# Vers une caractérisation de la courbe d'incertitude pour des graphes portant des signaux

Bastien PASDELOUP, Vincent GRIPON, Grégoire MERCIER, Dominique PASTOR

Telecom Bretagne, UMR CNRS Lab-STICC 655 Avenue du Technopole, 29200 Plouzané nom.prenom@telecom-bretagne.eu

**Résumé** – Le traitement de signal sur graphes est un domaine récent visant à généraliser les outils classiques du traitement de signal, afin d'analyser des signaux évoluant sur des domaines complexes. Ces domaines sont représentés par des graphes pour lesquels on peut calculer une matrice appelée Laplacien normalisé [3]. Il a été montré que les valeurs propres de ce Laplacien correspondent aux fréquences du domaine de Fourier en traitement de signal classique [2]. Ainsi, le domaine fréquentiel n'est pas identique pour tout graphe support des signaux. Une conséquence est qu'il n'y a pas de généralisation non triviale du principe d'incertitude d'Heisenberg, indiquant qu'un signal ne peut être à la fois localisé dans le domaine temporel et dans le domaine fréquentiel. Une manière de généraliser ce principe, introduite par Agaskar & Lu en [1], consiste à déterminer une courbe servant de borne inférieure au compromis entre précision dans le domaine du graphe et précision dans le domaine spectral. L'objectif de ce papier est de proposer une caractérisation des signaux atteignant cette courbe, pour une classe de graphes plus générique que celle étudiée par Agaskar & Lu.

**Abstract** – Signal processing on graphs is a recent research domain that aims at generalizing classical tools in signal processing, in order to analyze signals evolving on complex domains. Such domains are represented by graphs, for which one can compute a particular matrix, called the normalized Laplacian [3]. It was shown that the eigenvalues of this Laplacian correspond to the frequencies of the Fourier domain in classical signal processing [2]. Therefore, the frequential domain is not the same for every support graph. A consequence of this is that there is no non-trivial generalization of Heisenberg's uncertainty principle, that states that a signal cannot be localized both in the time domain and in the frequency domain. A way to generalize this principle, introduced by Agaskar & Lu in [1], consists in determining a curve that represents a lower bound on the compromise between precision in the graph domain and precision in the spectral domain. The aim of this paper is to propose a characterization of the signals achieving this curve, for a larger class of graphs than the one studied by Agaskar & Lu.

### **1** Introduction

En traitement de signal classique, un signal peut être vu comme une série temporelle, associant à chaque instant une intensité. Le support de l'information est alors unidimentionnel, et est représenté par l'axe du temps. Un des principaux objectifs du traitement de signal sur graphes est d'étendre les outils classiques pour l'étude des signaux à de nouveaux signaux, associés à des topologies plus complexes. Le portage d'outils tels que le produit de convolution, la translation d'un signal ou la transformée de Fourier [2] a été rendu possible grâce à la correspondance entre fréquences et valeurs propres d'une certaine matrice associée au graphe. Un signal est donc vu comme un vecteur associant une intensité à chaque nœud du graphe, et possédant une décomposition spectrale selon les vecteurs propres d'une matrice particulière dépendant du graphe.

Dans cette optique de transposition de résultats du traitement de signal classique au traitement de signal sur graphes, Agaskar & Lu [1] ont proposé une adaptation du principe d'incertitude d'Heisenberg [4] qui énonce qu'un signal ne peut pas être localisé à la fois dans les domaines temporel et fréquentiel. Les auteurs ont montré qu'il existe un tel compromis en traitement de signal sur graphes. Ainsi, un signal sur un graphe ne peut pas être localisé à la fois dans les domaines graphique et spectral.

Plus en détails, Agaskar & Lu ont montré que le compromis entre précision spectrale et précision graphique était dépendant du graphe porteur du signal. Ils ont donc introduit des courbes d'incertitude représentant, pour un graphe donné et un nœud choisi, les couples  $\langle \text{précision graphique, précision} spectrale \rangle$  étant des optima de Pareto. De plus, ils ont montré que toute courbe d'incertitude est convexe, et ont illustré pour un exemple particulier qu'elle pouvait être décrite par une portion d'ellipse.

L'objectif de ce papier est d'étendre leurs travaux, en proposant une méthode pour fournir une expression analytique de la courbe d'incertitude de tout graphe. L'obtention d'une telle équation permet de limiter l'espace de recherche des signaux atteignant la courbe d'incertitude, *i.e* atteignant un optimum de Pareto en termes de précision spectrale et graphique.

Ce document est organisé comme suit : en Section 2, nous présentons plus en détails le traitement de signal sur graphes, et introduisons les notations nécessaires à la compréhension de nos travaux. En Section 3, nous détaillons la notion d'incertitude sur graphe telle qu'introduite en [1]. Enfin, en Section 4, nous étendons les résultats d'Agaskar & Lu et proposons une méthode pour caractériser les signaux atteignant la courbe d'incertitude et en fournir une expression analytique. La Section 5 conclut ce document, et propose des extensions à nos travaux.

#### 2 Le traitement de signal sur graphes

En traitement de signal sur graphes, le support des signaux n'est pas uniquement le temps, mais peut être plus complexe. Pour représenter un tel support, nous introduisons la notion de graphes :

**Définition 1 (graphe)** – Un graphe (simple, non-orienté)  $\mathcal{G}$  est un triplet  $\langle \mathcal{V}, \mathcal{E}, \mathbf{W} \rangle$ , où  $\mathcal{V} = \{1 \dots N\}$  est un ensemble de nœuds,  $\mathcal{E} = \mathcal{V} \times \mathcal{V}$  est un ensemble d'arêtes, et  $\mathbf{W}$  est une matrice représentant les poids associés aux arêtes :  $\forall u, v \in$  $\mathcal{V} : (u, v) \in \mathcal{E} \Leftrightarrow \mathbf{W}_{u,v} = \mathbf{W}_{v,u} \neq 0.$ 

Un signal porté par un graphe est un vecteur associant une intensité à chaque nœud du graphe :

**Définition 2 (signal sur graphe) –** Un signal  $\boldsymbol{x} = [x_1 \dots x_N]$ sur un graphe  $\mathcal{G}$  à N nœuds est un vecteur de  $\mathbb{R}^N$ . Sans perdre en généralité, nous étudions dans ce document des signaux normalisés, i.e tels que  $\|\boldsymbol{x}\|_2 = 1$ .

La Figure 1 illustre un signal sur un graphe ayant une topologie en étoile. Les intensités par nœud sont représentées par des couleurs :



FIGURE 1 – Exemple de signal porté par un graphe étoilé. L'intensité du signal sur chaque nœud est donnée par l'échelle de couleurs.

Une autre matrice donnant des informations sur le graphe est son Laplacien normalisé [3]. C'est un opérateur différentiel analogue au Laplacien intervenant par exemple dans les équations de diffusion de la chaleur ou de l'analyse harmonique :

**Définition 3 (Laplacien normalisé)** – Le Laplacien normalisé  $L_W$  associé à un graphe  $\mathcal{G}$  de matrice de poids W est défini par  $L_W \triangleq I - D^{-\frac{1}{2}} W D^{-\frac{1}{2}}$ , où D est une matrice diagonale donnant les degrés des nœuds (i.e  $D_{i,i} = \sum_k W_{i,k}$ ), et I est une matrice identité de taille N.

Une propriété de la matrice  $\mathbf{L}_{\mathbf{W}}$  est qu'elle est définie par combinaison linéaire de matrices réelles et symétriques. En conséquence, elle est elle-même réelle et symétrique, et peut donc être diagonalisée dans une base orthonormale. On notera les vecteurs propres de cette base  $\mathcal{X}_{\mathbf{L}_{\mathbf{W}}} = [\boldsymbol{\chi}_1 \dots \boldsymbol{\chi}_N]$  et les valeurs propres associées  $\Lambda_{\mathbf{L}_{\mathbf{W}}} = [\lambda_1 \leq \cdots \leq \lambda_N]$ .

Shuman *et al.* ont montré [2] qu'il existe une correspondance entre les fréquences de la base de Fourier en traitement de signal classique et les valeurs propres  $\Lambda_{L_W}$  du Laplacien normalisé. Cette propriété a permis de définir pour le traitement de signal sur graphes des outils tels que la transformée de Fourier, la convolution de signaux ou encore la transformée en ondelettes [2].

Dans ce contexte de transposition des outils du traitement de signal classique au traitement de signal sur graphes, Agaskar & Lu ont montré [1] qu'un signal sur un graphe ne pouvait pas être à la fois localisé en temps et en fréquence. Le détail de leurs travaux fait l'objet de la section suivante.

## 3 Le principe d'incertitude appliqué au traitement de signal sur graphes

Le principe d'incertitude d'Heisenberg indique qu'un signal x ne peut pas être à la fois localisé en temps et en fréquence. Plus exactement, il existe une expression analytique décrivant ce compromis :

$$\Delta_t^2(\boldsymbol{x})\Delta_\omega^2(\boldsymbol{x}) \ge \frac{1}{4} \tag{1}$$

où  $\Delta_t^2(x)$  est l'étalement temporel du signal, et  $\Delta_{\omega}^2(x)$  est son étalement fréquentiel.

Afin de transposer cette notion au traitement de signal sur graphes, Agaskar & Lu [1] proposent de définir des notions analogues à  $\Delta_t^2(\boldsymbol{x})$  et  $\Delta_{\omega}^2(\boldsymbol{x})$ . Le graphe support du signal étant défini pour généraliser le domaine temporel à des topologies plus complexes, l'étalement temporel trouve son analogue dans une notion d'*étalement graphique* :

**Définition 4 (Etalement graphique) –** Soit un signal x porté par un graphe  $\mathcal{G} = \langle \mathcal{V}, \mathcal{E}, W \rangle$ . Soit  $u_0 \in \mathcal{V}$ . L'étalement graphique  $\Delta^2_{\mathcal{G}, u_0}(x)$  du signal autour du nœud  $u_0$  est défini<sup>1</sup> par :

$$\Delta_{\mathcal{G},u_0}^2(\boldsymbol{x}) \triangleq \sum_{u \in \mathcal{V}} d(u_0, u) x_u^2 = \boldsymbol{x}^\top \boldsymbol{P}_{u_0} \boldsymbol{x}$$
(2)

où d est une fonction de distance, et  $P_{u_0}$  est la matrice diagonale des distances à  $u_0$ . Dans la définition originale [1], les auteurs utilisent pour d la distance géodésique au carré entre les nœuds. Ce choix a été discuté en [5].

De plus, la correspondance entre domaine de Fourier et valeurs propres du Laplacian normalisé ayant été établie, Agaskar & Lu définissent un *étalement spectral* pour le signal [1] :

**Définition 5 (Etalement spectral) –** Soit un signal x porté par un graphe  $\mathcal{G} = \langle \mathcal{V}, \mathcal{E}, \mathbf{W} \rangle$ . L'étalement spectral  $\Delta_s^2(\mathbf{x})$  du signal est défini par :

$$\Delta_s^2(\boldsymbol{x}) \triangleq \sum_{n=1}^N \lambda_n \hat{\boldsymbol{x}}_n^2 = \boldsymbol{x}^\top \boldsymbol{L}_{\boldsymbol{W}} \boldsymbol{x}$$
(3)

<sup>1.</sup> Nous rappelons que nous considérons des signaux normalisés. L'équation originale en [1] inclut une normalisation par la norme de x qui n'est pas nécessaire ici.

où  $\Lambda_{L_W} = [\lambda_1 \leq \cdots \leq \lambda_N]$  sont les valeurs propres de  $L_W$ , et  $\hat{x} = [\hat{x}_1 \dots \hat{x}_N] \triangleq [\chi_1^\top x \dots \chi_N^\top x]$  est la transformée de Fourier [2] de x sur le graphe. On remarquera que  $\Delta_s^2(x)$  est minimal pour un signal ayant toute son énergie sur la valeur propre  $\lambda_1 = 0$ , ce qui est caractéristique des signaux diffusés un grand nombre de fois sur un graphe non biparti.

Informellement, ces deux notions quantifient la localité d'un signal dans le domaine graphique ou dans le domaine spectral. Par exemple, pour un signal  $\boldsymbol{x} = [1; 0 \dots 0]$ , nous obtenons  $\Delta_{\mathcal{G},1}^2(\boldsymbol{x}) = 0$ , ce qui correspond bien à un signal parfaitement localisé autour du nœud 1 dans le domaine graphique. De même, un signal ayant toute son énergie sur une unique valeur propre  $\lambda_i$  aura un étalement spectral égal à  $\lambda_i$ .

Agaskar & Lu ont montré que le compromis entre précision d'un signal dans le domaine graphique et précision dans le domaine spectral est dépendant de la topologie du graphe. Ainsi, il n'est pas possible de définir une expression analytique analogue à (1) qui soit non triviale et valable pour tout graphe.

Afin d'étudier le principe d'incertitude d'Heisenberg appliqué aux signaux portés par des graphes, il convient donc de fixer un graphe  $\mathcal{G}$ , et un nœud  $u_0$  servant de référence pour l'étalement graphique. On peut ensuite déterminer une *courbe d'incertitude* représentant les meilleurs compromis possibles :

**Définition 6 (courbe d'incertitude) –** *La courbe d'incertitude*  $\gamma_{u_0}$  associée à un graphe  $\mathcal{G}$ , pour  $u_0$  choisi, est définie par :

$$\gamma_{u_0}(\alpha) \triangleq \min_{\boldsymbol{x}} \Delta^2_{\mathcal{G}, u_0}(\boldsymbol{x}) \ \text{tel que } \Delta^2_s(\boldsymbol{x}) = \alpha \qquad (4)$$

et peut être tracée en faisant varier  $\alpha$  de 0 à 1 [1].

Nous choisissons pour le reste de cette section d'étudier un graphe étoilé non pondéré (voir Figure 1) et fixons  $u_0$  comme étant le nœud central. Pour ce graphe, la courbe d'incertitude est présentée en Figure 2. Contrairement à la plupart des graphes, il est possible de montrer que cette courbe est la même pour tout graphe étoilé quelle que soit la valeur de N:



FIGURE 2 – Courbe d'incertitude associée à un graphe étoilé. Le nœud central du graphe est choisi comme  $u_0$ , et on utilise la distance géodésique au carré pour le calcul de  $\Delta_{G,u_0}^2$ .

Cette approche pour tracer  $\gamma_{u_0}$  requiert de parcourir l'espace des signaux afin de déterminer ceux qui minimisent  $\Delta^2_{\mathcal{G},u_0}$  pour  $\Delta^2_s$  fixé. On notera  $\tilde{\boldsymbol{x}}$  un vecteur vérifiant la courbe d'incertitude, *i.e* tel que  $\langle \Delta^2_{\mathcal{G},u_0}(\tilde{\boldsymbol{x}}), \Delta^2_s(\tilde{\boldsymbol{x}}) \rangle$  appartient à  $\gamma_{u_0}$ .

Pour un graphe étoilé, en choisissant  $u_0$  comme nœud de référence, Agaskar & Lu ont montré que tout  $\tilde{x}$  est de la forme  $[\tilde{x}_1; \tilde{x}_2 \dots \tilde{x}_2]$ . Sans perdre en généralité, il est donc possible de décrire  $\tilde{x}$  par un paramètre  $\theta : \tilde{x} = [\cos \theta; \frac{\sin \theta}{\sqrt{N-1}} \dots \frac{\sin \theta}{\sqrt{N-1}}]$ . On peut donc obtenir une expression analytique de la courbe d'incertitude, en utilisant le fait que les signaux solution sont inclus dans un cercle. Ainsi, pour un graphe étoilé,  $\gamma_{u_0}$  correspond à la portion inférieure gauche d'une ellipse d'équation :

$$(\Delta_s^2 - 1)^2 + (2\Delta_{\mathcal{G},u_0}^2 - 1)^2 = 1.$$
(5)

Pour des graphes plus complexes, les auteurs utilisent un algorithme approché appelé *algorithme sandwich* afin de tracer une approximation arbitrairement précise de  $\gamma_{u_0}$ .

# 4 Réduction de l'espace de recherche des signaux solutions de la courbe d'incertitude

Afin de réduire l'espace de recherche pour la résolution de l'équation (4), nous voulons pouvoir caractériser les signaux réalisant la courbe d'incertitude. Afin de proposer une telle caractérisation, nous étendons la démarche menée par Agaskar & Lu en Annexe C de [1]. Cette démarche avait été initiée pour montrer que les signaux  $\tilde{x}$  pour un graphe complet ou pour un graphe étoilé (pour  $u_0$  nœud central) avaient une forme particulière. L'objectif de cette section est de généraliser l'approche de Agaskar & Lu en [1] pour la rendre applicable à une plus vaste classe de graphes. On se propose de prouver la Propriété 1 :

**Propriété 1.** Soit une matrice  $M(\alpha) \triangleq P_{u_0} - \alpha L_W$ , pour  $\alpha$  fixé. Si  $M(\alpha)$  est de la forme :

$$\boldsymbol{M}(\alpha) = \left(\begin{array}{c|c} \boldsymbol{A} & \boldsymbol{B} \\ \hline \boldsymbol{C} & \boldsymbol{D} \end{array}\right) \tag{6}$$

où :

- *A est une matrice carrée de dimension j.*
- **B** est une matrice constante par ligne, i.e  $\mathbf{B} = \mathbf{y} \mathbf{1}_j^\top$  pour  $\mathbf{y}$  un vecteur quelconque et  $\mathbf{1}_j$  un vecteur constant égal à 1 de dimension j.
- *C* est une matrice quelconque.

 $[\tilde{x}_1]$ 

- D est une matrice carrée circulante de dimension k. alors  $\tilde{x}$  est de la forme :

$$\dots \tilde{x}_j; \underbrace{\tilde{x}_{j+1} \dots \tilde{x}_{j+1}}_{k \text{ fois}} ].$$
(7)

Pour les besoins de la preuve, on rappelle la Propriété 2 [1] :

**Propriété 2.** Tout signal  $\tilde{x}$  réalisant la courbe d'incertitude est un vecteur propre associé à la plus faible valeur propre d'une matrice  $M(\alpha) \triangleq P_{u_0} - \alpha L_W$ .

Démonstration de la Propriété 1.

Comme D est circulante, on a que 1<sub>k</sub> est vecteur propre de D. Par construction de M(α), D est symmétrique, donc décomposable dans une base orthonormale. Soient (χ<sub>i</sub>)<sub>1≤i≤k-1</sub> des vecteurs propres pour D orthogonaux à 1<sub>k</sub>, associés aux valeurs propres (λ<sub>i</sub>)<sub>1≤i≤k-1</sub>. Par construction, on a ⟨χ<sub>i</sub>|1<sub>k</sub>⟩ = 0 ∀i ∈ [1; k − 1].

- Pour tout i ∈ [1; k − 1], on construit un vecteur χ<sub>i</sub><sup>+</sup> ≜ [0<sub>j</sub><sup>T</sup>; χ<sub>i</sub><sup>T</sup>]<sup>T</sup>, où 0<sub>j</sub> est le vecteur nul de dimension j. En utilisant le fait que B est constante par ligne, on obtient que ∀i : M(α)χ<sub>i</sub><sup>+</sup> = λ<sub>i</sub>χ<sub>i</sub><sup>+</sup>. Par conséquent, tout χ<sub>i</sub><sup>+</sup> est vecteur propre de M(α). En applicant la méthodologie en Annexe C de [1] (inégalité de Rayleigh), on obtient que le vecteur propre associé à la plus faible valeur propre de M(α) doit être orthogonal à χ<sub>i</sub><sup>+</sup> ∀i ∈ [1; k−1].
- En notant que les j premières composantes des vecteurs χ<sup>+</sup><sub>i</sub> sont nulles, et par application de la Propriété 2, on obtient que tout vecteur x̃ réalisant la courbe d'incerti-tude est de la forme donnée en (7).

*Remarque* 1. L'application de la Propriété 1 peut être faite de manière récursive sur **A**, permettant de raffiner la caractérisation des  $\tilde{x}_1 \dots \tilde{x}_j$  en (7) et de réduire l'espace des solutions.

Pour illustrer la caractérisation des signaux  $\tilde{x}$ , nous considérons à nouveau un graphe étoilé, mais choisissons à présent  $u_0$ comme étant un des nœuds *feuille* du graphe. On obtient que  $\mathbf{M}(\alpha)$  peut être décomposée de manière à faire apparaître une sous-matrice carrée circulante **D** de dimension N - 2. Par la Propriété 1 on obtient que, pour ce graphe et ce choix de nœud de référence,  $\tilde{x}$  est de la forme  $[\tilde{x}_1; \tilde{x}_2; \tilde{x}_3 \dots \tilde{x}_3]$ .

Afin de donner une expression analytique de l'espace des signaux solution, on peut considérer l'ensemble des signaux de norme unitaire définis sur une hypersphère de dimension M, où M est le nombre de composantes distinctes dans la caractérisation du signal. Dans l'exemple considéré, on peut donc réduire l'ensemble des solutions potentielles aux signaux caractérisés par deux paramètres  $\theta$  et  $\phi$  de la façon suivante :  $[\cos \theta; \sin \theta \cos \phi; \frac{\sin \theta \sin \phi}{\sqrt{N-2}}].$ 

La Figure 3 représente les paires  $\langle \Delta^2_{\mathcal{G},u_0}(\tilde{x}), \Delta^2_s(\tilde{x}) \rangle$  pour des signaux  $\tilde{x}$  obtenus par sous-échantillonnage de l'espace des signaux, en itérant sur des valeurs possibles de  $\theta$  et  $\phi$  sur l'intervalle  $[0; 2\pi]$  avec un pas de 0.05. De plus, les courbes d'incertitude obtenues par l'algorithme d'approximation *sandwich* sont présentées, et coïncident bien avec la frontière inférieure de l'ensemble des solutions.

### 5 Conclusion

Dans ce document, nous avons présenté une extension de la méthode introduite par Agaskar & Lu en [1] afin de caractériser les signaux réalisant la courbe d'incertitude pour un graphe donné et un nœud de référence choisi. Alors que l'approche initiale ne s'applique que pour les graphes complets et étoilés (uniquement quand le nœud de référence est celui au centre de l'étoile), notre méthode peut caractériser la structure des signaux solution pour une plus large classe de graphes. Pour illustrer nos résultats, nous avons tracé un sous-ensemble des signaux potentiellement solution pour un graphe étoilé en choisissant une feuille comme nœud de référence, et les résultats concordent bien avec ceux de l'algorithme approché.



FIGURE 3 – Sous-échantillonnage de l'espace des signaux potentiellement solution pour un graphe étoilé, selon le choix du nœud  $u_0$ . La bordure inférieure gauche de chaque ensemble est la courbe d'incertitude. L'approximation par l'algorithme sandwich concorde avec les résultats obtenus.

Deux extensions sont envisagées pour nos travaux. La première est d'améliorer notre méthode afin de réduire l'espace de recherche pour une classe de graphes encore plus large. La seconde est d'étudier l'impact du choix du nœud de référence sur la courbe d'incertitude d'un graphe, afin de proposer une notion de courbe minimale pour le graphe et de permettre la comparaison de différents graphes.

#### Remerciements

Ce travail a été partiellement financé par le Conseil de Recherche Européen par le biais du programme *European Union's* Seventh Framework Programme (FP7/2007-2013) / ERC grant agreement n° 290901.

### Références

- Ameya Agaskar et Yue M. Lu. A Spectral Graph Uncertainty Principle. CoRR, abs/1206.6356, 2012.
- [2] David I. Shuman, Sunil K. Narang, Pascal Frossard, Antonio Ortega et Pierre Vandergheynst. *The Emerging Field of Signal Processing on Graphs : Extending High-Dimensional Data Analysis to Networks and Other Irregular Data Domains*. CoRR, abs/1211.0053, 2013.
- [3] Fan R. K. Chung. *Spectral Graph Theory*. American Mathematical Society, 1997.
- [4] Gerald B. Folland et Alladi Sitaram. *The uncertainty principle : A mathematical survey*. Journal of Fourier Analysis and Applications, Birkhäuser-Verlag, 1997.
- [5] Bastien Pasdeloup, Réda Alami, Vincent Gripon et Michael Rabbat. Toward An Uncertainty Principle For Weighted Graphs. Soumis à EUSIPCO, http:// arxiv.org/abs/1503.03291, 2015.