



Analyse harmonique sur graphes dirigés et applications : de l'analyse de Fourier aux ondelettes

Harry Sevi

► To cite this version:

Harry Sevi. Analyse harmonique sur graphes dirigés et applications : de l'analyse de Fourier aux ondelettes. Analyse fonctionnelle [math.FA]. Université de Lyon, 2018. Français. NNT : 2018LY-SEN068 . tel-02024654

HAL Id: tel-02024654

<https://theses.hal.science/tel-02024654>

Submitted on 19 Feb 2019

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.



Numéro National de Thèse : 2018LYSEN068

THESE de DOCTORAT DE L'UNIVERSITE DE LYON

opérée par
l'Ecole Normale Supérieure de Lyon

Ecole Doctorale N° 52
Ecole Doctorale de Physique et d'Astrophysique de Lyon

Discipline : Physique

Soutenue publiquement le 22/11/2018, par :

Harry SEVI

Analyse harmonique sur graphes dirigés et applications : de l'analyse de Fourier aux ondelettes

Devant le jury composé de :

BORGNAT, Pierre	Directeur de recherche	ENS de Lyon	Directeur
FADILI, Jalal	Professeur des Universités	ENSICAEN	Examineur
GRIBONVAL, Rémi	Directeur de recherche	Inria Rennes	Rapporteur
MELOT, Clothilde	Maître de Conférences	Aix Marseille Université	Examinatrice
MICHEL, Olivier	Professeur	INP Grenoble	Rapporteur
PICARD, Dominique	Professeure	LPMA	Examinatrice
RILLING, Gabriel	Ingénieur-Chercheur	CEA Saclay	Co-encadrant

Table des matières

1	Introduction	8
I	Analyse harmonique sur graphes non dirigés	13
2	Généralités sur l'analyse harmonique sur graphes non dirigés	14
1	Motivation	15
2	Notions sur la théorie des graphes non dirigés	16
3	Généralités sur la théorie des opérateurs	17
4	Fondamentaux du traitement de signal sur graphes	19
4.1	Signaux sur graphes	19
4.2	Filtrage sur graphes	20
5	Opérateurs fondamentaux sur graphes non dirigés	21
5.1	Marche aléatoire sur graphes	21
5.2	Laplaciens de graphe	23
6	Transformées de Fourier sur graphes non dirigés	26
7	Analyse de Fourier sur graphes non dirigés	27
7.1	Régularité de signaux sur graphes	27
7.2	Analyse fréquentielle sur graphe non dirigé	29
8	Analyse multirésolution sur graphes non dirigés	30
8.1	La transformée en ondelettes continue classique	30
8.2	Transformée en ondelettes spectrale	32
9	Conclusion	34
II	Analyse harmonique sur graphes dirigés	35
3	État de l'art sur les propositions existantes en analyse de Fourier sur graphes dirigés	36
1	Notions sur la théorie des graphes dirigés	37
2	L'analyse de Fourier basée sur la matrice d'adjacence	38
2.1	L'analyse de Fourier selon Sandryhaila et Moura	39
2.2	L'analyse de Fourier selon Mhaskar	41

3	L'analyse de Fourier via une notion de variation dirigée. . . .	42
3.1	L'analyse de Fourier selon Sardellitti et coll.	42
3.2	L'analyse de Fourier selon Shafipour et coll.	43
4	Discussion	44
III Contributions		46
4	Analyse harmonique sur graphes dirigés	47
1	Présentation du cadre théorique	48
2	Filtrage sur graphes dirigés	49
2.1	Définitions	49
2.2	Stabilité des filtres sur graphes	50
3	Opérateurs fondamentaux sur graphes dirigés	57
3.1	Marche aléatoire sur graphes dirigés	57
3.2	Généralisation de marches aléatoires	59
3.3	Laplaciens sur graphes dirigés	60
3.4	Opérateurs fondamentaux sur graphes dirigés et espaces de Hilbert	62
4	Transformée de Fourier sur graphes dirigés	63
5	Analyse de Fourier sur graphes dirigés	64
5.1	Régularité de signaux sur graphes	64
5.2	Analyse fréquentielle sur graphes dirigés	65
5.3	Filtres sur graphes de type marche aléatoire	67
5.4	Analyse de Fourier sur les groupes finis : un point de vue traitement de signal sur graphe	68
6	Applications	71
6.1	Apprentissage semi-supervisé sur des graphes dirigés via régularisation de type ℓ_2	72
6.2	Modélisation du signal sur des graphes dirigés par filtrage	78
7	Analyses multi-résolution sur graphes dirigés	90
7.1	Transformée en ondelettes redondantes sur graphes dirigés	91
7.2	Transformée en ondelettes décimée sur graphes dirigés	94
7.3	Applications	98
5	Conclusion et Perspectives	105
Bibliographie		110

Table des figures

4.1	Évolution des valeurs propres ν_1 et ν_2 en fonction de l'angle entre les projecteurs spectraux $\mathbf{G}_1$ et $\mathbf{G}_2$	56
4.2	Distribution des valeurs propres de la matrice de marche aléatoire d'un graphe toroïdal dirigé $\mathcal{T}_{6,4}$. Valeurs propres sur la ligne verte pointillée : $\Re(\lambda^g) = -0.25, \forall \lambda^g \in \Lambda_g$. Valeurs propres sur la ligne orange pointillée : $\Re(\lambda^o) = 0.25, \forall \lambda^o \in \Lambda_o$	70
4.3	Représentation bidimensionnelle des vecteurs propres $\xi_{28,49}$ et $\xi_{28,50}$ de la matrice de transition du toroïdal dirigé $\mathcal{T}_{54,36}$. L'axe est l'étiquette d'un noeud dans $\mathbf{C}_m$ (resp. $\mathbf{C}_m$) dans la direction x (resp. direction y)	71
4.4	Apprentissage semi-supervisé sur $\mathcal{G}'$ des blogs politiques de la campagne présidentielle américaine de 2004. Les notations L_{norm} et $L_{\text{norm,sym}}$ correspondent respectivement à $\mathcal{L}$ et $\mathcal{L}_{\text{sym}}$	77
4.5	(Cas 1) Gauche : Reconstruction du signal sur graphe $\mathbf{f}'_0$ sur le sous-graphe $\mathcal{G}'$. Droite : Reconstruction du signal sur graphe $\mathbf{f}'_0$ sur la version symétrique de $\mathcal{G}'$	81
4.6	(Cas 2) Gauche : Reconstruction du signal sur graphe $\mathbf{f}'_0$ sur le sous-graphe $\mathcal{G}'$. Droite : Reconstruction du signal graphe $\mathbf{f}'_0$ sur la version symétrique de $\mathcal{G}'$	82
4.7	(Cas 1.1) Reconstruction du signal sur graphe $\mathbf{f}'_0$ sur le sous-graphe $\mathcal{G}'$	84
4.8	(Cas 1.2) Reconstruction du signal sur graphe $\mathbf{f}'_0$ sur le sous-graphe $\mathcal{G}'$	84
4.9	(Cas 2.1) Reconstruction du signal sur graphe $\mathbf{f}'_0$ sur le sous-graphe $\mathcal{G}'$	85
4.10	(Cas 2.2) Reconstruction du signal sur graphe $\mathbf{f}'_0$ sur le sous-graphe $\mathcal{G}'$	86
4.11	Reconstruction du signal sur graphe $\mathbf{f}_0$ sur $\mathcal{G}$	88
4.12	(Cas 1) Reconstruction du signal sur graphe $\mathbf{f}_0$ sur $\mathcal{G}$	89
4.13	(Cas 2) Reconstruction du signal sur graphe $\mathbf{f}_0$ sur $\mathcal{G}$	89
4.14	Fonctions d'échelle orthogonal et biorthogonal sur un graphe cyclique dirigé $\mathcal{C}_{256}$	98

4.15	Ondelettes orthogonales et biorthogonales sur le graphe cyclique dirigé $\mathcal{C}_{256}$	99
4.16	Fonctions d'échelle et ondelettes orthogonales sur $\mathcal{G}_2$	100
4.17	Fonctions d'échelle et ondelettes biorthogonales sur $\mathcal{G}_2$	101
4.18	Apprentissage semi-supervisé sur la plus grande composante connexe du graphe des blogs politiques.	103

Remerciements

Je tiens tout d'abord à remercier mes superviseurs, Gabriel Rilling et Pierre Borgnat. Tout au long de ces trois années, j'ai énormément appris à vos côtés, par les nombreuses discussions que nous avons pu avoir. Vous avez fait preuve d'une grande patience et liberté lorsqu'il s'agissait d'écrire, de structurer mes idées, d'explorer les idées qui me tenaient à cœur.

Je remercie chaleureusement Rémi Gribonval et Olivier Michel d'avoir accepté de relire mon manuscrit. Je remercie Dominique Picard, Jalal Fadili et Clothide Mélot qui m'ont fait l'honneur de participer à mon jury de thèse.

Cette thèse a été l'accomplissement de trois années intenses, mais également de trois années où je me suis bien amusé. Je remercie la ville de Grenoble, où j'ai effectué ma thèse, et dont le cadre m'a permis d'être disponible et libre pour penser. C'est également une ville où j'ai noué de belles amitiés. Je tiens à remercier ma bande d'amis de Grenoble, c'est-à-dire Basile, Nicolas, Damien, Gabriel, Sami, Naïm, Fanny, Margot, Zag ainsi que Ben, Mahoussi, Zack (Team Str808) pour tous les grands moments passés ensemble. Je tiens à remercier particulier Thibaud Bernard Helis (aka TBH) pour tous les moments qu'on a passé à Grenoble entre promenades et discussions entre maths, art, philosophie, religion et Booba bien sûr !.

Durant ces trois années de thèse, j'ai également pu mal voyager pour mes travaux de recherche. Je tiens à remercier mon ami de Philadelphie, Josh pour les quelques semaines assez incroyables que nous avons passé ensemble entre sessions musique à Fishtown, après midi avec Hodgy Beats, soirées dans des lieux cachées de Philly, et d'autres histoires rocambolesques !

Je tiens à remercier ma LA Team, Julien, Abby, Alex, Chris et Jackson pour tous les moments que nous avons partagé entre les sessions studios, les soirées au Low End Theory, au Hollywood Bowl à écouter Herbie Hancock et Kamashi Washington et à rouler dans la ville.

Je tiens à remercier également Benjamin pour son toute aide, son soutien et les discussions que nous avons pu avoir durant mon séjour à Los Angeles et pour les lunettes qu'il m'a eu pour l'éclipse, ça a été super pratique, tout le monde venait me les prêter dans DTLA en allant à USC !

Je tiens à remercier Antonio Ortega de m'avoir accueilli à USC et d'avoir "challenge" mes idées. Merci à Aamir, mon collègue de USC pour sa disponibilité et avec qui j'ai pu avoir d'enrichissantes discussions.

Je remercie ma Math Team Romain, Randall, Yanis, Neil et Guillaume pour toute l'aide qui m'ont apporté durant ma thèse et pour leur grand soutien.

Je remercie mes amis d'enfance Irwin, Justinien, Simon, Théodore, Alex, Felix et Raphael ainsi que mes grands amis Nitany et David pour leur soutien.

Je souhaite exprimer toute ma reconnaissance à mes parents, mes frères et mes sœurs qui m'ont soutenu depuis le début et qu'ont toujours encouragé à faire ce que je souhaitais et donner le meilleur de moi-même.

Enfin, je tiens à remercier Charlotte pour son soutien sans faille, sa patience et ses encouragements.

Chapitre 1

Introduction

Dans un monde où les données disponibles à des fins scientifiques ou sociales s'accumulent à un rythme exponentiel, la gestion, l'exploitation et l'analyse de ce torrent de données sont devenues l'un des défis de notre époque. Certaines d'entre elles prennent la forme de graphes, structures qui apparaissent dans divers domaines : les neurosciences [1], les réseaux sociaux [2], l'internet des objets [3] et les crypto-monnaies [4, 5] pour n'en nommer que quelques-uns. Par conséquent, il existe une explosion de l'intérêt pour les méthodes d'analyse de structures à domaine irrégulier tel que les graphes et les réseaux [6, 7]. Il est donc nécessaire d'élaborer des approches mathématiques et informatiques efficaces pour traiter et analyser ces graphes et les données qu'ils contiennent.

Parmi les méthodes existantes, un grand nombre implique le laplacien (de graphe) [8, 9, 10, 11, 12] :

- En tant qu'objet fondamental en mathématiques et physique, le laplacien est connu et utilisé, à travers son étude spectrale pour extraire des propriétés géométriques pertinentes d'une variété [13] ou d'un graphe [14] et joue un rôle crucial dans les applications d'apprentissage automatique comme le partitionnement spectral [12].
- Dans l'espace continu, les fonctions propres de l'opérateur de Laplace-Beltrami obtenues par sa décomposition spectrale, constituent la généralisation de la base de Fourier pour les variétés [13, 15].
- Dans le cadre de l'espace discret, les vecteurs propres du laplacien de graphe forment une base de Fourier orthonormale pour les fonctions sur graphes [16, 17] lorsque les graphes sont considérés comme une discrétisation ou comme un échantillonnage d'une variété [18, 19, 10, 11]. La valeur propre associée à chaque vecteur propre est liée à la notion de fréquence [16].

Grâce à ces propriétés, le laplacien de graphe devient l'objet créant le lien fondamental entre la théorie spectrale des graphes et le traitement du signal à travers le cadre mathématique du "traitement du signal sur graphes" [16, 17].

Ce cadre vise à étendre les concepts de traitement du signal traditionnel, tel que le filtrage ou l'échantillonnage à des fonctions définies sur les sommets d'un graphe.

Tout d'abord développé dans le cadre des graphes non dirigés, le traitement de signal sur graphes considère le laplacien de graphe (combinatoire ou normalisé) [14] comme élément central. Le laplacien de graphe est un opérateur symétrique et semi-défini positif. Par le théorème spectral, le laplacien de graphe admet une base orthonormale dont les vecteurs propres peuvent être considérés comme les modes de Fourier sur graphe non dirigé et dont les valeurs propres respectives sont associées à une notion de fréquence.

L'objectif de cette thèse est d'apporter des éléments de réponse à l'extension du cadre du traitement de signal sur graphes au cas des graphes dirigés. L'extension du traitement de signal sur graphe dirigé est due au fait que la plupart des structures de graphes que l'on retrouve dans les domaines scientifiques ou sociaux sont dirigées. Dans le domaine des neurosciences, les informations sensorielles circulent dans le cerveau de neurone à neurone à travers des synapses. L'existence d'une synapse d'un neurone a à un neurone b n'implique pas l'existence d'une synapse dans le sens opposé (de b vers a). Dans le cas des réseaux sociaux, en particulier le réseau Twitter, les informations diffusées sont souvent dirigées d'un émetteur de l'information (*followee* en anglais) vers le receveur de l'information (*follower* en anglais) et non nécessairement dans le sens opposé. Par conséquent, l'analyse des informations associées à ces types de graphes dirigés nécessite également l'élaboration d'approches mathématiques et donc au passage d'une extension du cadre du traitement de signal sur graphes au cas des graphes dirigés.

Formellement, l'utilisation directe de l'élément central du traitement de signal sur graphes, c'est-à-dire le laplacien de graphe, n'est plus appropriée. Un graphe dirigé est naturellement représenté par une matrice d'adjacence non symétrique. Il est toujours possible de définir naïvement un laplacien de graphe dans le cas dirigé, mais les propriétés spectrales associées au laplacien sur graphe non dirigé telles que l'obtention d'une base orthonormale et de valeurs propres positives ou nulles réelles ne sont plus vérifiées dans le cas des graphes dirigés. Il n'existe pas de consensus simple quant à une définition d'un laplacien pour graphe dirigé et dont la variation des vecteurs propres soit liée à une notion de fréquence. C'est le principal défi du traitement de signal sur graphes dirigés et de cette thèse : quels sont les opérateurs de référence associés à une analyse de Fourier qui permettraient probablement de créer ce fameux lien entre les graphes dirigés et non dirigés ?

Un petit nombre de constructions de bases de Fourier sur graphes dirigés [20, 21, 22, 23] ont été proposées ces dernières années. Parmi ces constructions, celle proposée par [20] se fonde sur la construction d'une base de Fourier à partir des vecteurs propres de la matrice d'adjacence d'un graphe dirigé. Les approches [23, 22] se fondent sur la construction d'une base de Fourier orthonormale déterminée à travers la résolution d'un problème d'optimi-

sation non convexe faisant intervenir une mesure de variation particulière. Enfin, l’approche [21] se fonde sur la décomposition polaire de la matrice d’adjacence d’un graphe dirigé.

Dans cette thèse, l’opérateur de marche aléatoire associé à une marche aléatoire sur graphes [24, 25] est proposé comme opérateur de référence pour étendre le cadre du traitement de signal sur graphes dirigés. Tout comme le laplacien de graphe, l’opérateur de marche aléatoire est un opérateur lié à la notion de diffusion. Par ailleurs, il est aussi bien défini sur graphe dirigé que non dirigé. Sous hypothèse que l’opérateur de marche aléatoire soit diagonalisable, il admet potentiellement des vecteurs propres complexes conjugués ainsi que des valeurs propres complexes conjuguées associées.

Notre cadre repose sur le fait que l’analyse variationnelle du vecteur propre de l’opérateur de marche aléatoire à travers son énergie de Dirichlet [24] associée est liée à la partie réelle de sa valeur propre [26]. Par conséquent, l’énergie de Dirichlet d’un vecteur propre donné de l’opérateur de marche aléatoire sur un graphe dirigé peut être considérée comme sa fréquence, en utilisant la même analogie que dans [16], cette fois-ci pour les graphes dirigés. Par cette analyse en fréquence, nous sommes en mesure de proposer une base de Fourier pour des fonctions définies sur les sommets d’un graphe fortement connexe dirigé. Par ailleurs, notre approche est cohérente avec l’analyse de Fourier dans le cas non dirigé.

En considérant dorénavant l’opérateur de marche aléatoire comme un opérateur de référence et ses vecteurs propres associés comme une base de Fourier pour fonctions définies sur graphe dirigé, la dernière partie de notre analyse harmonique sur graphe dirigé consiste à construire une analyse multirésolution.

En effet, les graphes complexes aussi bien dirigés que non dirigés présentent des structures à des échelles différentes. Motivées par l’efficacité de l’analyse multirésolution en ondelettes dans le traitement de signal traditionnel [27] et la dimension à différentes échelles des graphes et des données sur graphes, un certain nombre de constructions multirésolution sur graphes ont émergé ces dernières années [28, 29, 30, 31, 32, 33, 34, 35, 36, 37] dans le cadre du traitement de signal sur graphe non dirigé. Dans le cas dirigé, une seule analyse multirésolution redondante a été récemment proposée [21] généralisant le cadre théorique introduit dans [31]. Face aux approches existantes d’analyse harmonique sur les graphes dirigés (c.à-d. analyse de Fourier et analyse multirésolution), nous proposons donc une nouvelle analyse harmonique construite autour de l’opérateur de marche aléatoire et dont les constructions multirésolution généralisent les approches bien établies des ondelettes spectrales [32, 31] et des ondelettes de diffusion [30, 38] sur les graphes non dirigés.

La thèse est organisée en deux parties. La première partie comprenant le chapitre 2 apporte une vue d’ensemble des résultats pertinents de l’analyse harmonique, c’est-à-dire l’analyse de Fourier et les constructions multiréso-

lution sur graphe non dirigé.

La seconde partie de la thèse est constituée des chapitres 3 et 4. Dans le chapitre 3, nous abordons les notions de l'état de l'art des différentes approches de transformée de Fourier pour des fonctions définies sur les graphes dirigés [20, 21, 22, 23]. Nous discutons également des éventuelles limites de ces approches.

Le chapitre 4 correspond finalement à notre contribution où nous présentons une nouvelle analyse harmonique pour des fonctions définies sur graphes dirigés. Nos contributions sont les suivantes :

1. **Stabilité de filtres sur graphes.** Nous étudions la stabilité de filtres sur graphes en proposant d'abord des bornes sur le conditionnement sur des filtres en section 2.2.1 et ensuite en proposant des méthodes d'approximation de filtres robustes vis-à-vis d'une perturbation additive d'un signal sur graphe en section 2.2.2.
2. **Construction de bases de Fourier sur graphes dirigés.** Nous proposons l'opérateur de marche aléatoire associé à une marche aléatoire sur un graphe dirigé comme un opérateur de référence. Nous déterminons une analyse en fréquence des vecteurs propres de l'opérateur de marche aléatoire et déterminons une caractérisation variationnelle des vecteurs propres de l'opérateur de marche aléatoire à la proposition 4.7.
3. **Construction d'analyses multirésolution sur graphes dirigés.** Nous proposons une transformée en ondelettes redondantes spectrale sur graphes dirigés à la section 7.1. Cette construction étend le cadre des ondelettes spectrales sur graphes non dirigés [32] et des trames d'ondelettes polynomiales [31]. Nous proposons une transformée en ondelettes à échantillonnage critique à la section 7.2 généralisant le cadre des ondelettes de diffusion [30, 38] au cas dirigé.
4. **Efficience du cadre théorique à travers des applications.** Le développement de notre analyse harmonique sur graphes dirigés nous amène à considérer des problèmes d'apprentissage semi-supervisés avec régularisation de type ℓ^2 en section 6.1 et régularisation ℓ^1 en section 7.3.3 et de modélisation de signaux par filtrage sur graphes dirigés en section 6.2 mettant en évidence l'efficacité de notre cadre théorique.

Ce manuscrit de thèse est une présentation détaillée de certains travaux déjà publiés et d'autres en cours de publication.

Publications dans des journaux

- Sevi, H., Rilling, G., Borgnat, P. (2018). Harmonic analysis on directed graphs and applications : from Fourier analysis to wavelets. Submitted to ACHA.

- Couillet, R., Wainrib, G., Sevi, H., Ali, H. T. (2016). The asymptotic performance of linear echo state neural networks. JMLR. Accepted.

Publications dans des conférences

- Sevi, H., Rilling, G., Borgnat, P. (2018). Modeling signals over directed graphs through filtering. GLOBALSIP, Anaheim.
- Sevi, H., Rilling, G., Borgnat, P. (2017). Multiresolution analysis of functions on directed networks. Wavelet and Sparsity XVII, San Diego.
- Sevi, H., Rilling, G., Borgnat, P. (2017). Analyse fréquentielle et filtrage sur graphes dirigés. GRETSI, Juan-les-Pins.

Exposés oraux

- Sevi, H., Rilling, G., Borgnat, P. (2018). Frequency interpretation of random walk subspaces on directed graphs (Poster session). 3rd Graph signal processing workshop, EPFL, Lausanne.
- Sevi, H., Rilling, G., Borgnat, P. (2016). On the Conditioning Properties of Linear Shift Invariant Operators on Graphs (Poster session). 1st Graph signal processing workshop, UPenn, Philadelphie.

Première partie

Analyse harmonique sur
graphes non dirigés

Chapitre 2

Généralités sur l'analyse harmonique sur graphes non dirigés

Dans ce chapitre, nous abordons les notions fondamentales de l'analyse harmonique sur graphes non dirigés. Dans un premier temps, nous rappellerons des notions essentielles de théorie des graphes non dirigés et les notions fondamentales du traitement de signal sur graphes tels que la notion de filtrage sur graphes. Dans un second temps, nous introduirons les opérateurs essentiels de l'analyse harmonique sur graphes non dirigés tels que l'opérateur de marche aléatoire et le laplacien de graphe et ses variantes et nous mentionnons leurs propriétés respectives. Cela nous amènera à introduire la notion de transformée de Fourier sur graphes et son analyse fréquentielle. Enfin, nous évoquerons la notion d'analyse multirésolution sur graphes non dirigés en présentant la construction d'ondelettes spectrales sur graphes dont la construction s'inspire de la construction en ondelettes continue dans le traitement de signal classique.

Sommaire

1	Motivation	15
2	Notions sur la théorie des graphes non dirigés	16
3	Généralités sur la théorie des opérateurs	17
4	Fondamentaux du traitement de signal sur graphes	19
4.1	Signaux sur graphes	19
4.2	Filtrage sur graphes	20
5	Opérateurs fondamentaux sur graphes non dirigés	21
5.1	Marche aléatoire sur graphes	21
5.2	Laplaciens de graphe	23
6	Transformées de Fourier sur graphes non dirigés	26
7	Analyse de Fourier sur graphes non dirigés	27
7.1	Régularité de signaux sur graphes	27
7.2	Analyse fréquentielle sur graphe non dirigé	29
8	Analyse multirésolution sur graphes non dirigés	30
8.1	La transformée en ondelettes continue classique	30
8.2	Transformée en ondelettes spectrale	32
9	Conclusion	34

1 Motivation

L'analyse harmonique est l'étude de la décomposition d'une fonction quelconque $f : [0, 2\pi] \rightarrow \mathbb{R}$ en fonctions sinusoïdales. L'analyse harmonique est l'un des piliers des mathématiques dû à son impact dans de nombreux domaines. Récemment, motivé par la nécessité d'analyser des données associées à des domaines non euclidiens tels que des variétés ou des graphes et par l'efficacité de l'analyse harmonique sur des domaines réguliers (e.g. images, séries temporelles), un nouveau cadre théorique a émergé utilisant les concepts fondamentaux de l'analyse harmonique afin d'étudier les données sur des graphes. L'idée principale est simple : les sinus et cosinus interviennent naturellement comme les fonctions propres de l'opérateur laplacien (ou laplacien) $-\Delta = -d^2/dx^2$ sur l'intervalle $[0, 2\pi]$ et sont également les modes de Fourier. Le laplacien peut être considéré comme l'écart d'une fonction de sa moyenne locale, par exemple, en considérant l'approximation aux différences finies pour le laplacien dans $\mathbb{R}^2$:

$$-\Delta f(x, y) \approx \frac{4}{\varepsilon^2} \left(f(x, y) - \frac{f(x + \varepsilon, y) + f(x - \varepsilon, y) + f(x, y + \varepsilon) + f(x, y - \varepsilon)}{4} \right).$$

De manière générale, nous pouvons également considérer des fonctions vivant sur des domaines plus complexes tels que les variétés riemanniennes.

En effet, les fonctions propres de l'opérateur de Laplace-Beltrami (ou laplacien de variété) constituent les modes de Fourier sur la variété. Par ailleurs, la discrétisation de la variété nous amène à considérer une approximation de celle-ci sous forme d'un graphe où l'approximation de l'opérateur de Laplace-Beltrami sur ce graphe est appelée le laplacien de graphe. Ainsi, les vecteurs propres du laplacien de graphe constituent l'approximation des fonctions propres de l'opérateur de Laplace-Beltrami et forment une base de type Fourier [18, 19] sur graphe. Ces considérations ont ainsi motivé l'émergence d'une analyse harmonique sur graphes et la motivation du choix de l'opérateur laplacien comme élément-clé de l'analyse harmonique sur des graphes.

2 Notions sur la théorie des graphes non dirigés

Un graphe est un objet mathématique composé d'un ensemble de sommets reliés entre eux par des arêtes. Tout au long de ce chapitre, nous considérons uniquement le cas des graphes non dirigés. Cette section contient les principales définitions relatives aux graphes non dirigés [39, 40].

Définition 2.1. *Un graphe simple (c'est-à-dire sans arêtes multiples) non dirigé et non pondéré est un couple $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ où $\mathcal{V}$ constitue l'ensemble des sommets et $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ l'ensemble des arêtes, chaque arête étant une paire de sommets non nécessairement distincts. Si u et v sont des sommets, nous représentons une arête reliant les sommets u et v par $e = (u, v)$.*

Si $e = (u, v)$ forme une arête, alors nous disons que u et v sont adjacents ou que v est voisin de u . Ceci est représenté par $u \sim v$. Dans cette thèse, nous étudierons exclusivement les graphes **finis**, c'est-à-dire des graphes dont l'ensemble des sommets et l'ensemble des arêtes est fini. Ainsi le terme "graphe" voudra toujours dire "graphe simple fini". Nous définissons par $|\mathcal{V}|$, la cardinalité de l'ensemble des sommets $\mathcal{V}$, c'est-à-dire le nombre d'éléments dans $\mathcal{V}$. Le graphe $\mathcal{G}$ peut être représenté à travers sa matrice d'adjacence que nous définissons de la manière suivante.

Définition 2.2 (Représentation algébrique d'un graphe non dirigé et non pondéré). *Soit $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ un graphe simple non dirigé et non pondéré où $\mathcal{V} = (v_1, \dots, v_N)$ et de cardinalité $|\mathcal{V}| = N$. La matrice d'adjacence $\mathbf{A} = \{a_{ij}\}_{1 \leq i, j \leq N} \in \mathbb{R}^{N \times N}$ se rapportant à cet ensemble de sommets est la matrice $N \times N$ booléenne avec :*

$$a_{ij} = \begin{cases} 1 & \text{si } (v_i, v_j) \in \mathcal{E} \\ 0 & \text{sinon.} \end{cases}$$

et la relation :

$$a_{ij} = a_{ji}, \quad \forall i, j \in \{1, \dots, N\}. \quad (2.1)$$

Nous noterons que la relation (2.1) caractérise la notion de symétrie et est exprimée sous forme matricielle par la relation :

$$\mathbf{A}^\top = \mathbf{A},$$

où $\mathbf{A}^\top$ est la matrice transposée de $\mathbf{A}$.

Plus généralement, nous pouvons considérer le cas des graphes non dirigés et pondérés. Nous définissons un graphe non dirigé et pondéré de la manière suivante.

Définition 2.3. Soit $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ un graphe non dirigé et non pondéré et $w : \mathcal{V} \times \mathcal{V} \rightarrow \mathbb{R}_+$ une application qui associe une valeur $w_{ij} = w(v_i, v_j)$ à chaque arête (v_i, v_j) . L'application w est appelée fonction de poids. Le triplet $\mathcal{G} = (\mathcal{V}, \mathcal{E}, w)$ est donc un graphe non dirigé et pondéré.

Un graphe non dirigé et pondéré $\mathcal{G} = (\mathcal{V}, \mathcal{E}, w)$ peut être entièrement représenté par sa matrice d'adjacence $\mathbf{W} = \{w_{ij}\}_{1 \leq i, j \leq N} \in \mathbb{R}_+^{N \times N}$. Si $w_{ij} = 0$, pour un couple $(i, j) \in \{1, \dots, N\}^2$, alors il n'existe pas d'arête entre les sommets v_i et v_j . Nous définissons le degré (pondéré) d'un sommet $v_i \in \mathcal{V}$ comme la somme des poids des arêtes associées à v_i , c'est-à-dire $d_i = \sum_{j=1}^N w_{ij}$. Si le graphe $\mathcal{G}$ est non pondéré, le degré d'un sommet est simplement le nombre de sommets voisins à ce sommet. Dans ce chapitre, nous supposons les poids de la matrice d'adjacence $\mathbf{W}$ positifs ou nuls.

Nous définissons la notion de chaîne dans un graphe non dirigé de la manière suivante.

Définition 2.4. Soit $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ un graphe non dirigé. Une chaîne reliant un sommet v_i à un sommet v_j est définie par une suite finie d'arêtes consécutives, reliant v_i à v_j .

Nous définissons un graphe non dirigé connexe de la manière suivante.

Définition 2.5. Un graphe non dirigé $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ est connexe si chaque sommet est accessible à partir de n'importe quel autre. Autrement dit, si pour tout couple de sommets distincts $(v_i, v_j) \in \mathcal{V} \times \mathcal{V}$ il existe une chaîne reliant v_i à v_j .

3 Généralités sur la théorie des opérateurs

Nous rappelons les notions essentielles de la théorie des opérateurs en dimension finie nécessaires à la compréhension des prochaines sections [41, 42].

Dans cette section, $\mathbf{H}$ et $\mathbf{H}'$ sont deux espaces de Hilbert de dimension $\dim \mathbf{H} = N < \infty$, $\dim \mathbf{H}' = M < \infty$ et $\mathbf{e} = (e_1, \dots, e_N)$ est une base ortho-normale de $\mathbf{H}$. Pour tout vecteur donné $x \in \mathbf{H}$, ses coordonnées par rapport à

e seront notées $x_1, \dots, x_N$, c'est-à-dire $x = x_1 e_1 + \dots + x_N e_N$ et $\mathbf{x}$ le vecteur colonne correspondant :

$$\mathbf{x} = (x_1, \dots, x_N)^\top.$$

Nous définissons la notion d'opérateur de la manière suivante.

Définition 2.6. *Toute application linéaire continue $T : \mathbf{H} \rightarrow \mathbf{H}'$ s'appelle un opérateur. L'espace vectoriel $\mathcal{L}(\mathbf{H}, \mathbf{H}')$ des applications linéaires continues de $\mathbf{H}$ dans $\mathbf{H}'$ est l'espace des opérateurs de $\mathbf{H}$ dans $\mathbf{H}'$.*

Par souci de simplicité, supposons $\mathbf{H} = \mathbf{H}'$. Un opérateur linéaire $T : \mathbf{H} \rightarrow \mathbf{H}$ est représenté par rapport à la base orthonormale $\mathbf{e}$ par une matrice $\mathbf{T} \in \mathbb{C}^{N \times N}$ tel que les $t_{ij} \in \mathbb{C}$ de $\mathbf{T}$ s'expriment de la manière suivante :

$$t_{ji} = \langle e_j, T e_i \rangle.$$

Nous définissons la notion d'opérateur adjoint de la manière suivante.

Définition 2.7. *Soit $T \in \mathcal{L}(\mathbf{H}, \mathbf{H})$. Alors il existe un unique opérateur $T^* \in \mathcal{L}(\mathbf{H}, \mathbf{H})$ tel que :*

$$\langle T x, y \rangle_{\mathbf{H}} = \langle x, T^* y \rangle_{\mathbf{H}} \quad \forall x, y \in \mathbf{H}.$$

L'opérateur T^ s'appelle l'adjoint de T .*

En dimension finie, l'opérateur T^* , est représenté par la matrice conjuguée et transposée $\overline{\mathbf{T}}^\top$ de $\mathbf{T}$ (ou seulement $\mathbf{T}^\top$ si $\mathbf{T} \in \mathbb{R}^{N \times N}$).

Définition 2.8. *Si $T \in \mathcal{L}(\mathbf{H}, \mathbf{H})$ est tel que $T = T^*$, on dit que l'opérateur T est autoadjoint (ou hermitien).*

Les opérateurs autoadjoints ont des propriétés particulièrement importantes que nous examinons.

Théorème 2.1 (Théorème spectral des opérateurs auto-adjoints [43]). *Soit $\mathbf{H}$ un espace de Hilbert de dimension $\dim \mathbf{H} = N < \infty$ et $T \in \mathcal{L}(\mathbf{H}, \mathbf{H})$ un opérateur auto-adjoint. Alors il existe une base orthonormale $\{f_n\}_{n=1}^N$ de $\mathbf{H}$ et une suite de réels $\{\lambda_n\}_{n=1}^N$ tels que pour tout $n \in \{1, \dots, N\}$, $T f_n = \lambda_n f_n$ et pour tout $x \in \mathbf{H}$:*

$$T x = \sum_{n=1}^N \lambda_n \langle x, f_n \rangle f_n.$$

Nous définissons une isométrie de la manière suivante.

Définition 2.9. *Soient $\mathbf{H}, \mathbf{H}'$ deux espaces de Hilbert. Une transformation linéaire $V : \mathbf{H} \mapsto \mathbf{H}'$ est une isométrie si et seulement si :*

$$\|V(x - y)\|_{\mathbf{H}'} = \|x - y\|_{\mathbf{H}}, \quad \forall x, y \in \mathbf{H}.$$

Nous définissons l'entrelacement entre deux opérateurs linéaires de la manière suivante.

Définition 2.10. Soient H, H' deux espaces de Hilbert. Une transformation linéaire inversible bornée $V \in \mathcal{L}(H, H')$ entrelace un opérateur linéaire $M \in \mathcal{L}(H, H)$ avec un opérateur linéaire $S \in \mathcal{L}(H', H')$ si :

$$VM = SV.$$

Les opérateurs linéaires M et S sont alors dits similaires.

Nous introduisons maintenant la notion d'opérateur semi-défini positif.

Définition 2.11. Soit $T \in \mathcal{L}(H, H)$ un opérateur auto-adjoint. L'opérateur T est dit semi-défini positif si :

$$\langle x, Tx \rangle \geq 0, \quad \forall x \in H.$$

4 Fondamentaux du traitement de signal sur graphes

Dans cette section, nous introduisons les aspects fondamentaux du traitement de signal sur graphes [16, 17, 44].

4.1 Signaux sur graphes

Soit $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ un graphe non dirigé et $f : \mathcal{V} \rightarrow \mathbb{C}$ une fonction définie sur l'ensemble des sommets $\mathcal{V} = (v_1, \dots, v_N)$. Nous définissons un signal sur graphe, noté $\mathbf{f}$ comme la représentation sous forme de vecteur colonne de la fonction f appliquée en chacun des sommets $v_i \in \mathcal{V}$, c'est-à-dire :

$$\mathbf{f} = (f(v_1), \dots, f(v_N))^T \in \mathbb{C}^N.$$

Nous définissons la notion d'espace de fonctions définies sur les sommets d'un graphe arbitraire de la manière suivante.

Définition 2.12 ([45]). Soit $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ un graphe non dirigé et $\mu : \mathcal{V} \rightarrow [0, \infty)$ une fonction sur $\mathcal{V}$ considérée comme une mesure sur $\mathcal{V}$ en posant $\mu(U) = \sum_{x \in U} \mu(x)$, $U \subset \mathcal{V}$. Pour $p \in [1, \infty)$, nous notons $\ell^q(\mathcal{V}, \mu)$, l'espace de fonctions $f : \mathcal{V} \rightarrow \mathbb{C}$ tel que

$$\|\mathbf{f}\|_{\ell^q(\mu, \mathcal{V})} = \begin{cases} \left(\sum_{x \in \mathcal{V}} |f(x)|^q \mu(x) \right)^{1/q} < \infty, & q \in [1, \infty). \\ \max_{x \in \mathcal{V}} |f(x)| \mu(x) < \infty, & q = \infty. \end{cases}$$

Nous supposons les fonctions sur $\mathcal{G}$ définies sur l'espace $\ell^2(\mathcal{V}, \mu)$. L'espace $\ell^2(\mathcal{V}, \mu)$ est l'espace de Hilbert de fonctions définies sur l'ensemble des sommets $\mathcal{V}$ de $\mathcal{G}$ muni du produit scalaire suivant :

$$\langle \mathbf{f}, \mathbf{g} \rangle_\mu = \sum_{x \in \mathcal{V}} \overline{f(x)} g(x) \mu(x),$$

pour tout $\mathbf{f}, \mathbf{g} \in \ell^2(\mathcal{V}, \mu)$ et où $\overline{f(x)}$ est le complexe conjugué $f(x)$.

4.2 Filtrage sur graphes

Un opérateur linéaire sur un graphe non dirigé $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ de cardinalité $|\mathcal{V}| = N$ est représenté par une matrice $\mathbf{H} \in \mathbb{C}^{N \times N}$ qui agit sur un signal sur graphe $\mathbf{s} \in \mathbb{C}^N$ et produit un signal sur graphe $\tilde{\mathbf{s}} \in \mathbb{C}^N$ selon le produit matrice vecteur :

$$\tilde{\mathbf{s}} = \mathbf{H}\mathbf{s}.$$

Dans le cadre du traitement de signal sur graphes et inspiré du traitement de signal classique, on s'intéresse principalement à un type de filtres sur graphes qui commutent avec un opérateur de référence sur graphe non dirigé $\mathbf{R}$, c'est-à-dire :

$$\mathbf{H}\mathbf{R} = \mathbf{R}\mathbf{H}.$$

Nous définissons un filtre sur graphe $\mathbf{H}$ comme une somme finie polynomiale d'un opérateur de référence $\mathbf{R}$, c'est-à-dire :

$$\mathbf{H} = \sum_{t=0}^T h_t \mathbf{R}^t, \quad h_t \in \mathbb{C}, \quad \forall t = 0, \dots, T.$$

Le théorème suivant établit, sous une condition particulière, qu'un filtre sur graphe commutant avec un opérateur de référence donné peut s'exprimer comme un polynôme de celui-ci.

Théorème 2.2. *Soit $\mathbf{R}$ un opérateur linéaire de référence sur un graphe non dirigé $\mathcal{G}$ que nous supposons diagonalisable. Nous supposons également les polynômes caractéristique et minimal de $\mathbf{R}$ égaux, c.-à-d. $p_{\mathbf{R}}(x) = m_{\mathbf{R}}(x)$. Alors un filtre sur graphe $\mathbf{H}$ qui commute avec $\mathbf{R}$ s'exprime comme une somme finie polynomiale de $\mathbf{R}$, c'est-à-dire :*

$$\mathbf{H} = \sum_{t=0}^T h_t \mathbf{R}^t, \quad h_t \in \mathbb{C}, \quad \forall t = 0, \dots, T.$$

Le théorème 2.2 indique que si un filtre $\mathbf{H}$ jouit des trois conditions suivantes :

1. Commute avec un opérateur de référence $\mathbf{R}$.
2. Diagonalisable.
3. Chacun des espaces propres est de dimension un.

alors le filtre peut s'exprimer comme une somme finie polynomiale de $\mathbf{R}$. Ce type de filtre sur graphe est appelé filtre *linéairement invariant* par rapport à $\mathbf{R}$ (*linear shift invariant* en anglais) par Sandryhaila et Moura [44]. Cette notion d'invariance linéaire par rapport à un opérateur de référence sur graphe non dirigé semble assez restrictive dans la mesure où de nombreux opérateurs de référence sur graphes n'admettent pas systématiquement des espaces

propres de dimension égale à 1 mais éventuellement des espaces propres de dimension supérieure à 1. Par contre, il sera toujours possible de construire un filtre sur graphe comme une somme finie polynomiale d'un opérateur de référence. Un tel filtre est toujours invariant par rapport à son opérateur de référence. En revanche un filtre qui commute avec un opérateur de référence ne respectant pas l'une des trois conditions énoncées ci-dessus ne peut pas toujours s'exprimer comme un polynôme de cet opérateur de référence.

5 Opérateurs fondamentaux sur graphes non dirigés

L'analyse harmonique pour fonctions définies sur les sommets d'un graphe non dirigé est principalement construite autour de l'opérateur laplacien ou de ses variantes. Dans cette section, nous introduisons ces opérateurs fondamentaux ainsi que leurs propriétés nécessaires au cadre de l'analyse harmonique sur graphe non dirigé.

5.1 Marche aléatoire sur graphes

L'opérateur de marche aléatoire est un des opérateurs centraux de l'analyse harmonique sur graphes [24] non dirigés notamment comme élément de base dans le cadre de la réduction de dimension sur graphes [11] et de la construction d'ondelettes sur graphes [30]. Dans cette section, nous introduisons les définitions et les propriétés essentielles de l'opérateur de marche aléatoire sur graphes non dirigés [46, 47, 48, 25, 49].

Définition 2.13. *Une marche aléatoire sur un graphe non dirigé et pondéré $\mathcal{G} = (\mathcal{V}, \mathcal{E}, w)$ est une chaîne de Markov homogène et réversible $\mathcal{X} = (X_n)_{n \geq 0}$ munie d'un espace d'états fini $\mathcal{V}$ et dont les probabilités de transition sont proportionnelles aux poids des arêtes. Notamment les entrées de $\mathbf{P} = \{p(x, y)\}_{x, y \in \mathcal{V}}$ de la matrice de transition de $\mathcal{X}$ sur $\mathcal{V}$ sont :*

$$p(x, y) = \mathbb{P}(X_{n+1} = y | X_n = x).$$

Les entrées de $\mathbf{P}$ sont des probabilités, il en suit que :

$$p(x, y) \geq 0 \quad \text{et} \quad \sum_{y \in \mathcal{V}} p(x, y) = 1, \quad \forall x, y \in \mathcal{V}.$$

La matrice de transition $\mathbf{P} \in \mathbb{R}^{N \times N}$ peut s'exprimer de la manière suivante :

$$\mathbf{P} = \mathbf{D}^{-1} \mathbf{W},$$

où $\mathbf{D} = \text{diag}\{d_1, \dots, d_n\}$ est la matrice diagonale des degrés des sommets et $\mathbf{W}$ la matrice d'adjacence pondérée.

Pour tout signal sur graphe $f : \mathcal{V} \mapsto \mathbb{C}$, mesure $\mu : \mathcal{V} \rightarrow [0, \infty)$ et états $x, y \in \mathcal{V}$, on adopte les conventions suivantes [49] :

$$\begin{aligned}\mathbf{P}\mathbf{f}(x) &= \sum_{y \in \mathcal{V}} p(x, y)f(y), \\ \mu\mathbf{P}(y) &= \sum_{x \in \mathcal{V}} \mu(x)p(x, y).\end{aligned}$$

où $\mathbf{f}$ et μ sont les représentations vectorielles de f et μ respectivement.

Considérons l'application de la matrice de transition $\mathbf{P} \in \mathbb{R}^{N \times N}$ à une fonction $\mathbf{f}$ à sa droite. La valeur de $\mathbf{P}\mathbf{f}$ au sommet i correspond à la somme des valeurs de $\mathbf{f}$ aux sommets voisins de i , pondérée par les valeurs de la ligne i de $\mathbf{P}$. Comme la somme des valeurs sur une ligne de $\mathbf{P}$ est toujours égale à un, appliquer $\mathbf{P}$ sur une fonction à sa droite correspond à calculer pour chaque sommet du graphe une moyenne des valeurs de la fonction $\mathbf{f}$ aux sommets voisins. Étant donnée l'importance du moyennage local en traitement du signal, cela suggère que l'opérateur de marche aléatoire $\mathbf{P}$ peut être naturellement un élément central du traitement du signal sur graphes.

Nous définissons l'irréductibilité d'une marche aléatoire sur graphe de la manière suivante.

Définition 2.14. *On dit qu'une marche aléatoire $\mathcal{X}$ est irréductible si pour tout $x, y \in \mathcal{V}$, la probabilité partant de x d'atteindre y est strictement positive, autrement dit :*

$$\forall x, y \in \mathcal{V}, \exists m < \infty : \mathbb{P}(X_{n+m} = y | X_n = x) > 0.$$

Dans le cas des graphes connexes non dirigés, la marche aléatoire associée est irréductible.

Nous définissons la périodicité de la manière suivante.

Définition 2.15. *Soit $\mathcal{X}$ une marche aléatoire sur $\mathcal{G}$. La période d'un état $x \in \mathcal{V}$ est :*

$$\varrho(x) = \text{pgcd}\{n \in \mathbb{N}^+ : p^{(n)}(x, x) > 0\}.$$

Ainsi partant de x , la chaîne de Markov peut retourner à l'état x seulement en un multiple près de la période ϱ . L'état x est apériodique si $\varrho(x) = 1$ et périodique si $\varrho(x) > 1$.

Ayant défini les notions d'irréductibilité et de périodicité, nous sommes en mesure de définir la notion d'ergodicité d'une marche aléatoire.

Définition 2.16. *Soit $\mathcal{X}$ une marche aléatoire sur $\mathcal{G}$. La marche aléatoire $\mathcal{X}$ est ergodique si elle est irréductible et apériodique.*

Nous énonçons la proposition concernant la distribution stationnaire d'une marche aléatoire ergodique sur $\mathcal{G}$.

Proposition 2.1 ([25]). *Soit $\mathcal{G}$ un graphe non dirigé à espace d'états fini $\mathcal{V}$. Si une marche aléatoire $\mathcal{X}$ munie de sa matrice de transition $\mathbf{P}$ est ergodique, c.-à-d. irréductible et apériodique, les mesures $\mathbf{P}^n(x, \cdot)$ convergent vers le vecteur ligne $\boldsymbol{\pi} = (\pi(v_1), \dots, \pi(v_N)) \in \mathbb{R}_+^N$ quand n tend vers l'infini c.-à-d. l'unique distribution stationnaire. Notamment, $\boldsymbol{\pi}\mathbf{P} = \boldsymbol{\pi}$, avec :*

$$\sum_{i=1}^N \pi(v_i) = 1, \quad \pi(v_i) \geq 0.$$

Définition 2.17. *Une marche aléatoire est réversible par rapport à une mesure π si :*

$$\pi(x)p(x, y) = \pi(y)p(y, x), \quad \forall x, y \in \mathcal{V}.$$

Les marches aléatoires sur graphes non dirigés sont réversibles. Si nous supposons la marche aléatoire $\mathcal{X}$ ergodique, alors la distribution stationnaire associée π admet une forme analytique. Par conséquent, la marche aléatoire associée admet une unique distribution stationnaire π définie par $\pi(x) = d_x/d_{\mathcal{G}}$ où d_x est le degré du sommet x et $d_{\mathcal{G}} = \sum_{x \in \mathcal{V}} d_x$ [49].

5.2 Laplaciens de graphe

Comme motivé à la section 1, le laplacien de graphe constitue la pierre angulaire de l'analyse harmonique sur graphe non dirigé. Dans cette section, nous présentons le laplacien de graphe, ses variantes ainsi que leurs propriétés [14].

5.2.1 Laplacien combinatoire

Soit $\mathcal{G} = (\mathcal{V}, \mathcal{E}, w)$ un graphe non dirigé et pondéré avec l'ensemble des sommets $\mathcal{V} = (v_1, \dots, v_N)$ et de cardinalité $|\mathcal{V}| = N$. Le graphe $\mathcal{G}$ est représenté par sa matrice d'adjacence $\mathbf{W} \in \mathbb{R}_+^{N \times N}$. Le laplacien combinatoire sur $\mathcal{G}$ est représenté par la matrice $\mathbf{L} \in \mathbb{R}^{N \times N}$ définie de la manière suivante :

$$\mathbf{L} = \mathbf{D} - \mathbf{W}, \tag{2.2}$$

où la matrice $\mathbf{D}$ est la matrice diagonale des degrés des sommets de $\mathcal{G}$, c'est-à-dire $\mathbf{D} = \text{diag}\{d_1, \dots, d_N\}$. Les éléments de $\mathbf{L}$ sont donnés par :

$$\mathbf{L}_{ij} = \begin{cases} d_i & \text{si } v_i = v_j, \\ -w_{ij} & \text{si } (v_i, v_j) \in \mathcal{E} \\ 0 & \text{sinon.} \end{cases}$$

5.2.2 Laplacien normalisé

Le laplacien normalisé sur un graphe non dirigé et pondéré $\mathcal{G} = (\mathcal{V}, \mathcal{E}, w)$ est représenté par la matrice $\mathcal{L} \in \mathbb{R}^{N \times N}$ définie de la manière suivante :

$$\mathcal{L} = \mathbf{I} - \mathbf{D}^{-1/2} \mathbf{W} \mathbf{D}^{-1/2}. \quad (2.3)$$

Les éléments de $\mathcal{L}$ sont donnés par :

$$\mathcal{L}_{ij} = \begin{cases} 1 & \text{si } v_i = v_j, \\ -\frac{w_{ij}}{\sqrt{d_i d_j}} & \text{si } (v_i, v_j) \in \mathcal{E} \\ 0 & \text{sinon.} \end{cases}$$

Les laplaciens normalisé (2.3) et combinatoire (2.2) sont reliés de la manière suivante

$$\mathcal{L} = \mathbf{D}^{-1/2} \mathbf{L} \mathbf{D}^{-1/2}.$$

5.2.3 Laplacien marche aléatoire

Le laplacien marche aléatoire sur un graphe simple non dirigé et pondéré $\mathcal{G} = (\mathcal{V}, \mathcal{E}, w)$ dénoté $\mathbf{L}_{\text{RW}}$ est défini de la manière suivante :

$$\mathbf{L}_{\text{RW}} = \mathbf{I} - \mathbf{D}^{-1} \mathbf{W} = \mathbf{I} - \mathbf{P}. \quad (2.4)$$

Les laplaciens marche aléatoire (2.4) et combinatoire (2.2) sont reliés de la manière suivante :

$$\mathbf{L}_{\text{RW}} = \mathbf{D}^{-1} \mathbf{L}.$$

Les laplaciens marche aléatoire et normalisé (2.3) sont également reliés de la manière suivante :

$$\mathcal{L} = \mathbf{D}^{1/2} \mathbf{L}_{\text{RW}} \mathbf{D}^{-1/2}.$$

5.2.4 Propriétés

Les opérateurs laplaciens combinatoire (2.2), normalisé (2.3) et marche aléatoire (2.4) sont munis d'avantageuses propriétés que nous abordons dans cette section-ci.

Proposition 2.2. *Soit $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ un graphe non dirigé. Les opérateurs laplaciens combinatoire et normalisé définis en (2.2) et (2.3) sont auto-adjoints pour l'espace $\ell^2(\mathcal{V})$. L'opérateur laplacien marche aléatoire défini en (2.4) est quant à lui auto-adjoint pour l'espace $\ell^2(\mathcal{V}, \pi)$ où π est l'unique distribution stationnaire associée à la marche aléatoire ergodique $\mathcal{X}$.*

En dimension finie, la proposition 2.2 signifie que les représentations matricielles des opérateurs laplaciens combinatoire (normalisé) et marche aléatoire sont symétriques respectivement dans les espaces $\ell^2(\mathcal{V})$ et $\ell^2(\mathcal{V}, \pi)$.

Proposition 2.3 ([50]). *Le laplacien combinatoire et son équivalent normalisé sont semi définis positifs par rapport au produit scalaire induit par la mesure de comptage. Le laplacien marche aléatoire est quant à lui semi défini positif par rapport au produit scalaire associé à la mesure π .*

Les propositions 2.2 et 2.3 permettent d'introduire la proposition suivante.

Proposition 2.4 ([51]). *Le laplacien combinatoire et son équivalent normalisé sont tous deux auto-adjoints et semi définis positifs. Par le théorème spectral, ces laplaciens admettent respectivement des bases orthonormales $\{\phi_k\}_{k=1}^N$ et $\{\varphi_k\}_{k=1}^N$ dans $\ell^2(\mathcal{V})$ associées aux valeurs propres $\{\theta_k\}_{k=1}^N$ et $\{\vartheta_k\}_{k=1}^N$ positives ou nulles satisfaisant la relation :*

$$\begin{aligned}\mathbf{L}\phi_k &= \theta_k \phi_k, \\ \mathcal{L}\varphi_k &= \vartheta_k \varphi_k.\end{aligned}$$

Le laplacien marche aléatoire admet une base orthonormale $\{\chi_k\}_{k=1}^N$ dans $\ell^2(\mathcal{V}, \pi)$ et admet les mêmes valeurs propres que celles de $\mathcal{L}$.

Démonstration. Les laplaciens combinatoire et normalisé, respectivement $\mathbf{L}$ et $\mathcal{L}$ sont des opérateurs auto-adjoints et par le théorème spectral (théorème 2.1), il existe respectivement des bases orthonormales $\{\phi_k\}_{k=1}^N$ et $\{\varphi_k\}_{k=1}^N$ dans $\ell^2(\mathcal{V})$ associées à l'ensemble de valeurs propres $\{\theta_k\}_{k=1}^N$ et $\{\vartheta_k\}_{k=1}^N$ telles que :

$$\begin{aligned}\mathbf{L}\phi_k &= \theta_k \phi_k, \\ \mathcal{L}\varphi_k &= \vartheta_k \varphi_k.\end{aligned}$$

Par ailleurs, on en déduit par la proposition 2.3 que toutes les valeurs propres $\{\theta_k\}_{k=1}^N$ et $\{\vartheta_k\}_{k=1}^N$ sont positives ou nulles.

Montrons que le laplacien marche aléatoire $\mathbf{L}_{\text{RW}}$ et le laplacien normalisé $\mathcal{L}$ possèdent les mêmes valeurs propres.

Soit $\chi_i = \mathbf{\Pi}^{-1/2} \varphi_i$ où $\mathbf{\Pi} = \text{diag}\{\pi(v_1), \dots, \pi(v_N)\}$, la matrice diagonale de la mesure stationnaire. Nous avons $\mathbf{\Pi}$ proportionnel à $\mathbf{D}$. Le vecteur χ_i est un vecteur propre de $\mathbf{L}_{\text{RW}}$. En effet nous avons :

$$\mathbf{L}_{\text{RW}} \chi_i = \mathbf{\Pi}^{-1/2} \mathbf{\Pi}^{1/2} \mathbf{L}_{\text{RW}} \mathbf{\Pi}^{-1/2} \varphi_i \propto \mathbf{\Pi}^{-1/2} \mathcal{L} \varphi_i = \vartheta_i \chi_i.$$

Par conséquent, le laplacien marche aléatoire a les mêmes valeurs propres que le laplacien normalisé.

Par ailleurs, la famille de vecteurs propres $\{\chi_k\}_{k=1}^N$ forme une base orthonormée pour $\ell^2(\mathcal{V}, \pi)$. En effet, nous avons :

$$\delta_{ij} = \langle \varphi_i, \varphi_j \rangle = \langle \mathbf{\Pi}^{1/2} \chi_i, \mathbf{\Pi}^{1/2} \chi_j \rangle = \langle \chi_i, \chi_j \rangle_\pi.$$

□

6 Transformées de Fourier sur graphes non dirigés

Soit $\mathbf{L}$ (resp. $\mathcal{L}$) le laplacien combinatoire (resp. normalisé) défini sur un graphe non dirigé $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ de cardinalité $|\mathcal{V}| = N$. La matrice $\mathbf{L}$ (resp. $\mathcal{L}$) est symétrique semi-définie positive. Étant en dimension finie, elle est diagonalisable, c'est-à-dire que $\mathbf{L}$ (resp. $\mathcal{L}$) admet la décomposition en valeurs propres suivante :

$$\mathbf{L} = \mathbf{\Xi} \mathbf{\Theta} \mathbf{\Xi}^\top, \quad (\text{resp. } \mathcal{L} = \mathbf{\Phi} \mathbf{\Lambda} \mathbf{\Phi}^\top)$$

où $\mathbf{\Xi} = [\xi_1, \dots, \xi_N]$ (resp. $\mathbf{\Phi} = [\phi_1, \dots, \phi_N]$) est une matrice orthonormale de vecteurs propres et $\mathbf{\Theta} = \text{diag}\{\theta_1, \dots, \theta_N\}$ (resp. $\mathbf{\Lambda} = \text{diag}\{\lambda_1, \dots, \lambda_N\}$) est la matrice de valeurs propres non nécessairement distinctes. Étant donné un signal sur le graphe $\mathbf{s}$, sa *transformée de Fourier* notée $\hat{\mathbf{s}} = [\hat{s}_1, \dots, \hat{s}_N]$ est :

$$\hat{\mathbf{s}} = \mathbf{\Xi}^\top \mathbf{s} \quad (\text{resp. } \hat{\mathbf{s}} = \mathbf{\Phi}^\top \mathbf{s}).$$

Les valeurs $\{\hat{s}_j\}_{j=1}^N$ caractérisent le contenu du signal sur graphe $\mathbf{s}$ dans le domaine de Fourier sur graphe et sont appelés coefficients de Fourier.

La matrice de vecteurs propres de $\mathbf{L}$ est habituellement choisie comme base de type Fourier pour des signaux définis sur graphes non dirigés. De manière analogue, la matrice de vecteurs propres du laplacien normalisé peut également être choisie comme une base de type Fourier. Par conséquent, il existe dans l'état de l'art plusieurs bases de Fourier pour signaux sur graphes non dirigés.

Nous verrons dans les prochaines sections que le choix de la base de Fourier est en partie associé au problème traité de traitement de signal sur graphe et à l'espace de signaux sur graphe considéré. Contrairement au traitement de signal classique, l'unicité de la base de Fourier sur graphe n'est pas garantie comme nous le verrons également à la section 4 du chapitre 4.

La *transformée de Fourier inverse* correspondant à la matrice de changement de base $\mathbf{\Xi}$ est exprimée de la manière suivante :

$$\mathbf{s} = \mathbf{\Xi} \hat{\mathbf{s}}.$$

Elle reconstruit le signal originel en formant une combinaison linéaire des vecteurs propres du laplacien pondérés par les coefficients de Fourier associés.

Dans la littérature du traitement de signal sur graphe, les bases de type Fourier utilisées sont celles associées au laplacien combinatoire ou normalisé dû à leurs symétries et ainsi à l'application du théorème spectral sur ces opérateurs. Cependant, les vecteurs propres du laplacien marche aléatoire ne sont habituellement pas utilisés comme base de type Fourier sur graphe non dirigé. Cela est dû au fait que sa base ne soit pas orthogonale dans $\ell^2(\mathcal{V})$. Néanmoins, le laplacien marche aléatoire admet une base orthonormale dans $\ell^2(\mathcal{V}, \pi)$. Ainsi les vecteurs propres de $\mathbf{L}_{\text{RW}}$ constituent une base de Fourier qui peut être choisie orthonormale pour des signaux sur graphes appartenant

à l'espace $\ell^2(\mathcal{V}, \pi)$, tandis que la matrice de vecteurs propres de $\mathbf{L}$ (resp. $\mathcal{L}$) est une base de type Fourier orthonormale pour des signaux sur graphes appartenant à l'espace $\ell^2(\mathcal{V})$.

7 Analyse de Fourier sur graphes non dirigés

Soit $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ un graphe non dirigé de cardinalité $|\mathcal{V}| = N$. Dans la section précédente, les vecteurs propres du laplacien et de ses variantes ont été proposés comme base(s) de Fourier sur graphe non dirigé. Nous aimerions caractériser respectivement l'aspect fréquentiel de ces bases de Fourier. Une manière de caractériser l'aspect fréquentiel d'un signal sur graphe est d'en mesurer la variation ou la régularité.

7.1 Régularité de signaux sur graphes

Le comportement d'un signal défini sur graphe non dirigé peut être analysé en mesurant sa régularité ou sa variation. Nous définissons tout d'abord la "longueur du gradient" d'un signal sur graphe non dirigé de la manière suivante.

Définition 2.18 ([24]). *Soit $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ un graphe non dirigé et $\nu : \mathcal{E} \rightarrow [0, \infty)$ une mesure positive définie sur l'ensemble des arêtes $\mathcal{E}$. La "longueur du gradient" d'un signal sur graphe $f : \mathcal{V} \rightarrow \mathbb{C}$ situé au sommet $x \in \mathcal{V}$ sous la mesure ν est :*

$$|\nabla \mathbf{f}(x)| = \left(\frac{1}{2} \sum_{y \in \mathcal{V}, x, y \in \mathcal{E}} |f(x) - f(y)|^2 \nu(x, y) \right)^{1/2}.$$

Intuitivement, la longueur du gradient du graphe mesure la variation d'un signal sur graphe autour d'un sommet donné. La longueur du gradient du graphe est donc une mesure de régularité locale. Nous introduisons la variation quadratique d'un signal sur graphe comme mesure de régularité globale.

Définition 2.19. *Soit $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ un graphe non dirigé et $\nu : \mathcal{E} \rightarrow [0, \infty)$ une mesure positive définie sur l'ensemble des arêtes $\mathcal{E}$. La variation quadratique d'un signal sur graphe $f : \mathcal{V} \rightarrow \mathbb{C}$ sur $\mathcal{G}$ est :*

$$\mathcal{TV}(\mathbf{f}) = \frac{1}{2} \sum_{(x, y) \in \mathcal{E}} \nu(x, y) |f(x) - f(y)|^2.$$

Comme nous pouvons le remarquer, nous avons $\|\nabla \mathbf{f}(x)\|_2^2 = \mathcal{TV}(\mathbf{f})$.

Proposition 2.5 ([14]). *Soit $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ un graphe non dirigé muni d'une fonction de poids $w : \mathcal{V} \times \mathcal{V} \rightarrow \mathbb{R}^+$. La variation quadratique associée d'un signal sur graphe $f : \mathcal{V} \rightarrow \mathbb{C}$ sur $\mathcal{G}$ est :*

$$\begin{aligned}\mathcal{TV}_{\mathbf{L}}(f) &= \frac{1}{2} \sum_{x,y \in \mathcal{E}} w(x,y) |f(x) - f(y)|^2, \\ &= \langle \mathbf{f}, \mathbf{L}\mathbf{f} \rangle.\end{aligned}$$

Comme le montre la proposition 2.5, la variation quadratique permettant de mesurer la variation globale d'un signal sur graphe est une forme quadratique faisant intervenir le laplacien combinatoire $\mathbf{L}$. Une seconde variation quadratique par rapport à une autre pondération a également été proposée.

Proposition 2.6 ([14]). *Soit $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ un graphe non dirigé muni d'une fonction de poids $w : \mathcal{V} \times \mathcal{V} \rightarrow \mathbb{R}^+$. La variation quadratique normalisée associée d'un signal sur graphe $f : \mathcal{V} \rightarrow \mathbb{C}$ sur $\mathcal{G}$ est :*

$$\begin{aligned}\mathcal{TV}_{\mathcal{L}}(f) &= \frac{1}{2} \sum_{x,y \in \mathcal{E}} w(x,y) \left| \frac{f(x)}{\sqrt{d_x}} - \frac{f(y)}{\sqrt{d_y}} \right|^2, \\ &= \langle \mathbf{f}, \mathcal{L}\mathbf{f} \rangle.\end{aligned}$$

Cette seconde variation quadratique est une forme quadratique faisant intervenir le laplacien normalisé $\mathcal{L}$. Une nouvelle variation quadratique associée à une pondération différente est également proposée.

Proposition 2.7 ([25]). *Soit $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ un graphe non dirigé muni d'une fonction de poids $w : \mathcal{V} \times \mathcal{V} \rightarrow \mathbb{R}^+$. Soit $\mathcal{X}$ une marche aléatoire ergodique sur $\mathcal{G}$ munie de sa matrice de transition $\mathbf{P}$ avec distribution stationnaire π . La variation quadratique associée à la matrice de transition $\mathbf{P}$ d'un signal sur graphe $f : \mathcal{V} \rightarrow \mathbb{C}$ sur $\mathcal{G}$ est :*

$$\begin{aligned}\mathcal{TV}_{\mathbf{L}_{\text{RW}}}(\mathbf{f}) &= \frac{1}{2} \sum_{(x,y) \in \mathcal{E}} p(x,y) \pi(x) |f(x) - f(y)|^2, \\ &= \langle \mathbf{f}, \mathbf{L}_{\text{RW}}\mathbf{f} \rangle_{\pi}.\end{aligned}$$

Cette dernière variation quadratique $\mathcal{TV}_{\mathbf{L}_{\text{RW}}}$ admet une forme quadratique associée au produit scalaire de l'espace $\ell^2(\mathcal{V}, \pi)$ et fait intervenir le laplacien marche aléatoire $\mathbf{L}_{\text{RW}}$. Les variations quadratiques $\mathcal{TV}_{\mathbf{L}}$ et $\mathcal{TV}_{\mathcal{L}}$ admettent toutes deux des formes quadratiques par rapport à l'espace $\ell^2(\mathcal{V})$.

Nous introduisons maintenant le quotient de Rayleigh comme fonctionnelle de la variation quadratique.

Définition 2.20. *Soit $\mathbf{H}$ un espace de Hilbert de dimension $\dim \mathbf{H} = N < \infty$ et $T \in \mathcal{L}(\mathbf{H}, \mathbf{H})$ un opérateur auto-adjoint. Le quotient de Rayleigh de T est :*

$$\mathcal{R}_T(x) = \frac{\langle x, Tx \rangle}{\langle x, x \rangle}, \quad \forall x \in \mathbf{H}.$$

Par la définition du quotient de Rayleigh 2.20, nous pouvons introduire les quotients de Rayleigh associés aux variations quadratiques $\mathcal{TV}_{\mathbf{L}}, \mathcal{TV}_{\mathcal{L}}, \mathcal{TV}_{\mathbf{L}_{\text{RW}}}$ de la manière suivante :

$$\mathcal{R}_{\mathbf{L}}(\mathbf{f}) = \frac{\mathcal{TV}_{\mathbf{L}}(\mathbf{f})}{\|\mathbf{f}\|^2}, \quad \mathcal{R}_{\mathcal{L}}(\mathbf{f}) = \frac{\mathcal{TV}_{\mathcal{L}}(\mathbf{f})}{\|\mathbf{f}\|^2}, \quad \mathcal{R}_{\mathbf{L}_{\text{RW}}}(\mathbf{f}) = \frac{\mathcal{TV}_{\mathbf{L}_{\text{RW}}}(\mathbf{f})}{\|\mathbf{f}\|_{\pi}^2}.$$

7.2 Analyse fréquentielle sur graphe non dirigé

Après avoir introduit différentes variations quadratiques sur graphe non dirigé faisant intervenir les différents laplaciens ainsi que leurs quotients de Rayleigh associés, nous sommes en mesure d'analyser la variation des vecteurs propres des bases de Fourier obtenues par la diagonalisation du laplacien combinatoire ou de ses variantes.

Théorème 2.3 (Théorème min-max de Courant Fisher). *Soient $\mathbf{M}$ une matrice symétrique, $\{\lambda_i\}_{i=0,\dots,N-1}$ l'ensemble de ses valeurs propres rangées par ordre croissant et $\{\mathbf{v}_i\}_{i=0,\dots,N-1}$ une famille orthonormée de vecteurs propres associés. Par le principe du Min-Max, on a :*

$$\lambda_i = \inf_{\mathbf{v} \perp \text{vect}(\mathbf{v}_0, \dots, \mathbf{v}_{i-1})} \mathcal{R}_{\mathbf{M}}(\mathbf{v}).$$

Le théorème 2.3 permet ainsi la caractérisation variationnelle des valeurs propres d'une matrice symétrique.

Proposition 2.8. *Soit $\mathbf{L}$ le laplacien combinatoire défini sur un graphe non dirigé $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ de cardinalité $|\mathcal{V}| = N$. Soit $\phi \in \mathbb{R}^N$ un vecteur propre quelconque de $\mathbf{L}$ associé à la valeur propre réelle θ . Le quotient de Rayleigh associé est :*

$$\mathcal{R}_{\mathbf{L}}(\phi) = \theta.$$

La variation d'un vecteur propre ϕ du laplacien $\mathbf{L}$ mesurée par son quotient de Rayleigh associé $\mathcal{R}_{\mathbf{L}}(\phi)$ est directement égale à sa valeur propre correspondante associée θ . Par analogie à l'analyse de Fourier classique, ϕ correspond au mode de type Fourier sur le graphe $\mathcal{G}$ et θ à la notion de fréquence associée. De manière similaire, nous pouvons également caractériser de manière variationnelle les vecteurs propres de $\mathcal{L}$ et $\mathbf{L}_{\text{RW}}$.

Proposition 2.9. *Soit $\mathcal{L}$ le laplacien normalisé défini sur un graphe non dirigé $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ de cardinalité $|\mathcal{V}| = N$. Soit $\varphi \in \mathbb{R}^N$ un vecteur propre quelconque de $\mathcal{L}$ associé à la valeur propre réelle ϑ . Le quotient de Rayleigh associé est :*

$$\mathcal{R}_{\mathcal{L}}(\varphi) = \vartheta.$$

Proposition 2.10. *Soit $\mathbf{L}_{\text{RW}}$ le laplacien combinatoire défini sur un graphe non dirigé $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ de cardinalité $|\mathcal{V}| = N$. Soit $\chi \in \mathbb{R}^N$ un vecteur propre quelconque de $\mathbf{L}_{\text{RW}}$ associé à la valeur propre réelle ϑ . Le quotient de Rayleigh associé est :*

$$\mathcal{R}_{\mathbf{L}_{\text{RW}}}(\chi) = \vartheta.$$

8 Analyse multirésolution sur graphes non dirigés

Une multitude d'approches multi-échelles sur graphes non dirigés ont émergé ces dernières années. Ces constructions incluent des ondelettes sur graphes non pondérés [28] qui définissent les ondelettes en termes de diffusion autour de chaque sommet, des ondelettes définies par lifting [52], des ondelettes de type Haar sur graphes [29, 53, 34, 35], des ondelettes sur graphes construites à partir du concept d'auto-encodeurs [33] ou récemment la construction d'ondelettes sur graphes à partir de forêts aléatoires [37].

Dans cette section, nous présentons la construction d'ondelettes spectrales sur graphes proposée par Hammond et al. [32] et inspirée des trames de diffusion polynomiales de Maggioni et Mhaskar [31]. Il s'agit d'une construction d'ondelettes et de fonctions d'échelle par filtrage. Tout d'abord, nous rappelons les éléments nécessaires de la transformée en ondelettes continue dans le cas du traitement de signal classique. Ces éléments motiveront la construction d'ondelettes spectrales sur graphes.

8.1 La transformée en ondelettes continue classique

Une ondelette ψ est une fonction de $L^2(\mathbb{R})$ intégrable de moyenne nulle qui oscille localement :

$$\int_{-\infty}^{\infty} \psi(t) dt = \hat{\psi}(0) = 0.$$

qui peut être dilatée/contractée (facteur d'échelle $e \in \mathbb{R}_+^*$) et translatée (paramètre de localisation $u \in \mathbb{R}$) :

$$\psi_{u,e}(t) = \frac{1}{e} \psi\left(\frac{t-u}{e}\right).$$

La transformée en ondelettes continue du signal f (avec une ondelette dilatée/contractée de e et translatée de u) est donnée par :

$$Wf(u, e) = \int_{-\infty}^{\infty} \frac{1}{e} \psi^*\left(\frac{t-u}{e}\right) f(t) dt = \langle \psi_{u,e}, f \rangle$$

Le signal f peut-être reconstruit à partir de $Wf(u, e)$ en employant la relation suivante :

$$f(t) = \frac{1}{C_\psi} \int_0^{+\infty} \int_{-\infty}^{+\infty} Wf(u, e) \psi\left(\frac{t-u}{e}\right) du \frac{de}{e^2}, \quad (2.5)$$

avec

$$C_\psi = \int_0^{+\infty} \frac{|\hat{\psi}(\omega)|^2}{\omega} d\omega.$$

Cette constante C_ψ doit être finie. C'est ce que l'on appelle la "condition d'admissibilité". Cette condition implique que l'ondelette soit de moyenne

nulle. La représentation intégrale permettant de reconstruire f (2.5) fait intervenir des ondelettes correspondant à toutes les localisations temporelles et à tous les facteurs d'échelle.

Cette méthode de construction de la transformée en ondelettes se déroule en construisant les ondelettes directement dans le domaine du signal par dilatation/contraction et translation. Cependant, l'application directe de cette construction aux graphes est problématique. En effet, pour une fonction donnée définie sur les sommets d'un graphe, il n'est pas évident de savoir comment définir $\psi(ex)$ car x est un sommet du graphe et qu'il n'y a pas d'interprétation de ex pour un scalaire réel e . Par ailleurs, la translation $t - u$ pose également problème. L'approche employée pour surmonter cet obstacle est de travailler dans le domaine de Fourier du signal. En effet, dans la transformée en ondelettes continue classique, la dilatation/contraction peut s'effectuer dans le domaine de Fourier. Ainsi, l'expression résultante fournit une base pour un procédé analogue sur graphes.

Pour un facteur d'échelle e fixe, la transformée en ondelettes peut être interprétée comme un opérateur prenant la fonction f et retournant la fonction $T^e f(u) = W_f(u, e)$. En d'autres termes, nous considérons le paramètre de translation comme la variable indépendante de la fonction retournée par l'opérateur T^e . Si nous posons :

$$\bar{\psi}_e(x) = \frac{1}{e} \psi^* \left(-\frac{x}{e} \right). \quad (2.6)$$

Nous voyons que cet opérateur est obtenu par convolution, c.-à-d. :

$$\begin{aligned} T^e f(u) &= \int_{-\infty}^{\infty} \frac{1}{e} \psi^* \left(\frac{x-u}{e} \right) f(x) dx, \\ &= \int_{-\infty}^{\infty} \bar{\psi}_e(u-x) f(x) dx, \\ &= (\bar{\psi}_e \star f)(u). \end{aligned}$$

En prenant la transformée de Fourier et en appliquant le théorème de convolution, on obtient :

$$\widehat{T^e f}(\omega) = \hat{\bar{\psi}}_e(\omega) \hat{f}(\omega).$$

En utilisant les propriétés de facteurs d'échelle de la transformée de Fourier et (2.6), nous avons :

$$\hat{\bar{\psi}}_e(\omega) = \hat{\psi}^*(e\omega).$$

En combinant ceux-ci et en inversant la transformée nous obtenons :

$$T^e f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \exp(i\omega x) \hat{\psi}(e\omega) \hat{f}(\omega) d\omega. \quad (2.7)$$

Dans l'expression de la transformée en ondelettes (2.7), le facteur d'échelle e n'apparaît que dans l'argument de $\hat{\psi}(e\omega)$ montrant que l'opération de dilatation/contraction peut être complètement effectuée dans le domaine de Fourier. L'expression (2.7) montre que la transformée de Fourier à chaque échelle e peut être considérée comme un opérateur multiplicatif dans le domaine de Fourier, déterminé par des filtres qui sont obtenus à partir de la dilatation/contraction d'un filtre de base $\hat{\psi}^*(\omega)$. En écrivant $\delta_u(x) = \delta(x-u)$, on a :

$$(T^e \delta_u)(x) = \frac{1}{e} \psi^* \left(\frac{u-x}{e} \right).$$

8.2 Transformée en ondelettes spectrale

L'équivalent de la transformée de Fourier pour des fonctions définies sur les sommets d'un graphe ayant été définie, il est maintenant possible de définir la transformée en ondelettes spectrale sur graphes.

8.2.1 Ondelettes sur graphes

Soit $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, un graphe simple non dirigé de cardinalité $|\mathcal{V}| = N$. Nous définissons des ondelettes sur graphe de la manière suivante.

Définition 2.21 ([32]). *Étant donnée une fonction g se comportant comme un filtre passe-bande, c.-à-d. $g(0) = 0$ et $\lim_{x \rightarrow \infty} g(x) = 0$, on définit un opérateur d'ondelette de dilatation s , noté $\mathbf{T}_g^s$ comme une fonction du laplacien $\mathbf{L}$ défini sur $\mathcal{G}$, c'est-à-dire :*

$$\mathbf{T}_g^s = g(s\mathbf{L}) = g(s\mathbf{\Xi}\mathbf{\Theta}\mathbf{\Xi}^T) = \mathbf{\Xi}g(s\mathbf{\Theta})\mathbf{\Xi}^T.$$

Une ondelette spectrale de dilatation s et centrée en un sommet v_n est obtenue en appliquant le filtre $\mathbf{T}_g^s$ à un δ -Kronecker au sommet v_n , c'est-à-dire :

$$\psi_{n,s} = \mathbf{T}_g^s \delta_n.$$

Enfin, étant donné un signal sur graphe $\mathbf{f}$, les coefficients d'ondelettes associés sont définis de la manière suivante :

$$W_f(n, s) = \langle \psi_{n,s}, \mathbf{f} \rangle = (\mathbf{T}_g^s \mathbf{f})(n).$$

La transformée en ondelettes spectrale est obtenue à partir des opérateurs d'ondelettes qui sont des fonctions du laplacien obtenues en utilisant la représentation spectrale de l'opérateur (qui est l'équivalent de la transformée de Fourier sur graphe). Le facteur d'échelle s agit sur les valeurs propres, c'est-à-dire les fréquences de Fourier sur graphe.

8.2.2 Fonctions d'échelle sur graphes

Nous introduisons l'opérateur de fonction d'échelle qui est analogue à l'opérateur d'ondelettes défini précédemment.

Définition 2.22 ([32, 31]). *Étant donnée une fonction h se comportant comme un filtre passe bas, c.-à-d. $h(0) > 0$ et $\lim_{x \rightarrow \infty} h(x) = 0$, on définit un opérateur de fonction d'échelle de dilatation s , noté T_h^s comme une fonction du laplacien $\mathbf{L}$ définie sur $\mathcal{G}$, c'est-à-dire :*

$$\mathbf{T}_h^s = h(s\mathbf{L}) = \mathbf{\Xi}h(s\mathbf{\Theta})\mathbf{\Xi}^\top.$$

Une fonction d'échelle spectrale de dilatation s et centrée en un sommet v_n est obtenue en appliquant le filtre $\mathbf{T}_g^s$ à un δ -Kronecker au sommet v_n , c'est-à-dire :

$$\phi_{n,s} = \mathbf{T}_h^s \delta_n.$$

Enfin, étant donné un signal sur graphe $\mathbf{f}$, les coefficients d'échelle associés sont définis de la manière suivante :

$$W_f(n, s) = \langle \phi_{n,s}, \mathbf{f} \rangle = (\mathbf{T}_h^s \mathbf{f})(n).$$

L'introduction des fonctions d'échelle permet d'assurer une restauration stable du signal sur graphe $\mathbf{f}$ à partir des coefficients d'ondelettes lorsque le facteur d'échelle s est échantillonné à un nombre discret de valeurs t_j .

8.2.3 Transformée en ondelettes sur graphes

Ayant défini les opérateurs d'ondelettes et de fonctions d'échelle sur graphes, nous sommes en mesure d'introduire la notion de transformée en ondelettes spectrale sur graphes.

Définition 2.23 ([32]). *Étant donné :*

- *un ensemble d'échelles croissantes $\{t_j\}_{j=1}^J$.*
- *une fonction passe bande g .*
- *une fonction passe bas h .*
- *un signal sur graphe $\mathbf{f}$.*

la transformée en ondelettes spectrale de $\mathbf{f}$ est caractérisée par la matrice $\mathbf{K} \in \mathbb{R}^{N(J+1) \times N}$ définie de la manière suivante :

$$\mathbf{K}\mathbf{f} = [(\mathbf{T}_g^{t_1} \mathbf{f})^\top, (\mathbf{T}_g^{t_2} \mathbf{f})^\top, \dots, (\mathbf{T}_g^{t_J} \mathbf{f})^\top, (\mathbf{T}_h^{t_J} \mathbf{f})^\top]^\top.$$

Lorsque les filtres sont bien choisis, c.-à-d. de manière à définir une transformée inversible, l'ensemble des fonctions d'échelles et ondelettes constitue une trame.

9 Conclusion

Dans ce chapitre, nous avons rappelé les aspects fondamentaux de l'analyse harmonique pour des signaux sur graphes non dirigés. Nous avons rappelé que les vecteurs propres des opérateurs laplaciens combinatoire et normalisé forment tous deux des bases de type Fourier orthonormales pour l'espace de Hilbert $\ell^2(\mathcal{V})$. Nous avons également rappelé que les vecteurs propres du laplacien marche aléatoire comme base de type Fourier qui a la particularité d'être orthonormale sur l'espace de Hilbert $\ell^2(\mathcal{V}, \pi)$. Ainsi, il est possible de construire plusieurs transformées de Fourier sur graphes non dirigés à partir de la décomposition spectrale des opérateurs laplaciens sur graphes non dirigés. Nous avons également rappelé l'aspect fréquentiel des bases de Fourier sur graphes dirigés en étudiant leurs variations sur graphes. Enfin, nous avons présenté une construction d'ondelettes et de fonctions d'échelles par filtrage introduite par [32].

L'analyse harmonique sur graphes non dirigés étant bien établie, est-il possible d'étendre ce cadre pour l'analyse harmonique de signaux sur graphes dirigés ? Il s'avère que l'extension de ce cadre n'est pas directe. Un graphe dirigé est naturellement représenté par une matrice d'adjacence non symétrique. Il est toujours possible de définir naïvement un laplacien de graphe dans le cas dirigé, mais les propriétés spectrales associées au laplacien sur graphe non dirigé tel que l'obtention d'une base orthonormale et de valeurs propres positives ou nulles réelles ne sont plus vérifiées dans le cas des graphes dirigés.

Étendre l'analyse harmonique aux signaux sur graphes dirigés est le but de cette thèse. Récemment, des approches ont été proposées pour étendre l'analyse harmonique au cadre dirigé. Dans la partie II, nous présenterons l'état de l'art de l'analyse de Fourier sur graphes dirigés à travers deux approches dominantes. La partie III sera dédiée à notre contribution à la fois théorique et expérimentale à l'analyse harmonique pour signaux sur graphes dirigés.

Deuxième partie

Analyse harmonique sur
graphes dirigés

Chapitre 3

État de l'art sur les propositions existantes en analyse de Fourier sur graphes dirigés

Dans ce chapitre, nous abordons les récentes avancées du problème de construction de bases de type Fourier pour des fonctions définies sur les sommets d'un graphe dirigé. Contrairement au cas non dirigé où la construction de base de Fourier sur graphes à partir des vecteurs propres du laplacien de graphe ou ses équivalents est bien établie, le cas dirigé est plus discuté. Durant ces dernières années, un faible nombre de bases de Fourier ont été proposées. Nous pouvons regrouper ces constructions en deux types : celles construites à partir d'un opérateur de référence sur graphes dirigés et celles construites à partir de nouvelles mesures de variation sur graphes. Tout d'abord, nous présentons les notions essentielles relatives aux graphes dirigés. Dans un second temps, nous discutons du premier cas et introduisons les constructions et l'analyse de Fourier basées sur la matrice d'adjacence du graphe dirigé. Enfin, nous abordons les méthodes de construction associées au second cas.

Sommaire

1	Notions sur la théorie des graphes dirigés	37
2	L'analyse de Fourier basée sur la matrice d'adjacence	38
2.1	L'analyse de Fourier selon Sandryhaila et Moura	39
2.2	L'analyse de Fourier selon Mhaskar	41
3	L'analyse de Fourier via une notion de variation dirigée.	42
3.1	L'analyse de Fourier selon Sardellitti et coll.	42
3.2	L'analyse de Fourier selon Shafipour et coll.	43
4	Discussion	44

1 Notions sur la théorie des graphes dirigés

De manière analogue à la section 2 du chapitre 2, nous introduisons les notions fondamentales relatives aux graphes dirigés.

Définition 3.1. *Un graphe simple (c.-à.d. sans arêtes multiples) dirigé et non pondéré est un couple $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ où $\mathcal{V}$ constitue l'ensemble des sommets et $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ l'ensemble des arcs, chaque arc étant une paire de sommets non nécessairement distincts. Si u et v sont des sommets, nous représentons un arc reliant les sommets u et v par $e = (u, v)$ où u est appelé l'origine de (u, v) et v est appelé l'extrémité.*

Comme dans le cas des graphes non dirigés, nous considérons exclusivement le cas des graphes finis. Nous définissons $|\mathcal{V}|$, la cardinalité de l'ensemble des sommets $\mathcal{V}$, c'est-à-dire le nombre d'éléments dans $\mathcal{V}$. Le graphe $\mathcal{G}$ peut être représenté à travers sa matrice d'adjacence que nous définissons de la manière suivante.

Définition 3.2 (Représentation algébrique d'un graphe dirigé et non pondéré). *Soit $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ un graphe simple non dirigé et non pondéré où $\mathcal{V} = (v_1, \dots, v_N)$ et de cardinalité $|\mathcal{V}| = N$. La matrice d'adjacence $\mathbf{A} = \{a_{ij}\}_{1 \leq i, j \leq N} \in \mathbb{R}^{N \times N}$ se rapportant à cet ensemble de sommets est la matrice $N \times N$ booléenne avec :*

$$a_{ij} = \begin{cases} 1 & \text{si } (v_i, v_j) \in \mathcal{E} \\ 0 & \text{sinon.} \end{cases}$$

Contrairement aux graphes non dirigés, la condition de symétrie eq. (2.1) n'est plus vérifiée dans le cas dirigé.

Plus généralement, nous pouvons considérer le cas des graphes dirigés et pondérés. Nous définissons un graphe dirigé et pondéré de la manière suivante.

Définition 3.3. Soit $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ un graphe dirigé et non pondéré et $w : \mathcal{V} \times \mathcal{V} \rightarrow \mathbb{R}$ est une application qui associe une valeur $w(v_i, v_j)$ à chaque arc (v_i, v_j) . L'application w est appelé fonction de poids. Le triplet $\mathcal{G} = (\mathcal{V}, \mathcal{E}, w)$ est un graphe dirigé et pondéré.

Un graphe dirigé et pondéré $\mathcal{G} = (\mathcal{V}, \mathcal{E}, w)$ peut être entièrement représenté par sa matrice d'adjacence $\mathbf{W} = \{w_{ij}\}_{1 \leq i, j \leq N} \in \mathbb{R}^{N \times N}$ où $w_{ij} = w(v_i, v_j)$. Par convention, si $w_{ij} = 0$, pour tout $i, j \in \{1, \dots, N\}$, alors il n'existe pas d'arc entre les sommets v_i et v_j . Nous définissons respectivement le degré sortant et le degré entrant d'un sommet $v_i \in \mathcal{V}$ par $d_i^+ = \sum_{j=1}^N w_{ij}$ et $d_i^- = \sum_{j=1}^N w_{ji}$. Nous utiliserons uniquement le degré sortant que nous noterons d_i pour tout $v_i \in \mathcal{V}$.

Nous définissons la notion de chemin, concept analogue à la notion de chaîne (définition 2.4) dans un graphe dirigé.

Définition 3.4. Soit $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ un graphe dirigé. Un chemin reliant un sommet v_i à un sommet v_j est défini par une suite finie d'arcs consécutifs, reliant v_i à v_j .

Nous définissons la connexité pour un graphe dirigé de la manière suivante.

Définition 3.5. Un graphe dirigé $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ est connexe, si son graphe non dirigé correspondant est connexe.

Nous définissons la notion de graphe fortement connexe de la manière suivante.

Définition 3.6. Un graphe dirigé $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ est fortement connexe si chaque sommet est accessible à partir de n'importe quel autre. Autrement dit, si pour tout couple de sommets distincts $(v_i, v_j) \in \mathcal{V} \times \mathcal{V}$ il existe un chemin de v_i à v_j et un chemin de v_j à v_i .

2 L'analyse de Fourier basée sur la matrice d'adjacence

L'analyse de Fourier construite à partir de la matrice d'adjacence d'un graphe dirigé a été introduite tout d'abord par Sandryhaila et Moura [20] dans la cadre du "traitement de signal sur graphes" [16, 44]. Cette analyse de Fourier est effectuée à partir de la diagonalisation de la matrice d'adjacence. Récemment une alternative a été proposée par Mhaskar fondée sur la décomposition en valeur singulières de la matrice d'adjacence [21]. Dans un premier temps, nous présentons les aspects principaux de l'analyse de Fourier selon Sandryhaila et Moura.

2.1 L'analyse de Fourier selon Sandryhaila et Moura

Soit $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ un graphe dirigé et non pondéré de cardinalité $|\mathcal{V}| = N$ caractérisé par sa matrice d'adjacence $\mathbf{A} \in \mathbb{R}^{N \times N}$. Dans leur cadre théorique, la matrice d'adjacence est l'analogue sur graphe de l'opérateur de décalage dans le traitement de signal classique et est appelée opérateur de décalage sur graphe (ou *graph shift* en anglais).

2.1.1 La transformée de Fourier sur graphe

Soit $\mathbf{A}$ la matrice d'adjacence sur $\mathcal{G}$. Dans ce cadre théorique, la matrice d'adjacence $\mathbf{A}$ est considérée comme l'opérateur de référence sur graphe dirigé. La matrice d'adjacence $\mathbf{A}$ admet une décomposition de Jordan, c'est-à-dire :

$$\mathbf{A} = \mathbf{V}\mathbf{J}\mathbf{V}^{-1}.$$

où $\mathbf{V}$ est la matrice formée des vecteurs de Jordan et $\mathbf{J}$ est la forme réduite de Jordan de $\mathbf{A}$.

Étant donné un signal sur graphe $\mathbf{s}$, sa *transformée de Fourier* notée $\hat{\mathbf{s}} = [\hat{s}_1, \dots, \hat{s}_N]$ est :

$$\hat{\mathbf{s}} = \mathbf{V}^{-1}\mathbf{s}.$$

Les valeurs $\{\hat{s}_j\}_{j=1}^N$ de la transformée de Fourier du signal sur graphe $\mathbf{s}$ caractérisent le contenu en fréquence du signal $\mathbf{s}$. La *transformée de Fourier inverse* sur graphe est donnée par :

$$\mathbf{s} = \mathbf{V}\hat{\mathbf{s}}.$$

Elle reconstruit le signal original à partir de ses contenus en fréquence en construisant une combinaison linéaire des composantes en fréquence pondérées par les coefficients de la transformée de Fourier du signal.

Si la matrice d'adjacence $\mathbf{A}$ est diagonalisable et si tous ses sous-espaces propres sont de dimension 1, alors la base de Fourier est unique (à un scalaire multiplicatif près pour chaque vecteur propre). Dans le cas contraire, c'est-à-dire s'il existe un espace propre de dimension supérieure à 1, il n'y a plus unicité de la base de Fourier.

2.1.2 Analyse de Fourier sur graphe

Sandryhaila et Moura déterminent une interprétation fréquentielle des vecteurs de Jordan de la matrice d'adjacence $\mathbf{A}$ en étudiant leurs variations à travers l'introduction d'une notion de régularité de signal sur graphe.

2.1.2.1 Régularité de signaux sur graphes Le comportement d'un signal sur graphe dirigé peut être analysé en étudiant sa régularité. Dans ce cadre, la régularité est analysée à partir de la mesure de variation suivante.

Définition 3.7 ([44, 20]). Soit $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ un graphe dirigé et non pondéré de cardinalité $|\mathcal{V}| = N$ caractérisé par sa matrice d'adjacence $\mathbf{A} \in \mathbb{R}^{N \times N}$. La variation totale sur un graphe d'un signal sur graphe $\mathbf{s}$ est définie de la manière suivante :

$$\text{TV}_{\mathcal{G}}(\mathbf{s}) = \|\mathbf{s} - \mathbf{A}^{norm} \mathbf{s}\|_1.$$

où la version normalisée de la matrice d'adjacence, $\mathbf{A}^{norm}$ est définie de la manière suivante :

$$\mathbf{A}^{norm} = \frac{1}{|\lambda_{\max}|} \mathbf{A}.$$

où $\lambda_{\max}$ est la valeur propre maximale de $\mathbf{A}$.

2.1.2.2 Analyse fréquentielle sur graphe La proposition suivante établit un lien entre la régularité des vecteurs propres de la matrice d'adjacence $\mathbf{A}$ et les valeurs propres associées.

Proposition 3.1. Soit $\mathbf{v}$ un vecteur propre de la matrice d'adjacence $\mathbf{A}$ associé à la valeur propre λ . La variation totale de $\mathbf{v}$ par rapport à la mesure de régularité introduite à la définition 3.7 est :

$$\text{TV}(\mathbf{v}) = \left| 1 - \frac{\lambda}{|\lambda_{\max}|} \right| \|\mathbf{v}\|_1. \quad (3.1)$$

Dans ce cadre, la variation des vecteurs propres est directement reliée aux valeurs propres associées. Lorsque $\mathbf{A}$ n'est pas diagonalisable, Sandryhaila et coll. proposent de continuer à utiliser la variation déf. 3.7 pour comparer les variations des vecteurs propres généralisés formant la matrice $\mathbf{V}$ de la forme de Jordan. Toutefois, le lien entre la variation d'un vecteur propre généralisé $\mathbf{v}$ et la valeur propre correspondante (3.1) n'est valable que si $\mathbf{v}$ est un vecteur propre (i.e. non généralisé).

Ordre des fréquences sur graphes A partir de cette interprétation fréquentielle des vecteurs propres de la matrice d'adjacence d'un graphe dirigé, Sandryhaila et Moura déterminent un ordre des fréquences sur graphe dirigé. Le théorème suivant établit l'ordre relatif de deux fréquences sur graphe réelles distinctes.

Théorème 3.1 ([20]). Soient deux valeurs propres distinctes $\lambda_m, \lambda_n \in \mathbb{R}$ de la matrice d'adjacence $\mathbf{A}$ avec les vecteurs propres correspondants $\mathbf{v}_m$ et $\mathbf{v}_n$. Si les valeurs propres sont ordonnées comme :

$$\lambda_m < \lambda_n,$$

alors la variation de leur vecteurs propres satisfait la relation suivante :

$$\text{TV}_{\mathcal{G}}(\mathbf{v}_m) > \text{TV}_{\mathcal{G}}(\mathbf{v}_n).$$

Par le théorème 3.1, si un graphe admet un spectre réel et ses valeurs propres sont ordonnées de la manière suivante :

$$\lambda_0 > \lambda_1 > \dots > \lambda_{M-1},$$

alors λ_0 correspond à la plus basse fréquence et λ_{M-1} à la plus haute. Le théorème suivant étend le théorème 3.1 et établit un ordre relatif entre deux fréquences distinctes correspondant à des valeurs propres complexes.

Théorème 3.2 ([20]). *Soient deux valeurs propres distincts $\lambda_m, \lambda_n \in \mathbb{C}$ de la matrice d'adjacence $\mathbf{A}$. Soient $\mathbf{v}_m$ et $\mathbf{v}_n$ les vecteurs propres correspondants. Les variations totales de ces vecteurs propres satisfont :*

$$\text{TV}_{\mathcal{G}}(\mathbf{v}_m) < \text{TV}_{\mathcal{G}}(\mathbf{v}_n),$$

si la valeur propre λ_m est plus proche de $|\lambda_{max}|$ sur le plan complexe que la valeur propre λ_n .

Le théorème 3.2 montre donc que les fréquences d'un graphe dirigé dont la matrice d'adjacence admet un spectre complexe sont ordonnées en fonction de leur distance par rapport à λ_{max} .

2.2 L'analyse de Fourier selon Mhaskar

Plus récemment, Mhaskar propose dans [21] une analyse en fréquence pour des fonctions définies sur graphes dirigés construite à partir de la matrice d'adjacence d'un graphe dirigé et pondéré.

2.2.1 Transformée de Fourier sur graphe

Soit $\mathcal{G} = (\mathcal{V}, \mathcal{E}, w)$ un graphe dirigé et pondéré représenté par sa matrice d'adjacence $\mathbf{W} \in \mathbb{R}^{N \times N}$. La matrice d'adjacence admet la décomposition en valeurs singulières suivante :

$$\mathbf{W} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^*,$$

où $\mathbf{U}, \mathbf{V} \in \mathbb{C}^{N \times N}$ sont des matrices unitaires (orthogonales dans $\mathbb{R}$) et $\mathbf{\Sigma} = \text{diag}\{\sigma_1, \dots, \sigma_N\}$ est la matrice diagonale des valeurs singulières.

Dans ce cadre, Mhaskar choisit la matrice de vecteurs singuliers à gauche $\mathbf{U}$ comme sa base de Fourier sur graphe dirigé. Étant donné un signal sur graphe $\mathbf{s}$, sa *transformée de Fourier* est :

$$\hat{\mathbf{s}} = \mathbf{U}^* \mathbf{s},$$

où $\mathbf{U}^*$ est la matrice de transformée de Fourier sur graphe. La *transformée de Fourier inverse* sur graphe est donnée par :

$$\mathbf{s} = \mathbf{U} \hat{\mathbf{s}}.$$

2.2.1.1 Analyse de Fourier sur graphe Pour toute matrice d'adjacence, nous avons la décomposition polaire suivante :

$$\mathbf{W} = \mathbf{R}\mathbf{Z},$$

où $\mathbf{R}$ est une matrice symétrique semi-définie positive de même rang que $\mathbf{W}$ et $\mathbf{Z}$ une matrice réelle unitaire. La décomposition en valeurs singulières de $\mathbf{W}$ peut également être réécrite comme sa décomposition polaire de la manière suivante :

$$\mathbf{W} = \mathbf{U}\Sigma\mathbf{V}^* = \mathbf{U}\Sigma\mathbf{I}_N\mathbf{V}^* = \mathbf{U}\Sigma\mathbf{U}^*\mathbf{U}\mathbf{V}^* = \mathbf{R}\mathbf{Z}, \quad (3.2)$$

avec $\mathbf{R} = \mathbf{U}\Sigma\mathbf{U}^*$ et $\mathbf{Z} = \mathbf{U}\mathbf{V}^*$.

Comme nous avons montré à l'équation (3.2), la matrice $\mathbf{R} = \mathbf{U}\Sigma\mathbf{U}^*$ est une matrice symétrique semi-définie positive dont les valeurs propres $\{\sigma_1, \dots, \sigma_N\}$ sont les valeurs singulières de $\mathbf{W}$ et $\mathbf{U} = [\mathbf{u}_1, \dots, \mathbf{u}_N]$ la matrice unitaire des vecteurs singuliers à gauche de $\mathbf{W}$. Nous avons donc, d'après le théorème 2.3 :

$$\mathcal{R}_{\mathbf{R}}(\mathbf{u}_k) = \frac{\langle \mathbf{u}_k, \mathbf{R}\mathbf{u}_k \rangle}{\langle \mathbf{u}_k, \mathbf{u}_k \rangle} = \sigma_k, \quad \forall k = 1, \dots, N$$

Par conséquent, dans ce cadre, les vecteurs propres de $\mathbf{R}$ (ou vecteurs singuliers de $\mathbf{W}$) correspondent aux modes de type Fourier sur $\mathcal{G}$ et l'ensemble $\{\sigma_1, \dots, \sigma_N\}$ correspond aux fréquences associées.

À partir de la décomposition polaire de la matrice d'adjacence de $\mathbf{W}$ et l'interprétation fréquentielle des vecteurs singuliers de $\mathbf{W}$ à travers la matrice $\mathbf{R}$, Mhaskar est en mesure d'étendre son cadre d'analyse multirésolution de trames de diffusion polynomiales [31] aux cas dirigé [21].

3 L'analyse de Fourier via une notion de variation dirigée.

Cette section aborde les dernières approches de construction de bases de Fourier sur graphes dirigés nécessitant l'utilisation de méthodes de résolution de problèmes d'optimisation non convexes. Récemment, deux approches ont été introduites : celle de Sardellitti et coll. [23] et celle de Shafipour et coll. [22]. Ces approches sont similaires et se fondent sur la minimisation de l'extension continue du poids de la coupe du graphe [54, 55]. Nous aborderons plus en détail les principaux aspects des deux approches dans les sous-sections suivantes.

3.1 L'analyse de Fourier selon Sardellitti et coll.

Soit $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ un graphe dirigé caractérisé par sa matrice d'adjacence $\mathbf{A}$. Sardellitti et coll. proposent dans [23] de construire une base de Fourier sur

graphes dirigés fondée sur la définition de variation suivante :

$$f(\mathbf{x}) = \sum_{i,j=1}^N a_{ji}[x_i - x_j]_+ = \text{GDV}(\mathbf{x}), \quad (3.3)$$

où $[x]_+$ désigne la partie positive de x : $[x]_+ = \max\{x, 0\}$.

Dans leur cadre, cette fonction est appelée variation dirigée du graphe (*graph directed variation* ou GDV en anglais) qui capture la directionnalité des arcs de $\mathcal{G}$. Formellement cette notion de variation introduite par Sardellitti et coll. représente l'extension convexe continue de Lovász du poids de la coupe du graphe [23, 54, 55] et l'article [23] motive ce choix.

3.1.1 Construction de la base de Fourier sur graphes dirigés

Soit $\mathbf{X} = [\mathbf{x}_1, \dots, \mathbf{x}_N] \in \mathbb{R}^{N \times N}$ l'ensemble des vecteurs de la base recherchée. Selon [23], la base de Fourier sur graphe dirigé est définie comme l'ensemble de vecteurs orthonormaux dont la somme des variations dirigées (3.3) est minimale, c'est-à-dire :

$$\begin{aligned} \min_{\mathbf{X} \in \mathbb{R}^{N \times N}} \quad & \text{GDV}(\mathbf{X}) = \sum_{k=1}^N \text{GDV}(\mathbf{x}_k) \\ \text{s.c.} \quad & \mathbf{X}^\top \mathbf{X} = \mathbf{I}, \quad \mathbf{x}_1 = b\mathbf{1}. \end{aligned}$$

où $\mathbf{x}_1$ est le vecteur constant et de norme unité, c.-à-d. $\|\mathbf{x}_1\|^2 = 1$.

Bien que la fonction objective soit convexe, le problème d'optimisation est non convexe en raison de la contrainte d'orthogonalité. Afin de résoudre ce problème d'optimisation, Sardellitti et coll. suggèrent des stratégies d'optimisation [56, 57] permettant de résoudre le problème non convexe et non différentiable.

3.2 L'analyse de Fourier selon Shafipour et coll.

Shafipour et coll. proposent également de construire une base de Fourier orthonormale sur graphes dirigés [22] fondée sur l'extension de Lovász du poids de la coupe du graphe dirigé (3.3). La différence notable avec l'approche [23] est que les modes de Fourier sont créés pour avoir des fréquences réparties de manière quasi-uniforme dans tout le spectre de fréquences sur graphe et ainsi mieux capturer les basses, moyennes et hautes fréquences.

Formellement, Shafipour et coll. proposent de construire une base orthonormale $\mathbf{U} = [\mathbf{u}_1, \dots, \mathbf{u}_N] \in \mathbb{R}^{N \times N}$ pour les fonctions sur graphes dirigés. Afin de couvrir tout le spectre des variations, le premier vecteur est défini comme celui minimisant (3.3), i.e. le vecteur constant $\mathbf{u}_1 = \mathbf{u}_{\min}$ de norme

$\|\mathbf{u}_1\| = 1$. Le dernier vecteur est $\mathbf{u}_N = \mathbf{u}_{\max}$ où $\mathbf{u}_{\max}$ est défini de la manière suivante :

$$\mathbf{u}_{\max} = \operatorname{argmax}_{\|\mathbf{u}\|=1} \text{GDV}(\mathbf{u}).$$

Cela permet ainsi de couvrir l'ensemble des variations possibles des signaux, depuis le mode de fréquence 0, constant, à celui de plus haute fréquence sur le graphe.

Ensuite, pour déterminer les autres vecteurs de $\mathbf{U}$, ils introduisent la fonction de dispersion spectrale :

$$\delta(\mathbf{U}) = \sum_{i=1}^{N-1} [\text{GDV}(\mathbf{u}_{i+1}) - \text{GDV}(\mathbf{u}_i)]^2.$$

qui mesure le degré d'étalement des fréquences correspondantes sur le spectre des fréquences.

3.2.1 Construction de la base de Fourier sur graphes dirigés

La base de Fourier sur graphes dirigés est construite à partir de la résolution du problème d'optimisation suivant consistant à minimiser la fonction de dispersion spectrale :

$$\begin{aligned} \min_{\mathbf{U}} \quad & \sum_{i=1}^N [\text{GDV}(\mathbf{u}_{i+1}) - \text{GDV}(\mathbf{u}_i)]^2 \\ \text{s.c.} \quad & \mathbf{U}^\top \mathbf{U} = \mathbf{I}_N, \\ & \mathbf{u}_1 = \mathbf{u}_{\min}, \\ & \mathbf{u}_N = \mathbf{u}_{\max}. \end{aligned}$$

Tout comme l'approche [23], le problème d'optimisation est également non convexe, dû aux contraintes d'orthogonalité. Afin de résoudre ce problème d'optimisation, Shafipour et coll. ont également développé des méthodes pour résoudre ces problèmes d'optimisation avec contraintes d'orthogonalité [58, 22].

4 Discussion

Dans ce chapitre, nous avons passé en revue les différentes approches existantes de construction de transformée de Fourier sur graphes dirigés. Ces approches ont été réparties en deux catégories : les approches basées sur la matrice d'adjacence [20, 21] et les approches basées sur la GDV (3.3) [23, 22]. Dans la littérature du traitement de signal sur graphes, les approches [23, 22]

sont considérées comme des méthodes alternatives à la construction d’une transformée de Fourier sur graphes dirigés de Sandryhaila et Moura [20].

Les deux approches [23, 22] montrent leurs limites face à l’approche de Sandryhaila et Moura.

La méthode de Sardellitti et coll. [23] se limite aux signaux réels sur graphes. Contrairement à l’approche de Sandryhaila et Moura s’appliquant sur les signaux complexes sur graphes, leur approche ne parvient pas à déterminer les exponentielles complexes comme base de Fourier dans le cas du graphe cyclique. Sardellitti et coll. admettent par ailleurs que l’approche de Sandryhaila et Moura [20] est plus appropriée pour concevoir une théorie de filtrage de signaux sur graphes.

Dans [22], Shafipour et coll. proposent également la construction d’une base de Fourier sur graphes dirigés en déterminant un ensemble de vecteurs orthonormaux suite à la minimisation d’un problème d’optimisation non convexe. La spécificité de l’approche [22] est que les modes de Fourier obtenus couvrent l’ensemble de la gamme de fréquences et sont répartis aussi uniformément que possible dans le domaine spectral du graphe pour mieux capturer les basses, moyennes et hautes fréquences.

Concernant l’approche de Sandryhaila et Moura [44], nous pouvons relever certains problèmes. La matrice de changement de base n’est généralement pas orthogonale et cela peut poser certains problèmes (notre approche pose les mêmes problèmes et nous en discutons à la section 4). Dans cette approche, les valeurs propres sont complexes. Cela correspond dans leur cadre à une rotation de phase en plus d’une variation d’amplitude. (Nous avons également les mêmes problèmes et nous en discutons en section 5.2).

Dans le chapitre suivant (chapitre 4) nous proposerons une analyse de Fourier basée sur l’opérateur de marche aléatoire sur graphes dirigés. Dans les applications, nous comparerons les performances de notre approche à l’approche de Sandryhaila et Moura, approche alternative la plus appropriée en théorie du signal sur graphes.

Troisième partie

Contributions

Chapitre 4

Analyse harmonique sur graphes dirigés

Dans ce chapitre, nous introduisons une nouvelle analyse harmonique pour des fonctions définies sur les sommets d'un graphe dirigé fortement connexe où l'opérateur de marche aléatoire est la pièce angulaire. Dans un premier temps, nous considérons l'ensemble des vecteurs propres de l'opérateur de marche aléatoire comme une base non orthogonale du type Fourier pour des fonctions sur des graphes dirigés. Nous avons déterminé une interprétation fréquentielle en établissant un lien entre la variation des vecteurs propres de l'opérateur de marche aléatoire obtenue à partir de leur énergie de Dirichlet et la partie réelle de leurs valeurs propres associées. À partir de cette base de Fourier appropriée, nous sommes en mesure d'aller plus loin et de construire des analyses multi-échelles sur des graphes dirigés. Nous proposons dans ce chapitre aussi bien une transformée en ondelettes redondante qu'une transformée en ondelettes décimée en prolongeant le cadre des ondelettes de diffusion de Coifman et Maggioni pour des graphes dirigés. Le développement de notre analyse harmonique sur des graphes dirigés nous amène ainsi à considérer à la fois des problèmes d'apprentissage semi-supervisé et des problèmes de modélisation de signaux sur des graphes appliqués sur des graphes dirigés mettant en lumière l'efficacité de notre cadre.

Sommaire

1	Présentation du cadre théorique	48
2	Filtrage sur graphes dirigés	49
2.1	Définitions	49
2.2	Stabilité des filtres sur graphes	50
3	Opérateurs fondamentaux sur graphes dirigés	57
3.1	Marche aléatoire sur graphes dirigés	57
3.2	Généralisation de marches aléatoires	59
3.3	Laplaciens sur graphes dirigés	60
3.4	Opérateurs fondamentaux sur graphes dirigés et espaces de Hilbert	62
4	Transformée de Fourier sur graphes dirigés	63
5	Analyse de Fourier sur graphes dirigés	64
5.1	Régularité de signaux sur graphes	64
5.2	Analyse fréquentielle sur graphes dirigés	65
5.3	Filtres sur graphes de type marche aléatoire	67
5.4	Analyse de Fourier sur les groupes finis : un point de vue traitement de signal sur graphe	68
6	Applications	71
6.1	Apprentissage semi-supervisé sur des graphes dirigés via régularisation de type ℓ_2	72
6.2	Modélisation du signal sur des graphes dirigés par filtrage	78
7	Analyses multi-résolution sur graphes dirigés	90
7.1	Transformée en ondelettes redondantes sur graphes dirigés	91
7.2	Transformée en ondelettes décimée sur graphes dirigés	94
7.3	Applications	98

1 Présentation du cadre théorique

Dans cette section, nous introduisons le cadre théorique nécessaire à la compréhension du chapitre 4.

Soit $\mathcal{G} = (\mathcal{V}, \mathcal{E}, w)$ un graphe pondéré et dirigé où $\mathcal{V} = \{v_1, \dots, v_N\}$ est l'ensemble des sommets de cardinalité $|\mathcal{V}| = N$ et $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ est l'ensemble des arcs. Chaque arc $e_{ij} = (v_i, v_j)$ représente un lien dirigé du sommet v_i au sommet v_j . La fonction de poids $w : \mathcal{V} \times \mathcal{V} \rightarrow \mathbb{R}_+$ satisfait les conditions suivantes :

$$\begin{cases} w(v_i, v_j) > 0 & \text{si } e_{ij} \in \mathcal{E}, \\ w(v_i, v_j) = 0 & \text{si } e_{ij} \notin \mathcal{E}. \end{cases}$$

Le graphe $\mathcal{G}$ est représenté par sa matrice d'adjacence pondérée $\mathbf{W} = \{w_{ij}\}_{1 \leq i, j \leq N} \in \mathbb{R}_+^{N \times N}$ où $w_{ij} = w(v_i, v_j)$. Dans le cadre théorique de ce chapitre, le degré d'un sommet $v_i \in \mathcal{V}$, noté d_i est associé exclusivement à son degré sortant d_i^+ défini en section 1 du chapitre 3.

Une limitation de notre cadre est qu'il s'applique seulement dans le cas des graphes dirigés et pondérés avec des poids positifs ou nuls c.-à-d. $\mathbf{W} = \{w_{ij}\}_{1 \leq i, j \leq N} \in \mathbb{R}_+^{N \times N}$. D'autres approches telles que celle de Sandryhaila et Moura [44, 20] présentée en section 2.1 du chapitre 3 ou de Bauer [59] s'appliquent dans le cas des graphes dirigés avec des poids positifs et négatifs, par exemple dans le cas des neurosciences où l'on modélise les interactions entre les neurones par un graphe dirigé et dont chaque arc a un poids soit positif négatif représentant le caractère excitateur ou inhibiteur de la synapse. Cependant ces approches ne sont pas compatibles avec celles présentées ici.

Nous supposons à travers les sections théoriques de ce chapitre que le graphe dirigé $\mathcal{G}$ est **fortement connexe** et que les opérateurs de marche aléatoire sont **diagonalisables**. Ces deux conditions sont essentielles pour notre cadre théorique. Nous reviendrons en détail tout au long de ce chapitre sur ces deux conditions. Si le graphe dirigé n'est pas fortement connexe, la méthode de construction de la matrice de Google [60] à partir de la matrice d'adjacence originelle peut être appliquée afin d'assurer une connectivité forte du graphe résultant.

De manière analogue au chapitre 2, on définit par $f : \mathcal{V} \rightarrow \mathbb{C}$, une fonction définie sur les sommets d'un graphe dirigé $\mathcal{G}$ fortement connexe et $\mathbf{f}$ le signal sur graphe correspondant. De même, nous supposons tout au long du chapitre que les signaux sur graphes dirigés sont définis sur l'espace $\ell^2(\mathcal{V}, \mu)$ défini au chapitre 2, section 4.1.

2 Filtrage sur graphes dirigés

Dans la section 4.2 du chapitre 2, nous avons introduit des définitions et propriétés basiques du filtrage sur graphes non dirigés. Ces propriétés sont également valables dans le cas du filtrage sur graphes dirigés. Dans la section suivante, nous abordons le problème de la stabilité de filtres sur graphes dirigés.

2.1 Définitions

Soit $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ un graphe dirigé. Supposons un opérateur de référence sur graphe $\mathbf{R}$ diagonalisable dans $\mathbb{C}$. D'après la définition de filtre que nous avons retenue, un filtre sur graphe $\mathbf{H}$ peut s'exprimer comme une somme finie polynomiale d'un opérateur de référence sur graphe $\mathbf{R}$:

$$\mathbf{H} = \sum_{t=0}^T h_t \mathbf{R}^t, \quad h_t \in \mathbb{C}, \quad \forall t = 0, \dots, T.$$

Soit $\{\mathfrak{E}_j\}_{j=1}^m$ l'ensemble des espaces propres de $\mathbf{R}$ et $\{\mathbf{E}_j\}_{j=1}^m$ l'ensemble des projecteurs obliques correspondants, caractérisé par $\mathbf{E}_i \mathfrak{E}_j = \delta_{ij} \mathfrak{E}_j$. Un filtre sur graphe par rapport à un opérateur de référence diagonalisable $\mathbf{R}$ est une combinaison linéaire des projecteurs spectraux :

$$\mathbf{H} = \sum_{j=1}^m \gamma_j \mathbf{E}_j, \quad \gamma_j \in \mathbb{C}, \quad \mathbf{E}_j \in \mathbb{C}^{N \times N}, \quad \forall j = 1, \dots, m, \quad (4.1)$$

où γ_j est la valeur propre de $\mathbf{H}$ associée à l'espace propre $\mathfrak{E}_j$.

2.2 Stabilité des filtres sur graphes

Dans cette section, nous nous intéressons à la stabilité des filtres sur graphes. Généralement, nous constatons une forte sensibilité au bruit sur les filtres construits à partir des opérateurs de référence sur graphe dirigés. Les espaces propres des opérateurs de référence sur graphes dirigés ne sont pas orthogonaux et par conséquent ceux des filtres sur graphes linéairement invariant à ces opérateurs de référence ne le sont pas non plus. Un moyen d'évaluer la stabilité des filtres est de considérer une perturbation d'un signal sur graphe (que l'on suppose additive pour faciliter les calculs) et de regarder comment cela affecte le résultat.

Dans cette section, notre objectif est double. Nous souhaitons :

- Établir des bornes sur le conditionnement d'un filtre sur graphe.
- Modifier un filtre sur graphe pour améliorer sa robustesse par rapport à l'ajout d'un bruit additif sur un signal sur graphe.

2.2.1 Conditionnement de filtres sur graphes

Définition 4.1. *Le conditionnement d'une matrice non singulière $\mathbf{U}$, noté $\kappa(\mathbf{U})$ est :*

$$\kappa(\mathbf{U}) = \|\mathbf{U}\| \|\mathbf{U}^{-1}\|,$$

où $\|\cdot\|$ désigne la norme spectrale (ou la norme de Frobenius).

Un conditionnement large (typiquement $\kappa(\mathbf{U}) > 10^3$) indique un mauvais conditionnement de la matrice.

Proposition 4.1 (Caractérisation du conditionnement spectral (norme 2)). *Soit $\mathbb{R}^N$ muni de la norme euclidienne $\|\cdot\|$. Soit $\mathbf{U}$ une matrice non singulière. On note $\kappa_2(\mathbf{U})$, le conditionnement associé à la norme induite par la norme euclidienne sur $\mathbb{R}^N$.*

1. *Soit $\mathbf{U} \in \mathbb{R}^{N \times N}$ une matrice non singulière. On note σ_N /resp. σ_1 / la plus grande/resp. la plus petite/ valeur propre de $\mathbf{U}^\top \mathbf{U}$ (notons que $\mathbf{U}^\top \mathbf{U}$ est une matrice symétrique définie positive). Alors, nous avons :*

$$\kappa_2(\mathbf{U}) = \sqrt{\frac{\sigma_N}{\sigma_1}}.$$

2. Si de plus $\mathbf{U}$ est une matrice symétrique définie positive, alors nous avons :

$$\kappa_2(\mathbf{U}) = \frac{\lambda_N}{\lambda_1}.$$

où λ_N /resp. λ_1 / la plus grande/resp. la plus petite/ valeur propre de $\mathbf{U}$.

Le théorème suivant établit une borne supérieure du conditionnement d'un filtre sur graphe $\mathbf{H}$.

Proposition 4.2. *Soit $\mathbf{R}$ un opérateur de référence sur graphe, supposé diagonalisable, c.-à-d. $\mathbf{R} = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^{-1}$. Soit $\mathbf{H}$ un filtre sur graphe tel que $\mathbf{H} = \mathbf{Q}\mathbf{\Gamma}\mathbf{Q}^{-1}$. Une borne supérieure du conditionnement en norme spectrale de $\mathbf{H}$, noté $\kappa_2(\mathbf{H})$ est*

$$\kappa_2(\mathbf{H}) \leq \kappa_2^2(\mathbf{Q}) \frac{\gamma_{\max}}{\gamma_{\min}}.$$

où $\gamma_{\max}$ [resp. $\gamma_{\min}$] est le module de la valeur propre de $\mathbf{H}$ de plus grand [resp. petit] module.

Démonstration.

$$\begin{aligned} \kappa_2(\mathbf{H}) &= \|\mathbf{H}\|_2 \|\mathbf{H}^{-1}\|_2 \\ &= \|\mathbf{Q}\mathbf{\Gamma}\mathbf{Q}^{-1}\|_2 \|\mathbf{Q}\mathbf{\Gamma}^{-1}\mathbf{Q}^{-1}\|_2 \\ &\leq (\|\mathbf{Q}\|_2 \|\mathbf{Q}^{-1}\|_2)^2 \|\mathbf{\Gamma}\|_2 \|\mathbf{\Gamma}^{-1}\|_2 \\ &\leq \kappa_2^2(\mathbf{Q}) \frac{\gamma_{\max}}{\gamma_{\min}}. \end{aligned}$$

□

Le corollaire suivant établit une borne supérieure du conditionnement d'un filtre sur graphe $\mathbf{H}$ sous l'hypothèse que ses valeurs propres soient contenues dans une boule.

Corollaire 4.1. *Soit $\mathbf{R}$ un opérateur de référence sur graphe, supposé diagonalisable, c.-à-d. $\mathbf{R} = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^{-1}$. Soit $\mathbf{H}$ un filtre sur graphe tel que $\mathbf{H} = \mathbf{Q}\mathbf{\Gamma}\mathbf{Q}^{-1}$. Nous supposons toute valeur propre $\gamma \in \mathbf{\Gamma}$ appartient à la boule fermée de centre $\mu \in \mathbb{C}$ et de rayon $\eta|\mu|$, $\eta \in (0, 1)$ c.-à-d. $\gamma_k \in \mathcal{B}(\mu, \eta|\mu|)$, pour tout k . Une borne supérieure du conditionnement en norme spectrale de $\mathbf{H}$, $\kappa_2(\mathbf{H})$ est*

$$\kappa_2(\mathbf{H}) \leq \kappa_2(\mathbf{Q}) \frac{1 + \eta\kappa_2(\mathbf{Q})}{1 - \eta}. \quad (4.2)$$

Démonstration.

$$\begin{aligned}
\|\mathbf{H}\|_2 &= \|\mathbf{Q}\mathbf{\Gamma}\mathbf{Q}^{-1}\|_2 \\
&= \|\mathbf{Q}(\mathbf{\Gamma} - \mu\mathbf{I} + \mu\mathbf{I})\mathbf{Q}^{-1}\|_2 \\
&= \|\mathbf{Q}(\mathbf{\Gamma} - \mu\mathbf{I})\mathbf{Q}^{-1} + \mu\mathbf{I}\|_2 \\
&\leq \|\mathbf{Q}(\mathbf{\Gamma} - \mu\mathbf{I})\mathbf{Q}^{-1}\|_2 + |\mu| \\
&\leq \kappa_2(\mathbf{Q})\eta|\mu| + |\mu|, \\
\|\mathbf{H}\|_2 &\leq |\mu|(1 + \eta\kappa_2(\mathbf{Q})).
\end{aligned} \tag{4.3}$$

$$\begin{aligned}
\|\mathbf{H}^{-1}\|_2 &= \|\mathbf{Q}\mathbf{\Gamma}^{-1}\mathbf{Q}^{-1}\|_2 \\
&= \|\mathbf{Q}(\mathbf{\Gamma} - \mu\mathbf{I} + \mu\mathbf{I})^{-1}\mathbf{Q}^{-1}\|_2 \\
&= \|\mathbf{Q}(\mathbf{B} + \mu\mathbf{I})^{-1}\mathbf{Q}^{-1}\|_2, \quad \mathbf{B} = \mathbf{\Gamma} - \mu\mathbf{I} \\
&= \|\mathbf{Q}(\mathbf{B} - (-\mu)\mathbf{I})^{-1}\mathbf{Q}^{-1}\|_2, \\
&= \|\mathbf{Q}(\mathbf{B} - \beta\mathbf{I})^{-1}\mathbf{Q}^{-1}\|_2, \quad \beta = -\mu, \\
&\leq \kappa_2(\mathbf{Q})\|(\mathbf{B} - \beta\mathbf{I})^{-1}\|_2 \\
&\leq \kappa_2(\mathbf{Q})\|(-\beta)^{-1}(\mathbf{I} - \omega\mathbf{B})^{-1}\|_2, \quad \omega = -1/\beta \\
&\leq \kappa_2(\mathbf{Q})|\beta^{-1}| \sum_{k=0}^{\infty} |\omega^k| \|\mathbf{\Gamma} - \mu\mathbf{I}\|^k_2 \\
&\leq \kappa_2(\mathbf{Q})|\beta^{-1}| \sum_{k=0}^{\infty} |\omega^k| |\eta\mu|^k \\
\|\mathbf{H}^{-1}\|_2 &\leq \kappa_2(\mathbf{Q}) \frac{1}{|\mu|} (1 - \eta)^{-1}.
\end{aligned} \tag{4.4}$$

En multipliant les équations (4.3) et (4.4) entre elles, nous obtenons l'inégalité désirée (4.2). $\square$

Nous pouvons également définir des bornes supérieures du conditionnement d'un filtre sur graphe $\mathbf{H}$ en norme de Frobenius. Nous avons la relation suivante [61] :

$$\frac{\kappa_F(\mathbf{H})}{n} \leq \kappa_2(\mathbf{H}) \leq \kappa_F(\mathbf{H}),$$

où $\kappa_F(\mathbf{H}) = \|\mathbf{H}\|_F \|\mathbf{H}^{-1}\|_F$ est le nombre de conditionnement en norme de Frobenius.

Proposition 4.3. *Soit $\mathbf{R}$ un opérateur de référence sur graphe, supposé diagonalisable, c.-à-d. admettant une décomposition spectrale $\mathbf{R} = \sum_{j=1}^m \lambda_j \mathbf{G}_j$, où $\mathbf{G}_j$ est le projecteur spectral de rang supérieur ou égal à un, associé à la valeur propre λ_j . Soit $\mathbf{H}$ un filtre sur graphe qui s'exprime comme*

$\mathbf{R} = \sum_{j=1}^m \gamma_j \mathbf{G}_j$. Une borne supérieure du conditionnement en norme de Frobenius $\kappa_F(\mathbf{H})$ est :

$$\kappa_F(\mathbf{H}) \leq \left(\sum_{j=1}^m \gamma_j \|\mathbf{G}_j\|_F \right) \left(\sum_{k=1}^m \frac{1}{\gamma_k} \|\mathbf{G}_k\|_F \right). \quad (4.5)$$

Démonstration.

$$\begin{aligned} \kappa_F(\mathbf{H}) &= \|\mathbf{H}\|_F \|\mathbf{H}^{-1}\|_F, \\ \|\mathbf{H}\|_F &\leq \sum_{j=1}^m \gamma_j \|\mathbf{G}_j\|_F, \end{aligned} \quad (4.6)$$

$$\|\mathbf{H}^{-1}\|_F \leq \sum_{k=1}^m \frac{1}{\gamma_k} \|\mathbf{G}_k\|_F, \quad (4.7)$$

En multipliant les équations (4.6) et (4.7) entre elles, nous obtenons l'inégalité désirée (4.5). $\square$

2.2.2 Approximation de filtres sur graphes

Dans cette section, nous nous concentrons sur la modification de filtres sur graphe dirigé pour améliorer leur robustesse.

2.2.2.1 Formulation du problème Soit $\mathbf{H} \in \mathbb{R}^{N \times N}$ un filtre sur graphe dirigé construit à partir d'un opérateur de référence sur graphe dirigé $\mathbf{R}$. Soit un signal sur graphe $\mathbf{x} \in \mathbb{R}^{N \times N}$. Nous cherchons à construire un filtre sur graphe $\tilde{\mathbf{H}}$ qui minimise la quantité suivante :

$$\min_{\tilde{\mathbf{H}}} \mathbb{E} \left(\|\tilde{\mathbf{H}} \tilde{\mathbf{s}} - \mathbf{H} \mathbf{s}\|_2^2 \right). \quad (4.8)$$

où $\tilde{\mathbf{s}} = \mathbf{s} + \boldsymbol{\varepsilon}$ constitue le signal sur graphe bruité avec $\boldsymbol{\varepsilon} \sim \mathcal{N}(0, \boldsymbol{\Sigma}_{\varepsilon})$.

Nous proposons de résoudre le problème (4.8) pour deux types de condition sur le filtre $\tilde{\mathbf{H}}$:

1. $\tilde{\mathbf{H}}$ est un opérateur sur graphe quelconque sans contrainte.
2. $\tilde{\mathbf{H}}$ est un filtre construit à partir $\mathbf{R}$ de la manière suivante :

$$\tilde{\mathbf{H}} = \sum_{j=1}^m \nu_j \mathbf{G}_j.$$

2.2.2.2 Résolution du problème Nous résolvons le problème (4.8) selon les deux cas.

Cas 1 : pas de contraintes sur $\tilde{\mathbf{H}}$ La solution du problème (4.8) est $\tilde{\mathbf{H}} = \mathbf{H}\mathbf{W}$ où $\mathbf{W}$ est égale à :

$$\mathbf{W} = (\mathbf{s}\mathbf{s}^\top)(\mathbf{s}\mathbf{s}^\top + \Sigma_\varepsilon)^{-1}. \quad (4.9)$$

L'opérateur $\tilde{\mathbf{H}}$ n'est pas linéairement invariant à $\mathbf{R}$ et l'opérateur $\mathbf{W}$ s'apparente à un filtre de Wiener adapté aux signaux sur graphes.

Cas 2 : $\tilde{\mathbf{H}} = \sum_{j=1}^m \nu_j \mathbf{G}_j$ Dans ce second cas, nous réécrivons le problème d'optimisation (4.8) de la manière suivante :

$$\min_{\nu_j, j=1, \dots, m} \mathbb{E} \left[\left\| \left(\sum_{j=1}^m (\nu_j - \gamma_j) \mathbf{G}_j \mathbf{s} + \sum_{j=1}^m \nu_j \mathbf{G}_j \boldsymbol{\varepsilon} \right) \right\|_2^2 \right] \quad (4.10)$$

Nous cherchons à déterminer les valeurs propres $\nu_j, j = 1, \dots, m$ minimisant le problème (4.10) et nous gardons les projecteurs spectraux inchangés.

La solution du problème (4.10) est :

$$\boldsymbol{\nu} = (\mathbf{M}_s + \mathbf{M}_\varepsilon)^{-1} \mathbf{M}_s \boldsymbol{\gamma}, \quad (4.11)$$

où nous avons :

$$\mathbf{M}_s = \{\text{Tr}(\mathbf{G}_k^* \mathbf{G}_l \mathbf{s} \mathbf{s}^\top)\}_{k,l}, \quad \mathbf{M}_\varepsilon = \{\text{Tr}(\mathbf{G}_k^* \mathbf{G}_l \Sigma_\varepsilon)\}_{k,l}.$$

2.2.2.3 Exemple Dans cette section, nous abordons un exemple simple de modification d'un filtre sur graphe dirigé à trois sommets et tentons de comprendre comment les valeurs propres obtenues modifient le filtre. Soit $\mathbf{R} \in \mathbb{R}^{3 \times 3}$ un opérateur de référence sur graphe que l'on suppose diagonalisable, c.-à-d.

$$\mathbf{R} = \mathbf{X} \boldsymbol{\Lambda} \mathbf{X}^{-1},$$

où $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3]$ est la matrice de vecteurs propres, $\boldsymbol{\Lambda} = \text{diag}\{\lambda_1, \lambda_2, \lambda_3\}$ la matrice diagonale de valeurs propres. On pose également $\mathbf{Y} = \mathbf{X}^{-1}$ tel que :

$$\mathbf{y}_k^* \mathbf{x}_\ell = \delta_{k\ell}.$$

$\mathbf{R}$ est diagonalisable en trois espaces propres $\mathcal{E}_1, \mathcal{E}_2, \mathcal{E}_3$. de dimension un. Nous supposons que les sous-espaces propres $\mathcal{E}_1$ et $\mathcal{E}_2$ forment un angle quelconque tandis que les sous-espaces propres $\mathcal{E}_1, \mathcal{E}_3$ sont orthogonaux et les espaces propres $\mathcal{E}_2, \mathcal{E}_3$ également. L'opérateur de référence $\mathbf{R}$ admet ainsi la décomposition spectrale suivante :

$$\mathbf{R} = \sum_{k=1}^3 \lambda_k \mathbf{G}_k.$$

où les projecteurs spectraux $\mathbf{G}_k, k = 1, \dots, 3$ associés à leurs espaces propres respectifs $\mathcal{E}_k, k = 1, \dots, 3$ sont de rang 1 exprimés de la manière suivante :

$$\mathbf{G}_k = \mathbf{x}_k \mathbf{y}_k^*, \quad \forall k = 1, \dots, 3.$$

Un filtre sur graphe $\mathbf{H}$ construit à partir de $\mathbf{R}$ s'exprime de la manière suivante :

$$\mathbf{H} = \sum_{k=1}^3 \gamma_k \mathbf{G}_k, \quad \gamma_k \in \mathbb{C}$$

Nous cherchons à déterminer le filtre modifié $\tilde{\mathbf{H}}$ qui minimise le problème (4.10). Nous supposons le signal sur graphe considéré $\mathbf{s}$ colinéaire au vecteur propre $\mathbf{x}_1$, c.-à-d. $= \alpha \mathbf{x}_1$. Par ailleurs, nous supposons le bruit $\varepsilon \sim \mathcal{N}(0, \sigma_\varepsilon \mathbf{I})$. Les matrices $\mathbf{M}_s$ et $\mathbf{M}_\varepsilon$ de la solution (4.11) sont :

$$\mathbf{M}_\varepsilon = \sigma_\varepsilon [\mathbf{D}_1 + \text{Tr}(\mathbf{G}_1^* \mathbf{G}_2) (\mathbf{e}_1 \mathbf{e}_2^\top + \mathbf{e}_2 \mathbf{e}_1^\top)]$$

où $\mathbf{D}_1$ est

$$\mathbf{D}_1 = \text{diag}\{\|\mathbf{G}_1\|_F^2, \|\mathbf{G}_2\|_F^2, \|\mathbf{G}_3\|_F^2\},$$

et

$$\mathbf{M}_s = \alpha^2 \|\mathbf{x}_1\|^2 \mathbf{e}_1 \mathbf{e}_1^\top = \rho \mathbf{e}_1 \mathbf{e}_1^\top, \quad \rho = \alpha^2 \|\mathbf{x}_1\|^2. \quad (4.12)$$

où $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ est la base canonique dans $\mathbb{R}^3$. L'expression $\text{Tr}(\mathbf{G}_1^* \mathbf{G}_2)$ peut se réécrire de la manière suivante :

$$\begin{aligned} \text{Tr}(\mathbf{G}_1^* \mathbf{G}_2) &= \langle \mathbf{x}_1, \mathbf{x}_2 \rangle \langle \mathbf{y}_1, \mathbf{y}_2 \rangle. \\ &= \cos(\mathbf{x}_1, \mathbf{x}_2) \|\mathbf{x}_1\|_2 \|\mathbf{x}_2\|_2 \cos(\mathbf{y}_1, \mathbf{y}_2) \|\mathbf{y}_1\|_2 \|\mathbf{y}_2\|_2. \\ &= C \cos(\mathbf{x}_1, \mathbf{x}_2) \cos(\mathbf{y}_1, \mathbf{y}_2), \quad C = \|\mathbf{x}_1\|_2 \|\mathbf{x}_2\|_2 \|\mathbf{y}_1\|_2 \|\mathbf{y}_2\|_2. \end{aligned}$$

Ainsi l'expression $\text{Tr}(\mathbf{G}_1^* \mathbf{G}_2)$ est le produit des angles entre espaces propres et espaces propres "complémentaires", c'est-à-dire les angles entre les images des projecteurs et les angles entre leurs noyaux. L'angle entre les sous-espaces propres $\mathcal{E}_1$ et $\mathcal{E}_2$ est caractérisé par le cosinus entre les projecteurs $\mathbf{G}_1$ et $\mathbf{G}_2$, c'est à dire :

$$\begin{aligned} \cos(\mathbf{G}_1, \mathbf{G}_2) &= \frac{\langle \mathbf{G}_1, \mathbf{G}_2 \rangle_F}{\|\mathbf{G}_1\|_F \|\mathbf{G}_2\|_F} = \frac{\text{Tr}(\mathbf{G}_1^* \mathbf{G}_2)}{\|\mathbf{G}_1\|_F \|\mathbf{G}_2\|_F}. \\ &= \cos(\mathbf{x}_1, \mathbf{x}_2) \cos(\mathbf{y}_1, \mathbf{y}_2). \end{aligned}$$

La résolution du problème (4.10) donne les valeurs propres suivantes :

$$\begin{aligned} \nu_1 &= \frac{\rho}{\rho + \sigma_\varepsilon \frac{\|\mathbf{G}_1\|_F^2}{u}} \gamma_1 \\ \nu_2 &= \frac{\|\mathbf{G}_1\|_F}{\|\mathbf{G}_2\|_F} \frac{\rho(1-u)}{\cos(\mathbf{G}_1, \mathbf{G}_2)(\rho u + \sigma_\varepsilon \|\mathbf{G}_1\|_F^2)} \gamma_1 \\ \nu_3 &= 0. \end{aligned}$$

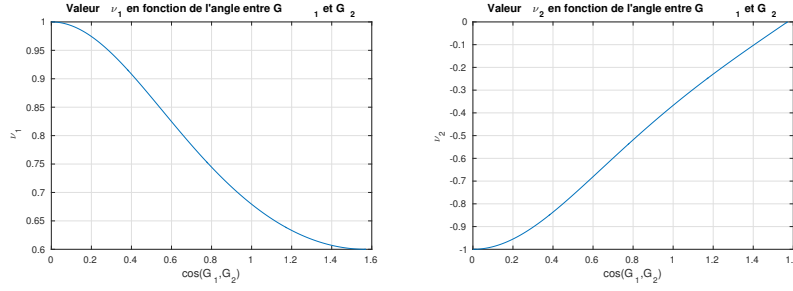


FIGURE 4.1 – Évolution des valeurs propres ν_1 et ν_2 en fonction de l'angle entre les projecteurs spectraux $\mathbf{G}_1$ et $\mathbf{G}_2$.

où u est :

$$u = 1 + \frac{1}{\tan^2(\mathbf{G}_1, \mathbf{G}_2)}. \quad (4.13)$$

Comme nous pouvons le remarquer, la valeur propre ν_3 est nulle. Cela était attendu étant donné que le signal sur graphe $\mathbf{s}$ vit dans le sous-espace propre engendré par $\mathbf{x}_1$, c'est-à-dire $\mathcal{E}_1$ et que les espaces propres $\mathcal{E}_1$ et $\mathcal{E}_3$ sont orthogonaux. Les seules valeurs propres non nulles sont les valeurs propres ν_1 et ν_2 . Elles font toutes deux intervenir l'angle entre les sous-espaces propres $\mathcal{E}_1$ et $\mathcal{E}_2$ à travers u introduit à l'équation (4.13). Par ailleurs elles sont proportionnelles à la valeur propre γ_1 dû au fait que le signal $\mathbf{s}$ est colinéaire à $\mathbf{x}_1$. Grâce aux expressions analytiques des valeurs propres ν_1, ν_2, ν_3 , nous pouvons étudier le comportement des valeurs propres en fonction de l'angle formé par les sous-espaces propres $\mathcal{E}_1$ et $\mathcal{E}_2$.

La figure 4.1 nous montre l'évolution des valeurs propres ν_1 et ν_2 en fonction du cosinus entre les projecteurs spectraux $\mathbf{G}_1$ et $\mathbf{G}_2$. Lorsque le cosinus entre les projecteurs spectraux $\mathbf{G}_1$ et $\mathbf{G}_2$ tend vers 0 ($\text{angle}(\mathbf{G}_1, \mathbf{G}_2) \rightarrow \pi/2$), c'est-à-dire lorsque les espaces propres $\mathcal{E}_1$ et $\mathcal{E}_2$ sont orthogonaux, nous avons :

$$\begin{aligned} \nu_1 &= \frac{\rho}{\rho + \sigma_\varepsilon \|\mathbf{G}_1\|_F^2} \gamma_1. \\ \nu_2 &= 0. \\ \nu_3 &= 0. \end{aligned}$$

Ainsi la valeur propre ν_2 devient nulle, étant donné que le signal sur graphe $\mathbf{s}$ vit dans le sous-espace propre $\mathcal{E}_1$ et que les espaces propres $\mathcal{E}_1$ et $\mathcal{E}_2$ tendent à devenir orthogonaux.

Lorsque le cosinus entre les projecteurs spectraux $\mathbf{G}_1$ et $\mathbf{G}_2$ tend vers 1 ($\text{angle}(\mathbf{G}_1, \mathbf{G}_2) \rightarrow 0$), c'est-à-dire lorsque les espaces propres $\mathcal{E}_1$ et $\mathcal{E}_2$ sont

colinéaires, nous avons :

$$\begin{aligned}\nu_1 &= \gamma_1. \\ \nu_2 &= -\frac{\|\mathbf{G}_1\|_F}{\|\mathbf{G}_2\|_F}\gamma_1. \\ \nu_3 &= 0.\end{aligned}$$

Ce cas est le cas extrême où l'opérateur est mal conditionné. Les valeurs propres ν_1 et ν_2 se "compensent", c'est-à-dire que ν_2 est l'opposé de ν_1 mais pondéré par le ratio des normes de Frobenius de $\mathbf{G}_1$ et $\mathbf{G}_2$. Par ailleurs le paramètre du bruit σ_ε disparaît car $u \rightarrow \infty$ quand $\mathcal{E}_1$ et $\mathcal{E}_2$ deviennent colinéaires.

3 Opérateurs fondamentaux sur graphes dirigés

Dans cette section, nous introduisons les opérateurs linéaires fondamentaux ainsi que leurs propriétés afin de développer une analyse harmonique sur graphes dirigés.

3.1 Marche aléatoire sur graphes dirigés

Les propriétés générales des marches aléatoires ont été introduites au chapitre 2, section 5.1. Ces propriétés introduites pour le cas non dirigé s'appliquent également dans le cas dirigé. Dans cette section, nous complétons les notions apportées précédemment sur les marches aléatoires par des propriétés supplémentaires relatives aux graphes dirigés. Nous définissons une marche aléatoire sur un graphe dirigé de la manière suivante.

Définition 4.2. *Une marche aléatoire sur un graphe dirigé $\mathcal{G} = (\mathcal{V}, \mathcal{E}, w)$ fortement connexe est une chaîne homogène de Markov $\mathcal{X} = (X_n)_{n \geq 0}$ non nécessairement réversible munie d'un espace d'états fini $\mathcal{V}$ et dont les probabilités de transition sont proportionnelles aux poids des arcs (supposés positifs, cf section 1 de ce chapitre).*

La matrice de transition $\mathbf{P} \in \mathbb{R}^{N \times N}$ d'un graphe dirigé $\mathcal{G}$ fortement connexe peut s'exprimer de la manière suivante :

$$\mathbf{P} = \mathbf{D}^{-1}\mathbf{W},$$

où $\mathbf{D} = \text{diag}\{d_1, \dots, d_n\}$ est la matrice diagonale des degrés sortants des sommets et $\mathbf{W}$ est la matrice d'adjacence pondérée.

En tant qu'opérateur appliqué à des signaux sur graphes dirigés, la matrice de transition $\mathbf{P}$ réalise comme dans le cas non dirigé une opération de moyennage local. En effet, la valeur de $\mathbf{P}\mathbf{f}$ au sommet i correspond à la somme des valeurs de $\mathbf{f}$ aux sommets connectés à i , pondérée par les valeurs

de la ligne i de $\mathbf{P}$. Une particularité est que les sommets correspondant aux poids non nuls de la ligne i de $\mathbf{P}$ sont les sommets *enfants* de i , i.e. ceux qui sont destinations d'arcs provenant de i . De ce fait, en appliquant $\mathbf{P}$ à un signal $\mathbf{f}$, l'information à un sommet donné de $\mathbf{f}$ est propagée en direction de ses parents et non de ses enfants, ce qui peut paraître contre-intuitif. En pratique, il sera donc important de tenir compte de cette observation pour orienter les arcs du graphe de telle manière que la direction de propagation de l'information corresponde à l'application.

3.1.1 Irréductibilité

La notion d'irréductibilité est similaire au cas non dirigé, c'est-à-dire pour toute marche aléatoire $\mathcal{X}$, nous avons :

$$\forall x, y \in \mathcal{V}, \exists m < \infty : \mathbb{P}(X_{n+m} = y | X_n = x) > 0.$$

L'irréductibilité de la chaîne de Markov $\mathcal{X}$ est équivalente à la forte connexité du graphe $\mathcal{G}$.

3.1.2 Réversibilité

Soit $\mathcal{X} = (X_n)_{n \geq 0}$ une chaîne de Markov discrète et $M > 0$ l'horizon de temps fini. Nous définissons la chaîne de Markov retournée dans le temps $\mathcal{X}^* = (X_n^*) = (X_{M-n})$ pour $n = 0, \dots, M$. Soit $\mathbf{P}^* = \{p^*(x, y)\}_{x, y \in \mathcal{V}}$ la matrice de transition associée à la chaîne de Markov $\mathcal{X}^*$. Elle vérifie :

$$p^*(x, y) = \mathbb{P}(X_{n+1}^* = y | X_n^* = x) = \frac{\mathbb{P}(X_{M-n-1} = y)}{\mathbb{P}(X_{M-n} = x)} p(y, x), \quad \forall x, y \in \mathcal{V}.$$

Supposons que $\mathcal{X}$ est ergodique avec distribution stationnaire π . Alors la chaîne de Markov retournée dans le temps $\mathcal{X}^*$ est également ergodique, de distribution stationnaire π et les entrées $p^*(x, y)$ de $\mathbf{P}^*$ sont

$$p^*(x, y) = \frac{\pi(y)}{\pi(x)} p(y, x), \quad \forall x, y \in \mathcal{V},$$

ou dans sa version matricielle

$$\mathbf{P}^* = \mathbf{\Pi}^{-1} \mathbf{P}^\top \mathbf{\Pi}, \quad (4.14)$$

où $\mathbf{\Pi} = \text{diag}\{\pi(v_1), \dots, \pi(v_N)\}$ est la matrice diagonale de la distribution stationnaire.

La marche aléatoire ergodique $\mathcal{X}$ avec sa matrice de transition $\mathbf{P}$ est réversible si et seulement si nous avons la relation suivante :

$$\mathbf{P} = \mathbf{P}^*.$$

Remarque 4.1. *Les marches aléatoires ergodiques sur des graphes non dirigés sont réversibles. Cela signifie que la matrice de transition associée à la marche aléatoire retournée dans le temps $\mathcal{X}^*$, $\mathbf{P}^*$ est égale à la matrice de transition $\mathbf{P}$ de la marche aléatoire originelle $\mathcal{X}$, c'est-à-dire $\mathbf{P}^* = \mathbf{P}$.*

Remarque 4.2. *Les marches aléatoires ergodiques définies sur graphes dirigés sont typiquement non réversibles. Néanmoins, des marches aléatoires réversibles peuvent être construites sur des graphes dirigés en modifiant la marche aléatoire ergodique non-réversible. Une telle modification est abordée dans la section 3.2.2.*

Introduisons l'espace de Hilbert $\ell^2(\mathcal{V}, \pi)$ muni de son produit scalaire.

$$\langle \mathbf{f}, \mathbf{g} \rangle_\pi = \sum_{x \in \mathcal{V}} \overline{f(x)} g(x) \pi(x).$$

Dans cet espace de Hilbert, $\mathbf{P}^*$ est l'adjoint de $\mathbf{P}$, c.-à-d. $\langle \mathbf{f}, \mathbf{P}\mathbf{g} \rangle_\pi = \langle \mathbf{P}^*\mathbf{f}, \mathbf{g} \rangle_\pi$, pour tout $\mathbf{f}, \mathbf{g} \in \ell^2(\mathcal{V}, \pi)$.

3.1.3 Périodicité

La notion de périodicité est donnée à la définition 2.15 du chapitre 2. Dans certains cas, il est possible qu'une chaîne de Markov $\mathcal{X}$ soit irréductible mais périodique et donc non-ergodique. Un moyen naturel de surmonter le problème sera considéré à la section 3.2.1.

3.2 Généralisation de marches aléatoires

Soit une marche aléatoire $\mathcal{X}$ sur un graphe dirigé $\mathcal{G}$, munie de sa matrice de transition $\mathbf{P}$. Nous sommes en mesure de construire de nouveaux types de marches aléatoires basés sur $\mathbf{P}$ avec des objectifs variés.

3.2.1 Marches aléatoires paresseuses [49]

La périodicité d'une marche aléatoire $\mathcal{X}$ peut être surmontée en considérant la version paresseuse de la marche aléatoire $\mathcal{X}$. La matrice de transition $\tilde{\mathbf{P}}$ associée à la marche aléatoire paresseuse $\tilde{\mathcal{X}}$, basée sur $\mathbf{P}$ est exprimée de la manière suivante :

$$\tilde{\mathbf{P}} = \frac{\mathbf{I} + \mathbf{P}}{2}.$$

La marche aléatoire paresseuse peut être vue comme une marche aléatoire sur un graphe dirigé modifié $\tilde{\mathcal{G}}$ où un arc provenant de chaque sommet de $\mathcal{G}$ à lui-même est ajouté avec un poids égal au degré du sommet dans $\mathcal{G}$. La marche aléatoire paresseuse $\tilde{\mathcal{X}}$ est toujours apériodique [51]. Ainsi lorsque $\mathcal{G}$ est fortement connexe $\tilde{\mathcal{X}}$ est également ergodique avec distribution stationnaire π vérifiant $\pi = \pi\tilde{\mathbf{P}} = \pi\mathbf{P}$.

Plus généralement, nous définissons par $\tilde{\mathcal{P}}$, la classe des matrices de transitions associée aux marches aléatoires paresseuses généralisées de la manière suivante :

$$\tilde{\mathcal{P}} = \left\{ \tilde{\mathbf{P}}_\gamma : \tilde{\mathbf{P}}_\gamma = (1 - \gamma)\mathbf{P} + \gamma\mathbf{I} \middle| \gamma \in [0, 1] \right\}.$$

Nous remarquons que les éléments $\tilde{\mathbf{P}}_\gamma \in \tilde{\mathcal{P}}$ partagent les mêmes espaces propres que $\mathbf{P}$. Par conséquent, les filtres sur graphe basés sur $\tilde{\mathbf{P}}_\gamma$ sont également des filtres sur graphe basés sur $\mathbf{P}$.

3.2.2 Réversibilisation additive

À partir d'une marche aléatoire $\mathcal{X}$, ergodique, de distribution stationnaire π , non-réversible et munie de sa matrice de transition $\mathbf{P}$, une réversibilisation additive de $\mathcal{X}$ peut être construite [62], notée $\bar{\mathcal{X}}$, dont la matrice de transition est la moyenne entre $\mathbf{P}$ et sa version retournée dans le temps $\mathbf{P}^*$:

$$\bar{\mathbf{P}} = \frac{\mathbf{P} + \mathbf{P}^*}{2}. \quad (4.15)$$

$\bar{\mathcal{X}}$ est donc une marche aléatoire avec la même unique distribution stationnaire π .

Plus généralement, nous définissons par $\bar{\mathcal{P}}$ la classe des combinaisons convexes entre la matrice de transition $\mathbf{P}$ et sa version retournée dans le temps $\mathbf{P}^*$ de la manière suivante :

$$\bar{\mathcal{P}} = \left\{ \bar{\mathbf{P}}_\alpha : \bar{\mathbf{P}}_\alpha = (1 - \alpha)\mathbf{P} + \alpha\mathbf{P}^* \middle| \alpha \in [0, 1] \right\}. \quad (4.16)$$

Nous noterons que les éléments $\bar{\mathbf{P}}_\alpha \in \bar{\mathcal{P}}$ partagent la même distribution stationnaire π , mais pas les mêmes espaces propres. Nous noterons également que toutes les marches aléatoires associées aux matrices de transition $\bar{\mathbf{P}}_\alpha \in \bar{\mathcal{P}}$ sont non réversibles excepté pour $\alpha = 1/2$, c'est-à-dire. $\bar{\mathbf{P}}_{1/2} = \bar{\mathbf{P}}$.

Remarque 4.3. *Étant donnée une marche aléatoire ergodique $(\mathcal{X}, \mathbf{P}, \pi)$, nous noterons que toutes les versions paresseuses et réversibilisées, respectivement $\tilde{\mathbf{P}}_\gamma \in \tilde{\mathcal{P}}$ et $\bar{\mathbf{P}}_\alpha \in \bar{\mathcal{P}}$ partagent la même distribution stationnaire π .*

Remarque 4.4. *Comme généralisation des marches aléatoires, nous pouvons également considérer la réversibilisation multiplicative de la marche aléatoire ergodique et non-réversible $\mathbf{P}$, c'est-à-dire $\mathbf{P}\mathbf{P}^*$ qui est aussi une chaîne de Markov réversible avec distribution stationnaire π . Elle est utile par exemple pour définir les bornes de convergences de chaînes de Markov non réversibles [62].*

3.3 Laplaciens sur graphes dirigés

Dans cette section, nous présentons les équivalents sur graphes dirigés des différentes versions de laplaciens [63, 64] évoquées au chapitre 2, section 5.2.

3.3.1 Laplacien normalisé sur graphe dirigé

Nous introduisons le laplacien normalisé sur un graphe dirigé $\mathcal{G}$ exprimé en fonction de la matrice de transition $\mathbf{P}$ d'une marche aléatoire ergodique $\mathcal{X}$ sur $\mathcal{G}$. Cet opérateur est défini de la manière suivante :

Définition 4.3 ([65, 66]). *Soit $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ un graphe dirigé de cardinalité $|\mathcal{V}| = N$. Soit $\mathcal{X}$ une marche aléatoire ergodique sur $\mathcal{G}$ munie de sa matrice de transition $\mathbf{P}$ et de sa distribution stationnaire π . Le laplacien normalisé sur $\mathcal{G}$ est défini de la manière suivante :*

$$\mathcal{L} = \mathbf{I}_N - \frac{\mathbf{\Pi}^{1/2} \mathbf{P} \mathbf{\Pi}^{-1/2} + \mathbf{\Pi}^{-1/2} \mathbf{P}^\top \mathbf{\Pi}^{1/2}}{2}, \quad (4.17)$$

où $\mathbf{I}_N$ est la matrice identité et $\mathbf{\Pi} = \text{diag}\{\pi(v_1), \dots, \pi(v_N)\}$ est la matrice diagonale de la distribution stationnaire.

3.3.2 Laplacien marche aléatoire

Une autre définition est le laplacien marche aléatoire $\mathbf{L}_{\text{RW}}$, défini de la manière suivante :

$$\mathbf{L}_{\text{RW}} = \mathbf{I} - \bar{\mathbf{P}}. \quad (4.18)$$

Le laplacien normalisé sur graphe dirigé $\mathcal{L}$ est relié au laplacien marche aléatoire à travers la relation suivante :

$$\mathcal{L} = \mathbf{\Pi}^{1/2} \mathbf{L}_{\text{RW}} \mathbf{\Pi}^{-1/2}. \quad (4.19)$$

Plus généralement, nous définissons le laplacien marche aléatoire sur graphe dirigé associé à sa matrice de transition $\bar{\mathbf{P}}_\alpha \in \bar{\mathcal{P}}$ par :

$$\mathbf{L}_{\text{RW},\alpha} = \mathbf{I} - \frac{\bar{\mathbf{P}}_\alpha + \bar{\mathbf{P}}_\alpha^*}{2}.$$

Dans la proposition suivante, nous montrons que le laplacien marche aléatoire $\mathbf{L}_{\text{RW},\alpha}$ est égal à $\mathbf{L}_{\text{RW}}$.

Proposition 4.4. *Pour tout $\bar{\mathbf{P}}_\alpha \in \bar{\mathcal{P}}$, nous avons :*

$$\mathbf{L}_{\text{RW},\alpha} = \mathbf{L}_{\text{RW}}, \quad \forall \alpha \in [0, 1]. \quad (4.20)$$

Démonstration.

$$\begin{aligned} \mathbf{L}_{\text{RW},\alpha} &= \mathbf{I} - \frac{\bar{\mathbf{P}}_\alpha + \bar{\mathbf{P}}_\alpha^*}{2} \\ &= \mathbf{I} - \frac{\alpha \mathbf{P} + (1 - \alpha) \mathbf{P}^* + \alpha \mathbf{P}^* + (1 - \alpha) \mathbf{P}}{2} \\ &= \mathbf{I} - \frac{\alpha(\mathbf{P} + \mathbf{P}^*) + (1 - \alpha)(\mathbf{P} + \mathbf{P}^*)}{2} \\ &= \mathbf{I} - \frac{\mathbf{P} + \mathbf{P}^*}{2} \\ &= \mathbf{L}_{\text{RW}}. \end{aligned} \quad \square$$

3.3.3 Laplacien combinatoire

Le laplacien combinatoire sur graphe dirigé $\mathbf{L}$ est défini de la manière suivante :

$$\mathbf{L} = \mathbf{\Pi} - \frac{\mathbf{\Pi}\mathbf{P} + \mathbf{P}^\top\mathbf{\Pi}}{2}. \quad (4.21)$$

Ce dernier est relié au laplacien marche aléatoire sur graphe dirigé à travers la relation suivante :

$$\mathbf{L} = \mathbf{\Pi} \mathbf{L}_{\text{RW}}.$$

Remarque 4.5. Comme mentionné dans la remarque 4.1, la matrice de transition d'une marche aléatoire d'un graphe non dirigé est égale à la matrice de transition de sa marche aléatoire retournée dans le temps. Par ailleurs, la distribution stationnaire π admet une forme analytique comme mentionné à la section 5.1 du chapitre 2. Par conséquent, les définitions (4.17), (4.18) et (4.21) pour graphes non dirigés prolongent les définitions usuelles abordées à la section 5.2 du chapitre 2 (à un scalaire multiplicatif près).

3.4 Opérateurs fondamentaux sur graphes dirigés et espaces de Hilbert

Soient $\ell^2(\mathcal{V})$ et $\ell^2(\mathcal{V}, \pi)$ les espaces de Hilbert associés respectivement à la mesure de comptage et à la mesure stationnaire π de la marche aléatoire $\mathcal{X}$. Soit $\varphi : \ell^2(\mathcal{V}) \rightarrow \ell^2(\mathcal{V}, \pi)$ l'application linéaire définie de la manière suivante :

$$\varphi : \mathbf{f} \rightarrow \mathbf{\Pi}^{-1/2} \mathbf{f}, \quad \forall \mathbf{f} \in \ell^2(\mathcal{V}). \quad (4.22)$$

La proposition suivante indique que l'application linéaire φ est une isométrie.

Proposition 4.5. La transformation linéaire $\varphi : \ell^2(\mathcal{V}) \mapsto \ell^2(\mathcal{V}, \pi)$ est une isométrie.

Démonstration. Étant donné $\mathbf{g}, \mathbf{h} \in \ell^2(\mathcal{V})$, et la transformation linéaire φ , nous avons par l'application de la définition (2.9) :

$$\langle \varphi(\mathbf{g}), \varphi(\mathbf{h}) \rangle_\pi = \langle \mathbf{\Pi}^{-1/2} \mathbf{g}, \mathbf{\Pi}^{-1/2} \mathbf{h} \rangle_\pi = \langle \mathbf{g}, \mathbf{h} \rangle. \quad \square$$

Introduisons l'opérateur linéaire $\mathbf{T} \in \ell^2(\mathcal{V})$ défini de la manière suivante :

$$\mathbf{T} = \mathbf{\Pi}^{1/2} \mathbf{P} \mathbf{\Pi}^{-1/2}. \quad (4.23)$$

La proposition suivante établit un lien d'entrelacement entre l'opérateur linéaire $\mathbf{T}$ et $\mathbf{P}$.

Proposition 4.6. L'application linéaire φ entrelace l'opérateur $\mathbf{T} \in \ell^2(\mathcal{V})$ et l'opérateur $\mathbf{P} \in \ell^2(\mathcal{V}, \pi)$.

Démonstration. En appliquant la définition 2.10, nous avons :

$$\varphi(\mathbf{T}) = \mathbf{\Pi}^{-1/2} \mathbf{T} = \mathbf{\Pi}^{-1/2} \mathbf{\Pi}^{1/2} \mathbf{P} \mathbf{\Pi}^{-1/2} = \mathbf{P} \mathbf{\Pi}^{-1/2}. \quad \square$$

Ainsi les opérateurs $\mathbf{T}$ et $\mathbf{P}$ sont similaires par rapport à φ . De manière similaire, le laplacien marche aléatoire $\mathbf{L}_{\text{RW}}$ et le laplacien normalisé $\mathcal{L}$ sont également similaires par rapport à φ .

4 Transformée de Fourier sur graphes dirigés

Soit $\mathbf{P}$ la matrice de transition d'une marche aléatoire ergodique $\mathcal{X}$. Nous supposons que $\mathbf{P}$ est diagonalisable, c'est-à-dire que $\mathbf{P}$ admet une décomposition en valeurs propres suivante :

$$\mathbf{P} = \mathbf{\Xi} \mathbf{\Theta} \mathbf{\Xi}^{-1},$$

où $\mathbf{\Xi} = [\boldsymbol{\xi}_1, \dots, \boldsymbol{\xi}_N]$ est une matrice de changement de base avec $\boldsymbol{\xi}_j \in \mathbb{C}^N, j = 1, \dots, N$ et $\mathbf{\Theta} = \text{diag}\{\vartheta_1, \dots, \vartheta_N\}$ est la matrice diagonale des valeurs propres, non nécessairement distinctes. Étant donné un signal sur graphe $\mathbf{s}$, sa *transformée de Fourier* notée $\hat{\mathbf{s}} = [\hat{s}_1, \dots, \hat{s}_N]$ est :

$$\hat{\mathbf{s}} = \mathbf{\Xi}^{-1} \mathbf{s}.$$

Les valeurs $\{\hat{s}_j\}_{j=1}^N$ caractérisent le contenu du signal sur graphe $\mathbf{s}$ dans le domaine de Fourier sur graphe.

La transformée de Fourier sur graphe n'est pas définie de manière unique comme cela dépend du choix de la matrice de changement de base $\mathbf{\Xi}$. Tout d'abord, tous les vecteurs propres sont définis par $\boldsymbol{\xi}_j$ à un scalaire multiplicatif près. Ceci peut être facilement résolu en normalisant les vecteurs et en rendant par exemple le premier coefficient non nul de chaque vecteur propre réel et positif. Deuxièmement, chaque fois qu'il y a des espaces propres dont la dimension est supérieure à 1, il y a des choix infinis pour les orientations des vecteurs de la base dans ces espaces propres. Il n'y a généralement pas de solution à ce deuxième problème. Cela implique que les valeurs des coefficients de Fourier individuels $\hat{s}_j$ dans ces espaces propres sont plutôt dénuées de sens. Cependant, elle n'affecte que les cas avec des espaces propres de dimension supérieure à 1, ce qui n'est pas si courant dans les applications.

La *transformée de Fourier inverse* correspondant à la matrice de changement de base $\mathbf{\Xi}$ est exprimée de la manière suivante

$$\mathbf{s} = \mathbf{\Xi} \hat{\mathbf{s}}.$$

Elle reconstruit le signal originel à partir du contenu du signal en formant une combinaison linéaire des modes de Fourier pondérés par leurs coefficients de Fourier respectifs.

Plus g n ralement, pour toute matrice de transition suppos e diagonalisable $\bar{\mathbf{P}}_\alpha \in \bar{\mathcal{P}}$, nous avons :

$$\bar{\mathbf{P}}_\alpha = \mathbf{\Xi}_\alpha \mathbf{\Theta}_\alpha \mathbf{\Xi}_\alpha^{-1}.$$

Par cons quent, nous pouvons construire une infinit  de bases de type Fourier $\{\mathbf{\Xi}_\alpha\}_{\alpha \in [0,1]}$ pour des signaux sur graphes dirig s. Parmi ces bases de Fourier, le cas $\alpha = 1/2$ est particuli rement int ressant dans la mesure o  $\bar{\mathbf{P}}_{1/2} = \bar{\mathbf{P}}$ est un op rateur auto-adjoint pour $\ell^2(\mathcal{V}, \pi)$. Cela implique qu'il existe une matrice de changement de base orthonormale pour $\ell^2(\mathcal{V}, \pi)$, $\mathbf{\Xi}_{1/2}$, qui donne le th or me suivant.

Th or me 4.1 (Th or me de Parseval g n ralis ). * tant donn  une matrice de changement de base $\mathbf{\Xi}_{1/2}$ de $\bar{\mathbf{P}}$ qui est orthonormale pour $\ell^2(\mathcal{V}, \pi)$, la transform e de Fourier correspondante est un op rateur isom trique de $\ell^2(\mathcal{V}, \pi)$   $\ell^2(\{1, \dots, N\})$:*

$$\langle \mathbf{\Xi}_{1/2} \mathbf{x}, \mathbf{\Xi}_{1/2} \mathbf{y} \rangle_\pi = \langle \mathbf{x}, \mathbf{y} \rangle, \quad \forall \mathbf{x}, \mathbf{y} \in \ell^2(\{1, \dots, N\}). \quad (4.24)$$

D monstration. $\mathbf{\Xi}_{1/2}^\top \mathbf{\Pi} \mathbf{\Xi}_{1/2} = \mathbf{I} \implies \langle \mathbf{\Xi}_{1/2} \mathbf{x}, \mathbf{\Xi}_{1/2} \mathbf{y} \rangle_\pi = \mathbf{x}^\top \mathbf{\Xi}_{1/2}^\top \mathbf{\Pi} \mathbf{\Xi}_{1/2} \mathbf{y} = \mathbf{x}^\top \mathbf{y} = \langle \mathbf{x}, \mathbf{y} \rangle.$ $\square$

Une autre propri t  int ressante de $\bar{\mathbf{P}}$ est que ses espaces propres sont les m mes que ceux du laplacien marche al atoire $\mathbf{L}_{RW}$ (4.18).

5 Analyse de Fourier sur graphes dirig s

Dans cette section, nous proposons une nouvelle analyse sur graphes dirig s diff rente des approches [20, 23, 22, 21] abord es au chapitre 3. Notre approche se fonde sur l' tude de variations des vecteurs propres de l'op rateur de marche al atoire. Avant d'aborder cette analyse de Fourier plus en profondeur, nous introduisons les  l ments nous permettant d' tudier la variation de signaux sur des graphes dirig s.

5.1 R gularit  de signaux sur graphes

Tout comme sur graphe non dirig , le comportement d'un signal d fini sur graphe dirig  peut  galement  tre analys  en mesurant sa r gularit . Nous d finissons la "longueur du gradient" d'un signal sur graphe dirig  de la mani re suivante.

D finition 4.4 ([24]). *Soit $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ un graphe dirig  et $\nu : \mathcal{E} \rightarrow [0, \infty)$ une mesure positive d finie sur l'ensemble des arcs $\mathcal{E}$. La "longueur du gradient" d'un signal sur graphe $f : \mathcal{V} \rightarrow \mathbb{C}$ situ e au sommet $x \in \mathcal{V}$ sous la mesure ν est :*

$$|\nabla \mathbf{f}(x)| = \left(\frac{1}{2} \sum_{y \in \mathcal{V}, (x,y) \in \mathcal{E}} |f(x) - f(y)|^2 \nu(x,y) \right)^{1/2}.$$

Nous introduisons l'énergie de Dirichlet comme une mesure de régularité globale d'un signal sur un graphe arbitraire fortement connexe.

Définition 4.5 ([25]). *L'énergie de Dirichlet d'un signal sur graphe $f : \mathcal{V} \rightarrow \mathbb{C}$ associée à une marche aléatoire ergodique munie d'une matrice de transition $\mathbf{P}$ et de distribution stationnaire π sur un graphe fortement connexe $\mathcal{G}$ est :*

$$\begin{aligned} \mathcal{D}_{\pi, \mathbf{P}}^2(\mathbf{f}) &= \frac{1}{2} \sum_{(x,y) \in \mathcal{E}} \pi(x)p(x,y)|f(x) - f(y)|^2, \\ &= \langle \mathbf{f}, \mathbf{L}_{\text{RW}} \mathbf{f} \rangle_{\pi}. \end{aligned} \quad (4.25)$$

Comme nous pouvons le remarquer, $\mathcal{D}_{\pi, \mathbf{P}}^2(\mathbf{f}) = \|\nabla \mathbf{f}\|_{2, \nu}^2$ où $\nu(x,y) = \pi(x)p(x,y)$ avec $p(x,y)$ l'entrée (x,y) de la matrice de transition $\mathbf{P}$ et $\pi(x)$ la distribution stationnaire au sommet x . Pour toute matrice de transition $\bar{\mathbf{P}}_{\alpha} \in \bar{\mathcal{P}}$, l'énergie de Dirichlet est identique :

$$\mathcal{D}_{\pi, \bar{\mathbf{P}}_{\alpha}}^2(\mathbf{f}) = \mathcal{D}_{\pi, \mathbf{P}}^2(\mathbf{f}) = \langle \mathbf{f}, \mathbf{L}_{\text{RW}} \mathbf{f} \rangle_{\pi}.$$

Nous introduisons également le quotient de Rayleigh d'un signal sur graphe $f : \mathcal{V} \rightarrow \mathbb{C}$ associé à l'énergie de Dirichlet $\mathcal{D}_{\pi, \mathbf{P}}^2(\mathbf{f})$

$$\mathcal{R}_{\pi, \mathbf{P}}(\mathbf{f}) = \frac{\mathcal{D}_{\pi, \mathbf{P}}^2(\mathbf{f})}{\|\mathbf{f}\|_{\pi}^2}.$$

Pour tout $\bar{\mathbf{P}}_{\alpha} \in \bar{\mathcal{P}}$, nous avons :

$$\mathcal{R}_{\pi, \mathbf{P}}(\mathbf{f}) = \mathcal{R}_{\pi, \bar{\mathbf{P}}_{\alpha}}(\mathbf{f}).$$

5.2 Analyse fréquentielle sur graphes dirigés

La proposition suivante établit un lien clé entre la régularité des vecteurs propres de la matrice de transition $\mathbf{P}$ de la marche aléatoire ergodique $\mathcal{X}$ et leurs valeurs propres associées.

Proposition 4.7. *Soit $\xi \in \mathbb{C}^N$ un vecteur propre d'une matrice de transition $\mathbf{P}$ d'une marche aléatoire ergodique $\mathcal{X}$, avec distribution stationnaire π , associé à la valeur propre $\vartheta \in \mathbb{C}$. Le quotient de Rayleigh de ξ est donné par :*

$$\mathcal{R}_{\pi, \mathbf{P}}(\xi) = 1 - \Re(\vartheta),$$

où $\Re(\vartheta)$ est la partie réelle de $\vartheta \in \mathbb{C}$.

Démonstration.

$$\begin{aligned}
\mathcal{R}_{\pi, \mathbf{P}}(\xi) &= \frac{1}{\|\xi\|_\pi^2} \left(\langle \xi, \xi \rangle_\pi - \frac{1}{2} \langle \xi, \mathbf{P}\xi \rangle_\pi - \frac{1}{2} \langle \xi, \mathbf{P}^*\xi \rangle_\pi \right), \\
&= \frac{1}{\|\xi\|_\pi^2} \left(\|\xi\|_\pi^2 - \frac{\vartheta + \bar{\vartheta}}{2} \|\xi\|_\pi^2 \right), \\
\mathcal{R}_{\pi, \mathbf{P}}(\xi) &= [1 - \Re(\vartheta)]. \quad \square
\end{aligned}$$

La dernière proposition nous indique donc que la régularité d'un vecteur propre quelconque ξ de $\mathbf{P}$, décrit par son quotient de Rayleigh est associée à la partie réelle de sa valeur propre respective ϑ . Plus généralement, la régularité d'un vecteur propre quelconque ξ_α de $\mathbf{P}_\alpha \in \mathcal{P}$ est caractérisée exactement de la même manière en proposition 4.7, c'est-à-dire :

$$\mathcal{R}_{\pi, \mathbf{P}}(\xi_\alpha) = 1 - \Re(\vartheta_\alpha).$$

Donc nous sommes en mesure d'associer à chaque vecteur propre ξ de $\mathbf{P}$ et plus généralement à tout vecteur propre ξ_α d'une matrice de transition donnée $\mathbf{P}_\alpha \in \bar{\mathcal{P}}$, une valeur ω caractérisant sa variation que nous appelons fréquence exprimée intuitivement de la manière suivante :

$$\omega = 1 - \Re(\vartheta), \quad \omega \in [0, 2]. \quad (4.26)$$

Sur le contenu des sous espaces propres associés à la marche aléatoire

Les valeurs propres d'une matrice de transition ergodique $\mathbf{P}$ associées à une marche aléatoire sur un graphe dirigé fortement connexe $\mathcal{G}$ sont soit réelles, soit des paires complexes conjuguées $(\vartheta, \bar{\vartheta}) \in \mathbb{C}^2$, $\vartheta = \alpha + i\beta$. Dans ce qui suit, nous examinons comment ces espaces propres individuels peuvent être compris les uns par rapport aux autres.

Considérons d'abord les paires d'espaces propres liés aux valeurs propres conjuguées complexes $\alpha \pm i\beta$. En supposant que les espaces propres ont une dimension égale à 1, les vecteurs propres peuvent être choisis comme complexes conjugués $\xi, \bar{\xi} \in \mathbb{C}^N$. Ces vecteurs propres partagent la même fréquence $1 - \alpha$ et peuvent être considérés comme une généralisation des modes de Fourier aux fréquences ω et $-\omega$ en analyse de Fourier classique unidimensionnelle. On peut également définir les modes de Fourier correspondants en valeur réelle :

$$\xi^{\cos} = \frac{\xi + \bar{\xi}}{2}, \quad \xi^{\sin} = \frac{\xi - \bar{\xi}}{2i},$$

qui généralisent les fonctions cosinus et sinus pour représenter les signaux sur graphes. Comme pour la transformée de Fourier discrète, il y a aussi des fréquences pour lesquelles le sous-espace de fréquence a une dimension égale à un (la fréquence zéro et éventuellement la fréquence 1/2). Le cas des valeurs

propres réelles dans le traitement de signal sur graphe peut être considéré comme une généralisation de ces valeurs.

Lorsque les espaces propres ont une dimension supérieure à 1, la situation est similaire à l'analyse de Fourier classique multidimensionnelle où la fréquence est un vecteur $\boldsymbol{\omega} = [\omega_1, \dots, \omega_n]^\top$.

Considérons maintenant deux valeurs propres avec la même partie réelle mais des parties imaginaires non conjuguées différentes. En poursuivant l'analogie avec l'analyse de Fourier classique, en deux dimensions les modes de Fourier avec une fréquence donnée ω peuvent être considérés comme des combinaisons linéaires de, disons, modes de Fourier vertical et horizontal avec la même fréquence, qui sont deux *différentes orientations*. Ceci est illustré dans la section 5.4.2 sur le graphe toroïdal dirigé.

5.3 Filtres sur graphes de type marche aléatoire

Dans notre cadre théorique, l'opérateur de référence est l'opérateur de marche aléatoire $\mathbf{P}$. En supposant $\mathbf{P}$ diagonalisable, nous pouvons représenter un filtre sur graphe $\mathbf{H}$ comme une combinaison linéaire des projecteurs spectraux $\{\mathbf{E}_k\}_{k=1}^m$ associés à $\mathbf{P}$, comme défini à l'équation (4.1), c.-à-d. $\mathbf{R} = \mathbf{P}$,

$$\mathbf{H} = \sum_{k=1}^m \gamma_k \mathbf{E}_{\vartheta_k}, \quad \gamma_k \in \mathbb{C}, \quad k = 1, \dots, m.$$

Nous définissons également un filtre sur graphe de type marche aléatoire $\mathbf{H}_\omega$ de la manière suivante :

$$\mathbf{H}_\omega = \sum_{\omega \in \boldsymbol{\omega}} \tau_\omega \mathbf{S}_\omega, \quad \tau_\omega \in \mathbb{C}.$$

où chaque τ_ω est la réponse dépendant seulement de la fréquence $\omega \in \boldsymbol{\omega}$ et $\mathbf{S}_\omega$ est le projecteur associé à la fréquence ω défini de la manière suivante :

$$\mathbf{S}_\omega = \sum_{\vartheta: 1 - \Re(\vartheta) = \omega} \mathbf{E}_\vartheta. \quad (4.27)$$

Soit $h : \boldsymbol{\omega} \mapsto \mathbb{R}$ (ou $\mathbb{C}$) la réponse en fréquence associé à la fréquence $\omega \in \boldsymbol{\omega}$. Nous sommes également en mesure de construire un filtre sur graphe de type marche aléatoire avec une réponse en fréquence $\mathbf{H}_\omega$

$$\mathbf{H}_\omega = \sum_{\omega \in \boldsymbol{\omega}} h(\omega) \mathbf{S}_\omega, \quad \omega \in [0, 2]. \quad (4.28)$$

Cette dernière définition permet de jouer avec la réponse en fréquence en la translatant, la contractant ou la dilatant suivant l'axe de fréquence sur graphe.

5.4 Analyse de Fourier sur les groupes finis : un point de vue traitement de signal sur graphe

Afin de montrer la cohérence de notre analyse de Fourier par rapport au cas du traitement de signal classique, nous décrivons les propriétés des versions dirigées de deux objets bien connus : le graphe cyclique associé au groupe cyclique $(\mathbb{Z}/n\mathbb{Z})$ et le graphe toroïdal étant le produit direct de groupes cycliques c.-à-d. $\mathbb{T} = (\mathbb{Z}/n_1\mathbb{Z}) \oplus \cdots \oplus (\mathbb{Z}/n_r\mathbb{Z})$ [67].

5.4.1 Analyse de Fourier sur le graphe cyclique dirigé

Un graphe cyclique dirigé $\mathcal{C}_N$ est un graphe avec N sommets qui forme un cycle simple à travers tous les sommets et dont tous les arcs sont dirigés dans la même direction. Le graphe cyclique dirigé $\mathcal{C}_N$ est représenté par sa matrice d'adjacence $\mathbf{C}_N$ qui est circulante.

$$\mathbf{C}_N = \begin{bmatrix} 0 & 0 & \dots & 0 & 1 \\ 1 & 0 & \dots & 0 & 0 \\ 0 & \ddots & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & 0 & 0 \\ 0 & \dots & 0 & 1 & 0 \end{bmatrix}$$

Comme la matrice des degrés sortants de $\mathcal{C}_N$ est $\mathbf{D} = \mathbf{I}_N$, nous définissons une marche aléatoire $\mathcal{X}$ sur $\mathcal{C}_N$ dont la matrice de transition est caractérisée par sa matrice d'adjacence, c.-à-d. $\mathbf{P} = \mathbf{C}_N$. La matrice de transition $\mathbf{P}$ est diagonalisable, c.-à-d. $\mathbf{P} = \mathbf{\Xi}\mathbf{\Theta}\mathbf{\Xi}^{-1}$ avec $\mathbf{\Theta} = \text{diag}\{\vartheta_1, \dots, \vartheta_N\} \in \mathbb{C}^{N \times N}$ la matrice diagonale des valeurs propres où :

$$\vartheta_k = e^{2\pi i(k-1)/N}, \quad k = 1, \dots, N.$$

et $\mathbf{\Xi} = [\boldsymbol{\xi}_1, \dots, \boldsymbol{\xi}_N] \in \mathbb{C}^{N \times N}$ est la base de Fourier discrète

$$\boldsymbol{\xi}_k = \frac{1}{\sqrt{N}} [1, \vartheta_k^1, \dots, \vartheta_k^{N-1}]^\top \quad k = 1, \dots, N.$$

La marche aléatoire $\mathcal{X}$ est irréductible et périodique. Comme notre analyse fréquentielle n'est applicable que dans le cadre de marches aléatoires ergodiques, nous avons besoin que $\mathcal{X}$ soit apériodique. Afin de surmonter ce problème de périodicité, nous considérons la marche aléatoire paresseuse d'ordre γ , c'est-à-dire $\tilde{\mathcal{X}}$ est caractérisée par sa matrice de transition $\tilde{\mathbf{P}}_\gamma \in \tilde{\mathcal{P}}$ définie à la section 3.2.1 du chapitre 4. Par conséquent, le quotient de Rayleigh associé à un vecteur propre donné $\boldsymbol{\xi}_k$ est donné par :

$$\mathcal{R}_{\pi, \tilde{\mathbf{P}}_\gamma}(\boldsymbol{\xi}_k) = (1 - \gamma) \left[1 - \cos \left(\frac{2\pi(k-1)}{N} \right) \right] = \omega_k, \quad k = 1, \dots, N.$$

Ainsi, l'ordre fréquentiel $\{\omega_k\}_{k=1}^N$ associé aux vecteurs propres $\{\xi_k\}_{k=1}^N$ par le quotient de Rayleigh coïncide exactement avec l'approche classique du traitement de signal.

En prenant la limite $\gamma \rightarrow 0$, on pourrait définir des fréquences pour la marche aléatoire non paresseuse. Plus généralement, notre analyse de Fourier marche parfaitement avec les marches aléatoires irréductibles mais périodiques un utilisant le même genre de limite.

5.4.2 Analyse de Fourier sur le graphe toroïdal dirigé

Un graphe toroïdal dirigé $\mathcal{T}_{m,n}$ est le produit cartésien de deux graphes cycliques dirigés $\mathcal{C}_m$ et $\mathcal{C}_n$, c'est-à-dire $\mathcal{T}_{m,n} = \mathcal{C}_m \square \mathcal{C}_n$. Nous introduisons les définitions nécessaires à la compréhension de la section.

Définition 4.6 ([68]). Soient $\mathcal{G} = (\mathcal{U}, \mathcal{E})$ et $\mathcal{H} = (\mathcal{V}, \mathcal{F})$ deux graphes dont les ensembles de sommets sont respectivement $\mathcal{U} = (u_1, \dots, u_n)$ et $\mathcal{V} = (v_1, \dots, v_m)$. Le produit cartésien entre $\mathcal{G}$ et $\mathcal{H}$ est le graphe $\mathcal{G} \square \mathcal{H}$ avec l'ensemble de sommets $\mathcal{U} \times \mathcal{V}$ pour lesquels les sommets $x = (u_i, v_j)$ et $y = (u_p, v_q)$ sont adjacents si et seulement si $u_i = u_p$ et $(v_j, v_q) \in \mathcal{F}$ ou $v_j = v_q$ et $(u_i, u_p) \in \mathcal{E}$.

Proposition 4.8 ([68]). Soient $\mathcal{G} = (\mathcal{U}, \mathcal{E})$ et $\mathcal{H} = (\mathcal{V}, \mathcal{F})$ deux graphes. La matrice d'adjacence du produit cartésien $\mathcal{G} \square \mathcal{H}$, notée $\mathbf{A}_{\mathcal{G} \square \mathcal{H}}$ est

$$\mathbf{A}_{\mathcal{G} \square \mathcal{H}} = \mathbf{A}_{\mathcal{G}} \otimes \mathbf{I}_{|\mathcal{V}|} + \mathbf{I}_{|\mathcal{U}|} \otimes \mathbf{A}_{\mathcal{H}}.$$

où $\otimes$ symbolise le produit de Kronecker.

Lemme 4.1. Soient $\lambda_1, \dots, \lambda_n$ les valeurs propres de $\mathbf{A}_{\mathcal{G}}$ et $\mu_1, \dots, \mu_m$ les valeurs propres de $\mathbf{A}_{\mathcal{H}}$. Alors les valeurs propres de $\mathbf{A}_{\mathcal{G} \square \mathcal{H}}$ sont toutes les valeurs propres $\lambda_i + \mu_j$ pour tout $1 \leq i \leq n$ et $1 \leq j \leq m$. Par ailleurs, si u et v sont des vecteurs propres de $\mathbf{A}_{\mathcal{G}}$ et $\mathbf{A}_{\mathcal{H}}$ avec valeurs propres λ et μ respectivement, alors le vecteur $\mathfrak{w} = u \otimes v$ est un vecteur propre de $\mathbf{A}_{\mathcal{G} \square \mathcal{H}}$ avec valeur propre $\lambda + \mu$.

Graphe toroïdal dirigé

Soit $\mathcal{T}_{m,n} = \mathcal{C}_m \square \mathcal{C}_n$ le graphe toroïdal dirigé caractérisé par sa matrice d'adjacence $\mathbf{A}_{\mathcal{T}_{m,n}} = \mathbf{C}_m \otimes \mathbf{I}_n + \mathbf{I}_m \otimes \mathbf{C}_n$. Les valeurs propres de $\mathbf{C}_m$ sont $\lambda_k = e^{2i\pi(k-1)/m}$ pour $k \in \llbracket 1, m \rrbracket$ et les vecteurs propres sont indiqués par $[\phi_1, \dots, \phi_m]$. De même, les valeurs propres de $\mathbf{C}_n$ sont $\mu_\ell = e^{2i\pi(\ell-1)/n}$ pour $\ell \in \llbracket 1, n \rrbracket$ et les vecteurs propres sont indiqués par $[\psi_1, \dots, \psi_n]$. Le graphe toroïdal dirigé $\mathcal{T}_{m,n}$ est un graphe régulier de degré 2. Par conséquent, la matrice des degrés sortants de $\mathcal{T}_{m,n}$ est $\mathbf{D}_{\mathcal{T}_{m,n}} = 2\mathbf{I}_{m \times n}$. Nous définissons donc une marche aléatoire $\mathcal{X}$ sur $\mathcal{T}_{m,n}$ avec la matrice de transition diagonalisable $\mathbf{P} = \mathbf{D}_{\mathcal{T}_{m,n}}^{-1} \mathbf{A}_{\mathcal{T}_{m,n}}$. Nous spécifions ses propriétés spectrales dans le lemme suivant :

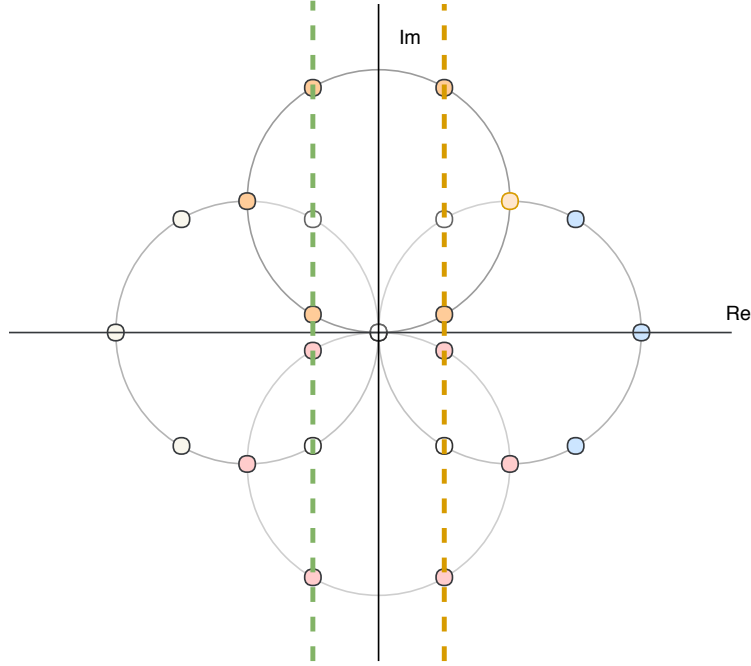


FIGURE 4.2 – Distribution des valeurs propres de la matrice de marche aléatoire d'un graphe toroïdal dirigé $\mathcal{T}_{6,4}$. Valeurs propres sur la ligne verte pointillée : $\Re(\lambda^g) = -0.25, \forall \lambda^g \in \Lambda_g$. Valeurs propres sur la ligne orange pointillée : $\Re(\lambda^o) = 0.25, \forall \lambda^o \in \Lambda_o$.

Lemme 4.2. Soit $\mathcal{T}_{m,n}$ un graphe toroïdal dirigé et $\mathcal{X}$ la marche aléatoire associée à une matrice de transition diagonalisable $\mathbf{P}$. Le graphe toroïdal dirigé $\mathcal{T}_{m,n}$ est un graphe dirigé régulier de degré 2. Par conséquent, le spectre de $\mathbf{P}$ forme l'ensemble

$$\sigma(\mathbf{P}) = \left\{ \zeta_{i,j} = \frac{\lambda_i + \mu_j}{2}, i \in \llbracket 1, m \rrbracket, j \in \llbracket 1, n \rrbracket \right\},$$

avec les vecteurs propres associés

$$\left\{ \phi_i \otimes \psi_j, i \in \llbracket 1, m \rrbracket, j \in \llbracket 1, n \rrbracket \right\}.$$

Comme dans la section 5.4.1, la marche aléatoire associée $\mathcal{X}$ est irréductible et périodique. Afin de surmonter ce problème de périodicité, nous considérons une marche aléatoire paresseuse d'ordre γ , $\tilde{\mathcal{X}}$, avec matrice de transition $\tilde{\mathbf{P}}_\gamma \in \tilde{\mathcal{P}}$ et distribution stationnaire π .

La figure 4.2 illustre la distribution des valeurs propres de la matrice de transition d'un graphe toroïdal dirigé. Comme nous pouvons l'observer, la distribution des valeurs propres est particulière. Nous désignons par Λ_g ,

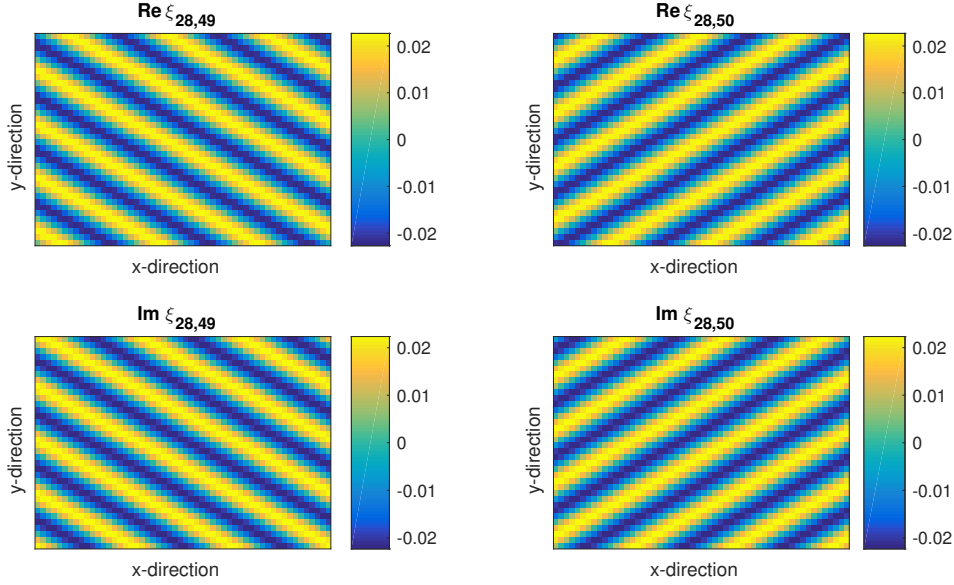


FIGURE 4.3 – Représentation bidimensionnelle des vecteurs propres $\xi_{28,49}$ et $\xi_{28,50}$ de la matrice de transition du toroïdal dirigé $\mathcal{T}_{54,36}$. L'axe est l'étiquette d'un noeud dans $\mathbf{C}_m$ (resp. $\mathbf{C}_m$) dans la direction x (resp. direction y) .

l'ensemble des valeurs propres se situant sur la ligne pointillée verte et Λ_o l'ensemble des valeurs propres se situant sur la ligne pointillée orange. Les ensembles Λ_g , et Λ_o , incluent respectivement des valeurs propres ayant la même partie réelle et différentes parties imaginaires c.-à-d. $\Re(\Lambda_g) = -0.25$ et $\Re(\Lambda_o) = 0.25$. Afin d'illustrer les idées discutées à la section 5.2, nous exposons sur la figure 4.3 la représentation bidimensionnelle de certains vecteurs propres de la matrice de transition d'un graphe toroïdal dirigé $\mathcal{T}_{54,36}$. Les vecteurs propres sélectionnés $\xi_{28,49}$ et $\xi_{28,50}$ sont associés à des valeurs propres ayant la même partie réelle, c'est-à-dire $\Re(\lambda(\xi_{28,49})) = \Re(\lambda(\xi_{28,50}))$ mais dont les parties imaginaires sont non conjuguées, c'est-à-dire $\Im(\lambda(\xi_{28,49})) \neq \pm \Im(\lambda(\xi_{28,50}))$. Analytiquement, les quotients de Rayleigh des vecteurs propres sélectionnés sont égaux, comme expliqué dans la section 5.2. Comme on peut le voir sur la figure 4.3, les vecteurs propres sélectionnés ont une différence d'orientation. Cette différence d'orientation est associée à l'influence des parties imaginaires des valeurs propres des vecteurs propres sélectionnés.

6 Applications

Cette section est consacrée à l'illustration de notre cadre théorique à travers les exemples d'apprentissage semi-supervisé et de modélisation de

signal sur graphes dirigés. Dans notre cadre théorique, il s'avère que dans les applications suivantes, nous pouvons travailler soit dans l'espace de Hilbert $\ell^2(\mathcal{V})$ soit dans l'espace $\ell^2(\mathcal{V}, \pi)$ où π est la mesure stationnaire d'une marche aléatoire ergodique sur graphe (dirigé ou non). Par ailleurs, par les résultats obtenus à la section 3.4, il existe une isométrie φ nous permettant de passer de $\ell^2(\mathcal{V})$ à $\ell^2(\mathcal{V}, \pi)$. Par souci de simplicité, nous nous ramènerons toujours à $\ell^2(\mathcal{V})$. Il est plus pratique de travailler dans $\ell^2(\mathcal{V})$ car le produit scalaire associé correspond au produit scalaire usuel entre des vecteurs de $\mathbb{C}^N$. Soit un signal donné. On décide dans quel espace vit le signal. On fait donc un choix arbitraire entre le fait que le signal vit dans un espace associé à la mesure de comptage ou bien dans un espace associé à la mesure stationnaire d'une marche aléatoire. Si on suppose que le signal vit dans $\ell^2(\mathcal{V}, \pi)$, on convertit le signal dans $\ell^2(\mathcal{V})$ en utilisant l'application φ^{-1} définie de la manière suivante :

$$\varphi^{-1} : \mathbf{f} \mapsto \mathbf{\Pi}^{1/2} \mathbf{f}, \quad \forall \mathbf{f} \in \ell^2(\mathcal{V}, \pi). \quad (4.29)$$

On effectue toutes les opérations nécessaires dans $\ell^2(\mathcal{V})$. Nous retournons ensuite dans l'espace $\ell^2(\mathcal{V}, \pi)$ dans lequel le signal vivait originellement. Dans $\ell^2(\mathcal{V})$, l'opérateur $\mathbf{T}$ défini en (4.23) correspond à l'opérateur de marche aléatoire et $\mathcal{L}$ est l'opérateur tel que $\langle \mathbf{f}, \mathcal{L} \mathbf{f} \rangle$ est l'énergie de Dirichlet.

6.1 Apprentissage semi-supervisé sur des graphes dirigés via régularisation de type ℓ_2

Nous discutons du problème d'apprentissage semi-supervisé sur graphe dirigé avec terme de régularisation de type ℓ_2 . Le problème suivant vise à montrer l'efficacité et la pertinence de l'énergie de Dirichlet définie en (4.25) comme terme de régularisation pour les signaux définis sur des graphes dirigés. Nous formulons le problème d'apprentissage semi-supervisé sur graphe de la manière suivante.

6.1.1 Définition du problème

Soit $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ un graphe dirigé fortement connexe où $\mathcal{V} = \{v_j\}_{j=1}^N$ est l'ensemble des sommets, $\mathcal{E}$ est l'ensemble des arcs et de cardinalité $|\mathcal{V}| = N$. Le graphe dirigé $\mathcal{G}$ est représenté par la matrice d'adjacence $\mathbf{W} \in \mathbb{R}_+^{N \times N}$. Soit $\mathcal{X}$ la marche aléatoire sur $\mathcal{G}$. La marche aléatoire $\mathcal{X}$ est représentée par son opérateur de marche aléatoire $\mathbf{P} \in \mathbb{R}_+^{N \times N}$ d'unique distribution stationnaire π . Nous introduisons également $\mathcal{G}_{\text{sym}}$ la version symétrisée du graphe dirigé $\mathcal{G}$. Le graphe non dirigé $\mathcal{G}_{\text{sym}}$ est représenté par sa matrice d'adjacence $\mathbf{W}_{\text{sym}} = (\mathbf{W} + \mathbf{W}^\top)/2 \in \mathbb{R}_+^{N \times N}$. Soit $\mathcal{X}_{\text{sym}}$ la marche aléatoire sur $\mathcal{G}_{\text{sym}}$. La marche aléatoire $\mathcal{X}_{\text{sym}}$ est représentée par son opérateur de marche aléatoire $\mathbf{P}_{\text{sym}} \in \mathbb{R}_+^{N \times N}$ avec unique distribution stationnaire π_{sym} .

Soit y une fonction étiquette définie sur l'ensemble $\mathcal{V}$ du graphe $\mathcal{G}$ comme $y : \mathcal{V} \rightarrow \{-1, 1\}$. On peut aussi voir un signal sur graphe $\mathbf{y}$, écrit comme le vecteur :

$$\mathbf{y} = [y(v_1), \dots, y(v_N)]^\top.$$

Etant donné les valeurs des étiquettes $y(v_k)_{\{1 \leq k \leq N\}}$ sur seulement un sous-ensemble de sommets étiquetés $\mathcal{U} \subset \mathcal{V}$, le but est d'estimer les étiquettes des sommets non marqués restants. Il s'agit du problème d'apprentissage semi-supervisé sur graphes [69, 9]. Nous introduisons la formulation du problème dans la section suivante.

6.1.2 Résolution du problème

Le problème d'apprentissage semi-supervisé sur graphes est formulé de la manière suivante :

$$\operatorname{argmin}_{\mathbf{f} \in \mathbb{R}^N} \{c \|\mathbf{M}_l(\mathbf{f} - \mathbf{y})\|^2 + \varrho_1 \|\mathbf{M}_u \mathbf{f}\|^2 + \varrho_2 \mathcal{S}(\mathbf{f})\}, \quad c, \varrho_1, \varrho_2 > 0. \quad (4.30)$$

Le terme $\|\mathbf{M}_l(\mathbf{f} - \mathbf{y})\|^2$ est le terme de fidélité aux données étiquetées où $\mathbf{M}_l = \operatorname{diag}\{m_1, \dots, m_N\}$ est une matrice diagonale avec $m_i = \mathbb{1}_{v_i \in \mathcal{U}}$, pour tous les $i = 1 \dots, N$ (donc 0 sur les sommets avec étiquettes inconnues). Dans le second terme $\|\mathbf{M}_u \mathbf{f}\|^2$ apparaît $\mathbf{M}_u$ défini par $\mathbf{M}_u = \mathbf{I}_N - \mathbf{M}_l$ et le troisième terme, $\mathcal{S}(\mathbf{f})$ est le terme de régularisation variationnelle défini ainsi

$$\mathcal{S}(\mathbf{f}) = \langle \mathbf{f}, \mathbf{X} \mathbf{f} \rangle, \quad \mathbf{X} \in \mathbb{R}^{N \times N}$$

Le problème d'apprentissage semi-supervisé sur graphes est formulé de la même manière pour le graphe dirigé $\mathcal{G}$ ou sa version symétrisée $\mathcal{G}_{\text{sym}}$, excepté que le terme de régularisation variationnelle $\mathcal{S}(\mathbf{f})$ diffère. Le problème (4.30) est quadratique et convexe et admet donc une solution analytique. Dans la prochaine section, nous présentons trois approches permettant de résoudre le problème (4.30).

6.1.3 Méthodes

Dans cette section, nous présentons trois approches afin de résoudre le problème d'apprentissage semi-supervisé sur graphes dirigés.

Méthode 1

Ici, nous introduisons une modification de (4.30), précédemment étudiée par [69]. Dans cette approche, nous supposons le signal sur graphe $\mathbf{y} \in \ell^2(\mathcal{V})$. Le problème à résoudre est formulé de la manière suivante :

$$\operatorname{argmin}_{\mathbf{f} \in \ell^2(\mathcal{V})} \{c \|\mathbf{M}_l(\mathbf{f} - \mathbf{y})\|^2 + c \|\mathbf{M}_u \mathbf{f}\|^2 + \varrho_2 \mathcal{S}_{\mathcal{L}}(\mathbf{f})\}, \quad c, \varrho_2 > 0 \quad (4.31)$$

où $\mathcal{S}_{\mathcal{L}}(\mathbf{f}) = \langle \mathbf{f}, \mathcal{L}\mathbf{f} \rangle$ est l'énergie de Dirichlet dans $\ell^2(\mathcal{V})$.

La solution de (4.31), pour $\mathbf{f} \in \ell^2(\mathcal{V})$ est donc donnée par :

$$\mathbf{f}^* = \text{signe}[(\mathbf{I} + \gamma\mathcal{L})^{-1}\mathbf{M}_l\mathbf{y}], \quad \gamma = \varrho_2/c, \quad \mathbf{f}^* \in \ell^2(\mathcal{V}). \quad (4.32)$$

Nous décidons également de traiter la même approche, sous la condition où nous travaillons avec la version symétrisée du graphe dirigé $\mathcal{G}$, c.-à-d. $\mathcal{G}_{\text{sym}}$.

Le problème dans le cas du graphe symétrisé $\mathcal{G}_{\text{sym}}$ est :

$$\underset{\mathbf{f} \in \ell^2(\mathcal{V})}{\text{argmin}} \{c\|(\mathbf{M}_l(\mathbf{f} - \mathbf{y}))\|^2 + c\|\mathbf{M}_u\mathbf{f}\|^2 + \varrho_2\mathcal{S}_{\mathcal{L}_{\text{sym}}}(\mathbf{f})\}, \quad c, \varrho_1, \varrho_2 > 0, \quad (4.33)$$

où $\mathcal{S}_{\mathcal{L}_{\text{sym}}}(\mathbf{f}) = \langle \mathbf{f}, \mathcal{L}_{\text{sym}}\mathbf{f} \rangle$ et $\mathcal{L}_{\text{sym}}$ est le laplacien normalisé associé :

$$\mathcal{L}_{\text{sym}} = \mathbf{I} - \mathbf{\Pi}_{\text{sym}}^{1/2}\mathbf{P}_{\text{sym}}\mathbf{\Pi}_{\text{sym}}^{-1/2}, \quad \mathbf{\Pi}_{\text{sym}} = \text{diag}\{\pi_{\text{sym}}(v_1), \dots, \pi_{\text{sym}}(v_N)\}.$$

La solution de (4.33), pour $\mathbf{f} \in \ell^2(\mathcal{V})$ est donc donnée par :

$$\mathbf{f}^* = \text{signe}[(\mathbf{I} + \gamma\mathcal{L}_{\text{sym}})^{-1}\mathbf{M}_l\mathbf{y}], \quad \gamma = \varrho_2/c, \quad \mathbf{f}^* \in \ell^2(\mathcal{V}). \quad (4.34)$$

Méthode 2

Dans cette méthode, nous utilisons la formulation (4.31), excepté que le signal sur graphe $\mathbf{y} \in \ell^2(\mathcal{V}, \pi)$. Nous définissons $\tilde{\mathbf{y}} \in \ell^2(\mathcal{V})$ de la manière suivante :

$$\tilde{\mathbf{y}} = \varphi^{-1}(\mathbf{y}) = \mathbf{\Pi}^{1/2}\mathbf{y}, \quad \mathbf{y} \in \ell^2(\mathcal{V}, \pi).$$

Il nous suffit de résoudre problème (4.31) avec $\tilde{\mathbf{f}}$. Si nous remplaçons $\tilde{\mathbf{y}}$ dans l'équation (4.31), nous obtenons le problème d'optimisation suivant :

$$\underset{\tilde{\mathbf{f}} \in \ell^2(\mathcal{V})}{\text{argmin}} \{c\|\mathbf{M}_l(\tilde{\mathbf{f}} - \tilde{\mathbf{y}})\|^2 + c\|\mathbf{M}_u\tilde{\mathbf{f}}\|^2 + \varrho_2\mathcal{S}_{\mathcal{L}}(\tilde{\mathbf{f}})\}, \quad c, \varrho_2 > 0 \quad (4.35)$$

avec $\tilde{\mathbf{f}} = \varphi^{-1}(\mathbf{f})$. En remplaçant $\tilde{\mathbf{y}}$ par son expression, nous obtenons :

$$\underset{\mathbf{f} \in \ell^2(\mathcal{V}, \pi)}{\text{argmin}} \{c\|\mathbf{M}_l(\mathbf{f} - \mathbf{y})\|_{\pi}^2 + c\|\mathbf{M}_u\mathbf{f}\|_{\pi}^2 + \varrho_2\mathcal{S}_{\mathbf{L}_{\text{RW}}}(\mathbf{f})\}, \quad c, \varrho_2 > 0 \quad (4.36)$$

où $\mathcal{S}_{\mathbf{L}_{\text{RW}}}(\mathbf{f}) = \langle \mathbf{f}, \mathbf{L}_{\text{RW}}\mathbf{f} \rangle_{\pi}$, $\mathbf{f} \in \ell^2(\mathcal{V}, \pi)$.

La solution du problème (4.35) est donc :

$$\tilde{\mathbf{f}}^* = \text{signe}[(\mathbf{I} + \gamma\mathcal{L})^{-1}\mathbf{M}_l\tilde{\mathbf{y}}], \quad \gamma = \varrho_2/c, \quad \tilde{\mathbf{f}}^* \in \ell^2(\mathcal{V}). \quad (4.37)$$

Le signal sur graphe $\mathbf{y}$ appartenant à $\ell^2(\mathcal{V}, \pi)$, nous appliquons sur $\tilde{\mathbf{f}}^*$ l'application isométrique φ afin de revenir dans $\ell^2(\mathcal{V}, \pi)$. Par conséquent le résultat final est :

$$\mathbf{f}_1^* = \text{signe} [\varphi((\mathbf{I} + \gamma\mathcal{L})^{-1}\mathbf{M}_l\tilde{\mathbf{y}})].$$

Si nous remplaçons $\tilde{\mathbf{y}}$ par son expression, nous obtenons :

$$\mathbf{f}^* = \text{signe} [(\mathbf{I} + \gamma\mathbf{L}_{\text{RW}})^{-1}\mathbf{M}_l\mathbf{y}], \quad \gamma = \varrho_2/c, \quad \mathbf{f}^* \in \ell^2(\mathcal{V}, \pi). \quad (4.38)$$

Tout comme la méthode 1, nous décidons également de traiter la méthode 2, sous la condition où nous travaillons avec la version symétrisée du graphe dirigé $\mathcal{G}$.

Nous avons donc le problème suivant :

$$\underset{\tilde{\mathbf{f}} \in \ell^2(\mathcal{V})}{\text{argmin}} \{c\|\mathbf{M}_l(\tilde{\mathbf{f}} - \tilde{\mathbf{y}})\|^2 + c\|\mathbf{M}_u\tilde{\mathbf{f}}\|^2 + \varrho_2\mathcal{S}_{\mathcal{L}_{\text{sym}}}(\tilde{\mathbf{f}})\}, \quad c, \varrho_2 > 0 \quad (4.39)$$

avec $\tilde{\mathbf{f}} = \varphi^{-1}(\mathbf{f})$. En remplaçant $\tilde{\mathbf{y}}$ par son expression, nous obtenons :

$$\underset{\mathbf{f} \in \ell^2(\mathcal{V}, \pi_{\text{sym}})}{\text{argmin}} \{c\|\mathbf{M}_l(\mathbf{f} - \mathbf{y})\|_{\pi_{\text{sym}}}^2 + c\|\mathbf{M}_u\mathbf{f}\|_{\pi_{\text{sym}}}^2 + \varrho_2\mathcal{S}_{\mathbf{L}_{\text{RW}, \text{sym}}}(\mathbf{f})\}, \quad c, \varrho_2 > 0 \quad (4.40)$$

où $\mathcal{S}_{\mathbf{L}_{\text{RW}, \text{sym}}}(\mathbf{f}) = \langle \mathbf{f}, \mathbf{L}_{\text{RW}, \text{sym}}\mathbf{f} \rangle$, $\mathbf{f} \in \ell^2(\mathcal{V}, \pi_{\text{sym}})$ et $\mathbf{L}_{\text{RW}, \text{sym}}$ est le Laplacien marche aléatoire associé défini de la manière suivante :

$$\mathbf{L}_{\text{RW}, \text{sym}} = \mathbf{I} - \mathbf{P}_{\text{sym}}.$$

La résultat final est donc :

$$\mathbf{f}^* = \text{signe} [\varphi((\mathbf{I} + \gamma\mathcal{L}_{\text{sym}})^{-1}\mathbf{M}_l\tilde{\mathbf{y}})].$$

Méthode 3

Nous comparons les méthodes 1 et 2 à la méthode proposée en traitement de signal sur graphe [20]. Le même problème est résolu pour les graphes dirigés en utilisant (4.30), écrit en [20] de la manière suivante :

$$\underset{\mathbf{f}}{\text{argmin}} \{c\|\mathbf{M}_l(\mathbf{f} - \mathbf{y})\|^2 + \varrho_2\mathcal{S}_{\mathbf{R}_M}(\mathbf{f})\}, \quad c, \varrho_2 > 0, \quad (4.41)$$

où $\mathcal{S}_{\mathbf{R}_M}(\mathbf{f}) = \langle \mathbf{f}, \mathbf{R}_M\mathbf{f} \rangle$ et $\mathbf{R}_M = (\mathbf{I} - \mathbf{W}^{\text{norm}})^\top(\mathbf{I} - \mathbf{W}^{\text{norm}})$, $\mathbf{W}^{\text{norm}}$ étant une version normalisée de la matrice adjacence telle que sa norme spectrale soit égale à 1.

La solution de (4.41) est donnée par :

$$\mathbf{f}^* = \text{signe} [(\mathbf{M}_l + \gamma\mathbf{R}_M^{-1})\mathbf{y}], \quad \gamma = \varrho_2/c, \quad \mathbf{f}^* \in \ell^2(\mathcal{V}).$$

Nous résumons les différentes méthodes dans la table 4.1.

Méthode	$\mathcal{G}$	$\mathcal{G}_{\text{sym}}$
Méthode 1	$\mathcal{L}$	$\mathcal{L}_{\text{sym}}$
Méthode 2	$\mathbf{L}_{\text{RW}}$	$\mathbf{L}_{\text{RW},\text{sym}}$
Méthode 3	$\mathbf{R}_M$ [44]	

TABLE 4.1 – Table des différents opérateurs impliqués dans les méthodes d'apprentissage semi-supervisé sur graphe selon $\mathcal{G}$ et $\mathcal{G}_{\text{sym}}$.

6.1.4 Expériences

Considérons le jeu de données des blogs politiques de la campagne présidentielle américaine de 2004 [70]. L'ensemble de données se compose de 1224 blogs politiques où chaque blog politique est associé à une orientation politique, qu'elle soit républicaine ou démocrate. Cet ensemble de données est modélisé par un graphe $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ où chaque sommet $v \in \mathcal{V}$ est associé à un blog et un arc entre deux sommets $\{v_i, v_j\}$ indique la présence d'un lien du blog associé au sommet v_i vers le sommet v_j . Les orientations politiques des blogs sont modélisées par un signal sur graphe $\mathbf{f}_0 = \{f_0(v_1), \dots, f_0(v_N)\}$ où $f_0(v_i) \in \mathcal{O}$ où chaque sommet est associé à une étiquette appartient au jeu $\mathcal{O} = \{-1, 1\}$ caractérisant son orientation politique, l'orientation politique associée aux Démocrates étant "+1" et l'orientation politique associée aux Républicains étant "-1". Le graphe dirigé $\mathcal{G}$ associé aux blogs politiques n'est pas fortement connexe. Notre cadre théorique nous amène à ne considérer que des graphes fortement connexes. Par conséquent, nous considérons le plus grand sous-graphe fortement connexe de $\mathcal{G}$, noté par $\mathcal{G}' = (\mathcal{V}', \mathcal{E}')$ qui est composé de $|\mathcal{V}'| = N' = 793$ sommets (donc environ 65 % de l'ensemble de sommets $\mathcal{V}$) et son signal sur graphe associé $\mathbf{f}'_0$. Les performances respectives des approches (4.31), (4.35), (4.41) sur $\mathcal{G}'$ et de leurs équivalents sur sa version symétrisée, c.-à-d. (4.33) et (4.39) sont comparées en figure 4.4.

Simulations numériques

La figure 4.4 affiche la performance des approches semi-supervisées sur graphe (4.31), (4.35), (4.41), (4.33) et (4.39). Le taux de performance est une moyenne de 500 réalisations calculée en déterminant le paramètre $\gamma \in [0, 10]$ sur l'ensemble des données d'entraînement donnant le meilleur taux de performance. Tout d'abord, nous remarquons que pour toutes les proportions d'étiquettes connues, les taux de performance des approches semi-supervisées basées sur $\mathcal{L}$, $\mathbf{L}_{\text{RW}}$ et ses homologues symétriques $\mathcal{L}_{\text{sym}}$, $\mathbf{L}_{\text{RW},\text{sym}}$ sont nettement meilleurs que les taux de performance de l'approche semi-supervisée basée sur $\mathbf{R}_M$. En haut à gauche de la figure 4.4, nous comparons les approches semi-supervisées impliquant les opérateurs $\mathcal{L}$, $\mathbf{L}_{\text{RW}}$ et $\mathbf{R}_M$. Le taux

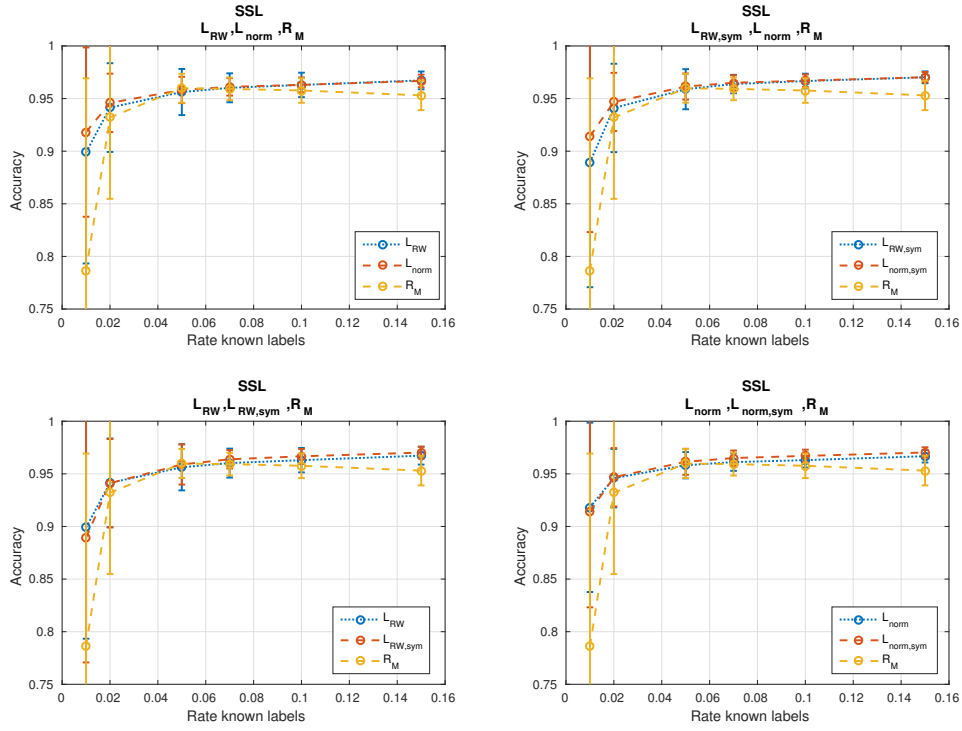


FIGURE 4.4 – Apprentissage semi-supervisé sur $\mathcal{G}'$ des blogs politiques de la campagne présidentielle américaine de 2004. Les notations L_{norm} et $L_{norm,sym}$ correspondent respectivement à $\mathcal{L}$ et $\mathcal{L}_{sym}$.

de performance de l'approche semi-supervisée associée à $\mathcal{L}$ est le meilleur comparativement aux approches impliquant $\mathbf{L}_{RW}, \mathbf{R}_M$ pour de faibles taux d'étiquettes connues. De plus, l'approche du taux de performance associée à $\mathcal{L}$ est légèrement meilleure que l'approche impliquant $\mathbf{L}_{RW}$. En haut à droite de la figure 4.4, nous comparons les approches semi-supervisées impliquant les opérateurs $\mathcal{L}_{sym}, \mathbf{L}_{RW,sym}$ et $\mathbf{R}_M$. De même, le taux de performance de l'approche semi-supervisée associée à $\mathcal{L}_{sym}$ est légèrement meilleur que celui de l'approche impliquant $\mathbf{L}_{RW}$. En bas à gauche de la figure 4.4, nous comparons les approches semi-supervisées impliquant les opérateurs $\mathbf{L}_{RW}, \mathbf{L}_{RW,sym}$ et $\mathbf{R}_M$. Les taux de performance entre l'approche impliquant $\mathbf{L}_{RW}$ et $\mathbf{L}_{RW,sym}$ sont presque identiques pour toute proportion d'étiquettes connues ainsi qu'en bas à droite de la figure 4.4 comparant les approches semi-supervisées impliquant $\mathbf{L}_{RW}, \mathbf{L}_{RW,sym}$ et $\mathbf{R}_M$.

La comparaison des taux de performance des approches semi-supervisées impliquant les différents opérateurs listés dans la table 4.1 associées à cet exemple nous amène aux conclusions suivantes :

- Les approches d'apprentissage semi supervisées impliquant des énergies de Dirichlet basées sur $\mathcal{L}$ and $\mathbf{L}_{RW}$ (et leurs équivalents symétriques) donnent de meilleurs taux de performance que l'approche proposée par Sandryhaila et Moura [44] basée sur $\mathbf{R}_M$. Par conséquent, des approches d'apprentissage semi-supervisé sur des graphes dirigés aux énergies de Dirichlet basées sur l'opérateur de marche aléatoire semblent plus appropriées.
- L'espace de Hilbert où l'on considère que le signal sur graphe vit a une influence sur les taux de performance du problème semi-supervisé. Ici, considérer le signal sur graphe $\mathbf{y}$ appartenant à $\ell^2(\mathcal{V})$, ou considérer de manière équivalente l'opérateur $\mathcal{L}$ dans le terme de régularisation donne une meilleure performance que considérer le signal sur graphe $\mathbf{y}$ appartenant à $\ell^2(\mathcal{V}, \pi)$ ou considérer de manière équivalente l'opérateur $\mathcal{L}_{RW}$ dans le terme de régularisation.

6.2 Modélisation du signal sur des graphes dirigés par filtrage

Dans cette section, nous examinons une façon de modéliser les relations entre les valeurs d'un signal sur graphe à l'aide d'un filtre sur graphe. Le modèle est exprimé sous la forme d'un filtre sur graphe qui prend quelques valeurs du signal sur graphe et reconstruit les autres valeurs. Une application possible d'un tel modèle pourrait être la compression avec perte du signal, où seule la connaissance du graphe, des coefficients du filtre sur graphe et de quelques valeurs de signal permettrait de reconstruire l'ensemble du signal. La différence majeure avec la précédente application est que nous supposons le signal sur graphe entier connu.

6.2.1 Formulation du problème

Sachant un signal aléatoire sur graphe $\mathbf{y}$, nous souhaitons déterminer les valeurs manquantes du signal en utilisant un filtre sur graphe $\mathbf{H}$, exprimé comme une somme polynomiale finie d'un opérateur de référence $\mathbf{R} \in \mathbb{R}^{N \times N}$. Autrement dit :

$$\mathbf{H} = \sum_{k=0}^K \theta_k \mathbf{R}^k, \quad \theta_k \in \mathbb{R}, \quad k = 0, \dots, K. \quad (4.42)$$

Les coefficients du filtre sur graphe sont appris en résolvant le problème d'optimisation suivant :

$$\hat{\boldsymbol{\theta}} = \underset{\boldsymbol{\theta} = \{\theta_k\}_{k=0}^K \in \mathbb{R}^{K+1}}{\operatorname{argmin}} \mathbb{E} \left[\|\mathbf{f}_0 - \sum_{k=0}^K \theta_k \mathbf{R}^k \mathbf{y}\|_{\mu}^2 \right], \quad (4.43)$$

où $\mathbf{y} = [y_1, \dots, y_N]^\top$ est une variable aléatoire multivariée indexée sur l'ensemble des sommets $\mathcal{V}$. Les variables aléatoires $y_j, j = 1, \dots, N$ peuvent être exprimées en $y_j = \varepsilon_j f_{0j}$, où les ε_j sont supposés indépendants et Bernoulli distribués avec les paramètres $\delta_j : \varepsilon_j \sim \operatorname{Ber}(\delta_j)$. La variable aléatoire $\bar{\varepsilon}$ définie comme suit :

$$p = \frac{1}{N} \sum_{i=1}^N \varepsilon_i, \quad (4.44)$$

est la proportion des valeurs connues de $\mathbf{f}_0$, toutes collectées en $\mathbf{y}$. Ce signal $\mathbf{y}$ peut être vu comme un échantillonnage aléatoire du signal sur graphe $\mathbf{f}_0$ et $\mathbf{H}\mathbf{y}$ comme une reconstruction de $\mathbf{f}_0$ à partir de ces échantillons. Le problème d'optimisation (4.43) correspond à la minimisation de l'erreur de reconstruction avec un tel ensemble d'échantillons aléatoires.

6.2.2 Résolution du problème

La solution de (4.43) est :

$$\hat{\boldsymbol{\theta}} = \mathbf{Z}^{-1} \mathbf{M} \mathbf{Q}^\top \mathbf{f}_0,$$

où $\mathbf{M} = \operatorname{diag}\{\mu_1, \dots, \mu_N\}$ est la matrice diagonale de la mesure associée μ , $\mathbf{Z}$ la matrice où chaque entrée $Z_{k\ell} \in \mathbf{Z}$ est exprimée par :

$$Z_{k\ell} = \operatorname{Tr}([\mathbf{R}^k]^\top \mathbf{M} \mathbf{R}^\ell \mathbb{E}(\mathbf{y} \mathbf{y}^\top)), \quad \forall \{k, \ell\} \in \llbracket 0, K \rrbracket^2,$$

et $\mathbf{Q} = [\mathbf{q}_0, \dots, \mathbf{q}_K] \in \mathbb{R}^{N \times (K+1)}$ où chaque vecteur $\mathbf{q}_j$ est :

$$\mathbf{q}_j = \mathbf{R}^j \mathbb{E}(\mathbf{y}), \quad \forall j \in \llbracket 0, K \rrbracket.$$

6.2.3 Une exploration de la précision de la modélisation avec différents opérateurs sur graphes

La qualité de la reconstruction selon l'approche précédente dépend de plusieurs facteurs : le choix de l'opérateur de référence $\mathbf{R}$, l'ordre polynomial du filtre K et la stratégie d'échantillonnage aléatoire définie par le paramètre de Bernoulli δ_k . Ici, nous évaluons principalement l'influence de $\mathbf{R}$ et comparons deux stratégies d'échantillonnage possibles. A titre d'exemple, nous considérons également l'ensemble des données des blogs politiques de la campagne présidentielle américaine de 2004 [70], décrit à la section 6.1.4. Les orientations politiques des blogs sont modélisées par un signal sur graphe $\mathbf{f}_0 = \{f_0(v_1), \dots, f_0(v_N)\}$ où chaque sommet v_i a un label $f_0(v_i) \in \{-1, 1\}$ caractérisant son orientation politique. De même que dans la précédente application, nous ne considérons que les graphes fortement connexes. Ainsi nous considérerons le grand sous-graphe fortement connexe de $\mathcal{G}$, noté $\mathcal{G}' = (\mathcal{V}', \mathcal{E}')$ qui est fait de $|\mathcal{V}'| = N' = 793$ sommets et son signal associé $\mathbf{f}'_0$.

Nous comparerons la précision de reconstruction à l'aide d'une proportion d'étiquettes correctement reconstruites obtenues par :

$$\hat{\mathbf{f}} = \text{signe}(\hat{\mathbf{H}}\mathbf{y}).$$

6.2.3.1 Modélisation du signal sur $\mathcal{G}'$ Dans cette section, seule la plus grande composante fortement connexe du graphe $\mathcal{G}'$ est utilisée (voir dans 6.2.4.1 comment étudier le même problème sur $\mathcal{G}$).

Résolvons le problème (4.43) en apprenant un filtre sur graphe polynomial avec $K = 10$ sur $\mathcal{G}'$. Nous considérons différents opérateurs de référence possibles : ceux pour les graphes dirigés, dans l'ensemble $\{\mathbf{A}_{norm}, \mathbf{P}, \bar{\mathbf{P}}, \mathbf{T}, \bar{\mathbf{T}}\}$ où $\mathbf{A}_{norm}$ est une version normalisée de la matrice adjacence dont la norme spectrale est égale à un [20], $\mathbf{P}$ l'opérateur de marche aléatoire, $\bar{\mathbf{P}}$ sa réversible additive, $\mathbf{T} = \mathbf{\Pi}^{1/2}\mathbf{P}\mathbf{\Pi}^{-1/2}$ l'équivalent de $\mathbf{P}$ en $\ell^2(\mathcal{V})$ et $\bar{\mathbf{T}} = \mathbf{\Pi}^{1/2}\bar{\mathbf{P}}\mathbf{\Pi}^{-1/2}$; et ceux pour les graphes non dirigés (ou symétriques), dans l'ensemble $\{\mathbf{A}_{sym}, \mathbf{P}_{sym}, \mathbf{T}_{sym}\}$ où $\mathbf{A}_{sym}$ est la version normalisée de la matrice adjacence symétrique $(\mathbf{A} + \mathbf{A}^\top)/2$, $\mathbf{P}_{sym}$ l'opérateur de marche aléatoire basé sur $\mathbf{A}_{sym}$ et $\mathbf{T}_{sym} = \mathbf{\Pi}_{sym}^{1/2}\mathbf{P}_{sym}\mathbf{\Pi}_{sym}^{-1/2}$.

Les variables aléatoires $y_j, j = 1, \dots, N'$ sont réparties dans l'un des deux cas suivants :

1. Variables aléatoires $y_j = \varepsilon_j f'_0(v_j)$ où $\varepsilon_j \sim \text{Ber}(p)$ avec p la proportion des étiquettes connues.
2. Variables aléatoires $y_j = \varepsilon_j f'_0(v_j)$ où $\varepsilon_j \sim \text{Ber}(\alpha\pi_j)$, tel que $\sum_j \mathbb{E}(\varepsilon_j) = pN'$ avec p le pourcentage des étiquettes connues.

Les proportions d'étiquettes correctement reconstruites sont mesurées pour différentes valeurs p en utilisant 500 réalisations de $\mathbf{y}$.

On récapitule les différents cas traités dans la table 4.2.

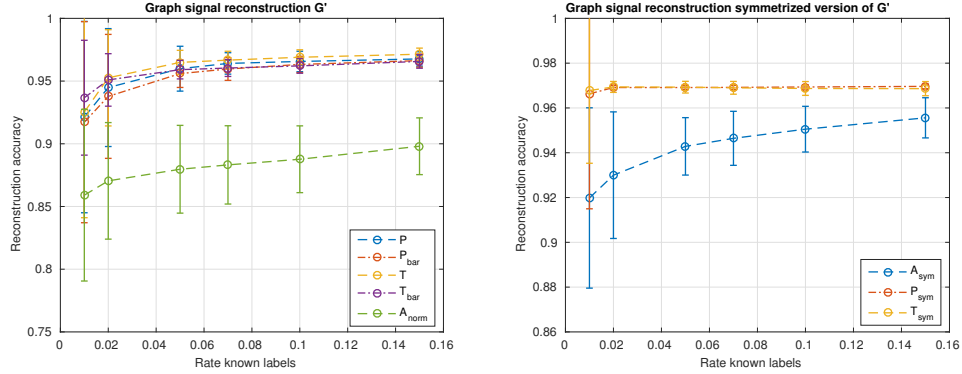


FIGURE 4.5 – (Cas 1) Gauche : Reconstruction du signal sur graphe $\mathbf{f}'_0$ sur le sous-graphe $\mathcal{G}'$. Droite : Reconstruction du signal sur graphe $\mathbf{f}'_0$ sur la version symétrique de $\mathcal{G}'$. .

Cas	Distribution $\mathbf{y}$	Opérateurs de référence
Cas 1	$y_j = \varepsilon_j f'_0(v_j), \varepsilon_j \sim \text{Ber}(p)$	$\mathbf{W}^{norm}, \mathbf{P}, \bar{\mathbf{P}}, \mathbf{T}, \bar{\mathbf{T}}, \mathbf{W}_{sym}^{norm}, \mathbf{P}_{sym}, \mathbf{T}_{sym}$
Cas 2	$y_j = \varepsilon_j f'_0(v_j), \varepsilon_j \sim \text{Ber}(\alpha\pi_j), \sum_j \mathbb{E}(\varepsilon_j) = pN'$	$\mathbf{W}^{norm}, \mathbf{P}, \bar{\mathbf{P}}, \mathbf{T}, \bar{\mathbf{T}}, \mathbf{W}_{sym}^{norm}, \mathbf{P}_{sym}, \mathbf{T}_{sym}$

TABLE 4.2 – Table des différents cas sur le graphe dirigé $\mathcal{G}'$.

Simulations numériques : Cas 1

Nous évaluons la performance de reconstruction du signal sur graphe $\mathbf{f}'_0$ à partir du signal sur graphe aléatoire $\mathbf{y} = [y_1, \dots, y'_N]$ où $y_j = \varepsilon_j f'_0(v_j)$ où $\varepsilon_j \sim \text{Ber}(p)$ avec p la proportion des étiquettes connues à la figure 4.5.

Pour toutes les proportions p d'étiquettes connues, le taux moyen d'étiquettes correctement reconstruites dans le signal sur graphe $\mathbf{f}'_0$ d'un filtre sur graphe polynomial appris de $\mathbf{P}$ ou $\bar{\mathbf{P}}$ est nettement meilleur que celui basé sur un filtre sur graphe polynomial appris de $\mathbf{A}_{norm}$ si l'on considère le sous-graphe $\mathcal{G}'$ ou sa version symétrisée. De plus, la performance de reconstruction utilisant un filtre basé sur $\mathbf{P}$ est légèrement meilleure que celle utilisant un filtre basé sur $\bar{\mathbf{P}}$.

La performance de reconstruction utilisant un filtre basé sur $\mathbf{T}$ est légèrement meilleure que celle basée sur un filtre sur graphe appris de $\mathbf{P}$ ou $\bar{\mathbf{P}}$ pour des proportions p d'étiquettes connues élevées. Pour de petites proportions p d'étiquettes connues, le taux moyen d'étiquettes correctement reconstruites à partir d'un filtre sur graphe polynomial appris de $\bar{\mathbf{T}}$ est légèrement meilleur

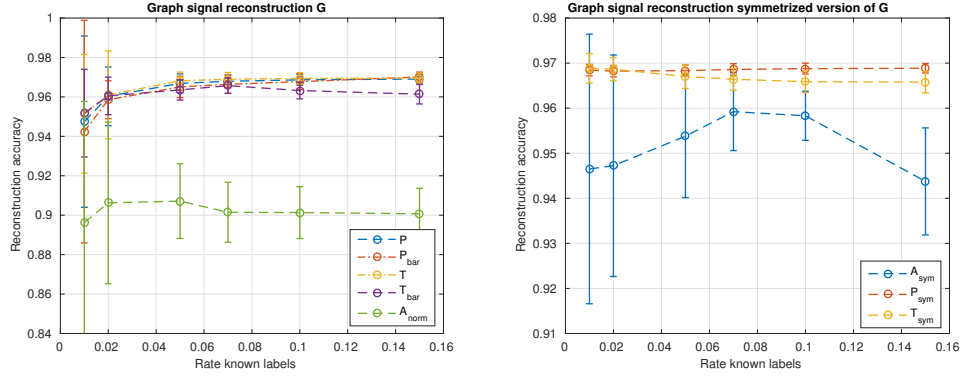


FIGURE 4.6 – (Cas 2) Gauche : Reconstruction du signal sur graphe $\mathbf{f}'_0$ sur le sous-graphe $\mathcal{G}'$. Droite : Reconstruction du signal graphe $\mathbf{f}'_0$ sur la version symétrique de $\mathcal{G}'$.

que ceux basés sur un filtre sur graphe appris de $\mathbf{P}$ ou $\bar{\mathbf{P}}$.

Nous remarquons également que la performance de reconstruction utilisant un filtre basé sur $\mathbf{T}_{sym}$ est identique à celle utilisant un filtre basé sur $\mathbf{P}_{sym}$ pour toutes les proportions p des étiquettes connues.

Simulations numériques : Cas 2

La performance de reconstruction du signal sur graphe $\mathbf{f}'_0$ à partir du signal sur graphe aléatoire $\mathbf{y} = [y_1, \dots, y_{N'}]$ où $y_j = \varepsilon_j f'_0(v_j)$ où $\varepsilon_j \sim \text{Ber}(\alpha\pi_j)$, tel que $\sum_j \mathbb{E}(\varepsilon_j) = pN'$ est évaluée à la figure 4.6. Nous observons une performance significativement meilleure en utilisant un filtre basé sur $\mathbf{P}$ ou $\bar{\mathbf{P}}$ qu'en utilisant $\mathbf{A}_{norm}$ dans le cas où nous considérons soit le sous-graphe $\mathcal{G}'$ ou sa version symétrique et une performance légèrement meilleure avec $\mathbf{P}$ plutôt que $\bar{\mathbf{P}}$. Pour toutes les proportions p d'étiquettes connues, le taux moyen d'étiquettes correctement reconstruites est identique en utilisant soit un filtre basé sur $\mathbf{T}$ ou $\bar{\mathbf{T}}$ sont meilleurs que l'utilisation d'un filtre basé sur $\mathbf{P}$ ou $\bar{\mathbf{P}}$. Nous remarquons que la performance de reconstruction basée sur $\mathbf{A}_{norm}$ a une grande variabilité et diminue même pour p plus grand.

Cas 1 vs. Cas 2

La performance de reconstruction utilisant un filtre basé sur $\mathbf{P}$ ou $\bar{\mathbf{P}}$ est légèrement meilleure lorsque les échantillons du signal $\mathbf{y}$ sont sélectionnés selon une distribution proportionnelle à la distribution fixe π (cas 2) que lorsque la distribution est uniforme (cas 1). C'est aussi le cas en utilisant un filtre basé sur $\mathbf{T}$ ou $\bar{\mathbf{T}}$. Ceci est attendu car les sommets avec des valeurs π_k plus grandes correspondent à des blogs mieux connectés, qui sont susceptibles d'avoir une influence sur un plus grand nombre d'autres blogs. Bien

que la performance de reconstruction utilisant un filtre basé sur $\mathbf{A}_{norm}$ soit meilleure dans le cas 2 que dans le cas 1, la grande variabilité et la faible performance à p grand dans le cas 2 suggèrent que ces filtres fournissent des modèles du signal généralement plus mauvais. La meilleure performance de reconstruction utilisant un filtre basé sur $\mathbf{T}$ ou $\bar{\mathbf{T}}$ par rapport à l'utilisation d'un filtre basé sur $\mathbf{P}$ ou $\bar{\mathbf{P}}$ suggère que le signal sur graphe semble plus approprié pour vivre dans $\ell^2(\mathcal{V})$ que $\ell^2(\mathcal{V}, \pi)$.

6.2.3.2 Modélisation du signal sur $\mathcal{G}'$ avec $\bar{\mathbf{P}}_\alpha \in \bar{\mathcal{P}}$ Ici nous considérons le même problème que dans la section précédente mais comparons maintenant les performances en utilisant les opérateurs de référence $\bar{\mathbf{P}}_\alpha \in \bar{\mathcal{P}}$ et leurs équivalents $\bar{\mathbf{T}}_\alpha = \mathbf{\Pi}^{1/2} \bar{\mathbf{P}}_\alpha \mathbf{\Pi}^{-1/2} \in \ell^2(\mathcal{V})$. En d'autres termes, nous comparons la performance si le signal sur graphe $\mathbf{f}'_0$ appartient à $\ell^2(\mathcal{V}, \pi)$ ou $\ell^2(\mathcal{V})$. Les variables aléatoires $y_j, j = 1, \dots, N'$ sont réparties dans l'un des deux cas suivants :

1. Variables aléatoires $y_j = \varepsilon_j f'_0(v_j)$ où $\varepsilon_j \sim \text{Ber}(p)$ avec p la proportion des étiquettes connues.
2. Variables aléatoires $y_j = \varepsilon_j f'_0(v_j)$ où $\varepsilon_j \sim \text{Ber}(\alpha \pi_j)$, tel que $\sum_j \mathbb{E}(\varepsilon_j) = pN'$ avec p le pourcentage des étiquettes connues.

On récapitule les différents cas traités dans la table 4.3.

Cas	Distribution $\mathbf{y}$	Opérateurs de référence
Cas 1.1	$y_j = \varepsilon_j f'_0(v_j), \varepsilon_j \sim \text{Ber}(p)$	$\bar{\mathbf{P}}_\alpha \in \bar{\mathcal{P}}$
Cas 1.2	$y_j = \varepsilon_j f'_0(v_j), \varepsilon_j \sim \text{Ber}(p)$	$\bar{\mathbf{T}}_\alpha, \forall \alpha \in [0, 1]$
Cas 2.1	$y_j = \varepsilon_j f'_0(v_j), \varepsilon_j \sim \text{Ber}(\alpha \pi_j), \sum_j \mathbb{E}(\varepsilon_j) = pN'$	$\bar{\mathbf{P}}_\alpha \in \bar{\mathcal{P}}$
Cas 2.2	$y_j = \varepsilon_j f'_0(v_j), \varepsilon_j \sim \text{Ber}(\alpha \pi_j), \sum_j \mathbb{E}(\varepsilon_j) = pN'$	$\bar{\mathbf{T}}_\alpha, \forall \alpha \in [0, 1]$

TABLE 4.3 – Table des différents cas traités.

Simulations numériques : Cas 1.1

Nous évaluons la performance de reconstruction de $\mathbf{f}'_0$ en fonction de α dans la figure 4.7. Pour de petites proportions p d'étiquettes connues, on remarque que la performance de reconstruction utilisant un filtre basé sur $\bar{\mathbf{P}}, \alpha \simeq 0.2$ est la meilleure alors que pour des proportions supérieures p , la performance de reconstruction utilisant un filtre basé sur $\bar{\mathbf{P}}, \alpha \simeq 0.3$ est la

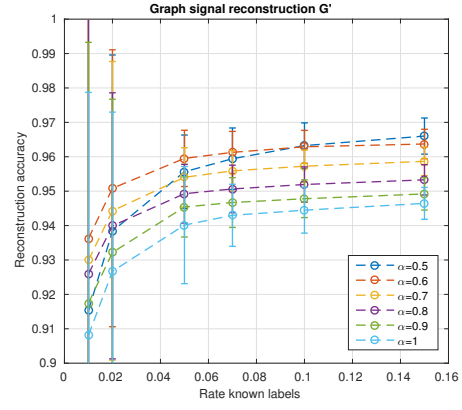
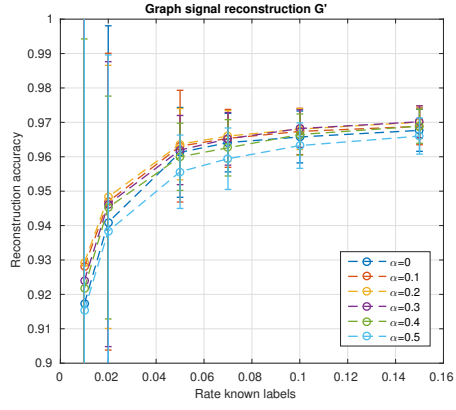


FIGURE 4.7 – (Cas 1.1) Reconstruction du signal sur graphe $\mathcal{f}'_0$ sur le sous-graphe $\mathcal{G}'$.

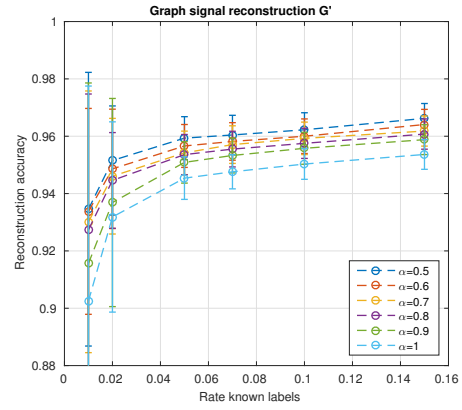
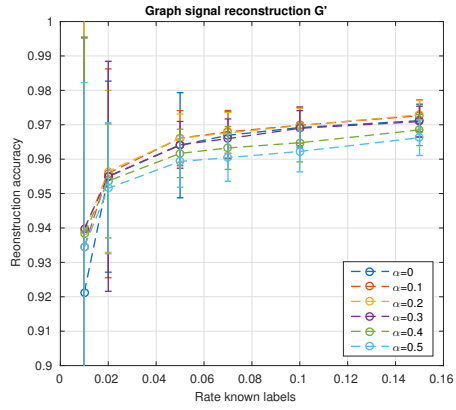


FIGURE 4.8 – (Cas 1.2) Reconstruction du signal sur graphe $\mathcal{f}'_0$ sur le sous-graphe $\mathcal{G}'$.

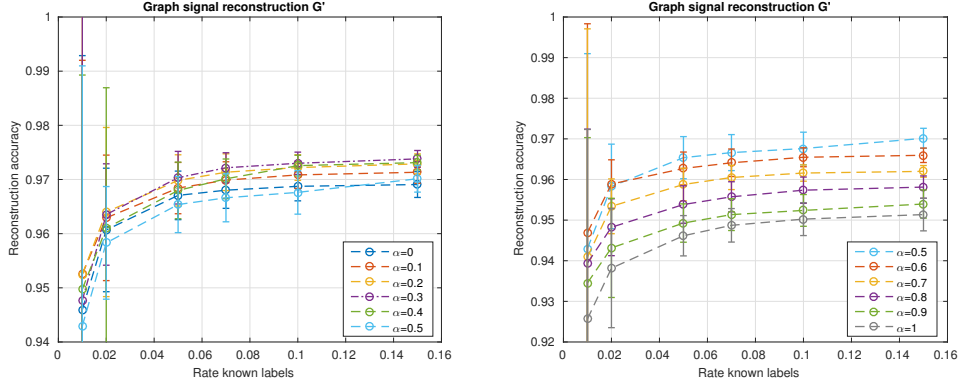


FIGURE 4.9 – (Cas 2.1) Reconstruction du signal sur graphe $\mathbf{f}'_0$ sur le sous-graphe $\mathcal{G}'$.

meilleure. Nous remarquons une moins bonne performance globale lorsque α passe de $\alpha = 0.5$ à $\alpha = 1$.

Simulations numériques : Cas 1.2

Nous évaluons la performance de reconstruction de $\mathbf{f}'_0$ en fonction de α dans la figure 4.8. Les performances de reconstruction utilisant un filtre basé sur $\bar{\mathbf{T}}_\alpha, \alpha \simeq 0.2$ sont les meilleures sauf pour le taux $p = 0.01$. Nous remarquons une moins bonne performance globale lorsque α passe de $\alpha = 0.5$ à $\alpha = 1$.

Cas 1.1 vs Cas 1.2

Pour toutes les proportions p d'étiquettes connues, les performances de reconstruction sont légèrement meilleures en utilisant des filtres basés sur $\bar{\mathbf{T}}_\alpha$ qu'en utilisant des filtres basés sur $\bar{\mathbf{P}}_\alpha$. De plus, l'étude des deux cas précédents suggère que, pour ce signal, les filtres sur graphe basés sur $\bar{\mathbf{P}}_\alpha$ donnent de meilleurs résultats lorsque la combinaison convexe de $\mathbf{P}$ et $\mathbf{P}^*$ implique davantage la marche aléatoire avant, c.-à-d. $\alpha \in (0, 0.5)$, que celle renversée dans le temps, c.-à-d. $\alpha \in (0.5, 1)$.

Simulations numériques : Cas 2.1

Nous évaluons la performance de reconstruction de $\mathbf{f}'_0$ en fonction de α en figure 4.9. Nous remarquons que les performances de reconstruction utilisant un filtre basé sur $\bar{\mathbf{P}}_\alpha, \alpha \simeq 0.3$ sont les meilleures sauf pour les taux $p = \{0.01, 0.02\}$. Nous remarquons une moins bonne performance globale lorsque α passe de $\alpha = 0.5$ à $\alpha = 1$. Ceci suggère que, pour ce signal, les filtres sur graphe basés sur $\bar{\mathbf{P}}_\alpha$ fonctionnent mieux lorsque la combinaison convexe

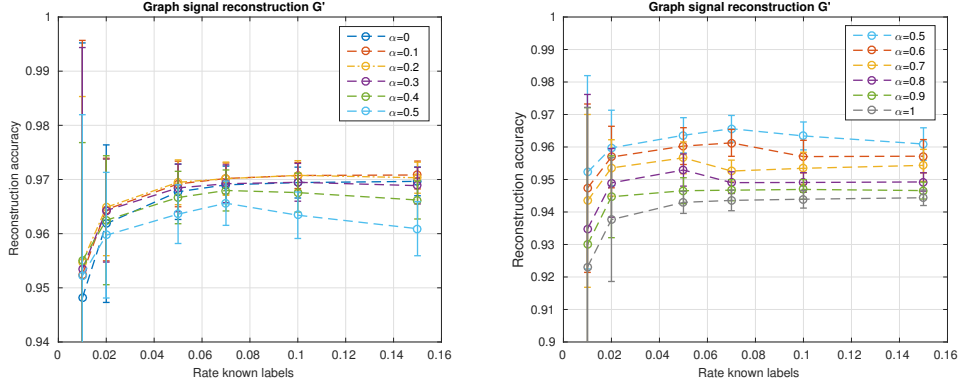


FIGURE 4.10 – (Cas 2.2) Reconstruction du signal sur graphe $\mathbf{f}'_0$ sur le sous-graphe $\mathcal{G}'$.

de $\mathbf{P}$ et $\mathbf{P}$ implique plus la marche aléatoire avant, c'est-à-dire $\alpha \in (0, 0.5)$, que celle renversée dans le temps, c'est-à-dire $\alpha \in (0.5, 1)$.

Simulations numériques : Cas 2.2

Nous évaluons la performance de reconstruction de $\mathbf{f}_0 \in \ell^2(\mathcal{V})$ en fonction de α dans la figure 4.10. Pour toutes les proportions p d'étiquettes connues, les performances de reconstruction utilisant un filtre basé sur $\bar{\mathbf{P}}, \alpha \simeq 0.3$ sont les meilleures. Comme dans les cas précédents, nous constatons une performance globale plus faible lorsque α passe de $\alpha = 0.5$ à $\alpha = 1$.

6.2.4 Cas 2.1 vs. Cas 2.2

les performances de reconstruction sont légèrement meilleures en utilisant des filtres basés sur $\bar{\mathbf{P}}_\alpha$ qu'en utilisant des filtres basés sur $\bar{\mathbf{T}}_\alpha$ lorsque les échantillons dans $\mathbf{y}$ sont sélectionnés parmi une distribution proportionnelle à la distribution fixe π que lorsque la distribution est uniforme. Inversement, les performances de reconstruction sont légèrement meilleures en utilisant des filtres basés sur $\bar{\mathbf{T}}_\alpha$ qu'en utilisant des filtres basés sur $\bar{\mathbf{P}}_\alpha$ lorsque la distribution en $\mathbf{y}$ est uniforme. Par conséquent, la distribution des échantillons a une influence sur la performance de reconstruction par l'utilisation de filtres appris basés soit sur $\bar{\mathbf{P}}_\alpha$ ou $\bar{\mathbf{T}}_\alpha$.

6.2.4.1 Modélisation du signal sur le graphe complet $\mathcal{G}$. Nous considérons maintenant le même problème sur l'ensemble du graphe $\mathcal{G}$, et nous devons traiter sa propriété de ne pas être fortement connexe, comme l'exige notre cadre.

Cas	Distribution $\mathbf{y}$	Opérateurs de référence
Cas 1	$y_j = \varepsilon_j f'_0(v_j), \varepsilon_j \sim \text{Ber}(p)$	$\mathbf{A}^{norm}, \mathbf{P}_\epsilon, \mathbf{P}_G, \bar{\mathbf{P}}_\epsilon, \bar{\mathbf{P}}_G, \mathbf{T}_\epsilon, \mathbf{T}_G, \bar{\mathbf{T}}_\epsilon, \bar{\mathbf{T}}_G$

TABLE 4.4 – Table du cas traité.

En effet, certains sommets de $\mathcal{G}$ ont un degré sortant égal à zéro et nous ne pouvons pas créer de matrice de transition directement à partir de la matrice d'adjacence. Pour surmonter ce problème, nous proposons deux approches qui rendent le graphe fortement connexe et la marche aléatoire associée ergodique :

1. Perturbation de rang un : à partir de la matrice d'adjacence originelle $\mathbf{W}$, on construit une nouvelle matrice d'adjacence $\mathbf{W}_\epsilon$ comme suit :

$$\mathbf{W}_\epsilon = \mathbf{W} + \epsilon \mathbf{J},$$

où $\mathbf{J} = \mathbf{1}\mathbf{1}^\top/N$ est une matrice de rang un et ϵ est petit. La perturbation faible de la matrice adjacence $\mathbf{W}$ par la matrice $\mathbf{J}$ assure que la marche aléatoire sur $\mathcal{G}$ est ergodique avec la mesure stationnaire π_ϵ et sa matrice de transition associée $\mathbf{P}_\epsilon$ bien définie. Pour nos expériences, nous choisissons $\epsilon = 10^{-4}$.

2. Construction de la matrice de Google de $\mathcal{G}$ [60]. Cela se fait en deux étapes. Tout d'abord, nous construisons une matrice d'adjacence $\tilde{\mathbf{W}}$ à partir de $\mathbf{W}$ en ajoutant un pour tous les sommets pendants, c'est-à-dire les sommets sans arêtes. De $\tilde{\mathbf{W}}$, nous pouvons construire la matrice de transition $\mathbf{S}$. Deuxièmement, nous définissons la matrice Google $\mathbf{P}_G$ comme suit :

$$\mathbf{P}_G = (1 - \gamma)\mathbf{S} + \gamma\mathbf{J}.$$

avec $\gamma = 0.85$ [60] et distribution stationnaire π_G .

Les variables aléatoires $y_j, j = 1 \dots, N$ sont distribuées comme $y_j = \varepsilon_j f'_0(v_j), \varepsilon_j \sim \text{Ber}(p)$. Pour chaque proportion d'étiquettes connues p , les bons taux de reconstruction sont obtenus en calculant une moyenne de 500 signaux sur graphe de réalisation du signal $\mathbf{y}$. Nous considérons l'ensemble des opérateurs de référence $\{\mathbf{A}^{norm}, \mathbf{P}_\epsilon, \mathbf{P}_G, \bar{\mathbf{P}}_\epsilon, \bar{\mathbf{P}}_G\}$ et l'ensemble $\{\mathbf{T}_\epsilon, \mathbf{T}_G, \bar{\mathbf{T}}_\epsilon, \bar{\mathbf{T}}_G\}$.

On récapitule les différents cas traités dans la table 4.4.

Simulations numériques

La performance de reconstruction est montrée dans la figure 4.11. Pour toutes les valeurs p , la performance de reconstruction utilisant des filtres

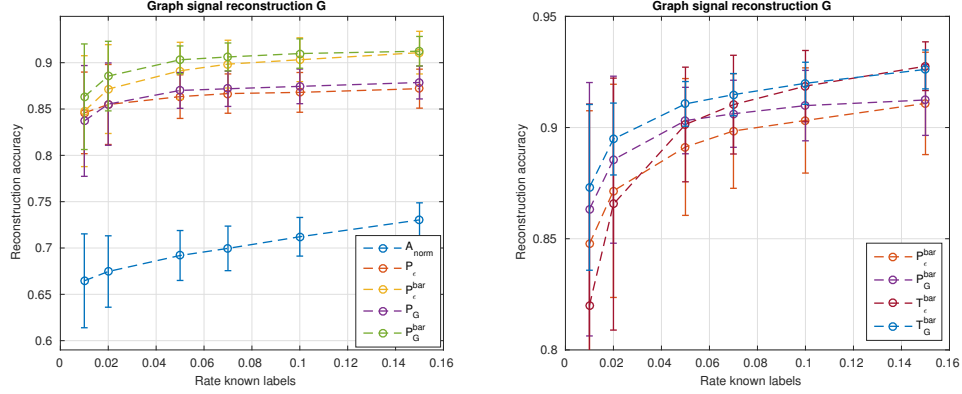


FIGURE 4.11 – Reconstruction du signal sur graphe $\mathbf{f}_0$ sur $\mathcal{G}$.

basés sur $\mathbf{A}_{norm}$ est significativement moins bonne qu'avec tous les autres opérateurs de référence. Par rapport aux autres opérateurs de référence, $\mathbf{A}_{norm}$ est le seul qui peut être utilisé sans modifier le graphe pour le rendre fortement connecté. Pourtant, ses performances sont toujours pauvres.

Parmi les autres opérateurs de référence, on note une performance nettement meilleure généralement en utilisant les réversibilisations $\bar{\mathbf{P}}_\epsilon$ et $\bar{\mathbf{P}}_G$ par rapport aux marches aléatoires irréversibles $\mathbf{P}_\epsilon$ et $\mathbf{P}_G$. De plus, la meilleure performance de reconstruction est obtenue en utilisant un filtre basé sur $\bar{\mathbf{T}}_G$ et la performance de reconstruction en utilisant un filtre basé sur $\bar{\mathbf{T}}_\epsilon$ est identique à $\bar{\mathbf{T}}_G$ pour un taux supérieur de labels connus. Ceci diffère des résultats sur $\mathcal{G}'$ où les deux auraient des performances similaires.

Enfin, la performance de la reconstruction ne dépend pas autant de l'approche que nous utilisons pour rendre le graphe fortement connexe. Il semble toutefois que l'approche de Google surpasse légèrement l'approche de la perturbation de rang un, tant pour les marches aléatoires non réversibles que pour leurs réversibilisations.

Modélisation de signal sur $\mathcal{G}$ avec $\bar{\mathbf{P}}_\alpha \in \bar{\mathcal{P}}$

Ici nous considérons le même problème que dans la section précédente mais comparons maintenant la performance en utilisant les opérateurs de référence $\bar{\mathbf{P}}_\alpha \in \bar{\mathcal{P}}$ et leurs équivalents respectifs $\bar{\mathbf{T}}_\alpha = \mathbf{\Pi}^{1/2} \bar{\mathbf{P}}_\alpha \mathbf{\Pi}^{-1/2} \in \ell^2(\mathcal{V})$. Les variables aléatoires $y_j, j = 1 \dots, N$ sont distribuées comme $y_j = \varepsilon_j f_0(v_j), \varepsilon_j \sim \text{Ber}(p)$.

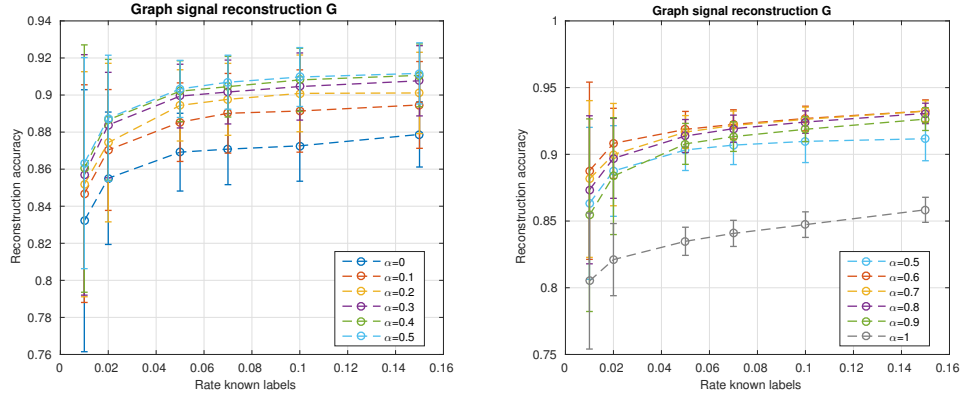


FIGURE 4.12 – (Cas 1) Reconstruction du signal sur graphe f_0 sur $\mathcal{G}$.

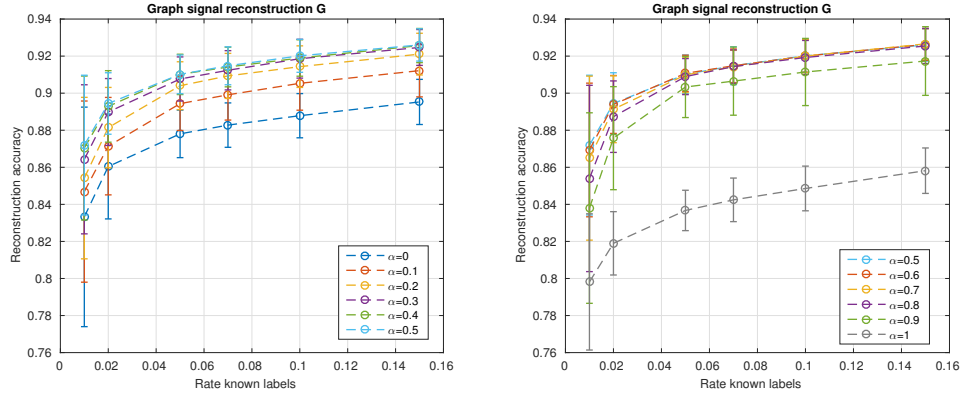


FIGURE 4.13 – (Cas 2) Reconstruction du signal sur graphe f_0 sur $\mathcal{G}$.

Cas	Distribution y	Opérateurs de référence
Cas 1	$y_j = \varepsilon_j f_0(v_j), \varepsilon_j \sim \text{Ber}(p)$	$\bar{\mathbf{P}}_\alpha \in \bar{\mathcal{P}}$
Cas 2	$y_j = \varepsilon_j f_0(v_j), \varepsilon_j \sim \text{Ber}(p)$	$\bar{\mathbf{T}}_\alpha, \forall \alpha \in [0, 1]$

TABLE 4.5 – Table des différents cas traités.

Simulations numériques : Cas 1

Nous évaluons la performance de reconstruction de $\mathbf{f}_0$ en fonction de α dans la figure 4.12. Pour toutes les proportions p d'étiquettes connues, les performances de reconstruction utilisant un filtre basé sur $\bar{\mathbf{P}}_\alpha$, $\alpha \simeq 0.6$ sont les meilleures. Nous remarquons aussi que la performance de reconstruction utilisant un filtre avec $\alpha \in \{0.7, 0.8\}$ est meilleure que $\alpha \simeq 0.5$. Ceci suggère que, pour ce signal, les filtres sur graphe basés sur $\bar{\mathbf{P}}_\alpha$ fonctionnent mieux lorsque la combinaison convexe de $\mathbf{P}$ et $\bar{\mathbf{P}}$ implique plus la marche aléatoire renversée dans le temps, c'est-à-dire $\alpha \in (0.6, 0.8)$, que celle avant, soit $\alpha \in (0, 10.5)$.

Simulations numériques : Cas 2

Nous évaluons la performance de reconstruction de $\mathbf{f}_0$ en fonction de α dans la figure 4.13. Pour toutes les proportions p d'étiquettes connues, les performances de reconstruction utilisant un filtre basé sur $\bar{\mathbf{T}}_\alpha$, $\alpha \simeq 0.5$ sont les meilleures. Nous remarquons une moins bonne performance globale lorsque α passe de $\alpha = 0.5$ à $\alpha = 1$.

Cas 1 vs Cas 2

Pour toutes les proportions p d'étiquettes connues, les performances de reconstruction utilisant des filtres basés sur $\bar{\mathbf{T}}_\alpha$ sont meilleures que celles utilisant des filtres basés sur $\bar{\mathbf{P}}_\alpha$.

7 Analyses multi-résolution sur graphes dirigés

Dans les sections précédentes, nous avons proposé une base de Fourier sur des graphes dirigés comme une base de vecteurs propres de la matrice de transition d'une marche aléatoire ergodique sur le graphe et nous avons établi une notion de fréquence en étudiant la variation des vecteurs propres de la marche aléatoire. Nous sommes ainsi capables de construire des analyses multi-échelles pour fonctions sur graphes dirigés.

Dans un premier temps, nous proposons des trames d'ondelettes construites à partir de bancs de filtres sur graphe d'analyse et de synthèse. Cette construction multi-échelle est étroitement liée à la construction des trames polynomiales de diffusion [31] et des ondelettes spectrales sur graphes [32]. Dans un second temps, nous proposons une construction d'ondelettes échantillonnées de manière critique sur graphes dirigés qui généralise le cadre des ondelettes de diffusion [30]. En utilisant la même remarque faite en début de section 6, nos constructions en ondelettes sur graphes se font pour des fonctions dans l'espace $\ell^2(\mathcal{V})$.

7.1 Transformée en ondelettes redondantes sur graphes dirigés

Dans cette section, nous proposons une transformée en ondelettes redondantes sur graphes dirigés. Nous suivons la construction des ondelettes spectrales sur graphes non dirigés [32] présentée en section 8.2 du chapitre 2 et des trames de diffusion polynomiales [31]. La nouveauté est la construction de filtres à réponse fréquentielle comme combinaison linéaire de projecteurs spectraux de la marche aléatoire.

Nous introduisons les éléments nécessaires à la construction des ondelettes redondantes sur graphes dirigés.

7.1.1 Cadre théorique

Soit $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ un graphe dirigé fortement connexe de cardinalité $|\mathcal{V}| = N$. Sur $\mathcal{G}$, on définit une marche aléatoire $\mathcal{X}$ supposée ergodique de distribution stationnaire π et représentée par sa matrice de transition $\mathbf{P} \in \mathbb{R}^{N \times N}$. Nous introduisons $\mathbf{T} = \mathbf{\Pi}^{1/2} \mathbf{P} \mathbf{\Pi}^{-1/2} \in \ell^2(\mathcal{V})$, une matrice similaire à $\mathbf{P}$. Nous supposons $\mathbf{T}$ diagonalisable. Nous définissons un opérateur passe-bas de la manière suivante.

Définition 4.7. *Un opérateur passe-bas dilaté de t , $\mathbf{H}_t^{\mathcal{G}} \in \mathbb{R}^{N \times N}$ sur un graphe dirigé fortement connexe $\mathcal{G}$ est défini par :*

$$\mathbf{H}_t^{\mathcal{G}} = \sum_{\ell=1}^p h(t\omega_{\ell}) \mathbf{S}_{\omega_{\ell}}, \quad t \in \mathbb{N}.$$

avec $h : [0, \infty) \rightarrow \mathbb{R}$, une fonction avec une réponse en fréquence passe-bas, ω_{ℓ} la fréquence sur graphe et $\mathbf{S}_{\omega_{\ell}}$ le projecteur associé à la fréquence ω_{ℓ} (4.27).

La fonction d'échelle dilatée de t et translatée de y est définie par :

$$\mathbf{h}_{t,y}^{\mathcal{G}} = \mathbf{H}_t^{\mathcal{G}} \delta_y,$$

où δ_y est la fonction delta Kronecker au sommet $y \in \mathcal{V}$.

Nous définissons un opérateur passe-bande de la manière suivante.

Définition 4.8. *Un opérateur passe-bande dilaté de t , $\mathbf{G}_t^{\mathcal{G}} \in \mathbb{R}^{N \times N}$ sur un graphe dirigé fortement connexe $\mathcal{G}$ est défini de la manière suivante :*

$$\mathbf{G}_t^{\mathcal{G}} = \sum_{\ell=1}^p g(t\omega_{\ell}) \mathbf{S}_{\omega_{\ell}}, \quad t \in \mathbb{N}.$$

avec $g : [0, \infty) \rightarrow \mathbb{R}$, une fonction avec une réponse en fréquence passe-bande, ω_{ℓ} la fréquence sur graphe et $\mathbf{S}_{\omega_{\ell}}$ le projecteur associé à la fréquence ω_{ℓ} (4.27).

La fonction d'ondelette dilatée de t et translatée de y est défini par :

$$\mathbf{g}_{t,y}^{\mathcal{G}} = \mathbf{G}_t^{\mathcal{G}} \delta_y,$$

où δ_y est la fonction delta Kronecker au sommet $y \in \mathcal{V}$.

Ayant défini les opérateurs passe-bas et passe-bande sur graphes dirigés, nous sommes en mesure de construire des bancs de filtres d'analyse et de synthèse sur graphes dirigés.

Définition 4.9. Nous définissons un banc de filtres de synthèse $\mathcal{K}$ comme l'ensemble d'un filtre passe-bas dilaté de t_J , $\mathbf{H}_{t_J}^{\mathcal{G}}$ et d'une série de filtres passe bande dilatés de manière croissante $\{t_j\}_{j=1}^J$, $\{\mathbf{G}_{t_j}\}_{j=1}^J$:

$$\mathcal{K} = \{\mathbf{H}_{t_J}^{\mathcal{G}}, \mathbf{G}_{t_1}^{\mathcal{G}}, \dots, \mathbf{G}_{t_J}^{\mathcal{G}}\}.$$

Définition 4.10. Nous définissons un banc de filtres d'analyse $\tilde{\mathcal{K}}$ comme l'ensemble d'un filtre dilaté de t_J , $\tilde{\mathbf{H}}_{t_J}^{\mathcal{G}}$ et d'une série de filtres dilatés de manière croissante $\{t_j\}_{j=1}^J$, $\{\tilde{\mathbf{G}}_{t_j}\}_{j=1}^J$:

$$\tilde{\mathcal{K}} = \{\tilde{\mathbf{H}}_{t_J}^{\mathcal{G}}, \tilde{\mathbf{G}}_{t_1}^{\mathcal{G}}, \dots, \tilde{\mathbf{G}}_{t_J}^{\mathcal{G}}\},$$

où

$$\tilde{\mathbf{H}}_{t_J}^{\mathcal{G}} = \sum_{\ell=1}^p \tilde{h}(t\omega_{\ell}) \mathbf{S}_{\omega_{\ell}} \in \mathbb{R}^{N \times N},$$

et

$$\tilde{\mathbf{G}}_{t_j}^{\mathcal{G}} = \sum_{\ell=1}^p \tilde{g}(t\omega_{\ell}) \mathbf{S}_{\omega_{\ell}} \in \mathbb{R}^{N \times N}, \quad \forall j = 1, \dots, J$$

Nous définissons également $\tilde{\mathbf{h}}_{t,k}$ et $\tilde{\mathbf{g}}_{t,k}$ comme des vecteurs ligne de la manière suivante :

$$\tilde{\mathbf{h}}_{t,k} = \delta_k^{\top} \tilde{\mathbf{H}}_t^{\mathcal{G}}, \quad \tilde{\mathbf{g}}_{t,k} = \delta_k^{\top} \tilde{\mathbf{G}}_t^{\mathcal{G}}.$$

Proposition 4.9. Sachant les bancs de filtres de synthèse et d'analyse, respectivement $\mathcal{K}$ et $\tilde{\mathcal{K}}$, la condition de reconstruction parfaite :

$$\mathbf{H}_{t_J}^{\mathcal{G}} \tilde{\mathbf{H}}_{t_J}^{\mathcal{G}} + \sum_{j=1}^J \mathbf{G}_{t_j}^{\mathcal{G}} \tilde{\mathbf{G}}_{t_j}^{\mathcal{G}} = \mathbf{I} \quad (4.45)$$

est garantie si et seulement si :

$$h(t_J\omega_{\ell})\tilde{h}(t_J\omega_{\ell}) + \sum_{j=1}^J g(t_j\omega_{\ell})\tilde{g}(t_j\omega_{\ell}) = 1, \quad \forall \ell \in \{1, \dots, p\}.$$

Démonstration.

$$\begin{aligned} \mathbf{H}_{t_J}^{\mathcal{G}} \tilde{\mathbf{H}}_{t_J}^{\mathcal{G}} + \sum_{j=1}^J \mathbf{G}_{t_j}^{\mathcal{G}} \tilde{\mathbf{G}}_{t_j}^{\mathcal{G}} &= \sum_{\ell, \ell'=1}^p h(t_J\omega_{\ell})\tilde{h}(t_J\omega_{\ell'}) \mathbf{S}_{\omega_{\ell}} \mathbf{S}_{\omega_{\ell'}} + \sum_{j=1}^J \sum_{\ell, \ell'=1}^p g(t_j\omega_{\ell})\tilde{g}(t_j\omega_{\ell'}) \mathbf{S}_{\omega_{\ell}} \mathbf{S}_{\omega_{\ell'}} \\ &= \sum_{\ell=1}^p h(t_J\omega_{\ell})\tilde{h}(t_J\omega_{\ell}) \mathbf{S}_{\omega_{\ell}} + \sum_{j=1}^J \sum_{\ell=1}^p g(t_j\omega_{\ell})\tilde{g}(t_j\omega_{\ell}) \mathbf{S}_{\omega_{\ell}} \\ &= \sum_{\ell=1}^p \left[h(t_J\omega_{\ell})\tilde{h}(t_J\omega_{\ell}) + \sum_{j=1}^J g(t_j\omega_{\ell})\tilde{g}(t_j\omega_{\ell}) \right] \mathbf{S}_{\omega_{\ell}} \end{aligned}$$

Par conséquent, la condition (4.45) est garantie si et seulement si

$$h(t_J \omega_\ell) \tilde{h}(t_J \omega_\ell) + \sum_{j=1}^J g(t_j \omega_\ell) \tilde{g}(t_j \omega_\ell) = 1 \quad \forall \ell = 1 \dots, p. \quad \square$$

Nous définissons une trame de la manière suivante.

Définition 4.11. Soit $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ un graphe dirigé. Une famille dénombrable d'éléments $\{\mathbf{f}_k\}_{k=1}^n \in \ell^2(\mathcal{V})$ est dite être une trame si pour tout signal sur graphe $\mathbf{f} \in \ell^2(\mathcal{V})$ nous avons :

$$A \|\mathbf{f}\|^2 \leq \sum_{k=1}^n |\langle \mathbf{f}, \mathbf{f}_k \rangle|^2 \leq B \|\mathbf{f}\|^2,$$

pour des constantes $0 < A \leq B < \infty$ qui sont appelées les bornes de la trame.

Nous introduisons la matrice rectangulaire $\mathbf{K} = (\mathbf{H}_{t_J}^{\mathcal{G}}, \mathbf{G}_{t_1}^{\mathcal{G}}, \dots, \mathbf{G}_{t_J}^{\mathcal{G}}) \in \mathbb{R}^{N \times N(J+1)}$ où :

$$\mathbf{K} = \left(\mathbf{h}_{t_J,1}^{\mathcal{G}}, \dots, \mathbf{h}_{t_J,N}^{\mathcal{G}}, \mathbf{g}_{t_1,1}^{\mathcal{G}}, \dots, \mathbf{g}_{t_J,1}^{\mathcal{G}}, \dots, \mathbf{g}_{t_J,N}^{\mathcal{G}} \right) \in \mathbb{R}^{N \times N(J+1)}.$$

Nous introduisons la matrice rectangulaire $\tilde{\mathbf{K}} = (\tilde{\mathbf{H}}_{t_J}^{\mathcal{G}}, \tilde{\mathbf{G}}_{t_1}^{\mathcal{G}}, \dots, \tilde{\mathbf{G}}_{t_J}^{\mathcal{G}})^{\top} \in \mathbb{R}^{N(J+1) \times N}$ où :

$$\tilde{\mathbf{K}} = \begin{pmatrix} \tilde{\mathbf{h}}_{t_J,1}^{\mathcal{G}} \\ \vdots \\ \tilde{\mathbf{h}}_{t_J,N}^{\mathcal{G}} \\ \tilde{\mathbf{g}}_{t_1,1}^{\mathcal{G}} \\ \vdots \\ \tilde{\mathbf{g}}_{t_J,1}^{\mathcal{G}} \\ \vdots \\ \tilde{\mathbf{g}}_{t_J,N}^{\mathcal{G}} \end{pmatrix} \in \mathbb{R}^{N(J+1) \times N}. \quad (4.46)$$

Proposition 4.10. Soient $\mathbf{K}$ et $\tilde{\mathbf{K}}$ respectivement les bancs de filtres de synthèse et d'analyse. Nous assumons que les bancs de filtres $\mathbf{K}$ et $\tilde{\mathbf{K}}$ vérifient la condition de reconstruction parfaite. Par conséquent, $\tilde{\mathbf{K}}$ est une trame pour $\ell^2(\mathcal{V})$ avec borne inférieure de la trame $1/\|\mathbf{K}\|^2$ et borne supérieure de la trame $\|\tilde{\mathbf{K}}\|^2$.

Démonstration. Nous avons :

$$\|\tilde{\mathbf{K}} \mathbf{f}\|^2 = \sum_{k=1}^N |\langle \tilde{\mathbf{h}}_{t_J,k}^{\mathcal{G}}, \mathbf{f} \rangle|^2 + \sum_{j=1}^J \sum_{k=1}^N |\langle \tilde{\mathbf{g}}_{t_j,k}^{\mathcal{G}}, \mathbf{f} \rangle|^2.$$

Par conséquent, nous pouvons écrire :

$$\begin{aligned} A\|\mathbf{f}\|^2 &\leq \|\tilde{\mathbf{K}}\mathbf{f}\|^2 \leq B\|\mathbf{f}\|^2, \quad \forall \mathbf{f} \\ \Leftrightarrow A\|\mathbf{f}\|^2 &\leq \sum_{k=1}^N |\langle \tilde{\mathbf{h}}_{t_J,k}^{\mathcal{G}}, \mathbf{f} \rangle|^2 + \sum_{j=1}^J \sum_{k=1}^N |\langle \tilde{\mathbf{g}}_{t_j,k}^{\mathcal{G}}, \mathbf{f} \rangle|^2 \leq B\|\mathbf{f}\|^2, \quad \forall \mathbf{f} \end{aligned}$$

où $A = 1/\|\mathbf{K}\|^2$ et $B = \|\tilde{\mathbf{K}}\|^2$.

La partie droite est obtenue par l'inégalité de Cauchy-Schwarz appliquée sur $\|\tilde{\mathbf{K}}\mathbf{f}\|^2$. La partie gauche est obtenue par le fait que $\|\mathbf{f}\|^2 = \|\mathbf{K}\tilde{\mathbf{K}}\mathbf{f}\|^2 \leq \|\mathbf{K}\|^2 \|\tilde{\mathbf{K}}\mathbf{f}\|^2$ et application de l'inégalité de Cauchy-Schwarz. Par conséquent, $\tilde{\mathbf{K}}$ est une trame. $\square$

Définition 4.12 (Décomposition en ondelettes d'un signal sur graphe dirigé). *Un signal sur graphe dirigé donné $\mathbf{f} \in \ell^2(\mathcal{V})$ peut s'exprimer de la manière suivante :*

$$\mathbf{f} = \sum_{k=1}^N \langle \mathbf{f}, \tilde{\mathbf{h}}_{t_J,k} \rangle \mathbf{h}_{t_J,k} + \sum_{j=1}^J \sum_{k=1}^N \langle \mathbf{f}, \tilde{\mathbf{g}}_{t_j,k} \rangle \mathbf{g}_{t_j,k}.$$

7.2 Transformée en ondelettes décimée sur graphes dirigés

7.2.1 Notations

Nous rappelons d'abord des notations nécessaires et utiles pour une compréhension claire de la construction des ondelettes de diffusion [30, 71]. Deux notations sont introduites : l'une pour la représentation des transformations linéaires sous forme de matrices, et l'autre pour la représentation des ensembles de vecteurs sous forme de matrices dont les colonnes correspondent à un vecteur de l'ensemble donné. Ces notations ont été utilisées dans [30, 71] sauf que nous adoptons une convention en vecteur colonne alors qu'une convention en vecteur ligne a été utilisée dans [30, 71].

Soit $\mathbf{V}_0 = \ell^2(\mathcal{V})$ l'espace des fonctions définies sur les sommets d'un graphe dirigé $\mathcal{G} = (\mathcal{V}, \mathcal{E})$. Si $\mathbf{L}$ est une transformation linéaire de $\mathbf{V}_0$ sur $\mathbf{V}_0$, $[\mathbf{L}]_{B_1}^{B_2}$ indique la matrice représentant la transformation linéaire $\mathbf{L}$ par rapport à la base du domaine de départ B_1 et à la base du domaine d'arrivée B_2 . Un ensemble de vecteurs X représentés dans une base donnée B sera écrit sous la forme matricielle $[X]_B$ où les colonnes de $[X]_B$ sont les coordonnées des vecteurs X dans le système de coordonnées défini par B . Généralement, $[B_1]_{B_2} = [\mathbf{I}]_{B_1}^{B_2}$ représente les vecteurs de base B_1 en fonction de la base B_2 . On notera que pour une base donnée B si les bases de départ et d'arrivée sont identiques, $[\mathbf{I}]_B^B = \mathbf{I}$ la matrice identité. Nous abuserons également de cette notation si B_2 couvre un sous-espace de l'espace couvert par B_1 . Dans ce cas, $[\mathbf{I}]_{B_1}^{B_2}$ représentera le projecteur orthogonal de l'espace généré par B_1 sur l'espace généré par B_2 . Si B_1 et B_2 sont deux bases, les représentations de X dans B_1 et B_2 sont liées comme suit : $[X]_{B_2} = [\mathbf{I}]_{B_1}^{B_2} [X]_{B_1}$.

Si B_1, B_2, B_3, B_4 sont des bases arbitraires, la représentation matricielle de $\mathbf{L}$ par rapport au domaine de départ B_1 et au domaine d'arrivée B_2 des différentes bases est exprimée de la manière suivante :

$$[\mathbf{L}]_{B_1}^{B_2} = [\mathbf{I}]_{B_4}^{B_2} [\mathbf{L}]_{B_3}^{B_4} [\mathbf{I}]_{B_2}^{B_3}.$$

De plus, si B_1 et B_2 sont des vecteurs linéairement indépendants qui ne couvrent pas tout l'espace $\mathbf{V}_0$ alors nous utiliserons toujours la notation $[\mathbf{L}]_{B_1}^{B_2}$, mais dans ce cas elle représente la restriction de la transformation linéaire $\mathbf{L}$ au domaine de départ et d'arrivée couvrant B_1 et B_2 .

7.2.2 Ondelettes de diffusion

La construction des ondelettes de diffusion [30, 71] permet une analyse multirésolution des fonctions sur graphes généralisant le concept d'analyse multirésolution classique [27]. Le point de départ de cette construction est un opérateur de diffusion $\mathbf{T}$. Comme pour l'analyse multirésolution classique, les ondelettes de diffusion sont caractérisées par une famille de sous-espaces imbriqués $\mathbf{V}_0 \supseteq \mathbf{V}_1, \dots, \mathbf{V}_j \supseteq \dots$, où chaque sous-espace $\mathbf{V}_j$ est représenté par une fonction d'échelle Φ_j . Le complément de $\mathbf{V}_{j+1}$ en $\mathbf{V}_j$ s'appelle $\mathbf{W}_j$ est couvert par un ensemble d'ondelettes de diffusion Ψ_j . Nous précisons que notre construction est adaptée également au cas biorthogonal [71].

7.2.3 Construction

La construction des ondelettes de diffusion se déroule avec un opérateur de diffusion $\mathbf{T}$ défini sur un graphe dirigé fortement connexe $\mathcal{G} = (\mathcal{V}, \mathcal{E})$. Dans le cadre originel des ondelettes de diffusion de Coifman et Maggioni [30], le graphe est supposé non dirigé et ils suggèrent d'utiliser l'opérateur de marche aléatoire réversible $\mathbf{P}$.

Plus généralement, n'importe quel opérateur $\mathbf{T}$ peut être utilisé tant qu'il s'agit d'un opérateur associé à une réponse en fréquence passe-bas lorsqu'il est appliqué à des signaux sur graphe localisés. L'opérateur $\mathbf{T}$ doit se comporter comme un filtre passe-bas. Étant donné un graphe dirigé $\mathcal{G}$, un choix approprié est d'utiliser un opérateur $\mathbf{T} = \mathbf{\Pi}^{\frac{1}{2}} \mathbf{P} \mathbf{\Pi}^{-\frac{1}{2}}$ où $\mathbf{P}$ est l'opérateur de marche aléatoire associé à une réponse en fréquence passe-bas par rapport à la définition proposée dans la section 5. Si la matrice de transition de la marche aléatoire n'est pas nécessairement un opérateur passe-bas, ce sera par contre toujours le cas de la marche aléatoire paresseuse $\tilde{\mathbf{P}}_{1/2}$.

En supposant un ordre de sommets de $\mathcal{V}$, $\mathbf{T}$ est initialement représenté dans la base canonique $\Phi_0 = \{\delta_k\}_{k \in \mathcal{V}}$ de $\mathbf{V}_0$ où δ_k est la fonction delta de Kronecker appliquée au sommet $k \in \mathcal{V}$. En utilisant les notations de la section 7.2.1, $[\mathbf{T}]_{\Phi_0}^{\Phi_0}$ est la représentation matricielle de l'opérateur linéaire $\mathbf{T}$ par rapport à la base Φ_0 dans son domaine de départ et Φ_0 dans son domaine d'arrivée.

Les colonnes de la matrice $[\mathbf{T}]_{\Phi_0}^{\Phi_0}$ peuvent être décrites comme un ensemble de fonctions $\tilde{\Phi}_1 = \{\tilde{\phi}_{1,k}\}_{k \in \mathcal{V}}$ représentées dans la base Φ_0 . Chaque élément $\tilde{\phi}_{1,k}, k \in \mathcal{V}$ de $\tilde{\Phi}_1$ correspond à l'action de $\mathbf{T}$ sur la fonction de Kronecker δ au sommet k et est représenté comme $[\tilde{\phi}_{1,k}]_{\Phi_0} = [\mathbf{T}]_{\Phi_0}^{\Phi_0} \delta_k$.

La propriété passe-bas de $\mathbf{T}$ entraîne que chaque élément $\tilde{\phi}_{1,k}$ de $\tilde{\Phi}_1$ est une fonction localisée autour de son sommet k dont le support s'étend à ses voisins proches. En termes de construction d'ondelettes de diffusion, les éléments $\tilde{\phi}_{1,k}, k \in \mathcal{V}$ sont donc interprétés comme des fonctions *d'échelle*. Étant donné que le support de chaque élément $\tilde{\phi}_{1,k}$ couvre un petit voisinage autour du sommet k , ces éléments $\tilde{\phi}_{1,k}$ peuvent généralement être bien approchés par une combinaison linéaire des autres fonctions $\tilde{\phi}_{1,l}$ avec $l \neq k$. L'étape suivante de la construction est une étape de sélection du sous-ensemble de colonnes [72].

Le but est de trouver un petit nombre de colonnes de $\tilde{\Phi}_1$, formant un ensemble $\mathcal{C}$ tel que le $\|[\tilde{\Phi}_1]_{\Phi_0} - \Pi_{\mathcal{C}}[\tilde{\Phi}_1]_{\Phi_0}\|_{\beta}$ est aussi petit que possible où $\Pi_{\mathcal{C}}$ est la matrice de projection sur l'espace couvert par les colonnes $\mathcal{C}$ et β désigne la norme spectrale ou la norme de Frobenius.

Plus précisément, cette étape consiste à sélectionner un sous-ensemble $\mathcal{C} = \{\tilde{\phi}_{1,k}, k \in \mathcal{I}_1\}$ de $\tilde{\Phi}_1$, $|\mathcal{I}_1| \leq |\mathcal{V}|$ de sorte que le sous-espace défini représente le mieux possible les éléments de $\tilde{\Phi}_1$. En termes classiques de traitement du signal, cette étape est analogue à un sous-échantillonnage d'un ensemble de fonctions d'échelle dans la transformée en ondelettes discrètes classique. Coifman et Maggioni ont utilisé une approche gloutonne pour construire l'ensemble $\mathcal{I}_1$ de manière itérative en utilisant une procédure d'orthogonalisation de Gram-Schmidt modifiée [30].

Le sous-ensemble $\{\tilde{\phi}_{1,k}, k \in \mathcal{I}_1\}$ couvre un sous-espace $\mathbf{V}_1$ qui correspond au premier sous-espace d'approximation de l'analyse multirésolution. Nous désignerons donc $\{\tilde{\phi}_{1,k}, k \in \mathcal{I}_1\}$ comme Φ_1 , où Φ_1 est par définition une base de $\mathbf{V}_1$ qui est généralement non orthogonale. Les vecteurs de Φ_1 sont les fonctions d'échelle d'ordre 1.

Pour définir les échelles suivantes des fonctions d'échelle, nous considérons l'étape de compression¹ de l'opérateur de diffusion $\mathbf{T}$ sur le sous-espace $\mathbf{V}_1$, c'est-à-dire sa restriction sur $\mathbf{V}_1$. Ce dernier peut être représenté en Φ_1 comme

$$[\mathbf{T}]_{\Phi_1}^{\Phi_1} = [\mathbf{I}]_{\Phi_0}^{\Phi_1} [\mathbf{T}]_{\Phi_0}^{\Phi_0} [\mathbf{I}]_{\Phi_1}^{\Phi_0}.$$

où $[\mathbf{I}]_{\Phi_0}^{\Phi_1}$ représente la restriction de l'espace $\mathbf{V}_0$ à $\mathbf{V}_1$ (avec bases respectives Φ_0 et Φ_1) et $[\mathbf{I}]_{\Phi_1}^{\Phi_0}$ représentant le plongement de $\mathbf{V}_1$ dans $\mathbf{V}_0$.

L'espace d'approximation suivant $\mathbf{V}_2 \subset \mathbf{V}_1$ et sa base associée Φ_2 peuvent

1. Représenter $\mathbf{T}_1$ dans Φ_1 peut être vu comme de la compression avec perte. Fondamentalement, $[\mathbf{T}]_{\Phi_0}^{\Phi_0} \approx [\mathbf{I}]_{\Phi_1}^{\Phi_0} [\mathbf{T}]_{\Phi_1}^{\Phi_1} [\mathbf{I}]_{\Phi_0}^{\Phi_1}$.

être obtenus de la même manière que la définition de V_1 et sa base Φ_1 sauf que l'opérateur $\mathbf{T}^2$ est maintenant limité à V_1 plutôt que $\mathbf{T}$ sur V_0 .

Les colonnes de $[\mathbf{T}^2]_{\Phi_1}^{\Phi_1} \approx ([\mathbf{T}]_{\Phi_1}^{\Phi_1})^2$ peuvent être interprétées comme fonctions d'échelle à l'échelle 2, $\tilde{\Phi}_2 = \{\tilde{\phi}_{2,k}\}$, représenté dans la base Φ_1 . De ces fonctions, nous extrayons un sous-ensemble $\Phi_2 = \{\tilde{\phi}_{2,k}, k \in \mathcal{I}_2\}$ tel que toute fonction dans $\tilde{\Phi}_2$ est bien approchée par une combinaison linéaire de fonctions de Φ_2 .

Après j itérations de cette procédure, nous avons défini les sous-espaces d'approximation $V_j \subset V_{j-1} \subset \dots \subset V_1$ avec les bases correspondantes $\Phi_j, \Phi_{j-1}, \dots, \Phi_1$. A chaque étape, la base Φ_j est définie par sa représentation dans la base Φ_{j-1} basée sur la restriction de l'opérateur $\mathbf{T}^{2^j}$ à V_{j-1} . Pour représenter ces fonctions dans la base originale Φ_0 de V_0 nous pouvons utiliser :

$$[\Phi_j]_{\Phi_0} = [\mathbf{I}]_{\Phi_j}^{\Phi_0} = [\mathbf{I}]_{\Phi_0}^{\Phi_0} [\mathbf{I}]_{\Phi_1}^{\Phi_0} \dots [\mathbf{I}]_{\Phi_{j-1}}^{\Phi_{j-2}} [\mathbf{I}]_{\Phi_j}^{\Phi_{j-1}}.$$

Puisque chaque fonction dans Φ_0 est définie sur V_0 , chaque fonction sur V_j l'est aussi. Ainsi, toute fonction dans l'espace d'approximation V_j peut être étendue naturellement à l'espace entier V_0 .

En ce qui concerne la construction des ondelettes, nous proposons de construire les bases d'ondelettes Ψ_j pour les sous-espaces W_j en sélectionnant un sous-ensemble des colonnes de l'opérateur passe bande $[\mathbf{I}_{V_j} - \Phi_{j+1} \Phi_{j+1}^\dagger]_{\Phi_j}^{\Phi_j}$ qui est le projecteur orthogonal sur le complément V_{j+1} dans V_j . Les ondelettes capturent le détail perdu en passant de V_j à V_{j+1} . Comme notre cadre s'inscrit dans un cadre bi-orthogonal, nous mentionnons que la construction d'ondelettes dans le cadre biorthogonal, i.e dans [38] n'est pas abordée.

7.2.4 Généralisation avec opérateurs de moyennage arbitraires

Nous proposons une généralisation du cadre des ondelettes de diffusion qui permet de le combiner par exemple avec la construction des ondelettes présentée dans la section 7.1. L'idée est simplement de remplacer les puissances de $\mathbf{T}$ comme opérateurs de fonction d'échelle multirésolution par des filtres passe-bas plus arbitraires. Plus précisément, quand l'opérateur $\mathbf{T}^{2^j}$ est utilisé pour définir l'espace d'approximation V_{j+1} dans la section 7.2.3, nous proposons d'utiliser un filtre sur graphe passe-bas $\mathbf{H}_j$. Si les filtres sur graphe $\mathbf{H}_j$ correspondent aux opérateurs de fonction d'échelles définis en section 7.1 avec des échelles convenablement croissantes, cette approche permettrait de réduire la redondance des ensembles de fonctions d'échelle. De même, la même approche pourrait être utilisée pour réduire la redondance des fonctions d'ondelettes définies également en section 7.1.

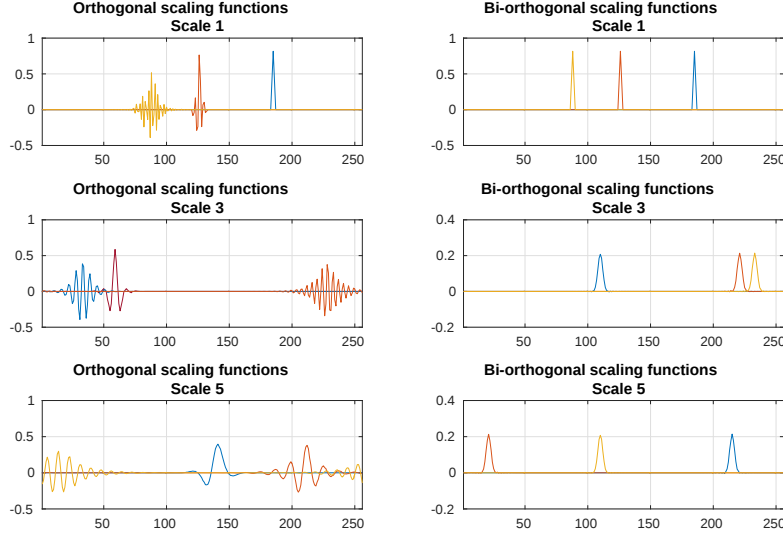


FIGURE 4.14 – Fonctions d’échelle orthogonal et biorthogonal sur un graphe cyclique dirigé $\mathcal{C}_{256}$.

7.3 Applications

Pour les sous-sections suivantes, les constructions multi-échelles commenceront avec l’opérateur

$$\mathbf{T} = \mathbf{\Pi}^{1/2} \bar{\mathbf{P}} \mathbf{\Pi}^{-1/2},$$

en tant qu’opérateur passe-bas. Comme mentionné dans la section 3.4, l’opérateur $\mathbf{T}$ est l’équivalent de $\bar{\mathbf{P}}$ dans l’espace de Hilbert $\ell^2(\mathcal{V})$ et est adapté aux applications pratiques.

7.3.1 Analyse multirésolution sur le graphe cyclique dirigé

Nous montrons un exemple d’analyse multirésolution sur le graphe cyclique $\mathcal{C}_N$ avec $N = 256$. Cette exemple a été analysé auparavant dans [30, 38]. La nouveauté est son adaptation au cas dirigé. Nous utilisons les mêmes hypothèses que dans la section 5.4.1. Nous construisons des analyses multirésolutions orthogonales et biorthogonales sur $\mathcal{C}_N$ à travers le cadre des ondelettes de diffusion appliquées sur les puissances dyadiques de $\mathbf{T}$, soit $\{\mathbf{T}^{2^j}\}_{j=1}^J$. Nous fixons le nombre d’échelles à $J = 6$.

La figure 4.14 montre les fonctions d’échelle orthogonales et biorthogonales aux échelles 1, 3 et 5. Pour chaque échelle, nous représentons 3 fonctions d’échelle.

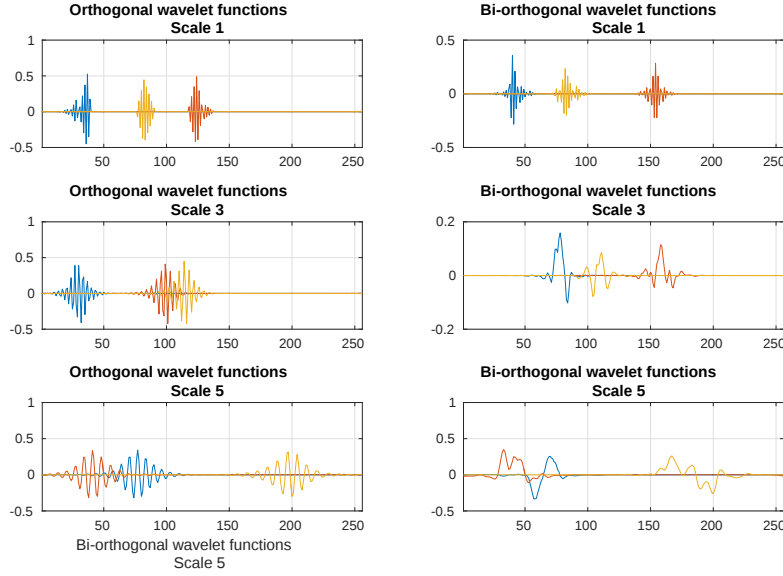


FIGURE 4.15 – Ondelettes orthogonales et biorthogonales sur le graphe cyclique dirigé $\mathcal{C}_{256}$.

La figure 4.15 montre les ondelettes orthogonales et biorthogonales aux échelles 1, 3 et 5.

Nous remarquons, pour chaque échelle observée, que le support des fonctions d'échelles orthogonales est plus large que le support des fonctions d'échelle biorthogonales. Nous notons également aux échelles 3 et 5 que les ondelettes biorthogonales ont un support légèrement plus petit que le support des ondelettes orthogonales. Par ailleurs, nous remarquons que les fonctions d'échelle orthogonales ont plus d'oscillations que leur versions biorthogonales. Cela signifie que les fonctions d'échelle et ondelettes ont une mauvaise localisation en fréquence. Les conclusions à propos de la localisation spatiale des fonctions d'échelle sont identiques aux conclusions dans [38]. La nouveauté ici est que nous travaillons sur le graphe cyclique dirigé et que nous montrons les ondelettes orthogonales et biorthogonales en plus de fonctions d'échelle orthogonales et biorthogonales.

7.3.2 Analyse multirésolution sur la version dirigée du graph Watts-Strogatz

Dans cette section, nous montrons un exemple d'analyse multirésolution sur la version dirigée du modèle Watts-Strogatz. Le modèle Watts-Strogatz [73] est un modèle de graphe aléatoire présentant les propriétés de petit monde, y compris les courtes longueurs moyennes de trajet et les clusters élevés [74].

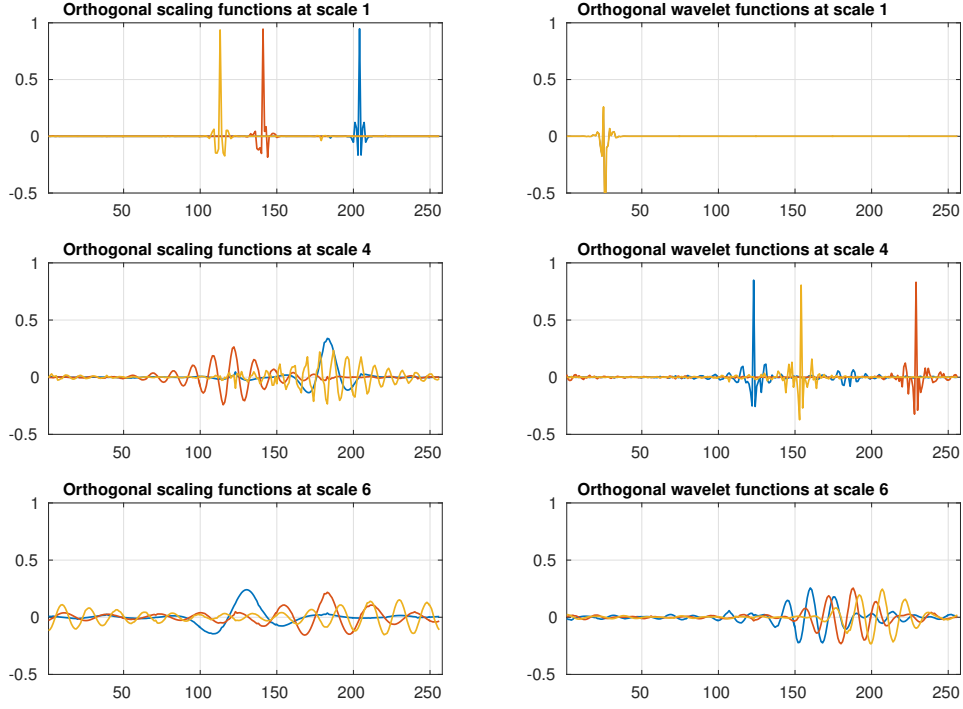


FIGURE 4.16 – Fonctions d'échelle et ondelettes orthogonales sur $\mathcal{G}_2$.

La construction d'un graphe dirigé provenant du modèle Watts-Strogatz commence avec un graphe cyclique dirigé à N sommets. Chaque sommet est connecté à ses k prochains sommets suivant la sens de l'orientation du graphe cyclique dirigé. L'arc qui relie i à chacun de ses plus proches voisins est rebranché au hasard avec une probabilité de p ou reste intact avec une probabilité de $1 - p$. Cette procédure est répétée cycliquement pour chaque sommet successif jusqu'à ce que le sommet i soit à nouveau sélectionné. Nous notons $\mathcal{G} \sim \text{DWS}(N, K, \beta)$ un graphe construit suivant la version dirigée du modèle Watts Strogatz avec N sommets, K plus proches voisins et une probabilité de rebranchement β .

Ici, nous faisons des analyses multi-résolutions orthogonales et biorthogonales sur un graphe $\mathcal{G}_2 \sim \text{DWS}(256, 2, 0.02)$. Les analyses multi-résolution sont réalisées dans le cadre des ondelettes de diffusion appliquées sur les puissances dyadiques de $\mathbf{T}$, soit $\{\mathbf{T}^{2^j}\}_{j=1}^J$. Nous fixons le nombre d'échelles à $J = 10$.

La figure 4.16 montre les fonctions d'échelle orthogonales et les ondelettes également à l'échelle 1, 4 et 6. A chaque échelle, nous représentons 3 ou 4 fonctions d'échelle et d'ondelettes.

La figure 4.17 montre quelques fonctions d'échelle biorthogonale et des ondelettes à l'échelle 1, 4 et 6 sur 10, localisées à différents sommets. Ils res-

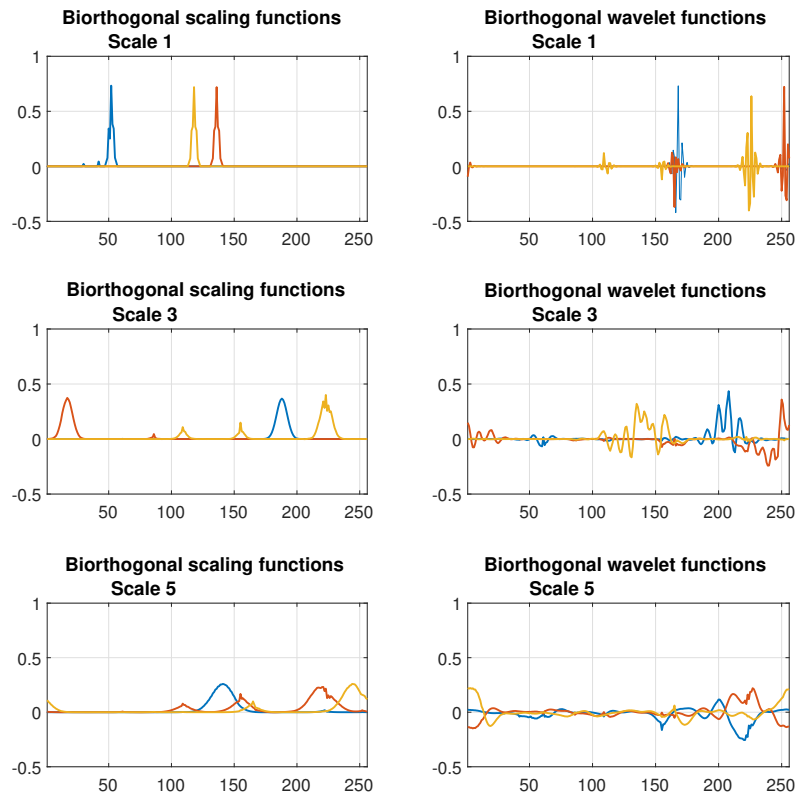


FIGURE 4.17 – Fonctions d'échelle et ondelettes biorthogonales sur $\mathcal{G}_2$.

semblent à ce que l'on attend des fonctions d'échelle (fonctions d'agrégation passe-bas) et des ondelettes (fonctions oscillantes localisées).

7.3.3 Apprentissage semi-supervisé sur graphes dirigés avec régularisation de type ℓ_1

Nous discutons de l'approche d'apprentissage semi-supervisé dans le cas de fonctions sur graphes dirigés avec régularisation ℓ_1 . Des approches précédentes d'apprentissage semi-supervisé avec régularisation ℓ_1 sur graphes non dirigés ont été étudiées à l'aide d'ondelettes spectrales sur graphes [75] ou d'ondelettes de type spline sur graphe à échantillonnage critique [76].

Notre but est de montrer que la performance du problème d'apprentissage semi-supervisé avec la régularisation ℓ_1 est compétitive par rapport aux approches existantes, par exemple le problème d'apprentissage semi-supervisé étudié dans la section 6. Nous discutons du problème d'apprentissage semi-supervisé suivant, toujours sur l'exemple de l'ensemble de données des blogs politiques [70]. Nous utilisons les mêmes notations décrites dans la section 6 et travaillons sur le sous-graphe $\mathcal{G}'$. Le problème d'apprentissage semi-supervisé avec régularisation de type ℓ_1 est :

$$\mathbf{w}^* = \underset{\mathbf{w}}{\operatorname{argmin}} \|\tilde{\mathbf{y}} - \mathbf{MK}\mathbf{w}\|_2^2 + \lambda \|\mathbf{w}\|_1, \quad \lambda \in \mathbb{R}^+. \quad (4.47)$$

Le signal sur graphe $\tilde{\mathbf{y}} = \mathbf{M}\mathbf{y}$ est le signal sur graphe partiellement étiqueté avec $\mathbf{M} = \{m_{ij}\}_{1 \leq i, j \leq N}$ l'opérateur de masquage, c'est-à-dire la matrice diagonale où $m_{ii} = \mathbb{1}_{v_i \in \mathcal{O}}$ où $\mathcal{O} \subset \mathcal{V}$ est le sous-ensemble des étiquettes connues. La matrice $\mathbf{K} = (\mathbf{H}_J, \mathbf{G}_1, \dots, \mathbf{G}_J)$ est le filtre de synthèse. Si on met $\mathbf{X} = \mathbf{MK}$, l'équation (4.47) peut être réécrite comme suit :

$$\mathbf{w}^* = \underset{\mathbf{w}}{\operatorname{argmin}} \|\tilde{\mathbf{y}} - \mathbf{X}\mathbf{w}\|_2^2 + \lambda \|\mathbf{w}\|_1, \quad \lambda \in \mathbb{R}.$$

Cette dernière formulation est connue sous le nom de régression LASSO [77]. La formulation (4.47) est identique à la formulation du problème de restauration du signal avec des transformées d'ondelettes redondantes [78], sauf que $\mathbf{K}$ est ici le banc filtre de synthèse pour les fonctions définies sur des graphes dirigés. Le problème d'apprentissage semi-supervisé avec régularisation ℓ_1 peut être résolu par exemple à partir de méthodes de gradient proximal [79, 80, 81]. En conséquence, le signal restauré indiqué par le symbole $\mathbf{f}^*$ est :

$$\mathbf{f}^* = \operatorname{signe}(\mathbf{K}\mathbf{w}^*).$$

Nous comparons les performances des formulations (4.47) et (4.30). La formulation (4.47) doit construire le banc de filtres de synthèse $\mathbf{K}$. Nous choisissons de construire un banc de filtres de synthèse à deux échelles $\mathbf{K} = (\mathbf{H}_2, \mathbf{G}_1, \mathbf{G}_2)$ où nous proposons une construction du noyau de chaleur [46]

$$\mathbf{H}_j = e^{-\mathcal{L}t_j}, \quad t_j = 2^j, \quad j = 1, 2.$$

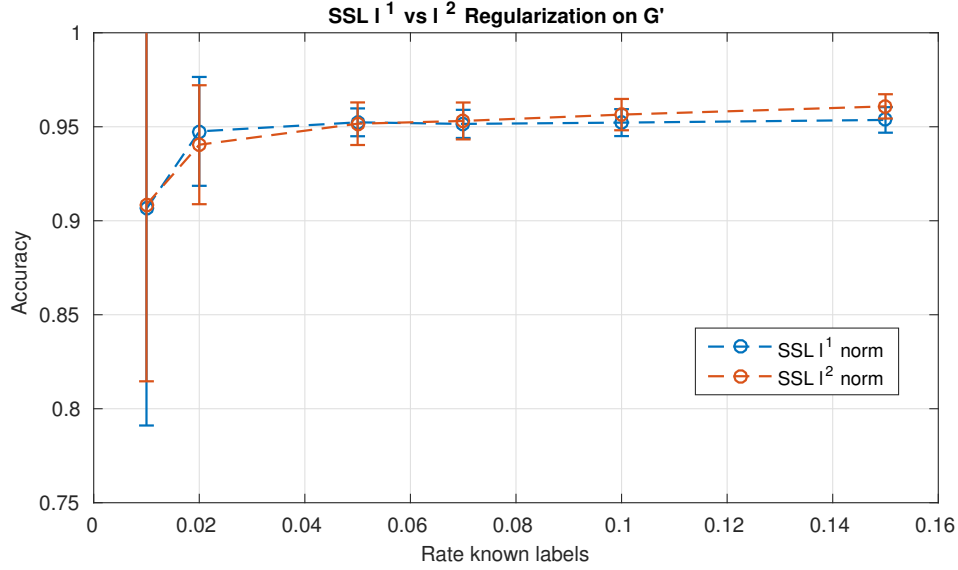


FIGURE 4.18 – Apprentissage semi-supervisé sur la plus grande composante connexe du graphe des blogs politiques.

avec $\mathcal{L}$ le laplacien normalisé sur graphe dirigé (4.17) et les filtres passe bandes sur graphe $\mathbf{G}_1, \mathbf{G}_2$ tels que

$$\mathbf{G}_2 = \mathbf{H}_1 - \mathbf{H}_2, \quad \mathbf{G}_1 = \mathbf{I} - \mathbf{H}_1.$$

La figure 4.18 montre la performance de l'apprentissage semi-supervisé entre la formulation (4.47) et (4.30) sur le plus grand graphe fortement connexe obtenu des blogs politiques américains. Le taux de performance est obtenu en faisant la moyenne de 200 réalisations et en déterminant les paramètres associés à chacune des approches donnant les meilleurs taux de performance.

Sur la figure 4.18, le taux de performance basé sur la pénalisation ℓ_1 des coefficients d'ondelettes redondantes sur les graphes dirigés est compétitif par rapport à la pénalisation ℓ_2 du terme Dirichlet pour tous les pourcentages d'étiquettes connues. Plus précisément, pour un faible taux d'étiquettes connues, le taux de performance basé sur la pénalisation ℓ_1 est légèrement meilleur que le problème ℓ_2 et comme le taux d'étiquettes connues augmente, le problème ℓ_1 (4.47) est concurrentiel comparativement au problème ℓ_2 (4.30). Par conséquent, l'approche d'apprentissage semi-supervisé basée sur la pénalisation ℓ_1 des coefficients d'ondelettes sur les graphes dirigés est prometteuse afin de reconstruire les fonctions sur les graphes dirigés avec un faible taux d'étiquettes connues.

Généralement, dans le traitement de signal classique les approches du

type ℓ_1 en ondelettes surpassent les méthodes en variation ℓ_2 pour des signaux réguliers par morceaux. Particulièrement, si nous considérons des ondelettes de Haar pour des signaux constant par morceaux. La moins bonne performance de la méthode ℓ_1 sur graphes (4.47) par rapport à la méthode ℓ_2 est probablement due au support assez large de nos fonctions d'ondelettes (bien plus large que les ondelettes de Haar). Afin d'améliorer les performances de la méthode (4.47), il pourrait être judicieux de construire de nouvelles ondelettes avec une meilleure caractérisation espace-fréquence, et de choisir le nombre d'échelles le plus approprié en fonction des caractéristiques du jeu de données.

Chapitre 5

Conclusion et Perspectives

Dans cette thèse, nous avons présenté une nouvelle analyse harmonique pour fonctions définies sur les sommets d'un graphe dirigé fortement connexe. Nous avons proposé une analyse en fréquence pour des fonctions définies sur un graphe dirigé. Cette analyse en fréquence est fondée sur l'étude de la variation des vecteurs propres de l'opérateur de marche aléatoire décrivant une marche aléatoire ergodique sur ce graphe dirigé. À partir de cette construction de bases de Fourier sur graphes dirigés, nous avons proposé une construction d'ondelettes redondantes généralisant le cadre des ondelettes spectrales sur graphes [32, 31] et une construction en ondelettes à échantillonnage critique étendant le cadre des ondelettes de diffusion [30, 38] au cas des graphes dirigés. Par ailleurs, notre analyse harmonique coïncide avec le cadre de l'analyse harmonique sur graphes non dirigés par le lien qui existe entre l'opérateur de marche aléatoire sur graphe non dirigé et les laplaciens de graphe correspondants. Enfin, nous avons illustré notre analyse harmonique à travers des exemples d'apprentissage semi-supervisé et de modélisation de signaux sur graphes dirigés et montré la pertinence de notre cadre face aux approches existantes.

Dans le cadre de cette thèse et plus généralement de l'analyse harmonique sur graphes, quelques perspectives de recherche peuvent être émises :

Perspectives théoriques

- **Analyse de Fourier et plongements de graphe dirigés.** La limite de graphes de taille infinie, en particulier des graphes dirigés est une perspective de recherche intéressante à aborder. En effet, l'étude de la limite de graphes non dirigés de taille infinie est déjà bien établie [82, 83, 84]. Une des questions intéressantes serait d'étudier le comportement asymptotique de l'analyse de Fourier lorsque le nombre de sommets d'un graphe dirigé tend asymptotiquement vers l'infini. Ces questions ont été déjà soulevées pour montrer la convergence des vecteurs propres et valeurs propres du laplacien de

graphe vers les fonctions propres et valeurs propres de l'opérateur de Laplace-Beltrami lorsque le nombre de sommets tend vers l'infini sur une variété sous-jacente [18]. Dans le cas des graphes dirigés, il serait intéressant d'étudier tout d'abord les cas simples du graphe cyclique dirigé et du graphe toroïdal dirigé vers leurs versions continues respectives, c'est-à-dire si on fait tendre le nombre de sommets vers l'infini, et d'établir un lien entre l'analyse de Fourier bien connue de ces objets continus avec leurs versions approximées par un graphe dirigé. À partir de l'étude théorique de ces cas, cela nous apporterait éventuellement des pistes pour étudier le cas de plongement de graphes dirigés.

Le plongement d'un graphe dirigé est une question supplémentaire qui pourrait faire suite à cette thèse, de la même manière que le plongement d'un graphe non dirigé dans une variété a été abordé précédemment [19]. Dans certains cas, il y a un intérêt à considérer les sommets d'un graphe non dirigé comme un échantillonnage d'une variété sous-jacente et les arcs comme les observations macroscopiques d'un noyau de diffusion entre des points voisins sur la variété. Dans le cas des graphes dirigés, le graphe n'est a priori pas un échantillonnage d'un objet bien précis. Néanmoins dans le cas où on suppose que l'objet est une variété, des approches [85, 86] ont discuté du plongement d'un graphe dirigé dans une variété en recouvrant les caractéristiques essentielles du processus d'échantillonnage : géométrie de la variété, distribution de l'échantillonnage et la directionnalité globale sous l'hypothèse de limite d'un échantillon infini. Il serait intéressant d'explorer ces pistes en profondeur et si possible de proposer une alternative, la question de plongement d'un graphe dirigé étant loin d'être résolue.

Une autre question importante serait de discuter de moyens de faire le lien entre discret et continu. Dans le cadre non dirigé, des éléments de réponses ont été apportés [18, 19] à travers la convergence entre le laplacien de graphe et l'opérateur de Laplace. Un élément de piste de recherche dans le cas dirigé pourrait être de partir de la marche aléatoire sur le graphe dirigé et d'en déterminer la limite continue. Un élément de réponse a été apporté par Hashimoto et al. [87, 88] pour des graphes dirigés de type KNN (*K-nearest neighbors directed graphs* en anglais) où la marche aléatoire convergerait vers un processus d'Îto.

- **Réduction (multirésolution) de graphes dirigés.** La réduction de graphes est une question de recherche particulièrement intéressante mais pour laquelle il n'existe pas de solution unique. En effet, dans le cadre de l'analyse harmonique sur graphes, plusieurs des constructions d'ondelettes proposent une analyse multirésolution de fonctions où les différentes échelles sont représentées sur le même graphe [31, 32, 26]. Il est donc intéressant de proposer une construction multirésolution

qui réduirait la dimension du signal (fonction sur le graphe) à chaque échelle ainsi que du graphe en même temps. Dans le cas des graphes bipartites, la méthode de sous-échantillonnage du signal sur un des graphes coset a été étudiée et validée dans [89] pour des constructions de bancs de filtres. Dans le cas des graphes arbitraires non dirigés, des analyses multirésolution ont été proposées soit en divisant par deux le nombre de sommets à chaque échelle puis en réduisant le graphe en utilisant la réduction de Kron du laplacien de graphe [90], soit en faisant une partition du graphe en sous-graphes connectés [35] puis en réduisant le graphe en créant des super-sommets étant l'agrégation du sous-graphe connecté. Mais ces approches ad hoc ne se généralisent pas simplement à d'autres constructions d'ondelettes. Dans le cas des ondelettes de diffusion [30, 38, 26], l'étape de compression de l'opérateur lié à la notion de diffusion s'apparente à une réduction en taille de l'opérateur et donc une réduction du graphe associé où chaque sommet à l'échelle J serait constitué d'une sorte de combinaison des sommets à l'échelle $J - 1$. Lorsque l'opérateur en question est un opérateur de marche aléatoire sur graphe non dirigé, il est possible de partitionner les signaux sur graphes et de réduire (donc sous-échantillonner) le nombre de sommets retenus, en se fondant sur la dynamique à travers les échelles de plus en plus stationnaire de la marche aléatoire sur le graphe non dirigé en question [91], ou de faire un sous-échantillonnage en utilisant les arbres couvrant de poids minimal issus de la marche aléatoire sur graphe [37].

Dans le cas de notre analyse multirésolution sur graphes dirigés, et en utilisant l'opérateur issu de la réversibilisation additive de la marche aléatoire ou son laplacien associé, il serait possible d'étendre les approches [91, 37, 90]. Par conséquent, les méthodes de réduction de graphes pourraient être éventuellement employées pour l'étape de "mise en commun" (*pooling* en anglais) dans la construction de réseaux de neurones convolutifs sur graphes dirigés. L'approche utilisant la réduction de Kron n'a pas été abordée pour le cas des réseaux convolutifs sur graphes non dirigés et serait une autre voie à explorer.

- **Localisation spatiale des ondelettes sur graphes dirigés.** La question de la localisation spatiale des ondelettes n'a pas pu être abordée dans cette thèse. Il s'agit d'une question théorique importante qui demanderait à être analysée. La question de la localisation spatiale des ondelettes a été abordée dans le cadre des ondelettes spectrales sur graphes non dirigés dans [32] et plus récemment dans [92]. Il ne semble pas possible d'étendre directement les propriétés de localisation spatiale des ondelettes sur graphes dirigés en utilisant comme opérateur de référence l'opérateur de marche aléatoire, excepté si les ondelettes spectrales sont construites à partir de la réversibilisation additive de la marche aléatoire sur le graphe dirigé. Par conséquent, analyser les

propriétés de localisation spatiale des ondelettes construites à partir de l'opérateur de marche aléatoire sur graphe dirigé reste un travail pour le futur.

Perspectives appliquées

- **Réseaux de neurones sur graphes dirigés.** Les réseaux de neurones sont des réalisations populaires de ces hiérarchies multi-couches profondes. Les modèles d'apprentissage profond sont particulièrement performants en ce qui concerne les signaux tels que la parole, les images ou la vidéo dans lesquels il existe une structure euclidienne. Ainsi dû à l'efficacité de l'apprentissage profond, il y a eu récemment un intérêt grandissant d'appliquer les concepts d'apprentissage profond aux signaux vivant sur des structures non euclidiennes tels que les graphes, de la même manière que l'on a pu constater l'efficacité des ondelettes pour des signaux vivant sur des structures euclidiennes. Ce cadre est appelé "apprentissage profond géométrique" [93] (*geometric deep learning* en anglais). La première construction de réseaux de neurones convolutifs sur graphes a été proposée par Bruna et coll. [94] utilisant la définition de convolution dans le domaine spectral et filtrage à partir de la décomposition spectrale du laplacien de graphe où les vecteurs propres constituent une base de Fourier sur graphe non dirigé. D'autres constructions ont été proposées, fondées sur la construction de filtres comme polynôme du laplacien de graphe [95] ou de l'opérateur de marche aléatoire [96]. Toutes ces constructions de réseaux de neurones convolutifs sur graphes ne s'appliquent jusqu'ici qu'aux graphes non dirigés. Par conséquent, un axe de recherche future serait d'étendre le cadre des réseaux de neurones convolutifs et de l'apprentissage profond en général au cas des graphes dirigés. Récemment, une construction de réseaux de neurones profond sur graphes dirigés a été proposée par Monti et coll. [97] qui resterait à comparer à un cadre utilisant des filtres comme dans [95] mais avec notre proposition d'opérateur de marche aléatoire comme opérateur de référence.

En effet, dans notre cadre théorique, nous avons proposé l'opérateur de marche aléatoire comme opérateur de référence pertinent sur graphe dirigé et nous avons également proposé la classe de combinaisons convexes entre l'opérateur de marche aléatoire et sa version retournée dans le temps $\bar{\mathcal{P}}$. Ainsi chaque élément de cette classe peut également être considéré comme un opérateur de référence dont les vecteurs propres constituent également une base de Fourier sur graphes. Il serait ainsi possible de proposer une nouvelle construction de réseaux convolutifs sur graphes dirigés à partir de $\bar{\mathcal{P}}$ et de la comparer avec la construction proposée par Monti et coll. [97].

Dans la totalité des approches de réseaux de neurones sur graphes non dirigés, les jeux de données sont des graphes de documents initialement dirigés. Ils considèrent exclusivement la version non dirigée de ces graphes. Ainsi les approches adaptées aux graphes dirigés sembleraient plus appropriées. Par ailleurs, la prédiction de liens dans un réseau social Twitter, l'apprentissage d'entités et des relations dans les "knowledge graphs" sont potentiellement des pistes d'application des réseaux de neurones sur graphes dirigés [98].

- **Apprentissage par renforcement et bandits sur graphes** Le domaine de l'apprentissage par renforcement sur graphes est enfin une dernière perspective où les travaux de cette thèse pourraient servir. Aujourd'hui l'apprentissage par renforcement est un domaine clé qui a permis une avancée majeure en apprentissage automatisé (voire en intelligence artificielle) avec des systèmes capables d'égaliser ou de surpasser les performances humaines sur des tâches précises, à l'exemple d'AlphaGo [99]. Le cadre de l'apprentissage par renforcement sur graphes a émergé il y a une dizaine d'années avec les travaux de Mahadevan et Maggioni, portant sur la résolution de processus markoviens décisionnels pour apprendre les politiques optimales fondées sur le laplacien sur graphe [100] et également sur une analyse multi-échelle des processus markoviens décisionnels [101]. Récemment, les travaux de Machado ont également porté sur une approche sur graphe de l'apprentissage par renforcement [102]. Il serait intéressant d'étendre le cadre de l'apprentissage par renforcement aux graphes dirigés grâce à l'approche développée dans cette thèse et ce serait approprié pour l'analyse de processus markoviens décisionnels.

L'apprentissage par renforcement s'appuie aussi sur des approches de bandits [103] sur graphes. L'usage de la théorie des bandits est répandu dans le cadre des systèmes de recommandation (e.g. Spotify, Netflix). Sur graphes non dirigés, l'approche des bandits sur graphes d'un point de vue spectral a été analysée par Valko et coll. [104] à travers l'analyse de Fourier des vecteurs propres du laplacien. Récemment, dans le cadre du traitement de données sur graphes, le développement d'algorithmes de bandits sur graphes a été publié [105]. Un exemple d'étude pourrait être la recommandation de produits à partir du graphe des achats d'Amazon (co-purchasing network en anglais) qui est un graphe dirigé. Il nous semble que l'adaptation d'algorithmes de bandits à partir de notre cadre théorique sur graphes dirigés pourrait être appropriée.

Bibliographie

- [1] Ed Bullmore and Olaf Sporns. Complex brain networks : graph theoretical analysis of structural and functional systems. *Nature Reviews Neuroscience*, 10(3) :186, 2009.
- [2] Peter J Carrington, John Scott, and Stanley Wasserman. *Models and methods in social network analysis*, volume 28. Cambridge university press, 2005.
- [3] Luigi Atzori, Antonio Iera, and Giacomo Morabito. The internet of things : A survey. *Computer networks*, 54(15) :2787–2805, 2010.
- [4] Christian Decker and Roger Wattenhofer. Information propagation in the bitcoin network. In *Peer-to-Peer Computing (P2P), 2013 IEEE Thirteenth International Conference on*, pages 1–10. IEEE, 2013.
- [5] Serguei Popov. The tangle. *cit. on*, page 131, 2016.
- [6] Marc Barthélemy. Spatial networks. *Physics Reports*, 499(1-3) :1–101, 2011.
- [7] Alain Barrat, Marc Barthelemy, and Alessandro Vespignani. *Dynamical processes on complex networks*. Cambridge university press, 2008.
- [8] Mikhail Belkin and Partha Niyogi. Laplacian eigenmaps for dimensionality reduction and data representation. *Neural computation*, 15(6) :1373–1396, 2003.
- [9] Mikhail Belkin, Irina Matveeva, and Partha Niyogi. Regularization and semi-supervised learning on large graphs. In *International Conference on Computational Learning Theory*, pages 624–638. Springer, 2004.
- [10] Evarist Giné, Vladimir Koltchinskii, et al. Empirical graph laplacian approximation of laplace–beltrami operators : Large sample results. In *High dimensional probability*, pages 238–259. Institute of Mathematical Statistics, 2006.
- [11] Ronald R Coifman and Stéphane Lafon. Diffusion maps. *Applied and computational harmonic analysis*, 21(1) :5–30, 2006.
- [12] Ulrike Von Luxburg. A tutorial on spectral clustering. *Statistics and computing*, 17(4) :395–416, 2007.

- [13] Steven Rosenberg. *The Laplacian on a Riemannian manifold : an introduction to analysis on manifolds*, volume 31. Cambridge University Press, 1997.
- [14] Fan RK Chung. *Spectral graph theory*. Number 92. American Mathematical Soc., 1997.
- [15] Olivier Lablée. *Spectral Theory in Riemannian Geometry*. 2015.
- [16] 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 domains. *IEEE Signal Processing Magazine*, 30(3) :83–98, 2013.
- [17] Antonio Ortega, Pascal Frossard, Jelena Kovačević, José MF Moura, and Pierre Vandergheynst. Graph signal processing. *arXiv preprint arXiv :1712.00468*, 2017.
- [18] Mikhail Belkin and Partha Niyogi. Convergence of laplacian eigenmaps. In *Advances in Neural Information Processing Systems*, pages 129–136, 2007.
- [19] Matthias Hein, Jean-Yves Audibert, and Ulrike von Luxburg. Graph laplacians and their convergence on random neighborhood graphs. *Journal of Machine Learning Research*, 8(Jun) :1325–1368, 2007.
- [20] Aliaksei Sandryhaila and Jose MF Moura. Discrete signal processing on graphs : Frequency analysis. *IEEE Trans. Signal Processing*, 62(12) :3042–3054, 2014.
- [21] Hrushikesh N Mhaskar. A unified framework for harmonic analysis of functions on directed graphs and changing data. *Applied and Computational Harmonic Analysis*, 2016.
- [22] Rasoul Shafipour, Ali Khodabakhsh, Gonzalo Mateos, and Evdokia Nikolova. A digraph fourier transform with spread frequency components. *arXiv preprint arXiv :1705.10821*, 2017.
- [23] Stefania Sardellitti, Sergio Barbarossa, and Paolo Di Lorenzo. On the graph fourier transform for directed graphs. *IEEE Journal of Selected Topics in Signal Processing*, 11(6) :796–811, 2017.
- [24] Thierry Coulhon and Alexander Grigoryan. Random walks on graphs with regular volume growth. *Geometric and Functional Analysis*, 8(4) :656–701, 1998.
- [25] Ravi Montenegro, Prasad Tetali, et al. Mathematical aspects of mixing times in markov chains. *Foundations and Trends® in Theoretical Computer Science*, 1(3) :237–354, 2006.
- [26] Harry Sevi, Gabriel Rilling, and Pierre Borgnat. Multiresolution analysis of functions on directed networks. In *Wavelets and Sparsity XVII*,

- volume 10394, page 103941Q. International Society for Optics and Photonics, 2017.
- [27] Stephane G Mallat. A theory for multiresolution signal decomposition : the wavelet representation. *IEEE transactions on pattern analysis and machine intelligence*, 11(7) :674–693, 1989.
 - [28] Mark Crovella and Eric Kolaczyk. Graph wavelets for spatial traffic analysis. In *INFOCOM 2003. Twenty-Second Annual Joint Conference of the IEEE Computer and Communications. IEEE Societies*, volume 3, pages 1848–1857. IEEE, 2003.
 - [29] Arthur D Szlam, Mauro Maggioni, Ronald R Coifman, and James C Bremer. Diffusion-driven multiscale analysis on manifolds and graphs : top-down and bottom-up constructions. In *Wavelets XI*, volume 5914, page 59141D. International Society for Optics and Photonics, 2005.
 - [30] Ronald R Coifman and Mauro Maggioni. Diffusion wavelets. *Applied and Computational Harmonic Analysis*, 21(1) :53–94, 2006.
 - [31] M Maggioni and HN Mhaskar. Diffusion polynomial frames on metric measure spaces. *Applied and Computational Harmonic Analysis*, 24(3) :329–353, 2008.
 - [32] David K Hammond, Pierre Vandergheynst, and Rémi Gribonval. Wavelets on graphs via spectral graph theory. *Applied and Computational Harmonic Analysis*, 30(2) :129–150, 2011.
 - [33] Raif Rustamov and Leonidas J Guibas. Wavelets on graphs via deep learning. In *Advances in neural information processing systems*, pages 998–1006, 2013.
 - [34] Jeff Irion and Naoki Saito. Hierarchical graph laplacian eigen transforms. *JSIAM Letters*, 6 :21–24, 2014.
 - [35] Nicolas Tremblay and Pierre Borgnat. Subgraph-based filterbanks for graph signals. *IEEE Transactions on Signal Processing*, 64(15) :3827–3840, 2016.
 - [36] Bin Dong. Sparse representation on graphs by tight wavelet frames and applications. *Applied and Computational Harmonic Analysis*, 42(3) :452–479, 2017.
 - [37] Luca Avena, Fabienne Castell, Alexandre Gaudillière, and Clothilde Mélot. Intertwining wavelets or multiresolution analysis on graphs through random forests. *arXiv preprint arXiv :1707.04616*, 2017.
 - [38] James C Bremer, Ronald R Coifman, Mauro Maggioni, and Arthur D Szlam. Diffusion wavelet packets. *Applied and Computational Harmonic Analysis*, 21(1) :95–112, 2006.
 - [39] Reinhard Diestel. *Graph theory (Graduate texts in mathematics)*, volume 173. Springer Heidelberg, 2005.

- [40] John Adrian Bondy, Uppaluri Siva Ramachandra Murty, et al. *Graph theory with applications*, volume 290. Citeseer, 1976.
- [41] Carlos S Kubrusly. Hilbert space operators. In *Hilbert Space Operators*, pages 13–22. Springer, 2003.
- [42] Tosio Kato. *Perturbation theory for linear operators*, volume 132. Springer Science & Business Media, 2013.
- [43] Haim Brezis. *Functional analysis, Sobolev spaces and partial differential equations*. Springer Science & Business Media, 2010.
- [44] Aliaksei Sandryhaila and José MF Moura. Discrete signal processing on graphs. *IEEE transactions on signal processing*, 61(7) :1644–1656, 2013.
- [45] Delio Mugnolo. *Semigroup methods for evolution equations on networks*. Springer.
- [46] Martin T Barlow. *Random walks and heat kernels on graphs*, volume 438. Cambridge University Press, 2017.
- [47] László Lovász. Random walks on graphs. *Combinatorics, Paul erdos is eighty*, 2(1-46) :4, 1993.
- [48] David Aldous and Jim Fill. Reversible markov chains and random walks on graphs, 2002.
- [49] Pierre Brémaud. *Markov chains : Gibbs fields, Monte Carlo simulation, and queues*, volume 31. Springer Science & Business Media, 2013.
- [50] Bojan Mohar, Y Alavi, G Chartrand, and OR Oellermann. The laplacian spectrum of graphs. *Graph theory, combinatorics, and applications*, 2(871-898) :12, 1991.
- [51] David A Levin and Yuval Peres. *Markov chains and mixing times*, volume 107. American Mathematical Soc., 2017.
- [52] Maarten Jansen, Guy P Nason, and Bernard W Silverman. Multiscale methods for data on graphs and irregular multidimensional situations. *Journal of the Royal Statistical Society : Series B (Statistical Methodology)*, 71(1) :97–125, 2009.
- [53] Matan Gavish, Boaz Nadler, and Ronald R Coifman. Multiscale wavelets on trees, graphs and high dimensional data : Theory and applications to semi supervised learning. In *ICML*, pages 367–374, 2010.
- [54] László Lovász. Submodular functions and convexity. In *Mathematical Programming The State of the Art*, pages 235–257. Springer, 1983.
- [55] Francis Bach et al. Learning with submodular functions : A convex optimization perspective. *Foundations and Trends® in Machine Learning*, 6(2-3) :145–373, 2013.
- [56] Rongjie Lai and Stanley Osher. A splitting method for orthogonality constrained problems. *Journal of Scientific Computing*, 58(2) :431–449, 2014.

- [57] Weiqiang Chen, Hui Ji, and Yanfei You. An augmented lagrangian method for 1-regularized optimization problems with orthogonality constraints. *SIAM Journal on Scientific Computing*, 38(4) :B570–B592, 2016.
- [58] Zaiwen Wen and Wotao Yin. A feasible method for optimization with orthogonality constraints. *Mathematical Programming*, 142(1-2) :397–434, 2013.
- [59] Frank Bauer. Normalized graph laplacians for directed graphs. *Linear Algebra and its Applications*, 436(11) :4193–4222, 2012.
- [60] Amy N Langville and Carl D Meyer. *Google’s PageRank and beyond : The science of search engine rankings*. Princeton University Press, 2011.
- [61] Jean-Paul Chehab and Marcos Raydan. Geometrical properties of the frobenius condition number for positive definite matrices. *Linear Algebra and its Applications*, 429(8-9) :2089–2097, 2008.
- [62] James Allen Fill. Eigenvalue bounds on convergence to stationarity for nonreversible markov chains, with an application to the exclusion process. *The annals of applied probability*, pages 62–87, 1991.
- [63] Aamir Anis, Akshay Gadde, and Antonio Ortega. Efficient sampling set selection for bandlimited graph signals using graph spectral proxies. *IEEE Transactions on Signal Processing*, 64(14) :3775–3789, 2016.
- [64] Benjamin Girault, Antonio Ortega, and Shrikanth Narayanan. Irregularity-aware graph fourier transforms. *arXiv preprint arXiv :1802.10220*, 2018.
- [65] Fan Chung. Laplacians and the cheeger inequality for directed graphs. *Annals of Combinatorics*, 9(1) :1–19, 2005.
- [66] Steve Butler. Interlacing for weighted graphs using the normalized laplacian. *Electronic Journal of Linear Algebra*, 16(1) :8, 2007.
- [67] Audrey Terras. *Fourier analysis on finite groups and applications*, volume 43. Cambridge University Press, 1999.
- [68] Ali Kaveh. *Optimal analysis of structures by concepts of symmetry and regularity*. Springer, 2013.
- [69] Dengyong Zhou, Jiayuan Huang, and Bernhard Schölkopf. Learning from labeled and unlabeled data on a directed graph. In *Proceedings of the 22nd international conference on Machine learning*, pages 1036–1043. ACM, 2005.
- [70] Lada A Adamic and Natalie Glance. The political blogosphere and the 2004 us election : divided they blog. In *Proceedings of the 3rd international workshop on Link discovery*, pages 36–43. ACM, 2005.

- [71] Mauro Maggioni, James C Bremer, Ronald R Coifman, and Arthur D Szlam. Biorthogonal diffusion wavelets for multiscale representations on manifolds and graphs. In *Wavelets XI*, volume 5914, page 59141M. International Society for Optics and Photonics, 2005.
- [72] Ali Civril. *Column Subset Selection for approximating data matrices*. PhD thesis, Rensselaer Polytechnic Institute, 2009.
- [73] Duncan J Watts and Steven H Strogatz. Collective dynamics of ‘small-world’ networks. *nature*, 393(6684) :440, 1998.
- [74] MEJ Newman. Mejl newman, *siam rev.* 45, 167 (2003). *Siam Rev.*, 45 :167, 2003.
- [75] David I Shuman, Mohammadjavad Faraji, and Pierre Vandergheynst. Semi-supervised learning with spectral graph wavelets. In *Proceedings of the International Conference on Sampling Theory and Applications (SampTA)*, number EPFL-CONF-164765, 2011.
- [76] Venkatesan N Ekambaram, Giulia Fanti, Babak Ayazifar, and Kannan Ramchandran. Wavelet-regularized graph semi-supervised learning. In *Global Conference on Signal and Information Processing (GlobalSIP), 2013 IEEE*, pages 423–426. IEEE, 2013.
- [77] Robert Tibshirani. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 267–288, 1996.
- [78] Ivan W Selesnick and Mario AT Figueiredo. Signal restoration with overcomplete wavelet transforms : Comparison of analysis and synthesis priors. In *Wavelets XIII*, volume 7446, page 74460D. International Society for Optics and Photonics, 2009.
- [79] Patrick L Combettes and Jean-Christophe Pesquet. Proximal splitting methods in signal processing. In *Fixed-point algorithms for inverse problems in science and engineering*, pages 185–212. Springer, 2011.
- [80] Amir Beck and Marc Teboulle. A fast iterative shrinkage-thresholding algorithm for linear inverse problems. *SIAM journal on imaging sciences*, 2(1) :183–202, 2009.
- [81] Stephen Becker, Jérôme Bobin, and Emmanuel J Candès. NESTA : A fast and accurate first-order method for sparse recovery. *SIAM Journal on Imaging Sciences*, 4(1) :1–39, 2011.
- [82] László Lovász. *Large networks and graph limits*, volume 60. American Mathematical Soc., 2012.
- [83] Christian Borgs, Jennifer T Chayes, László Lovász, Vera T Sós, and Katalin Vesztegombi. Convergent sequences of dense graphs i : Subgraph frequencies, metric properties and testing. *Advances in Mathematics*, 219(6) :1801–1851, 2008.

- [84] Yosra Hafiene, Jalal Fadili, and Abderrahim Elmoataz. The nonlocal p-laplacian evolution problem on graphs : The continuum limit. In *International Conference on Image and Signal Processing*, pages 370–377. Springer, 2018.
- [85] Dominique Perrault-Joncas and Marina Meila. Estimating vector fields on manifolds and the embedding of directed graphs. *arXiv preprint arXiv :1406.0013*, 2014.
- [86] Saman Mousazadeh and Israel Cohen. Embedding and function extension on directed graph. *Signal Processing*, 111 :137–149, 2015.
- [87] Tatsunori Hashimoto, Yi Sun, and Tommi Jaakkola. From random walks to distances on unweighted graphs. In *Advances in neural information processing systems*, pages 3429–3437, 2015.
- [88] Tatsunori Hashimoto, Yi Sun, and Tommi Jaakkola. Metric recovery from directed unweighted graphs. In *Artificial Intelligence and Statistics*, pages 342–350, 2015.
- [89] Sunil K Narang and Antonio Ortega. Perfect reconstruction two-channel wavelet filter banks for graph structured data. *IEEE Transactions on Signal Processing*, 60(6) :2786–2799, 2012.
- [90] David I Shuman, Mohammad Javad Faraji, and Pierre Vandergheynst. A multiscale pyramid transform for graph signals. *IEEE Transactions on Signal Processing*, 64(8) :2119–2134, 2016.
- [91] Stephane Lafon and Ann B Lee. Diffusion maps and coarse-graining : A unified framework for dimensionality reduction, graph partitioning, and data set parameterization. *IEEE transactions on pattern analysis and machine intelligence*, 28(9) :1393–1403, 2006.
- [92] Franziska Göbel, Gilles Blanchard, and Ulrike von Luxburg. Construction of tight frames on graphs and application to denoising. *arXiv preprint arXiv :1408.4012*, 2014.
- [93] Michael M Bronstein, Joan Bruna, Yann LeCun, Arthur Szlam, and Pierre Vandergheynst. Geometric deep learning : going beyond euclidean data. *IEEE Signal Processing Magazine*, 34(4) :18–42, 2017.
- [94] Joan Bruna, Wojciech Zaremba, Arthur Szlam, and Yann LeCun. Spectral networks and locally connected networks on graphs. *arXiv preprint arXiv :1312.6203*, 2013.
- [95] Michaël Defferrard, Xavier Bresson, and Pierre Vandergheynst. Convolutional neural networks on graphs with fast localized spectral filtering. In *Advances in Neural Information Processing Systems*, pages 3844–3852, 2016.
- [96] James Atwood and Don Towsley. Diffusion-convolutional neural networks. In *Advances in Neural Information Processing Systems*, pages 1993–2001, 2016.

- [97] Federico Monti, Karl Otness, and Michael M Bronstein. Motifnet : a motif-based graph convolutional network for directed graphs. *arXiv preprint arXiv :1802.01572*, 2018.
- [98] Peter W Battaglia, Jessica B Hamrick, Victor Bapst, Alvaro Sanchez-Gonzalez, Vinicius Zambaldi, Mateusz Malinowski, Andrea Tacchetti, David Raposo, Adam Santoro, Ryan Faulkner, et al. Relational inductive biases, deep learning, and graph networks. *arXiv preprint arXiv :1806.01261*, 2018.
- [99] David Silver, Aja Huang, Chris J Maddison, Arthur Guez, Laurent Sifre, George Van Den Driessche, Julian Schrittwieser, Ioannis Antonoglou, Veda Panneershelvam, Marc Lanctot, et al. Mastering the game of go with deep neural networks and tree search. *nature*, 529(7587) :484, 2016.
- [100] Sridhar Mahadevan and Mauro Maggioni. Proto-value functions : A laplacian framework for learning representation and control in markov decision processes. *Journal of Machine Learning Research*, 8(Oct) :2169–2231, 2007.
- [101] Mauro Maggioni and Sridhar Mahadevan. A multiscale framework for markov decision processes using diffusion wavelets. *University of Massachusetts, Massachusetts*, 2006.
- [102] Marlos C Machado, Marc G Bellemare, and Michael Bowling. A laplacian framework for option discovery in reinforcement learning. *arXiv preprint arXiv :1703.00956*, 2017.
- [103] Sébastien Bubeck, Nicolo Cesa-Bianchi, et al. Regret analysis of stochastic and nonstochastic multi-armed bandit problems. *Foundations and Trends® in Machine Learning*, 5(1) :1–122, 2012.
- [104] Michal Valko, Rémi Munos, Branislav Kveton, and Tomáš Kocák. Spectral bandits for smooth graph functions. In *International Conference on Machine Learning*, pages 46–54, 2014.
- [105] Kaige Yang and Laura Toni. Graph-based recommendation system. *arXiv preprint arXiv :1808.00004*, 2018.