

Toward An Uncertainty Principle For Weighted Graphs

Bastien Pasdeloup, Réda Alami, Vincent Gripon, Michael Rabbat

To cite this version:

Bastien Pasdeloup, Réda Alami, Vincent Gripon, Michael Rabbat. Toward An Uncertainty Principle For Weighted Graphs. EUSIPCO 2015 : 23rd European Signal Processing Conference, Aug 2015, Nice, France. pp.1496 - 1500, 10.1109/EUSIPCO.2015.7362633. hal-01270655

HAL Id: hal-01270655 <https://hal.science/hal-01270655v1>

Submitted on 8 Feb 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

TOWARD AN UNCERTAINTY PRINCIPLE FOR WEIGHTED GRAPHS

Bastien Pasdeloup, Reda Alami, Vincent Gripon ´ ∗

Telecom Bretagne UMR CNRS Lab-STICC name.surname@telecom-bretagne.eu *Michael Rabbat*

McGill University ECE dept. name.surname@mcgill.ca

ABSTRACT

The uncertainty principle states that a signal cannot be localized both in time and frequency. With the aim of extending this result to signals on graphs, Agaskar & Lu introduce notions of graph and spectral spreads. They show that a graph uncertainty principle holds for some families of unweighted graphs. This principle states that a signal cannot be simultaneously localized both in graph and spectral domains. In this paper, we aim to extend their work to weighted graphs. We show that a naive extension of their definitions leads to inconsistent results such as discontinuity of the graph spread when regarded as a function of the graph structure. To circumvent this problem, we propose another definition of graph spread that relies on an inverse similarity matrix. We also discuss the choice of the distance function that appears in this definition. Finally, we compute and plot uncertainty curves for families of weighted graphs.

Index Terms— Signal processing on graphs, uncertainty principle, weighted graphs.

1. INTRODUCTION

In classical signal processing holds an uncertainty principle stating that a signal cannot be localized both in time and frequency domains [1]. This tradeoff is defined by the following equation

$$
\Delta_t^2 \Delta_\omega^2 \ge \frac{1}{4} \tag{1}
$$

in which Δ_t^2 is the *time spread* of the signal and Δ_ω^2 its *frequency spread*.

Graph signal processing [2] is a generalization of classical Fourier analysis in which the support for the signal is not necessarily a uniform sampling in time but may be a more complex structure, represented as a graph. This emerging domain has received a lot of interest recently [3–5] and has been applied to fields such as image denoising [2] and social networks [6].

In the context of signal processing on graphs, [7] introduces a spectral graph uncertainty principle analog to (1), stating that a signal on a graph cannot be localized both in the graph domain and in the spectral domain. For a given signal, the authors propose notions of *graph spread* around a node u_0 , that we denote by $\Delta_{\mathcal{G},u_0}^2$, and *spectral spread* that we denote by Δ_s^2 . They show that for a fixed node u_0 and any signal **x** on a graph, $(\Delta^2_{\mathcal{G},u_0}(\mathbf{x}), \Delta^2_{s}(\mathbf{x}))$ is higher than a certain curve called *uncertainty curve*. The authors then plot this curve for some particular unweighted graphs for which an equation can be determined, and propose an efficient algorithm to estimate it for any unweighted graph.

In this paper, we aim to extend the results of [7] to weighted graphs. We first review the uncertainty principle for unweighted graphs in Section 2. Then, we show in Section 3 that a naive use of the method introduced in [7] leads to inconsistent results when applied to weighted graphs, and propose a new definition for the graph spread. Additionally, we discuss in Section 4 the choice of the distance function that appears in our definition of graph spread. Finally, in Section 5, we use our definition to plot uncertainty curves for some weighted graphs, using various distance functions, and conclude in Section 6.

2. UNCERTAINTY PRINCIPLE FOR UNWEIGHTED GRAPHS

2.1. Context and definitions

In this document, we consider a connected, simple graph $\mathcal{G} =$ $(V, \mathcal{E}, \mathbf{W})$ composed of a set of $|V| = N$ nodes, a set of edges \mathcal{E} , and a matrix W. Without loss of generality, we label the nodes using integers (*i.e* $V = \{1...N\}$). In the definition of G, W is a symmetric matrix of real values such that $W_{u,v}$ denotes the weight associated with edge $(u, v) \in \mathcal{E}$. In the particular case of unweighted graphs, W is the binary adjacency matrix of \mathcal{G} .

A signal **x** on a graph $\mathcal G$ is a set of real values associated with the nodes of V. Mathematically, $\mathbf{x} = {\mathbf{x}(1) \dots \mathbf{x}(N)}$ is a vector of \mathbb{R}^N . Figure 1 depicts an example of graph carrying a signal.

A signal **x** is said to be smooth on a graph G if nearby

[∗]This work was supported by the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013) / ERC grant agreement n^o 290901.

Fig. 1: Example of graph carrying a signal x. The value of x associated with each node is described by a color according to the given scale.

nodes carry similar values of signal. Such a measure of smoothness is given by the discrete *p*-Dirichlet form [2] of the signal:

$$
S_{\mathcal{G},p}(\mathbf{x}) \triangleq \frac{1}{p} \sum_{u \in \mathcal{V}} \left(\sum_{v \in \mathcal{V} \text{ s.t. } (u,v) \in \mathcal{E}} \mathbf{W}_{u,v}(\mathbf{x}(v) - \mathbf{x}(u))^p \right)^{\frac{1}{p}}.
$$
 (2)

A smooth signal is associated with a low $S_{\mathcal{G},p}(\mathbf{x})$ value. In particular, $S_{\mathcal{G},p}(\mathbf{x}) = 0$ if and only if **x** is constant.

When interpreting the significance of G with respect to a signal x, (2) gives us that *W is analogous to a similarity between nodes*, with the noticeable exception of $W_{u,u} = 0$. More generally, a zero value in W indicates the absence of an edge in $\mathcal G$. As a consequence $\mathcal E$ is redundant with W and can be dropped from the definition of \mathcal{G} .

The normalized Laplacian L_W of W [8] is a difference operator analogous to the Laplacian operator arising for example in the study of heat diffusion, wave propagation, and harmonic analysis. It is defined by

$$
\mathbf{L}_{\mathbf{W}} \triangleq \mathbf{I} - \mathbf{D}^{-\frac{1}{2}} \mathbf{W} \mathbf{D}^{-\frac{1}{2}}
$$
 (3)

where **D** is the diagonal matrix of nodes degrees. Since **D** and W are both symmetric real matrices, L_W can be diagonalized and described by its orthonormal eigenvectors \mathcal{X}_{L_W} = $(\boldsymbol{f}_1 \dots \boldsymbol{f}_N)$ and associated eigenvalues $\Lambda_{\mathbf{L_W}} = (\lambda_1 \leq \cdots \leq \lambda_N)$ λ_N).

2.2. Notions of spreads for unweighted graphs

The notions of graph and spectral spreads introduced in this paper are an extension of [7]. In the following paragraphs we recall their definitions.

The graph spread $\Delta_{\mathcal{G},u_0}^2(\mathbf{x})$ of a signal **x** around a given node u_0 is defined by

$$
\Delta_{\mathcal{G},u_0}^2(\mathbf{x}) \triangleq \frac{1}{\|\mathbf{x}\|_2^2} \sum_{u \in \mathcal{V}} d_{geo(\mathbf{W})}^2(u_0, u)\mathbf{x}(u)^2 \tag{4}
$$

where **x**(*u*) is the value of **x** at node *u*, and $d_{geo(W)}^2(u_0, u)$ is the squared geodesic distance $-i.e$ the squared length of the shortest path – between u_0 and u using weights matrix W. Informally, this definition of $\Delta_{\mathcal{G},u_0}^2(\mathbf{x})$ quantifies the distance from u_0 to signal **x**. It allows us to introduce a notion of

locality of the signal in \mathcal{G} : the smaller the graph spread is, the more **x** is concentrated around u_0 .

The spectral spread $\Delta_s^2(\mathbf{x})$ of **x** is defined by

$$
\Delta_s^2(\mathbf{x}) \triangleq \frac{1}{\|\mathbf{x}\|_2^2} \sum_{n=1}^N \lambda_n \widehat{x}_n^2 \tag{5}
$$

where $\hat{\mathbf{x}} = (\hat{x}_1 \dots \hat{x}_N) \triangleq (\mathbf{f}_1^T \mathbf{x} \dots \mathbf{f}_N^T \mathbf{x})$ is the graph Fourier
transform [2] of x. Note that Δ^2 is minimized for a signal have transform [2] of **x**. Note that Δ_s^2 is minimized for a signal having all its energy on frequency $\lambda = 0$. This situation occurs when a signal has been inifinitely diffused on a non-bipartite graph using the diffusion matrix $T = D^{-\frac{1}{2}}WD^{-\frac{1}{2}}$. As a matter of fact, due to the property that the eigenvalues of T are all in $[-1, 1]$, with exactly one equal to 1, and due to the definition of L_W , an inifinitely diffused signal has all its energy on the smallest eigenvalue of L_W .

One can show that for any signal x on an unweighted graph G , there exists a relation between (4) and (5) such that any pair $(\Delta^2_{\mathcal{G},u_0}(\mathbf{x}), \Delta^2_{s}(\mathbf{x}))$ is constrained from below by a certain curve γ_{u_0} . Figure 2 depicts the uncertainty curve for some chosen graphs of 100 nodes. Additional examples of uncertainty curves are proposed in [7].

Fig. 2: Examples of uncertainty curves for some unweighted graphs of 100 nodes. For the star graph, the middle node (*i.e* the node connected to all others) is chosen as u_0 .

It is shown in [7] that any uncertainty curve intersects the horizontal axis at exactly one location $(1, 0)$ obtained for a signal **x** localized at node u_0 . Moreover, the curve reaches a spectral spread of 0 for $\mathbf{x} = f_{1}$ [7]. In this case, the associated graph spread is equal to $f_1^{\top} P^2 f_1$, where $\mathbf{P} = \text{diag}(d_{geo(\mathbf{W})}(u_0, u)).$

In the remainder of this document, we consider unit-norm signals to simplify the reasoning. Therefore, (4) becomes

$$
\Delta_{\mathcal{G},u_0}^2(\mathbf{x}) \triangleq \sum_{u \in \mathcal{V}} d_{geo(\mathbf{W})}^2(u_0, u)\mathbf{x}(u)^2 \tag{6}
$$

and (5) becomes

$$
\Delta_s^2(\mathbf{x}) \triangleq \sum_{n=1}^N \lambda_n \hat{x}_n^2 .
$$
 (7)

3. TOWARDS AN UNCERTAINTY PRINCIPLE FOR WEIGHTED GRAPHS

In this section we aim to extend the definitions of [7] to weighted graphs. First, we show that a naive use of (6) leads to inconsistent results such as discontinuity of the graph spread when regarded as a function of \mathcal{G} .

3.1. Discontinuity of the graph spread for weighted graphs

Let us consider the graph G in Figure 3 in which u_0 is fixed and x is equally distributed among the nodes.

Fig. 3: Example weighted graph (a) for which we want to compute the uncertainty curve, and associated matrix of weights (b). We consider a signal x equally shared among the nodes (*i.e* $\forall w \in \mathcal{V} : \mathbf{x}(w) = \frac{1}{\sqrt{2}}$ $\overline{3}$).

Using (6), we obtain that $\Delta_{\mathcal{G},u_0}^2(\mathbf{x}) = d_{geo(\mathbf{W})}^2(u_0, u)\mathbf{x}(u)^2 +$ $d_{geo(W)}^2(u_0, v) \mathbf{x}(v)^2 = \frac{\varepsilon^2 + 1}{3} \xrightarrow[\varepsilon \to 0]{} \frac{1}{3}$. It seems reasonable to expect that as ε tends to $0, \Delta^2_{\mathcal{G},u_0}$ tends to the limit case where $\varepsilon = 0$. In particular, $\Delta_{\mathcal{G},u_0}^2$ should be robust to measurement noise in scenarios where \bf{W} is not perfectly known.

Figure 4 depicts the matrix of weights associated with the limit graph \mathcal{G}' .

Fig. 4: Matrix of weights (b) representing the limit of Figure 3b when $\varepsilon \longrightarrow 0$, and associated graph \mathcal{G}' (a). The edge (u_0, u) has been removed since $W_{u_0, u} = 0$.

Again, we use (6) to compute the graph spread for \mathcal{G}' around u_0 . With this graph, we obtain that $\Delta_{\mathcal{G}',u_0}^2(\mathbf{x}) = \frac{10}{3}$, leading to a discontinuity of $\mathcal{G} \mapsto \Delta^2_{\mathcal{G},u_0}$.

Remark: Looking closely at the above mentioned example, we point out that there is a misuse of W in the definition of $\Delta_{\mathcal{G},u_0}^2$. As a matter of fact (2) gives us that **W** is a similarity matrix, whereas (6) uses it as a distance matrix. More generally we expect the graph spread to grow with the distance between nodes in a graph, that is to say as the similarity decreases. In the next subsection we propose a generic frame-

work for a rectified definition of the graph spread in the case of weighted graphs.

3.2. Expected behavior of a graph spread

In order to define a new notion of graph spread that does not lead to unexpected behaviors as in Section 3.1, we present some desired properties on $\Delta^2_{\mathcal{G},u_0}$.

We expect from a graph spread notion that it captures the locality of a signal **x** in the graph domain. In other words, for a fixed node u_0 , the graph spread around u_0 should measure the extent to which the signal x is concentrated around u_0 . To achieve this, we would like to ensure the following properties:

- $\Delta_{\mathcal{G},u_0}^2(\mathbf{x})$ should be small if **x** is localized around u_0 , and should increase as the distance between u_0 and the nodes carrying x increases.
- Additionally, the only situation leading to $\Delta_{\mathcal{G},u_0}^2(\mathbf{x})=0$ should be when the signal is entirely localized on u_0 .
- A third desired property is that the graph spread should be *similar* for graphs with *similar* weights (continuity of $\mathcal{G} \mapsto \Delta^2_{\mathcal{G},u_0}$).

Moreover it appears to us that the choice of the geodesic distance in (6) is arbitrary. In order to be compliant with the previously enumerated properties, we characterize the class of acceptable functions d:

- a) $\forall u, v \in \mathcal{V} : d(u, v) \geq 0.$
- b) $\forall u, v \in \mathcal{V} : d(u, v) = 0 \Leftrightarrow u = v.$
- c) d is continuous, and if we increase $W_{u,v}$ for a single edge (u, v) , then $\forall u', v' \in \mathcal{V} : d(u', v')$ does not increase.

Remark: The geodesic distance $d_{geo(W)}^2$ based on **W** is not compliant with c) (not continuous and increasing with W).

4. EXAMPLES OF COMPLIANT DISTANCES FOR GRAPH SPREAD

In this section we present two choices of distances compliant with the previously introduced properties. Other distances can be used, such as the *resistance distance* [9] or the *commute time distance* [10]. We choose to focus on two simpler examples for the sake of clarity.

4.1. Inverse similarity matrix

The distance introduced in this subsection is a simple rectified version of (6) and is compatible with it in the case of unweighted graphs. We introduce a new matrix \overline{S} :

$$
\forall u, v \in \mathcal{V} : \bar{\mathbf{S}}_{u,v} \triangleq \begin{cases} \infty & \text{if } \mathbf{W}_{u,v} = 0 \\ 0 & \text{if } u = v \\ \frac{1}{\mathbf{W}_{u,v}} & \text{otherwise} \end{cases} \tag{8}
$$

We propose to use it instead of W in (6) .

Remark: The choice of taking the inverse is arbitrary and could be replaced by other functions. Basically, any decreasing distance function associating an infinite distance to unconnected pairs of nodes can be used. Standard alternatives are Gaussian kernels, as shown later in Section 5.2. In some cases weighted similarity graphs are constructed from distance graphs and in such cases it appears more natural to use the latter directly instead of estimating it back from W. Some examples of such graphs are given in the next section.

We now show that the squared geodesic distance using $\bar{\mathbf{S}}, d_{geo(\bar{\mathbf{S}})}^2$, is compliant with the three properties enounced in Section 3.2:

- a) is trivially true, since $d_{geo(\bar{S})}^2(u, v)$ features a square.
- b) is ensured for any couple of nodes by construction of \overline{S} .
- c) is in most cases trivial. The only concern is when an edge is removed from G . Such a scenario occurs in the case where the similarity between two nodes u and v becomes zero. By definition of \bar{S} , this corresponds to a distance between u and v that diverges to infinity. It is obvious that eventually the shortest paths of \bar{S} do not include this edge.

With this function, the definition of graph spread in (6) now becomes

$$
\Delta_{\mathcal{G},u_0}^2(\mathbf{x}) \triangleq \sum_{u \in \mathcal{V}} d_{geo(\bar{\mathbf{S}})}^2(u_0, u)\mathbf{x}(u)^2.
$$
 (9)

4.2. Diffusion distance

Another distance function we study in this paper is the diffusion distance, as defined in [11]. Given a graph adjacency matrix W and its associated (non-normalized) Laplacian matrix L_W [8], d_{diff} is defined in matrix form for some constant parameter α as follows:

$$
\forall u, v \in \mathcal{V} : d_{diff}(u, v) \triangleq ||(\mathbf{I} + \alpha \mathbf{L}_{\mathbf{W}})^{-1}(\mathbf{x}_{u} - \mathbf{x}_{v})|| \quad (10)
$$

where x_u is a unit-norm signal having only one non-zero value on node u.

One can show that d_{diff} verifies the three desired properties presented in Section 3.2. In the remaining of the document, we set $\alpha = 1$ and use the l_2 norm.

5. RESULTS FOR CLASSICAL WEIGHTED GRAPHS

In this section we introduce several classical weighted graphs and plot their uncertainty curves considering both inverse similarity matrix and diffusion distance. Curves are plotted using the Sandwich algorithm introduced in [7]. By comparing the resulting curves to known uncertainty curves [7] obtained for graphs such as the ring or star graphs, one can evaluate the *amount of uncertainty* associated to the graph under study.

5.1. Random graph

We call random graph a graph which adjacency symmetric matrix W is such that each non-null coordinate $W_{u,v}$ is drawn uniformly between 0 and 1. Using the previously introduced distance functions, we plot in Figure 5 the uncertainty curves for such families of graphs. The curves are normalized such that the graph spread associated with $\Delta_s^2(\mathbf{x}) = 0$ is at most equal to 1 for each distance function used.

Fig. 5: Examples of uncertainty curves for some randomly weighted families of graphs of 10 nodes. The curves are computed for the two distance functions $d_{geo(\bar{S})}$ and d_{diff} . Mean uncertainty curves for 100 random weights.

It is interesting to notice that the choice of the distance does not impact the relative order of the curves. Additionally, the intersection between the uncertainty curves associated to the star and complete graphs is kept when switching the distance function. The main difference is the smoothness of the curves. Using d_{diff} tends to produce uncertainty curves that are more regular than when using $d_{geo(\bar{S})}$.

5.2. Gaussian kernel

We consider graphs obtained using a Gaussian kernel. The idea is to build a distance graph and to apply a Gaussian kernel to all weights to obtain W. The Gaussian kernel has two parameters α and β and is defined as follows:

$$
g: x \mapsto \alpha \exp(-\beta x^2) . \tag{11}
$$

We consider a set S of N sensors uniformly distributed in a 1×1 square. We define a symmetric matrix **E** as follows. Fix some radius r such that if two sensors u and v are at Euclidean distance $d_{euc}(u, v)$ less than r, then $\mathbf{E}_{u,v} = \frac{d_{euc}(u, v)}{\max_{v} d_{v}(u, v)}$ $\max_{u',v' \in \mathcal{V}} d_{euc}(u',v')$ and $\mathbf{E}_{u,v} = 0$ otherwise. W is then defined by applying g to each cell of E.

Figure 6 depicts the mean uncertainty curves for random geometric graphs. When computing the uncertainty curve using the squared geodesic distance $d_{geo(E)}^2$, we directly use the matrix of Euclidean distances E, and do not retrieve it from W (see remark in Section 4.1). The curves are normalized so that no value of $\Delta_{\mathcal{G},u_0}^2$ exceeds 1 for each distance function.

Fig. 6: Uncertainty curves for random geometric graphs of 10 nodes. Parameters (α, β, r) are respectively fixed to $(1, 1, 0.3)$. Mean uncertainty curves for 100 random graphs.

Additionally, we apply the same Gaussian kernel to semilocalized graphs. We use the same graph as presented in [2] (Example 2). Such a graph is obtained by connecting pixels of the 32×32 cameraman image to their eight neighbors, weighting connections using g over the difference of intensity of pixels to obtain W. This method for constructing graphs for images has been previously used for example in [12]. Figure 7a depicts the picture from which the graph is extracted. Figure 7b shows the associated uncertainty curves using the distances $d_{geo(\bar{S})}^1$ and d_{diff} .

Fig. 7: Computation of the normalized uncertainty curves (b) associated to the image (a), for the introduced distance functions. Parameters (α, β) are respectively fixed to $(1, 1)$.

6. CONCLUSION

In this work, we have extended the notion of uncertainty on graphs introduced in [7] to weighted graphs, and pointed out important properties of the distance function used in the definition of graph spread. We have shown the applicability of our work on classical families of graphs, as well as on semilocalized graphs that are encountered in real-life use-cases.

A direction of our future work will focus on side aspects, such as determining a way to efficiently choose the node used as u_0 in the computation of $\Delta_{\mathcal{G},u_0}^2$ to perform better comparisons of uncertainty curves. We will also investigate the properties linked to the uncertainty curve of a given graph.

REFERENCES

- [1] Gerald B. Folland and Alladi Sitaram, "The uncertainty principle: A mathematical survey," *Journal of Fourier Analysis and Applications*, vol. 3, no. 3, 1997.
- [2] David I. Shuman, Sunil K. Narang, Pascal Frossard, Antonio Ortega, and Pierre Vandergheynst, "The emerging field of signal processing on graphs: Extending high-dimensional data analysis to networks and other irregular data domains," *CoRR*, vol. abs/1211.0053, 2013.
- [3] David K. Hammond, Pierre Vandergheynst, and Rémi Gribonval, "Wavelets on graphs via spectral graph theory," *Applied and Computational Harmonic Analysis*, vol. 30, no. 2, 2011.
- [4] Sunil K. Narang and Antonio Ortega, "Perfect reconstruction two-channel wavelet filter-banks for graph structured data," *CoRR*, vol. abs/1106.3693, 2011.
- [5] David I Shuman, Benjamin Ricaud, and Pierre Vandergheynst, "Vertex-frequency analysis on graphs," 2013.
- [6] Michael Rabbat and Vincent Gripon, "Towards a spectral characterization of signals supported on smallworld networks," in *ICASSP 2014 : IEEE International Conferences on Acoustics, Speech and Signal Processing*, IEEE, Ed., 2014.
- [7] Ameya Agaskar and Yue M. Lu, "A spectral graph uncertainty principle," *CoRR*, vol. abs/1206.6356, 2012.
- [8] Fan R. K. Chung, *Spectral Graph Theory*, American Mathematical Society, 1997.
- [9] D.J. Klein and M. Randi, "Resistance distance," *Journal of Mathematical Chemistry*, vol. 12, no. 1, 1993.
- [10] James A Albano and David W Messinger, "Euclidean commute time distance embedding and its application to spectral anomaly detection," in *SPIE Defense, Security, and Sensing*. International Society for Optics and Photonics, 2012, pp. 83902G–83902G.
- [11] Santiago Segarra, Weiyu Huang, and Alejandro Ribeiro, "Diffusion and superposition distances for signals supported on networks," *CoRR*, vol. abs/1411.7443, 2014.
- [12] Sunil K Narang, Yung Hsuan Chao, and Antonio Ortega, "Graph-wavelet filterbanks for edge-aware image processing," in *Statistical Signal Processing Workshop (SSP), 2012 IEEE*. IEEE, 2012.

¹Contrary to the study of random geometric graphs, we do not directly use a matrix of distances D associated to the difference of pixels intensity, but retreive \bar{S} from W using (8). As a matter of fact, two adjacent pixels with identical intensity result in a distance of 0 if using D, and would cause the discontinuity problem previously introduced. A solution to cope with this problem is to add an ε noise to all edge weights. However, this leads to hard to visualize curves. Therefore, for the sake of comprehension, we use $d_{geo(\bar{S})}^2$ and not $d_{geo(D)}^2$.