



HAL
open science

Pilotage des urgences par analyse et exploitation géométrique et topologique des flux de patients

Guillaume Bouleux

► **To cite this version:**

Guillaume Bouleux. Pilotage des urgences par analyse et exploitation géométrique et topologique des flux de patients. Géométrie différentielle [math.DG]. INSA Lyon; Université Claude Bernard Lyon 1, 2022. tel-03621471

HAL Id: tel-03621471

<https://hal.science/tel-03621471>

Submitted on 28 Mar 2022

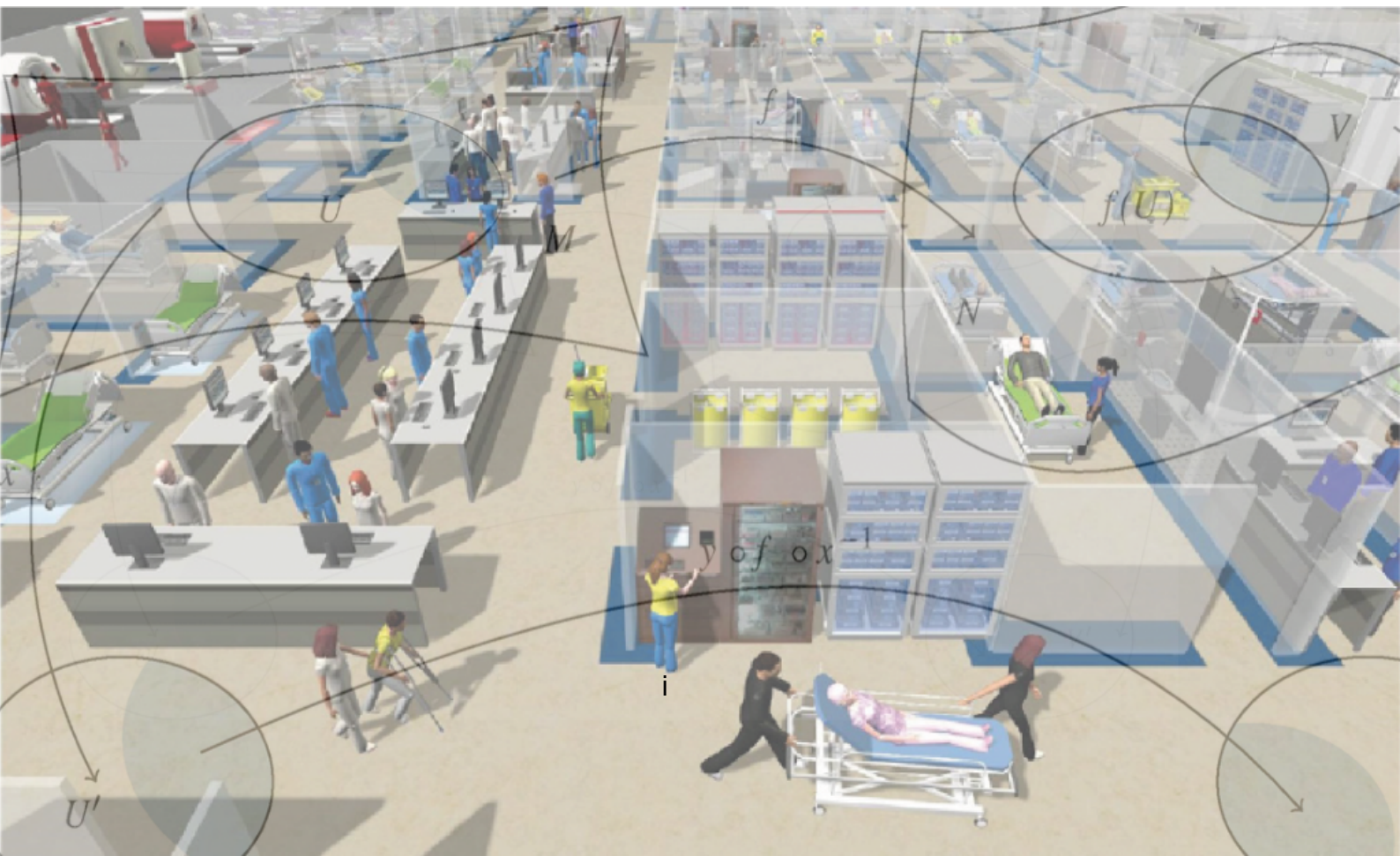
HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

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.

Pilotage des urgences par analyse et exploitation géométrique et topologique des flux de patients

Mémoire pour l'obtention de l'HDR

G. Bouleux



HABILITATION A DIRIGER DES RECHERCHES

Présentée devant

L'Institut National des Sciences Appliquées de Lyon
et l'Université Claude Bernard LYON I

**Pilotage des urgences par analyse et exploitation géométrique et
topologique des flux de patients**

Par

Guillaume BOULEUX

Soutenue le 21 janvier 2022 devant le jury

Président	Silvère Bonnabel, PU, Mines-ParisTech
Rapporteurs	Maria Di Mascolo, DR CNRS, G-SCOP Nicolas Le Bihan, DR CNRS, GIPSA-lab Louise Travé-Massuyès, DR CNRS, LAAS
Examineurs	Valérie Dos Santos Martins, MCF-HDR, LAGEPP Antoine Duclos, PU, RESHAPE Damien Trentesaux, PU, LAMIH Vincent Cheutet, PU, DISP

Sommaire

Résumé des points forts développés	1
I Curriculum vitæ-Enseignement	3
1 Curriculum Vitæ	5
1.1 Parcours Académique	6
1.2 Postes Occupés	7
1.3 Production Scientifique	7
1.3.1 Revues internationales avec comité de lecture	7
1.3.2 Articles de conférence internationale avec comité de lecture	8
1.3.3 Articles de conférence nationale avec comité de lecture	10
1.3.4 Articles de revue soumis	10
1.4 Encadrements	11
1.4.1 Doctorants	11
1.4.2 Stages M2	11
1.5 Contrats de Recherche	11
1.6 Rayonnement et Vulgarisation	12
1.6.1 Rayonnement National et International	12
1.6.2 Vulgarisation et valorisation	12
1.7 Responsabilités	13
1.8 Travail d'expertise	13
2 Activités d'Enseignement	15
2.1 Modules de Formation Créés	16
2.2 Cours Dispensé en DUT 1ère Année	17
2.3 Cours Dispensé en DUT 2e Année	18
2.4 Cours Dispensés aux Licences Pro. R&T	18
2.5 Encadrement de projets tuteurés et de stages	18
II Synthèse des travaux de recherche	19
Avant Propos	21
Mon positionnement	23
3 Prévenir la tension aux urgences pédiatriques	25
3.1 Un outil statistique simple	26
3.2 Dilatation d'opérateurs	28
3.2.1 Principes théoriques-le côté accessible	28
3.2.2 Principes théoriques-le côté mathématiques	29
3.2.3 Construction des matrices de Dilatation	31
3.2.4 Une implémentation pas à pas par l'exemple	31
3.3 Comparaison et caractérisation intrinsèque des processus	34
3.3.1 Indicateur intrinsèque par l'espace des formes	34
3.3.2 Indicateur intrinsèque par la topologie algébrique	37
3.4 Mise en place des indicateurs	45
3.4.1 L'indicateur précoce basé sur la topologie	45
3.4.2 L'application web	49
3.4.3 Retour d'expérience	51
3.5 Contributions scientifiques de cette partie	51

4	Anticiper l'engorgement aux urgences adultes en déterminant le flux de patients le plus impactant	53
4.1	Contexte et objectif	54
4.2	Le clustering de séries temporelles	55
4.2.1	Trajectoires	56
4.2.2	Les variétés testées	56
4.2.3	Les distances et calcul de barycentres sur les variétés	58
4.2.4	Le benchmark	60
4.3	Application aux flux de patients adultes	64
4.4	Contributions scientifiques de cette partie	64
5	Sur l'organisation des urgences adultes	67
5.1	Contexte et hypothèses	68
5.2	Le caractère auto-similaire du flux de patient	68
5.2.1	L'analyse du flux	69
5.2.2	Modélisation par des cascades multiplicatives	72
5.3	Le groupe de renormalisation pour l'organisation	74
5.3.1	Principes généraux	74
5.3.2	L'opérateur de Koopman	76
5.4	Contributions scientifiques de cette partie	80
III	Bilan et projet	81
6	Vers l'infini et l'au-delà !	83
6.1	Dans la continuité	84
6.2	Le Projet de Gestion de crises humanitaires	84
6.2.1	Modélisation d'une crise	86
6.2.2	Réseau temporel - Signature et Résilience	87
6.3	Le Projet Jumeau Numérique des urgences	88
6.4	Résolution de problèmes de recherche opérationnelle par l'approche des flots sur les variétés	92
	Références	110

Résumés

Version Française

Dans ce manuscrit, j'aborde la résolution des problématiques d'encombrement ou de tension dans les services d'accueil des urgences par la constitution d'outils d'aide à la décision pour le pilotage des services. Je m'intéresse donc à proposer des indicateurs précoces et préventifs de venues anormales au sein des services et à caractériser les problèmes organisationnels au travers des flux de patients.

J'ai tout d'abords traité le cas des urgences pédiatriques en proposant un indicateur statistique simple. Puis, j'ai proposé de modéliser plus finement le processus du flux de patients par un processus périodiquement corrélé. Cette modélisation m'a permis de proposer de nouvelles représentations de ces processus sur des groupes de Lie au travers de la théorie de la dilatation d'opérateurs et de pouvoir ainsi utiliser la géométrie différentielle et la topologie algébrique pour de la classification. Je montre également au travers de ces représentations l'existence d'une information géométrique et topologique intrinsèque à la classe des processus. La topologie algébrique est par la suite utilisée pour la création de nouveaux indicateurs précoces de venues anormales aux urgences pédiatriques.

Une fois la question de tension aux urgences pédiatriques traitée, je me suis naturellement tourné vers la problématique des urgences adultes pour lesquelles, le flux de patients impactant les urgences n'est pas forcément bien déterminé. J'ai choisi d'aborder cette problématique par l'angle du clustering de séries temporelles, où les séries temporelles sont les séries chronologiques des codes diagnostics des patients. Il est montré dans cette partie que le clustering des séries temporelles est grandement amélioré lorsque les séries sont représentées dans des espaces de plus grande dimension. En particulier, la variété Stiefel est la variété sur laquelle doivent être projetées les séries temporelles. Le framework de clustering donnant ensuite les meilleures performances est celui composé de UMAP pour la réduction de dimension puis de HDBSCAN pour le clustering. Ces résultats sont validés sur plusieurs benchmarks de séries temporelles.

Il restait à traiter le caractère organisationnel du service d'accueil des urgences en tant que système complet. C'est ce qui est développé dans une troisième partie. Bien que les travaux ne soient pas encore publiés dans leur totalité, j'expose ma vision de l'organisation au sein d'un service d'accueil des urgences. L'approche que j'ai choisie de prendre est tout d'abords celle de la modélisation du flux de patients sous la caractéristique d'auto-similarité et des cascades multiplicatives puis de la quantification organisationnelle par l'étude des interactions massives grâce à l'approche du groupe de renormalisation.

Pour finir, je décris les projets en cours, qu'ils soient sur la thématique des crises humanitaires sous des approches de réseaux complexe multi-couches, de jumeau numérique des urgences et la compréhension organisationnelle du système au travers des systèmes multi-agents ou encore de recherche opérationnelle pour la résolution de problèmes de programmation linéaire par les approches de flots de gradient sur des variétés et de double crochet de Lie.

Version Anglaise

In this manuscript, I approach the resolution of congestion or tension problems in emergency departments by the constitution of decision support tools for the management of the departments. I am therefore interested in proposing early and preventive indicators of abnormal arrivals within the services and in characterizing the organizational problems through the patient flows.

I first dealt with the case of pediatric emergencies by proposing a basic statistical indicator. Then, I proposed to model more finely the patient flow process by a periodically correlated one. This modeling allowed me to propose new representations of these processes on Lie groups through the theory of dilation of operators and to use differential geometry and algebraic topology for classification. I also show through these representations the existence of geometric and topological information intrinsic to the class of processes. The algebraic topology is then used for the creation of new early indicators of abnormal visits to pediatric emergencies.

Once the issue of tension in pediatric emergencies was addressed, I naturally turned to the problem of adult emergencies for which the flow of patients impacting the emergencies is not necessarily well determined. I chose to approach this problem from the angle of time series clustering, where the time series are the time series of the patients' diagnostic codes. It is shown in this section that time series clustering is greatly improved when the series are represented in higher dimensional spaces. In particular, the Stiefel manifold is the manifold on which the time series should be projected. The clustering framework giving then the best performances is the one composed of UMAP for dimension reduction and HDBSCAN for clustering. These results are validated on several time series benchmarks.

It remained to treat the organizational character of the emergency service as a complete system. This is developed in a third part. Although the work is not yet published in its entirety, I present my vision of the organization within an emergency department. The approach I have chosen to take is firstly that of modeling the flow of patients under the characteristic of self-similarity and multiplicative cascades, and then of organizational quantification by studying massive interactions thanks to the renormalization group approach.

Finally, I describe the current projects, whether they are on the theme of humanitarian crises under the approaches of complex multi-layer networks, digital twin of emergencies and organizational understanding of the system through multi-agent systems or operational research for the resolution of linear programming problems by the approaches of gradient flows on manifolds and double Lie bracket.



Curriculum vitæ-Enseignement

Curriculum Vitæ



- Maître de conférences (classe normale)
- Université Jean Monnet (IUT de Roanne)
- Section 61 : Génie informatique, automatique et traitement du signal
- Laboratoire Décision et Information pour les Systèmes de Production (DISP, EA-4570)
- Adresse Professionnelle : Bât. Léonard de Vinci, 20 Av. Jean Capelle, 69621, Villeurbanne
- Téléphone : +33 6 08 98 99 98
- Adresse électronique : guillaume.bouleux@{insa-lyon.fr, univ-st-etienne.fr}
- Page web : <https://sites.google.com/site/guillaumebouleux/Home>

Sommaire

1.1	Parcours Académique	6
1.2	Postes Occupés	7
1.3	Production Scientifique	7
1.3.1	Revue internationale avec comité de lecture.	7
1.3.2	Articles de conférence internationale avec comité de lecture	8
1.3.3	Articles de conférence nationale avec comité de lecture	10
1.3.4	Articles de revue soumis.	10
1.4	Encadrements	11
1.4.1	Doctorants	11
1.4.2	Stages M2	11
1.5	Contrats de Recherche	11
1.6	Rayonnement et Vulgarisation.	12
1.6.1	Rayonnement National et International	12
1.6.2	Vulgarisation et valorisation	12
1.7	Responsabilités	13
1.8	Travail d'expertise	13

1.1. Parcours Académique

2005-2007	Thèse de doctorat en Traitement du Signal soutenue le 04/12/2007 <i>Traitement d'antenne adapté aux modèles linéaires intégrant une interférence structurée. Application aux signaux mécaniques</i>		
	Etablissement	Laboratoire(s) d'accueil	Directeur de Thèse
	Université Jean Monnet	LSS-SUPELEC/LASPI	François Guillet
	Composition du jury	Président : M. Jean Pierre Delmas, Pr. 61 Rapporteurs : M. Pascal Larzabal, Pr. 61 M. Jean-Yves Tourneret, Pr. 61 Examineurs : Mme. Sylvie Marcos, DR-CNRS M. Roland Badeau, MCF 61 M. Mohamed El Badaoui, MCF 61 Directeurs : M. François Guillet, Pr. 61 M. Rémy Boyer, MCF 61	
2004-2005	DEA Traitement des Images et du Signal (TIS), Laboratoire ETIS ENSEA, université de Cergy-Pontoise		
2000-2004	Ingénieur ENSEA - spécialité Signal temps Réel et Communications		
1999-2000	Classe préparatoire ATS - ENREA, Clichy		
1997-1999	DUT GEII, IUT de Cachan		

1.2. Postes Occupés

2014-	MCF université Jean Monnet, IUT de Roanne - Laboratoire DISP
2008-	MCF université Jean Monnet, IUT de Roanne - Laboratoire LASPI
2008	Post-Doc de 6 mois Université Uppsala, Suede - Direction : Petre Stoica
2006-2008	ATER mi-temps, IUT de Roanne - Laboratoire LASPI

1.3. Production Scientifique

Revues :

- 12 articles de revue internationale avec comité de lecture publiés (depuis décembre 2007).
- 3 articles de revue internationale avec comité de lecture soumis.

Conférences :

- 22 articles de conférence internationale avec comité de lecture et actes publiés.
- 6 articles de conférence nationale (GRETSI) avec comité de lecture et actes publiés.

1.3.1. Revues internationales avec comité de lecture

- [CBC22] Aurélie Charles, Guillaume Bouleux, and Chantal Chérifi. “Understanding the situation at a glance : the powerful value added of complex networks to analyse humanitarian operations”. In: *International journal of disaster risk reduction* (2022).
- [Mor+22] Benoît Morel et al. “Evaluating the increased burden in Adult Emergency Departments of cardio-respiratory illness visits during flu and bronchiolitis outbreaks in the pediatric population by a retrospective multicentric time series analysis”. In: *JMIR Public Health and Surveillance* (2022). DOI: 10.2196/25532.
- [PBC21] Clément Péalat, Guillaume Bouleux, and Vincent Cheutet. “Improved Time Series Clustering Based on New Geometric Frameworks”. In: *Pattern Recognition, Elsevier* (2021). DOI: 10.1016/j.patcog.2021.108423.
- [BDM19a] Guillaume Bouleux, Maël Dugast, and Eric Marcon. “Information Topological Characterization of Periodically Correlated Processes by Dilation Operators”. In: *IEEE Transactions on Information Theory* 65.10 (2019), pp. 6484–6495. ISSN: 15579654. DOI: 10.1109/TIT.2019.2923217.
- [Dug+19] Maël Dugast et al. “Improving Health Care Management Through Persistent Homology of Time-Varying Variability of Emergency Department Patient Flow”. In: *IEEE Journal of Biomedical and Health Informatics* 23.5 (2019), pp. 2174–2181. ISSN: 21682208. DOI: 10.1109/JBHI.2018.2882748.

- [DBM18b] Maël Dugast, Guillaume Bouleux, and Eric Marcon. “Representation and characterization of nonstationary processes by dilation operators and induced shape space manifolds”. In: *Entropy* 20.9 (2018), p. 717. ISSN: 10994300. DOI: 10.3390/e20090717. URL: <http://www.mdpi.com/1099-4300/20/9/717>.
- [BB17] Guillaume Bouleux and Rémy Boyer. “Sparse-based estimation performance for partially known overcomplete large-systems”. In: *Signal Processing* 139 (2017), pp. 70–74. ISSN: 01651684. DOI: 10.1016/j.sigpro.2017.04.010.
- [BMM15] Guillaume Bouleux, Eric Marcon, and Olivier Mory. “Early Index for Detection of Pediatric Emergency Department Crowding”. In: *Biomedical and Health Informatics, IEEE Journal of* 19.6 (2015), pp. 1929–1936. ISSN: 2168-2194. DOI: 10.1109/JBHI.2014.2350996.
- [Bou13a] Guillaume Bouleux. “Oblique projection pre-processing and TLS application for diagnosing rotor bar defects by improving power spectrum estimation”. In: *Mechanical Systems and Signal Processing* 41.1-2 (2013), pp. 301–312. ISSN: 08883270. DOI: 10.1016/j.ymsp.2013.06.018. URL: <http://www.sciencedirect.com/science/article/pii/S0888327013003014>.
- [Bou13b] Guillaume Bouleux. “Prior Knowledge Optimum Understanding by Means of Oblique Projectors and Their First Order Derivatives”. In: *IEEE Signal Processing Letters* 20.3 (2013), pp. 205–208. ISSN: 10709908. DOI: 10.1109/LSP.2013.2238928.
- [Wir+12] Petter Wirfält et al. “Optimal prior knowledge-based direction of arrival estimation”. In: *IET Signal Processing* 6.8 (2012), pp. 731–742. ISSN: 17519675. DOI: 10.1049/iet-spr.2011.0453.
- [BB08b] Rémy Boyer and Guillaume Bouleux. “Oblique Projections for Direction-Of-Arrival Estimation with Prior Knowledge”. In: *IEEE Trans. on Signal Processing* 56.4 (2008), pp. 1374–1387. ISSN: 1053587X. DOI: 10.1109/TSP.2007.909348.

1.3.2. Articles de conférence internationale avec comité de lecture

- [PBC20] Clément Pealat, Guillaume Bouleux, and Vincent Cheutet. “Improved time-series clustering with UMAP dimension reduction method”. In: *Proceedings - International Conference on Pattern Recognition*. Milan, 2020, pp. 5658–5665. ISBN: 9781728188089. DOI: 10.1109/ICPR48806.2021.9412261.
- [BB19] Guillaume Bouleux and Frédéric Barbaresco. “Dilation Operator Approach and Square Root Velocity Transform for Time/Doppler Spectra Characterization on SU(n)”. In: *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. Vol. 11712 LNCS. 2019, pp. 31–38. ISBN: 9783030269791. DOI: 10.1007/978-3-030-26980-7_4.
- [PBC19] Clément Pealat, Guillaume Bouleux, and Vincent Cheutet. “Extracting most impacting emergency department patient flow by embedding laboratory-confirmed and clinical diagnosis on the stiefel manifold”. In: *2019 IEEE EMBS International Conference on Biomedical and Health Informatics, BHI 2019 - Proceedings*. Chicago, 2019. ISBN: 9781728108483. DOI: 10.1109/BHI.2019.8834507.
- [DBM18a] Maël Dugast, Guillaume Bouleux, and Eric Marcon. “Information topological characterization for non-stationary random processes”. In: *Entropy 2018 – From Physics to Information Sciences and Geometry*. Barcelona, 2018.
- [Emi+18] Sid Ahmed Emine et al. “Emergency Department Overcrowding Detection by a Multifractal Analysis”. In: *IFAC Symposium on Biological and Medical Systems*. Vol. 51. 27. Sao paulo, 2018, pp. 104–109. DOI: 10.1016/j.ifacol.2018.11.659.

- [Sol+18] Gregory Soler et al. "Emergency Department Admissions Overflow Modeling by a Clustering of Time Evolving Clinical Diagnoses". In: *IEEE International Conference on Automation Science and Engineering*. Vol. 2018-Augus. 2018, pp. 365–370. ISBN: 9781538635933. DOI: 10.1109/COASE.2018.8560585.
- [Cho+15] Vincent Choqueuse et al. "Voltage sags estimation in three-phase systems using Unconditional Maximum Likelihood estimation". In: *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*. Vol. 2015-Augus. IEEE. 2015, pp. 2869–2873. ISBN: 9781467369978. DOI: 10.1109/ICASSP.2015.7178495.
- [WJB13] Petter Wirfält, Magnus Jansson, and Guillaume Bouleux. "Prior-exploiting direction-of-arrival algorithm for partially uncorrelated source signals". In: *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*. 2013, pp. 3972–3976. ISBN: 9781479903566. DOI: 10.1109/ICASSP.2013.6638404.
- [BK11] Guillaume Bouleux and Thameur Kidar. "Total Least Squares application for diagnosing rotor bar defects by improving power spectrum estimation". In: *8th International Conference on Condition Monitoring and Machinery Failure Prevention Technologies 2011, CM 2011/MFPT 2011*. Vol. 1. 2011, pp. 165–176. ISBN: 9781618390141.
- [BKG11] Guillaume Bouleux, Thameur Kidar, and François Guillet. "An improper random vector approach for ESPRIT and unitary ESPRIT frequency estimation". In: *7th International Workshop on Systems Signal Processing and their Applications (WOSSPA)*. 2011, pp. 299–302. ISBN: 9781457706905. DOI: 10.1109/WOSSPA.2011.5931478.
- [Wir+11a] Petter Wirfält et al. "Subspace-based frequency estimation utilizing prior information". In: *IEEE Workshop on Statistical Signal Processing Proceedings*. 2011, pp. 533–536. ISBN: 9781457705700. DOI: 10.1109/SSP.2011.5967751.
- [Wir+11b] Petter Wirfält et al. "Prior knowledge-based direction of arrival estimation". In: *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*. 2011, pp. 2540–2543. ISBN: 9781457705397. DOI: 10.1109/ICASSP.2011.5947002.
- [BSB09b] Guillaume Bouleux, Petre Stoica, and Remy Boyer. "An optimal prior knowledge-based DoA estimation method". In: *European Signal Processing Conference*. 2009, pp. 869–873.
- [El +09] Mohammed Nabil El Korso et al. "Sequential estimation of the range and the bearing using the Zero-Forcing MUSIC approach". In: *European Signal Processing Conference*. 2009, pp. 1404–1408.
- [BB08a] Guillaume Bouleux and Rémy Boyer. "First-order analysis of the prior-based knowledge minnorm algorithm". In: *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*. Las Vegas, 2008, pp. 2489–2492. ISBN: 1424414849. DOI: 10.1109/ICASSP.2008.4518153.
- [Bou+08] Guillaume Bouleux et al. "A subspace-based rejection method for detecting bearing fault in asynchronous motor". In: *Proceedings of 2008 International Conference on Condition Monitoring and Diagnosis, CMD 2008*. Beijing, 2008, pp. 171–174. ISBN: 9781424416219. DOI: 10.1109/CMD.2008.4580256.
- [BB07a] Guillaume Bouleux and Rémy Boyer. "A lower bound for sequential estimators". In: *2007 2nd IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing, CAMPSAP*. 2007, pp. 277–280. ISBN: 9781424417148. DOI: 10.1109/CAMSP.2007.4498019.
- [BB07c] Guillaume Bouleux and Rémy Boyer. "Zero-forcing based sequential music algorithm". In: *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*. Vol. 3. Hawaii, 2007. ISBN: 1424407281. DOI: 10.1109/ICASSP.2007.366855.
- [BB06a] Guillaume Bouleux and Rémy Boyer. "Analysis of prior-subspace estimation schemes". In: *2006 IEEE Sensor Array and Multichannel Signal Processing Workshop Proceedings, SAM 2006*. Boston, 2006, pp. 526–530. ISBN: 1424403081. DOI: 10.1109/sam.2006.1706189.

- [BB06b] Guillaume Bouleux and Rémy Boyer. “Subspace-based estimation of Direction-Of-Arrival with known directions”. In: *European Signal Processing Conference*. Florence, 2006.
- [BB06c] Rémy Boyer and Guillaume Bouleux. “Cramér-Rao lower bound for prior-subspace estimation”. In: *2006 IEEE Sensor Array and Multichannel Signal Processing Workshop Proceedings, SAM 2006*. 2006, pp. 394–398. ISBN: 1424403081. DOI: 10.1109/SAM.2006.1677228.
- [BBA06] Rémy Boyer, Guillaume Bouleux, and Karim Abed-Meraim. “Common pole estimation with an Orthogonal Vector Method”. In: *European Signal Processing Conference*. Florence, 2006.

1.3.3. Articles de conférence nationale avec comité de lecture

- [BDM19b] Guillaume Bouleux, Maël Dugast, and Eric Marcon. “Représentation et comparaison des processus périodiquement corrélés sur des groupes de Lie par les opérateurs de Dilatation”. In: *XXVII colloque GRETSI*. Lille, 2019, pp. 4–7.
- [BB15] Guillaume Bouleux and Rémy Boyer. “Acquisition Comprimée intégrant une Connaissance Partielle du Modèle d’Observation: Analyse Asymptotique”. In: *XXV colloque GRETSI*. Lyon, 2015.
- [WBJ13] Petter Wirfält, Guillaume Bouleux, and Magnus Jansson. “Estimation de DOA sous l’hypothèse de sources partiellement corrélées”. In: *XXIV colloque GRETSI*. Brest, 2013.
- [BSB09a] Guillaume Bouleux, Peter Stoica, and Rémy Boyer. “Une méthode optimale pour l’estimation des directions d’arrivées basée sur des connaissances a priori”. In: *XXII colloque GRETSI*. Dijon, 2009, p. 120.
- [Kor+09] Mohammed Nabil E L Korso et al. “Estimation séquentielle des paramètres de localisation en champ proche à l’aide de l’approche Zero-Forcing”. In: *XXII colloque GRETSI*. Dijon, 2009, pp. 2–5.
- [BB07b] Guillaume Bouleux and Rémy Boyer. “Estimation des Directions D’Arrivées incorporant un a priori : Algorithmes et Variances Théoriques algorithmes d’estimation et analyse asymptotique”. In: *XXI colloque GRETSI*. Troyes, 2007, pp. 11–14.

1.3.4. Articles de revue soumis

- [Liu+22a] Yinling Liu et al. “An agent-based architecture of the Digital Twin for an Emergency Department”. In: *Health and Technology* (2022).
- [Liu+22b] Yinling Liu et al. “Hybrid Simulation Modeling of Emergency Departments for Resource Scheduling”. In: *Heal. Syst.* (2022).
- [PBC22] Clément Péalat, Guillaume Bouleux, and Vincent Cheutet. “Using a manifold-based approach to extract clinical codes associated with winter respiratory viruses at an emergency department”. In: *Health Care Management Science* (2022).

1.4. Encadrements

1.4.1. Doctorants

Étudiants Ce = co-encadrement, Cd = co-direction	Directeur(s)	Financement	Date Début	Date Fin	Publications
Maël DUGAST (Ce, 70%)	E. Marcon (30%)	ARS Auvergne-Rhône-Alpes, APP innovants	01/11/2015	21/12/2018	[DBM18a]; [BDM19a], [Dug+19]; [DBM18b]
Sid'Ahmed Emine (Cd, 50%)	E. Marcon (25%), A.Haouba (25%)	Co-tutelle Mauritanie, Ministère des affaires étrangères	01/09/2016	31/12/2021	[Emi+18]
Clément Pealat (Cd, 50%)	V. Cheutet (50%)	APP PREPS- DGOS, Ministère de la santé	01/12/2018	01/12/2021	[PBC19]; [PBC20] [PBC21]; [PBC22]
Johan Leydet	G. Bouleux (40%)	Bourse ED InfoMaths	01/10/2021		

1.4.2. Stages M2

Étudiants MASTER Co-encadrés	Co-encadrant(s)	Financement	Date Début	Date Fin
Thameur Kidar (100%)		Interne Laboratoire	01/02/2011	31/07/2011
Fernand Girardin (30%)	A. Charles (30%), C. Cherifi (40%)	Interne Laboratoire	01/02/2020	31/07/2020
Amine Filali (40%)	A. Charles (30%), C. Cherifi (30%)	Interne Laboratoire	01/03/2021	31/08/2021

1.5. Contrats de Recherche

Je suis le co-investigateur avec Eric MARCON, d'un projet de valorisation portant sur la détection précoce de phénomènes épidémiques/pandémiques saisonniers à partir des données des passages aux urgences. Une expérimentation a été financée à hauteur de 35K€ par l'Agence Régional de la Santé (ARS) Rhône-Alpes. Ce projet a concerné environ 4 personnes.

Je suis co-investigateur avec Eric MARCON du projet « GO Plan hivernal en pédiatrie » financé à hauteur de 152 400€ sur appel d'offre projets innovants de l'ARS Rhône-Alpes. Ce projet concerne environ 5 personnes. Financement d'un doctorant.

Je suis le responsable scientifique et investigateur principal du projet « Indicateur moyen d'attentes aux urgences » financé à hauteur de 32K€ par le ministère de la Santé, des affaires sociales et du droit des femmes. Ce projet est en collaboration avec la société CERCLH et regroupe environ 5 personnes.

Je suis le responsable scientifique du projet PREPS (Direction Générale de l'Offre de SOIN, DGOS) PRED AFLU à hauteur de 384K€. Ce projet concerne les labo-

ratoires VirPath (HCL, Lyon 1), GIMAP (CHU Saint Etienne, UJM) et DISP (INSA Lyon) et regroupe environ 15 personnes. Financement d'un doctorant et d'un post-doc ; sous-traitance de la partie informatique à une entreprise spécialisée dans le développement WEB.

1.6. Rayonnement et Vulgarisation

1.6.1. Rayonnement National et International

Titulaire de la PEDR depuis 2019.

J'ai co-organiser et présidé une session spéciale à la conférence MOSIM 2014 s'intitulant :

« *Gestion des situations de crises dans les systèmes de production de soins* »

J'ai co-organisé et présidé une session spécial à la conférence MOSIM 2016 s'intitulant :

« *Pilotage des systèmes de production de soins en milieu incertain* »

J'ai co-organisé et présidé une session spéciale à la conférence IEEE BHI 2019 s'intitulant:

« *Decision-support computing by data-driven and AI-based approaches for healthcare* »

Exposé invité à la journée automatique et santé organisée par le club EEA (gregory.faraut@ens-paris-saclay.fr) le JEUDI 24 Octobre 2019

Exposé invité à la journée action Jumeau Numérique du GDR MACS organisée par Vincent Cheutet (vincent.cheutet@insa-lyon.fr) le Lundi 28 Juin 2021.

Je suis membre du comité technique de programme des conférences HEALTHINFO 2020 et 2021.

Je suis membre du comité d'organisation de la conférence ROADEF 2022 à LYON.

Je suis membre du comité scientifique de la conférence GISEH 2022.

1.6.2. Vulgarisation et valorisation

Une WEB application est mise à disposition des urgences pédiatriques du CHU de Saint Etienne, du CHU de Grenoble, du CH de Valence et du CH de Villefranche sur Saône afin d'aider à la mise en place du plan hivernal Bronchiolite <http://predaflu.chu-st-etienne.fr>. Cette WEB application sera la base pour l'application cible du projet PREPS PREDAFLU; pour la mise en place d'une organisation adaptée pour les urgences adultes lors des épidémies virales hivernales.

J'ai exposé le projet PREDAFLU sur invitation à la Paris Healthcare Week (porte de versailles) en 2017.

le projet PREDAFLU a fait sujet d'un reportage France 3 Saint Etienne : <https://www.youtube.com/watch?v=VYRmlpDA8pg>

Le projet PREDAFLU a fait sujet d'un article dans le CHU Mag de Saint Etienne : <https://fr.calameo.com/read/004523752f46304d6c5f0>

1.7. Responsabilités

Enseignement

Responsable des stages de la formation R&T de l'IUT de Roanne de 2010 à 2021. Afin de faciliter le suivi des stagiaires, la notation et la communication entre les étudiants et les référents IUT, j'ai créé un site web des stages : <http://217.128.29.189/>

Je suis responsable et coordinateur national du nouveau parcours ROM (Réseaux d'Opérateurs et Multimédia) pour la mise en place du Bachelor Universitaire de Technologie (BUT ex DUT).

Je suis membre du comité de sélection pour le département R&T de l'IUT de Roanne, j'ai participé à un comité de sélection pour un poste dans le département R&T.

Je suis membre du comité de sélection 61-63 de l'Université Jean Monnet, j'ai participé à 3 comités de recrutement pour des postes d'ATER.

Recherche

Responsable de l'organisation des séminaires du LASPI entre 2010 et 2013.

J'ai été co-Responsable de l'axe Pilotage des Systèmes de Production de biens et de services (PSP, axe regroupant environ 14 enseignants-chercheurs) du laboratoire DISP de 2016 à 2020.

Je suis co-responsable de l'axe Gestion et Optimisation des Opérations (G20, axe regroupant environ 15 enseignants-chercheurs) du laboratoire DISP depuis 2020.

Je suis Membre nommé du Conseil de Laboratoire DISP depuis le 01/01/2016.

Je suis Membre du Comité de Direction du Laboratoire DISP depuis le 01/01/2016.

1.8. Travail d'expertise

J'ai effectué l'expertise d'un dossier lors de l'appel à projets PHRIP - DGOS PF4 en 2021.

2

Activités d'Enseignement



Enseigner en Réseaux et Télécoms (R&T) dans un IUT demande une grande faculté d'adaptation. En presque 15 années le programme pédagogique nationale a déjà évolué par deux fois et nous devons une nouvelle fois faire face à grand bouleversement pédagogique pour la mise en œuvre du BUT (Bachelors universitaire de Technologie). Le département R&T de l'IUT de

Roanne s'étant ouvert en 2006, année de mon arrivée à l'IUT de Roanne, j'ai participé à la création complète de certains modules.

Sommaire

2.1	Modules de Formation Créés	16
2.2	Cours Dispensé en DUT 1ère Année.	17
2.3	Cours Dispensé en DUT 2e Année	18
2.4	Cours Dispensés aux Licences Pro. R&T	18
2.5	Encadrement de projets tuteurés et de stages	18

2.1. Modules de Formation Créés

J'ai ainsi créé :

- un module CM, TD, TP de 30h sur les transmissions analogiques.

Il s'agit d'introduire les concepts de la transmission de type analogique. Introduire les notions de modulations, quelles soient d'amplitude ou de fréquence et prolonger avec les modulations d'amplitudes en quadrature. En associant ce module au module sur l'électronique pour les télécommunications, le cours ainsi que les TPs se sont enrichis, permettant aux étudiants de trouver un lien entre ces matières.

- un module CM, TD, TP de 30h pour l'introduction à l'algèbre linéaire et aux graphes.

L'idée dans ce module est d'introduire les concepts de linéarités, puis d'application linéaire en concomitance avec les premières définitions sur les matrices : dimension, opérateurs, rang, pour la résolution de systèmes linéaires N équations P inconnues. Au travers de ces résolutions la diagonalisation est pratiquée en TD (comme prolongement) et l'estimation des directions d'arrivées (DDA) sur un réseau équidistant de capteurs linéairement espacés par l'algorithme Root-MUSIC est réalisée par les étudiants en TP. Les graphes sont introduits et l'algorithme de recherche du plus court chemin Dijkstra sert de support à la compréhension et à l'application.

- un module CM, TD, TP de 15h pour l'introduction à l'algèbre linéaire et à la résolution de systèmes d'équations.

Ce module est un module aujourd'hui obsolète dû au changement du programme pédagogique national en 2013. Il s'agissait d'introduire les matrices et les espaces vectoriels. Les étudiants réalisaient une application de traitement du signal : le filtrage en exploitant la construction et l'application de la matrice de Sylvester.

- un module CM, TD, TP de 90h sur l'accès pour les particuliers et entreprises au cœur de réseaux et sur les technologies des cœurs de réseaux.

Ce module est en réalité une refonte du module réseaux opérateurs et réseaux étendus de l'ancien programme. J'avais la charge de 30h sur les technologies xDSL. Aujourd'hui je suis devenu le seul spécialiste et seul enseignant référent/compétent à l'IUT de Roanne dans les technologies liées au cœur de réseau. J'ai suivi une formation CCNA dispensée par les collègues de R&T de Nancy puis une formation CISCO sur le MPLS. J'ai petit à petit acquis suffisamment de connaissance pour pouvoir réaliser une maquette dite « opérateur » dans le département. Cette maquette permet aux étudiants en TD ou TP de configurer des DSLAM (ADSL) et OLT (FFTH), de trouver quelles peuvent être les causes lorsqu'un abonné n'a plus de service FTTH ou ADSL. Les étudiants s'intéressent aux boucles SDH en configurant les ADM, ils mettent en œuvre le cœur de réseau MPLS et les VPN L3 via VRF et BGP, Il y a beaucoup d'heures de travail et d'apprentissage pour ce module dont encore une fois, je suis le seul à mettre en œuvre.

- Un module CM, TD, TP de 60h sur l'électronique pour les télécommunications.

Ce module n'existe plus depuis le nouveau programme, c'est un module axé sur la Phase Locked Loop (PLL). Il fallait mettre en place la transformée de Laplace, les systèmes bouclés, la stabilité via les différents critères : Nyquist, Bode, ..., les oscillateurs sinusoïdaux puis enfin donner les principes fondamentaux de la PLL. Les TP étaient vraiment tournés vers les applications réelles des télécommunications (émission/réception par exemple).

- Un module CM, TD, TP de 10h d'anglais technique pour les licences pro R&T.

Le responsable de la licence pro R&T m'a demandé de prendre en charge un module d'anglais technique. Ce module permet aux étudiants de suivre un cours technique sur le réseau en anglais, de faire des TP avec comme unique consigne un fichier audio en anglais. Chaque étudiant doit créer un TP qu'il présente comme tuto en anglais à l'ensemble des étudiants. Les étudiants doivent être en capacité de reproduire le TP, ils posent des questions lorsqu'ils ont des difficultés.

2.2. Cours Dispensé en DUT 1ère Année

Intitulé du cours	CM (H)	TD (H)	TP (H)
Transmissions Analogiques	5	10	24
Algèbre Linéaire	5	8	3
Mathématiques : Bases			10

2.3. Cours Dispensé en DUT 2e Année

Intitulé du cours	CM (H)	TD (H)	TP (H)
Électronique pour les télécommunications	12	24	24
Réseaux Opérateurs	6	8	28
Algèbre linéaire et Graphes	6	20	6
Accès au cœur de réseau	9	6	15
Technologies du cœur de réseau	15	12	33

2.4. Cours Dispensés aux Licences Pro. R&T

Intitulé du cours	CM (H)
Réseaux avancés et réseaux opérateurs	48
Anglais technique	10

2.5. Encadrement de projets tuteurés et de stages

Encadrements	Durées pour les étudiants (H)	Heures affectées par la composante (H)	Nb moyen d'étudiants suivis	Total heures (H)
Projets de 2e année	60	6	4	24
Stages de 2e année	500	2	7	14
Licences pro	200	10	3	30



Synthèse des travaux de recherche

Avant Propos



Après avoir soutenu ma thèse en décembre 2007 suite à deux années de travaux, des collaborations avec P. Stoica (université d'Uppsala, Suède) ainsi qu'avec M. Jansson (université KTH, Suède) se sont créées. La thématique principale de ces collaborations a concerné le traitement d'antenne. Nous avons revisité des méthodes d'estimation des directions d'arrivées (DDA) pour intégrer des hypothèses fortes sur le modèle statistique des observations et ainsi proposer des estimateurs beaucoup plus performants. J'ai également pu collaborer avec le laboratoire LSS (université Paris XI, supelec) via Rémy Boyer sur ces mêmes thématiques. Ces travaux ont été valorisés par 2 articles de revue internationale [Wir+12]; [Bou13a], 4 articles de conférence internationale à comité de lecture et actes [BSB09b]; [Wir+11b]; [Wir+11a]; [WJB13], 3 articles de conférence nationale à comité de lecture et acte [WBJ13]; [Kor+09]; [BSB09a].

Ces collaborations étant issues de mes travaux de thèse, elles sont aujourd'hui révolues, même si de manière ponctuelle, certains outils connexes comme l'utilisation de la théorie des matrices aléatoires de grandes dimensions peuvent donner lieu à des publications, 1 article de revue internationale [BB17] et 1 article de conférence nationale à comité de lecture et acte [BB15] ont été publiés.

Jusqu'en 2014 j'ai eu été membre de l'axe traitement du Signal du LASPI, UJM-Saint Etienne. La thématique affichée par cet axe concerne l'utilisation de la propriété de cyclostationarité des signaux à des fins de diagnostic de machines tournantes. Au travers de ces considérations, 1 article de revue internationale [Bou13a] et 4 articles de conférence internationale à comité de lecture et actes ont été publiés [Bou+08]; [BKG11]; [BK11]; [Cho+15]. Ces travaux proposent des méthodes et des estimateurs statistiques pour l'estimation fréquentielle de défauts. Ils ont permis de poursuivre la collaboration avec P. Wirfält et M. Jansson (Université KTH, Suède) et de faire naître une collaboration avec Vincent Choqueuse (Université Brest Occidentale).

Lors de mon intégration au laboratoire DISP en 2014, j'ai pu m'intéresser au domaine de la santé et m'impliquer dans une problématique liée aux hôpitaux sous tension. J'ai dû adapter mon champs disciplinaire initial afin de répondre à cette nouvelle thématique.

J'ai tout d'abords été sollicité par E. Marcon pour répondre à une problématique soulevée par les urgentistes pédiatriques. En effet en 2012 l'épidémie de Bronchiolite, causée principalement par le Virus Respiratoire Syncytial (VRS), a été nettement plus précoce et d'une très forte ampleur cette année-là. Si les responsables de service des urgences, et autres cadres de santé pilotaient grossièrement l'organisation, cette épidémie sans précédent jusqu'à lors, a causé une très forte désorganisation. Les responsables d'urgences pédiatriques ont alors mobilisé l'Agence Régional de

Santé (ARS) afin d'apporter une solution pour prévenir ces arrivées massives aux urgences. C'est donc dans ce cadre que j'ai commencé mon travail collaboratif avec le CHU de Saint-Etienne et ses urgences pédiatriques. Nous avons obtenu le financement du projet « **GO Plan Hivernal en Pédiatrie (GO PHP)** » par l'Agence Régionale de Santé Auvergne-Rhône-Alpes sur appel à projet innovant d'un montant total de 156 K€ qui a financé la thèse de **Maël DUGAST**. L'épidémie de bronchiolite étant jusqu'en 2020, une épidémie saisonnière, le flux de patients peut se modéliser sous une forme de processus périodiquement corrélé. Son étude m'a donc permis d'approfondir la connaissance et la modélisation de cette classe de processus par des outils de géométrie différentielle et de topologie algébrique, et a conduit en la création de premiers détecteurs précoces d'arrivées aux urgences. Cette partie sera développée dans le chapitre 3.

Au cours des discussions avec le consortium des médecins urgentistes et autres virologues, nous nous sommes aperçus que les résultats obtenus par la thèse de Maël DUGAST, qui se focalisaient sur les urgences pédiatriques pouvaient être utilisés pour les patients des urgences adultes. Les réflexions qui ont découlés nous ont permis de construire un nouveau projet et d'obtenir le financement du projet **PREPS-DGOS (Ministère de la Santé) PREDAFLU**, sur appel à projet générique d'un montant total de 384 K€. Ce projet finance entre autre la thèse de **Clément PEALAT**. L'idée principale de ce projet, est d'utiliser les premières alertes associées au flux de patients enfant pour la pathologie de Bronchiolite en conjonction des données virales des laboratoires pour les virus influenza (grippe) et VRS. Un flux de patients est déterminé en fonction des différents codes diagnostic obtenus et est ensuite monitoré pour un meilleur pilotage du service. Nous abordons cette problématique au travers du clustering de séries temporelles injectées sur des variétés différentielles appropriées et sera développé au chapitre 4.

Ces résultats fortement encourageants ont initié un certain nombre de projets. En particulier, nous avons participé à la contribution au tableau de bord de la stratégie nationale de santé via l'appel à projet de recherche pour la définition et l'élaboration technique d'indicateurs synthétiques. Nous avons obtenu le Lot n°1 : Analyse de l'indicateur « **Temps d'attente moyen dans les services hospitaliers d'accueil des urgences** » pour un montant total de 35 K€ et financé par le ministère de la santé. L'étude du flux de patients adultes des urgences m'a permis d'entreprendre des recherches sur la modélisation de ce flux par un processus auto-similaire voire multi-fractal et d'aborder la notion d'indicateur d'attente ou de congestion pour les urgences au travers des outils liés au Groupe de Renormalisation. Cette thématique sera abordée dans le chapitre 5 et correspond aux travaux de thèse de **Sid'Ahmed EMINE**.

Mon positionnement



Si la littérature est abondante concernant les effets de la désorganisation des services d'urgence en Amérique du Nord, les mêmes effets sont observés en Europe, et notamment en France. Il semble également que la littérature soit très vague quant à la définition de ces effets. Aussi, par souci de clarté, j'ai opté pour la définition donnée dans [Hoo+08], qui caractérise la rupture ou la tension des urgences comme un « encombrement », « crowding » ou encore « overcrowding » dans la littérature anglophone. La définition du crowding proposée à l'origine par l'American College of Emergency Physicians, est la suivante : « *Il y a encombrement lorsque le besoin identifié de services d'urgence dépasse les ressources disponibles pour la prise en charge des patients dans le service d'urgence, l'hôpital, ou les deux* ». Une revue de littérature des définitions et autres indicateurs sont discutés dans la thèse de F. Kadri [Kad14]. À partir de cette définition, j'ai donc interprété le crowding comme un phénomène qui implique l'interaction entre l'offre et la demande. Chaque service hospitalier ayant une capacité spécifique, le problème de l'encombrement se répercute instantanément sur l'organisation.

Au cours des dernières décennies, de nombreuses recherches se sont penchées sur les moyens de résoudre ces problèmes d'organisation. L'une d'entre elles consiste à assimiler les hôpitaux à des entreprises classiques dans lesquelles les patients sont considérés comme des produits. Cela dématérialise les patients et ne prend pas en compte toutes les contraintes des patients, mais cela permet de modéliser et de résoudre le problème. Par exemple, la modélisation par événements discrets [HA08]; [Hun+07]; [AGM09]; [Ken+21]; [Cub+20], fournit de bonnes solutions pour réduire les temps d'attente des patients, améliorer l'ordonnancement ou minimiser le nombre de chambres. L'utilisation de ce type d'approche pour traiter le problème permet de trouver des solutions à des situations complexes, dans lesquelles de nombreux paramètres interagissent de manière potentiellement non linéaire. Dans le contexte de la simulation, on retrouve également les travaux dédiés à la réduction du temps d'attente des patients et au dimensionnement des services, naturellement traités par le formalisme de la théorie des files d'attente.

Pour estimer la quantité nécessaire de ressources à l'état stationnaire d'un service d'accueil des urgences, la théorie des files d'attente, chaîne de Markov et processus de décision markovien ont montré une certaine efficacité [AHK06]; [PPQ08]; [ACL06]; [STM15]; [TWL12]; [LXL14]; [DGP15]. Ces approches montrent que le flux de patients peut être amélioré pour une meilleure prise en charge des patients. Dans ces modèles, les arrivées aux urgences sont modélisées par un processus de Poisson qui nous le verrons, n'est pas forcément la meilleure modélisation.

Enfin, ces mêmes problématiques sont abordées par le biais des systèmes multi-agents [YF17]; [BRA15]; [Dak+08]; [SBA17]; [Tab+13] qui de plus en plus sont mêlés

à des algorithmes d'optimisation pour augmenter la performance des systèmes d'aide à la décision.

Le flux de patients est donc un paramètre essentiel, et lorsqu'il est modélisé à l'aide de lois exponentielles ou équivalentes, il ne reflète pas la réalité. C'est une démonstration faite par exemple dans [MRV02] où les lois Poissonnienne et Lognormale sont mise à défaut. Cela suggère implicitement que la connaissance du flux de patients pourrait réellement aider à développer des solutions. En outre, le flux de patients est la cause et l'origine des effets de l'encombrement. Cette idée a été introduite par [Eit+10] ou [Pec+13]; [Pec13]; ce dernier fait référence au travail précédent de [Hal+06], dans lequel le flux de patients aux urgences a été étudié pour prédire le nombre de lits dans les unités d'hospitalisation. La problématique de prédiction est donc une thématique largement développée aussi [ABB09]; [Lin89]. Elle se base le plus souvent sur une modélisation du flux de patients par un modèle ARMA, SARIMA, ou encore GARCH. Le problème organisationnel dans les hôpitaux peut être alors réduit à l'étude du flux de patients aux urgences. Ce point a été illustré par [Sta+10].

Pour faire le point, les problématiques liées à l'organisation ou au pilotage de systèmes d'accueil des urgences sont traités par une modélisation du flux de patients sous l'angle

- des systèmes à évènements discrets avec des hypothèses de loi Poissonniennes pour les arrivées
- par des files d'attentes avec généralement les mêmes hypothèses pour les inter-arrivées
- par les systèmes multi-agents
- par modélisation statistique des arrivées sous la forme de modèles auto-régressifs et de leurs dérivés.

Ayant, due à ma formation, un peu plus d'accointances avec une approche statistique de la modélisation du flux, j'ai donc abordé la résolution des problématiques d'encombrement ou de tension dans les services d'accueil des urgences par la constitution d'outils d'aide à la décision pour le pilotage (et non pour le dimensionnement) via l'approche statistique. La décision proposée pour le pilotage n'est pas issue de modèles mathématiques des services, ni de la modélisation du flux sous sa forme simulée mais issus de la dynamique, ou variabilité de ce dernier. Je ne cherche donc pas à caractériser le système mais plutôt d'en comprendre les effets, et d'utiliser le flux de patients comme un indicateur de la tension, et de décision.

3

Prévenir la tension aux urgences pédiatriques



Afin de répondre aux sollicitations des urgentistes pédiatriques souhaitant avoir un outil de détection précoce de venues massives aux urgences, je me suis intéressé à la structure du flux de patients. Je développerai donc dans ce chapitre la notion de dilatation d'opérateur pour la représentation du processus stochastique associée et les outils de géométrie différentielle et de topologie algébrique utilisés. Ces outils servent d'une part à caractériser de manière intrinsèque au

travers des trajectoires des opérateurs de dilatation sur un groupe de Lie le processus lui-même, et d'autre part à proposer des détecteurs précoces.

Sommaire

3.1	Un outil statistique simple	26
3.2	Dilatation d'opérateurs	28
3.2.1	Principes théoriques-le côté accessible	28
3.2.2	Principes théoriques-le côté mathématiques	29
3.2.3	Construction des matrices de Dilatation	31
3.2.4	Une implémentation pas à pas par l'exemple.	31
3.3	Comparaison et caractérisation intrinsèque des processus.	34
3.3.1	Indicateur intrinsèque par l'espace des formes.	34
3.3.2	Indicateur intrinsèque par la topologie algébrique	37
3.4	Mise en place des indicateurs	45
3.4.1	L'indicateur précoce basé sur la topologie	45
3.4.2	L'application web	49
3.4.3	Retour d'expérience	51
3.5	Contributions scientifiques de cette partie	51

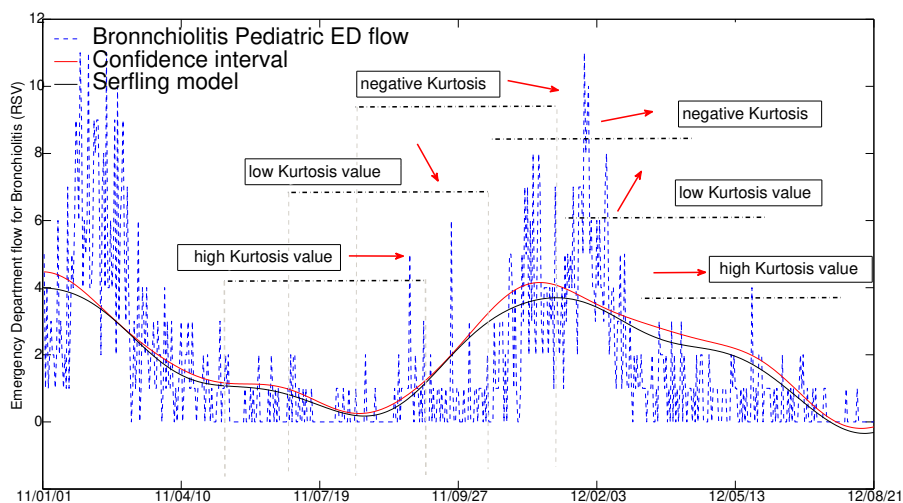
3.1. Un outil statistique simple

Construire un détecteur précoce de venues massives aux urgences peut s'apparenter à la problématique du *change point détection*. Si l'on considère que les admissions aux urgences sont un phénomène prédictible la plupart du temps, lors d'une épidémie, les venues deviennent non-prédictibles et un phénomène anormal apparaît. La détection de cette anormalité représente le changement d'état du phénomène associé aux arrivées des patients et caractérise le début d'une épidémie. C'est à partir de ce constat que j'ai donc appliqué des méthodes existantes pour la détection de changement d'état. En particulier, la méthode dite du *Serfling* est une méthode employée par les organismes de surveillance sanitaire depuis quelques années maintenant. Il s'agit d'approximer par un processus multi-périodiques le flux d'admissions aux urgences des patients avec l'idée de capturer la périodicité et les variations hautes fréquences sans pouvoir capturer le coté outlier dans les données qui représente l'aspect non-prédictif de l'épidémie. Ainsi, le modèle paramétrique obtenue est comparé à la courbe réelle des admissions. Lorsque la courbe réelle dépasse celle du modèle, il est considéré que l'épidémie débute. Le modèle sinusoïdale multi-périodique utilisée est le suivant

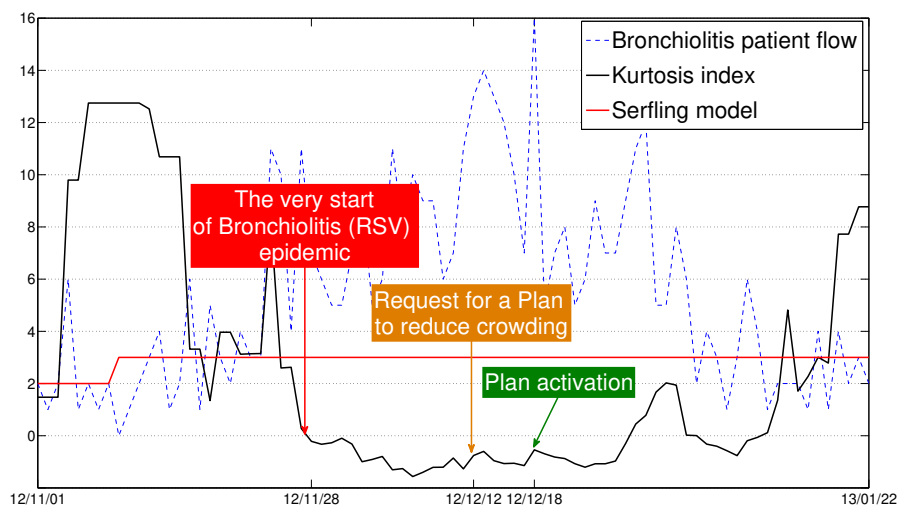
$$\hat{x}(t) = \hat{\beta}_0 + \hat{\beta}_1 t + \hat{\beta}_2 \cos\left(\frac{2\pi}{\hat{T}_1} t\right) + \dots + \hat{\beta}_{n_p+2} \cos\left(\frac{2\pi}{\hat{T}_{n_p}} t\right). \quad (3.1)$$

Dans [BMM15] est discuté de l'estimation des différents paramètres du modèle. Cette procédure généralement employée pour des données hebdomadaires ne peut s'appliquer aux admissions aux urgences sur la base de données quotidiennes dues à de trop fortes variations des admissions. J'ai donc proposé d'améliorer la détection en analysant statistiquement la série des erreurs de prédiction du modèle. L'indicateur statistique largement employé en maintenance prédictive sur les machines tournantes est le kurtosis. J'ai donc mis en place un détecteur basé sur l'évolution du Kurtosis des erreurs de prédictions. Les résultats sont exposés dans [BMM15]. Ci-après en Figure 3.1, une image du fonctionnement global du détecteur ainsi qu'une image montrant le retard d'activation du plan hivernal des urgences pédiatriques du CHU de Saint-Etienne par rapport à la détection faite par l'outil statistique. Sur cet exemple, l'outil n'était pas encore mis en place. En se positionnant du côté temps réel et opérationnel, on peut notamment s'apercevoir que l'activation du plan hivernal par la direction du CHU a été donnée presque 3 semaines après la détection, alors même que l'affluence aux urgences était en diminution et que l'épidémie était proche de sa fin.

Ces résultats positifs m'ont ensuite amené à modéliser plus finement ce processus d'admissions. Je me suis alors tourné vers des éléments théoriques sur les processus périodiquement corrélés, m'amenant doucement à la théorie de la dilatation des opérateurs grâce à un ouvrage de T. Constantinescu [Con95], majeur dans mes recherches.



(a)



(b)

Figure 3.1: En haut, la figure explicative du fonctionnement du Kurtosis sur le flux d'arrivées des patients avec diagnostic de bronchiolite sur 1 année. En bas, illustration de l'intérêt de l'indicateur de Kurtosis pour la mise en place du plan hivernal en pédiatrie.

3.2. Dilatation d'opérateurs

L'analyse et/ou la représentation des processus non stationnaires est abordée depuis quatre ou cinq décennies à présent par l'analyse temps-échelle/temps-fréquence [Fla99]; [Aug+13] ou encore par la représentation de Fourier lorsque les processus appartiennent à la classe des processus périodiquement corrélés (PC) [Nap16]; [HM07a]. D'autres approches, moins connues, utilisent les coefficients de corrélation partiels (parcours) [Lam05]; [DL03]. L'un des avantages à utiliser les parcours réside dans leur forte relation avec la mesure du processus par la relation point-à-point qui les lie avec les coefficients de corrélations [Des03a]; [YAB10]. Ils apparaissent par conséquent explicitement dans la décomposition en polynôme orthogonaux de la mesure [Bin11]; [Sim09], et sont les éléments pour la construction de matrices de dilatation qui apparaissent dans les matrices Cantero Moral et Velazquez (CMV) / Géronimus, Gragg et Teplyaev (GGT) [Sim07], dans le problème des flots de Schur [AGR91], sont également vues dans la littérature en tant qu'opérateurs d'évolution [Sim09] ou opérateur de décalage [Mas78], et apparaissent enfin dans la représentation d'espace-état [Con95]; [Des96]. Les références données précédemment ne sont qu'un échantillon faible de la définition et de l'utilisation de cet opérateur qui, nous venons de la constater, apparaît dans un nombre important de thématiques différentes. La découverte et la compréhension de cet objet mathématique a donc été une tâche importante tant sa définition, souvent théorique est plurielle. Le côté positif étant que cet objet me semble aujourd'hui un objet relativement central.

Quel angle choisir pour présenter cet objet ... difficile de faire un choix ! Je propose donc une explication simplifiée et intuitive qui ne devrait pas nuire à la compréhension et un partie pris basé sur la représentation de groupe.

3.2.1. Principes théoriques-le côté accessible

La théorie de la dilatation s'inspire de la théorie des opérateurs [Sz+10], qui relie la mesure et les opérateurs unitaires du processus. Dans sa version la plus simple, la théorie de la dilatation correspond à la dilatation de Naimark [Sz+10], et stipule qu'étant donné une séquence de coefficients de corrélation $\{R_i\}_{i \in \mathbb{N}}$, il existe une matrice unitaire W telle que

$$R_i \triangleq (1 \ 0 \ 0 \ \dots) W^i (1 \ 0 \ 0 \ \dots)^T$$

où \cdot^T dénote la transposition. On s'aperçoit au passage que les puissances successives de l'opérateur de dilatation conduisent à l'estimation des coefficients de corrélations. Lorsque le processus n'est pas stationnaire, sa matrice de corrélation associée n'est plus structurée Toeplitz, un ensemble de matrices est alors requis [Con95] et l'expression précédente devient

$$R_{i,j} \triangleq (1 \ 0 \ 0 \ \dots) U_{i+1} U_{i+2} \dots U_j (1 \ 0 \ 0 \ \dots)^T$$

Les matrices U_i sont théoriquement comprises comme des matrices de rotation car elles traduisent l'orthogonalisation et la renormalisation de l'espace de Hilbert dans lequel se représente la matrice de corrélation/covariance. Elles sont théoriquement infinies donc mais deviennent finies lorsque la séquence des coefficients de corrélation est elle-même finie. Dans ce cas particulier, les matrices U_i appartiennent à $SO(n)$ ou $SU(n)$, le groupe spécial orthogonal ou unitaire, respectivement, et la mesure du processus est totalement décrite par l'ensemble des U_i . En conséquence, la mesure du processus est magnifiquement caractérisée pour le cas nonstationnaire, par une trajectoire échantillonnée induite par les matrices de dilatation sur un groupe de Lie (sans démonstration, les groupes $SO(n)$ et $SU(n)$ sont également des variétés, conférant ainsi cette structure de groupe de Lie). Lorsque le processus est périodiquement corrélé, la séquence de parcos est périodique, ainsi que la séquence de matrices de dilatation, ce qui donne une courbe fermée sur les groupes de représentation.

3.2.2. Principes théoriques-le côté mathématiques

Le parti pris que je propose ici est celui qui est guidé par l'approche de [Wer77a]. Dans cet article, la dilatation de la mesure spectrale est présentée de la manière suivante. Soit $\{X_t\}_{t \in T}$ un processus stationnaire vivant sur un groupe T , il existe une représentation unitaire U de T sur $\mathcal{H}_X = sp\{X_t, t \in T\}$, appelé le domaine temporel [HKF04]; [Hur91]; [HM07a]; [MS89] et tel que $X_t = U_t X_e$, où e est l'élément d'identité du groupe T . Dans ce contexte, l'opérateur U est appelé opérateur de décalage, il décrit comment passer de tout vecteur X_{t-1} de \mathcal{H}_X à X_t . Avec maintenant G le dual de T , la transformée de Fourier de la mesure spectrale F de $\{X_t\}_{t \in T}$ est donnée par

$$R_\tau = \int_G [\tau, g] F(dg), \quad \tau \in T \quad (3.2)$$

où $[\tau, g]$ est la valeur d'un caractère g sur τ . Lorsque $T = \mathbb{R}$, les caractères sont les fonctions $g : x \mapsto e^{2k\pi x}$, $k \in \mathbb{R}$, G est isomorphe à \mathbb{R} et on obtient la transformée de Fourier habituelle. Enfin, il existe un espace de Hilbert $\mathcal{H} = \mathcal{H}_X$ et une représentation unitaire U_g telle que

$$\begin{aligned} R_\tau = \langle X_\tau, X_e \rangle &= \langle U_\tau X_e, X_e \rangle \\ &= X_e^* U_\tau X_e = \int_G [\tau, g] X_e^* E(dg) X_e, \end{aligned} \quad (3.3)$$

où E est la mesure spectrale de U_g . Par l'unicité de la transformée de Fourier, nous avons donc

$$F(\Delta) = X_e^* E(\Delta) X_e \quad \Delta \in \Sigma \quad (3.4)$$

avec Σ une σ -algèbre fixée de T^* . Cette expression signifie finalement que la mesure spectrale du processus $\{X_t\}_{t \in T}$ est incluse dans une mesure spectrale plus grande et située en dans la partie supérieure gauche de cette dernière, de sorte que

$$E(\Delta) = \begin{pmatrix} F(\Delta) & * \\ * & * \end{pmatrix}, \quad (3.5)$$

$E(\Delta)$ est donc la dilatation spectrale de $F(\Delta)$. Nous reconnaissons ici la dilatation de Naimark, qui peut également être formulée du point de vue des opérateurs par le théorème fondamental suivant

Theorem 1 (Dilatation de Naimark) *Une suite d'opérateurs $\{R_n\}_{n \in \mathbb{N}}$ sur un espace de Hilbert \mathcal{H} est définie positive, i.e. $\sum_{0,0}^{\infty, \infty} \langle R_{i-j} h_i, h_j \rangle \geq 0$ pour toutes séquences $\{h_i\}_0^\infty$ si et seulement si il existe une isométrie U sur $\mathcal{K} \subset \mathcal{H}$ tel que*

$$R_{ij} = P_{\mathcal{H}} U|_{\mathcal{H}}^{j-i}, \quad (3.6)$$

où $P_{\mathcal{H}}$ représente la projection orthogonale sur \mathcal{H} et $|_{\mathcal{H}}$ indique une réduction de dimension à \mathcal{H} . En outre, si

$$\mathcal{K}_R = \bigvee_{n \geq 0} U^n \mathcal{H}_n \quad (3.7)$$

alors U est unique à un isomorphisme près.

Fondamentalement, (3.6) implique que la séquence de corrélation entière est obtenue par puissance successive de l'opérateur de dilatation unitaire U . En fait, en reliant (3.3) et (3.6), on voit que l'opérateur de dilatation unitaire U n'est autre que l'opérateur de décalage. En parcourant la littérature, nous avons trouvé de nombreuses références traitant de l'utilisation du théorème de Naimark pour la caractérisation des processus stochastiques. Par exemple, [Sz+10] ou [TY72] abordent la théorie de la dilatation par des considérations de mesure spectrale alors que [Fra86]; [FF90]; [KB86]; [Kai82a] utilise les suites définies positives comme point de départ pour développer la dilatation de Naimark. Enfin, les lecteurs trouveront des preuves explicites de la correspondance entre l'opérateur de décalage et la dilatation de Naimark dans [Mas78] ou [Wer77b].

Je finirai cette partie en donnant la structure de la matrice de dilatation unitaire lorsqu'un processus est non stationnaire. Dans ce cas, la séquence d'opérateurs du Théorème 1 n'est plus expliquée par un seul opérateur unitaire U . En fait, il suffit de penser que pour chaque temps d'origine, il existe une dilatation unitaire spécifique. C'est le concept de non-stationnarité. Les statistiques du processus ne dépendent pas seulement du décalage mais aussi du moment où elles sont calculées. A chaque instant, il existe une composante spectrale différente, qui donne lieu à des dilatations unitaires différentes. En gardant la même notation que celle adoptée pour la définition dans le domaine temporel de l'opérateur de décalage, il existe à chaque instant un nouvel opérateur de décalage. Alors, $X_t = U_t X_{t-1} = \prod_{i=0}^t U_i X_e$. La généralisation formelle de la dilatation unitaire de Naimark est donc donnée par la décomposition de Kolmogorov qui s'écrit

Theorem 2 (Décomposition de Kolmogorov) *Soit R un noyau défini positif. Alors, il existe un espace de Hilbert \mathcal{H} et des applications $U_i, i \in \mathbb{Z}$ telles que*

$$\begin{aligned} R_{i,j} &= P_{\mathcal{H}} U_i U_{i-1} \cdots U_{j-1} |_{\mathcal{H}} \quad \forall i, j \in \mathbb{Z} \\ \mathcal{K}_R &= \bigvee_{n \in \mathbb{Z}} V^n \mathcal{H}_n \end{aligned}$$

avec $V^i = \prod_{i=0}^{i-1} U_i$. Une séquence d'opérateurs est alors obtenue par le produit des dilatations unitaires $U_i, i \in \mathbb{Z}$.

3.2.3. Construction des matrices de Dilatation

L'idée, principalement exposée dans [Con95] consiste en la paramétrisation des matrices de corrélations par une ou des séquences de nombres, appelés séquence de choix, coefficients de Geronimo ou Verblunsky, paramètres de Schur ou encore coefficients de réflexion partiels (parcours) selon le contexte. Pour simplifier l'étude, nous nous plaçons dans le cadre des matrices définies positives. Il existe en fait une bijection entre l'ensemble des matrices définies positives et les parcours, comme cela a pu être noté par Levinson, Burg, et plus récemment par Lambert-Lacroix ou Barbaresco, dans le cas de signaux stationnaires. Cette paramétrisation dans sa forme la plus simple associe les coefficients de réflexion partiels que nous nommerons $\{\Gamma_{i,j}\}_{i<j}$, à la matrice de corrélation. A partir des $\{\Gamma_{i,j}\}_{i<j}$ on forme les rotations de Givens :

$$G_{j-k}(\Gamma_{k,k+l}) = I \oplus \begin{pmatrix} \Gamma_{k,k+l} & D_{\Gamma_{k,k+l}^*} \\ D_{\Gamma_{k,k+l}} & -\Gamma_{k,k+l}^* \end{pmatrix} \oplus I \quad (3.8)$$

avec I la matrice identité et $D_{\Gamma_{i,j}} = (I - \Gamma_{(i,j)}^* \Gamma_{i,j})$ l'opérateur de défaut de la contraction $\Gamma_{i,j}$. En notant :

$$U_i = G_{j-i}(\Gamma_{i,i+1}) G_{j-i}(\Gamma_{i,i+2}) \cdots G_{j-i}(\Gamma_{i,j}) \quad (3.9)$$

on peut alors écrire :

$$R_{i,j} = P_1 U_i U_{i-1} \cdots U_{j-1} |_{H_1} \quad (3.10)$$

où $P_1 = (1 \ 0 \ \cdots)^T$ c'est-à-dire la projection orthogonale sur la première ligne, et H_1 est la restriction à la première colonne. On retrouve alors la relation énoncée précédemment $R_{i,j} = (1 \ 0 \ \cdots) U_{i+1} U_{i+2} \cdots U_j (1 \ 0 \ \cdots)^T$, correspondant à la décomposition de Kolmogorov. La structure d'une matrice de dilatation U_i est telle que

$$U_i = \begin{pmatrix} \Gamma_1 & D_{1^*} \Gamma_2 & D_{1^*} D_{2^*} \Gamma_3 & \cdots \\ D_1 & -\Gamma_1^* \Gamma_2 & -\Gamma_1^* D_{2^*} \Gamma_3 & \cdots \\ 0 & D_2 & -\Gamma_2^* \Gamma_3 & \cdots \\ 0 & 0 & D_3 & \cdots \\ 0 & 0 & 0 & \cdots \\ 0 & 0 & 0 & \cdots \end{pmatrix}.$$

U_i appartient à la classe de matrices dites Hessenberg supérieures à coefficients sous-diagonaux positifs. Bien évidemment, lorsque le processus est stationnaire, les calculs sont grandement simplifiés. Ainsi, on a $U = U_i, \forall i, \Gamma_{i,j} = \Gamma_{j-i}$.

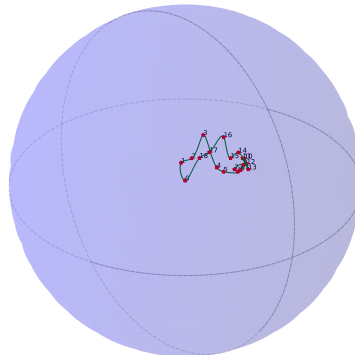
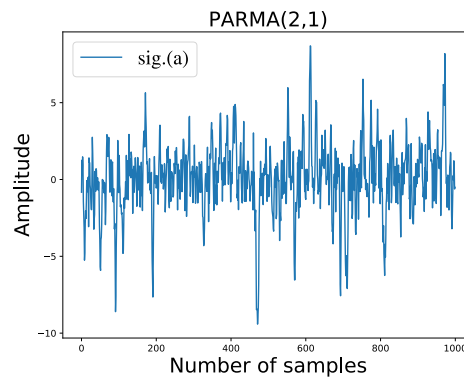
3.2.4. Une implémentation pas à pas par l'exemple

Afin d'illustrer plus pratiquement cette démarche théorique, nous proposons ici d'étudier cinq processus périodiquement corrélés dont la périodicité est égale à $T = 20$ échantillons pour chacun des processus. Le package R `perARMAR` ainsi que la fonction incluse `perpacfs` sont utilisés pour calculer ensuite les 20 séquences de 3 parcours (choix arbitraire ici) nécessaires pour la construction des matrices de dilatation. Pour rappel, comme $X_{t+T} = U_t X_t$ pour un processus périodiquement corrélé de période T , seules les T séquences apportent l'information spectrale. Enfin, l'estimation de trois parcours conduisant à une représentation sur $SO(4)$, nous ne gardons que les

deux premiers pour une visualisation sur $SO(3)$. Les T matrices de dilatation sous-unitaires U_i sont ensuite obtenues grâce à l'ensemble $\{\Gamma_{i,j}\}_{i=1,j=1}^{i=20,j=2}$ de parcours et en appliquant la relation

$$U_i = \begin{pmatrix} -\Gamma_{i,1} & \sqrt{1 - |\Gamma_{i,1}|^2} & & \\ \sqrt{1 - |\Gamma_{i,1}|^2} & \Gamma_{i,1} & & \\ & & 1 & \\ & & & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & & & \\ -\Gamma_{i,2} & \sqrt{1 - |\Gamma_{i,2}|^2} & & \\ \sqrt{1 - |\Gamma_{i,2}|^2} & \Gamma_{i,2} & & \\ & & 1 & \end{pmatrix} \quad \forall i = 1, \dots, 20 \quad (3.11)$$

La Figure 3.2 ci-après illustre cette démarche.



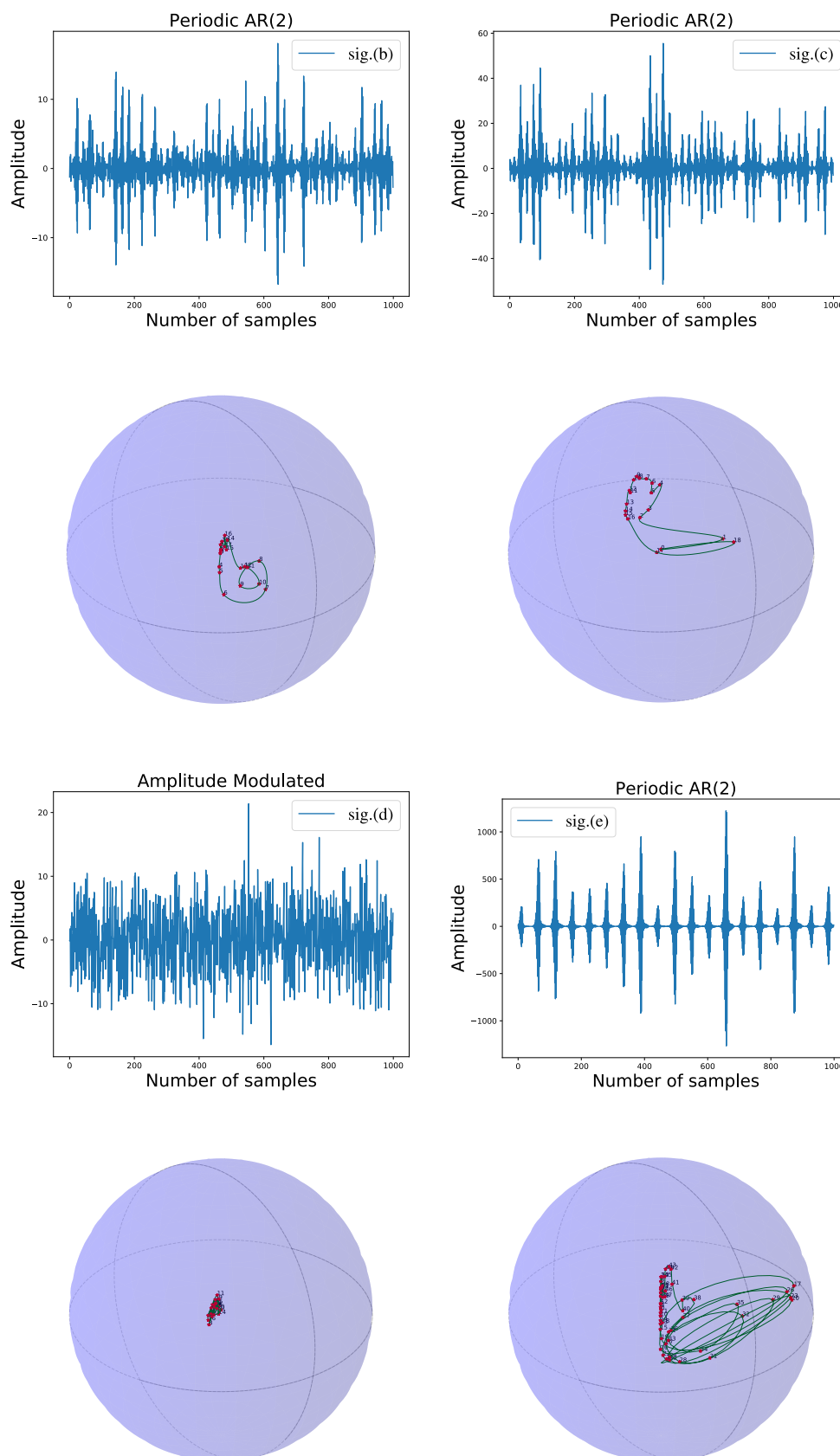


Figure 3.2: Une réalisation de chacun des 5 signaux des classes PAR(2), Modulation d'Amplitude, PARMA(2,1) et PAR(2). Chacun de ces signaux est corrélé périodiquement avec une période $T = 20$ échantillons parmi les 1000 échantillons disponibles. Les *sig.(b)* et *sig.(c)* ont des coefficients générateurs différents, *sig.(e)* a les mêmes coefficients générateurs que celui de *sig.(c)* mais avec une période de $T = 54$. Un exemple de trajectoire formée par les matrices de dilatation dans $SO(3)$ pour les signaux associés. La trajectoire est représentée à l'intérieur de la boule de rayon π . Pour tracer la courbe, nous avons d'abord interpolé avec des splines dans l'algèbre de Lie et utilisé ensuite la carte exponentielle pour revenir sur $SO(3)$.

Cette figure nous apporte quelques éléments somme toute intéressant. On peut constater que chacune des classes de processus générés à une trajectoire ou une forme différente. On distinguera en particulier, la modulation d'amplitude $sig.(d)$, dont la trajectoire dénote un processus presque stationnaire et le processus AR(2) avec une période $T = 54$ ($sig.(e)$) qui possède un grand nombres de cercles. Les processus AR(2) générés avec des coefficients générateurs différents ($sig.(b)$ et $sig.(c)$) présentent quant à eux une forme très similaire avec 2 cercles plus ou moins distincts. L'approche de la matrice de dilatation est donc bien une approche qui porte de l'information et semble-t-il une information intrinsèque à la classe de processus. Je noterai que la recherche d'information intrinsèque des processus par caractérisation de trajectoires issues des parcours a été développé presque en parallèle par A. Le Brigant [Bri17] sans connaissance de ma part à l'origine. Toutefois, les résultats obtenus dans sa thèse ne sont pas probants et ne permettent pas l'extraction d'une information intrinsèque. La méthodologie qu'elle a employée et développée pour comparer les trajectoires sur des espaces de formes, a été en revanche une très forte source d'inspiration et m'a permis de découvrir un champ thématique et une communauté que je ne connaissais pas.

Disposant de trajectoires (échantillonnées) sur des groupes de Lie ($SO(n)$ dans notre cas), la question restait de pouvoir montrer ce côté « information intrinsèque ». Pour initier la réponse j'ai donc proposé de comparer les trajectoires en termes de forme, et montrer leur différence donc. Puis plus profondément encore, j'ai décidé de montrer par la topologie algébrique et l'homologie persistante que chacune des trajectoires avait bien une topologie associée à la classe d'un processus. Je vais donc aborder ces deux points maintenant.

3.3. Comparaison et caractérisation intrinsèque des processus

3.3.1. Indicateur intrinsèque par l'espace des formes

Nous nous intéressons donc aux courbes reposant sur une variété, et plus précisément sur un groupe de Lie. Notons c une telle courbe, $c : [0, 1] \in SO(n)$. Pour étudier la structure géométrique de ces objets, considérons l'ensemble des courbes de $SO(n)$ défini par $\mathcal{M} = \{c \in \mathcal{C}^\infty([0, 1], SO(n)) : c'(t) \neq 0 \forall t\}$. Lorsque nous comparons deux courbes, il est naturel de définir une distance entre ces courbes qui ne dépendent pas de leurs paramétrisations ou de leurs rotations. Cela revient à reparamétriser les courbes par un difféomorphisme croissant $\phi \in \mathcal{D} : [0, 1] \rightarrow [0, 1]$ et impose que la métrique sur \mathcal{M} soit invariante par reparamétrisation. Cette propriété conduit à la relation d'équivalence

$$c_0 \sim c_1 \Leftrightarrow \exists \phi \in \mathcal{D} : c_0 = c_1 \circ \phi. \quad (3.12)$$

pour deux courbes c_0, c_1 de \mathcal{M} . De cette relation nous obtenons naturellement une définition d'un espace quotient de \mathcal{M} , généralement nommé l'espace des formes et s'écrivant

$$\mathcal{S} = \mathcal{M} / \sim, \text{ ou } \mathcal{S} = \mathcal{M} / \mathcal{D}. \quad (3.13)$$

On en déduit une distance sur l'espace des formes à partir de celle définie sur \mathcal{M}

$$d_S([c_0], [c_1]) = \inf_{\phi \in \mathcal{D}} d_{\mathcal{M}}(c_0, c_1 \circ \phi), \quad (3.14)$$

où $[c_0], [c_1]$ sont respectivement les représentants des classes d'équivalences de c_0 et c_1 .

Métrique et Distance sur \mathcal{S} puis \mathcal{M}

Nous nous proposons maintenant de donner quelques informations sur le choix d'une métrique pertinente sur \mathcal{M} afin de comparer les différentes courbes fermées. Comme l'espace de base est un groupe de Lie, nous utilisons l'approche de Celledoni, mais beaucoup d'autres travaux ont été proposés pour tenir compte de la structure (ou non) homogène de la variété de base. Dans [BBM16] il est montré que malheureusement nous ne pouvons utiliser la simple métrique L^2 sur \mathcal{M} ou \mathcal{S} car elle conduit à une distance nulle. Il est dans ce cas impossible de différencier 2 formes de courbes différentes qu'elles appartiennent à \mathcal{M} ou \mathcal{S} . Pour pallier cette difficulté, la famille des métriques élastiques, dérivée de la métrique de Sobolev, a été étudiée. Dans le cas d'un espace euclidien \mathbb{R}^n , elle admet l'expression :

$$g_c^{a,b}(u, v) = \int (a^2 \langle D_t u^N, D_t v^N \rangle + b^2 \langle D_t u^T, D_t v^T \rangle) \|c'\| dt,$$

où $D_t u = u' / \|c'\|$, $D_t u^T = \langle D_t u, w \rangle w$, avec $w = c' / \|c'\|$ and $D_t u^N = D_t u - D_t u^T$. Dans notre cas il est montré que $a, b = 1, 1/2$ permet d'avoir une équivalence avec une transformation à la fois simple d'interprétation et d'implémentation, la TSRV. La TSRV, adaptée à notre cas, paramétrise la courbe c par son vecteur vitesse et son point d'ancrage puis transporte les nouvelles courbes appartenant aux espaces tangents successifs par translation à droite, tous en même point de référence, qui sera l'identité e de $SO(n)$. C'est une isométrie de \mathcal{M} vers l'espace des courbes lisses. Sa définition est

$$F_{Lie} : \mathcal{M} \longrightarrow SO(n) \times L^2([0, 1], \mathfrak{g}) \quad (3.15)$$

$$F_{Lie}(c)(t) = (e, q(t)) = \left(e, \frac{T_c^{c(t) \rightarrow I}(c'(t))}{\sqrt{\|c'(t)\|}} \right) \quad (3.16)$$

où \mathfrak{g} est l'algèbre de Lie, $\|\cdot\|$ une norme induite par une métrique invariante à droite sur $SO(n)$, et où $T_c^{c(t) \rightarrow I}$ correspond au transport de $c(t)$ vers l'identité suivant la courbe c . Avec cette transformation, qui correspond donc à une métrique de Sobolev du premier ordre, la distance voulue en (3.14) conduit à résoudre le problème d'optimisation :

$$d_S([c_0], [c_1]) = \inf_{\phi \in \mathcal{D}} \left(\int_0^1 \|q_0(t) - q_1(\phi(t)) \sqrt{\phi'(t)}\|^2 \right)^{1/2}, \quad (3.17)$$

qui est résolu par un algorithme de descente de gradient traditionnel ou une programmation linéaire dynamique [CES15]. Nous obtenons aussi une expression familière pour l'interpolation géodésique entre deux courbes c_0 et c_1 , exprimée dans leur domaine TSRV :

$$\mathcal{Q}(s) = F_{Lie}^{-1}((1-s)F_{Lie}(c_0) + sF_{Lie}(c_1)) \text{ pour } s \in [0, 1].$$

Nous avons maintenant tous les ingrédients pour décrire la procédure de caractérisation et de comparaison de processus nonstationnaires :

1. **Input:** Un ensemble de matrices de rotations $\{U_i\}_i$;
2. Injecter l'ensemble des matrices $\{U_i\}$ dans leurs algèbre de Lie respectives $\{V_i\}$ via l'application exponentiel inverse;
3. Faire une interpolation par splines entre les matrices V_i [HP04]; [Mai20];
4. Revenir sur la variété de base $SO(n)$ via l'application exponentielle ;
5. Translater la courbe de manière à satisfaire la condition $c(0) = e$ puis faire la transformation SRV donnée par (3.16) ;
6. Calculer la distance définie par (3.17).
Le problème d'optimisation (non-crossing matching graph) est résolu par un programme dynamique ;
7. **Output:** Distance entre les courbes sur \mathcal{M} , et interpolation géodésique entre ces courbes.

Quelques distances

Nous avons implémenté la procédure précédente et déterminé quelles étaient les distances entre les trajectoires des classes de processus sur $SO(3)$, $SO(4)$ et $SO(5)$. Les distances sont rapportées à l'intérieur de la Table 3.1. Clairement, les distances entre les formes des courbes caractérisant la mesure spectrale de chaque processus révèlent une certaine proximité spectrale entre les processus référencés. Le signal intéressant est celui qui est indexé par la lettre (e) et représente un signal PAR(2) avec exactement les mêmes paramètres de modèle que celui du signal de référence, mais avec une périodicité de 54 points au lieu de 20. La courbe associée à ce processus a de nombreux lacets dans sa représentation, ce qui fait que sa distance est la plus grande sur $SO(3)$, mais lorsque la dimension augmente, sa distance au processus (b) devient la plus faible. Ceci est particulièrement intéressant. Il semble qu'une hiérarchie se mette en place à partir de $SO(4)$ où le processus le plus proche au processus PAR(2) soit le processus PAR(2) de mêmes paramètres générateurs mais avec une périodicité différente, suivi par le processus modulé en amplitude. Le processus qui semble le plus éloigné dans cette hiérarchie semble être le processus de classe PARMA(2,1). Dans [DBM18b], nous avons également comparé les distances en fonction de l'interpolation par splines des trajectoires, mais les résultats restent identiques et la hiérarchie reste conservée. Je préciserai enfin que les interpolations de géodésiques entre les courbes fermées représentatives des mesures spectrales des processus sont données dans [DBM18b].

Modèle de Signal affiché dans la Figure 3.2	Distance au Signal (b)		
	SO(3)	SO(4)	SO(5)
(a)	4.19	187.42	621.51
(d)	5.72	97.19	526.36
(e)	31.63	41.78	298.64

Table 3.1: Distances entre les processus (a),(d),(e) et le processus de référence (b) de la Figure 3.2 calculées par (3.17)

3.3.2. Indicateur intrinsèque par la topologie algébrique

La base même de l'homologie est de caractériser des formes à n dimensions (disons des formes discrètes) par le nombre de trous ou cavités à k dimensions, $k < n$. Contrairement à l'homotopie, qui utilise la notion de boucles, *i.e.*, un chemin fermé qui commence et se termine au même point, l'homologie utilise la notion de cycles, *i.e.* un chemin fermé sans considération de son origine. Pour déterminer le nombre de trous dimensionnels d'une forme, les points sont joints ensemble pour former ce qu'on appelle une cellule dans le contexte de la topologie algébrique. Lorsque les cellules sont collées ensemble, elles forment un complexe, une sorte de squelette de la forme. Il existe plusieurs possibilités pour définir une cellule, et donc un complexe. Pour répondre à nos besoins, nous avons choisi le complexe simplicial qui répond aux objectifs de simplicité d'interprétation et de calcul. Le squelette de la forme est alors décrit par des complexes simpliciaux (simplexes) à plusieurs dimensions, dont les formes sont illustrées par la Fig. 3.3. L'idée derrière l'homologie est de rechercher ensuite le chemin fermé créé par la composition des complexes simpliciaux. Pour mettre être un peu plus concret, un trou ou une cavité à 1 dimension, que nous associerons à un cercle, est caractérisé par la somme des segments (1-simplexe) du squelette pour former une trajectoire fermée (un cycle). Si maintenant cette trajectoire fermée n'est pas une partie (une frontière) d'une autre forme constituée par la composition de triangles (2-simplexes), alors on affirme que la trajectoire fermée représente un trou à 1 dimension pour la forme. Dans ce cas, elle appartient au premier groupe d'homologie, noté H_1 et l'invariant topologique associé au nombre de trous unidimensionnels indépendants pour la forme est défini comme le nombre de Betti, $\beta_1 = \text{rank}(H_1)$. La forme est finalement décrite par ses nombres de Betti $(\beta_0, \beta_1, \dots, \beta_k) \forall k < n$.

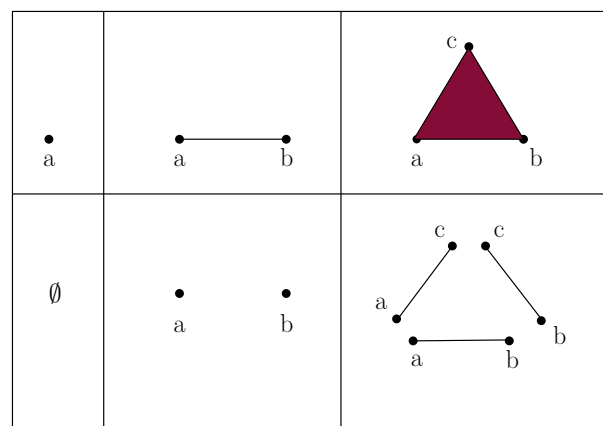


Figure 3.3: Un n -simplexe est l'enveloppe convexe de $(n+1)$ points dans un espace à n dimensions. C'est la généralisation d'un triangle. En haut (de gauche à droite) : Un 0-simplexe est un point, c'est (a) dans le graphe. Un 1-complexe est un segment, c'est (a, b) (où a et b sont les sommets qui sont à l'extrémité du segment), un 2-complexe est un triangle (a, b, c) dans le plan. En bas (de gauche à droite) : les faces correspondantes.

Un petit exemple pour fixer les idées

Donnons un exemple simple pour illustrer ces concepts. Dans la Fig. 3.3 le triangle $v = (a, b, c)$ a des faces (ab), (bc), et (ca) qui les segments (1-simplexe) le constituant.

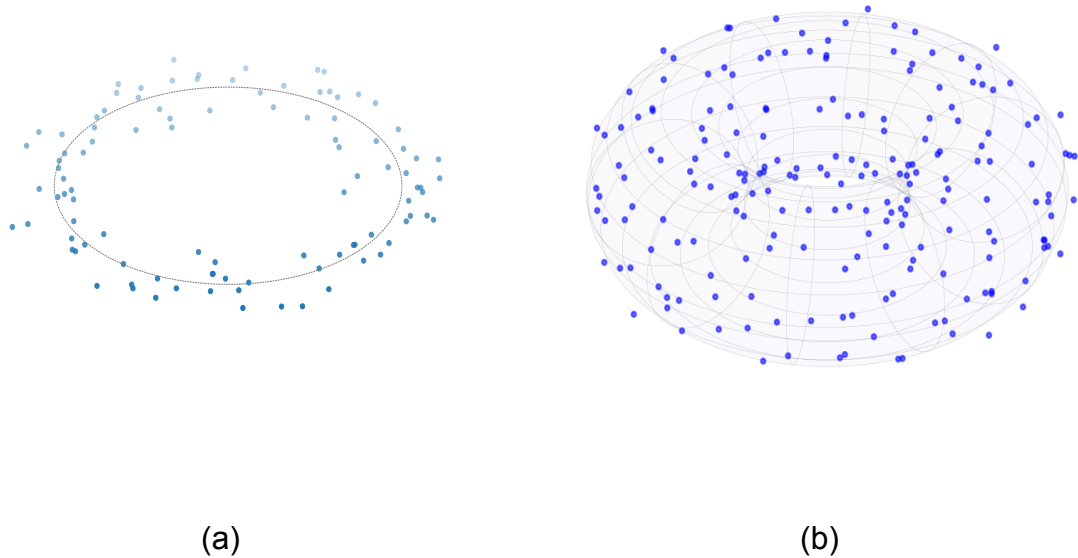


Figure 3.4: (a) points répartis le autour d'un cercle, les nombres de Betti sont $(\beta_0 = 1, \beta_1 = 1)$, (b) points répartis autour d'un tore, les nombres de Betti sont $(\beta_0 = 1, \beta_1 = 2, \beta_2 = 1)$.

La bordure du triangle est alors $(ab) + (bc) + (ca)$ qui définit un chemin fermé. Par conséquent, $(ab) + (bc) + (ca)$ est un cycle, mais c'est aussi une bordure du 2-simplexe, le triangle v lui-même. Nous en concluons que $(ab) + (bc) + (ca)$ ne représentent pas un trou.

Pour terminer cette partie d'illustration, nous proposons au lecteur d'observer la Fig. 3.4. Ce graphique présente deux formes discrètes matérialisées par la distribution de points le long d'un cercle, Fig.3.4-(a) ou d'un tore, Fig. 3.4-(b). Un point seul est une composante connectée et possède donc un nombre de Betti $\beta_0 = 1$. Les nombres de Betti associés à la Fig. ??-(a) sont $(\beta_0 = 1, \beta_1 = 1)$, il y a donc une composante connectée et un trou à 1 dimension. Enfin, les nombres de Betti obtenus pour des points répartis le long du tore de la Figure ??-(b) sont $(\beta_0 = 1, \beta_1 = 2, \beta_2 = 1)$, *i.i.* une composante connectée, tout chemin fermé dessiné sur la surface peut être réduit à un point, deux cercles, les générateurs intrinsèques, et une cavité.

Une explication plus théorique

Considérons S comme un ensemble de points de \mathbb{R}^n et fixons $k \leq n$. Soit pour l'exemple $\sigma = (v_0, v_1, \dots, v_k)$ le k -simplexe constitué par l'enveloppe convexe de ses $(k+1)$ sommets. Alors, un 0-simplexe est un point de S , un 1-simplexe est un segment joignant deux points de S , un 2-simplexe est un triangle rempli, un 3-simplexe est un tétraèdre rempli et ainsi de suite. Les faces d'un k -simplexe correspondent aux k -simplexe donné par $(v_1, \dots, v_{i-1}, v_{i+1}, \dots, v_k)$. La Fig. 3.3 donne une représentation visuelle de ces concepts.

Definition 1 (complexe simplicial) *L'ensemble des k -simplexes disjoints ou se croisant sur une face commune est appelé complexe simplicial et est noté \mathcal{C}_k . Il*

est souvent considéré comme la représentation squelettique (représentations de type maillage) d'un nuage de points de données.

Nous rappelons ici que la question fondamentale soulevée par la théorie de l'homologie est de déterminer les caractéristiques intrinsèques des trous multi-dimensionnels des nuages de points sous-jacents. Il s'agit de déterminer comment ces trous multi-dimensionnels doivent être trouvés lorsque le complexe simplicial est obtenu. Nous devons d'abord définir comment les k -simplexes se composent.

Definition 2 (k -chaîne) *La somme formelle des k -simplexes dans C_k , appelée k -chaîne, est dotée d'une structure de groupe lorsqu'on définit l'addition commutative interne $i + j = (i \cup j) - (i \cap j)$ pour tous les $i, j \in C_k$. Ceci équivaut à une addition dans le champ \mathbb{Z}_2 .*

Afin d'illustrer ceci, si deux 0-simplexes (a) , et (b) sont disjoints, alors $(a) + (b) = (b) + (a)$ alors que $(a) + (a) = (b) + (b) = 0$. Enfin, il reste à savoir quand une k -chaîne participe ou non à la définition d'un trou multi-dimensionnels. Pour ce faire, nous avons besoin de la définition de ce que l'on appelle frontière ou bordure et d'un cycle.

Definition 3 (Frontières ou bordures et cycles) *La frontière d'un simplexe σ dans C_k est la somme formelle des faces (de dimensions $k - 1$) de σ . Elle définit un homomorphisme de groupe $\partial_k : C_k \rightarrow C_{k-1}$. Une k -chaîne u est alors une k -frontière s'il existe $(k + 1)$ -chaîne v telle que $\partial_{k+1}(v) = u$, c'est-à-dire que $v \in \text{Im} \partial_{k+1}$. Or une k -chaîne u est un cycle si $\partial_k(u) = \emptyset$, ce qui signifie que $u \in \text{Ker}(\partial_k)$.*

On peut maintenant définir le k -ième groupe d'homologie comme étant

$$H_k = \text{Ker}(\partial_k) / \text{Im}(\partial_{k+1}). \quad (3.18)$$

Ce sont tous les éléments qui sont des cycles (trous multi-dimensionnels) mais pas des frontières. En utilisant à nouveau l'exemple de la section 3.3.2 mais de manière plus formelle ; on obtient que l'application frontière ∂_2 du triangle $v = (a, b, c)$ est $\partial_2(v) = (ab) + (bc) + (ca)$ et l'application frontière ∂_1 est $\partial_1((ab) + (bc) + (ca)) = (a) + (b) + (b) + (c) + (c) + (a) = \emptyset$. La k -chaîne $u = (ab) + (bc) + (ca)$ est donc un cycle puisque $\partial_1(u) = \emptyset$ et est également une frontière, en effet $\partial_2(v) = u$. Alors $\partial_2 \circ \partial_1 = \emptyset$ et u n'est pas un élément du premier groupe d'homologie H_1 .

L'homologie persistante

Ayant maintenant une meilleure idée des concepts mathématiques de l'homologie, nous abordons la discussion sur l'homologie persistante. L'homologie persistante a été appliquée pour la classification de séries temporelles, [Ume17], la reconnaissance de formes d'images [Car+08], la biologie [TZH15], la phylogénétique [CLR16], la science du langage [Por+18], pour la surveillance du cerveau [Pia+18] et dans divers autres domaines comme [Ott+17]. Comme vu dans les sections précédentes, le groupe d'homologie est obtenu grâce à la comparaison de simplexes à différentes dimensions. Les invariants, les nombres de Betti par exemple, dépendent donc directement de la dimension du complexe simplicial. Si l'on dispose de n points et

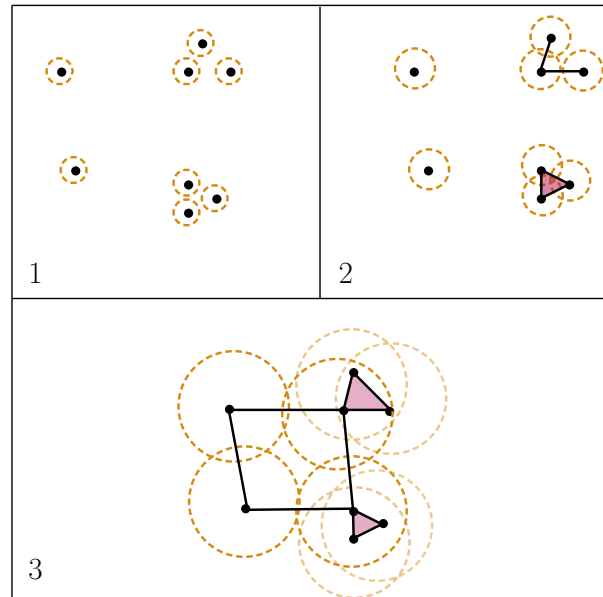


Figure 3.5: Une séquence imbriquée de complexes simpliciaux est créée en augmentant le rayon des boules représentées en orange. Dès que deux boules se croisent, les points en leur centre sont reliés. Lorsque trois boules se croisent, on forme le triangle, et ainsi de suite. Au stade 3, un trou apparaît. Ce n'est pas encore la limite d'un simplexe. Ce trou disparaîtra lorsque ses quatre sommets seront reliés deux à deux.

que l'on choisit de construire un complexe à n -dimensions, certains invariants cachés peuvent ne pas être observés car les cycles seront nécessairement des frontières. A contrario, si nous choisissons de déterminer les invariants avec un nombre $k < n$ de points, des invariants de haute dimension pourraient être manqués. C'est précisément pour cela que l'homologie persistante est utilisée, elle calcule l'homologie à différentes échelles, *i.e.* avec des dimensions de complexes simpliciaux variables, et garde une trace de l'évolution des groupes d'homologie. Cela correspond à construire un complexe filtré avec la séquence $\mathcal{C}_1 \subset \mathcal{C}_2 \cdots \subset \mathcal{C}_n$, avec \mathcal{C}_n le complexe simplicial maximum. Une fois encore, la littérature est extrêmement vaste sur le choix à faire pour le processus de filtration. Cela sort largement du cadre de ce mémoire et nous laissons les lecteurs intéressés explorer davantage la littérature pour les comparaisons et les propriétés de chacun des filtrages. Dans notre cas, nous avons choisi la filtration de Vietoris-Rips pour sa simplicité de mise en œuvre et sa robustesse. Cette méthode consiste à fixer la dimension du complexe simplicial par un nombre ρ . Les simplexes (v_1, \dots, v_k) sont ensuite formés dès que $\|v_i - v_j\| < \rho, \forall i, j$. Cette opération est reproduite avec un ρ croissant jusqu'à atteindre le complexe simplicial maximal contenant toutes les données. Cette méthode est brièvement illustrée dans la Fig. 3.5. Plus de détails sur le calcul peuvent être trouvés dans [ELZ02]; [EM] et [ZC05].

Au cours de la construction de la filtration de Vietoris-Rips, des trous apparaissent et ensuite disparaissent en raison de la dimension croissante du complexe. Ce phénomène est représenté dans un diagramme appelé diagramme de persistance. Il code le moment de la naissance et le moment de la mort du trou multi-dimensionnel. Pour fixer les idées, nous invitons le lecteur à examiner la Fig. 3.6. Ce tracé permet de voir pour trois ρ différents l'évolution du complexe simplicial et nous espérons qu'il

illustre pourquoi certains trous apparaissent à une certaine dimension et disparaissent ensuite. Le diagramme de persistance correspondant est également fourni dans ce graphique. L'exemple jouet étant unidimensionnel, seuls les 0-simplexes et les 1-simplexes sont calculés. Cela donne les trous d'ordre 0 et les trous d'ordre 1 affichés par la Fig. 3.6. Le résultat le plus important donné par cette figure est l'existence d'un point solitaire aux coordonnées $(0.45, 1.2)$. Il matérialise pour $\rho = 0.45$ la naissance d'un trou à 1 dimension, et la mort de ce trou pour $\rho = 1.2$, ces faits sont explicitement montrés par la Fig. 3.6.

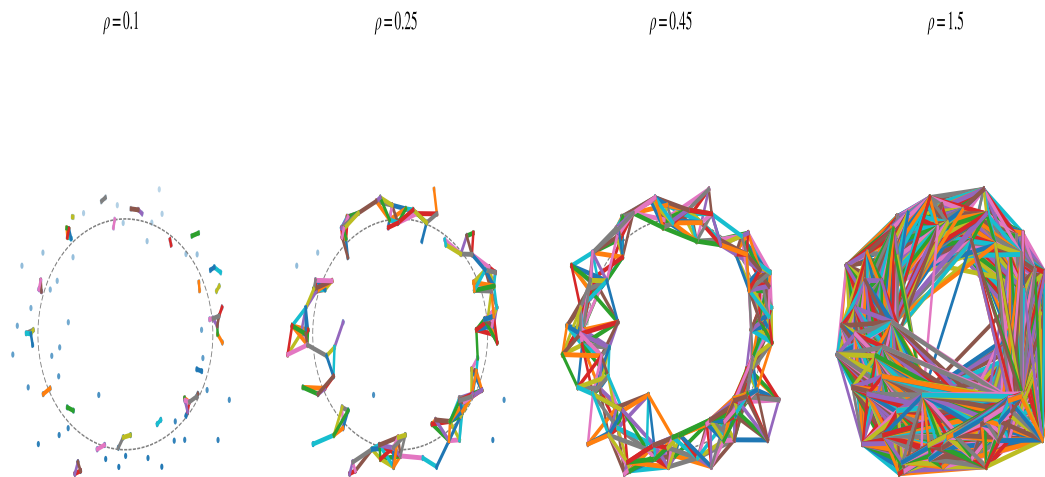


Figure 3.6: De gauche à droite, une séquence croissante de complexes simpliciaux pour des points distribués le long d'un cercle. Le nombre ρ est la distance pour la filtration de Vietoris-Rips.

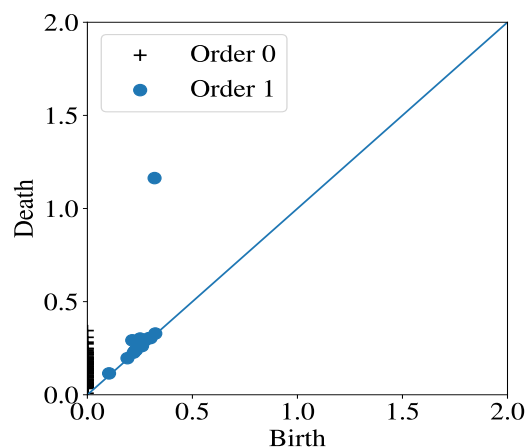


Figure 3.7: Diagramme de persistance du cercle bruité de la Fig. 3.6, les axes correspondent aux étapes ρ de la filtration.

Nous souhaitons caractériser topologiquement les formes obtenues par les matrices de dilatation sur $SO(n)$. Afin de calculer l'homologie persistante sur ce

groupe de Lie avec la filtration de Vietoris-Rips, nous avons utilisé la métrique $d(A, B) = \|\log(A^T B)\|$, où $\|\cdot\|$ désigne la norme de Frobenius classique [Bou13c]. Ces explications théoriques faites, nous proposons maintenant d'utiliser l'homologie persistante et le diagramme de persistance sous-jacent. Nous cherchons ainsi à montrer que selon leur non-stationnarité, les processus PC ont des caractéristiques topologiques intrinsèques.

Je vais maintenant présenter les différents résultats que nous avons obtenus en appliquant l'homologie persistante sur les trajectoires de matrices de dilatation. Tous les résultats et analyses proposés se basent sur les cinq processus définies en 3.2. Nous avons affiché quelques diagrammes de persistance dans [BB19] afin de guider notre réflexion. Les résultats obtenus nous indiquent, au travers de quelques représentations seulement par processus, une certaine stabilité dans l'information topologique sur $SO(3)$. Sur cette représentation, nous montrons qu'il n'est pas nécessaire de rechercher les diagrammes de persistance de second ordre et concluons que dans $SO(3)$ la trajectoire des matrices de dilatation appartient à un plan, ces trajectoires sont donc orientées au sens des matrices de rotation. Les quelques diagrammes pris au hasard, nous confirment également que les classes de processus ont des formes précises. Fort de ces résultats, nous avons lancé 500 réalisations de Monté-Carlo de chacun des processus et moyenné les résultats.

Information donnée par les nombres de Betti

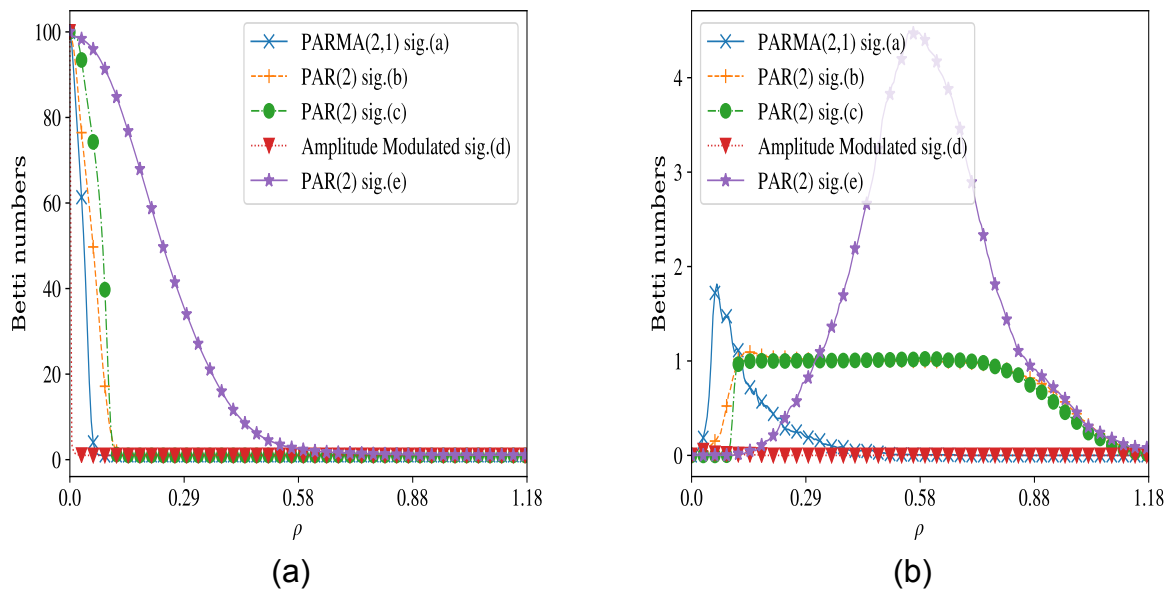


Figure 3.8: Séquences de nombres de Betti moyennés associées à (a) un diagramme persistant d'ordre 0, (b) un diagramme persistant d'ordre 1.

De manière similaire à l'approche de [Giu+15], nous avons analysé la séquence des nombres de Betti. Nous présentons les résultats dans la Fig. 3.8, pour le diagramme d'ordre 0 Fig. 3.8-(a) et le diagramme d'ordre 1 Fig. 3.8-(b). Les différentes courbes proposées sont les courbes moyennes sur les 500 réalisations des

Model	β_0	β_1	β_2
PARMA(2,1)	≥ 1	2	0
PAR(2), $T = 20$	≥ 1	1	0
PAR(2), $T = 54$	≥ 4	4	0
Amplitude Modulated	1	0	0

Table 3.2: Les nombres de Betti en fonction de la classe des processus.

séquences de Betti. La Figure 3.8-(b), qui est une réelle confirmation et un résultat jamais présenté dans la littérature, est pour moi exceptionnelle. Elle montre tout ce dont nous avons besoin pour conclure sur les caractéristiques topologiques. Nous avons un processus presque stationnaire, le *sig.(d)* qui n'a par conséquent aucune boucle dans sa représentation sur $SO(3)$. Nous avons deux modèles PAR(2) avec des coefficients générateurs différents, ce sont les *sig.(b)* et *sig.(c)* qui admettent la même séquence de nombres de Betti moyennés. Nous avons un modèle PAR(2) avec les mêmes coefficients générateurs que celui de *sig.(b)* mais avec une période $T = 54$ au lieu de $T = 20$, c'est le *sig.(e)*. Ce processus présente le plus grand nombre de boucles dans sa représentation et enfin, nous avons le *sig.(a)* qui est un modèle PARMA(2,1). Sa séquence de nombres de Betti moyennés est sans aucun doute différente des autres séquences. Il n'y a donc aucune ambiguïté pour discerner les classes de processus, et même plus, si les coefficients générateurs sont différents, la séquence des nombres de Betti moyennés dit qu'il s'agit de la même classe. Enfin, et pour finir avec cette partie, l'analyse de la Fig. 3.8-(a) montre le nombre de composantes connectées. Comme on peut le voir, le *sig.(e)* montre un grand nombre de composantes connectées, ceci est lié à la forte dispersion de ses points. Pour tous les autres, le nombre de composantes connectées, qui est certainement supérieur à un, est le même ; à l'exception encore une fois du *sig.(d)* pour lequel il n'y a qu'une seule composante connectée. Pour résumer ces résultats, j'ai proposé une table empirique des nombres de Betti pour chacun des processus. Il s'agit de la Table 3.2.

Information donnée par l'entropie persistante

Si les résultats précédents sont probants, ces résultats sont issus de 500 réalisations moyennées. Nous nous sommes donc posé la question de l'homogénéité des diagrammes de persistance sur lesquels sont basés les analyses. Toujours dans la même optique de savoir si une information intrinsèque à la classe de processus était disponible, l'idée a été de savoir si les mesures spectrales et les matrices de dilatation étaient contraintes dans leur topologie. Un moyen simple pour y parvenir a été d'implémenter l'indicateur d'homologie persistante normalisée définie par

$$H(p) = -\frac{1}{\log(N_p)} \sum_{j=1}^{N_p} p_j \log(p_j) \quad (3.19)$$

où $p_j = \frac{l_j}{L}$ avec $l_j = y_j - x_j$ (la durée de vie), $L = \sum_{j=1}^{N_p} l_j$, et N_p le nombre de points dans le diagramme de persistance, et de regarder ensuite la distribution de cette entropie sur l'ensemble de tous les diagrammes. L'idée n'est pas tant d'analyser la valeur de l'entropie persistante normalisée par diagramme, qui est difficile d'interprétation mais d'analyser la distribution de cette entropie. Plus la distri-

bution tendra vers une distribution de Dirac et plus les diagrammes de persistance seront homogènes. Cela sous-entend que la structure topologique des trajectoires est récurrente et montre ainsi l'existence d'une contrainte structurelle liée à la classe du processus.

Voyons maintenant la distribution de l'entropie persistante normalisée pour les processus étudiés. Nous avons représenté ces distributions sur la Fig. 3.9. Comme d'habitude jusqu'à présent, la Fig. 3.9-(a) est associée au diagramme de persistance d'ordre 0, et la Fig. 3.9-(b) au diagramme d'ordre 1. Point assez inattendu, les distributions de l'entropie persistante normalisée sont toutes disjointes pour les diagrammes de persistance d'ordre 0. Points positifs sur le côté séparable des processus, mais aussi sur le fait que chacune des matrices de dilatation semblent être contraintes dans leur distance relative (connectivité). L'analyse de la distribution de l'entropie persistante normalisée associée au diagramme d'ordre 1 montre le fait que les $sig.(b)$ et $sig.(c)$ ont un faible nombre de points dans le diagramme et que la boucle est clairement identifiée par un seul point avec une durée de vie élevée. Cela ne semble pas être le cas pour le $sig.(e)$, pour lequel il y a certainement beaucoup de points dans le diagramme avec plus de variabilité. Le $sig.(d)$ a presque la même distribution d'entropie persistante normalisée, ce qui révèle sa grande volatilité et la difficulté d'estimer une trou ou une boucle dans la trajectoire. Ce que nous concluons d'après ces résultats c'est que l'ensemble des matrices de dilatation suivent une évolution (une rotation) contrainte par le processus. Plus il y a de matrices et de boucles, plus la topologie admet une certaine variabilité, sans doute un léger morphisme de la trajectoire.

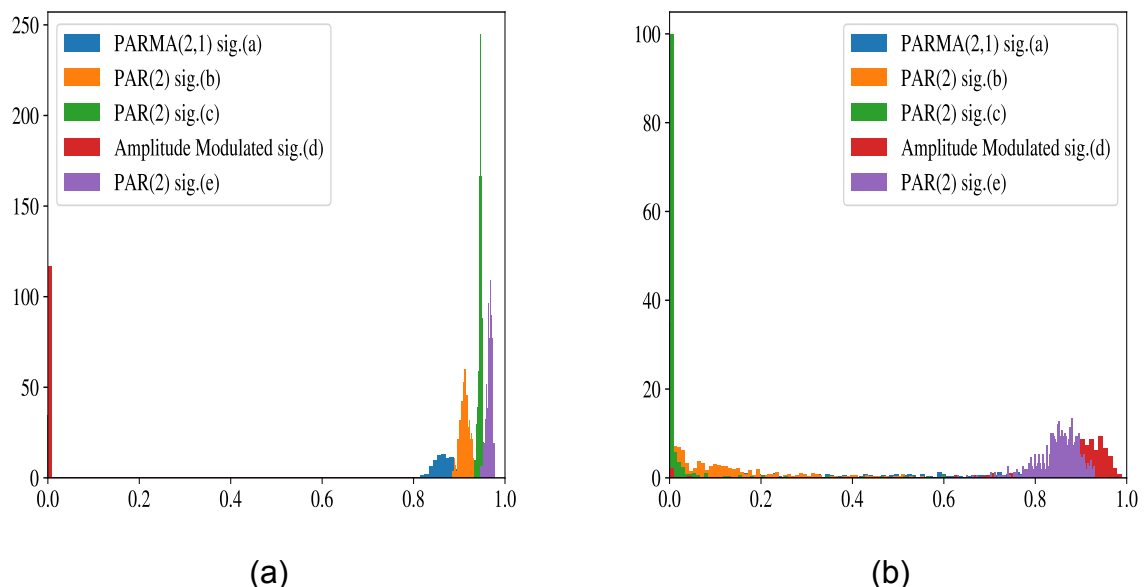
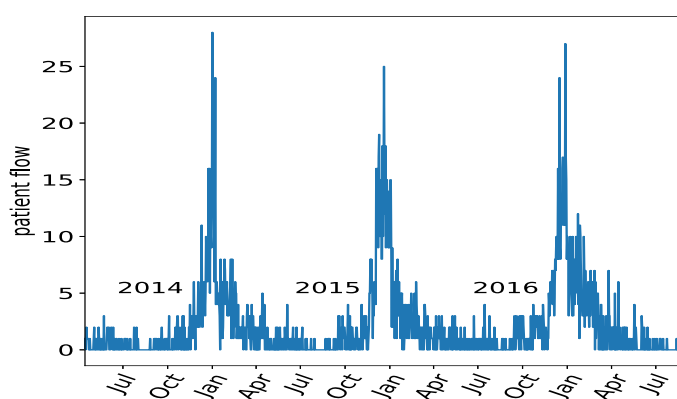


Figure 3.9: Histogrammes de l'entropie persistante normalisée associés à (a) un diagramme persistant d'ordre 0 et (b) un diagramme persistant d'ordre 1.

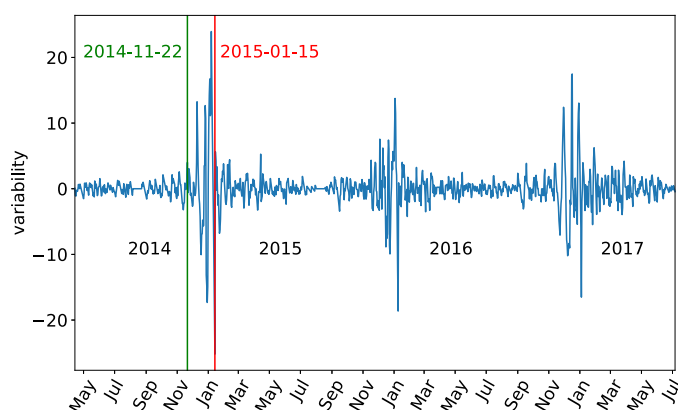
3.4. Mise en place des indicateurs

3.4.1. L'indicateur précoce basé sur la topologie

Une fois ces considérations théoriques sur les classes de processus périodiquement corrélés faites, je suis me tourné vers le coté applicatif. L'objectif étant d'améliorer la détection précoce de venues massives aux urgences pédiatriques afin d'améliorer la mise en place du plan hivernal définie par l'ARS, nous avons cherché comment la description de la non-stationnarité périodique par les matrices de dilatation pouvaient nous y amener. L'idée a été de trouver la distorsion de la non-stationnarité périodique du flux de patients à appliquer afin que le départ à la non-stationnarité soit le plus marqué. Cela a fait sujet de la publication [Dug+19]. Un nouveau détecteur est ainsi né et basé sur les étapes décrites ci-après.



(a)



(b)

Figure 3.10: En haut, flux d'admissions pour les patients diagnostiqués bronchiolite pendant 3 années. En bas représentation d'un exemple de la variabilité des admissions dont la valeur moyenne est nulle.

La distorsion du flux de patients

Distordre la non-stationnarité périodique du flux de patients a été possible grâce à une technique employée pour l'analyse auto-similaire des processus, la Detrended Fluctuation Analysis (DFA) [Dug+19]; [Emi+18]. Cette technique nous permet d'extraire la non-stationnarité périodique et d'en contrôler la distorsion, *i.e.* l'amplitude, et la durée des fortes variabilités. Un exemple est donné par la Figure 3.10 où figure en haut un flux d'admissions aux urgences pédiatriques pour symptômes de bronchiolite et en bas l'extraction de la non-stationnarité périodique.

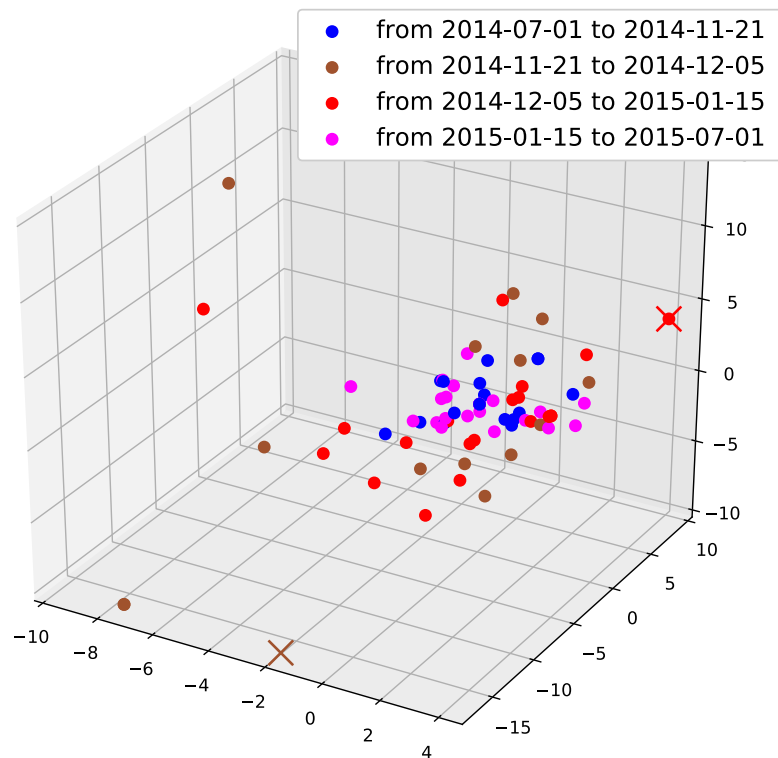


Figure 3.11: Trajectoire échantillonnée dans \mathbb{R}^3 de la non-stationnarité périodique de la Figure 3.10.

Pour créer le nouveau détecteur nous approximons la trajectoire dynamique de l'extraction de la non-stationnarité par une intégration de type Hankel [BK86]; [PBC19]; [PH15]. Chacune des colonnes de la matrice ainsi obtenue est alors associée à un point dans un espace euclidien. La collection des points forme le nuage de points, représentatifs de la trajectoire. Pour une fenêtre glissante de 3 échantillons, nous avons tracé la trajectoire échantillonnée obtenue (Fig. 3.11). Sur ce graphique, nous avons coloré quatre périodes pour mettre en évidence des faits majeurs. En particulier, les périodes du 2014-07-01 au 2014-11-21 et du 2015-01-15 au 2015-07-01 présentent des points ayant le même comportement. Ils sont concen-

trés à l'intérieur d'une boule et semblent tous être presque également espacés. Ceci est vraiment intéressant car ces deux périodes sont associées à des faibles arrivées aux urgences et il n'y a pas d'épidémie de VRS à cette époque. Au contraire, pour la période du 2014-11-21 au 2014-12-05, les points ne sont plus également espacés mais ils s'étalent dans l'espace. Cette période représente le début de l'épidémie de VRS et correspond alors à des arrivées anormalement élevées aux urgences. Pour la dernière période, qui a débuté le 2014-12-05 et s'est terminée le 2015-01-15, certains des points ont des coordonnées élevées et d'autres sont concentrés à l'intérieur d'une boule aux coordonnées faibles. Cette période correspond à la diminution de l'épidémie et indique donc une diminution des admissions aux urgences. De cette représentation nous pouvons donc envisager la détection du début de l'épidémie.

Pour ce faire, nous avons en temps réel caractérisé la topologie de ce nuage de points par l'homologie persistante et déterminer un indicateur principal

$$\text{Bar}_{\max}(k) = \max_{\mathcal{D}_k} \{l_i\}, i = 1, \dots, N_{p_k} \quad (3.20)$$

$$(3.21)$$

où $l_j = y_j - x_j$ représente toujours la durée de vie dans le diagramme de persistance et N_{p_k} le nombre de points du nuage à l'instant temporel d'évaluation k . L'ensemble $\mathcal{D}_k = \{(x_i, y_i)\}_{i=1}^{N_{p_k}}$ représente le diagramme de persistance calculé à l'instant k . Nous avons appliqué cet indicateur sur le flux de patients des urgences et confronté ses performances par rapport à quelques méthodes largement plébiscitées dans le domaine de la production de soins ou de biens manufacturiers. La Figure 3.12 ci-après permet d'apprécier les très bonnes performances de l'indicateur que nous proposons relativement aux performances des autres méthodes.

Le gain apporté par ce nouvel indicateur est pluriel. D'une part il a une plus grande précocité que l'indicateur initialement proposé (Détection par le Kurtosis) et d'autre part il fournit une alarme unique ne laissant aucune ambiguïté dans la prise de décision.

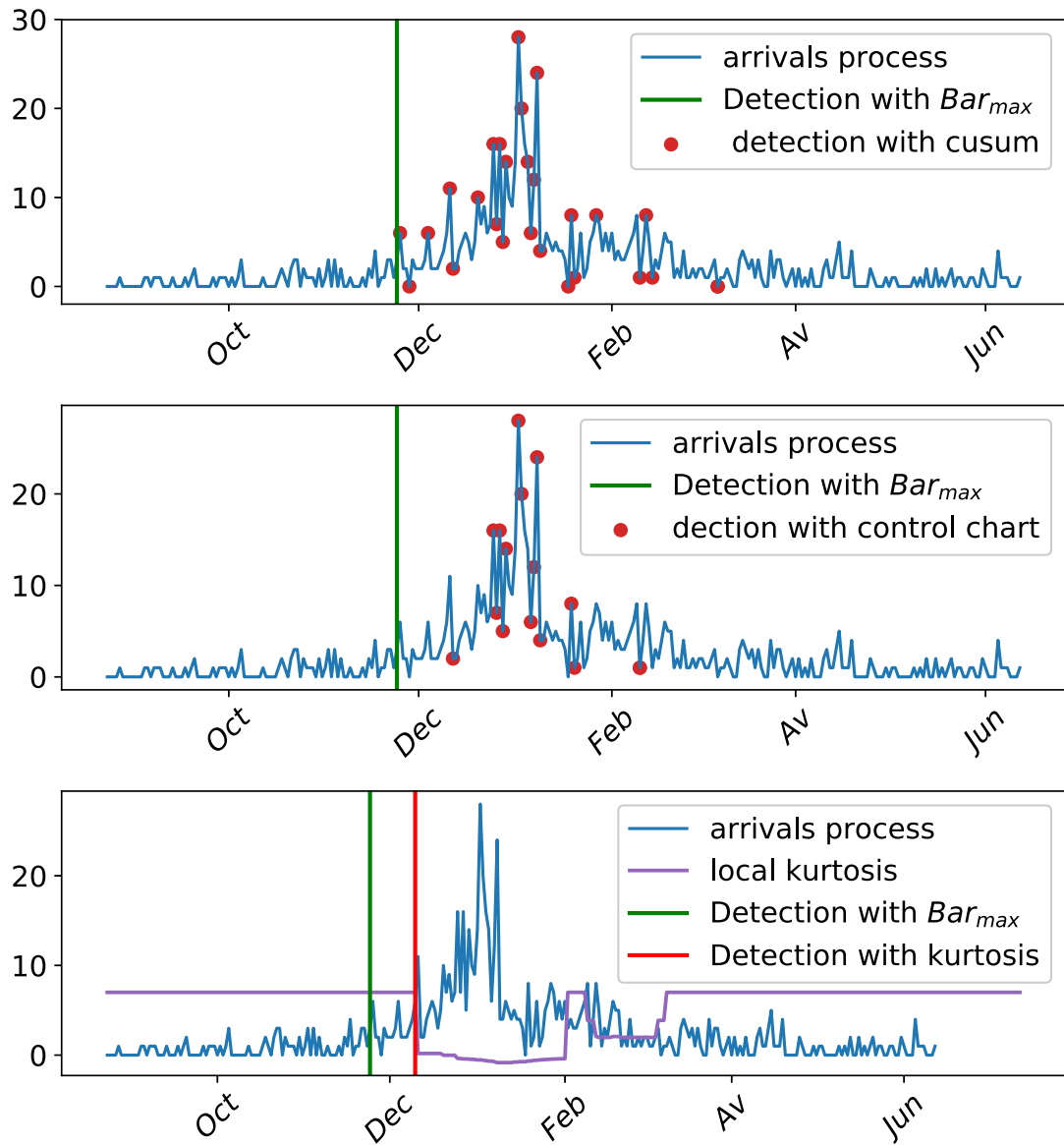


Figure 3.12: En haut, comparaison entre l'indicateur $Bar_{max}(k)$ et la détection des points de changement CUSUM. Au milieu, comparaison entre l'indicateur $Bar_{max}(k)$ et la détection des points de changement par carte de contrôle (Shewhart). En bas, comparaison entre l'indicateur $Bar_{max}(k)$ et la méthode du kurtosis local.

3.4.2. L'application web

L'ensemble des outils nécessaires au suivi et à la définition des alertes a été implémenté sur un site web hébergé par le CHU de Saint Etienne à l'adresse <http://predaflu.chu-st-etienne.fr> en 2019. Ce site récupère chaque jour les données nécessaires à l'affichage des flux de patients. Les indicateurs mis en œuvre sont ceux publiés dans les revues internationales [BMM15]; [Dug+19].

Les médecins autorisés à consulter le dashboard de l'application, ont la vision de leur flux en temps réel (journée) ainsi que trois niveaux d'alertes, figure 3.13 et figure 3.14. Une alerte très précoce (50% de confiance, bleu clair sur le graphique), une alerte moyennement précoce (80% de confiance, bleu médium sur le graphique) ainsi qu'une alerte épidémique (100% de confiance, bleu foncé sur le graphique) marquant un flux d'admissions aux urgences pédiatriques trop anormal. Les deux premières alertes sont construites par l'homologie persistante de la non-stationnarité périodique du flux de patients alors que l'alerte à 100% est donnée par l'analyse du Kurtosis de l'erreur de prédiction du flux. Une courbe de prévision de l'activité est également proposée et permet une gestion long terme de l'unité, elle correspond au modèle de Serfling.

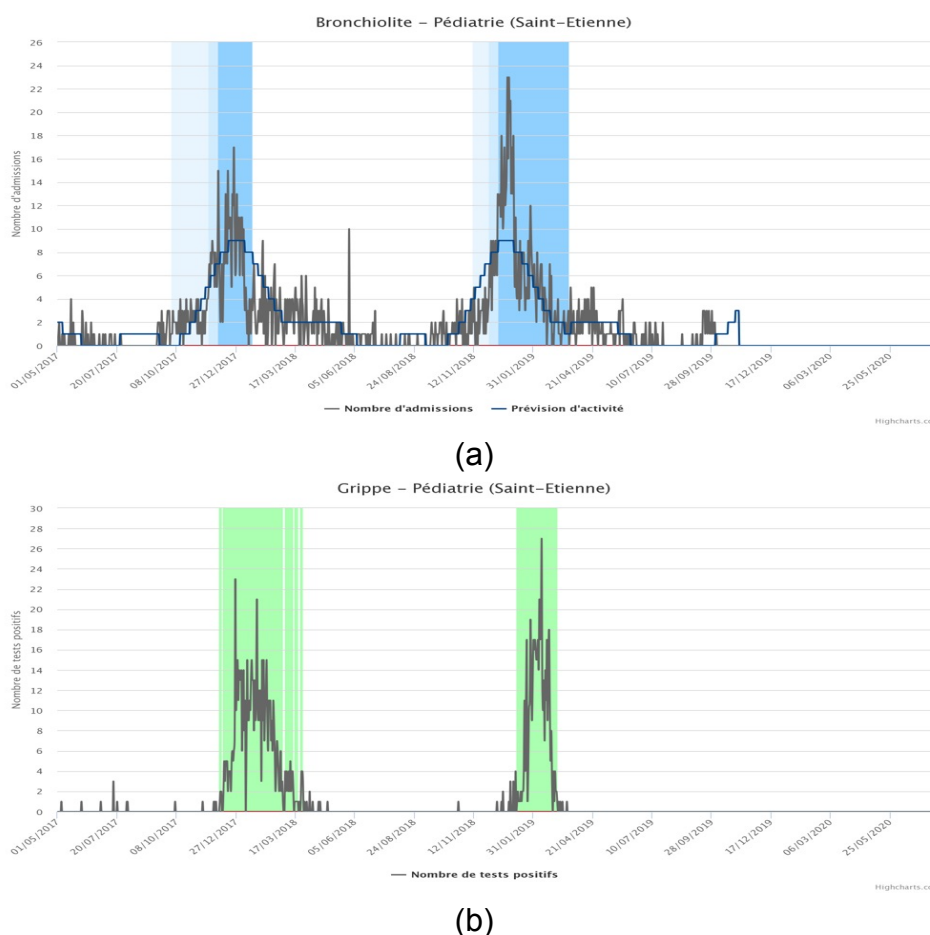


Figure 3.13: Images du tableau de bord de l'application PREDAFLU. En haut, flux quotidien des patients diagnostiqués bronchiolite, avec la prédiction, et les 3 niveaux d'alertes. En bas, tests quotidiens positifs à l'Influenza (grippe) avec la détection en vert.

Un graphique de présence de la grippe par analyse virologique est également proposé figure 3.13-(b). Ce graphique dépend des centres d'études, ceux ne disposant pas de laboratoire virologique voient leur courbe remplacée par le flux de patients enfants diagnostiqués grippe. Ce flux est ensuite ajusté en fonction des informations virologiques. Pour les établissements ayant un laboratoire et/ou ceux pratiquant les tests de diagnostic rapides, les données sont affichées. En temps réel, les médecins peuvent juger de la présence du virus grippal dans leur établissement, un indicateur épidémique est également fourni, des lors qu'au moins 2 virus Influenza pendant 3 jours sont détectés.

Enfin, un graphique du flux des entrées aux urgences adultes pour cause de grippe est également proposé. Ce dernier graphique trace le flux de patients âgés (+60 ans) des urgences adultes correspondant aux patients dont les symptômes sont les plus proches de ceux du virus grippal. Le sens de « au plus proche » est défini par la thèse de Clément PEALAT et sera donc abordé dans le prochain chapitre.

La DSM moyenne journalière du flux de patients est également affichées, elle permet déjà d'avoir l'indication aux médecins urgentistes de l'encombrement de leurs services. Grâce aux niveaux de coloration correspondant à la présence du VRS (bleu), à la présence de la grippe (vert) et à la présence conjointe des deux virus (rouge), les médecins du CHU de saint Etienne ont pu noter que la DMS des personnes âgées avec certaines pathologies cardio-respiratoires commençait à augmenter pendant l'activité du VRS chez l'enfant et pas nécessairement lorsque le virus Grip-pal est avéré. Les recommandations et préconisation d'organisation seront donc déclenchées préventivement afin de palier à cette problématique d'encombrement des urgences adultes.

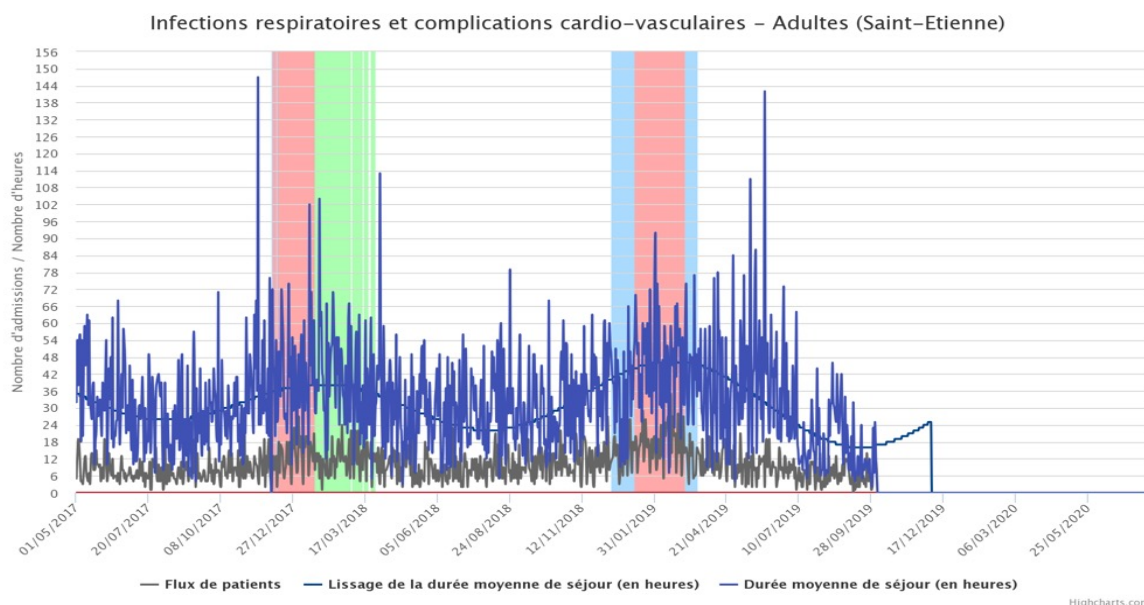


Figure 3.14: Représentation du flux quotidien de patients adultes de plus de 60 ans pour les pathologies déterminées par les algorithmes de clustering du chapitre 4. Pour ces patients, est affichée également la durée de séjour ainsi qu'un lissage de cette durée.

3.4.3. Retour d'expérience

BRONCHIOLITE ET GRIPPE : QUAND LE CHU DE SAINT-ÉTIENNE ANTICIPE SUR L'AFFLUENCE AUX URGENCES

RECHERCHE & INNOVATION

Afin d'anticiper les épidémies hivernales, le Dr Olivier Mory, chef du service d'Urgences pédiatriques, travaille sur un programme de recherche dénommé PREDAFLU, visant à construire un outil de prévision des pics d'affluence dans les services d'Urgences au CHU de Saint-Étienne. Grâce à l'aide des mathématiques, il a mis en évidence un lien fort entre les épidémies de bronchiolites du nourrisson et l'arrivée de personnes âgées en décompensation cardio-respiratoire aux Urgences adultes. Le projet a reçu le soutien du ministère de la Santé grâce à un financement PREPS.



Figure 3.15: Article du CHUMAG du CHU de Saint-Etienne.

Aujourd'hui et depuis sa mise en place, la direction du CHU active le plan hivernal en pédiatrie dès lors que la détection est faite (80% de confiance). Le pilotage du service en est donc amélioré et depuis, les urgences pédiatriques font face aux affluences massives. Ces bonnes pratiques organisationnelles ont trouvé écho principalement auprès du CHU de Grenoble et du CHU de Genève qui utilisent également en routine l'application web PREDAFLU. Les CH de Valence et Villefranche-sur-Saône, sont des centres moins réactifs qui ne pilotent pas complètement leur service grâce à l'outil, mais participent néanmoins à l'expérimentation.

Seuls les CHU de Saint-Etienne et de Grenoble ont un regard sur le flux de patients adultes. L'article [Mor+22] positionne l'intérêt de l'outil pour prévenir précocement les venues pour détresse cardio-respiratoire aux urgences adultes mais il reste encore un bout de chemin pour que les urgences adultes pilotent totalement le service grâce à cette application

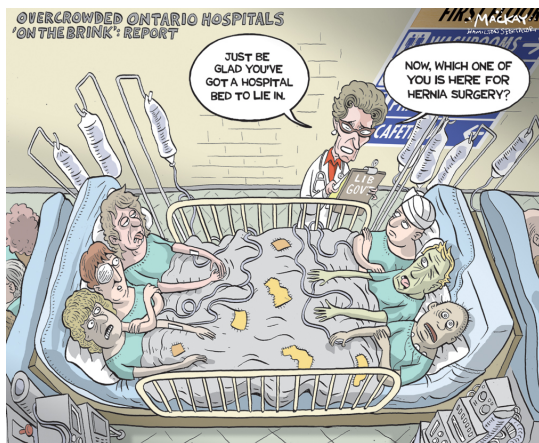
3.5. Contributions scientifiques de cette partie

Bilan scientifique



- 4 Articles de revues [BMM15]; [DBM18b]; [Dug+19]; [BDM19a]
- 1 Article de conférence Internationale à comité et acte de lecture [DBM18a]

Anticiper l'engorgement aux urgences adultes en déterminant le flux de patients le plus impactant



Nous l'avons vu dans le chapitre précédent, l'application PREDAFLU disponible pour les services d'accueil des urgences contient un graphique relatif aux urgences adultes. Les virus respiratoires hivernaux comme le VRS et l'Influenza, impactent sévèrement les services d'urgences d'adultes, beaucoup plus encore que les urgences pédiatriques. Chez l'adulte, ces virus atteignent plus

généralement une population âgée (+60 ans) avec des complications cardio-respiratoires. Toutefois, l'engorgement aux urgences adultes pendant la période de circulation de ces virus, peut également être provoqué par d'autres pathologies. Dans ce chapitre, je vais donc développer la manière dont nous avons extrait les codes diagnostics relatifs à la circulation des virus respiratoires et je m'attarderai sur les préconisations méthodologiques à employer.

Sommaire

4.1	Contexte et objectif	54
4.2	Le clustering de séries temporelles	55
4.2.1	Trajectoires	56
4.2.2	Les variétés testées	56
4.2.3	Les distances et calcul de barycentres sur les variétés	58
4.2.4	Le benchmark	60
4.3	Application aux flux de patients adultes	64
4.4	Contributions scientifiques de cette partie	64

4.1. Contexte et objectif

Même si les diagnostics cliniques, conséquence d'un virus respiratoire sont relativement bien documentés dans littérature médicale, l'engorgement dans les services d'accueil des urgences peut être dû à l'adjonction d'autres pathologies ou d'autres virus moins recherchés. Il est également bien documenté [SS13] que les virus respiratoires hivernaux (ou du moins jusqu'en 2020 ... et anté-COVID-19) impactent fortement le service d'accueil des urgences adultes avec une augmentation très nette de la durée moyenne de séjour pour une partie de la population. A la différence de l'enfant, la charge virale chez l'adulte est moindre ou plus délicate à détecter par de simples tests virologiques. De plus, les tests virologiques systématiques ont un coût important que l'hôpital ne peut prendre en charge. Seuls les patients hospitalisés sont testés afin d'éviter d'éventuelles contagions. Enfin, les laboratoires rendent leurs résultats disponibles de manière hebdomadaire ne permettant pas réellement d'avoir un monitoring fin. L'image de la présence virale donnée par la virologie n'est donc pas une image exacte de la présence des virus et ne permet pas d'identifier les patients dont les diagnostics peuvent correspondre à une conséquence de l'infection virale. Tout comme l'outil PREDAFLU, destiné de prime abords aux urgences pédiatriques, il est nécessaire de baser le pilotage des unités adultes ou à minima de fournir un dashboard précis de la situation, à partir des données d'entrées aux urgences. Il faut pouvoir extraire de l'ensemble des données d'entrées, le flux de patient impactant le plus l'organisation des urgences adultes lors de la présence des virus hivernaux.

Afin d'être précoce dans la détection de venues anormales aux urgences adultes, une hypothèse soutenue par les urgentistes pédiatriques est que l'infection virale chez l'enfant et notamment celle du VRS, intervient en amont des contaminations chez l'adulte. Or la détection virale chez l'enfant est bien plus simple et moins coûteuse. À partir des données d'entrées, nous disposons d'une base de multiples séries temporelles associées chacune à un code diagnostic. Lorsqu'un patient arrive, le médecin référence dans son dossier médical généralement un diagnostic principal et deux ou trois autres diagnostics secondaires. De même, nous disposons de séries temporelles associées à la détection du virus chez un enfant testé.

L'idée revient alors à utiliser l'information virale obtenue par test chez l'enfant comme une information non négligeable à utiliser pour l'adulte. Parmi toutes les approches possibles, j'ai choisi l'approche par clustering des séries temporelles qui permet facilement de se servir des séries temporelles des virus comme attracteur pour les séries temporelles associées aux codes diagnostics des adultes.

Les enjeux soulevés par cette approche se concentrent sur le choix de l'espace de représentation des séries temporelles et la modification des méthodes de clustering pour s'adapter à l'espace de représentation, *i.e.* la caractérisation des distances à utiliser sur ces espaces pour le clustering. La Figure 4.1 illustre l'idée générale. Une fois les codes diagnostics obtenus, il ne reste plus qu'à construire le flux de patients associé et à le monitorer.

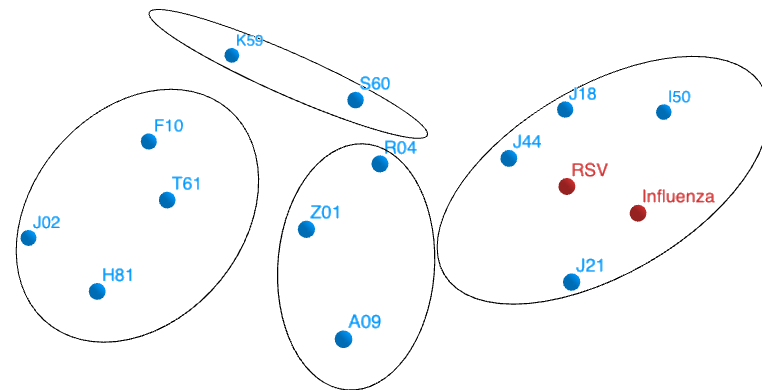


Figure 4.1: Illustration du principe de clustering des codes diagnostics CIM-10 lorsqu'ils sont représentés dans un espace de plus haute dimension.

4.2. Le clustering de séries temporelles

Généralement et le plus souvent [JLR20], le clustering de séries temporelles (univariées) s'applique directement sur les données brutes [KP10] avec comme distance la Dynamic Time Warping [OFC99] ou simplement la distance euclidienne. Bon nombre d'article de review des méthodes de clustering ne traitent pas d'autres approches et notamment pas celles se basant sur un aspect plus géométrique. Étant donné mes affinités mathématiques, j'ai rapidement choisi de représenter les séries temporelles sur un autre espace que celui de l'espace euclidien. De manière assez naturelle finalement, je me suis tout d'abord indirectement tourné vers le « *manifold learning* » en représentant les séries temporelles par un projecteur car conscient de l'efficacité accrue de ce type d'approches [LTS07]; [Tur+11] et guidé par la méthodologie présentée dans [OMN17]. C'est une démarche classique lorsqu'on traite les algorithmes d'estimation en sous-espace ou à haute-résolution. La méthodologie est simple et pour rappel, on construit la matrice de Hankel, on applique généralement une réduction de dimension de type PCA ou Kernel-PCA puis on recherche une base vectorielle. Cela a conduit aux premiers résultats publiés dans [Sol+18]; [PBC19]. Résultats encourageants mais qui nécessitaient d'aller un cran plus loin dans la compréhension de l'espace de représentation et de son influence pour le clustering.

En fait, le processus de clustering dans sa globalité est plutôt compliqué, le lecteur pourra en juger au fil du chapitre. Il commence par l'injection de la série temporelle univariée dans un espace ambiant de plus haute dimension. Il y a là en fait deux points importants qui ont concentré une bonne partie des efforts de recherche. Le premier réside dans la construction de la matrice de Hankel, qui n'est pas anodine alors que le second réside dans la projection de cette matrice dans un nouvel espace qui favorisera le clustering.

4.2.1. Trajectoires

Comme énoncé précédemment, la construction de la matrice de Hankel était pour moi une construction naturelle qui n'avait pas forcément d'intérêt à re-questionner. La matrice de Hankel est fortement employée pour approximer la covariance de la série à étudier et permet d'avoir une vision statistique de la série. Mais on fouillant un peu plus à partir de [BK86], un nouveau mot clé est apparu, celui de **Takens' Theorem**. Je suis toujours pantois de constater qu'un mot est parfois une clé pour ouvrir des portes jusqu'alors cachées ou fermées. Ce théorème de Takens' nous a amené à considérer différemment la série temporelle. Non plus la voir comme un élément d'un espace de Hilbert mais comme une réalisation d'un système dynamique sous-jacent. Et de voir que la construction de la matrice de Hankel est un moyen de représenter une trajectoire (convergeant vers un attracteur) dans l'espace des phases du système en question. Finalement, la construction de Hankel est généralisée en paramétrisant par un paramètre τ l'overlapping des fenêtres glissantes. Cela donne la matrice trajectoire T_y d'une série temporelle $y(t)$ [Sta+97]

$$T_y = \begin{bmatrix} y(0) & y(\tau) & \dots & y((m-1)\tau) \\ y(1) & y(2+\tau) & & \vdots \\ \vdots & \vdots & & \vdots \\ y(k) & y(k+\tau) & \dots & y(k+(m-1)\tau) \\ \vdots & \vdots & & \vdots \\ y(l-1-(m-1)\tau) & y(l-1-(m-2)\tau) & \dots & y(l-1) \end{bmatrix}. \quad (4.1)$$

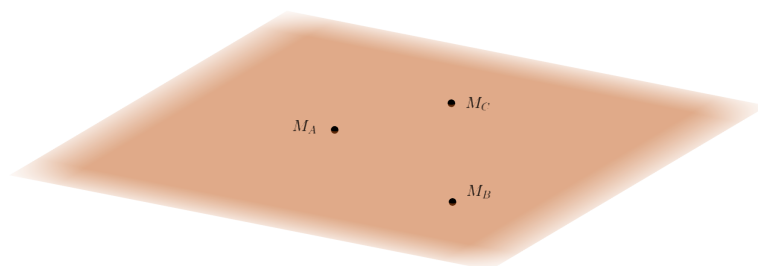
Pour une construction de type Hankel classique $\tau = 1$.

La valeur du paramètre τ , est bien évidemment essentiel et intervient fortement dans les résultats du clustering. Un article de journal est en cours de finalisation sur ce sujet, je débâterai des conséquences de cette valeur dans la partie 4.2.4 de ce chapitre.

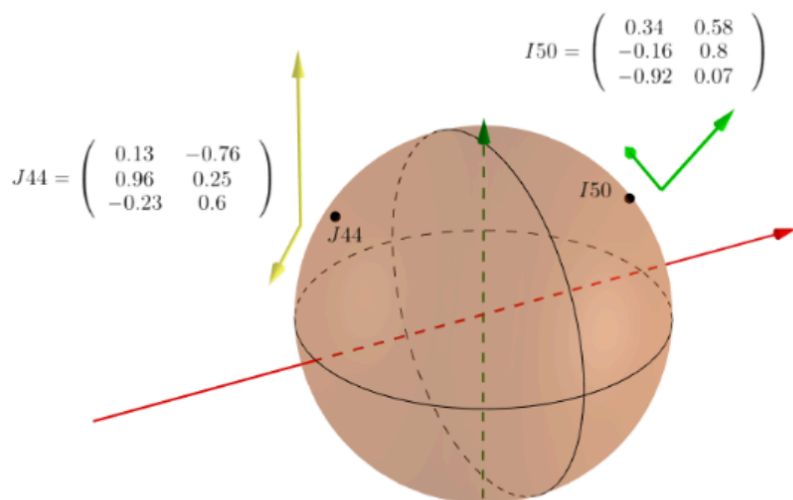
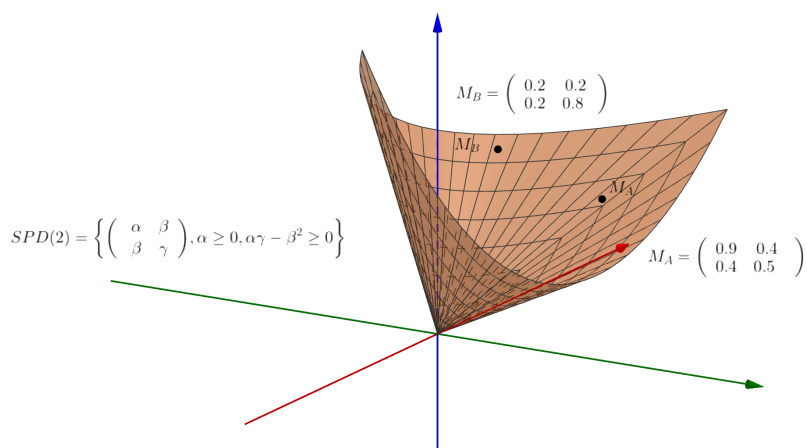
Si on considère que la matrice trajectoire T_y est un élément d'un espace ou d'un ensemble, il y a de forte chance pour que cet ensemble soit une variété différentielle. Or, tel quel, il est difficile de le montrer ou même de disposer ensuite d'une métrique qui sera nécessaire pour le clustering. Nous avons donc fait le choix de projeter la matrice T_y sur des ensembles connus, qui sont des variétés et donc les métriques sont bien définies.

4.2.2. Les variétés testées

Parmi les variétés différentielles disponibles, nous avons fait le choix de projeter sur la variété de Stiefel, la variété Sphère ainsi que sur la variété Euclidienne classique. Nous avons testé la variété des matrices définies positives au travers de la métrique de Gramm, mais pour des raisons de calculs nous n'avons pas pu appliquer cet projection au benchmark réalisé à la section 4.2.4. La Figure 4.2 illustre les différences entre les courbures des variétés de projection, et mettent clairement en évidence les différentes métriques sur ces variétés.



Flat Space, Euclidean geometry between matrices $\{M_i\}_{i=A,B,C}$



$$V_2(\mathbb{R}^3) = \{A \in \mathbb{R}^{3 \times 2} : A^T A = Id\}$$

Figure 4.2: Illustration de trois géométries de variétés différentes. En haut, la variété euclidienne. Au milieu, la variété des matrices semi-définies positives et enfin en bas la variété de Stiefel.

Ce choix de projection se justifie par les approches possibles vues précédemment. Une comparaison des séries temporelles sur une variété à courbure nulle de type euclidienne, une comparaison des séries temporelles par les statistiques du second ordre et la variété des matrices semi-définies positives et enfin une comparaison des séries temporelles par leur projecteurs associé et la variété de Stiefel des matrices orthogonales. La variété Sphère étant une comparaison des séries temporelles lorsque les dimensions des projecteurs à comparer sont différentes, elle est induite par la variété de Stiefel donc.

4.2.3. Les distances et calcul de barycentres sur les variétés

Nous n'avons pas proposé de nouvelle métrique sur les variétés considérées mais nous proposons un travail de synthèse qui regroupe clairement les expressions des distances ainsi que les approximations des exponentielles de Riemann ou Logarithme de Riemann à utiliser.

Parmi les algorithmes de clustering les plus performants et utilisés, nous avons choisi comme algorithme de comparaison l'algorithme de K-means. Si la recherche d'un barycentre est plutôt aisée dans le cas euclidien, elle est nettement moins intuitive lorsque l'on traite les déplacements sur des variétés. Il existe néanmoins un algorithme qui permet la recherche itérative d'un barycentre sous ces conditions, c'est l'algorithme de « *Karcher Mean* ». Il s'agit d'une descente de gradient sur la variété, où le point courant est injecté sur l'espace tangent. Le déplacement de ce point est ensuite calculé en moyennant l'ensemble des vitesses de déplacements aux points dont on cherche le barycentre. Le nouveau point est ensuite ramené sur la variété par la carte exponentielle. Lorsqu'il n'y a plus de déplacement, le gradient est nul et le barycentre atteint. L'idée générale de la procédure est illustrée par la Figure 4.3 et l'algorithme donné par 1.

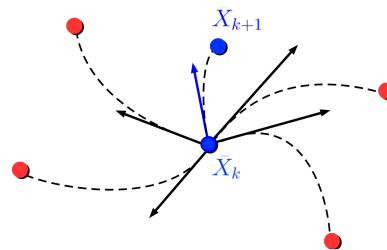


Figure 4.3: Concept du Karcher mean

Algorithm 1: Karcher mean

Input : $X_1, X_2, \dots, X_N \in \mathcal{M}$;
 $\mu_0 = X_1$
while $\mu_k \neq \mu_{k+1}$ **do**
 $\Delta\mu = \frac{1}{N} \sum_{i=1}^N \text{Log}_{\mu_k}(X_i)$;
 $\mu_{k+1} = \text{Exp}_{\mu_k}(\Delta\mu)$;
end
Output : μ

Le travail de synthèse des distances et des définitions des cartes exponentielle et Logarithme de Riemann est donné par la table 4.1 ci-après.

Table 4.1: Synthèse des distances et cartes Exponentielle et Logarithme

Géométrie	$V_p(\mathbb{R}^n)$	S_n	$\mathbb{R}^{n \times p}$
Distance	algo. [PBC21]	$d(X, Y) = \arccos(X^T Y)$	$d(X, Y) = (\sum_{k=0}^{n-1} \sum_{l=0}^{p-1} (x_{k,l} - y_{k,l})^2)^{\frac{1}{2}}$
Karcher mean		Descente de Gradient (algo. 1)	$X_1, \dots, X_n \in \mathbb{R}^{n \times p}, \bar{X} = \frac{1}{k} \sum_{i=1}^k X_i$
Exponentielle de Riemann	algo.2	$Exp_X(v) = \cos(v) \times x + \frac{\sin(v)}{ v } \times y$	$Exp_X(v) = X + v$
Logarithme de Riemann	algo.3	$Log_X(Y) = \frac{d(X,Y)}{\sqrt{1-(X^T Y)^2}} (Y - (X^T P)X)$	$Log_X(Y) = Y - X$

Nous donnons également les algorithmes qui permettent l'implémentation des cartes Exponentielle et Logarithme sur la variété de Stiefel [Zim16]; [PBC20].

Algorithm 2: Carte Exponentielle sur la variété de Stiefel

Input: $U \in V_p(\mathbb{R}^n)$, $\Delta \in T_U V_{n,p}$
 Décomposition de Δ : $\Delta = UU^T \Delta + (I - UU^T) \Delta$
 Décomposition QR de $(I - UU^T) \Delta$: $\Delta = UA + Q_E R_E$
 $\begin{bmatrix} M \\ N_E \end{bmatrix} = \exp\left(\begin{bmatrix} A & -R_E^T \\ R_E & 0 \end{bmatrix}\right) \begin{bmatrix} I_p \\ 0 \end{bmatrix}$
 Output: $Exp_U^S(\Delta) = UM + Q_E N_E$;

Algorithm 3: Carte Logarithme sur la variété de Stiefel

Input: $U \in V_p(\mathbb{R}^n)$, $\bar{U} \in V_p(\mathbb{R}^n)$, seuil τ ;
 Initialisation:
 $M = U^T \bar{U}$
 $QN = \bar{U} - UM$
 $V_0 = \begin{bmatrix} M & X_0 \\ N & Y_0 \end{bmatrix}$
 $\log(V_0) = \begin{bmatrix} A_0 & -B_0^T \\ B_0 & C_0 \end{bmatrix}$
while $\|C_k\|_2 \leq \tau$ **do**
 $\left| \begin{array}{l} \phi_k = \exp(-C_k) \\ V_{k+1} = V_k W_k \text{ où } W_k = \begin{bmatrix} I_p & 0 \\ 0 & \phi_k \end{bmatrix} \end{array} \right.$
end
 Output: $Log_U^S(\bar{U}) = UA_k + QB_k$

4.2.4. Le benchmark

La représentation des séries temporelles faites, les distances définies sur les variétés trouvées et synthétisées, nous avons pu adapter les algorithmes classiques de clustering travaillant sur l'espace euclidien mono-dimensionnel, aux variétés de Stiefel, Sphère et Euclidien matriciel. Les algorithmes qui nous ont servi de Gold Standard, sont les algorithmes de K-means, d'Agglomération Hiérarchique, et HDBSCAN. Les dimensions des matrices de trajectoire pouvant être grande, et afin de faciliter l'implémentation des distances de la table 4.1 il nous a été nécessaire de trouver une méthode efficace de réduction de dimension, autre que les PCA ou KPCA devant être adaptées au contexte des variétés. Lors de la soutenance de thèse de Maël Dugast, les échanges avec Pierre Baudot m'ont permis de découvrir un nouvel algorithme de réduction de dimension basé sur la conservation topologique du nuage de points initial. Nous avons donc exploré l'algorithme Uniform Manifold Approximation and Projection (UMAP) [MHM18] qui s'est montré extrêmement efficace et utile.

Nous avons montré dans [PBC20] que le couple UMAP pour la réduction de dimension et HDBSCAN pour la recherche de clusters était le plus performant sur la variété de Stiefel eu égard aux autres algorithmes de clustering, qu'ils soient utilisés seuls ou en conjonction avec UMAP. Dans cet article nous sommes uniquement montré que sur la variété de Stiefel, ce couple permettaient d'avoir les meilleures performances de clustering de séries temporelles issues de la base UCR Time Series Classification [Che+15]. Nous sommes allés un cran plus loin dans les tests en cherchant tout d'abord à connaître les performances de clustering des framework proposés (algorithme de clustering, réduction de dimensions+clustering, sur les variétés de Stiefel, Sphère et Euclidienne matricielle) lorsque les séries temporelles sont obtenues directement à partir de systèmes dynamiques puis nous avons testé l'ensemble de tous les framework sur la base UCR de séries temporelles.

Clustering de séries temporelles de systèmes dynamiques

Lorsque la série temporelle est issue d'un système dynamique, nous avons pris par exemple le système dynamique de Lorentz, mais de paramètres générateurs différents. On peut voir les différentes trajectoires sur la Figure 4.4. Le clustering des séries est totalement parfait pour le couple UMAP+HDBSCAN sur la variété de Stiefel. Les résultats sont ensuite moins bons pour la variété sphère, l'euclidien matriciel et catastrophique pour l'euclidien classique.

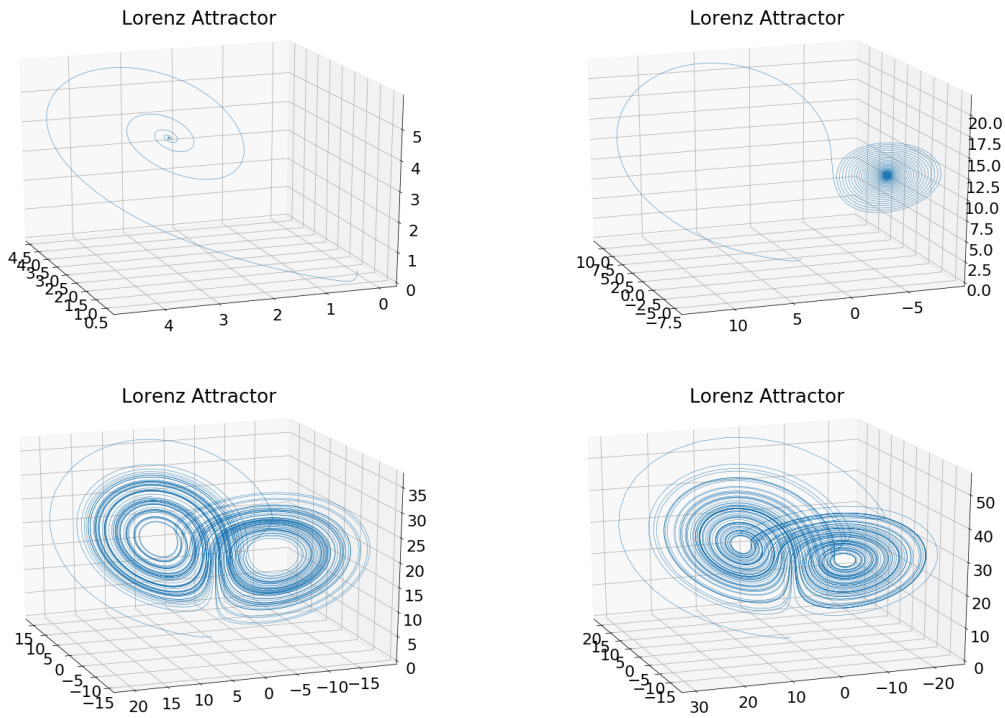


Figure 4.4

Nous avons ensuite testé quatre systèmes dynamiques différents, les systèmes de Lorenz, Rössler, Henon et Ikeda en construisant une large base de 400 séries temporelles. Un exemple des attracteurs de ces systèmes est donné en Figure 4.5.

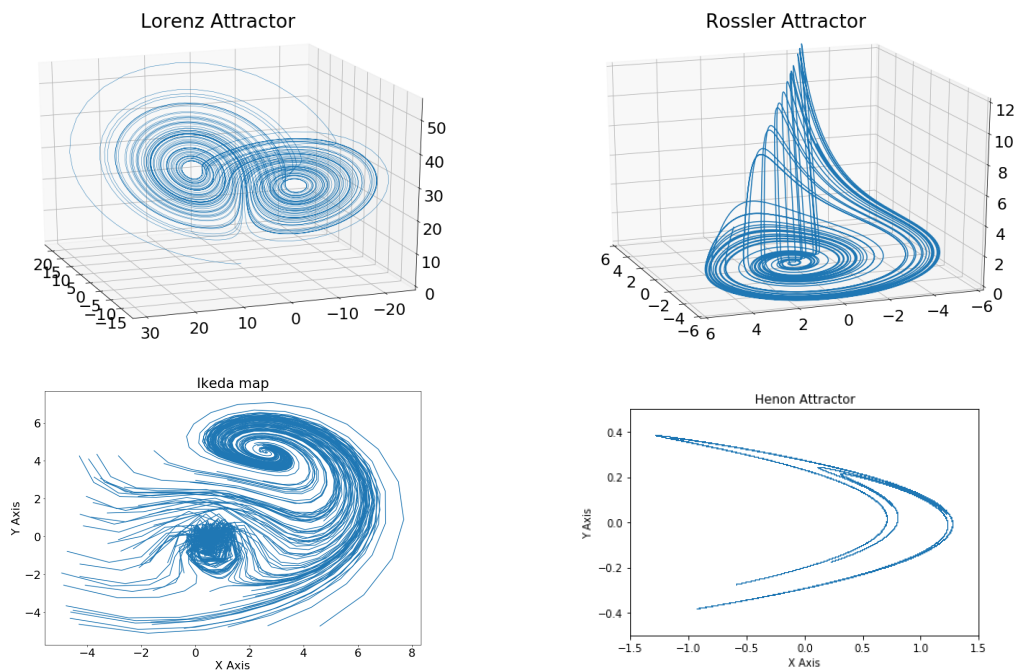


Figure 4.5

Dans 85% des cas, le couple UMAP+HDBSCAN sur la variété de Stiefel arrive à séparer chacune des classes de système dynamique, 80% des cas pour le couple UMAP+HDBSCAN sur la Sphère, 48% pour le couple UMAP+HDBSCAN sur l'euclidien matriciel et seulement 47% des cas pour l'euclidien classique.

Ces premiers résultats nous montrent bien l'utilité de représenter les séries temporelles dans des espaces de plus grande dimension. Dans le cadre où les séries temporelles sont issues de systèmes dynamiques, très clairement la représentation sur la variété de Stiefel permet de mieux exploiter la géométrie des attracteurs et donc de mieux séparer les trajectoires pour un meilleur clustering. Dans ce cadre précis, néanmoins, l'estimation de la trajectoire de l'attracteur du système doit être performante. Cela passe par l'estimation du temps de retard τ décrit en 4.1. Cette estimation a été réalisée par un algorithme basé sur l'information mutuelle [Guo+21]. Fort de ces très bons résultats, nous avons donc testé l'ensemble des frameworks sur la base de données internationale UCR times series classification.

Clustering de séries temporelles de la base UCR

Les résultats synthétisés de cette expérience sont reportés dans la Table 4.2 ci-après. Comme déjà mentionné, la construction de la représentation d'une série temporelle sur un espace de grande dimension est réalisée via le théorème de Takens' où la taille de la fenêtre glissante a une grande importance. Pour cette expérience, nous avons sélectionné les meilleurs résultats de clustering pour chacun des frameworks en faisant une recherche exhaustive de la valeur de τ dans 4.1. Cette recherche nous a indiqué que pour des séries temporelles (en tout cas pour certains modèles), la recherche d'un τ optimal était nécessaire mais sa valeur n'était pas vraiment en adéquation avec la valeur du τ donnée dans le contexte de systèmes dynamiques. Cette question reste ouverte même si nous avons déjà quelques résultats.

Du point de vue de la représentation géométrique, nous obtenons que :

- la variété de Stiefel est la meilleure représentation géométrique pour 47 % des bases de données. Cela signifie que pour 37 bases de données parmi les 79, l'un des 4 algorithmes de clustering utilisé sur la variété de Stiefel donne les meilleurs résultats.
- La sphère unité est la meilleure pour 44% des bases de données (35/79).
- L'espace $\mathbb{R}^{n \times p}$ est le meilleur pour 21,5% des bases de données (17/79).
- L'approche euclidienne sur les données brutes arrive en dernière position avec seulement 26% des bases de données bien clusterisées (21/79).

Une fois la représentation géométrique des séries temporelles analysée, nous laissons le lecteur se faire une idée sur ce qui pourrait être le meilleur framework de clustering à utiliser sur la variété de Stiefel. Ci-joint, quelques éléments d'aide :

- UMAP+HDBSCAN est le meilleur framework de clustering pour 58,2% des bases de données (46/79).
- K-means est capable de bien clusteriser les bases de données dans 31,6% des cas (25/79).
- l'algorithme hiérarchique est parvenu à clusteriser 29,1% des bases de données (23/79).
- HDBSCAN seul a obtenu un ratio de 20% de bon clustering (16/79).

Table 4.2: Caractéristiques principales des performances de clustering sur les 79 bases de séries temporelles. Les indicateurs de performance utilisés sont le v-measure score ainsi que le ARI score indiqué entre parenthèses.

Variété	$\mathbb{R}^{n \times p}$			
Algo. de Clustering	K-means	Hier.	HDB.	U.+HDB.
Moyenne	.240 (.220)	.224 (.152)	.217 (.151)	.300 (.226)
Std	.241 (.252)	.253 (.208)	.232 (.202)	.272 (.234)
Nbr. meilleur résultats	3 (4)	2 (3)	3 (4)	9 (7)
% meilleur résultats	3.8% (5.1)	2.5% (3.8)	3.8% (5.1)	11.4% (8.9)
Variété	S_n			
Algo. de Clustering	K-means	Hier.	HDB.	U.+HDB.
Moyenne	.303 (.228)	.254 (.199)	.227 (.178)	.322 (.252)
Std	.260 (.235)	.290 (.270)	.250 (.232)	.270 (.230)
Nbr. meilleur résultats	7 (8)	10 (5)	6 (6)	12 (6)
% meilleur résultats	8.9% (10.1)	12.7% (6.3)	7.6% (7.6)	15.2% (7.6)
Variété	$V_p(\mathbb{R}^n)$			
Algo. de Clustering	K-means	Hier.	HDB.	U.+HDB.
Moyenne	.245 (.166)	.220 (.223)	.190 (.164)	.376 (.270)
Std	.228 (.224)	.263 (.265)	0.219 (.224)	0.271 (.242)
Nbr. meilleur résultats	7 (4)	5 (5)	4 (4)	21 (14)
% meilleur résultats	8.9% (5.1)	6.3% (6.3)	5.1% (5.1)	26.6% (5.1)
Variété	Euclidien \mathbb{R}^l			
Algo. de Clustering	K-means	Hier.	HDB.	U.+HDB.
Moyenne	.281 (.210)	.242 (.138)	.215 (.131)	.293 (.183)
Std	.256 (.228)	.260 (.219)	.233 (.201)	.259 (.217)
Nbr. meilleur résultats	8 (8)	6 (0)	3 (0)	4 (1)
% meilleur résultats	17.7% (10.1)	7.6% (0)	3.8% (0)	5.1% (1.3)

Une partie de la conclusion est que l'algorithme HDBSCAN seul semble être le pire algorithme de clustering. Cependant, son utilisation avec UMAP semble vraiment très intéressante car les résultats sont grandement améliorés. L'utilisation de UMAP comme étape de pré-traitement de l'algorithme HDBSCAN semble donc très importante.

Enfin, les auteurs de [JLR20] ont effectué un travail de benchmark similaire et sur les mêmes bases de données. Ils ont testé deux distances, la distance euclidienne et la Distance Time Wrapping (DTW) ainsi que 5 algorithmes de clustering, K-means, K-medoids, Fuzzy C-means, Density Peaks et Agglomerative. La méthode K-means avec la distance euclidienne sur les données brutes semble être la meilleure méthode de clustering de leur benchmark (voir fig. 5 de l'article). Or nous avons montré que le framework UMAP+HDBSCAN sur la variété de Stiefel présentait clairement de meilleurs résultats que K-means sur les données brutes. Ce framework surpasse donc le meilleur proposé dans [JLR20]. En conclusion, aucune des méthodes de clustering proposées et testées dans [JLR20] ne semble meilleure que les frameworks géométriques que nous avons proposé et étudié.

4.3. Application aux flux de patients adultes

En outre, une série chronologique approximative peut être reconstruite à partir de ces clusters. Les codes CIM-10 obtenus à partir du cluster présentant le VRS et la grippe, ont été moyennés pour construire une nouvelle série chronologique, qui représente la signature des virus. La nouvelle série temporelle des admissions est représentée sur la Figure 4.6 pour Grenoble. Nous avons fait de même avec les résultats obtenus avec K-means et une distance euclidienne.

Avec le framework K-means euclidien, il est clair qu'en se reportant aux Figure 4.6-(a) et Figure 4.6-(c) que le schéma particulier des virus n'est pas visible et les admissions sont constantes sur l'année. Les clusters obtenus par cette méthode sont par conséquent trop imprécis.

A contrario le framework UMAP+HDSCAN sur la variété de Stiefel validé théoriquement offre de bons résultats en pratique également. Pour l'hôpital de Saint-Etienne, la reconstruction est imparfaite mais les séries temporelles obtenues reflètent bien la périodicité et une partie de l'amplitude des pics de présence virale. Pour l'hôpital de Grenoble, les résultats sont encore meilleurs. En particulier, la reconstruction de la grippe est très précise, Figure 4.6-(b), celle du VRS Figure 4.6-(d) très intéressante car c'est le virus le plus délicat à détecté chez l'adulte (assez peu recherché aussi ...).

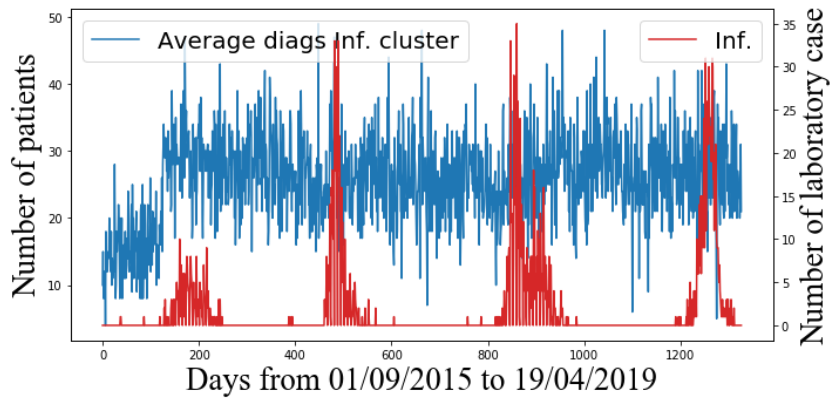
Cela valide pratiquement et opérationnellement l'avantage apporté par une représentation des séries temporelles sur un espace de grande dimension projeté ensuite sur la variété de Stiefel. En outre, ces reconstructions donnent une représentation jour par jour des codes CIM-10 associés à la présence quotidienne du virus. La détection des points de changement pourrait être appliquée chaque jour afin d'augmenter la capacité de l'hôpital à éviter les charges anormales dues aux virus de l'hiver.

4.4. Contributions scientifiques de cette partie

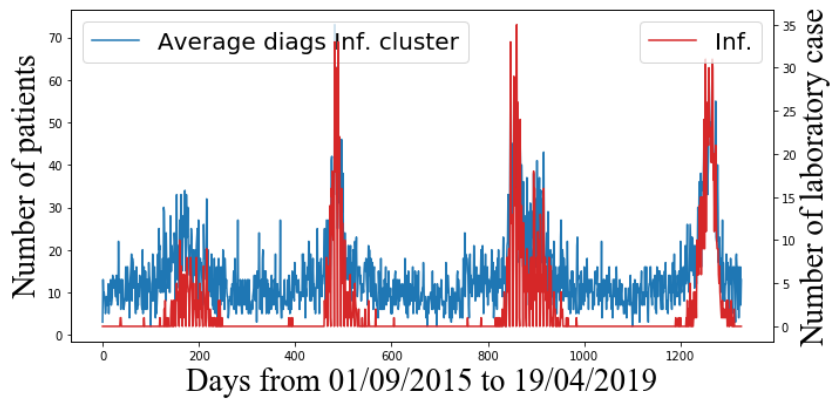
Bilan scientifique



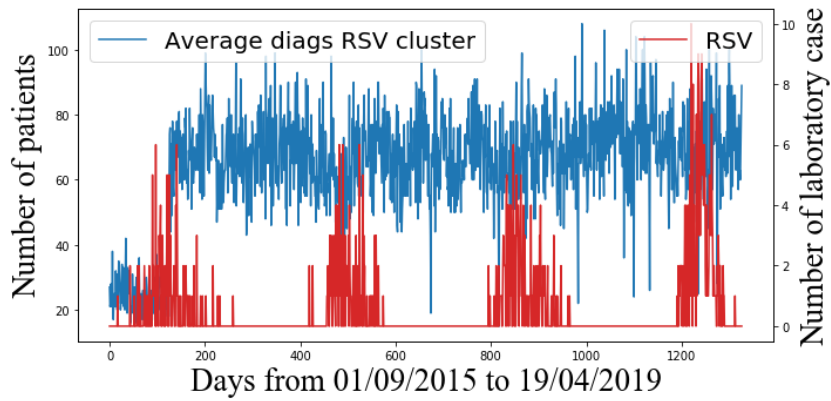
- 2 Articles de revue [[Mor+22](#)]; [[PBC21](#)]
 - 1 Article de revue soumis [[PBC22](#)]
 - 3 Articles de conférence Internationale à comité et acte de lecture [[Sol+18](#)]; [[PBC19](#)]; [[PBC20](#)]
-



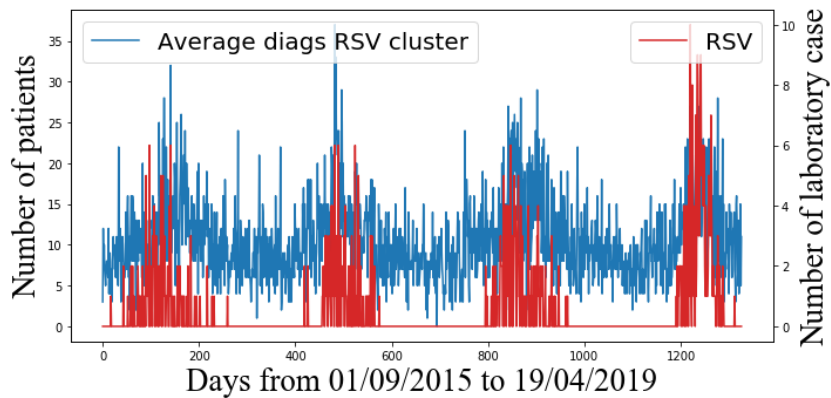
(a) Framework K-means euclidien - virus influenza



(b) Framework UMAP+HDSCAN sur la variété de Stiefel - virus influenza



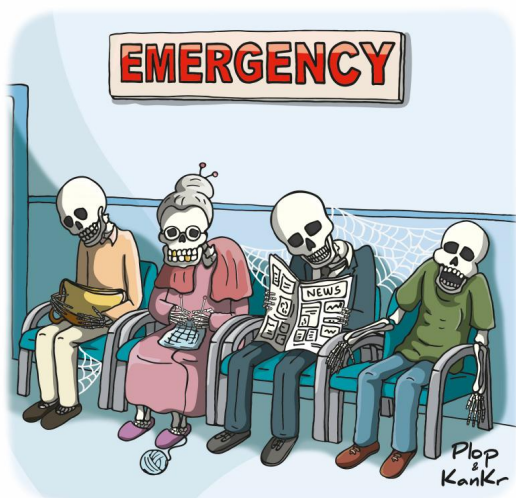
(c) Framework K-means euclidien - VRS



(d) Framework UMAP+HDSCAN sur la variété de Stiefel - VRS

Figure 4.6: Grenoble

Sur l'organisation des urgences adultes



Jusqu'à présent, je me suis essentiellement concentré sur l'information portée par le flux de patients, qu'il soit pédiatrique tout d'abord puis adulte, pour des pathologies bien définies. Rien dans ces travaux ne permet réellement de traiter la tension dans le service des urgences par une modélisation du flux total des patients. C'est en répondant à la sollicitation du ministère de la santé pour définir l'indicateur de temps d'attente aux urgences du tableau national, que

j'ai pu aborder la modélisation de ce flux. Sous une forme statistique dans un premier temps puis sous sa forme dynamique dans un second. Les notions de processus auto-similaire et groupe de renormalisation seront donc développées dans ce chapitre.

Sommaire

5.1	Contexte et hypothèses.	68
5.2	Le caractère auto-similaire du flux de patient	68
5.2.1	L'analyse du flux	69
5.2.2	Modélisation par des cascades multiplicatives	72
5.3	Le groupe de renormalisation pour l'organisation.	74
5.3.1	Principes généraux	74
5.3.2	L'opérateur de Koopman.	76
5.4	Contributions scientifiques de cette partie	80

5.1. Contexte et hypothèses

Définir l'attente pour un système de production, qu'il soit de biens ou de services, est abordé généralement par modélisation en file d'attente. Cela suggère un certain nombre d'hypothèses sur les loi inter-arrivées et de temps de service, et de dimensionner la file elle-même. L'attente, est de mon point de vue une caractéristique lié à une variabilité anormale d'un flux. Si nous prenons le cas des urgences, un patient qui arrive aux urgences avec sa/ses pathologie(s) passera un certain temps dans le service. En condition normale, les urgences voies donc un flux de patients parcourir les urgences. Le temps de passage d'un patient dans le service dépend alors de sa condition (disons sa/ses pathologie(s)) et des conditions de prise en charge (disons l'organisation au sein du service). En condition supposée anormale, *i.e.* sous une organisation ébranlée ou un flux anormal de pathologies par exemple, nous pouvons faire l'hypothèse que la trajectoire du temps de passage aux urgences est différente, avec une variabilité inhabituelle. Cette variabilité inhabituelle du temps que les patients passent aux urgences, est sous ces hypothèses, caractéristique de l'attente.

En prenant ce point de vue, le flux de passage et le temps de passage aux urgences, sont vus comme des flux ou des fluides se rapprochant de flux bien connus. On peut citer, sans être exhaustif, le flux de voitures ou encore le flux de trames ou de paquets dans les réseaux. Or, ces flux sont statistiquement bien modélisés aujourd'hui comme des processus auto-similaires.

Cette hypothèse faite, après quelques recherches bibliographiques, je n'ai trouvé que très peu de références qui abordaient les urgences, ou en relaxant, le système de santé sous cet aspect [PGW04]. Cet article est une des seules références traitant le flux de patients sous l'angle statistique. Je reviendrai un peu plus loin sur cette référence car elle a été décisive pour orienter mes recherches vers le groupe de renormalisation.

5.2. Le caractère auto-similaire du flux de patient

La toute première étape a donc été de vérifier cette hypothèse d'auto-similarité du flux de patients aux urgences. Cette hypothèse a été renforcée par le constat du comportement quotidien moyen des arrivées aux urgences par rapport aux flux quotidien moyen de voitures et du flux quotidien moyen de connexion à internet.

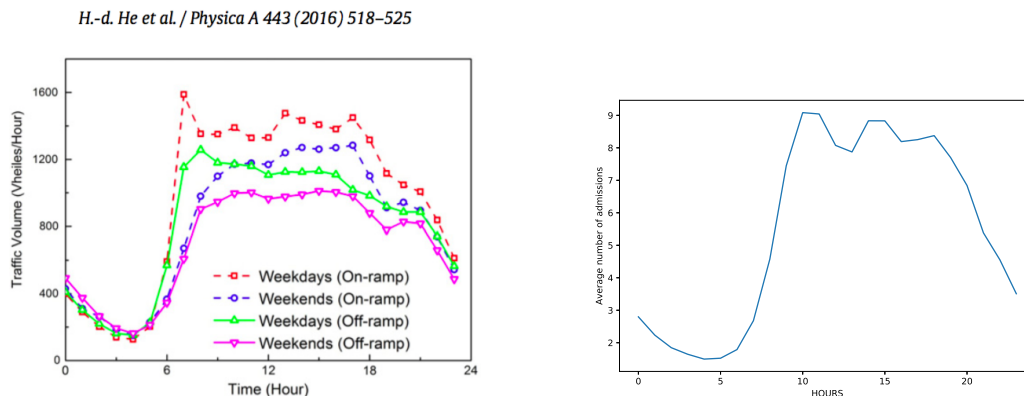


Figure 5.1: Distribution quotidienne des flux de voitures à gauche et des patients à droite.

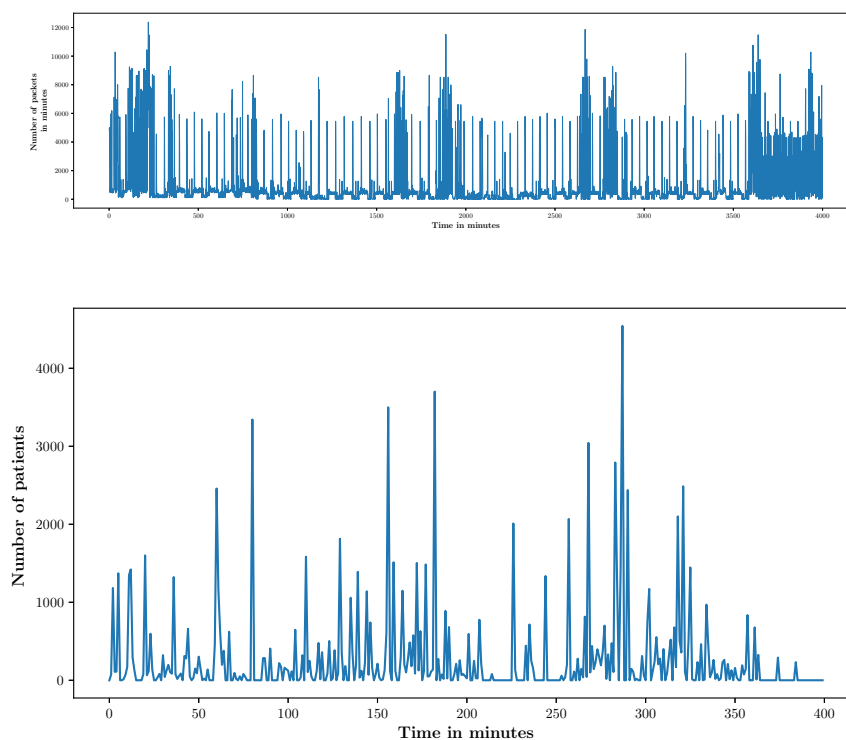


Figure 5.2: Flux de trafic internet en haut et flux d'arrivées des patients (CHU de Saint-Etienne) en bas.

Comme on peut le voir sur la Figure 5.1 ces distributions sont très similaires, et chacun de ces systèmes subit aléatoirement des effets de burst comme montré dans et par la Figure 5.2.

5.2.1. L'analyse du flux

L'hypothèse faite est donc que le flux des admissions ainsi que le temps de passage aux urgences sont des processus auto-similaires. Pour rappel, un processus auto-

similaire $X = \{X(t), t \geq 0\}$ est un processus stochastique qui respecte la relation

$$X(t) \stackrel{d}{=} \{c^H X(t/c)\} \text{ pour tout } c > 0. \quad (5.1)$$



- ↪ L'égalité est à comprendre au sens des distributions
- ↪ Le paramètre H prend des valeurs entre 0 et 1, et correspond au paramètre d'auto-similarité ou paramètre de Hurst
- ↪ H reflète une certaine dépendance :
 - pour $0 < H < 0.5$ le processus est anti-persistant
 - for $0.5 < H < 1$ le processus est persistant
 - for $H = 0.5$ le processus est Brownien
- ↪ La DFA peut servir d'estimateur du paramètre de Hurst

Dans [Emi+18] nous avons testé cette hypothèse et validé le caractère auto-similaire en réalisant différents tests. Tous ont montré que le processus des admissions et le temps de passage aux urgences sont auto-similaires avec une variation du paramètre de Hurst en fonction des échelles, laissant entendre que le processus est donc multi-fractionnaire.

Nous avons maintenant une notion de la dépendance long terme entre les incréments du processus, mais qu'en est-il de leur comportement local ?

La régularité locale de la trajectoire du processus stochastique $X = \{X_t, t \in [0, T]\}$ est communément étudiée via la notion d'exposant de Hölder. Cet exposant se définit par la relation

$$\alpha_q(s) = \sup \left\{ h : \frac{|X_t - X_s|}{|t - s|^h} \leq C \right\} \quad (5.2)$$

pour un ordre donné q .



- ↪ Notons que pour un processus auto-similaire, la dimension fractale, qui est la dimension de Hausdorff, et l'exposant de Hurst sont liés par la relation $H + D = n + 1$, pour une surface de dimension n .
- ↪ C'est un moyen de relier les propriétés locales et à long terme
- ↪ On distingue au moins deux approches principales pour caractériser la multi-fractilité

Méthode multi-échelles des moments statistiques

un processus stochastique auto-similaire X_t est multi-fractal lorsqu'il respecte la relation

$$E[|X_t|^q] = c(q)t^{\tau(q)+1}, \quad \forall t \in \mathcal{G} \text{ and } q \in \mathcal{Q} \quad (5.3)$$

où $0 < T < \infty$, \mathcal{G} et \mathcal{Q} sont des intervalles ouverts de $[0, T] \subset \mathcal{G}$ et $[0, 1] \subset \mathcal{Q}$. La fonction $\tau(q)$ est appelée fonction d'échelle.



- ↪ Si $\tau(q)$ vs q est une ligne droite \rightarrow le processus est mono-fractal
- ↪ Si $\tau(q)$ vs q est une fonction convexe \rightarrow le processus est multi-fractal.

L'adaptation de la méthode DFA peut être utilisée sur les moments empiriques, c'est la MF-DFA.

Le spectre de singularité

L'exposant de Hurst généralisé est la dimension fractale D . Tout d'abord, on prend le support isoHölder tel que

$$\mathbb{E}_q(h) = \{\alpha_q : h(t_0) = \alpha_q\}. \quad (5.4)$$

Comme vu précédemment, la dimension de Hausdorff de l'ensemble des points $E(h)$ donne la dimension de Hausdorff, également appelée exposant de Hurst généralisé $h(q)$, tel que

$$h(q) = D(h) = \dim_H(\mathbb{E}(h)) \quad (5.5)$$

qui définit

$$D(q) = \frac{\tau(q)}{q-1} = \frac{qh(q) - 1}{q-1} \quad (5.6)$$

Ensuite, le spectre de singularité est lié à l'exposant d'échelle par la transformée de Legendre.

$$\tau'(q) = h(q) + qh'(q) \quad (5.7)$$

Lorsque l'exposant de Hölder prend une valeur unique H en chaque point t_0 , le processus multi-fractal devient un processus mono-fractal.

une fois encore les tests menés dans [Emi+18] avec l'algorithme MF-DFA, ont conclu que les processus d'admissions et de temps d'attente étaient multi-fractals. Nous avons également proposé dans cet article un suivi temps-réel de l'exposant de Hurst afin de pouvoir cibler les moments où le processus des admissions devenait très persistant ou aléatoire. L'idée sous-jacente est que plus il y a de persistance avec un paramètre de Hurst élevé et plus l'existence d'un problème de congestion dans le service est probable, suggérant un dysfonctionnement organisationnel.

Ayant montré empiriquement le caractère multi-fractal des processus, et comme chacun d'entre-eux est un processus positif, nous nous sommes tournés vers la modélisation à base de cascades multiplicatives toujours motivés par l'analogie faite avec le trafic de données internet dans les réseaux.

Les éléments qui suivent à présent n'ont pas encore été publiés. Ils ont été développés par Sid'Ahmed Emine, ou par les élèves ingénieurs du département Génie Industriel, INSA-Lyon, via les instruments « Initiation à la Recherche » (INR) ou « Recherche en Génie Industriel » (RGI). Ce sont des modules dont les heures allouées sont au premier semestre de la 5e année.

5.2.2. Modélisation par des cascades multiplicatives

Aussi bien Sid'Ahmed Emine que certains élèves INR et RGI ont pu mettre en œuvre une modélisation du flux des arrivées et du temps de passage aux urgences par des cascades multiplicatives.

J'ai décidé de tester le modèle des cascades multiplicatives binomiales avec répartition aléatoire sur chaque échelle de la construction. Le principe est simple et rapide à mettre œuvre pour une première approche. Son principe, est de considérer un ensemble \mathcal{I}_0 sur l'intervalle $[0, 1]$ et tel que la mesure de cet ensemble $m(\mathcal{I}_0) = 1$. La mesure m est ensuite divisée en 2 intervalles $\mathcal{I}_{0,0}$ et $\mathcal{I}_{0,1}$ affectés d'une probabilité p et $1 - p$ choisie aléatoirement suivant une loi uniforme. La mesure m initiale est conservée et on a $m_{0,0}(\mathcal{I}_{0,0}) + m_{0,1}(\mathcal{I}_{0,1}) = 1$. A l'étape suivante, les deux ensembles $\mathcal{I}_{0,0}$ et $\mathcal{I}_{0,1}$ sont à leur tour décomposés en deux nouveaux intervalles pour former quatre intervalles $\mathcal{I}_{1,0}, \mathcal{I}_{1,1}, \mathcal{I}_{1,2}, \mathcal{I}_{1,3}$ de sorte que la mesure soit préservée. Ce procédé est répété k fois, k étant le nombre d'échelle que l'on souhaite considérer, nous noterons qu'à chaque étape la mesure est préservée. Au final, en se plaçant à la k^e itération de la cascade, la masse totale initiale est partagée sur des intervalles dyadiques $[i 2^{-k}, (i + 1) 2^{-k}]$, où $i = 0, \dots, k - 1$ et la mesure sur l'intervalle \mathcal{I}_k vaut

$$\prod_{i=1}^k = p^{n_0} (1 - p)^{n_1} \quad (5.8)$$

avec $n_0 + n_1 = k$ et n_0, n_1 le nombre de fois où la mesure est affectée par les probabilités p et $(1 - p)$. Un exemple visuel de cette procédure de construction est donné par la figure 5.3 ci-dessous.

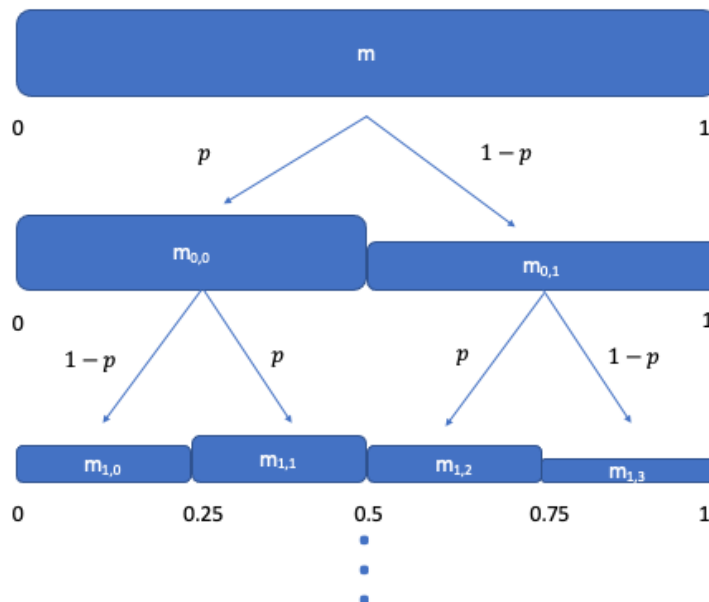


Figure 5.3: Représentation de la démarche des cascades multiplicatives binomiales, avec probabilité p constante pour chaque niveau de la cascade.

En analysant les résultats et en comparant avec le régime moyen des arrivées aux urgences au cours d'une journée Figure 5.1 nous avons raffiné quelque peu la modélisation. La distribution des admissions ayant des masses variant au cours d'un journée, j'ai voulu tester une discrétisation de cette distribution en construisant des cascades par morceaux. Après des essais empiriques pour déterminer les paramètres de probabilité et d'échelle, nous avons obtenu un modèle dont l'adéquation avec les distributions d'arrivées par heure est donnée en Figure 5.4.

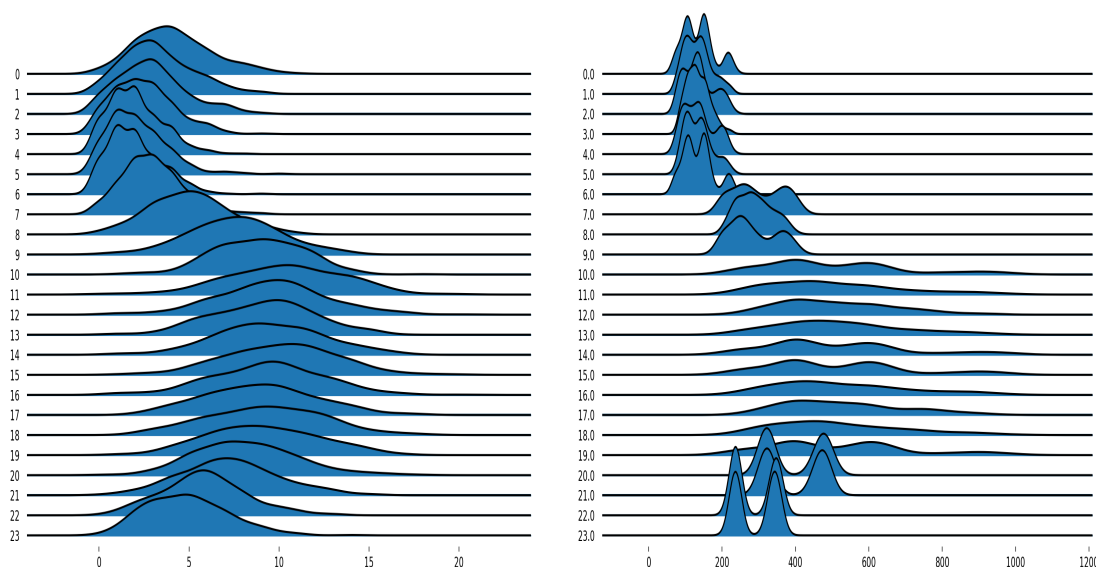


Figure 5.4: Distribution horaire du flux de patients au CHU de Saint-Etienne à gauche. Distribution horaire modélisée du flux de patients à droite.

Certes le résultat n'est pas encore au mieux, mais je noterai que par le modèle proposé, le flux horaire est proche du réel. La distribution moyenne quotidienne du flux de patient est elle-aussi respectée dû à la construction des cascades, et enfin la multi-fractionnalité du flux a été vérifiée. Il est évident que plus d'investigations sont nécessaires, tant pour l'estimation des paramètres, que sur le modèle de cascade lui-même, néanmoins ces premiers résultats encourageants, m'ont permis de considérer ce modèle théorique comme viable pour la suite des analyses et propositions.

Sous le spectre des cascades multiplicatives, bon nombre de méthodes d'estimation du paramètre de Hurst ou de Hölder ont été proposées. Elles permettent de donner une information temps-réelle de l'évolution de ce(s) paramètre(s) et ainsi donner une indication sur la congestion possible dans le système. Basé sur la méthodologie présentée par [MGD02] qui reprend l'algorithme Wavelet Transform Modulus Maxima (WTMM) bien connu, nous avons regardé quels pouvaient être les résultats sur les données simulées puis celles du CHU de Saint-Etienne. Encore une fois, même si ces premiers résultats sont perfectibles, ils n'en restent pas moins encourageants.

C'est en approfondissant la littérature autour de la modélisation fractale du flux de patients ou plus généralement du système de santé, qu'encore une fois j'ai été attiré par

un concept qui m'était jusqu'ici encore inconnu. À la suite de la lecture du dernier paragraphe de la partie discussion de [PGW04], le **groupe de renormalisation** est mentionné comme un outil presque usuel pour l'analyse des processus auto-similaires, ou même multi-fractals. J'ai donc commencé à approfondir ce concept au travers des références citées dans [PGW04] puis en cherchant ensuite à relier le caractère « géométrique et variété » au groupe de renormalisation, j'ai découvert les travaux d'A. Lesne [Les98] puis de M. Morrisson [Mor14] qui ont été révélateurs et une porte d'entrée vers ce monde qu'est la renormalisation.

5.3. Le groupe de renormalisation pour l'organisation

Finalement tout s'est très vite recoupé. La méthodologie DFA pour la recherche de dépendance long terme, le concept multi-échelles et la conservation des distributions moyennant un facteur de dilatation et surtout le principe même des cascades multiplicatives. Pour arriver à comprendre un peu mieux ces relations, je propose de présenter le groupe de renormalisation puis les pistes que nous avons explorées.

5.3.1. Principes généraux

Là encore, le concept nous amène vers des contrées obscures, liées ici à des problèmes de la physique quantique. Néanmoins, il faut retenir l'idée principale de l'utilisation du groupe de renormalisation. Il s'agit de pouvoir réduire les degrés de libertés d'un système à fortes interactions. C'est une technique permettant de trouver le comportement macroscopique d'un système complexe en le décomposant en une succession de problèmes comportant un nombre décroissant de parties microscopiques en interaction à différentes échelles. Le domaine d'application privilégié du groupe de renormalisation est donc celui d'un grand nombre d'éléments ayant une interaction complexe et importante. Même si c'est la physique quantique qui a guidé l'élaboration de cette technique, on peut sentir l'analogie avec un système d'accueil des urgences, où chacun des éléments de l'ensemble {soignants, patients, salles, lits, etc } interagissent de manière complexe, provoquant une dépendance sur le temps passage de chacun des patients. Il serait donc possible d'avoir une caractérisation théorique de l'organisation du service au travers de l'analyse d'un observable, le flux de patients.

Afin de faire la connexion avec notre flux de patients ou de temps de séjour et les approches d'analyse présentées précédemment, prenons l'approche de Sinai [Sin76]. Ce dernier part de l'espace vectoriel des réalisations et propose naturellement un opérateur de translation dans cet espace qui permet de quantifier la stationnarité. Ensuite il définit un semigroup multiplicatif, que je nommerai R_k , avec $k \in \mathbb{N}^*$ dont l'action sur la réalisation est la suivante

$$R_k X = k^{-\frac{1}{2}} \sum_{\substack{(i-1)\frac{T}{k} \leq \mathcal{J} < i\frac{T}{k} \\ i=1, \dots, k}} X_{\mathcal{J}}. \quad (5.9)$$

J'ai naturellement simplifiée l'expression pour une observation mono-dimensionnelle. Cette expression signifie simplement que la longueur de l'observation

X est subdivisée en petits intervalles de longueur k , dans chaque intervalle k on fait la somme 2 à 2 des observations et la somme est normalisée par la longueur de l'intervalle élevée la puissance $\frac{1}{2}$. L'opérateur R_k ainsi définie agit sur l'espace des distributions et le processus est auto-similaire lorsque sa distribution est invariante pour cet opérateur, autrement dit si $P \in \mathcal{D}$ est une distribution de l'espace des distributions (stationnaires) alors

$$R_k^* P = P \quad \forall k \in \mathbb{N}^*, P \in \mathcal{D} \tag{5.10}$$

avec R_k^* l'adjoint de R_k . Le processus auto-similaire X de distribution P est alors un invariant ou encore un point fixe pour le semigroup $\{R_k^*\}_{k \in \mathbb{N}^*}$.

Le version que je viens de présenter, s'apparente à la technique de « coarse-graining » et est plutôt rencontrée dans la littérature pour la recherche de l'opérateur de renormalisation R_k sous le contexte de « real space renormalization ».

Jusqu'ici, le groupe de renormalisation nous a permis de définir un processus auto-similaire comme un invariant obtenu par itération successive d'un opérateur. Existe-t-il d'autres invariants ? Peut-il y avoir des cas où la convergence n'existe pas ? Ce sont des questions qui peuvent avoir réponse en ne considérant plus simplement une seule réalisation, mais en essayant de comprendre l'action de la renormalisation sur l'ensemble des réalisations possibles et produites par le système des urgences. Cette considération m'a donc conduit à étudier l'incidence de l'opérateur sur le système sous sa forme dynamique. Un des avantages à cela est de se rapprocher de la théorie originelle qui considère l'opérateur de renormalisation comme un opérateur d'évolution dans des espaces similaires comme l'espace des hamiltoniens, des couplages, ou encore des théories. Dans ces espaces l'action de l'opérateur décrit alors une trajectoire, un flot dont une expression peut être formulée comme

$$R_b H(0) = H(T^n(0)) \tag{5.11}$$

avec le paramètre b un facteur général de normalisation et T , la transformation d'un point de l'espace à un autre après le coarse-grained effectué.

Pour résumer et illustrer ce qui vient d'être décrit dans cette partie, je propose de regarder la Figure 5.5 issue de [EV10] et quasiment reproduite en l'état.

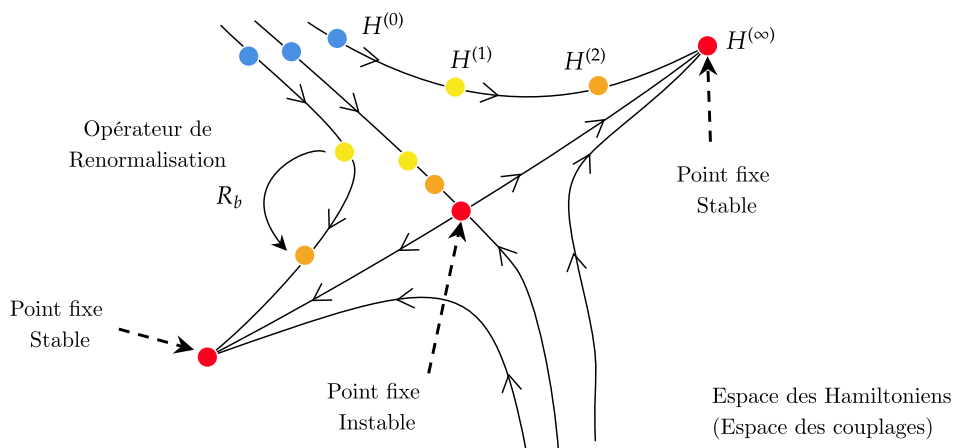


Figure 5.5: Principe du flots de l'opérateur de renormalisation.

Le champ des possibles s'ouvre alors, même si les points clés pour obtenir ces flots sont assez délicats. En cherchant comment nous pourrions construire ces espaces, un article intéressant de W. Redman [Red20] nous a fait découvrir l'opérateur de Koopman et son intérêt pour la renormalisation.

5.3.2. L'opérateur de Koopman

Partons d'un système dynamique s'écrivant

$$\frac{d}{dt}X(t) = f(X(t)), \quad (5.12)$$

où $X \in \mathbb{R}^N$ représente l'état du système et $f : \mathbb{R}^N \rightarrow \mathbb{R}^N$ est un champ vectoriel éventuellement non linéaire. La relation (5.12) induit une famille de cartes dépendant du temps $\{F_t\}_{t \in \mathbb{R}^+}$ satisfaisant

$$X(t) = F_t(X_0) \quad (5.13)$$

et qui donne l'état $X(t)$ au temps t en fonction de l'état initial X_0 . L'idée ensuite défendue par Koopman et Von Neumann repose sur la possibilité d'exprimer par un opérateur linéaire de dimension infinie la dynamique du système non linéaire. L'opérateur agit en conséquence sur les états du systèmes via un ensemble de fonctions observable g de sorte que le semigroup $\{\mathcal{K}_t\}_{t \in \mathbb{R}^+}$

$$\mathcal{K}_t g = g \circ F_t, \quad (5.14)$$

où \circ est l'opérateur de composition et tel que

$$\mathcal{K}_t g(X_0) = g(F_t(X_0)) = g(X_t). \quad (5.15)$$

L'opérateur de Koopman \mathcal{K}_t fait correspondre à la fonction g les valeurs qu'elle prendra après que la dynamique ait progressé d'un pas t , c'est bien un opérateur d'évolution sur l'espace des observables et non plus sur l'espace des états (ou espace des phases). En ce sens l'opérateur de Koopman est le dual de l'opérateur d'évolution Perron-Frobenius [KKS16].

Redman dans par exemple [Red20] montre effectivement l'analogie entre l'opérateur de renormalisation et l'opérateur de Koopman. Ainsi, l'opérateur de renormalisation est est opérateur de Koopman mais où les observables sont des fonctions des hamiltoniens et non plus les observables des fonctions d'états.

Convaincu que nous pourrions atteindre l'opérateur de renormalisation par le biais de l'opérateur de Koopman nous avons tout d'abords cherché à construire puis à utiliser cet opérateur sous la contrainte de ne disposer que d'une observation de la dynamique du service d'accueil des urgences. C'est en partie en croisant les recherches avec le clustering de séries temporelles que l'article de S. Hirsh [Hir+21] nous a permis de computer l'opérateur de Koopman dans ces conditions. Cet article donne d'ailleurs une interprétation supplémentaire de l'opérateur de Koopman comme étant un descripteur géométrique des courbures de la trajectoire de l'attracteur du système dynamique étudié. Outre cette interprétation, il nous a fait découvrir la Dynamic Mode Decomposition (DMD) et une version data-driven de son estimation par l'algorithme Hankel Alternative View Of Koopman (HAVOK).

DMD et HAVOK (et toutes leurs extensions ...)

L'algorithme DMD exact fonctionne comme suit [Kam+20]. Une matrice instantanée des états du système X^1 est construite ainsi que sa version décalée temporellement X^2

$$X^1 = \begin{pmatrix} | & | & \dots & | \\ X_1 & X_2 & \dots & X_{M-1} \\ | & | & & | \end{pmatrix}, X^2 = \begin{pmatrix} | & | & \dots & | \\ X_2 & X_3 & \dots & X_M \\ | & | & & | \end{pmatrix} \quad (5.16)$$

$$X^2 \approx KX^1 \quad (5.17)$$

L'algorithme estime ensuite une relation linéaire entre ces matrices de données

$$K = \arg \min_{\hat{K}} \|X^2 - \hat{K}X^1\|_F. \quad (5.18)$$

La norme F est la norme de Frobenius, et la solution de ce système d'équations est classiquement obtenue par

$$\hat{K} = X^2 X^{1\dagger} \quad (5.19)$$

où $X^{1\dagger}$ est la Moore-Penrose pseudoinverse de X^1 . Une fois la matrice \hat{K} obtenue, les modes propres sont calculés par diagonalisation, de sorte que

$$\hat{K} = \Phi \Lambda \Phi^{-1}. \quad (5.20)$$

Les vecteur colonnes de Φ^{-1} sont les DMD modes de \mathcal{K}_t (modes de Koopman) alors que les vecteurs colonne de Φ sont une approximation linéaire des fonctions propres de \mathcal{K}_t . Les modes de Koopman sont généralement renommés c_j car ils apparaissent dans la décomposition des observables dans la base des fonctions propres

$$(\mathcal{K}_t g)(X) = \sum_j^K c_j \lambda_i \phi_j X. \quad (5.21)$$

En pratique et dans notre contexte, nous ne disposons pas des états du système mais seulement d'un observable. Nous retrouvons maintenant le théorème de Takens vue dans le chapitre précédent qui par la construction d'une matrice de Hankel permet d'estimer la trajectoire de l'attracteur et d'avoir ainsi un accès aux états. L'algorithme HAVOK commence donc par la construction de cette matrice de Hankel avec

$$H = \begin{pmatrix} | & | & \dots & | \\ g(X_1) & g(X_2) & \dots & g(X_M) \\ | & | & & | \end{pmatrix} \quad (5.22)$$

Une SVD de la matrice de Hankel est faite et r vecteurs sont conservés. D'où

$$H_r = U_r \Sigma_r V_r^T. \quad (5.23)$$

On obtient une représentation de la trajectoire avec

$$V_r^T = (v_1 \ v_2 \ \dots \ v_n) \quad (5.24)$$

À ce stade nous serions tentés de créer les matrices V_1 et V_2 pour estimer \hat{K} mais les différents résultats ont montré que l'estimation était fortement incorrecte. Il est

nécessaire de contraindre cette estimation par la dernière colonne de V_r^T soit v_r . Ainsi $V_2 = V_{[1:r-1]}^T$ et $V_1 = V_{r-1}^T$ et l'estimation de \hat{K} s'écrit dans le cas discret

$$\mathcal{K}_t = \hat{K} = \frac{(V_2 V_1^\dagger - I)}{\Delta t} \quad (5.25)$$

avec I la matrice identité et Δt le pas entre 2 éléments de la trajectoire.

Nous avons donc mis en œuvre cet algorithme pour déterminer les modes de l'opérateur de Koopman associés à une simulation des admissions. Les paramètres pour construction de la matrice de Hankel ont été déterminés par des procédures classiques d'information mutuelle et de k-plus proche voisins. La Figure 5.6 montre deux trajectoires obtenues. La Figure 5.6-(a) est obtenue sur les données directement alors que la Figure 5.6-(b) est le résultat d'une interpolation des données simulées. Ces trajectoires montrent bien le caractère de périodicité dans le flux ainsi que l'aspect multi-échelles.

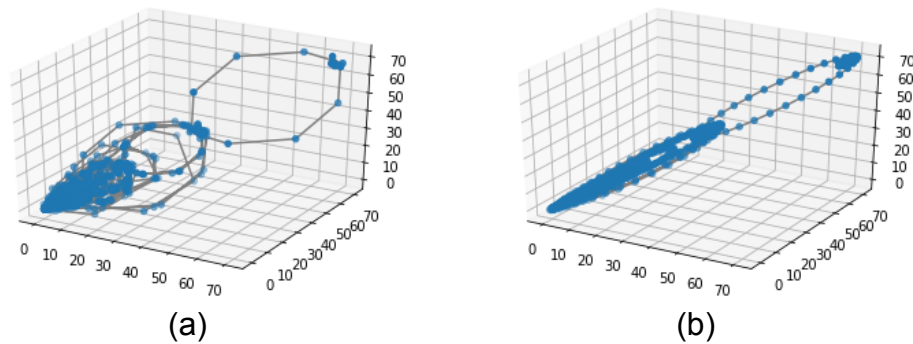


Figure 5.6: Trajectoires sur \mathbb{R}^3 du flux de patients modélisé. À gauche, sans interpolation, à droite avec une interpolation par splines.

Enfin, la matrice linéaire de Koopman, qui je le rappelle contient les courbures de la trajectoire sur ces éléments diagonaux, est donnée pour les deux trajectoires de la Figure 5.6. Sont affichés également trois modes estimés, Figure 5.7

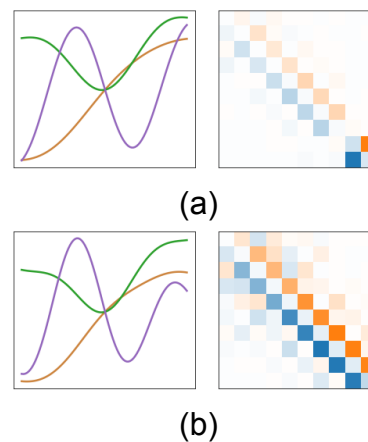


Figure 5.7: Opérateur de Koopman associés aux représentations de la Figure 5.6. À gauche 3 modes extraits, à droite, représentation de la matrice de l'opérateur.

Une fois l'opérateur de Koopman légèrement dompté, l'objectif a été de pouvoir obtenir un champ de vecteur de l'attracteur global. Nous n'avons pas pu encore aller au bout de cet objectif même si nous disposons a priori d'une expression du générateur infinitésimal $\mathcal{A}_{\mathcal{K}}$ du semigroup $\{\mathcal{K}_t\}_{t \in \mathbb{R}^+}$ comme [DTK20]

$$\mathcal{A}_{\mathcal{K}}X_j = \lim_{t \rightarrow 0} \frac{(\mathcal{K}_t - I)}{t} X_j \approx \frac{\ln \Lambda}{\Delta t} \Phi_i(X) c_j \quad (5.26)$$

ce qui permet d'avoir une approximation du champ de vecteur par

$$v(X) = \lim_{t \rightarrow 0} \frac{(\mathcal{K}_t - I)}{t} \Phi_i(X) C. \quad (5.27)$$

Avec $C = \{c_j\}_{j=1, \dots, k}$. Reste encore à pouvoir appliquer Koopman sur les observables des hamiltoniens effectifs. Nous avons pour cela explorer des pistes comme la connexion entre l'opérateur de Koopman et la structure symplectique de l'observable ou la théorie de l'information en lien avec la renormalisation.

Relation avec les formes symplectiques

Étant donné la dualité entre l'opérateur de Koopman et l'opérateur de Perron-Frobenius, il est assez clair que des connexions existent entre l'observable et les 2-forme symplectique des systèmes hamiltoniens. C'est en creusant cette idée que nous avons découvert le lien entre le hamiltonien du système et le noyau de l'opérateur de Koopman. Ce lien est explicitement développé dans [KKB19] et dont je rappelle ici les grandes lignes.

Partons de l'observable g des états, paramétrée en termes de positions q et de momenta p . L'évolution temporelle de l'observable est décrite par la dérivée totale par rapport au temps :

$$\frac{d}{dt}g = \frac{\partial g}{\partial x} \frac{\partial x}{\partial t} = \frac{\partial g}{\partial q} \frac{\partial q}{\partial t} + \frac{\partial g}{\partial p} \frac{\partial p}{\partial t} + \frac{\partial g}{\partial t}. \quad (5.28)$$

Le dernier terme disparaît, si nous supposons qu'il n'y a pas de dépendance temporelle explicite. Nous pouvons identifier les termes restants avec le crochet de Poisson (canonique).

$$\frac{d}{dt}g = \{g, \mathcal{H}\}. \quad (5.29)$$

Ainsi, l'évolution de tout observable sans dépendance temporelle explicite dans un système hamiltonien est donnée par son crochet de Poisson avec le hamiltonien. Nous savons de plus décrire la dynamique de l'observable par l'opérateur de Koopman

$$\frac{d}{dt}g = \mathcal{K}g = \{g, \mathcal{H}\}. \quad (5.30)$$

Pour les fonctions propres on obtient alors que

$$\frac{d}{dt}\phi = \{\phi, \mathcal{H}\} = \lambda\phi. \quad (5.31)$$

Ainsi, l'hamiltonien est une fonction propre de l'opérateur de Koopman associé à $\lambda = 0$, ce qui découle de

$$\frac{d}{dt}\mathcal{H} = \{\mathcal{H}, \mathcal{H}\} = \lambda\mathcal{H} = 0 \quad (5.32)$$

pour $\phi = \mathcal{H}$ et où le crochet de Poisson disparaît en raison de la structure symplectique. En appliquant du coarse-graining il serait donc possible de construire les hamiltoniens effectifs de cette manière. Cela reste encore à développer.

Géométrie de l'information

Enfin, la dernière piste explorée est celle de la théorie de l'information. En effet, le concept même de la renormalisation qui explore les interactions et dépendances (corrélations) en supprimant celles non nécessaires lors du passage à l'échelle, laisse suggérer une réflexion autour de la quantité d'information. Les articles sur lesquels j'ai appuyé cette réflexion intègrent la théorie de l'information et sa version géométrique [Mar20] en quantifiant la distance de séparation (distinguabilité) entre des paramètres des densités de probabilités des états pour chaque action de l'opérateur de renormalisation. Cela correspond à la thèse de A. Raju [Raj18] couplée à ses travaux initiaux [RMS18]. Ces travaux montrent en particulier le caractère contractant du coarse-graining sur les paramètres peu influant des densités pour chaque action de l'opérateur.

L'approche développée par les auteurs de [BO15] repose sur la distinguabilité entre les états eux-mêmes et met en lien le champ de vecteur sur l'espace des hamiltoniens et le champ de vecteur des états renormalisés. En particulier, les auteurs montrent l'existence d'un opérateur Ω (push-forward) entre les vecteurs tangents de l'espace des hamiltoniens et l'espace des états qui agit sur l'opérateur d'évolution de type Perron-Frobenius. Pour un état $\rho = e^H/Z$ de fonction de partition Z , et pour deux vecteurs tangents x et y en ρ

$$\langle x, y \rangle_\rho = \frac{\partial^2}{\partial t \partial s} S(\rho + tx + sy | \rho) \Big|_0 \quad (5.33)$$

avec $S(\rho || \rho') = \text{Trace} [\rho'(\log \rho' - \log \rho)]$ l'entropie relative quantifiant la distinguabilité des états ρ et ρ' . La distance entre deux vecteurs tangents s'approxime alors par

$$\text{Trace} [x \Omega_\rho^{-1}(y)]. \quad (5.34)$$

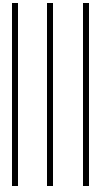
De fait la norme d'un vecteur a de l'espace des hamiltoniens permet de mesurer la distinguabilité entre les états renormalisés. Des indications de construction de l'opérateur ou du champ de vecteurs sur l'espace des hamiltoniens sont ensuite discutés, ce sont là encore des pistes qui nécessitent « un peu de travail ».

5.4. Contributions scientifiques de cette partie

Bilan scientifique



- 1 Article de conférence Internationale à comité et acte de lecture [Emi+18]



Bilan et projet

6

Vers l'infini et l'au-delà !



Nous l'avons vu tout au long de ce manuscrit, je suis très transversal aussi bien dans les thématiques que pour les outils employés ou développés pour mener à bien mes recherches. Depuis mon intégration au sein du DISP, le caractère de « service public » a largement motivé mes actions. Mon profil initial de traiteur de signal puis de mathématiques appliquées me permet d'aborder les thématiques d'aide à la décision, de pilotage ou d'information pour les systèmes de

production, sous un angle très différent que celui adopté par une majorité de la communauté de Génie Industriel.

Je pense avoir montré mon intérêt pour les questions fondamentales, ma capacité à répondre à ces questions, et surtout ma capacité à transférer technologiquement les avancées théoriques. Le temps ne s'arrêtant pas, ma curiosité restante encore intacte, je participe ou suis à l'initiative de nouveaux projets dont je vais maintenant décrire la motivation ainsi que mon positionnement.

Sommaire

6.1	Dans la continuité	84
6.2	Le Projet de Gestion de crises humanitaires	84
6.2.1	Modélisation d'une crise	86
6.2.2	Réseau temporel - Signature et Résilience	87
6.3	Le Projet Jumeau Numérique des urgences.	88
6.4	Résolution de problèmes de recherche opérationnelle par l'approche des flots sur les variétés	92

6.1. Dans la continuité

Il est évident que je vais et dois publier les différents résultats décrits dans le chapitre 5 dans un premier temps. Ensuite, il me reste tout de même à consolider certaines idées. Par exemple, en suivant la méthodologie décrite dans [MGD02] qui estime la cascade multiplicative sous l'hypothèse de probabilité de changement d'échelle gaussienne et non plus fixe. Ainsi, la probabilité p vue en 5.2.2 est interprétée comme une réalisation d'une loi à déterminer. Sous le prisme du Gaussian Mixture Model (GMM) il est possible d'estimer pour chaque agrégation (chaque niveau de la cascade) une distribution empirique de la probabilité p que nous verrons comme une variable aléatoire r . Pour chaque niveau de la cascade et sous l'hypothèse GMM alors la loi s'écrit

$$p(r) = \sum_{c=1}^N \pi_c \mathcal{N}(r|\mu_c, \sigma_c) \quad (6.1)$$

où μ_c est la moyenne, σ_c la variance (à estimer) et

$$\sum_{i=1}^N \pi_i = 1 \quad (6.2)$$

avec N le nombre de gaussiennes composant le modèle de mixture. On notera qu'une distribution de dirac est vue comme une gaussienne de variance presque nulle, soit

$$\delta(r) = \lim_{\sigma \rightarrow 0} \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{r^2}{2\sigma}} \quad (6.3)$$

ainsi, une distribution constante apparaissant (hypothèse d'une probabilité fixe) lors d'une échelle n'aura aucun effet sur l'algorithme.

Il est aussi nécessaire pour moi de terminer l'approche liée au groupe de renormalisation pour la caractérisation organisationnelle des urgences.

Lors d'un moment de oisiveté, j'aimerais aussi pouvoir continuer les travaux du chapitre 3, tant du côté de la dilatation et des parcours que du côté de la topologie des processus.

Je suis également impliqué dans de nouveaux projets débutants ou en cours. Je vais en faire état à présent.

6.2. Le Projet de Gestion de crises humanitaires

Chaque année, il y a en moyenne 150 millions d'êtres humains affectés par l'une des 400 crises humanitaires recensées. Ces crises sont le résultat de catastrophes naturelles telles que les tremblements de terre, les inondations ou encore les sécheresses. Les mouvements de population dûs à certains conflits armés sont eux aussi considérés comme des crises humanitaires majeures où bien souvent les états sont démunis face à l'afflux de réfugiés.

Les opérations logistiques dans le contexte de la réponse à une crise humanitaire sont

très différentes de ce qui se fait généralement dans le secteur privé. Différents auteurs se sont penchés sur la description de ce contexte particulier [Olo10]. La complexité de la situation lors d'une crise humanitaire rend sa vision globale difficile. Plusieurs facteurs exposés par exemple dans le Grand Bargain d'Istanbul en 2006 ou dans [Kun+17], expliquent cette complexité :

- Les données disponibles en temps réel sont souvent incomplètes et dynamiques. Ce travail de récolte et de mise à jour de données prend du temps et n'est donc pas toujours traité en priorité ;
- Les acteurs sont très variés (des petites ONG locales aux associations internationales, l'armée, le gouvernement, les transporteurs, les donateurs, etc). Leurs moyens et capacités de coordination ne sont pas toujours alignés ;
- Les appels de fonds, et donc le financement de la réponse, sont gérés en parallèle des opérations. La réponse humanitaire apportée aux personnes affectées dépend de la disponibilité de ces fonds. Le périmètre des actions menées et le nombre d'acteurs impliqués évoluent en conséquence ;
- L'aspect humain est prépondérant. Quelques jours de retard sur les livraisons peuvent avoir un impact en termes de vies humaines. Une inégalité de traitement entre différentes régions touchées peut générer des mouvements de population, donc plus de camps à ouvrir pour gérer ces migrants.

De ce fait, des problèmes de coordination pour la réponse sont souvent constatés [CLT10] ceci sans compter sur les choix d'actions de plus en plus guidés par la recherche de l'efficacité du rapport coût/vies sauvées qui conduit malheureusement à observer des sur-concentrations et un recouvrement d'interventions dans certaines zones, alors que d'autres sont délaissées.

A cela s'ajoute la dynamique du système. En effet, au cours même de la réponse à une crise humanitaire, la situation sur le terrain évolue. On peut citer entre autres, les répliques de tremblement de terre qui peuvent générer l'impraticabilité de certaines routes ou autres infrastructures logistiques, avant que la situation ne se débloque progressivement. De même, les besoins des personnes affectées peuvent évoluer au cours du temps, notamment lorsque les crises se complexifient (épidémies de choléra quelques semaines après un tremblement de terre). L'écosystème des acteurs est lui aussi sujet à évolution avec par exemple le retrait d'un acteur, la destruction de locaux ou encore l'endommagement du matériel. On comprend alors la difficulté d'atteindre des niveaux d'efficacité et d'efficacité optimaux pour la coordination d'une réponse.

Dans le contexte de la réponse à une crise humanitaire, qui nous l'avons vu, fait interagir un nombre important d'acteurs hétérogènes avec l'environnement logistique complexe, un indicateur de vulnérabilité doit être proposé afin de garantir la meilleure caractérisation de la vulnérabilité de la population, et ainsi permettre la meilleure résilience possible. C'est ce que l'équipe du DISP propose en abordant la question de la caractérisation de la vulnérabilité et de la coordination des acteurs par une approche de réseau de réseaux.

Le projet de recherche s'articule autour de deux grands axes, l'un porté par la modélisation d'une crise humanitaire par l'approche des réseaux de réseaux et de son exploitation topologique, alors que le second se concentre sur la dynamique du réseau pour la recherche de signatures, de résilience et d'aide à la décision.

6.2.1. Modélisation d'une crise

L'approche par réseaux est largement employée depuis maintenant quelques décennies pour traiter les interactions complexes des systèmes multi-agents, pour la modélisation des infrastructures routières, pour la représentation des interactions sociales ou encore pour les diffusions épidémiques ou les systèmes biologiques, mais finalement très peu pour la modélisation du système complexe que représente la réponse à une crise humanitaire ou l'environnement logistique d'une population [Gut+20].

Travail préliminaire

Sous le prisme de la crise humanitaire d'Haïti en 2010, un travail préliminaire [CBC22] a permis à l'équipe de constater l'intérêt d'une telle approche pour l'identification des vulnérabilités. A titre d'exemple donc, l'environnement logistique de la population d'Haïti après le tremblement de terre a été modélisé. L'infrastructure routière avec ses villes, les infrastructures aériennes, médicales ainsi que les ONG présentes sur site, ont été considérées dans un premier temps et ceci de manière isolée. Chacun de ces éléments représentent une couche particulière, *i.e.* un réseau propre ou un graphe au sens mathématique.

La représentation de l'environnement logistique complexe lors d'une crise humanitaire par l'approche des réseaux multi-couches permet rapidement de définir les communautés, *i.e.* les acteurs ou infrastructures étant fortement connectés entre eux pouvant signifier une faible vulnérabilité de la population habitant dans les villes associées. C'est par exemple le cas de Port-au-Prince qui a une forte connexion routière, la présence d'infrastructures aériennes et médicales et où certains ONG sont déjà présentes. D'autres villes du nord d'Haïti sont quant à elles beaucoup moins connectées, elles ont entre autres un faible nombre de structures médicales, pas de transport aérien ou peu de connexions routières. Elles apparaissent donc comme isolées et la population est considérée comme plus vulnérable.

Analyse topologique du réseau

De cette approche par réseaux (complexes) multi-couches, et outre l'aspect visuel, nous pouvons déjà apporter des éléments d'analyse quantitatif de fragilité par le calcul des centralités [NW99]; [Gut+20]. L'ensemble de ces premiers éléments nous permettrons déjà de consolider un indicateur :

- Permettant une analyse non plus segmentée et centrée sur une zone et son microcosme associé, mais en permettant une analyse plus globale et systémique;
- Permettant de mettre en exergue les zones non couvertes par les ONG;
- Donnant une indication sur les zones ("exo-zones") où des ressources logistiques doivent être mises en place par les humanitaires (stockage, transporteur de la capitale, etc).

L'identification de la vulnérabilité d'une zone est donc déterminée par un indicateur lié au calcul des centralités, mais ne se limite pas uniquement à ce dernier. Une analyse topologique précise du réseau par des approches d'homologies persistantes [Dug+19]; [BDM19a]; [MM20] la détermination de la connectivité du graphe par l'analyse du noyau de la matrice Laplacienne associée, ou encore la détermination et l'existence de composantes géantes [Li+19] sous le formalisme des réseaux

complexes sont autant d'indicateurs qui composeront la détermination précise de cet indicateur de vulnérabilité.

6.2.2. Réseau temporel - Signature et Résilience

Si nous considérons de plus les questions opérationnelles de préparation et d'intervention face à une crise, un ensemble de méthodes devront être proposées afin de simuler au mieux les scénarios liés aux causes de la crise (tremblement de terre, ouragan, conflits armés, etc). Les simulations seront enrichies dès lors qu'une signature des crises sera identifiée par son impact sur le réseau. En conjonction, de nouvelles méthodes d'estimation de la résilience du réseau devront également être proposée afin de permettre en amont d'une crise de proposer les meilleures préconisations logistiques et ainsi préparer au mieux les interventions. La question de la signature d'une crise ainsi que son caractère résilient pour la population sont des propriétés qui nécessitent d'introduire une notion de temporalité dans le réseau lui-même. D'ailleurs, depuis maintenant quelques années, un certain nombre de recherches traitent des réseaux complexes variables dans le temps.

Signature temporelle

La recherche d'une signature d'une crise, est un problème évidemment peu abordé dans la littérature, et nous proposons par exemple de déterminer cette signature comme une forme particulière de trajectoire sur l'espace des matrices symétriques semi-définies positives, représentant un cône convexe. Cette idée de forme peut être traitée par des approches liées aux espaces des formes [DBM18b] ou comme une trajectoire issue d'un système complexe dynamique [PBC21] ou enfin comme une morphologie directe de la crise, *i.e.* en créant une représentation informative de plus haut niveau du réseau dans un espace de morphologie [Car03].

Résilience

Des mécanismes de détermination de la résilience d'un réseau ont déjà été abordés dans la littérature dans différents domaines souvent liés à la distribution de services, comme la logistique, le transport aérien, ou encore la téléphonie mobile mais très peu, encore une fois, ne s'attaquent à la problématique de réseaux multi-couches dans le contexte d'une crise humanitaire. Ici aussi, l'équipe DISP souhaite enrichir les travaux proposés dans [Li+19] dans le contexte multi-couches en exploitant l'aspect topologique et élastique du réseau.

Modélisation du besoin dans un réseau

Finalement, lors de la gestion opérationnelle d'une crise, un organisme humanitaire peut être amené à choisir au mieux les quantités de ressources à distribuer et bien sûr à déterminer les lieux de stockage et de distribution en conséquence. Afin de permettre une meilleure aide à la décision, le besoin des populations, que nous traduirons par une demande, doit concrètement être matérialisé. Par l'approche réseau complexe multi-couches les communautés et a contrario les éléments isolés sont déterminés par un calcul topologique lié directement aux connexions et à la distance dans la représentation. Cette distance est dans notre exemple [CBC22] obtenue par l'emploi de l'algorithme de Force-layout Atlas2. D'autres algorithmes de force-layout ont déjà été proposés dans la littérature, mais ces derniers ne reflètent pas l'information liée

à la demande ou au besoin. Un des derniers objectifs de ce projet consiste donc à proposer un nouvel algorithme de force-layout capable de traduire la force d'attraction des zones en fonction du besoin. La conséquence directe d'une telle méthode, sera de permettre le dimensionnement en tonne et volume des ressources à allouer et pouvoir ensuite déterminer le positionnement des hubs logistiques.

A. Charles, C. Cherifi et moi-même avons obtenu un financement de l'école doctorale pour mener ce projet. Je serai le directeur de thèse de **Johan Leydet** dès mon HDR obtenue.

6.3. Le Projet Jumeau Numérique des urgences

Ce projet est un projet en cours et en collaboration avec H. Bril El Haouzi, W. Derigent du CRAN et T. Moyaux, V. Cheutet, L. Trilling du DISP. Une partie de la réponse à AAP ANR générique est reprise ici.

Une étude menée par le DISP via le projet HRP3 du Cluster GOSPI, a eu pour objectif de tester et de valider différentes stratégies de pilotage, afin de minimiser les durées de séjour et les temps d'attente des patients aux urgences. Pour ce faire un recueil de données sur le terrain a permis d'analyser les causes des temps d'attente en mesurant de façon précise les éléments suivants :

- Temps moyen de prise en charge total
- Temps moyen avant examen médical
- Temps moyen de l'examen médical
- Temps moyen après examen avant fin de prise en charge
- Temps moyen de sortie
- Temps moyen de la disponibilité des lits
- Temps moyen Examen complémentaire par type d'examen (Radiologie, scanner, Echographie, etc)
- Temps moyen avis spécialiste par type de spécialité

La complexité organisationnelle du SAU ainsi que sa nécessaire dépendance à la disponibilité des lits de l'hôpital, ne pouvait se comprendre ni s'expliquer par une telle étude. Si nous percevons bien que le temps de séjour aux urgences est un, voire même le paramètre à monitorer pour un pilotage efficace du SAU, il semble obligatoire de pouvoir modéliser ce système sous une forme dynamique permettant ainsi de modéliser les décisions prises par les médecins, internes, externes ou autres infirmières ainsi que leurs interactions.

Le fonctionnement actuel du SAU de Saint-Etienne repose sur la prise de décision décentralisée car chaque soignant sélectionne son prochain patient d'après le tableau de bord du SAU (Figure 6.1).

En première approximation (donc, sans distinguer ni internes, externes et médecins séniors, ni SAU adulte ou pédiatrique), ce choix repose sur le niveau de gravité du patient :

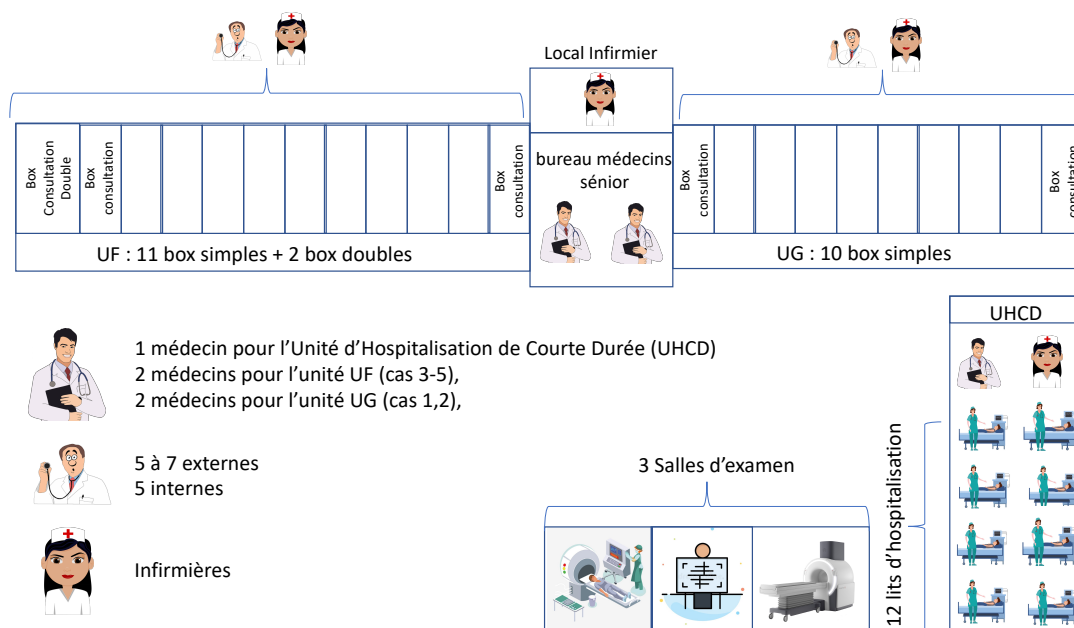


Figure 6.1: Un idée du serve d'accueil des urgences du CHU de Saint-Etienne.

- Niveaux de gravité 1 et 2 : Ces patients sont les plus urgents, donc l'Infirmier d'Accueil et d'Orientation doit leur trouver un médecin au plus vite. Pour cela, cet infirmier cherche un médecin dans leur bureau ou, s'ils sont déjà tous en consultation, dans les box ; dans ce second cas, la consultation en cours sera mise en pause pour que le médecin traite ce cas plus urgent.
- Niveaux de gravité 3 à 5 : Ces patients vont attendre en salle d'attente jusqu'à ce qu'un soignant les prenne en charge en faisant un compromis (propre à chaque soignant) entre niveau de gravité et respect de l'ordre d'arrivée. Plus précisément, les différentes catégories de soignants consultent toutes le même tableau de bord dans lequel elles sélectionnent leur prochain patient en fonction des soins nécessaires, condition du patient, etc.

L'optimisation du parcours patient est un sujet qui suscite un intérêt grandissant grâce aux approches orientées-données comme le Machine Learning (Deep Learning, Reinforcement Learning, etc.) [Cur+18]. D'autres approches abordent cette problématique en intégrant des systèmes multi-agents à des workflows [Ajm+19]. En secteur hospitalier, on appelle agent toute entité indépendante, intelligente et dynamique qui est amenée à prendre des décisions dans un cadre particulier où s'appliquent des règles. Les médecins, infirmières, patients et manipulateurs sont par conséquent qualifiés d'agents. Dans ce contexte, les agents sont en permanence amenés à interagir entre eux et à apprendre des expériences passées influençant ainsi la prise de décisions, tout comme l'environnement. Dans ce cas de figure, il s'agit donc de traiter le problème de manière décentralisée. A cet égard, un SAU est un cas particulier de système de production de services qui contraste avec les systèmes de production de biens et de services en général. En effet, ceux-ci sont traditionnellement contrôlés par un système de pilotage centralisé, c.-à-d. par une autorité centrale (AC) qui (i) reçoit des informations des agents (c.-à-d., les opérateurs d'une ligne de productions

ou, dans notre cas, les soignants), (ii) choisit quoi faire pour atteindre les objectifs du système de production et (iii) communique cette décision aux opérateurs. L'avantage d'un tel système de pilotage centralisé est son efficacité en l'absence de perturbations. Son inconvénient est son manque de réactivité lors des imprévus.

Dans ce contexte, l'objectif du projet est de proposer un jumeau numérique pour le SAU permettant à la fois de visualiser le comportement du service en temps quasi-réel mais aussi de prévoir et anticiper le comportement en ajoutant un moteur de simulation couplé à de l'optimisation permettant de répondre aux aléas inhérents à ce type de services, afin « d'asservir » et de contrôler l'indicateur temps de passage aux urgences.

Vu le comportement fortement dynamique du système complexe SAU, une étude des coopérations, synchronisations ou consensus entre les agents suscite toujours un grand intérêt, surtout du point de vue théorique. L'étude des systèmes dynamiques multi-agents fait appel à des branches des mathématiques diverses (théorie des graphes, topologie, processus stochastiques, théorie des jeux ...). Beaucoup de travaux se sont focalisés sur une caractérisation algébrique via le plus souvent une étude spectrale de la matrice Laplacienne du graphe représentant les interactions entre les agents. Je souhaite ici donner des invariants géométriques et/ou topologiques tels qu'ils ont été introduit par exemple par [Mun13]; [Mar18] pour caractériser le système coopératif entre les agents dans leur prise de décision. Ces approches sont en rupture par rapport aux études traditionnelles car l'information géométrique des interactions entre les agents est étudiée au moyen de la géométrie différentielle ou de l'homologie persistante. Le changement d'états des agents est donc régi par la structuration et les contraintes intrinsèques des interactions sur des variétés différentielles.

J'ai été complètement passionné par l'approche d'A. Cardon dans [Car04] qui relie le système multi-agents aux systèmes hamiltoniens. C'est une source réelle d'inspiration qui me permet de retrouver les outils vus par exemple dans le chapitre 5 au travers de la renormalisation. Ainsi, dans le projet Jumeau Numérique pour les urgences, je me positionne centralement entre la simulation pilotée par les données et le pilotage hybride avec les verrous scientifiques proposés comme suit :

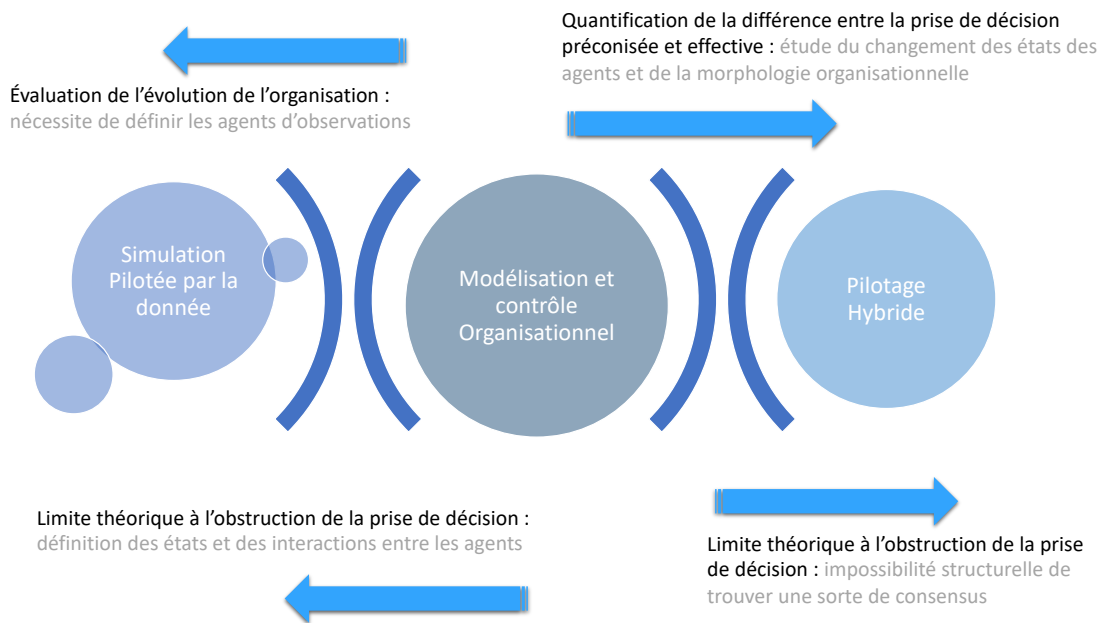


Figure 6.2

Modélisation et contrôle organisationnel

États des agents et de la morphologie organisationnelle



Approche par les systèmes dynamiques à interactions massives

- ↔ extraire une information topologique (homologie) déformation topologique
- ↔ extraire une information d'évolution du système et de ses régimes coordonnés ou non



Opérateurs Liouville, Hamiltonien, variété symplectique, groupe de renormalisation, entropie

Impossibilité structurelle de trouver une sorte de consensus

- ↔ Étude topologique et géométrique de l'espace des états (vue comme un espace des phases du système)



Homotopie, homologie, gradients sur les variétés pour définir les obstructions au consensus

6.4. Résolution de problèmes de recherche opérationnelle par l'approche des flots sur les variétés

Étant depuis 2016 co-responsable de l'axe Pilotage des Systèmes de Production (PSP) d'abord puis Gestion et Optimisation des Opérations (G20) ensuite du DISP, j'ai pu découvrir les problématiques de recherche opérationnelle. Parmi les problématiques traitées par certains membres de l'axe, les tournées de véhicules et les problèmes d'affectation ont attiré plus particulièrement mon attention.

Le problème du voyageur de commerce est un problème NP-complet extrêmement connu et qui suscite un intérêt majeur depuis des décennies. Ce problème d'optimisation apparaît dans de nombreux champs d'application comme pour la construction de cartes du génome, la gestion des télescopes, les tournées d'infirmières, etc. Dû à sa grande complexité, de nombreuses heuristiques ont été proposées par le passé. N'étant pas vraiment familier avec les algorithmes développés en programmation linéaire, j'ai cherché à comprendre le problème puis à savoir s'il n'était pas possible de pouvoir utiliser des algorithmes continus de descente de gradient pour répondre à ce problème. Évidemment il en existe ... ces approches proposent une relaxation continue du problème discret sous l'angle des systèmes dynamiques, que nous retrouvons donc encore une fois dans cette thématique !

Je suis rapidement tombé sur les travaux précurseurs de R. Brockett [BW91]; [Bro91a], d' A. Bloch [Blo90a]; [Blo90b] puis de W. Wong [Won95]; [Won05] qui tous formalisent le problème du Traveling Salesman Problem (TSP) sous la forme matricielle

$$P_{opt} = \arg \min_{P \in \mathcal{P}} \text{Trace} (D^T P^T A P) \quad (6.4)$$

où la matrice D est la matrice de distance entre les villes, la matrice A , la matrice d'adjacence du graphe complet associé au problème et où P est la matrice de permutation des villes. Ainsi la recherche de la solution du problème est déportée à la recherche d'une matrice de permutation P dans l'espace des matrices de permutation ou en concordance, dans l'espace des matrices de rotation $SO(n)$. Une piste que je dois explorer consiste à intégrer les matrices de permutation dans l'espace des matrices doublement stochastiques. Les avantages, dans ce cas, découlent du fait que les matrices de permutation sont les points extrêmes de l'ensemble des matrices doublement stochastiques et l'ensemble des matrices doublement stochastiques est l'enveloppe convexe de l'ensemble des matrices de permutation [BW91]. Bien entendu, $SO(n)$ n'est pas un polytope et n'a pas de sommets. Il s'agit cependant d'un groupe contenant l'ensemble des matrices de permutation de déterminant 1 comme sous-ensemble et sous-groupe. C'est aussi une variété compacte sans bord et les flux de gradients dans cette variété ont toujours au moins un point stationnaire stable. L'analogie avec les systèmes dynamiques provient de la recherche même de P sur cette variété. En effet, on a facilement que

$$\dot{P} = -\nabla f(P) \quad (6.5)$$

avec $f(P) = \text{Trace} (D^T P^T A P)$ et $\dot{P} = \lim_{t \rightarrow 0} \frac{P_{k+1} - P_k}{t}$ qui constitue bien un système dynamique, un flot sur la variété $SO(n)$. De fait je me suis tout d'abord intéressé à ce

groupe pour regarder les performances d'estimation guidé par les travaux de [ZP08]; [Sah+19] mais cela reste encore à approfondir car les premiers résultats obtenus montrent des problèmes de convergence.

Outre les aspects de recherche de matrice de permutation par double crochet de Lie comme énoncé par R. Brockett dans ses différents travaux, des résultats un peu plus récents d'A. Bloch et notamment ceux de [BGR13] m'intéressent particulièrement. Dans ce travail A. Bloch fait une jointure entre le flot de gradient défini par un double crochet de Lie et les équations d'un système de Toda. Le système de Toda est un système hamiltonien qui modélise le mouvement de n particules ayant une interaction exponentielle en fonction des distances les caractérisants sur un axe. Le hamiltonien est connu et s'écrit

$$H(x, y) = \frac{1}{2} \sum_{k=1}^n y_k^2 + \sum_{k=1}^{n-1} e^{x_k - x_{k+1}}. \quad (6.6)$$

Or il est montré que par le changement de variable de Flaschka [Fla74], le flot défini par le double crochet de poisson devient un flot de Lax dont les coefficients respectent les mêmes relations que l'estimation des parcours vus au chapitre 3. La relation est

$$\dot{L} = [B, L] = BL - LB \quad (6.7)$$

où L, B sont des matrices issues du changement de variable. Si $L \in \text{SO}(n)$ alors le système est complètement solvable. Ce formalisme permet d'être une passerelle importante avec les orbites co-adjointes des matrices de Lax L et donc nos matrices de permutations sous l'hypothèse que $P \in \text{SO}(n)$.

On peut voir ici encore comment à partir des approches des systèmes dynamiques, il est possible d'aborder un grand nombre de problématiques. Ayant déjà investi pas mal d'énergie dans l'apprentissage et la compréhension des problèmes liés aux orbites co-adjointes, les 2-forme Kostant-Kirilov-Souriau, la méthode des moments dans le contexte de géométrie différentielle, j'aimerais pouvoir maintenant les utiliser à bon escient.

Références

- [Des] F. Desbouvries. “Geometrical aspects of linear prediction algorithms”. In: (). url: <http://www-public.it-sudparis.eu/~desbouvr/papiers-web/mtns.pdf> (visited on 06/13/2016).
- [EM] Herbert Edelsbrunner and Dmitriy Morozov. *Persistent Homology: Theory and Practice*.
- [Get] R. K. Getoor. “The shift operator for non-stationary stochastic processes”. In: *Duke Math. J.* 23.1 (), pp. 175–187. issn: 0012-7094, 1547-7398. doi: 10.1215/S0012-7094-56-02316-X. url: <http://projecteuclid.org/euclid.dmj/1077466687> (visited on 01/16/2017).
- [HS] T. Hastie and W. Stuetzle. “Principal curves”. In: *Journal of the American Statistical association* 84.406 (), pp. 502–516.
- [CBC22] Aurélie Charles, Guillaume Bouleux, and Chantal Chérifi. “Understanding the situation at a glance : the powerful value added of complex networks to analyse humanitarian operations”. In: *International journal of disaster risk reduction* (2022).
- [Liu+22a] Yinling Liu et al. “An agent-based architecture of the Digital Twin for an Emergency Department”. In: *Health and Technology* (2022).
- [Liu+22b] Yinling Liu et al. “Hybrid Simulation Modeling of Emergency Departments for Resource Scheduling”. In: *Heal. Syst.* (2022).
- [Mor+22] Benoît Morel et al. “Evaluating the increased burden in Adult Emergency Departments of cardio-respiratory illness visits during flu and bronchiolitis outbreaks in the pediatric population by a retrospective multicentric time series analysis”. In: *JMIR Public Health and Surveillance* (2022). doi: 10.2196/25532.
- [PBC22] Clément Péalat, Guillaume Bouleux, and Vincent Cheutet. “Using a manifold-based approach to extract clinical codes associated with winter respiratory viruses at an emergency department”. In: *Health Care Management Science* (2022).
- [BSH21] Petar Bevanda, Stefan Sosnowski, and Sandra Hirche. “Koopman Operator Dynamical Models: Learning, Analysis and Control”. In: 871295 (2021). arXiv: 2102.02522. url: <http://arxiv.org/abs/2102.02522>.
- [Guo+21] Lihong Guo et al. “Assessing coupling dynamics from an ensemble of time series”. In: *Entropy* 17.4 (2021), pp. 1958–1970. issn: 10994300. doi: 10.3390/e17041958. arXiv: 1008.0539. url: <https://doi.org/10.1007/s41109-021-00370-x>.
- [Hir+21] Seth M. Hirsh et al. “Structured Time-Delay Models for Dynamical Systems with Connections to Frenet-Serret Frame”. In: 2021. arXiv: 2101.08344. url: <http://arxiv.org/abs/2101.08344>.
- [Ken+21] Ezra Kenny et al. “Patient flow simulation using historically informed synthetic data”. In: *Studies in Health Technology and Informatics*. Vol. 276. 2021, pp. 32–37. isbn: 9781643681689. doi: 10.3233/SHTI210007.
- [PBC21] Clément Péalat, Guillaume Bouleux, and Vincent Cheutet. “Improved Time Series Clustering Based on New Geometric Frameworks”. In: *Pattern Recognition, Elsevier* (2021). doi: 10.1016/j.patcog.2021.108423.
- [Cub+20] Cemre Cubukcuoglu et al. “A discrete event simulation procedure for validating programs of requirements: The case of hospital space planning”. In: *SoftwareX* 12 (2020). issn: 23527110. doi: 10.1016/j.softx.2020.100539.
- [DTK20] Felix Dietrich, Thomas N. Thiem, and Ioannis G. Kevrekidis. “On the koopman operator of algorithms”. In: *SIAM Journal on Applied Dynamical Systems* 19.2 (2020), pp. 860–885. issn: 15360040. doi: 10.1137/19M1277059. arXiv: 1907.10807.
- [Gut+20] Walter J Gutjahr et al. *Innovative approaches in humanitarian operations*. 2020. doi: 10.1007/s00291-020-00598-6. url: <https://doi.org/10.1007/s00291-020-00598-6>.

- [Isf+20] M. Nasr Isfahani et al. “Decreased Emergency Department Overcrowding by Discharge Lounge: A Computer Simulation Study”. In: *International Journal of Preventive Medicine* 11 (2020).
- [JLR20] A. Javed, B. Suk Lee, and D.M. Rizzo. “A benchmark study on time series clustering”. In: *Machine Learning with Applications* 1 (2020), p. 100001. issn: 2666-8270. doi: <https://doi.org/10.1016/j.mlwa.2020.100001>. url: <http://www.sciencedirect.com/science/article/pii/S2666827020300013>.
- [Kam+20] Mason Kamb et al. “Time-delay observables for koopman: Theory and applications”. In: *SIAM Journal on Applied Dynamical Systems* 19.2 (2020), pp. 886–917. issn: 15360040. doi: 10.1137/18M1216572. arXiv: 1810.01479.
- [Mai20] L. B. Maier. “Sparse Data Interpolation and Smoothing on Embedded Submanifolds”. In: *Journal of Scientific Computing* 84.1 (2020). issn: 15737691. doi: 10.1007/s10915-020-01268-z.
- [Mar20] Charles Michel Marle. “On Gibbs states of mechanical systems with symmetries”. In: *Journal of Geometry and Symmetry in Physics* 57 (2020), pp. 45–85. issn: 13145673. doi: 10.7546/jgsp-57-2020-45-85. arXiv: 2012.00582.
- [MM20] Elizabeth Munch and Elizabeth Munch. “Persistent Homology of Complex Networks for Dynamic State Detection”. In: March (2020). arXiv: arXiv:1904.07403v2.
- [PBC20] Clément Pealat, Guillaume Bouleux, and Vincent Cheutet. “Improved time-series clustering with UMAP dimension reduction method”. In: *Proceedings - International Conference on Pattern Recognition*. Milan, 2020, pp. 5658–5665. isbn: 9781728188089. doi: 10.1109/ICPR48806.2021.9412261.
- [Red20] William T Redman. “Renormalization group as a Koopman operator”. In: *Physical Review E* 101.6 (2020), pp. 1–5. issn: 24700053. doi: 10.1103/PhysRevE.101.060104.
- [Ajm+19] Faten Ajmi et al. “Agent-based dynamic optimization for managing the workflow of the patient’s pathway”. In: *Simulation Modelling Practice and Theory* 96 (2019). issn: 1569190X. doi: 10.1016/j.simpat.2019.101935.
- [BB19] Guillaume Bouleux and Frédéric Barbaresco. “Dilation Operator Approach and Square Root Velocity Transform for Time/Doppler Spectra Characterization on SU(n)”. In: *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. Vol. 11712 LNCS. 2019, pp. 31–38. isbn: 9783030269791. doi: 10.1007/978-3-030-26980-7_4.
- [BDM19a] Guillaume Bouleux, Maël Dugast, and Eric Marcon. “Information Topological Characterization of Periodically Correlated Processes by Dilation Operators”. In: *IEEE Transactions on Information Theory* 65.10 (2019), pp. 6484–6495. issn: 15579654. doi: 10.1109/TIT.2019.2923217.
- [BDM19b] Guillaume Bouleux, Maël Dugast, and Eric Marcon. “Représentation et comparaison des processus périodiquement corrélés sur des groupes de Lie par les opérateurs de Dilatation”. In: *XXVII colloque GRETSI*. Lille, 2019, pp. 4–7.
- [Dug+19] Maël Dugast et al. “Improving Health Care Management Through Persistent Homology of Time-Varying Variability of Emergency Department Patient Flow”. In: *IEEE Journal of Biomedical and Health Informatics* 23.5 (2019), pp. 2174–2181. issn: 21682208. doi: 10.1109/JBHI.2018.2882748.
- [KKB19] Eurika Kaiser, J. Nathan Kutz, and Steven L. Brunton. “Discovering Conservation Laws from Data for Control”. In: *Proceedings of the IEEE Conference on Decision and Control* 2018-Decem.i (2019), pp. 6415–6421. issn: 07431546. doi: 10.1109/CDC.2018.8618963. arXiv: 1811.00961.
- [Li+19] Wenguo Li et al. “Maximizing Network Resilience against Malicious Attacks”. In: *Scientific Reports* 9.1 (2019), p. 2261. issn: 20452322. doi: 10.1038/s41598-019-38781-7. url: <http://www.nature.com/articles/s41598-019-38781-7>.

- [PBC19] Clement Pealat, Guillaume Bouleux, and Vincent Cheutet. “Extracting most impacting emergency department patient flow by embedding laboratory-confirmed and clinical diagnosis on the stiefel manifold”. In: *2019 IEEE EMBS International Conference on Biomedical and Health Informatics, BHI 2019 - Proceedings*. Chicago, 2019. isbn: 9781728108483. doi: 10.1109/BHI.2019.8834507.
- [Sah+19] Tuhin Sahai et al. “Continuous relaxations for the traveling salesman problem”. In: *Nonlinear Dynamics* 97.4 (2019), pp. 2003–2022. issn: 1573269X. doi: 10.1007/s11071-019-05092-5. arXiv: 1702.05224. url: <https://doi.org/10.1007/s11071-019-05092-5>.
- [Cur+18] Catherine Curtis et al. “Machine Learning for Predicting Patient Wait Times and Appointment Delays”. In: *Journal of the American College of Radiology* 15.9 (2018). issn: 1558349X. doi: 10.1016/j.jacr.2017.08.021.
- [DBM18a] Maël Dugast, Guillaume Bouleux, and Eric Marcon. “Information topological characterization for non-stationary random processes”. In: *Entropy 2018 – From Physics to Information Sciences and Geometry*. Barcelona, 2018.
- [DBM18b] Maël Dugast, Guillaume Bouleux, and Eric Marcon. “Representation and characterization of nonstationary processes by dilation operators and induced shape space manifolds”. In: *Entropy* 20.9 (2018), p. 717. issn: 10994300. doi: 10.3390/e20090717. url: <http://www.mdpi.com/1099-4300/20/9/717>.
- [Emi+18] Sid Ahmed Emine et al. “Emergency Department Overcrowding Detection by a Multifractal Analysis”. In: *IFAC Symposium on Biological and Medical Systems*. Vol. 51. 27. Sao paulo, 2018, pp. 104–109. doi: 10.1016/j.ifacol.2018.11.659.
- [Mar18] Johan Markdahl. “A geometric obstruction to almost global synchronization on Riemannian manifolds”. In: *arXiv August 2018* (2018). issn: 23318422. arXiv: 1808.00862.
- [MHM18] Leland McInnes, John Healy, and James Melville. *UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction*. 2018. url: <http://arxiv.org/abs/1802.03426>.
- [Pia+18] Marco Piangerelli et al. “Topological classifier for detecting the emergence of epileptic seizures”. In: *BMC research notes* 11.1 (2018), p. 392.
- [Por+18] Alexander Port et al. “Persistent Topology of Syntax”. en. In: *Mathematics in Computer Science* 12.1 (Mar. 2018), pp. 33–50. issn: 1661-8270, 1661-8289. doi: 10.1007/s11786-017-0329-x. url: <https://link.springer.com/article/10.1007/s11786-017-0329-x> (visited on 05/06/2018).
- [Raj18] Archishman Raju. “Aspects of the Renormalization Group”. PhD thesis. 2018.
- [RMS18] Archishman Raju, Benjamin B Machta, and James P Sethna. “Information loss under coarse graining: A geometric approach”. In: *Physical Review E* 98.5 (2018), pp. 1–6. issn: 24700053. doi: 10.1103/PhysRevE.98.052112. arXiv: 1710.05787.
- [Sol+18] Gregory Soler et al. “Emergency Department Admissions Overflow Modeling by a Clustering of Time Evolving Clinical Diagnoses”. In: *IEEE International Conference on Automation Science and Engineering*. Vol. 2018-Augus. 2018, pp. 365–370. isbn: 9781538635933. doi: 10.1109/COASE.2018.8560585.
- [BB17] Guillaume Bouleux and Rémy Boyer. “Sparse-based estimation performance for partially known overcomplete large-systems”. In: *Signal Processing* 139 (2017), pp. 70–74. issn: 01651684. doi: 10.1016/j.sigpro.2017.04.010.
- [Bri17] Alice Le Brigant. “Probability on the spaces of curves and the associated metric spaces using information geometry; radar applications.” PhD thesis. Université de Bordeaux, 2017.
- [Kun+17] Nathan Kunz et al. “Relevance of humanitarian logistics research: best practices and way forward”. In: *International Journal of Operations & Production Management* 37.11 (2017), pp. 1585–1599. issn: 0144-3577. doi: 10.1108/ijopm-04-2016-0202.
- [Le 17] A. Le Brigant. “A discrete framework to find the optimal matching between manifold-valued curves”. In: *arXiv preprint arXiv:1703.05107* (2017). url: <https://arxiv.org/abs/1703.05107> (visited on 09/12/2017).

- [OMN17] Colin O'Reilly, Klaus Moessner, and Michele Nati. "Univariate and Multivariate Time Series Manifold Learning". In: *Knowledge-Based Systems* 133 (2017), pp. 1–16. issn: 09507051. doi: 10.1016/j.knosys.2017.05.026. url: <http://dx.doi.org/10.1016/j.knosys.2017.05.026>.
- [Ott+17] N. Otter et al. "A roadmap for the computation of persistent homology". In: *EPJ Data Science* 6.17 (2017), pp. 1–38.
- [SBA17] Manel Saad Saoud, Abdelhak Boubetra, and Safia Attia. *A multi-agent based modeling and simulation data management and analysis system for the hospital emergency department*. 2017. doi: 10.4018/IJHISI.2017070102.
- [Ume17] Yuhei Umeda. "Time Series Classification via Topological Data Analysis". en. In: *Transactions of the Japanese Society for Artificial Intelligence* 32.3 (2017).
- [YF17] M. Yousefi and R. P.M. Ferreira. "An agent-based simulation combined with group decision-making technique for improving the performance of an emergency department". In: *Brazilian Journal of Medical and Biological Research* 50.5 (2017), pp. 1–10. issn: 16784510. doi: 10.1590/1414-431X20175955.
- [BBM16] M. Bauer, M. Bruveris, and P. W. Michor. "Why Use Sobolev Metrics on the Space of Curves". en. In: *Riemannian Computing in Computer Vision*. DOI: 10.1007/978-3-319-22957-7_11. Springer, Cham, 2016, pp. 233–255. isbn: 978-3-319-22956-0 978-3-319-22957-7. url: https://link.springer.com/chapter/10.1007/978-3-319-22957-7_11 (visited on 10/02/2017).
- [CLR16] Pablo G. Cámara, Arnold J. Levine, and Raúl Rabadán. "Inference of Ancestral Recombination Graphs through Topological Data Analysis". en. In: *PLOS Computational Biology* 12.8 (Aug. 2016), e1005071. issn: 1553-7358. doi: 10.1371/journal.pcbi.1005071. url: <http://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1005071> (visited on 05/06/2018).
- [Hau16] S. Hauberg. "Principal Curves on Riemannian Manifolds". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 38.9 (Sept. 2016), pp. 1915–1921. issn: 0162-8828, 2160-9292. doi: 10.1109/TPAMI.2015.2496166. url: <http://ieeexplore.ieee.org/document/7312494/> (visited on 12/14/2016).
- [KKS16] Stefan Klus, Péter Koltai, and Christof Schütte. "On the numerical approximation of the Perron-Frobenius and Koopman operator". In: *Journal of Computational Dynamics* 3.1 (2016), pp. 51–79. issn: 21582505. doi: 10.3934/jcd.2016003. arXiv: 1512.05997.
- [Le 16] A. Le Brigant. "Computing distances and geodesics between manifold-valued curves in the SRV framework". In: *arXiv preprint arXiv:1601.02358* (2016). url: <https://arxiv.org/abs/1601.02358> (visited on 09/06/2017).
- [Lü+16] Linyuan Lü et al. "Vital nodes identification in complex networks". In: (July 2016). doi: 10.1016/j.physrep.2016.06.007. arXiv: 1607.01134. url: <http://arxiv.org/abs/1607.01134>
<http://dx.doi.org/10.1016/j.physrep.2016.06.007>.
- [Min+16] H. Q. Minh et al. "Kernel Methods on Approximate Infinite-Dimensional Covariance Operators for Image Classification". In: *arXiv preprint arXiv:1609.09251* (2016). url: <https://arxiv.org/abs/1609.09251> (visited on 10/18/2016).
- [Nap16] A. Napolitano. "Cyclostationarity: New trends and applications". In: *Signal Processing* 120 (2016), pp. 385–408. issn: 0165-1684. doi: <http://dx.doi.org/10.1016/j.sigpro.2015.09.011>. url: <http://www.sciencedirect.com/science/article/pii/S0165168415003138>.
- [RH16] R. Roy and F. Harper. "Floquet topological phases with symmetry in all dimensions". In: *arXiv:1610.06899 [cond-mat]* (Oct. 2016). arXiv: 1610.06899. url: <http://arxiv.org/abs/1610.06899> (visited on 01/16/2017).
- [YHW16] W.-H. Yang, S. H. Holan, and C. K. Wikle. "Bayesian Lattice Filters for Time-Varying Autoregression and Time-Frequency Analysis". en. In: *Bayesian Analysis* 11.4 (Dec. 2016), pp. 977–1003. issn: 1936-0975. doi: 10.1214/15-BA978. url: <http://projecteuclid.org/euclid.ba/1445263834> (visited on 04/05/2017).

- [Zim16] Ralf Zimmermann. “A matrix-algebraic algorithm for the Riemannian logarithm on the Stiefel manifold under the canonical metric”. en. In: *arXiv:1604.05054 [math]* (Apr. 2016). arXiv: 1604.05054. url: <http://arxiv.org/abs/1604.05054>.
- [Ajm+15] I. Ajmi et al. “Mapping patient path in the Pediatric Emergency Department: A workflow model driven approach”. In: *Journal of biomedical informatics* 54 (2015), pp. 315–28.
- [BEG15] M. Bauer, M. Eslitzbichler, and M. Grasmair. “Landmark-guided elastic shape analysis of human character motions”. In: *arXiv preprint arXiv:1502.07666* (2015). url: <https://arxiv.org/abs/1502.07666> (visited on 09/27/2017).
- [BRA15] N. Benhajji, D Roy, and D Anciaux. “Patient-centered multi agent system for health care”. In: *IFAC-PapersOnLine*. Vol. 28. 3. Elsevier Ltd., 2015, pp. 710–714. doi: 10.1016/j.ifacol.2015.06.166. url: <http://dx.doi.org/10.1016/j.ifacol.2015.06.166>.
- [BO15] Cédric Bény and Tobias J. Osborne. “Information-geometric approach to the renormalization group”. In: *Physical Review A - Atomic, Molecular, and Optical Physics* 92.2 (2015), pp. 1–6. issn: 10941622. doi: 10.1103/PhysRevA.92.022330. arXiv: 1206.7004.
- [BB15] Guillaume Bouleux and Rémy Boyer. “Acquisition Comprimée intégrant une Connaissance Partielle du Modèle d’Observation: Analyse Asymptotique”. In: *XXV colloque GRETSI*. Lyon, 2015.
- [BMM15] Guillaume Bouleux, Eric Marcon, and Olivier Mory. “Early Index for Detection of Pediatric Emergency Department Crowding”. In: *Biomedical and Health Informatics, IEEE Journal of* 19.6 (2015), pp. 1929–1936. issn: 2168-2194. doi: 10.1109/JBHI.2014.2350996.
- [CES15] E. Celledoni, M. Eslitzbichler, and A. Schmeding. “Shape analysis on Lie groups with applications in computer animation”. In: *arXiv preprint arXiv:1506.00783* (2015). url: <https://arxiv.org/abs/1506.00783> (visited on 09/12/2017).
- [Che+15] Yanping Chen et al. *The UCR Time Series Classification Archive*. www.cs.ucr.edu/~eamonn/time_series_data/. July 2015.
- [Cho+15] Vincent Choqueuse et al. “Voltage sags estimation in three-phase systems using Unconditional Maximum Likelihood estimation”. In: *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*. Vol. 2015-Augus. IEEE, 2015, pp. 2869–2873. isbn: 9781467369978. doi: 10.1109/ICASSP.2015.7178495.
- [DGP15] Christopher DeFlitch, Glenn Geeting, and Harold L. Paz. “Reinventing emergency department flow via healthcare delivery science”. In: *Health Environments Research and Design Journal* 8.3 (2015), pp. 105–115. issn: 19375867. doi: 10.1177/1937586715580949.
- [Giu+15] C. Giusti et al. “Clique topology reveals intrinsic geometric structure in neural correlations”. en. In: *Proceedings of the National Academy of Sciences* 112.44 (Nov. 2015), pp. 13455–13460. issn: 0027-8424, 1091-6490. doi: 10.1073/pnas.1506407112. url: <http://www.pnas.org/lookup/doi/10.1073/pnas.1506407112> (visited on 11/22/2017).
- [PH15] Jose A Perea and John Harer. “Sliding windows and persistence: An application of topological methods to signal analysis”. In: *Foundations of Computational Mathematics* 15.3 (2015), pp. 799–838.
- [STM15] Hummy Song, Anita L. Tucker, and Karen L. Murrell. “The diseconomies of queue pooling: An empirical investigation of emergency department length of stay”. In: *Management Science* 61.12 (2015), pp. 3032–3053. issn: 15265501. doi: 10.1287/mnsc.2014.2118.
- [TZH15] Chad M. Topaz, Lori Ziegelmeier, and Tom Halverson. “Topological Data Analysis of Biological Aggregation Models”. en. In: *PLOS ONE* 10.5 (May 2015), e0126383. issn: 1932-6203. doi: 10.1371/journal.pone.0126383. url: <http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0126383> (visited on 05/06/2018).
- [VS15] Hajnal Vass and Zsuzsanna K. Szabo. “Application of Queuing Model to Patient Flow in Emergency Department. Case Study”. In: *Procedia Economics and Finance* 32 (2015), pp. 479–487. issn: 22125671. doi: 10.1016/s2212-5671(15)01421-5.

- [Zha+15] Z. Zhang et al. "Video-based action recognition using rate-invariant analysis of covariance trajectories". In: *arXiv preprint arXiv:1503.06699* (2015). url: <https://arxiv.org/abs/1503.06699> (visited on 09/28/2017).
- [Bau+14] M. Bauer et al. "Constructing reparameterization invariant metrics on spaces of plane curves". In: *Differential Geometry and its Applications* 34 (2014), pp. 139–165. url: <http://www.sciencedirect.com/science/article/pii/S092622451400062X> (visited on 09/27/2017).
- [DSZ14] X. Duan, H. Sun, and Z. Zhang. "A Natural Gradient Descent Algorithm for the Solution of Lyapunov Equations Based on the Geodesic Distance". In: *Journal of Computational Mathematics* 32.1 (Jan. 2014), pp. 93–106. issn: 02549409. doi: 10.4208/jcm.1310-m4225. url: <http://www.global-sci.org/jcm/volumes/v32n1/pdf/321-93.pdf> (visited on 11/25/2016).
- [Gel+14] A. Gelman et al. *Bayesian data analysis*. Vol. 2. CRC press Boca Raton, FL, 2014. url: <http://amstat.tandfonline.com/doi/full/10.1080/01621459.2014.963405> (visited on 10/02/2017).
- [Kad14] Farid Kadri. "Contribution à la conception d'un système d'aide à la décision pour la gestion de situations de tension au sein des systèmes hospitaliers. Application à un service d'urgence". PhD thesis. Université de VALENCIENNES ET DU HAINAUT-CAMBRESIS, 2014, p. 216.
- [LXL14] Qiang Liu, Xiaolan Xie, and Ran Liu. "A queuing approximation method for capacity planning of emergency department with time-varying demand". In: *IEEE International Conference on Automation Science and Engineering*. Vol. 2014-Janua. 2014, pp. 558–563. doi: 10.1109/CoASE.2014.6899381.
- [Mor14] Margaret Morrison. "Complex systems and renormalization group explanations". In: *Philosophy of Science* 81.5 (2014), pp. 1144–1156. issn: 00318248. doi: 10.1086/677904.
- [ABL13] M. Arnaudon, F. Barbaresco, and Le Yang. "Riemannian Medians and Means With Applications to Radar Signal Processing". In: *IEEE Journal of Selected Topics in Signal Processing* 7.4 (Aug. 2013), pp. 595–604. issn: 1932-4553, 1941-0484. doi: 10.1109/JSTSP.2013.2261798. url: <http://ieeexplore.ieee.org/document/6514112/> (visited on 10/06/2017).
- [Aug+13] F. Auger et al. "Time-Frequency Reassignment and Synchrosqueezing: An Overview". In: *IEEE Signal Process. Mag.* 30.6 (Nov. 2013), pp. 32–41. issn: 1053-5888. doi: 10.1109/MSP.2013.2265316. url: <http://ieeexplore.ieee.org/document/6633061/>.
- [BGR13] Anthony M. Bloch, François Gay-Balmaz, and Tudor S. Ratiu. "Double bracket flows, toda flows and rigid body toda". In: *2013 51st Annual Allerton Conference on Communication, Control, and Computing, Allerton 2013* (2013), pp. 1567–1572. doi: 10.1109/Allerton.2013.6736714.
- [Bou13a] Guillaume Bouleux. "Oblique projection pre-processing and TLS application for diagnosing rotor bar defects by improving power spectrum estimation". In: *Mechanical Systems and Signal Processing* 41.1-2 (2013), pp. 301–312. issn: 08883270. doi: 10.1016/j.ymssp.2013.06.018. url: <http://www.sciencedirect.com/science/article/pii/S0888327013003014>.
- [Bou13b] Guillaume Bouleux. "Prior Knowledge Optimum Understanding by Means of Oblique Projectors and Their First Order Derivatives". In: *IEEE Signal Processing Letters* 20.3 (2013), pp. 205–208. issn: 10709908. doi: 10.1109/LSP.2013.2238928.
- [Bou13c] Nicolas Boumal. "Interpolation and regression of rotation matrices". In: *Geometric Science of Information*. Springer, 2013, pp. 345–352.
- [Mun13] Elizabeth Munch. "Applications of persistent homology to time varying systems". PhD thesis. 2013.
- [Pec13] J S Peck. "Using prediction to facilitate patient flow in a health care delivery chain". PhD thesis. MIT, 2013.

- [Pec+13] J S Peck et al. "Generalizability of a Simple Approach for Predicting Hospital Admission From an Emergency Department." In: *Academic emergency medicine : official journal of the Society for Academic Emergency Medicine* 20.11 (2013).
- [SS13] Dena L Schanzer and Brian Schwartz. "Impact of Seasonal and Pandemic Influenza on Emergency Department Visits, 2003–2010, Ontario, Canada". In: *Academic Emergency Medicine* 20.4 (Apr. 2013). Ed. by Michael J Mello, pp. 388–397. issn: 1069-6563. doi: 10.1111/acem.12111.
- [Tab+13] Manel Taboada et al. "Modeling, simulation and optimization of resources management in hospital emergency departments using the agent-based approach". In: *Advances in Computational Modeling Research: Theory, Developments and Applications*. 2013, pp. 1–31. isbn: 9781626180659.
- [WBJ13] Petter Wirfält, Guillaume Bouleux, and Magnus Jansson. "Estimation de DOA sous l'hypothèse de sources partiellement corrélées". In: *XXIV colloque GRETSI*. Brest, 2013.
- [WJB13] Petter Wirfält, Magnus Jansson, and Guillaume Bouleux. "Prior-exploiting direction-of-arrival algorithm for partially uncorrelated source signals". In: *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*. 2013, pp. 3972–3976. isbn: 9781479903566. doi: 10.1109/ICASSP.2013.6638404.
- [YW13] Yili L. and K. M. Wong. "Riemannian Distances for Signal Classification by Power Spectral Density". In: *IEEE Journal of Selected Topics in Signal Processing* 7.4 (Aug. 2013), pp. 655–669. issn: 1932-4553, 1941-0484. doi: 10.1109/JSTSP.2013.2260320. url: <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6509394> (visited on 09/07/2016).
- [BDT12] O. Besson, N. Dobigeon, and J.-Y. Tournet. "CS Decomposition Based Bayesian Subspace Estimation". In: *IEEE Transactions on Signal Processing* 60.8 (Aug. 2012), pp. 4210–4218. issn: 1053-587X. doi: 10.1109/TSP.2012.2197619.
- [Bin12] N.H. Bingham. "Szegő's theorem and its probabilistic descendants". en. In: *Probability Surveys* 9 (2012), pp. 287–324. issn: 1549-5787. doi: 10.1214/11-PS178. url: <http://www.i-journals.org/ps/viewarticle.php?id=178&layout=abstract> (visited on 07/06/2016).
- [TWL12] Kar Way Tan, Chao Wang, and Hoong Chuin Lau. "Through Dynamic Priority Queue". In: *IEEE International Conference on Automation Science and Engineering*. 2012, pp. 125–130. isbn: 9781467304306.
- [Wir+12] Petter Wirfält et al. "Optimal prior knowledge-based direction of arrival estimation". In: *IET Signal Processing* 6.8 (2012), pp. 731–742. issn: 17519675. doi: 10.1049/iet-spr.2011.0453.
- [BDT11] O. Besson, N. Dobigeon, and J.-Y. Tournet. "Minimum Mean Square Distance Estimation of a Subspace". In: *IEEE Transactions on Signal Processing* 59.12 (Dec. 2011), pp. 5709–5720. issn: 1053-587X. doi: 10.1109/TSP.2011.2166548.
- [Bin11] N. H. Bingham. "Szego's Theorem and its Probabilistic Descendants". In: (Aug. 2011). arXiv: 1108.0368. url: <http://arxiv.org/abs/1108.0368>.
- [BK11] Guillaume Bouleux and Thameur Kidar. "Total Least Squares application for diagnosing rotor bar defects by improving power spectrum estimation". In: *8th International Conference on Condition Monitoring and Machinery Failure Prevention Technologies 2011, CM 2011/MFPT 2011*. Vol. 1. 2011, pp. 165–176. isbn: 9781618390141.
- [BKG11] Guillaume Bouleux, Thameur Kidar, and François Guillet. "An improper random vector approach for ESPRIT and unitary ESPRIT frequency estimation". In: *7th International Workshop on Systems Signal Processing and their Applications (WOSSPA)*. 2011, pp. 299–302. isbn: 9781457706905. doi: 10.1109/WOSSPA.2011.5931478.
- [Sri+11] A. Srivastava et al. "Shape Analysis of Elastic Curves in Euclidean Spaces". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 33.7 (July 2011), pp. 1415–1428. issn: 0162-8828. doi: 10.1109/TPAMI.2010.184.

- [Tur+11] P. Turaga et al. "Statistical Computations on Grassmann and Stiefel Manifolds for Image and Video-Based Recognition". en. In: *IEEE Trans. Pattern Anal. Mach. Intell.* 33.11 (Nov. 2011), pp. 2273–2286. issn: 0162-8828, 2160-9292. doi: 10.1109/TPAMI.2011.52.
- [Wir+11a] Petter Wirfält et al. "Subspace-based frequency estimation utilizing prior information". In: *IEEE Workshop on Statistical Signal Processing Proceedings*. 2011, pp. 533–536. isbn: 9781457705700. doi: 10.1109/SSP.2011.5967751.
- [Wir+11b] Petter Wirfält et al. "Prior knowledge-based direction of arrival estimation". In: *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*. 2011, pp. 2540–2543. isbn: 9781457705397. doi: 10.1109/ICASSP.2011.5947002.
- [CLT10] Aurélie Charles, Matthieu Lauras, and Rolando Tomasini. "Collaboration networks involving humanitarian organisations - Particular problems for a particular sector". In: *IFIP Advances in Information and Communication Technology*. Vol. 336 AICT. 2010. doi: 10.1007/978-3-642-15961-9_18.
- [Eit+10] D R Eitel et al. "Improving Service Quality by Understanding Emergency Department Flow: A White Paper and Position Statement Prepared For the American Academy of Emergency Medicine". In: *The Journal of Emergency Medicine* 38.1 (2010), pp. 70–79. issn: 0736-4679.
- [EV10] G. Evenbly and G. Vidal. "Entanglement renormalization in free bosonic systems: Real-space versus momentum-space renormalization group transforms". In: *New Journal of Physics* 12 (2010), pp. 1–14. issn: 13672630. doi: 10.1088/1367-2630/12/2/025007. arXiv: 0801.2449.
- [KP10] V. Kavitha and M. Punithavalli. "Clustering Time Series Data Stream - A Literature Survey". In: *International Journal of Computer Science and Information Security* 8.1 (2010).
- [Kit+10] T. Kitagawa et al. "Topological characterization of periodically-driven quantum systems". In: *Physical Review B* 82.23 (Dec. 2010). arXiv: 1010.6126. issn: 1098-0121, 1550-235X. doi: 10.1103/PhysRevB.82.235114. url: <http://arxiv.org/abs/1010.6126> (visited on 01/16/2017).
- [Olo10] Richard Oloruntoba. "An analysis of the Cyclone Larry emergency relief chain: Some key success factors". In: *International Journal of Production Economics* 126.1 (2010), pp. 85–101. issn: 09255273. doi: 10.1016/j.ijpe.2009.10.013.
- [Sta+10] A S Stang et al. "Markers of Overcrowding in a Pediatric Emergency Department". In: *Academic Emergency Medicine* 17.2 (2010), pp. 151–156.
- [Sz+10] B. Sz.-Nagy et al. *Harmonic Analysis of Operators on Hilbert Space*. en. New York, NY: Springer New York, 2010. isbn: 978-1-4419-6093-1 978-1-4419-6094-8. url: <http://link.springer.com/10.1007/978-1-4419-6094-8> (visited on 07/19/2016).
- [YAB10] L. Yang, M. Arnaudon, and F. Barbaresco. "Riemannian median, geometry of covariance matrices and radar target detection". In: *Radar Conference (EuRAD), 2010 European*. Sept. 2010, pp. 415–418.
- [ABB09] G Abraham, G B Byrnes, and C A Bain. "Short-Term Forecasting of Emergency Inpatient Flow". In: *IEEE Transactions on Information Technology in Biomedicine* 13.3 (2009), pp. 380–388.
- [AGM09] A J Arenas, G Gonzalez-Parra, and J A Morano. "Stochastic modeling of the transmission of respiratory syncytial virus (RSV) in the region of Valencia, Spain". In: *Biosystems* 96.3 (2009), pp. 206–212.
- [Bar09] F. Barbaresco. "Interactions between symmetric cone and information geometries: Bruhat-tits and siegel spaces models for high resolution autoregressive doppler imagery". In: *Emerging Trends in Visual Computing*. Springer, 2009, pp. 124–163. url: http://link.springer.com/chapter/10.1007/978-3-642-00826-9_6 (visited on 07/05/2016).
- [BSB09a] Guillaume Bouleux, Peter Stoica, and Rémy Boyer. "Une méthode optimale pour l'estimation des directions d'arrivées basée sur des connaissances a priori". In: *XXII colloque GRETSI*. Dijon, 2009, p. 120.

- [BSB09b] Guillaume Bouleux, Petre Stoica, and Remy Boyer. "An optimal prior knowledge-based DoA estimation method". In: *European Signal Processing Conference*. 2009, pp. 869–873.
- [El +09] Mohammed Nabil El Korso et al. "Sequential estimation of the range and the bearing using the Zero-Forcing MUSIC approach". In: *European Signal Processing Conference*. 2009, pp. 1404–1408.
- [Kor+09] Mohammed Nabil E L Korso et al. "Estimation séquentielle des paramètres de localisation en champ proche à l' aide de l' approche Zero-Forcing". In: *XXII colloque GRETSI*. Dijon, 2009, pp. 2–5.
- [Sim09] B. Simon. *Orthogonal Polynomials on the Unit Circle Part1 and Part 2*. en. Vol. 54. American Mathematical Society, 2009.
- [VB09] H. L. Van Trees and K. L. Bell. "Covariance, Subspace, and Intrinsic CramrRao Bounds". In: (2009). url: http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5266021 (visited on 10/18/2016).
- [Bho08] W. Bhosri. "Geometric approach to prediction and periodic systems". PhD thesis. Purdue University, 2008. url: <http://search.proquest.com/openview/2e253bda5b68196fd22b8f9044411634/1?pq-origsite=gscholar&cbl=18750&diss=y> (visited on 05/04/2017).
- [BB08a] Guillaume Bouleux and Rémy Boyer. "First-order analysis of the prior-based knowledge minnorm algorithm". In: *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*. Las Vegas, 2008, pp. 2489–2492. isbn: 1424414849. doi: 10.1109/ICASSP.2008.4518153.
- [Bou+08] Guillaume Bouleux et al. "A subspace-based rejection method for detecting bearing fault in asynchronous motor". In: *Proceedings of 2008 International Conference on Condition Monitoring and Diagnosis, CMD 2008*. Beijing, 2008, pp. 171–174. isbn: 9781424416219. doi: 10.1109/CMD.2008.4580256.
- [BB08b] Rémy Boyer and Guillaume Bouleux. "Oblique Projections for Direction-Of-Arrival Estimation with Prior Knowledge". In: *IEEE Trans. on Signal Processing* 56.4 (2008), pp. 1374–1387. issn: 1053587X. doi: 10.1109/TSP.2007.909348.
- [Car+08] Gunnar Carlsson et al. "On the Local Behavior of Spaces of Natural Images". en. In: *International Journal of Computer Vision* 76.1 (Jan. 2008), pp. 1–12. issn: 0920-5691, 1573-1405. doi: 10.1007/s11263-007-0056-x. url: <https://link.springer.com/article/10.1007/s11263-007-0056-x> (visited on 05/06/2018).
- [Dak+08] Amani Daknou et al. "Toward a multi-agent model for the care of patients at the emergency department". In: *Proceedings of the 10th WSEAS international conference on Mathematical methods, computational techniques and intelligent systems*. 2008, pp. 264–269. isbn: 978-960-474-012-3. url: <http://portal.acm.org/citation.cfm?id=1562334.1562383>.
- [HA08] N R Hoot and D Aronsky. "Systematic Review of Emergency Department Crowding: Causes, Effects, and Solutions". In: *Annals of Emergency Medicine* 52.2 (2008), 126–136.e1.
- [Hoo+08] N R Hoot et al. "Forecasting Emergency Department Crowding: A Discrete Event Simulation". In: *Annals of emergency medicine* 52.2 (2008), pp. 116–125.
- [Les08] Annick Lesne. "Regularization , renormalization , and renormalization groups : relationships and epistemological aspects". In: 2008.
- [PPQ08] Jonathan Patrick, Martin L. Puterman, and Maurice Queyranne. "Dynamic multipriority patient scheduling for a diagnostic resource". In: *Operations Research* 56.6 (2008), pp. 1507–1525. issn: 0030364X. doi: 10.1287/opre.1080.0590.
- [ZP08] Michael M. Zavlanos and George J. Pappas. "A dynamical systems approach to weighted graph matching". In: *Automatica* 44.11 (2008), pp. 2817–2824. issn: 00051098. doi: 10.1016/j.automatica.2008.04.009. url: <http://dx.doi.org/10.1016/j.automatica.2008.04.009>.

- [Ars+07] Vincent Arsigny et al. “Geometric Means in a Novel Vector Space Structure on Symmetric Positive-Definite Matrices”. en. In: *SIAM Journal on Matrix Analysis and Applications* 29.1 (Jan. 2007), pp. 328–347. issn: 0895-4798, 1095-7162. doi: 10.1137/050637996. url: <http://epubs.siam.org/doi/abs/10.1137/050637996> (visited on 11/09/2016).
- [BB07a] Guillaume Bouleux and Rémy Boyer. “A lower bound for sequential estimators”. In: *2007 2nd IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing, CAMPSAP*. 2007, pp. 277–280. isbn: 9781424417148. doi: 10.1109/CAMPSAP.2007.4498019.
- [BB07b] Guillaume Bouleux and Rémy Boyer. “Estimation des Directions D’Arrivées incorporant un a priori : Algorithmes et Variances Théoriques algorithmes d’estimation et analyse asymptotique”. In: *XXI colloque GRETSI*. Troyes, 2007, pp. 11–14.
- [BB07c] Guillaume Bouleux and Rémy Boyer. “Zero-forcing based sequential music algorithm”. In: *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*. Vol. 3. Hawai, 2007. isbn: 1424407281. doi: 10.1109/ICASSP.2007.366855.
- [Fil+07] P. Fillard et al. “Clinical DT-MRI estimation, smoothing, and fiber tracking with log-Euclidean metrics”. In: *IEEE transactions on medical imaging* 26.11 (2007), pp. 1472–1482. url: http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=4359024 (visited on 11/09/2016).
- [Hun+07] G R Hung et al. “Computer modeling of patient flow in a pediatric emergency department using discrete event simulation.” In: *Pediatric emergency care* 23.1 (2007), pp. 5–10.
- [HM07a] H. L. Hurd and A. Miamee. *Periodically Correlated Random Sequences: Spectral Theory and Practice*. Wiley Series in Probability and Statistics. Wiley, 2007. isbn: 9780470182826. url: http://books.google.fr/books?id=%7B%5C_%7DTVwqWV9xNIC.
- [HM07b] H. L. Hurd and A. Miamee. “Review of Hilbert Spaces”. en. In: *Periodically Correlated Random Sequences*. DOI: 10.1002/9780470182833.ch3. John Wiley & Sons, Inc., 2007, pp. 45–66. isbn: 978-0-470-18283-3. url: <http://onlinelibrary.wiley.com/doi/10.1002/9780470182833.ch3/summary> (visited on 10/06/2017).
- [HM07c] H. L. Hurd and A. Miamee. “Stationary Random Sequences”. en. In: *Periodically Correlated Random Sequences*. DOI: 10.1002/9780470182833.ch4. John Wiley & Sons, Inc., 2007, pp. 67–131. isbn: 978-0-470-18283-3. url: <http://onlinelibrary.wiley.com/doi/10.1002/9780470182833.ch4/summary> (visited on 10/06/2017).
- [LTS07] Rui Li, Tai-Peng Tian, and Stan Sclaroff. “Simultaneous Learning of Nonlinear Manifold and Dynamical Models for High-dimensional Time Series”. In: *2007 IEEE 11th International Conference on Computer Vision*. Rio de Janeiro, Brazil: IEEE, 2007, pp. 1–8. doi: 10.1109/ICCV.2007.4409044.
- [Sim07] B. Simon. “CMV matrices: Five years after”. en. In: *Journal of Computational and Applied Mathematics* 208.1 (Nov. 2007), pp. 120–154. issn: 03770427. doi: 10.1016/j.cam.2006.10.033. url: <http://linkinghub.elsevier.com/retrieve/pii/S0377042706006315> (visited on 10/06/2017).
- [ACL06] Elif Akcali, Murray J. Côté, and Chin Lin. “A network flow approach to optimizing hospital bed capacity decisions”. In: *Health Care Management Science* 9.4 (2006), pp. 391–404. issn: 13869620. doi: 10.1007/s10729-006-0002-4.
- [BB06a] Guillaume Bouleux and Rémy Boyer. “Analysis of prior-subspace estimation schemes”. In: *2006 IEEE Sensor Array and Multichannel Signal Processing Workshop Proceedings, SAM 2006*. Boston, 2006, pp. 526–530. isbn: 1424403081. doi: 10.1109/sam.2006.1706189.
- [BB06b] Guillaume Bouleux and Rémy Boyer. “Subspace-based estimation of Direction-Of-Arrival with known directions”. In: *European Signal Processing Conference*. Florence, 2006.
- [BB06c] Rémy Boyer and Guillaume Bouleux. “Cramér-Rao lower bound for prior-subspace estimation”. In: *2006 IEEE Sensor Array and Multichannel Signal Processing Workshop Proceedings, SAM 2006*. 2006, pp. 394–398. isbn: 1424403081. doi: 10.1109/SAM.2006.1677228.

- [BBA06] Rémy Boyer, Guillaume Bouleux, and Karim Abed-Meraim. “Common pole estimation with an Orthogonal Vector Method”. In: *European Signal Processing Conference*. Florence, 2006.
- [Hal+06] R Hall et al. “Patient Flow: Reducing Delay in Healthcare Delivery”. In: ed. by Randolph W. Hall. Vol. 91. International Series in Operations Research & Management Science. Springer US, 2006, pp. 1–44.
- [Pen06] X. Pennec. “Intrinsic statistics on Riemannian manifolds: Basic tools for geometric measurements”. In: *Journal of Mathematical Imaging and Vision* 25.1 (2006), pp. 127–154. url: <http://link.springer.com/article/10.1007/s10851-006-6228-4> (visited on 09/07/2016).
- [PFA06] X. Pennec, P. Fillard, and N. Ayache. “A Riemannian framework for tensor computing”. In: *International Journal of Computer Vision* 66.1 (2006), pp. 41–66. url: <http://link.springer.com/article/10.1007/s11263-005-3222-z> (visited on 09/07/2016).
- [Tse06] M. C. Tseng. “Contractions, Matrix Parametrizations, and Quantum Information”. In: *arXiv:quant-ph/0610259* (Oct. 2006). arXiv: quant-ph/0610259. url: <http://arxiv.org/abs/quant-ph/0610259> (visited on 10/02/2017).
- [TR06] M. C. Tseng and V. Ramakrishna. “Dilation Theoretic Parametrizations of Positive Matrices with Applications to Quantum Information”. In: *arXiv:quant-ph/0610021* (Oct. 2006). arXiv: quant-ph/0610021. url: <http://arxiv.org/abs/quant-ph/0610021> (visited on 10/02/2017).
- [AHK06] S. W.M. Au-Yeung, P. G. Harrison, and W. J. Knottenbelt. “A queueing network model of patient flow in an accident and emergency department”. In: *ESM 2006 - 2006 European Simulation and Modelling Conference: Modelling and Simulation 2006*. 2006, pp. 60–67.
- [Lam05] S. Lambert-Lacroix. “Extension of Autocovariance Coefficients Sequence for Periodically Correlated Processes”. en. In: *Journal of Time Series Analysis* 26.3 (May 2005), pp. 423–435. issn: 1467-9892. doi: 10.1111/j.1467-9892.2004.00409.x. url: <http://online.library.wiley.com/doi/10.1111/j.1467-9892.2004.00409.x/abstract> (visited on 02/16/2017).
- [MM05a] J. G. Marcano and M. D. Morán. “The Arov-Grossman model and Burg’s entropy”. In: *Recent Advances in Applied Probability*. Springer, 2005, pp. 329–349. url: http://link.springer.com/chapter/10.1007/0-387-23394-6_15 (visited on 01/31/2017).
- [MM05b] P. W. Michor and D. Mumford. “Vanishing geodesic distance on spaces of submanifolds and diffeomorphisms”. In: *Doc. Math* 10 (2005), pp. 217–245. url: <http://www.emis.ams.org/journals/DMJDMV/vol-10/05.pdf> (visited on 10/02/2017).
- [Won05] W.S. Wong. “Solving combinatorial optimization problems by gradient flows”. In: 4. 2005, pp. 1494–1496. doi: 10.1109/cdc.1992.371166.
- [ZC05] A. Zomorodian and G. Carlsson. “Computing Persistent Homology”. en. In: *Discrete Comput Geom* 33.2 (Feb. 2005), pp. 249–274. issn: 0179-5376, 1432-0444. doi: 10.1007/s00454-004-1146-y. url: <https://link.springer.com/article/10.1007/s00454-004-1146-y> (visited on 11/16/2017).
- [Bro+04] E. Broszkiewicz-Suwaj et al. “On detecting and modeling periodic correlation in financial data”. en. In: *Physica A: Statistical Mechanics and its Applications* 336.1-2 (May 2004), pp. 196–205. issn: 03784371. doi: 10.1016/j.physa.2004.01.025. url: <http://linkinghub.elsevier.com/retrieve/pii/S0378437104000482> (visited on 05/17/2017).
- [Car04] Alain Cardon. *Etude de la conception et du controle comportemental d'une organisation massive d'agents*. Tech. rep. Paris: LIP6, 2004, pp. 1–61. url: <http://admiroutes.asso.fr/larevue/2001/22/Robot.pdf>.
- [HP04] Michael Hofer and Helmut Pottmann. “Energy-minimizing splines in manifolds”. In: *ACM Transactions on Graphics*. Vol. 23. 3. 2004. doi: 10.1145/1015706.1015716.

- [HKF04] H. Hurd, G. Kallianpur, and J. Farshidi. "Correlation and spectral theory for periodically correlated random fields indexed on \mathbb{Z}^2 ". In: *Journal of Multivariate Analysis* 90.2 (Aug. 2004), pp. 359–383. issn: 0047-259X. doi: 10.1016/j.jmva.2004.04.010. url: <http://www.sciencedirect.com/science/article/pii/S0047259X04000843> (visited on 06/13/2016).
- [PGW04] Pierre Philippe, Maria R. Garcia, and Bruce J. West. "Evidence of "essential uncertainty" in emergency-ward length of stay". In: *Fractals* 12.2 (2004), pp. 197–209. issn: 0218348X. doi: 10.1142/S0218348X04002495.
- [Car03] Alain Cardon. "Control and Behavior of a Massive Multi-agent System". In: *Innovative Concepts for Agent-Based Systems*. Ed. by Walt Truszkowski, Mike Hinchev, and Chris Rouff. Berlin, Heidelberg: Springer Berlin Heidelberg, 2003, pp. 46–60. isbn: 978-3-540-45173-0.
- [DL03] S. Dégerine and S. Lambert-Lacroix. "Characterization of the partial autocorrelation function of nonstationary time series". en. In: *Journal of Multivariate Analysis* 87.1 (Oct. 2003), pp. 46–59. issn: 0047259X. doi: 10.1016/S0047-259X(03)00025-3. url: <http://linkinghub.elsevier.com/retrieve/pii/S0047259X03000253> (visited on 07/07/2016).
- [Des03a] F. Desbouvries. "Non-Euclidean geometrical aspects of the Schur and Levinson-Szego algorithms". In: *IEEE Trans. on Information theory* 49.8 (2003), p. 346.
- [Des03b] F. Desbouvries. "Non-euclidean geometrical aspects of the schur and levinson-szego algorithms". en. In: *IEEE Transactions on Information Theory* 49.8 (Aug. 2003), pp. 1992–2003. issn: 0018-9448. doi: 10.1109/TIT.2003.814478. url: <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=1214077> (visited on 07/04/2016).
- [DL02] S. Dégerine and S. Lambert-Lacroix. "Evolutionary instantaneous spectrum associated with partial autocorrelation function". In: (2002). url: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=314361 (visited on 03/28/2017).
- [ELZ02] Edelsbrunner, Letscher, and Zomorodian. "Topological Persistence and Simplification". en. In: *Discrete & Computational Geometry* 28.4 (Nov. 2002), pp. 511–533. issn: 0179-5376, 1432-0444. doi: 10.1007/s00454-002-2885-2. url: <http://link.springer.com/10.1007/s00454-002-2885-2> (visited on 03/21/2018).
- [MRV02] Enric Monte, Josep Roca, and Lluís Vilardell. "On the self-similar distribution of the emergency ward arrivals time series". In: *Fractals* 10.4 (2002), pp. 413–427. issn: 0218348X. doi: 10.1142/S0218348X02001464.
- [MGD02] P. Murali Krishna, Vikram M Gadre, and Uday B Desai. "Modelling and control of broadband traffic using multiplicative multifractal cascades". In: *Sadhana - Academy Proceedings in Engineering Sciences* 27.PART 6 (2002), pp. 699–723. issn: 02562499. doi: 10.1007/bf02703360.
- [SCK02] A. H. Sayed, T. Constantinescu, and T. Kailath. "Recursive Construction of Multichannel Transmission Lines with a Maximum Entropy Property". In: *Codes, Graphs, and Systems*. Springer, 2002, pp. 259–290. url: http://link.springer.com/chapter/10.1007/978-1-4615-0895-3_15 (visited on 06/13/2016).
- [Fla99] P. Flandrin. In: *Time-Frequency/Time-Scale Analysis*. Vol. 10. Wavelet Analysis and Its Applications. Academic Press, 1999. doi: [https://doi.org/10.1016/S1874-608X\(99\)80002-8](https://doi.org/10.1016/S1874-608X(99)80002-8). url: <https://www.sciencedirect.com/science/article/pii/S1874608X9980028>.
- [NW99] M. E.J. Newman and D. J. Watts. "Scaling and percolation in the small-world network model". In: *Physical Review E - Statistical Physics, Plasmas, Fluids, and Related Interdisciplinary Topics* 60.6 (1999). issn: 1063651X. doi: 10.1103/PhysRevE.60.7332.
- [OFC99] Tim Oates, Laura Firoiu, and Paul R Cohen. "Clustering Time Series with Hidden Markov Models and Dynamic Time Warping". In: *Proceedings of the IJCAI-99 workshop on neural, symbolic and reinforcement learning methods for sequence learning*. 1999, pp. 17–21.

- [WPK99] M. West, R. Prado, and A. D. Krystal. "Evaluation and Comparison of EEG Traces: Latent Structure in Nonstationary Time Series". In: *Journal of the American Statistical Association* 94.448 (Dec. 1999), p. 1083. issn: 01621459. doi: 10.2307/2669922. url: <http://www.jstor.org/stable/2669922?origin=crossref> (visited on 04/05/2017).
- [EAS98] A. Edelman, T. A. Arias, and S. T. Smith. "The geometry of algorithms with orthogonality constraints". In: *SIAM journal on Matrix Analysis and Applications* 20.2 (1998), pp. 303–353. url: <http://epubs.siam.org/doi/abs/10.1137/S0895479895290954> (visited on 07/07/2016).
- [Hua+98] N. E. Huang et al. "The empirical mode decomposition and the Hilbert spectrum for non-linear and non-stationary time series analysis". In: *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*. Vol. 454. The Royal Society, 1998, pp. 903–995. url: <http://rspa.royalsocietypublishing.org/content/royprsa/454/1971/903.full.pdf> (visited on 10/06/2016).
- [Les98] Annick Lesne. *Renormalization Methods : Critical Phenomena, Chaos, Fractal Structures*. 1998, p. 349. isbn: 9780471966890.
- [PA98] X. Pennec and N. Ayache. "Uniform distribution, distance and expectation problems for geometric features processing". In: *Journal of Mathematical Imaging and Vision* 9.1 (1998), pp. 49–67. url: <http://link.springer.com/article/10.1023/A:1008270110193> (visited on 10/14/2016).
- [Lam97] S. Lambert. "Sur l'estimation autorégressive des processus périodiquement corrélés". In: *16° Colloque sur le traitement du signal et des images, FRA, 1997*. GRETSI, Groupe d'Etudes du Traitement du Signal et des Images, 1997. url: <http://documents.irevues.inist.fr/handle/2042/12659> (visited on 03/28/2017).
- [Sta+97] J. Stark et al. "Takens embedding theorems for forced and stochastic systems". In: *Nonlinear Analysis: Theory, Methods & Applications* 30.8 (1997), pp. 5303–5314. doi: 10.1016/S0362-546X(96)00149-6.
- [Des96] F. Desbouvries. "Unitary Hessenberg and state-space model based methods for the harmonic retrieval problem". In: *IEE Proc. - Radar, Sonar Navig.* 143.6 (1996), p. 346. issn: 13502395. doi: 10.1049/ip-rsn:19960853. url: http://digital-library.theiet.org/content/journals/10.1049/ip-rsn%7B%5C_%7D19960853.
- [Les96] Annick Lesne. "A comparative introduction to the renormalization methods used in statistical mechanics and for dynamical systems". In: *Actes du congrès: Renormalization in mathematics and physics*. 1996, pp. 1–39.
- [Th+96] Physique Th et al. "Regularization or renormalization group : a few words to avoid confusion". In: *Actes du congrès: Renormalization in mathematics and physics*. 1996, pp. 1–15.
- [Con95] T. Constantinescu. *Schur Parameters, Factorization and Dilation Problems*. en. Basel: Birkhäuser Basel, 1995. isbn: 978-3-0348-9910-9 978-3-0348-9108-0. url: <http://link.springer.com/10.1007/978-3-0348-9108-0> (visited on 06/13/2016).
- [Won95] W. S. Wong. "Matrix representation and gradient flows for NP-hard problems". In: *Journal of Optimization Theory and Applications* 87.1 (1995), pp. 197–220. issn: 00223239. doi: 10.1007/BF02192047.
- [Rao92] C. Radhakrishna Rao. "Information and the accuracy attainable in the estimation of statistical parameters". In: *Breakthroughs in statistics*. Springer, 1992, pp. 235–247. url: http://link.springer.com/chapter/10.1007/978-1-4612-0919-5_16 (visited on 11/15/2016).
- [AGR91] G. Ammar, W. Gragg, and L. Reichel. "Constructing a Unitary Hessenberg Matrix from Spectral Data". en. In: ed. by Gene H. Golub and Paul Van Dooren. NATO ASI Series 70. Springer Berlin Heidelberg, 1991, pp. 385–395. isbn: 978-3-642-75538-5 978-3-642-75536-1. url: http://link.springer.com/chapter/10.1007/978-3-642-75536-1_18 (visited on 07/08/2016).

- [Bro91a] R. W. Brockett. “Dynamical systems that sort lists, diagonalize matrices, and solve linear programming problems”. In: *Linear Algebra and Its Applications* 146.C (1991), pp. 79–91. issn: 00243795. doi: 10.1016/0024-3795(91)90021-N.
- [Bro91b] R. W. Brockett. “Dynamical systems that sort lists, diagonalize matrices, and solve linear programming problems”. In: *Linear Algebra and Its Applications* 146.C (1991), pp. 79–91. issn: 00243795. doi: 10.1016/0024-3795(91)90021-N.
- [BW91] Roger W. Brockett and Wing Shing Wong. “A Gradient Flow for the Assignment Problem”. In: *New Trends in Systems Theory*. 1991, pp. 170–177. doi: 10.1007/978-1-4612-0439-8_20.
- [Hur91] H. L. Hurd. “Correlation theory of almost periodically correlated processes”. In: *Journal of Multivariate Analysis* 37.1 (Apr. 1991), pp. 24–45. issn: 0047-259X. doi: 10.1016/0047-259X(91)90109-F. url: <http://www.sciencedirect.com/science/article/pii/S0047259X9190109F> (visited on 06/13/2016).
- [Blo90a] A. M. Bloch. “Steepest descent, linear programming, and Hamiltonian flows”. In: 1990, pp. 77–88. doi: 10.1090/conm/114/1097866.
- [Blo90b] Anthony M. Bloch. “The Kähler Structure of the Total Least Squares Problem, Brockett’s Steepest Descent Equations, and Constrained Flows”. In: *Realization and Modelling in System Theory*. 1990, pp. 83–88. doi: 10.1007/978-1-4612-3462-3_7.
- [FF90] C. Foias and A. E. Frazho. “A Geometric Approach to Positive Definite Sequences”. en. In: *The Commutant Lifting Approach to Interpolation Problems*. OT 44 Operator Theory: Advances and Applications 44. Birkhäuser Basel, 1990, pp. 497–546. isbn: 978-3-0348-7714-5 978-3-0348-7712-1. doi: 10.1007/978-3-0348-7712-1_15. url: http://link.springer.com/chapter/10.1007/978-3-0348-7712-1_15 (visited on 04/13/2017).
- [FFS90] C. Foias, A. E. Frazho, and P. J. Sherman. “A new approach for determining the spectral data of multichannel harmonic signals in noise”. In: *Mathematics of Control, Signals and Systems* 3.1 (1990), pp. 31–43. url: <http://link.springer.com/article/10.1007/BF02551354> (visited on 10/11/2016).
- [Lin89] Winston T. Lin. “Modeling and forecasting hospital patient movements: Univariate and multiple time series approaches”. In: *International Journal of Forecasting* 5.2 (1989), pp. 195–208. issn: 01692070. doi: 10.1016/0169-2070(89)90087-3.
- [MS89] A. Makagon and H. Salehi. “Notes on infinite dimensional stationary sequences”. In: *Probability Theory on Vector Spaces IV*. Springer, 1989, pp. 200–238. url: <http://link.springer.com/content/pdf/10.1007/BFb0083393.pdf> (visited on 10/11/2016).
- [FFS88] C. Foias, A. E. Frazho, and P. J. Sherman. “A geometric approach to the maximum likelihood spectral estimator for sinusoids in noise”. In: *IEEE Transactions on Information Theory* 34.5 (1988), pp. 1066–1070. url: http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=21232 (visited on 07/18/2016).
- [AG87] G. S. Ammar and W. B. Gragg. “The generalized Schur algorithm for the superfast solution of Toeplitz systems”. In: *Rational approximation and its applications in mathematics and physics*. Springer, 1987, pp. 315–330. url: <http://link.springer.com/10.1007%2FBFb0072474> (visited on 07/08/2016).
- [DI87] J.-M. Delosme and I. C. Ipsen. *Computing Partial Correlations from the Data Matrix*. Tech. rep. YALEU/DCS/RR-541. YALE UNIV NEW HAVEN CT DEPT OF COMPUTER SCIENCE, YALE UNIV NEW HAVEN CT DEPT OF COMPUTER SCIENCE, June 1987. url: <http://www.dtic.mil/docs/citations/ADA193904> (visited on 07/12/2017).
- [BK86] D S Broomhead and Gregory P King. “Extracting qualitative dynamics from experimental data”. In: *Phys. D Nonlinear Phenom.* 20.2 (1986), pp. 217–236. issn: 0167-2789. doi: [http://dx.doi.org/10.1016/0167-2789\(86\)90031-X](http://dx.doi.org/10.1016/0167-2789(86)90031-X). url: <http://www.sciencedirect.com/science/article/pii/016727898690031X>.

- [Des86] Uday B. Desai, ed. *Modelling and Application of Stochastic Processes*. en. Boston, MA: Springer US, 1986. isbn: 978-1-4612-9400-9 978-1-4613-2267-2. url: <http://link.springer.com/10.1007/978-1-4613-2267-2> (visited on 10/11/2016).
- [Dou+86] R. G. Douglas et al., eds. *Advances in Invariant Subspaces and Other Results of Operator Theory*. en. Basel: Birkhäuser Basel, 1986. isbn: 978-3-0348-7700-8 978-3-0348-7698-8. url: <http://link.springer.com/10.1007/978-3-0348-7698-8> (visited on 06/13/2016).
- [Fra86] A. E. Frazho. "On Stochastic Bilinear Systems". en. In: *Modelling and Application of Stochastic Processes*. Ed. by Uday B. Desai. DOI: 10.1007/978-1-4613-2267-2_9. Springer US, 1986, pp. 215–241. isbn: 978-1-4612-9400-9 978-1-4613-2267-2. url: http://link.springer.com/chapter/10.1007/978-1-4613-2267-2_9 (visited on 07/08/2016).
- [KB86] T. Kailath and A. M. Bruckstein. "Naimark Dilations, State-Space Generators and Transmission Lines". en. In: *Advances in Invariant Subspaces and Other Results of Operator Theory*. Ed. by R. G. Douglas et al. Operator Theory: Advances and Applications 17. DOI: 10.1007/978-3-0348-7698-8_13. Birkhäuser Basel, 1986, pp. 173–186. isbn: 978-3-0348-7700-8 978-3-0348-7698-8. url: http://link.springer.com/chapter/10.1007/978-3-0348-7698-8_13 (visited on 07/06/2016).
- [Tim86] D. Timotin. "Prediction theory and choice sequences: an alternate approach". In: *Advances in invariant subspaces and other results of operator theory*. Springer, 1986, pp. 341–352. url: http://link.springer.com/chapter/10.1007/978-3-0348-7698-8_24 (visited on 07/04/2016).
- [AC85] Gr. Arsene and T. Constantinescu. "The structure of the Naimark dilation and Gaussian stationary processes". en. In: *Integr equ oper theory* 8.2 (Mar. 1985), pp. 181–204. issn: 0378-620X, 1420-8989. doi: 10.1007/BF01202813. url: <http://link.springer.com/article/10.1007/BF01202813> (visited on 06/13/2016).
- [LK84] H. Lev-Ari and T. Kailath. "Lattice filter parameterization and modeling of nonstationary processes". In: *IEEE Transactions on Information Theory* 30.1 (Jan. 1984), pp. 2–16. issn: 0018-9448. doi: 10.1109/TIT.1984.1056849.
- [Kai82a] T. Kailath. "Time-Variant and Time-Invariant Lattice Filters for Nonstationary Processes." In: *outils et modele mathematiques pour l'automatique, l'analyse de systemes et le traitement du signal 2* (1982).
- [Kai82b] T. Kailath. *Time-Variant and Time-Invariant Lattice Filters for Nonstationary Processes*. Tech. rep. DTIC Document, 1982. url: <http://oai.dtic.mil/oai/oai?verb=getRecord&metadataPrefix=html&identifier=ADA116083> (visited on 06/13/2016).
- [Mas78] P. Masani. "Dilations as Propagators of Hilbertian Varieties". In: *SIAM J. Math. Anal.* 9.3 (June 1978), pp. 414–456. issn: 0036-1410. doi: 10.1137/0509027. url: <http://epubs.siam.org/doi/abs/10.1137/0509027> (visited on 01/16/2017).
- [Dav77] C. Davis. "Geometric approach to a dilation theorem". In: *Linear Algebra and its Applications* 18.1 (Jan. 1977), pp. 33–43. issn: 0024-3795. doi: 10.1016/0024-3795(77)90077-5. url: <http://www.sciencedirect.com/science/article/pii/0024379577900775> (visited on 12/08/2016).
- [Wer77a] A. Weron. "Stochastic Processes of Second Order with Values in Banach Space". In: *Transactions of the 7th Prague Conference on Information Theory, Statistical Decision Functions, Random Processes*, (1977), pp. 567–574.
- [Wer77b] A. Weron. "Stochastic Processes of Second Order with Values in Banach Spaces". en. In: *Transactions of the Seventh Prague Conference on Information Theory, Statistical Decision Functions, Random Processes and of the 1974 European Meeting of Statisticians*. Ed. by J. Kožešnik. Transactions of the Seventh Prague Conference on Information Theory, Statistical Decision Functions, Random Processes and of the 1974 European Meeting of Statisticians 7A. DOI: 10.1007/978-94-010-9910-3_60. Springer Netherlands, 1977, pp. 567–574. isbn: 978-94-010-9912-7 978-94-010-9910-3. url: http://link.springer.com/chapter/10.1007/978-94-010-9910-3_60 (visited on 06/13/2016).

- [Sin76] Yu. G. Sinai. “Self-Similar Probability Distributions”. In: *Theory of Probability & Its Applications* 21.1 (1976), pp. 64–80. issn: 0040-585X. doi: 10.1137/1121005.
- [Fla74] H. Flaschka. “The Toda lattice. II. Existence of integrals”. In: *Phys. Rev. B* 9 (4 Feb. 1974), pp. 1924–1925. doi: 10.1103/PhysRevB.9.1924. url: <https://link.aps.org/doi/10.1103/PhysRevB.9.1924>.
- [TY72] Hiroshi Takai and Hiroaki Yamada. “A note on the dilation theorems”. en. In: *Proceedings of the Japan Academy* 48.4 (1972), pp. 216–220. issn: 0021-4280. doi: 10.3792/pja/1195519699. url: <http://projecteuclid.org/euclid.pja/1195519699> (visited on 03/21/2018).
- [Ger61] I. L. Geronimus. *Orthogonal polynomials: estimates, asymptotic formulas, and series of polynomials orthogonal on the unit circle and on an interval*. engrus. Translation of Mnogochleny, ortogonalniye na okruzhnosti i na otrezke (romanized form). New York: Consultants Bureau, 1961.
- [Sze59] G. Szegő. *Orthogonal Polynomials*. en. American Mathematical Society, 1959. isbn: 978-0-8218-8952-7.
- [Ver36] S. Verblunsky. “On positive harmonic functions”. In: *Proceedings of the London Mathematical Society* 2.1 (1936), pp. 290–320. url: <http://plms.oxfordjournals.org/content/s2-40/1/290.full.pdf> (visited on 12/15/2016).