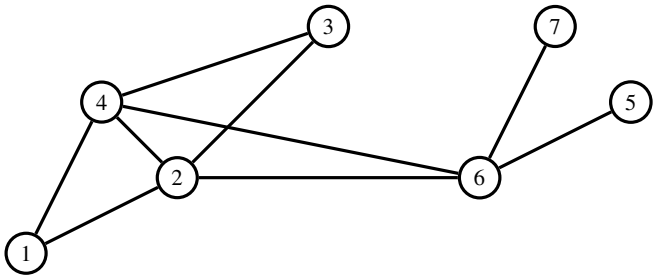


Fig. 1: Example of a graph containing 7 vertices $1,2,\ldots,7$. Vertices are depicted as circles and edges as lines connecting them.

$$\forall i,j \in \mathscr{V} : \mathbf{A}_{ij} = \begin{cases} 1 \text{ if } \{i,j\} \in \mathscr{E} \\ 0 \text{ otherwise.} \end{cases} \tag{1}$$

In the remainder of this chapter, we always consider vertices to be integers between 1 and N, such that we make no distinction between a vertex and its index.

An example graph is shown in Figure 1, and its adjacency matrix is:

$$\mathbf{A} = \begin{bmatrix} 0&1&0&1&0&0&0 \\ 1&0&1&1&0&1&0 \\ 0&1&0&1&0&0&0 \\ 1&1&1&0&0&1&0 \\ 0&0&0&0&0&1&0 \\ 0&1&0&1&1&0&1 \\ 0&0&0&0&0&1&0 \end{bmatrix}. \tag{2}$$

Note that the adjacency matrix of a graph is symmetric. This is because we chose edges as *pairs* (unordered sets) of vertices. Extended definitions of graphs have been proposed in the literature, including digraphs where edges are *couples* (ordered pairs) of vertices. Digraphs are particularly useful when representing oriented relations between vertices. As a consequence, adjacency matrices of digraphs are not necessarily symmetric. For the rest of the chapter we focus on symmetric graphs.

Other extensions include weighted graphs. With such graphs, the notion of edges is refined to take into account intensities. In this context, we introduce the weight matrix $\mathbf{W}$ which is such that:

$$\forall i,j \in \mathscr{V} : (\mathbf{W}_{ij} > 0) \Rightarrow (\mathbf{A}_{ij} = 1). \tag{3}$$

In other words, the weight matrix of a graph has the same support as the adjacency matrix of the graph. More precisely, it determines the weight of each edge in the graph. If the graph is unweighted, we conveniently adopt the convention that the weight matrix is exactly identical to the adjacency matrix of the graph.

2.2 Basic definitions on graphs

In this section, we introduce some definitions from graph theory that will be used in the following sections. More complete literature on graphs can be found for instance in [8].

Let us first introduce paths and walks on a graph.

Definition 1. A *walk* on a graph is a (possibly infinite) sequence of vertices, such that any two consecutive vertices form an edge in the graph. Getting back to the example of Figure 1, the sequence $(1,2,6,2,3,4)$ is a walk.

Definition 2. A *path* on a graph is a walk in which each consecutive (unordered) pair of vertices appears at most once. As such, paths are necessarily finite because there is only a finite number of possible pairs of vertices. The starting and ending vertices of a path are called its *extremities*. An example of a path for the graph in Figure 1 is $(1,2,4,6)$.

Paths and walks are often confused in the literature due to their very similar definitions. Walks are shorter to define, but paths may include cycles.

Definition 3. A *cycle* on a graph is a path with identical extremities. Not all graphs admit cycles. An example of a cycle for the graph in Figure 1 is $(1,2,4,1)$.

Definition 4. The *length* of a path is the number of vertices in the sequence minus 1. For example, the length of the path $(1,2,4,6)$ is 3. This is also the number of edges traversed in the path.

Definition 5. The *weight* of a path is the sum of the weights of edges formed by consecutive vertices in the path.

Weights and length are two separate notions for weighted graphs. Thanks to our previously mentioned convention, in the case of unweighted graphs they are identical.

Definition 6. A *connected* graph is a graph for which every pair of vertices are extremities of at least one path.

Definition 7. The *geodesic distance* $d_{\mathscr{G}}$ on a graph $\mathscr{G}$ is a function that associates a pair of vertices with the minimum weight of a path having these vertices as its extremities.

Based on these definitions, one can define classical families of graphs:

Definition 8. A *tree* is a connected graph that contains no cycle.

Definition 9. A *bipartite* graph is a graph that contains no cycle with odd length.

Definition 10. A *ring* graph is an unweighted graph with N vertices in which all edges appear in a single cycle of length N.

Definition 11. A *complete* graph is an unweighted graph containing all possible edges.

Definition 12. A *star* graph is an unweighted graph with N vertices and $N-1$ edges for which all edges have one extremity in common.

2.3 Meaning of edge weights

It is very important to understand that the convention we use here to put 0s in the matrices when there is no link is not without consequence. Indeed, we suppose a weight 0 is equivalent to the absence of an edge between the corresponding vertices. As such, weights should not represent quantities for which this is a contradiction.

For example, consider a graph in which vertices model cities and connection weights represent road distances between these cities. The absence of a connection between two cities could correspond to the absence of a direct road connecting them, in which case the distance should not be 0, but to the contrary $+\infty$. Such weights do not make sense with respect to our convention.

Now imagine a graph in which vertices are terminals on the Internet and weights represent the number of packets that directly travel between pairs of terminals. In such a graph, a connection weight 0 corresponds to the absence of a direct connection, or to a completely useless one. Such weights make sense with respect to our convention.

There are fine underlying theoretical reasons to explain why we choose this convention of 0s and 1s in the adjacency matrix, and it is mainly related to the fact that we suppose working with the regular linear algebra. Distance graphs, that we discussed before, are better processed using the tropical algebra in which $+\infty$ is a neutral element for the addition [7].

The weighted graphs we introduce in this document typically model the similarity of their corresponding vertices, and not their distances.

2.4 The graph Laplacian and its properties

Consider a connected graph $\mathscr{G}$ with N vertices together with its weight matrix $\mathbf{W}$. We call *strength*[1] of a vertex i the quantity:

$$s(i) = \sum_{j \in \mathscr{V}} \mathbf{W}_{ij} \, . \tag{4}$$

Definition 13. A graph is said to be *regular* if all of its vertices have the same strength.

The strengths of all vertices can be merged into a single diagonal matrix $\mathbf{S}$ called the *strength matrix*, such that:

$$\forall i, j \in \mathscr{V} : \mathbf{S}_{ij} = \begin{cases} s(i) & \text{if } i = j \\ 0 & \text{otherwise.} \end{cases} \tag{5}$$

[1] This quantity is often referred to as *degree* in the literature of GSP. In graph theory the degree refers to the *number* of neighbors of a given vertex whereas its strength takes into account the weights of corresponding edges. These two quantities are identical when considering unweighted graphs.

The strength matrix of a graph is especially useful to perform various kinds of normalizations on this graph. For instance, a common way to transform a graph into a Markov chain consists in considering the matrix $\mathbf{S}^{-1}\mathbf{W}$, of which sums of rows always equal 1.

The strength matrix of a graph can also be used to define the Laplacian of a graph:

Definition 14. The *Laplacian* of a graph $\mathscr{G}$ with weight matrix $\mathbf{W}$ and strength matrix $\mathbf{S}$ is:

$$\mathbf{L} = \mathbf{S} - \mathbf{W} \,. \tag{6}$$

As an example, the Laplacian of the example graph in Figure 1 is:

$$\mathbf{L} = \begin{bmatrix} 2 & -1 & 0 & -1 & 0 & 0 & 0 \\ -1 & 4 & -1 & -1 & 0 & -1 & 0 \\ 0 & -1 & 2 & -1 & 0 & 0 & 0 \\ -1 & -1 & -1 & 4 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & -1 & -1 & 0 \\ 0 & -1 & 0 & -1 & -1 & 4 & -1 \\ 0 & 0 & 0 & 0 & 0 & -1 & 1 \end{bmatrix} \tag{7}$$

Being a symmetric, real-valued matrix, the Laplacian of a graph with N vertices can be written as:

$$\mathbf{L} = \mathbf{F}\boldsymbol{\Lambda}\mathbf{F}^\top \,, \tag{8}$$

where $\mathbf{F}$ is such that $\mathbf{F}\mathbf{F}^{-1} = \mathbf{F}^{-1}\mathbf{F} = \mathbf{I}_N$, $\mathbf{I}_N$ being the identity matrix of dimension N, and $\boldsymbol{\Lambda}$ is a diagonal matrix of which diagonal elements are $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_N$. In other words, $\mathbf{F}$ is a matrix of eigenvectors and $\boldsymbol{\Lambda}$ is a diagonal matrix of eigenvalues, arranged in ascending order.

The Laplacian of a graph offers multiple interesting properties, as pointed out in the next propositions.

Proposition 1. *Let* $\mathbf{L}$ *be the Laplacian matrix of a graph* $\mathscr{G}$*. The vector* $\mathbf{1}$ *with all entries equal to* 1 *is an eigenvector of* $\mathbf{L}$ *associated with the eigenvalue* 0*.*

Proof. The proof is straightforward:

$$\forall i \in \mathscr{V} : (\mathbf{L}\mathbf{1})_i = (\mathbf{S}\mathbf{1})_i - (\mathbf{W}\mathbf{1})_i = s(i) - \sum_{j \in \mathscr{V}} \mathbf{W}_{ij} = 0 \,. \tag{9}$$

Proposition 2. *The eigenvalues of the Laplacian of a graph are all nonnegative.*

Proof. Suppose, for the sake of a contradiction, that some negative eigenvalue exists. Let us denote it λ and let $\mathbf{f}$ be an associated nonzero eigenvector. Let us look at one of the entries i of $\mathbf{f}$ such that $|\mathbf{f}_i|$ is maximum. We obtain:

$$s(i)\mathbf{f}_i - \sum_{j \in \mathscr{V}} \mathbf{W}_{ij}\mathbf{f}_j = \lambda \mathbf{f}_i \,, \tag{10}$$

which can be rewritten as:

$$\mathbf{f}_i - \sum_{j \in \mathscr{V}} \frac{\mathbf{W}_{ij}}{s(i)} \mathbf{f}_j = \frac{\lambda}{s(i)} \mathbf{f}_i \ . \tag{11}$$

Without loss of generality, we can suppose that $\mathbf{f}_i$ is positive. As such, the right term of this equality is negative. We conclude that:

$$\sum_{j \in \mathscr{V}}^{N} \frac{\mathbf{W}_{ij}}{s(i)} \mathbf{f}_j > \mathbf{f}_i \ , \tag{12}$$

which is a contradiction since the left part of this inequality is a weighted average of values of $\mathbf{f}_j$ that are by definition all less than or equal to $\mathbf{f}_i$.

From the two previous propositions we conclude that $\lambda_1 = 0$, in all cases.

Proposition 3. *The second eigenvalue of the Laplacian of a graph $\mathscr{G}$ is* 0 *if and only if the graph is not connected.*

Proof. It is immediate to see that the second eigenvalue is 0 if the graph is not connected. Indeed, fix some vertex $i \in \mathscr{V}$ and consider the set $\mathscr{V}_i \subset \mathscr{V}$ of vertices connected to i and the complement set $\overline{\mathscr{V}_i}$. Then the two linearly independent vectors obtained by putting 1s on coordinates in $\mathscr{V}_i$ and 0s in those of $\overline{\mathscr{V}_i}$, and conversely, are both eigenvectors associated with the eigenvalue 0.

Conversely, consider a non-constant eigenvector associated with eigenvalue 0. Denote it by $\mathbf{f}$, and let i be an index such that $|\mathbf{f}_i|$ is maximum. Since we have:

$$\mathbf{f}_i = \sum_{j \in \mathscr{V}}^{N} \frac{\mathbf{W}_{ij}}{s(i)} \mathbf{f}_j \tag{13}$$

is a weighted sum of the value of the neighbors of vertex i, we conclude that all its neighbors have the same value as i. By repeating this process, we conclude that any vertex connected to i has the same value in $\mathbf{f}$ as i. If the graph were connected, the obtained vector would be constant.

Proposition 4. *The quadratic form of the Laplacian* $\mathbf{L}$ *of a graph is such that:*

$$\mathbf{x}^\top \mathbf{L} \mathbf{x} = \sum_{\{i,j\} \in \mathscr{E}} \mathbf{W}_{ij} \left(\mathbf{x}_i - \mathbf{x}_j\right)^2 \ . \tag{14}$$

Proof. We use simple mathematics here:

$$\begin{aligned}\mathbf{x}^\top \mathbf{L}\mathbf{x} &= \mathbf{x}^\top (\mathbf{S}-\mathbf{W})\mathbf{x} \\ &= \sum_{i\in\mathscr{V}} s(i)\mathbf{x}_i^2 - \sum_{\substack{i,j\in\mathscr{V}\\ i<j}} 2\mathbf{W}_{ij}\mathbf{x}_i\mathbf{x}_j \\ &= \sum_{\substack{i,j\in\mathscr{V}\\ i<j}} \mathbf{W}_{ij}\left(\mathbf{x}_i^2+\mathbf{x}_j^2\right) - \sum_{\substack{i,j\in\mathscr{V}\\ i<j}} 2\mathbf{W}_{ij}\mathbf{x}_i\mathbf{x}_j \\ &= \sum_{\substack{i,j\in\mathscr{V}\\ i<j}} \mathbf{W}_{ij}\left(\mathbf{x}_i-\mathbf{x}_j\right)^2 . \end{aligned} \tag{15}$$

Some authors prefer to use the normalized Laplacian $\mathbf{Ł}$ instead of the Laplacian $\mathbf{L}$, where:

$$\mathbf{Ł} = \mathbf{S}^{-\frac{1}{2}}\mathbf{L}\mathbf{S}^{-\frac{1}{2}} . \tag{16}$$

Quadratic forms involving the normalized Laplacian satisfy a similar relationship:

$$\mathbf{x}^\top \mathbf{Ł}\mathbf{x} = \sum_{\{i,j\}\in\mathscr{E}} \mathbf{W}_{ij}\left(\frac{\mathbf{x}_i}{\sqrt{s(i)}} - \frac{\mathbf{x}_j}{\sqrt{s(j)}}\right)^2 . \tag{17}$$

The normalized Laplacian also satisfies Proposition 3, along with three others:

Proposition 5. *Consider the normalized Laplacian $\mathbf{Ł}$ of a connected graph $\mathscr{G}$, then the spectrum of $\mathbf{Ł}$ is between* 0 *and* 2*.*

Proof. Introduce the Rayleigh coefficient:

$$r(\mathbf{Ł},\mathbf{x}) = \frac{\mathbf{x}^\top \mathbf{Ł}\mathbf{x}}{\mathbf{x}^\top\mathbf{x}} . \tag{18}$$

It is thus sufficient to show that this coefficient is between 0 and 2 [13]. We obtain:

$$\begin{aligned}\frac{\mathbf{x}^\top \mathbf{Ł}\mathbf{x}}{\mathbf{x}^\top\mathbf{x}} &= \frac{\mathbf{x}^\top \mathbf{S}^{-\frac{1}{2}}\mathbf{L}\mathbf{S}^{-\frac{1}{2}}\mathbf{x}}{\mathbf{x}^\top\mathbf{x}} \\ &= \frac{\mathbf{x}^\top \mathbf{L}\mathbf{x}}{\left(\mathbf{S}^{\frac{1}{2}}\mathbf{x}\right)^\top \mathbf{S}^{\frac{1}{2}}\mathbf{x}} \\ &= \frac{\sum_{\substack{i,j\in\mathscr{V}\\ i<j}} \mathbf{W}_{ij}\left(\mathbf{x}_i-\mathbf{x}_j\right)^2}{\sum_{i\in\mathscr{V}} s(i)\mathbf{x}_i^2} . \end{aligned} \tag{19}$$

This quantity is clearly nonnegative. This concludes the proof since, for any real numbers a and b, $(a-b)^2 \leq 2(a^2+b^2)$.

Proposition 6. *The largest eigenvalue of the normalized Laplacian $\mathbf{Ł}$ of a connected unweighted graph $\mathscr{G}$ is* 2 *if and only if the graph is bipartite.*

Proof. The proof is omitted here. See [3] for details.

Proposition 7. *The first eigenvalue of the normalized Laplacian is always* 0*, and it has associated eigenvector* $\mathbf{f}$ *where:*

$$\forall i \in \mathscr{V} : \mathbf{f}_i = \sqrt{\frac{s(i)}{\sum_{j\in\mathscr{V}} s(j)}} \,. \tag{20}$$

Proof. Observe that

$$\mathbf{f} = \frac{\mathbf{S}^{\frac{1}{2}}\mathbf{1}}{\sqrt{\sum_{j\in\mathscr{V}} s(j)}} \,. \tag{21}$$

We have:

$$\begin{aligned}\left[\sqrt{\sum_{j\in\mathscr{V}} s(j)}\right] \text{Ł}\mathbf{f} &= \mathbf{S}^{\frac{1}{2}}\mathbf{1} - \mathbf{S}^{-\frac{1}{2}}\mathbf{W}\mathbf{1} \\ &= 0 \,. \end{aligned} \tag{22}$$

3 Graph signals and Fourier transform

3.1 Graphs and signals

As mentioned in the introduction, a graph is a convenient tool to model the topology of a signal. Consider a snippet of audio for instance, which is a continuous, smooth function of time. Typically, such signal is represented in computer memories using a regular sampling of time. It is thus a collection of values corresponding to the amplitude of the sound measured at distinct, regularly-spaced time steps. With no additional priors about the considered sound, it is reasonable to model smoothness by saying that two consecutive measurements are likely to be similar—at least, more than two measurements separated by more time. A natural representation of the topology of a sound is therefore obtained using a line graph, as depicted in Figure 2 (left). Considering an image instead, and following the same reasoning, a typical graph to model its topology would be a grid, as depicted in Figure 2 (right).

As such, we have a first correspondence between signals and graphs. If a signal is a vector $\mathbf{x} \in \mathbb{R}^N$, then its topology should be modeled by a graph containing N vertices, one per coordinate of the vector.

Of course there is no real interest in introducing graphs to represent the topology of images or sounds. These are just particular examples of topologies. In practice we are interested in other families of graphs, some of which we introduce in the following definitions.

Definition 15. An *Erdős–Rényi* graph with parameters N and P is an unweighted graph with N vertices that is obtained by drawing independently at random each edge with a Bernoulli random variable with parameter P.

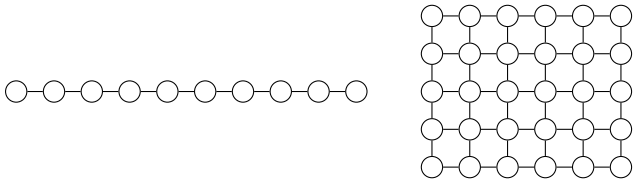


Fig. 2: Example of a line graph (left) and a grid graph (right) that are natural topologies to represent a sound or an image.

As such, an Erdős–Rényi graph with $P = 1$ is a complete graph.

Definition 16. A *random geometric* graph with parameters N and R is an unweighted graph with N vertices that is obtained by drawing uniformly at random 2D coordinates between 0 and 1 for the N vertices. Then vertices which coordinates are less than R apart, considering ℓ_2 norm, are connected through an edge.

Again, choosing $R \geq \sqrt{2}$ leads to a complete graph.

Erdős–Rényi graphs have interesting asymptotic properties, such as possibly being at the same time sparsely connected together with each pairs of vertices at very small distances. Random geometric graphs are often used to describe sensor networks, as they are built using underlying 2D coordinates. Although they may have irregular, complex structure, random geometric graphs often have properties which are very similar to those of a two-dimensional grid.

3.2 Sharpness

Given a signal $\mathbf{x} \in \mathbb{R}^N$, there are $2^{\binom{N}{2}}$ possible unweighted graphs to represent its topology, and an infinite number of weighted graphs. Understanding the relations between graphs and signals requires to quantify these relationships.

Since the graphs we introduce in this chapter typically model similarities between their vertices, it is natural to expect connected vertices to contain similar values. This can be measured with sharpness:

Definition 17. Consider a graph $\mathcal{G}$ with N vertices and normalized Laplacian $\mathbf{Ł}$, and a signal $\mathbf{x} \in \mathbb{R}^N$. The *sharpness*[2] of $\mathbf{x}$ on $\mathcal{G}$ is the quantity:

$$h(\mathbf{x}) = \mathbf{x}^\top \mathbf{Ł} \mathbf{x} \,. \tag{23}$$

Note that sharpness grows quadratically with the norm of a signal:

[2] This quantity is often denoted *smoothness* in the GSP literature. However, the lower this value, the smoother the signal on the graph.

$$\forall \mathbf{x} \in \mathbb{R}^N, \forall \alpha \in \mathbb{R} : \alpha^2 h(\mathbf{x}) = h(\alpha \mathbf{x}) \,, \tag{24}$$

such that we consider in the following the normalized sharpness:

$$\hbar(\mathbf{x}) = \frac{\mathbf{x}^\top \mathbf{Ł} \mathbf{x}}{\|\mathbf{x}\|_2^2} = h\left(\frac{\mathbf{x}}{\|\mathbf{x}\|_2}\right) \,, \tag{25}$$

for nonzero signals, with the convention that $\hbar(\mathbf{0}) = 0$.

More generally, signals with very similar values at well-connected vertices will present a smaller normalized sharpness than signals with significant differences between those vertices. As such, when the normalized sharpness is close to zero, it makes sense to say that the signal is aligned with the graph whereas when sharpness is large it is not. In other words, signals aligned with the graph are signals that are close (in terms of angle) to the span of the first eigenvectors of $\mathbf{Ł}$ (those with smallest eigenvalues).

3.3 Diffusion sequence

Definition 18. Consider a signal $\mathbf{x}$. We call *diffusion sequence* of $\mathbf{x}$ the sequence:

$$\left((\mathbf{I}_N - \mathbf{Ł})^t \mathbf{x}\right)_{t \in \mathbb{N}^*} \,. \tag{26}$$

In other words, the diffusion sequence takes a signal $\mathbf{x}$, and iteratively mixes it using the matrix $\mathbf{I}_N - \mathbf{Ł}$.

Proposition 8. *For any connected graph, the diffusion sequence of a signal* $\mathbf{x}$ *converges to the first eigenvector of its normalized Laplacian (the one associated with eigenvalue* 0*) if the graph is not bipartite.*

Proof. Denote $0 = \lambda_1 < \lambda_2 \leq \cdots \leq \lambda_N$ the eigenvalues of the normalized Laplacian $\mathbf{Ł}$ of a connected graph $\mathscr{G}$, and $\mathbf{f}_1, \ldots, \mathbf{f}_N$ the corresponding unit norm eigenvectors. Then for any unit norm signal $\mathbf{x}$, we obtain:

$$(\mathbf{I}_N - \mathbf{Ł})^t \mathbf{x} = (1-\lambda_1)^t \left(\mathbf{x}^\top \mathbf{f}_1\right) \mathbf{f}_1 + \cdots + (1-\lambda_N)^t \left(\mathbf{x}^\top \mathbf{f}_N\right) \mathbf{f}_N \,. \tag{27}$$

Because of Proposition 6, we have $\forall i \geq 2 : |1 - \lambda_i| < 1$, and thus:

$$\lim_t (\mathbf{I}_N - \mathbf{Ł})^t \mathbf{x} = \left(\mathbf{x}^\top \mathbf{f}_1\right) \mathbf{f}_1 \,. \tag{28}$$

In the case of a bipartite graph, it is interesting to see that in the above expression both $1 - \lambda_1 = 1$ and $1 - \lambda_N = -1$ do not vanish, leading to an almost alternative sequence.

In some sense, as we go through the diffusion sequence of a signal, it is increasingly more bound to the graph, with the extreme case of reaching a sharpness of

0 (using normalized Laplacian) after an infinite number of steps when the graph is not bipartite. In Figure 3, we draw the average normalized sharpness of randomly uniformly generated signals on a path graph, an Erdős–Rényi graph and a random-geometric graph as a function of the number of steps. The path graph is a bipartite graph, which explains why its curve does not seem to converge to 0. But interestingly, we see here that some graphs have faster convergence to 0 than others, which is due to the spectrum of their normalized Laplacian, more concentrated around 1.

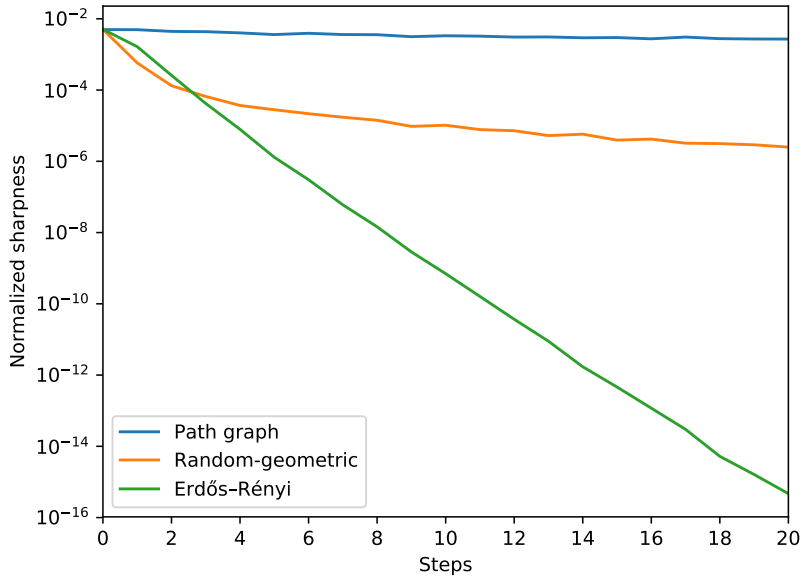


Fig. 3: Normalized sharpness of uniformly drawn random signals, as a function of the number of diffusion steps and for three graphs: path, Erdős–Rényi and random-geometric. All graphs contain exactly 200 vertices. Erdős–Rényi graphs were generated with $P = 0.07$ and random-geometric graphs with $R = 0.2$. Drawn curves were obtained by averaging over 100 graphs.

4 Graph Fourier Transform (GFT)

4.1 Analogy with discrete Fourier transform and definitions

In the case of a ring graph, vertices can be indexed so that the adjacency matrix becomes circulant [6]. For a ring graph with 7 vertices, it would look like:

$$\mathbf{A} = \begin{bmatrix} 0&1&0&0&0&0&1 \\ 1&0&1&0&0&0&0 \\ 0&1&0&1&0&0&0 \\ 0&0&1&0&1&0&0 \\ 0&0&0&1&0&1&0 \\ 0&0&0&0&1&0&1 \\ 1&0&0&0&0&1&0 \end{bmatrix}$$

Of course, it is also the case for its Laplacian matrix, which is:

$$\mathbf{L} = \begin{bmatrix} 2&-1&0&0&0&0&-1 \\ -1&2&-1&0&0&0&0 \\ 0&-1&2&-1&0&0&0 \\ 0&0&-1&2&-1&0&0 \\ 0&0&0&-1&2&-1&0 \\ 0&0&0&0&-1&2&-1 \\ -1&0&0&0&0&-1&2 \end{bmatrix},$$

and the normalized Laplacian is also circulant.

As a consequence, it is possible to obtain a mathematical closed form of the eigenvectors, which in this case are exactly discrete Fourier modes. In other words, consider a signal $\mathbf{x} \in \mathbb{R}^N$, and multiply it by the transpose matrix of eigenvectors $\mathbf{F}$ of the Laplacian of the ring graph, and you obtain the discrete Fourier transform of $\mathbf{x}$.

Following this analogy, the *graph Fourier transform* (GFT) can be defined.

Definition 19. The *Graph Fourier Transform (GFT)* of a signal $\mathbf{x}$ on a graph $\mathscr{G}$ is the operation:

$$\widehat{\mathbf{x}} = \mathbf{F}^\top \mathbf{x}\,. \tag{29}$$

The *Inverse Graph Fourier Transform (IGFT)* of a signal $\widehat{\mathbf{x}}$ on a graph $\mathscr{G}$ is the operation:

$$\mathbf{x} = \mathbf{F}\widehat{\mathbf{x}}\,. \tag{30}$$

4.2 Examples

In order to illustrate some of the previously introduced properties, we consider here a random-geometric graph $\mathscr{G}$, depicted in Figure 4. A uniformly random signal $\mathbf{x}$ has been initialized on its vertices. Its spectrum $\widehat{\mathbf{x}} = \mathbf{F}^\top \mathbf{x}$ (using the eigenvectors of the normalized Laplacian) is also shown.

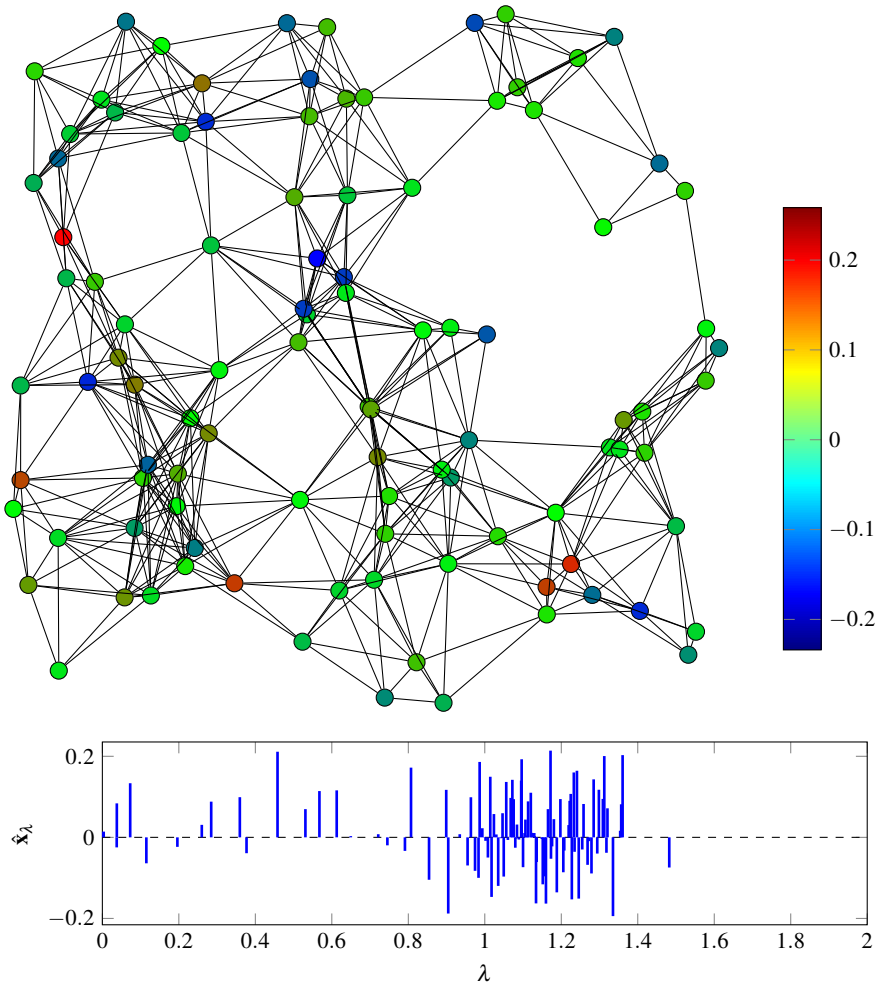


Fig. 4: Example of a random geometric graph on which a uniformly random signal has been initialized (top). The spectrum of this signal in the basis of the normalized Laplacian is also given (bottom). We use the notation $\widehat{\mathbf{x}}_\lambda$ to describe the signal amplitude at eigenvalue λ.

The graph Fourier transform provides two representations of a same signal on the graph, linearly related through the matrix of eigenvectors of the Laplacian matrix. When represented on the graph, the signal is a scalar value observed at each vertex, which generally represents an observed phenomenon to analyze. Its spectral representation gives insights of the properties of this signal, such as its sharpness on the graph or its bandwidth for instance.

Also, these eigenvectors being vectors of dimension N, they can be seen as signals on the graph, revealing some interesting properties of the associated eigenvalues. Figure 5 depicts some of these.

As illustrated, the eigenvectors associated with the lowest eigenvalues of the Laplacian matrix vary smoothly on the graph, while those associated with larger ones tend to model more localized, sharp variations. This illustrates the analogy between the eigenvalues of this matrix and some notion of frequencies in classical signal processing.

When considering a diffusion process, eigenvalues in the Laplacian spectrum that are close to 0 are located next to eigenvalue 1 in the spectrum of the diffusion matrix. As a consequence, they vanish in a slower way compared to Laplacian higher eigenvalues. Since the eigenvectors associated to these eigenvalues are highly localized patterns on the graph, a diffusion phenomenon on a graph can be understood as a process that smoothens signal values across the graph, similarly to heat diffusion in the classical settings. Figure 6 depicts a few diffusion steps of the signal in Figure 4.

5 Graph uncertainty principle

5.1 Classical uncertainty principle

In classical signal processing, the uncertainty principle states that a signal cannot be perfectly localized both in time and frequency [10]. More precisely, define the variance of a measurable function f as:

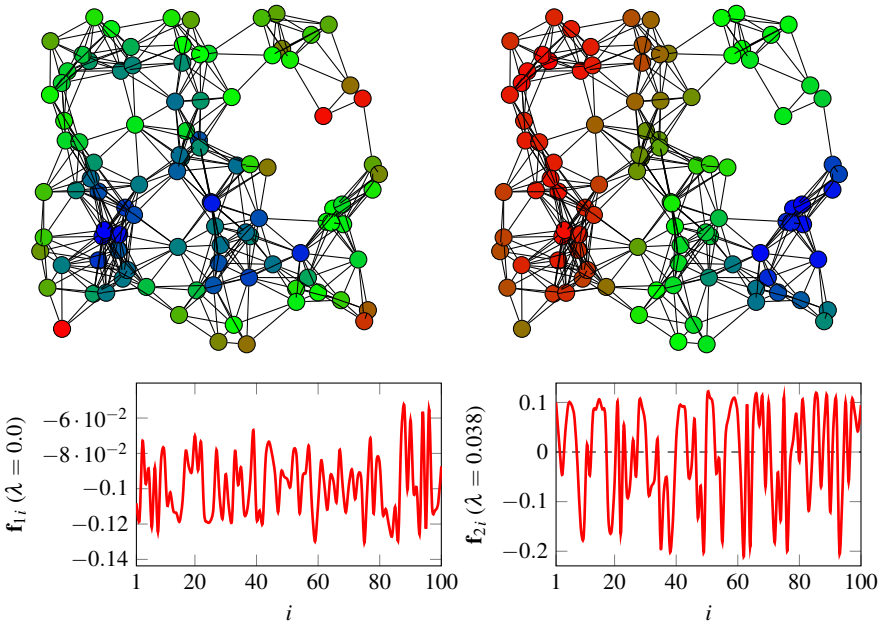
$$v(f) = \inf_{a\in\mathbb{R}} \int (x-a)^2 df \,. \tag{31}$$

Introduce the Fourier transform $\widehat{f}$ of f as:

$$\widehat{f} : \lambda \mapsto \int f(x) e^{-2\pi i x\lambda} dx \,. \tag{32}$$

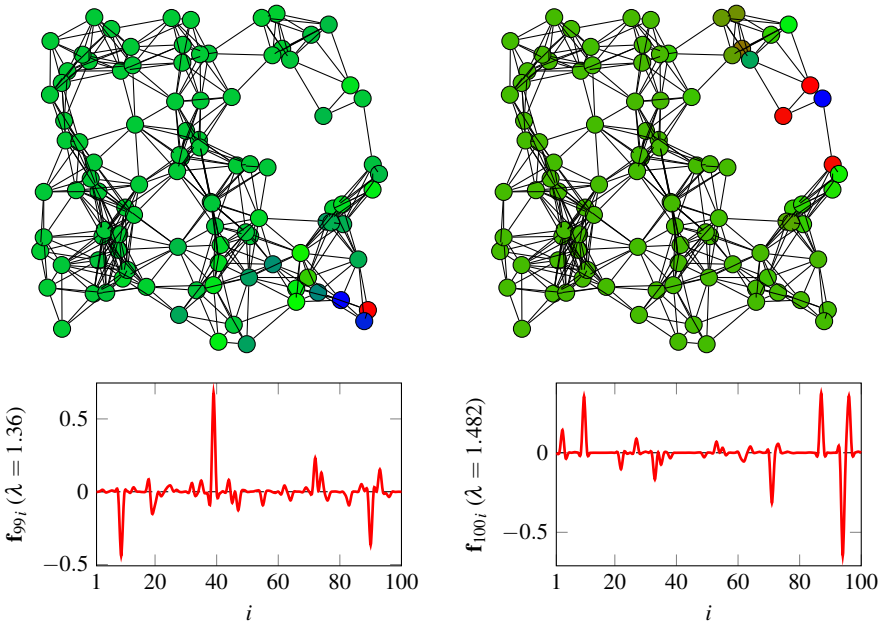
Then the uncertainty principle states that:

$$\forall f \in L^2(\mathbb{R}) : \int f = 1 \Rightarrow v(f)v(\widehat{f}) \geq \frac{1}{2} \,. \tag{33}$$



(a) First eigenvector

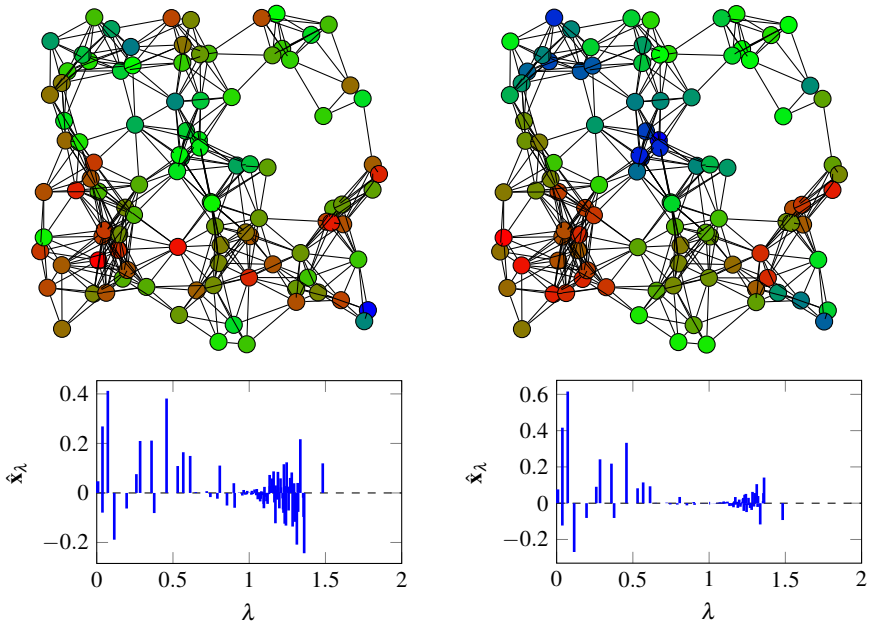
(b) Second eigenvector



(c) Penultimate eigenvector

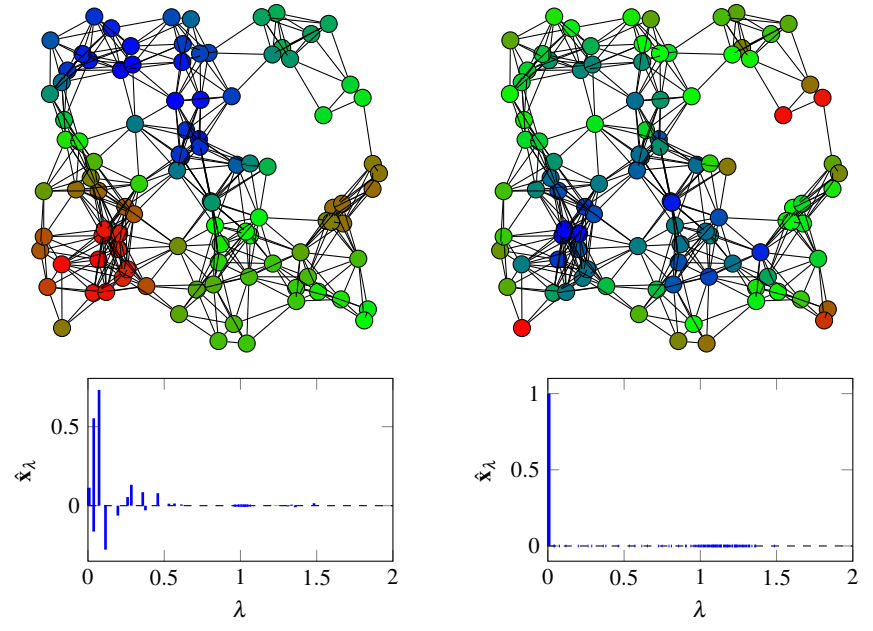
(d) Last eigenvector

Fig. 5: Representation of some of the eigenvectors of the normalized Laplacian (bottom) associated with the graph in Figure 4. Eigenvectors associated with lower eigenvalues correspond to low frequencies, and vary smoothly when used as signals on the graph (top). On the contrary, those associated with large eigenvalues feature some strong local variations when represented in the graph domain.



(a) $t = 1$

(b) $t = 2$



(c) $t = 50$

(d) $t = 1,000$

Fig. 6: Examples of diffusions of the signal in Figure 4, for various values of t, represented both on the graph (top) and in the spectral basis of the normalized Laplacian (bottom). As t increases, spectral components of the signal associated with eigenvalues close to 1 vanish faster than others. Eventually, as t grows to infinity, only the signal spectrum associated with eigenvalue 0 remains. As a consequence, the signal converges to the first eigenvalue of the normalized Laplacian, or its opposite depending on the sign of the initial signal component associated with eigenvalue 0. This can be observed in the present example by comparing the graph signal in (d) with the one in Figure 5 (a).

In other words, the variances of a function f and of its Fourier transform cannot both simultaneously be small (since their product is at least $\frac{1}{2}$). Note that the variance of a function tends to zero precisely when the function tends to a (possibly shifted) Dirac delta; i.e., when it is well localized.

An important consequence of the uncertainty principle is related to sampling: a function which is very localized in time necessarily is spread in the frequency domain; i.e., it has a wide bandwidth and thus must be sampled at a higher rate. Put differently, if one samples a very brief portion of a time function f, then one can only hope to reconstruct signals whose spectra have are widely spread. Conversely, if one samples a very narrow frequency band, one can only expect to reconstruct signals that are spread in time.

This is a very strong fundamental result. Consequently, an important literature has been developed in the past few years about finding counterparts of this result in the context of graph signal processing. In particular, it is of interest to know when one may hope to sample a graph signal at a few vertices (for instance) and hope to faithfully recover the signal value at other unobserved vertices.

5.2 Graph spread and spectral spread

Throughout this section, we consider the normalized Laplacian and not the Laplacian to define the Graph Fourier Transform of signals. The notions we use are based on those introduced in [1, 2]. We only consider connected graphs.

It is important to note that the uncertainty principle on graphs discussed in this chapter — chosen due to its tied relationship with Heisenberg's uncertainty principle — is only one of multiple proposed definitions in the literature. Different uncertainty principles have been proposed in [17, 21, 20]. Additional results on the presented uncertainty principle can be found for instance in [18, 15, 16, 5]. Here, we choose to focus on the fundamentals.

The spread of a signal in time is a measure of how different its values are after large delays. As a consequence, measuring the spread of a signal on a graph should take into account the variations of the values together with their relative distance on the graph. Following this lead, we introduce the graph spread.

Definition 20. The graph spread $\delta_{\mathscr{G}}(\mathbf{x})$ of a unit norm signal $\mathbf{x}$ on a connected graph $\mathscr{G}$ is defined by:

$$\delta_{\mathscr{G}}(\mathbf{x}) = \inf_{i \in \mathscr{V}} \sum_{j \in \mathscr{V}} d_{\mathscr{G}}(\{i,j\}) \mathbf{x}_j^2 \,. \tag{34}$$

Note that the graph spread is trivially nonnegative. Also, it can be 0 as stated in the following proposition.

Proposition 9. *The graph spread of a one-hot signal is 0.*

Proof. Consider i the vertex where the coordinate of the vector is nonzero.

Proposition 10. *The graph spread of a unit norm signal is upper bounded by the diameter of the graph:*

$$\max_{i,j\in\mathscr{V}} d_{\mathscr{G}}(\{i,j\}) \,. \tag{35}$$

Proof.

$$\inf_{i\in\mathscr{V}} \sum_{j\in\mathscr{V}} d_{\mathscr{G}}(\{i,j\})\mathbf{x}_j^2 \leq \max_{i,j\in\mathscr{V}} d_{\mathscr{G}}(\{i,j\}) \inf_{i\in\mathscr{V}} \sum_{j\in\mathscr{V}} \mathbf{x}_j^2 \,. \tag{36}$$

The spectral spread of a signal is a measure of how localized the spectrum is. Similarly to its definition in the classical case, it should take into account both the spread between frequencies and the corresponding values of the Fourier transform of the signal.

Definition 21. The *spectral spread* of a unit norm signal $\mathbf{x}$ on graph $\mathscr{G}$ is defined as:

$$\widehat{\boldsymbol{\delta}}_{\mathscr{G}}(\mathbf{x}) = \inf_{i\in\mathscr{V}} \sum_{j\in\mathscr{V}} \|\lambda_j - \lambda_i\|_2 \widehat{\mathbf{x}}_j^2 \,. \tag{37}$$

A particular case in the expression of the spectral spread is for $i = 1$ where the expression boils down to that of the sharpness. Thus in general the spectral spread is lesser than the sharpness of the signal, and as a corollary it is upper bounded by 2. It is also trivially lower-bounded by 0.

Definition 22. The uncertainty domain of a connected graph $\mathscr{G}$ is the set:

$$\mathscr{U}_{\mathscr{G}} = \left\{ \left(\boldsymbol{\delta}_{\mathscr{G}}(\mathbf{x}), \widehat{\boldsymbol{\delta}}_{\mathscr{G}}(\mathbf{x}) \right), \|\mathbf{x}\|_2 = 1 \right\} \,. \tag{38}$$

Proposition 11. *The uncertainty domain of a connected graph $\mathscr{G}$ is compact.*

Proof. By previous remarks it is bounded. It is also the image of a compact (the sphere of unit norm vectors) by continuous functions so it is closed.

Proposition 12. *Noticeable points of the uncertainty domain are the one-hot vectors, which have zero graph spread and nonzero spectral spread, and eigenvectors of the normalized Laplacian, which have zero spectral spread and nonzero graph spread.*

Proof. Consider a one-hot vector. Obviously its graph spread is zero. Its spectral spread cannot be zero as the graph is connected, and thus each row of the normalized Laplacian matrix contains at least two nonzero coordinates.

Conversely, each eigenvector contains at least two nonzero coordinates, and thus has a nonzero graph spread.

Of particular interest is the lower frontier of this compact set that characterizes the extent to which signals can simultaneously achieve low graph and spectral spreads.

Definition 23. The *uncertainty curve* of a connected graph $\mathcal{G}$ is the function

$$u : g \mapsto \inf \left\{ \widehat{\delta}_{\mathcal{G}}(\mathbf{x}), \|\mathbf{x}\|_2 = 1 \wedge \delta_{\mathcal{G}}(\mathbf{x}) = g \right\} , \tag{39}$$

defined on the interval $[0, \delta_{\mathcal{G}}(\lambda_1)]$.

5.3 Example graphs

5.3.1 Complete graphs

Proposition 13. *The graph spread of a unit norm signal on a complete graph $\mathcal{G}$ can be written:*

$$\delta_{\mathcal{G}}(\mathbf{x}) = 1 - \max_{i \in \mathcal{V}} \mathbf{x}_i^2 . \tag{40}$$

Proof.

$$\begin{aligned} \delta_{\mathcal{G}}(\mathbf{x}) &= \inf_{i \in \mathcal{V}} \sum_{j \in \mathcal{V}} d_{\mathcal{G}}(\{i,j\}) \mathbf{x}_j^2 \\ &= \inf_{i \in \mathcal{V}} \sum_{\substack{j \in \mathcal{V} \\ i \neq j}} \mathbf{x}_j^2 \end{aligned} \tag{41}$$

Proposition 14. *The normalized Laplacian of a complete graph has two eigenvalues: 0 and $\frac{N}{N-1}$.*

Proof. Denote N the number of vertices of the complete graph. First note that $\mathbf{1}$ is an eigenvector associated with eigenvalue 0. Now consider any vector $\mathbf{x}$ orthogonal to $\mathbf{1}$, that is to say its coordinates sum to 0. we obtain:

$$\begin{aligned} \text{Ł}\mathbf{x} &= \begin{bmatrix} 1 & -\frac{1}{N-1} & \cdots & -\frac{1}{N-1} \\ -\frac{1}{N-1} & 1 & \cdots & -\frac{1}{N-1} \\ \cdots & \cdots & \cdots & \cdots \\ -\frac{1}{N-1} & -\frac{1}{N-1} & \cdots & 1 \end{bmatrix} \mathbf{x} \\ &= \begin{bmatrix} -\frac{1}{N-1} & -\frac{1}{N-1} & \cdots & -\frac{1}{N-1} \\ -\frac{1}{N-1} & -\frac{1}{N-1} & \cdots & -\frac{1}{N-1} \\ \cdots & \cdots & \cdots & \cdots \\ -\frac{1}{N-1} & -\frac{1}{N-1} & \cdots & -\frac{1}{N-1} \end{bmatrix} \mathbf{x} + \frac{N}{N-1}\mathbf{x} \\ &= \frac{N}{N-1}\mathbf{x}. \end{aligned} \tag{42}$$

Proposition 15. *Consider a unit norm signal $\mathbf{x}$ and denote $\alpha_1 = \frac{\mathbf{x}^\top \mathbf{1}}{\sqrt{N}}$. On the complete graph we have:*

$$\widehat{\delta}_{\mathscr{G}}(\mathbf{x}) = \frac{N}{N-1} \min\{\alpha_1^2, 1-\alpha_1^2\} \,. \tag{43}$$

Proof. Using Proposition 14 we obtain that:

$$\begin{aligned} \widehat{\delta}_{\mathscr{G}}(\mathbf{x}) &= \inf_{i \in \mathscr{V}} \sum_{j \in \mathscr{V}} \|\lambda_j - \lambda_i\|_2 \widehat{\mathbf{x}}_j^2 \\ &= \frac{N}{N-1} \min \left\{ \alpha_1^2, \sum_{j=2}^{N} \widehat{\mathbf{x}}_j^2 \right\}, \end{aligned} \tag{44}$$

We conclude using the orthonormality of $\mathbf{F}$.

Proposition 16. *For sufficiently large number of vertices N, the uncertainty domain of complete graphs can be made arbitrarily close to* $(0,0)$.

Proof. Consider a one-hot vector. Its graph spread is 0 and its spectral spread is $\frac{1}{N-1}$.

Proposition 16 suggests there is no equivalent of the classical uncertainty principle for signals on complete graphs. This is not a surprising result as the complete graph is a degenerate topology in which all elements are identically close to all others.

5.3.2 Star graphs

Proposition 17. *Normalized Laplacians of star graphs admit only three eigenvalues: 0, 1 and 2.*

Proof. Proposition 7 gives us that the first eigenvector of a star graph is a vector containing values $\sqrt{\frac{1}{2N-2}}$ everywhere but at one coordinate where it is $\sqrt{\frac{1}{2}}$. Without loss of generality, let us suppose the coordinate where it is $\sqrt{\frac{1}{2}}$ is 1.

It is trivial to verify that vectors containing a 1 at coordinate $i \geq 2$ and a -1 at coordinate $i+1$ are linearly independent and eigenvectors associated with eigenvalue 1.

Finally, being bipartite, Proposition 6 gives us that 2 is an eigenvalue. We easily check that the corresponding eigenvector is the one that contains $-\sqrt{\frac{1}{2}}$ at coordinate 1 and $\sqrt{\frac{1}{2N-2}}$ everywhere else.

Proposition 18. *For sufficiently large number of vertices N, the uncertainty domain of star graphs can be made arbitrarily close to* $(0,0)$.

Proof. Following the notations introduced in the proof of Proposition 17, we consider a one-hot vector where the 1 is not at coordinate 1 but at some other coordinate j. Its graph spread is thus 0. Also, its Fourier transform is a vector containing $\sqrt{\frac{1}{2N-2}}$ at coordinates 1 and N.

Finally, choosing $j = i$ we obtain that the spectral spread is not greater than $2\sqrt{\frac{1}{2N-2}}$.

Similar to complete graphs, Proposition 18 is not surprising as star graphs also correspond to degenerate topologies.

5.3.3 Ring graphs

Proposition 19. *Eigenvectors of the normalized Laplacian of ring graphs associated with a nonzero eigenvalue can be chosen as uniformly sampled cosines and sines describing at least one period.*

Proof. Denote:

$$\mathbf{x}^k = \begin{pmatrix} \cos(0) \\ \cos\left(\frac{k2\pi}{N}\right) \\ \cos\left(\frac{k4\pi}{N}\right) \\ \cdots \\ \cos\left(\frac{(N-1)k2\pi}{N}\right) \end{pmatrix}. \tag{45}$$

Then:

$$\begin{aligned} \left(Ł\mathbf{x}^k\right)_i &= 2\cos\left(\frac{i2\pi k}{N}\right) - \cos\left(\frac{(i+1)2\pi k}{N}\right) - \cos\left(\frac{(i-1)2\pi k}{N}\right) \\ &= 2\left(1 - \frac{\cos(2\pi k)}{N}\right)\mathbf{x}^k_i. \end{aligned} \tag{46}$$

A very similar proof can be derived using sines instead of cosines.

Proposition 20. *There exist $N_0 \in \mathbb{Z}$ and $M > 0$ such that for any number of vertices $N \geq N_0$, the uncertainty domain of a ring graph is at least at a distance M from the set $\{(0,0)\}$.*

Proof. Fix $\alpha > 0$ and consider a unit-norm signal $\mathbf{x}$ with a graph spread of at most α:

$$\begin{aligned} \alpha &\geq \inf_{i\in\mathscr{V}} \sum_{j\in\mathscr{V}} d_{\mathscr{G}}(\{i,j\})\mathbf{x}_v^2 \\ &\geq \inf_{i\in\mathscr{V}} \sum_{\substack{j\in\mathscr{V} \\ i\neq j}} \mathbf{x}_j^2 \\ &= \inf_{i\in\mathscr{V}} \{1 - \mathbf{x}_i^2\} \end{aligned} \tag{47}$$

Denote i^* a value of i that reaches the minimum of the last quantity:

$$\mathbf{x}_{i^*}^2 \geq 1 - \alpha\,. \tag{48}$$

Now using Proposition 19 we obtain that the largest magnitude value in an eigenvector of a ring graph is of the order of $\frac{1}{\sqrt{N}}$. Thus, each value in $\widehat{\mathbf{x}}$ is at most of the order of $\frac{\sqrt{1-\alpha}+\sqrt{\alpha}}{\sqrt{N}}$. Because $\|\widehat{\mathbf{x}}\|_2 = 1$, we obtain that of the order of N of the coordinates of $\widehat{\mathbf{x}}$ are close to this maximum.

We conclude by observing the span of eigenvalues yielded in Proposition 19 and the definiton of $\widehat{\delta}_{\mathscr{G}}(\mathbf{x})$.

So, contrary to the previous examples of the complete graph and the star graph, in the case of a ring graph does exhibit a non-trivial uncertainty principle.

6 Conclusion

We introduced graphs and signals on graphs. We showed there are strong relations that tie signals to the graphs they are defined on. We considered several examples of graphs, either deterministic or randomized.

As when it comes to an uncertainty principle, we showed examples where such a principle holds and examples where it does not.

Better understanding the connections between graphs and this principle could lead to important developments in the field. In particular being able to characterize for which graphs the principle does not hold could allow very the rise of very efficient sampling strategies, able to capture very precise events both in the vertex and in the spectrum domains.

References

1. Agaskar, A., Lu Y.: Uncertainty principles for signals defined on graphs: Bounds and characterizations. IEEE International Conference on Acoustics, Speech and Signal Processing (2012).
2. Agaskar, A., Lu Y.: A spectral graph uncertainty principle. IEEE Transactions on Information Theory (2013).
3. Bauer, F., Jost, J.: Bipartite and neighborhood graphs and the spectrum of the normalized graph Laplacian. ArXiv preprint 0910.3118 (2009).
4. Bayram, E., Frossard, P., Vural, E., Alatan, A.: Analysis of Airborne LiDAR Point Clouds with Spectral Graph Filtering. IEEE Geoscience and Remote Sensing Letters (2018).
5. Di Lorenzo, P., Barbarossa, S., Banelli, P., Sardellitti, S.: Adaptive least mean squares estimation of graph signals. IEEE Transactions on Signal and Information Processing over Networks (2016).
6. Gray, Robert M.: Toeplitz and circulant matrices: A review. Foundations and Trends in Communications and Information Theory (2006).
7. Gripon, V.: Tropical graph signal processing. IEEE 51st Asilomar Conference on Signals, Systems, and Computers (2017).
8. Gross, J. L., Yellen, J.: Graph theory and its applications. Chapman and Hall/CRC (2005).
9. Hammond, D. K., Vandergheynst, P., Gribonval, R.: Wavelets on graphs via spectral graph theory. Applied and Computational Harmonic Analysis (2011).

10. Heisenberg, W.: Über den anschaulichen Inhalt der quantentheoretischen Kinematik und Mechanik. Zeitschrift für Physik 33 (1925).
11. Hosseinalipour, S., Wang, J., Dai, H., Wang, W.: Detection of Infections Using Graph Signal Processing in Heterogeneous Networks. IEEE Global Communications Conference (2017).
12. Huang, W., Bolton, T. A.W, Medaglia, J. D., Bassett, D. S., Ribeiro, A. and Van De Ville, D.: A Graph Signal Processing Perspective on Functional Brain Imaging. Proceedings of the IEEE (2018).
13. Lancaster, P., Tismenetsky, M.: The theory of matrices: with applications. Elsevier (1985).
14. Ortega, A., Frossard, P., Kovačević, J., Moura, J. M.F., Vandergheynst, P.: Graph signal processing: Overview, challenges, and applications. Proceedings of the IEEE (2018).
15. Pasdeloup, B., Alami, R., Gripon, V., Rabbat, M. G.: Toward an uncertainty principle for weighted graphs. IEEE 23rd European Signal Processing Conference (2015).
16. Pasdeloup, B., Gripon, V., Mercier, G., Pastor, D.: Towards a characterization of the uncertainty curve for graphs. IEEE International Conference on Acoustics, Speech and Signal Processing (2016).
17. Perraudin, N., Ricaud, B., Shuman, D., Vandergheynst, P.: Global and local uncertainty principles for signals on graphs. APSIPA Transactions on Signal and Information Processing (2018).
18. Rabbat, M. G. and Gripon, V.: Towards a spectral characterization of signals supported on small-world networks. IEEE International Conference on Acoustics, Speech and Signal Processing (2014).
19. 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 Processing Magazine (2013).
20. Teke, O., Vaidyanathan, P.P.: Uncertainty principles and sparse eigenvectors of graphs. IEEE Transactions on Signal Processing (2017).
21. Tsitsvero, M., Barbarossa, S., Di Lorenzo, P.: Signals on graphs: Uncertainty principle and sampling. IEEE Transactions on Signal Processing (2016).