

Mémoire d'habilitation à diriger des recherches

Inférence bayésienne et représentations parcimonieuses pour la décomposition en motifs

Vincent MAZET

Soutenue le 23 octobre 2019 devant le jury :

Josiane Zerubia	Présidente	Directrice de Recherche à l'Inria
Pierre Chainais	Rapporteur	Professeur à Centrale Lille
Jean-François Giovannelli	Rapporteur	Professeur à l'Université de Bordeaux
Christian Jutten	Rapporteur	Professeur à l'Université Grenoble Alpes
David Brie	Examineur	Professeur à l'Université de Lorraine
Christophe Collet	Garant	Professeur à l'Université de Strasbourg

Table des matières

Remerciements	5
Avant-propos	7
1 Synthèse des activités de recherche	9
1.1 Décomposition en motifs	9
1.2 Détection des bâtiments dans des images de télédétection	14
1.3 Détection de signaux ténus dans des données hyperspectrales fortement bruitées	18
1.4 Publications personnelles	23
2 Décomposition d'un signal spectroscopique en raies	27
2.1 Difficultés de la décomposition en raies	28
2.2 État de l'art	33
2.3 Méthodes développées	36
2.4 Conclusion	37
3 Décomposition conjointe d'une séquence de signaux spectroscopiques	39
3.1 État de l'art	41
3.2 Intérêt d'une décomposition conjointe	42
Modèle pour la décomposition séquentielle	43
Modèle pour la décomposition conjointe	43
Optimisation	44
Comparaison des deux modèles	44
3.3 Approche bayésienne	46
Modèle	46
Optimisation	47
Comparaison avec le modèle à dimension constante	49
Résultat sur des données réelles de photoélectrons	49
3.4 Approche par représentation parcimonieuse	52
Modèle	52
Optimisation	53
Résultats sur des données simulées	56
Résultats sur des données réelles de photoélectrons	56
Résultats sur des données réelles de cinématique de galaxies	56
3.5 Conclusion	58
4 Décomposition d'une image astronomique en structures physiques	61
4.1 Problématique	61
4.2 Modèle	63

4.3	Optimisation	68
4.4	Estimation	70
4.5	Résultats	71
4.6	Conclusion	73
5	Bilan et perspectives	75
5.1	Bilan général	75
5.2	Perspectives	76
	Bibliographie	81

Merci !

La rédaction des remerciements dans un mémoire est synonyme de fin de la rédaction. C'est donc toujours un plaisir de s'y atteler et de prendre le temps de penser à toutes les personnes qui m'ont soutenu.

Aussi, je remercie en premier lieu les trois rapporteurs de mon mémoire, Pierre Chainais, Jean-François Giovannelli et Christian Jutten d'avoir accepté de lire et d'évaluer mon travail. Je remercie également Josiane Zerubia pour avoir endossé le rôle de présidente de jury, et pour les nombreux commentaires et remarques dont elle m'a fait bénéficier. Merci à David Brie d'avoir été présent pour cette nouvelle étape de ma vie professionnelle, qui, je le rappelle, a débuté au rayon fruits et légumes d'un supermarché nancéien. Merci enfin à Christophe Collet d'avoir été garant de mon travail et de m'avoir donné toutes les opportunités de recherche dont j'ai bénéficié à Strasbourg. J'ai vraiment eu plaisir à vous accueillir comme membres de mon jury !

Je remercie chaleureusement mes doctorants pour le plaisir que j'ai eu de travailler avec eux pendant ces années et pour m'avoir fait découvrir de nombreux problèmes et outils scientifiques. Je remercie notamment Hassan et Jean-Baptiste d'avoir été présents à ma soutenance.

Je profite également de cet instant pour remercier tous ceux qui m'apportent tant de motivation dans mon travail d'enseignant qui est pour moi tout aussi important que mon travail de chercheur : mes étudiants bien sûr, mais également mes collègues enseignants et l'équipe de l'Idip pour toutes les discussions et les questionnements pédagogiques que nous avons eu.

Je remercie très sincèrement mes collègues et amis pour leur présence, leur soutien et les nombreuses attentions dont j'ai pu profiter lors de ma soutenance ; mais également pour leur amitié lors des autres moments : Adrien, Alex, Alice, Alix, Anastacia, Argheesh, Benjamin, Céline, Denis, Éléonore, Étienne, Étienne, Florent, Florian, François, Hassan, Hugo, Iris, Jean, Jean, Jean-Baptiste, Jennifer, Loïc, Nicolas, Ola, Pierre-Henri, Stan, Sylvain, Trinh et Vincent !

Je pense aussi à toutes les personnes qui ont contribué à me soutenir matériellement, que ce soit de façon musicale, flambée, ambrée ou chocolatée.

Enfin, je voudrais remercier mes parents ainsi que Mylène, Clément et Camille, d'avoir accepté que je consacre autant de temps sur ce mémoire au détriment du temps passé à leurs côtés.

Avant-propos

Ce mémoire décrit mes activités de recherches effectuées au sein du laboratoire ICube après mon doctorat. Ces activités portent sur la résolution de problèmes inverses en traitement du signal et des images sur des problématiques de détection et de segmentation en spectroscopie, en télédétection et en imagerie astronomique. J'ai principalement utilisé les modélisations bayésiennes ou les représentations parcimonieuses, avec des algorithmes d'optimisation ℓ_2 - ℓ_0 ou de Monte Carlo par chaînes de Markov.

Après mon doctorat en 2005, j'ai obtenu un poste de maître de conférences à l'Université de Strasbourg¹ rattaché à Télécom Physique Strasbourg pour l'enseignement et à l'équipe IMAGEs du laboratoire ICube pour la recherche. J'ai eu l'opportunité d'encadrer cinq thèses de doctorat et cinq stages de master. J'ai également eu la chance de porter deux projets scientifiques (un PEPS de l'INS2I et un projet jeune chercheur de l'ANR) sur la thématique principale de mes recherches et qui fait l'objet de trois chapitres de ce mémoire : la décomposition en motifs de signaux et d'images.

La décomposition en motifs consiste à détecter les « objets » qui sont présents dans un signal ou une image. Dans le cadre de mon travail, ces motifs se modélisent sous la forme d'une fonction paramétrique : le problème revient donc à l'estimation du nombre de motifs et de leurs paramètres. La contribution principale de mon travail a été de considérer le cas où l'on dispose de plusieurs signaux ou images à décomposer et dans lesquels les motifs sont liés entre eux. Ce problème apparaît dans des applications diverses, comme la spectroscopie de photoélectrons ou l'imagerie multispectrale astronomique. La difficulté est d'arriver à gérer dans un temps de calcul raisonnable les dépendances entre motifs ; cette dépendance existe à la fois entre les motifs d'un même signal (dépendance intra-signal) et les motifs appartenant à la même évolution (dépendance inter-signal, due par exemple à une acquisition des signaux à travers le temps).

Ce manuscrit se structure en cinq chapitres. Le chapitre 1 donne une vue d'ensemble de mes principales activités de recherche depuis 2006. Elles sont au nombre de trois : la décomposition en motifs de signaux spectroscopiques et d'images astronomiques, la détection de bâtiments en télédétection et la détection de très faibles signaux dans des données multispectrales astronomiques. Chacun de ces trois sujets a fait l'objet d'une ou deux thèse(s). Le reste du mémoire se concentre sur la décomposition en motifs, en distinguant trois aspects. Tout d'abord, le chapitre 2 étudie la décomposition d'un unique signal unidimensionnel, afin de faire un état de l'art des méthodes de décomposition et d'en identifier les difficultés. Ce chapitre fait donc une synthèse des difficultés et des réponses possibles pour traiter un seul signal, ce qui est le premier pas avant de traiter plusieurs signaux. Le chapitre 3 aborde le problème de la décomposition de plusieurs signaux, avec l'originalité que les motifs présents dans ces signaux sont liés entre eux. C'est un problème que l'on trouve en spectroscopie de photoélectrons et en imagerie astronomique. Ce chapitre expose donc les diverses solutions proposées pour résoudre ce problème original. Le chapitre 4 traite du problème

1. Toutes les structures dans lesquelles j'ai travaillé ont changé de nom depuis mon recrutement. Elles s'appelaient : Université Louis Pasteur (Université de Strasbourg), ENSPS (Télécom Physique Strasbourg), LSIIT (ICube), MIV (IMAGEs).

de la décomposition d'une image multispectrale, en le considérant comme la décomposition d'une séquence d'images à une bande. L'originalité majeure de cette méthode a été le couplage entre une approche bayésienne et un filtrage non linéaire fondé sur les arbres de composantes connexes. Enfin, le chapitre 5 conclut ce mémoire en dressant un bilan des travaux réalisés et en décrivant les perspectives de recherche envisagées.

Chapitre 1

Synthèse des activités de recherche

Ce chapitre résume mes principaux travaux de recherche depuis ma thèse, en 2005. Ces recherches ont été effectuées dans l'équipe IMAGeS¹ du laboratoire ICube, dont les domaines d'intérêt sont la modélisation et le traitement des images avec des approches qui couvrent la morphologie mathématique, la géométrie discrète, les approches statistiques et les méthodes d'apprentissage. Lors de mon recrutement en 2006, j'ai intégré le thème IPSEO² qui développe des méthodes statistiques d'analyse d'image pour l'astronomie et, plus ponctuellement, la télédétection. Ce thème a été créé en 2005 par Christophe Collet et concrétise une longue collaboration entre plusieurs observatoires astronomiques français qui avait été mise en place par Mireille Louys (ICube). En diversifiant mes activités, j'ai également intégré le thème MIBISA³ qui s'intéresse aux approches bayésiennes et à l'apprentissage machine. La richesse de cet environnement de travail a enrichi les compétences acquises lors de mon expérience doctorale et m'a permis d'étendre progressivement mes activités de recherche vers la modélisation markovienne, les représentations parcimonieuses, l'imagerie astronomique, etc. J'ai principalement travaillé sur trois thématiques : la décomposition en motifs de signaux spectroscopiques et d'images astronomiques (section 1.1), la détection de bâtiments dans des images de télédétection (section 1.2) et la détection de signaux ténus dans des images astronomiques (section 1.3). Dans ce mémoire, j'ai choisi de mettre l'accent sur la première thématique car elle a été récurrente dans mes activités de recherche ; les chapitres suivants en détaillent les différents aspects.

1.1 Décomposition en motifs

Contexte

En 2006, j'ai poursuivi la collaboration avec Bernd Vollmer (Observatoire astronomique de Strasbourg) qui avait été mise en place par Christophe Collet dans le cadre de la thèse de Farid Flitti (2002–2005). En collaboration avec Christophe Collet et Sylvain Faisan (ICube), j'ai développé de premières méthodes d'analyse dans un cadre bayésien, en particulier en utilisant les modèles markoviens. Quelques années plus tard, j'ai collaboré avec Charles Soussen (CRAN puis L2S) et El-Hadi Djermoune (CRAN) sur les méthodes d'approximation parcimonieuses. En parallèle, j'ai entamé en 2010, à mon initiative, une collaboration avec Lionel Poisson (LIDYL, CEA & CNRS) en spectroscopie de photoélectrons.

1. Images, modélisation, apprentissage, géométrie et statistique, porté par Fabrice Heitz.

2. *Image processing for space and Earth observations*, porté par Christophe Collet.

3. *Methods in Bayesian inference and statistical analysis*, porté par Erik-André Sauleau.

Deux thèses et quatre stages de M2 ont été effectués dans le cadre de ces recherches. Benjamin Perret (2007–2010) s’est intéressé à la décomposition d’images de galaxies en structures physiques. Sa thèse a été dirigée par Christophe Collet et Éric Slezak (Observatoire de la Côte d’Azur) et encadrée par Sébastien Lefèvre (ICube puis IRISA) et moi-même. Mon rôle s’est concentré sur la modélisation bayésienne et l’optimisation stochastique du problème. Plus tard, Hassan Mortada (2015–2018) a considéré le problème de la décomposition conjointe de signaux spectroscopiques. Sa thèse a été dirigée par Christophe Collet et Charles Soussen. Je suis intervenu dans cette thèse comme encadrant et responsable du projet ANR assurant son financement. Nous avons traité la décomposition en motifs à l’aide de méthodes d’approximation parcimonieuses, profitant de l’expertise de Charles Soussen et El-Hadi Djermoune dans ce domaine. Enfin, j’ai initié et encadré les stages de M2 de quatre étudiants : Fatma Abdelmoula (2009), Guillaume Dollé (2012, avec Sylvain Faisan), Kais Baccour (2016, avec Sylvain Faisan) et Argheesh Bhanot (2017).

Mes recherches sur ce sujet ont bénéficié de plusieurs financements. En premier lieu, la thèse de Benjamin Perret a été financée grâce à une allocation de recherche ministérielle. J’ai ensuite participé au projet DAHLIA (2009–2013)⁴ puis j’ai réussi à être porteur à la fois du projet SPECTRODEC (2011–2012)⁵ et du projet DSIM (2014–2018)⁶. Ce dernier projet a permis de financer la thèse de Hassan Mortada.

La décomposition en motifs a pour origine l’estimation des raies dans un signal spectroscopique. J’avais abordé ce sujet durant ma thèse de doctorat. C’est donc tout naturellement que j’ai continué ces travaux après mon recrutement en 2006, lorsque j’ai eu l’opportunité de travailler sur des observations multispectrales⁷ de galaxies, dans le but d’étudier leur cinématique. Deux différences majeures ont toutefois distingué ces nouvelles recherches de mes travaux de thèse. D’un point de vue applicatif, les données concernaient des images multispectrales astronomiques, c’est-à-dire un ensemble de plusieurs signaux. D’un point de vue méthodologique, un problème original apparaissait : les signaux constituant ces données n’étaient pas indépendants les uns des autres et il était donc très intéressant de prendre en compte cette dépendance pour améliorer l’estimation. Un problème similaire est apparu également dans le cadre de la décomposition multibande de galaxies et de la spectroscopie de photoélectrons.

Les méthodes développées ont été mises en œuvre dans trois applications.

- En cinématique de galaxies : comme les galaxies sont des masses de gaz en mouvement, les raies d’émission du gaz bougent en fonction de la vitesse de déplacement du gaz (c’est l’effet Doppler). En observant ces déplacements, les astronomes sont en mesure d’en déduire la cinématique de la galaxie, et donc d’apporter des éléments pour comprendre l’histoire de celles-ci. Une observation multispectrale permet donc d’obtenir, en chaque position spatiale, la vitesse de déplacement du gaz. Le traitement de ces données est décrit section 3.4.
- En décomposition multibande de galaxies, on cherche à modéliser une galaxie comme un ensemble de structures physiques (bulbe, disque, bras. . .). Les observations avec lesquelles nous avons travaillé sont des images multispectrales. Le chapitre 4 présente la méthode

4. DAHLIA est un projet financé par l’ANR et porté par Éric Slezak. Il a regroupé une vingtaine de personnes issues de quatre laboratoires (OCA, ICube, CRAL, IRAP).

5. SPECTRODEC était financé par l’INS2I dans le cadre d’un PEPS (projet exploratoire premier soutien) Rupture. Il a réuni neuf personnes de quatre laboratoires (ICube, CRAN, Observatoire de Strasbourg, LIDYL).

6. DSIM a été financé par l’ANR dans le cadre de l’appel jeunes chercheurs–jeunes chercheuses. Il a réuni une quinzaine de personnes issues de deux laboratoires (ICube et CRAN).

7. La distinction entre les adjectifs « multispectral » et « hyperspectral » n’est pas claire. Souvent, on considère qu’une image multispectrale est une image contenant au plus quelques dizaines de bandes relativement larges en longueur d’onde alors qu’une image hyperspectrale contient un très grand nombre de bandes très étroites. Par ailleurs, le terme « hyperspectral » est souvent utilisé à outrance. De ce fait, toutes les images $2D + \lambda$ sont dénommées « multispectrales » dans ce mémoire.

implémentée et les résultats obtenus.

- En spectroscopie de photoélectrons : lors d'une même expérience, plusieurs spectres sont acquis à différents instants pour aboutir à une séquence mettant en évidence les modifications du système étudié au niveau atomique. Le traitement de ces données est décrit dans les sections 3.3 et 3.4.

Les résultats ont fait l'objet d'un chapitre de livre [159] et ont été publiés dans sept articles de revue [10, 138, 146, 149, 153, 157, 177] (trois étant publiés dans des revues applicatives), cinq articles en conférence internationale [151, 155, 156, 174, 176] et cinq articles en conférence nationale [154, 158, 160, 152, 175].

Modèle général

Le problème de la décomposition en motifs, tel que je l'ai abordé, se décrit de la manière suivante. On dispose de S signaux \mathbf{y}_s qui sont organisés dans un espace particulier (le temps ou l'espace par exemple). On suppose que chacun d'eux est la somme de K_s motifs élémentaires $\varphi_{k,s}$. L'équation qui régit ce modèle s'écrit donc, en supposant les données perturbées par un bruit additif \mathbf{b}_s :

$$\forall s, \quad \mathbf{y}_s = \sum_{k=1}^{K_s} \varphi_{k,s} + \mathbf{b}_s. \quad (1.1)$$

On considère que les motifs $\varphi_{k,s}$ peuvent s'écrire sous la forme d'une fonction paramétrique. Ses paramètres sont inconnus, et nous considérons également les cas où ces paramètres sont liés les uns aux autres à travers les signaux \mathbf{y}_s . Plus précisément, nous supposons que les paramètres de deux signaux proches dans l'espace (temporel ou spatial) ont des valeurs proches. Cette hypothèse est fondamentale et c'est une originalité majeure du problème traité. Elle fait la spécificité des approches que j'ai développées et qui cherchent donc à prendre en compte cette dépendance dans la procédure d'estimation.

Le problème de la décomposition en motifs est très proche d'autres problèmes tels que la séparation de sources, le démélange spectral ou la déconvolution impulsionnelle. En effet, la séparation de sources se modélise également par l'équation (1.1) où les motifs $\varphi_{k,s}$ sont identiques pour un même k (à l'amplitude près). En général, il n'existe pas d'expression paramétrique de ces motifs. En déconvolution impulsionnelle, on traite le plus souvent un seul signal, donc $S = 1$. La caractéristique principale de la déconvolution impulsionnelle est toutefois que la somme $\sum_{k=1}^{K_s} \varphi_{k,s}$ s'écrit comme le produit de convolution d'un signal impulsionnel avec une réponse impulsionnelle. Cela revient à dire que les motifs $\varphi_{k,s}$ ont tous la même expression (quels que soient k et s) et ne dépendent que de deux paramètres qui sont leur amplitude et leur positions.

Par ailleurs, le cadre général du problème défini par l'équation (1.1) se structure en plusieurs sous-problèmes qui sont résumés dans la table 1.1 et qui se différencient par les caractéristiques suivantes :

- la dimension des motifs (ou de manière équivalente la dimension des signaux), qui correspond à la dimension de l'espace dans lequel se trouve les motifs $\varphi_{k,s}$. Par exemple en spectroscopie, les motifs sont unidimensionnels et correspondent aux raies d'émission alors qu'en imagerie on cherchera plutôt des motifs bidimensionnels ;
- la dimension de l'espace dans lequel se répartissent les S signaux. Par exemple, une séquence unidimensionnelle de spectres est un ensemble de plusieurs signaux spectroscopiques, et une image multispectrale correspond à un ensemble de signaux répartis dans un espace spatial à deux dimensions. Par extension, on peut considérer que lorsque les données ne contiennent

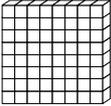
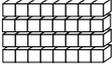
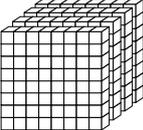
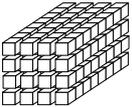
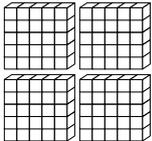
Motifs unidimensionnels	Motifs bidimensionnels
Signaux répartis dans un espace de dimension 0	
	
Exemple : signal spectroscopique Cas traité dans le chapitre 2 (décomposition d'un spectre chimique)	Exemple : image seule
Signaux répartis dans un espace de dimension 1	
	
Exemple : séquence de signaux Cas traité dans le chapitre 3 (séquence de spectres de photoélectrons)	Exemples : pile d'image, image 2D+t, image multispectrale Cas traité dans le chapitre 4 (décomposition de galaxies)
Signaux répartis dans un espace de dimension 2	
	
Exemple : image multispectrale, séquence bidimensionnelle de signaux Cas traité dans le chapitre 3 (cinématique de galaxies)	Exemple : IRM de diffusion

TABLE 1.1 – Différents problèmes et exemples modélisés par l'équation (1.1). Les illustrations schématisent les données : chaque petit cube représente un scalaire et ces cubes sont regroupés pour former des signaux unidimensionnels (colonne de gauche) ou bidimensionnels (colonne de droite).

qu'un unique signal, la dimensionnalité de l'espace est nulle. La répartition des signaux dans l'espace permet de définir le voisinage et le type d'évolution entre les paramètres des motifs. Il est remarquable de noter qu'une image multispectrale peut être vue comme un empilement d'images disposées dans \mathbb{N} (données bidimensionnelles dans un espace unidimensionnel), mais aussi comme un ensemble de spectres disposés dans \mathbb{N}^2 (données unidimensionnelles dans un espace bidimensionnel).

Résumé des travaux

Le problème de la décomposition d'un signal unidimensionnel a été le principal sujet de ma thèse [148]. Ce problème s'appliquait à un signal spectroscopique dont il fallait extraire les raies d'émission. Dans ma thèse, j'ai proposé deux modèles bayésiens dont l'estimation était obtenue à l'aide de méthodes de Monte Carlo par chaînes de Markov (MCMC). La décomposition d'un signal unidimensionnel est la base des autres problèmes listés dans la table 1.1 puisque les données sont les plus simples possible. Aussi, le chapitre 2 fait un état de l'art de ce problème et détaille les difficultés qui sont présentes dans les autres cas.

Le problème de la décomposition de plusieurs signaux unidimensionnels est présenté dans le chapitre 3. En fait, que les signaux soient disposés dans un espace à une ou deux dimensions importe peu puisque dans les deux cas on doit gérer un ensemble de signaux unidimensionnels liés entre eux. La première approche développée a consisté à décomposer les signaux un par un (de manière séquentielle), puis à effectuer un post-traitement pour suivre l'évolution des motifs. Comme on le montre dans la section 3.2, cette procédure n'est pas satisfaisante car les décompositions ne peuvent pas bénéficier des décompositions obtenus pour les signaux voisins. Aussi, les méthodes qui ont été développées dans la suite sont des méthodes qui décomposent l'ensemble des signaux de manière conjointe. Ainsi, dans la continuité de ma thèse, j'ai abordé ce problème en utilisant la modélisation bayésienne et les méthodes MCMC : cette approche est détaillée dans la section 3.3. J'ai notamment traité le cas où le nombre de motifs est inconnu en utilisant un algorithme MCMC trans-dimensionnel. Mais cette approche s'est avérée très complexe à mettre en œuvre pour être efficace. Aussi, nous avons profité de l'intérêt grandissant des représentations parcimonieuses pour les évaluer dans le cadre de la décomposition en motifs. Cette partie est traitée dans la section 3.4.

Le cas de la décomposition d'un signal bidimensionnel en motifs correspond à la détection de motifs dans une image. Beaucoup de travaux existent à ce sujet, je l'ai rapidement abordé lors du stage de master de Fatma Abdelmoulah (2009), mais sans approfondir le sujet.

Une séquence d'images est un bon modèle pour les images multispectrales. C'est ce cadre que nous avons utilisé pour effectuer la décomposition d'images de galaxies en motifs physiques, tels que le bulbe ou le disque. Ce problème est présenté dans le chapitre 4.

Enfin, le cas d'un ensemble d'images réparties dans un espace à deux dimensions est un modèle pour l'IRM de diffusion puisqu'on peut considérer les données comme un ensemble d'images indexées par une pondération de diffusion et une direction spatiale. C'est un modèle qui pourrait également s'appliquer aux assemblages de photos ou aux mosaïques d'images. Je n'ai pas traité ce cas.

1.2 Détection des bâtiments dans des images de télédétection

Contexte

Le Sertit (service régional de traitement d'image et de télédétection) est une structure locale qui produit de l'information géographique à partir d'images satellitaires. Il est chargé de deux types de missions :

- fournir des cartes pour aider la prise de décision chez ses clients (occupation du sol, caractérisation de la biodiversité, zones propices au développement des maladies tropicales, etc.);
- fournir des cartes commentées en quelques heures lors d'événements graves (séismes, inondations, feux de forêt, etc.).

Le Sertit a intégré le laboratoire ICube en septembre 2015 en tant que plate-forme. Il a donc été naturel que le thème IPSEO collabore avec le Sertit. Cette collaboration s'est développée dans le cadre du projet Kal-Haïti (kalideos.fr) dont l'objectif est de construire une base de données d'images de télédétection et de données complémentaires relatives au séisme de janvier 2010 en Haïti. L'un des objectifs du projet Kal-Haïti est de suivre l'état des bâtiments (en bon état, endommagé, en ruine...) en établissant et en gardant à jour une base de données. Évidemment, la mise à jour manuelle de la base de données à partir d'images de télédétection est très coûteuse, et le Sertit souhaitait une procédure plus automatique.

La collaboration entre le Sertit et le thème IPSEO s'est concrétisée par la thèse de Tran-Thanh Ngo (2012–2015), dirigée par Christophe Collet et co-encadrée par moi-même, avec le support de Paul de Fraipont (Sertit) pour l'expertise applicative. Cette thèse a été co-financée par l'Institut Carnot « Télécom et Société numérique » et l'ancienne région Alsace.

L'objectif de la thèse de Tran-Thanh Ngo [183] a été de développer une chaîne de traitements pour détecter les bâtiments, ouvrant ainsi la possibilité d'évaluer l'évolution de leur état. La première étape de cette chaîne de traitements a consisté à détecter l'ombre et la végétation dans les images. Nous avons donc sélectionné des indicateurs appropriés de l'ombre et de la végétation parmi ceux existants dans la littérature. Pour fusionner les résultats fournis par ces indicateurs, Tran-Thanh Ngo a proposé d'utiliser la théorie de Dempster-Shafer [62, 232]. Le résultat obtenu a été amélioré en utilisant les champs de Markov, qui sont un outil faisant partie du champ d'expertise de l'équipe IMAGeS. La seconde étape de la chaîne de traitements a consisté à segmenter l'image en super-pixels, puis nous avons évalué quelles règles de combinaison de ces super-pixels pouvaient finalement segmenter les bâtiments. Les choix effectués durant cette thèse ont été principalement guidés par l'aspect applicatif et ont permis plusieurs contributions dans ce sens. Les travaux ont été publiés en revue internationale [188] et nationale [187], ainsi qu'en conférence [184, 185, 186], dont deux de portée internationale.

Détection de l'ombre et la végétation

La détection de l'ombre ou de la végétation à partir d'images de télédétection bénéficie d'une littérature très abondante (lire par exemple [3, 191, 192, 234, 246, 252]). Cependant, détecter séparément ombre et végétation n'est pas optimal. En effet, un pixel de végétation ombragée sera considéré comme de la végétation par un algorithme de détection de végétation ; au contraire, il sera considéré comme de l'ombre par un algorithme de détection d'ombre. Par ailleurs, une analyse visuelle aboutira également à un problème cornélien : ce pixel est-il de l'ombre ou de la végétation ?

Aussi, nous avons proposé une méthode de détection *simultanée* de l'ombre et de la végétation, qui segmente l'image en trois classes : ombre, végétation et autre.

Le principe de l'approche proposée est de calculer trois indices issus de la littérature pour identifier (séparément) les ombres, la végétation, et les zones sombres. Les trois cartes obtenues, après segmentation binaire, servent de sources d'information à une méthode de fusion effectuée dans le cadre de la théorie de Dempster–Shafer et utilisant une régularisation par champs de Markov.

Une comparaison de différents indices d'ombre et de végétation de la littérature (notamment [3]), a permis de sélectionner deux indices qui donnent les meilleurs résultats sur nos données :

- l'indice c_3 de l'espace de couleur $c_1c_2c_3$ [228] pour la détection de l'ombre ;
- l'indice ExG pour la détection de la végétation [261].

La luminance L a par ailleurs également été prise en compte, car il est apparu qu'elle améliorerait les résultats de la segmentation. Ces trois indices fournissent trois cartes sur lesquelles est appliqué un seuillage de Otsu [197] afin d'obtenir une segmentation initiale des ombres, des végétations et des zones sombres.

Ces trois cartes binaires sont alors fusionnées en utilisant l'approche de Dempster–Shafer [62, 232]. On considère trois hypothèses correspondant aux trois classes de la segmentation. Les fonctions de masse sont modélisées par des gaussiennes : ce choix est motivé par des études empiriques [254]. Leur combinaison est réalisée en utilisant la règle de combinaison prudente de Denœux [63], qui permet de combiner des données dépendantes. Le résultat de la segmentation est obtenu en choisissant, à chaque pixel, l'hypothèse simple de plus grande probabilité pignistique⁸. Un exemple de classification obtenue est donné figure 1.1e.

Or, la fusion est effectuée au niveau du pixel : cela suppose implicitement que chaque pixel ne dépend pas de ses voisins, la procédure est donc très sensible au bruit. Pour cette raison, une régularisation par champ de Markov [87] est introduite dans l'étape de fusion. Notons que l'utilisation conjointe de la théorie de Dempster–Shafer et des champs de Markov dans le cadre de la classification à partir d'observations multisources a déjà été étudié [21]. Pour introduire l'a priori markovien dans la théorie de Dempster–Shafer, il suffit de considérer l'a priori comme une source d'information supplémentaire, ce qui revient à définir une fonction de masse « a priori ». Les fonctions de masse « a posteriori » sont alors calculées en effectuant la somme orthogonale de deux sources d'information évidentielles que sont les fonctions de masses « vraisemblables » et « a priori ». Ce sont ces nouvelles fonctions de masse qui sont utilisées pour effectuer la décision. Le résultat obtenu est donné figure 1.1f : l'effet de la régularisation est clairement visible par rapport au résultat de la figure 1.1e.

Les résultats obtenus avec cette approche ont été comparés avec des approches reconnues de détection de l'ombre [246] et de la végétation [234]. La conclusion est que l'approche proposée, en détectant simultanément l'ombre et la végétation, est meilleure en terme de précision totale [75] et de coefficient de corrélation de Matthews [147].

Détection des bâtiments

La détection de bâtiments à partir d'images de télédétection est également un problème pour lequel de nombreuses solutions ont été proposées, même si l'on ne dispose que d'images RVB [76, 135, 191, 192]. Dans notre cas, nous disposons d'une information supplémentaire : l'heure de prise

8. La décision dite pignistique est utilisée dans le cadre de la règle de Denœux, elle consiste à transformer les fonctions de masse en distributions de probabilité.

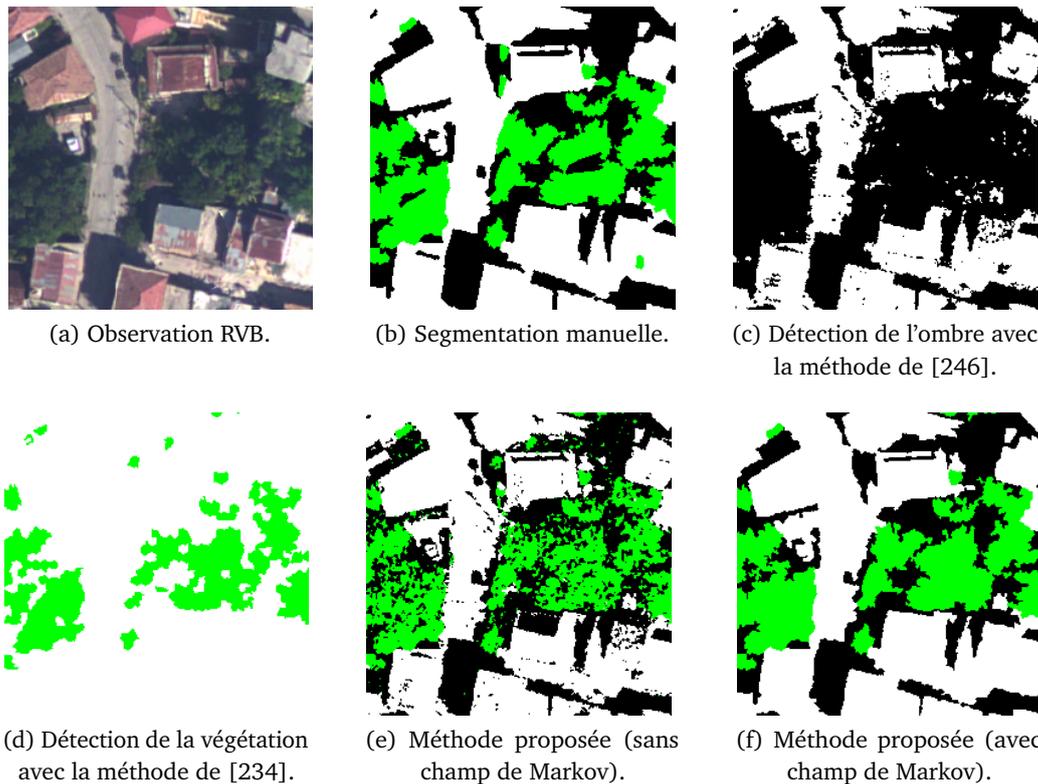


FIGURE 1.1 – Segmentation des ombres (noir), de la végétation (vert) et du reste (blanc).

de vue, et donc l'orientation de l'éclairage, est connue. En revanche, les observations sont des zones de densités très hétérogènes (du rural à l'urbain). Nous n'avons pas non plus accès à un modèle numérique d'élévation qui aurait pu apporter une information d'altimétrie. Aussi, nous formulons les hypothèses suivantes sur l'apparence des bâtiments :

- les bâtiments ont une couleur homogène (cette hypothèse permet de séparer les bâtiments dans les zones denses) ;
- les bâtiments créent une ombre portée ;
- les bâtiments sont plutôt de forme rectangulaire ou sont une combinaison de quelques rectangles.

Tout d'abord, la carte des ombres obtenue précédemment est débarrassée des ombres générées par autre chose que des bâtiments, comme les ombres dues à la végétation ou aux objets de petite taille. Ensuite, l'image originale, en excluant les zones de végétation qui ne sont pas censées contenir des bâtiments, est segmentée en super-pixels à l'aide de la méthode SLIC [2] (figure 1.2b). Les bâtiments seront des groupes de ces super-pixels. Les super-pixels sont classés en fonction de leur couleur dominante, forme et voisinage. Ceux ayant les mêmes caractéristiques, dont la fusion forme une forme rectangulaire, et voisins d'une zone d'ombre dans la direction opposée au soleil sont finalement fusionnés. Ces deux étapes de classification et de fusion sont répétées jusqu'à ce qu'aucune fusion ne soit possible. Un exemple de résultat obtenu à cette étape est visible figure 1.2c. Enfin, les bâtiments sont identifiés en approximant les régions fusionnées par un rectangle.

La méthode proposée a été comparée avec trois méthodes de l'état de l'art qui utilisent également l'information d'ombre [76, 135, 191]. Les résultats obtenus avec notre approche sont meilleurs, tant au niveau pixel qu'au niveau objet (cf. figure 1.3).

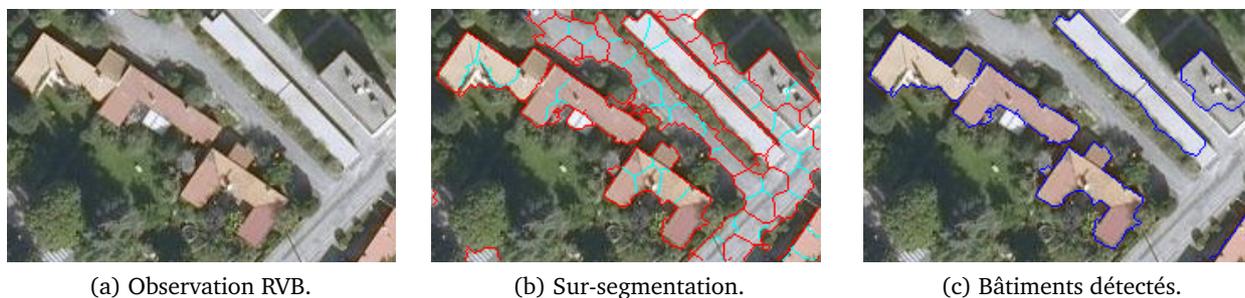


FIGURE 1.2 – Illustration des étapes de la détection des bâtiments.

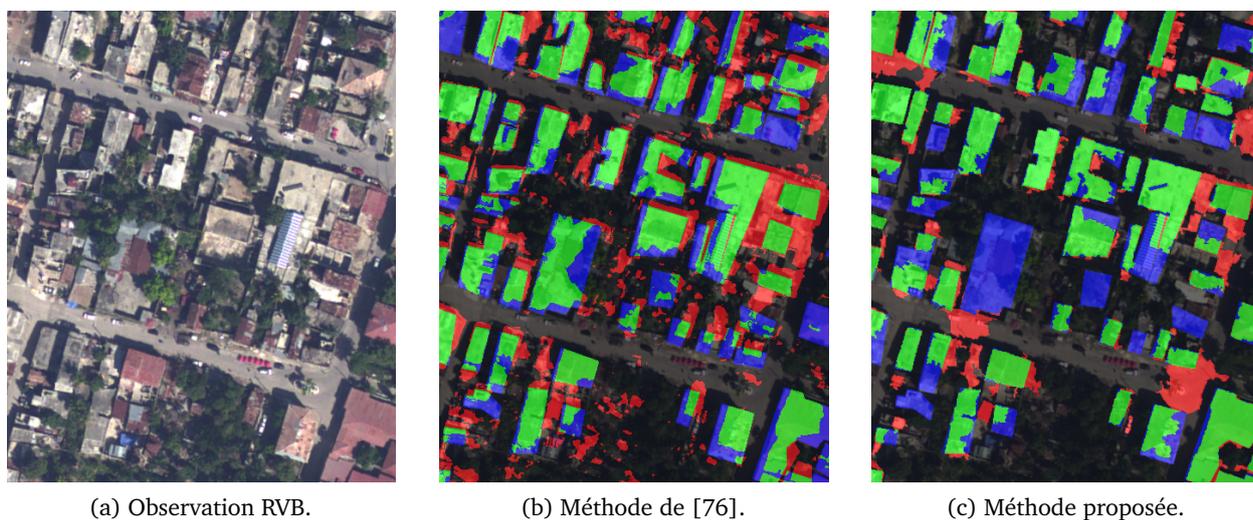


FIGURE 1.3 – Détection des bâtiments à partir de l'observation de gauche (pixels verts : vrai positifs ; pixels rouges : faux positifs ; pixels bleus : faux négatifs).

1.3 Détection de signaux ténus dans des données hyperspectrales fortement bruitées

Contexte

Le thème IPSEO collabore depuis plus de vingt ans avec la communauté astronomique. En particulier, notre collaboration avec l'équipe de Roland Bacon (CRAL) s'est inscrite sur dix ans et s'est réalisée dans le cadre de deux projets de recherche autour de l'instrument MUSE⁹.

Ainsi, nous avons été sollicité en 2008 pour intégrer le projet DAHLIA (2009–2012, porté par Éric Slezak et financé par l'ANR) dont l'objectif était d'explorer diverses techniques pour traiter les images que MUSE allait produire. Suite à ce projet et au moment des premières observations de MUSE, nous avons continué à collaborer avec le CRAL dans le cadre du projet MUSICOS (2014–2019, porté par Roland Bacon et financé par l'ERC). L'objectif de cette collaboration était de développer des algorithmes spécifiques de segmentation et de détection et les appliquer à des nouvelles observations. Ce projet a financé la thèse de Jean-Baptiste Courbot (2014–2017), dirigée par Christophe Collet et Roland Bacon, et encadrée par Emmanuel Monfrini (SAMOVAR) et moi-même.

Dans cette thèse, nous avons étudié le problème de la détection et de la segmentation de halos de gaz autour des galaxies lointaines. Ces halos sont des zones floues et noyées dans le bruit des images.

Tout d'abord, nous avons considéré la détection des halos en utilisant des tests d'hypothèses, dans la lignée des travaux effectués à Nice [199] pendant le projet DAHLIA. La nouveauté a été de tenir compte des aspects spatiaux et spectraux des signatures lumineuses cherchées, mais également de la similarité entre un pixel et plusieurs pixels de référence. Les propriétés statistiques de ces tests ont également été étudiées. Cette première partie a fait l'objet d'un article en revue [50] et d'un autre à ICASSP [49].

Ces premières approches ne donnant pas des résultats satisfaisants (cf. figure 1.4), nous avons alors considéré le problème de la segmentation d'images à l'aide de techniques markoviennes. Plus précisément, l'utilisation des modèles développés par l'équipe de Wojciech Pieczynski (SAMOVAR) a été proposée par Jean-Baptiste Courbot et Emmanuel Monfrini. Ainsi, un modèle de champ de Markov couple a été proposé pour la segmentation d'images convoluées et extrêmement bruitées. Puis, un modèle de champ de Markov triplet a permis de segmenter des images constituées de zones orientées en estimant à la fois les orientations et les classes. Enfin, un modèle d'arbres de Markov triplets a été proposé pour tenir compte du voisinage hiérarchique et spatial pour la segmentation d'images. Ces modèles nous apparaissent comme des alternatives à une représentation bayésienne classique mais sont très peu répandus. Nous avons initié durant la thèse de Jean-Baptiste Courbot une étude (toujours en cours) pour mesurer les différences et les similarités entre ces deux approches. Cette seconde partie a fait l'objet de deux articles en revue [54, 51], deux articles en conférence internationale [52, 55] et un article en conférence nationale [53].

Détection par tests d'hypothèses dans les images hyperspectrales

Un défi majeur de MUSE est de détecter les halos de gaz entourant les galaxies. Ces halos émettent dans les longueurs d'onde de l'hydrogène, l'une des raies d'émission caractéristique est nommée raie

9. MUSE (*multi unit spectroscopic explorer*) est un spectro-imageur intégral de champ installé depuis 2014 sur l'un des télescopes du VLT au Chili. Cet instrument produit des images multispectrales de 300×300 pixels sur presque 4000 longueurs d'ondes dans le visible et l'infrarouge proche.

Lyman-alpha. Notre objectif a donc été de détecter ces halos faiblement brillants et n'émettant que sur quelques longueurs d'onde (sur près de 4000 !) situés en périphérie d'une galaxie.

Nous supposons donc que le spectre d'émission de la galaxie et de son halo ne présente que quelques coefficients non nuls correspondant à la raie Lyman-alpha. Spatialement, l'intensité lumineuse s'atténue en s'éloignant du centre : seuls les spectres au centre de la galaxie peuvent donc être considérés brillants lorsque tous les autres sont ténus.

Plusieurs tests du rapport de vraisemblance généralisé ont été formulés, la plupart étant contraints à la manière du $\text{GLR}_{1s}^{(1D)}$ proposé dans [199]. Les tests incluent ou non des contraintes spatiales, d'observations multiples, de similarité et ou sont composites.

La première étape de l'approche développée est la détection des spectres brillants, grâce à un test du rapport de vraisemblance généralisé appliqué à chaque spectre \mathbf{y}_s de l'image. Sous l'hypothèse nulle, le spectre est supposé n'être que du bruit alors que sous l'hypothèse alternative, il est supposé être la somme d'un bruit avec un signal particulier μ_s . Dans le cadre de notre application, le signal μ_s est inconnu mais contraint à être 1-parcimonieux puisqu'il modélise la raie Lyman-alpha.

La seconde étape de l'approche développée consiste à effectuer une détection des spectres ténus et s'appuie sur la connaissance des spectres détectés lors de la première étape ; l'ensemble des sites de ces spectres est noté \mathcal{B} . Ici encore, un test du rapport de vraisemblance généralisé est appliqué à chaque spectre $s \notin \mathcal{B}$. Plusieurs variantes de ces tests sont testées :

- pour chaque site $s \notin \mathcal{B}$, on impose la contrainte que le signal μ_s est proportionnel à l'un des signaux μ_b ($b \in \mathcal{B}$) de la partie brillante. Il est possible de qualifier les probabilités de fausses alarmes de ce test ;
- une autre possibilité est d'imposer pour chaque site $s \notin \mathcal{B}$ la contrainte que le signal μ_s est proportionnel à l'ensemble des signaux μ_b ($b \in \mathcal{B}$) de la partie brillante. Cette combinaison est formulée comme un produit, afin de pouvoir qualifier les probabilités de fausses alarmes du test ;
- par ailleurs, nous avons formulé des tests en tenant compte de la réponse impulsionnelle de l'instrument ;
- une dernière variante consiste à combiner plusieurs observations d'un même champ, sans avoir à fusionner ces observations multiples.

Champs de Markov couples

Le problème de détection précédent peut être réécrit comme un cas particulier de segmentation en deux classes. La segmentation d'images multispectrales à l'aide de modèles bayésiens a fait l'objet d'un grand nombre de travaux [72, 136, 137, 217, 231, 262]. Tous reposent sur des modèles de champs de Markov cachés.

Nous avons souhaité explorer les modèles de Markov couples introduits il y a une vingtaine d'années par Pieczynski et Tebbache [212]. Ce modèle est fondé sur l'observation qu'un algorithme de segmentation, dans le cas d'une image modélisée par un champ de Markov caché, ne se base que sur l'hypothèse que $p(\mathbf{x}|\mathbf{y})$ est un processus de Markov. Cette hypothèse est vérifiée lorsque \mathbf{x} est markovien, mais aussi lorsque c'est le couple (\mathbf{x}, \mathbf{y}) qui est markovien. Cette seconde condition correspond au cas, plus général, des modèles de champs de Markov couple.

Comme on le voit sur la figure 1.4, l'a priori markovien apporte une homogénéité spatiale de l'estimation. C'est sur ce point que le modèle proposé diffère de ses alternatives, qui ne prennent l'information de voisinage en compte que localement. Il est possible de calculer deux estimateurs,

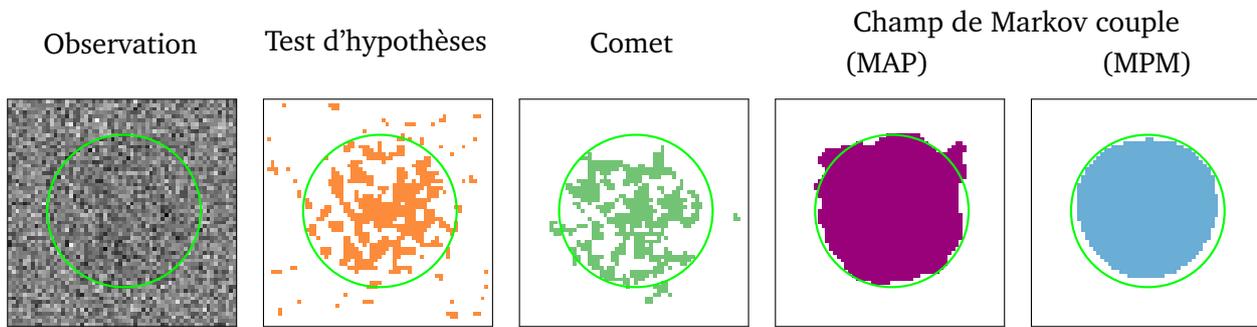


FIGURE 1.4 – Exemples de résultats sur une observation synthétique avec un rapport signal-à-bruit de -14 dB. Le cercle vert représente le contour de la vérité terrain. Les résultats obtenus avec le test d'hypothèse et le champ de Markov couple (pour les deux estimateurs) sont comparés à la méthode Comet [12] développée par Raphaël Bacher dans le cadre du projet MUSICOS.

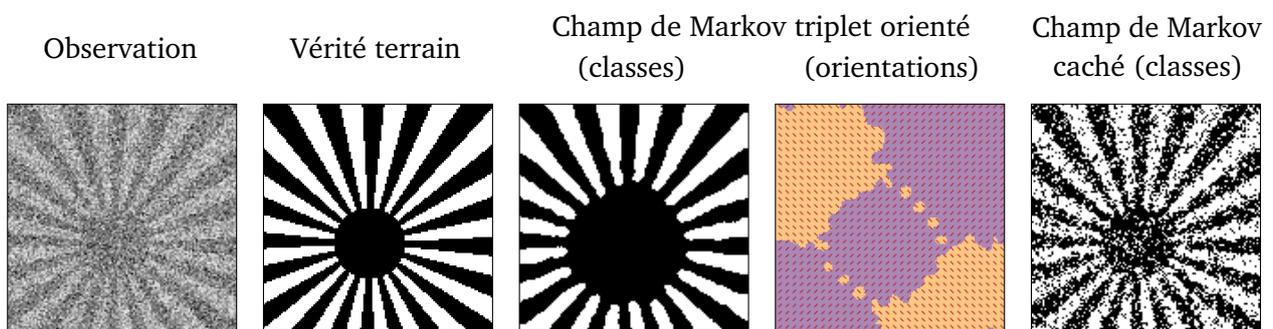


FIGURE 1.5 – Segmentation d'une image synthétique avec un rapport signal-à-bruit de -6 dB. L'estimation des classes au sens du MPM obtenue avec le champ de Markov triplet orienté est comparé à la vérité terrain et à l'estimation obtenue avec un champ de Markov caché.

à savoir le maximum a posteriori (MAP) et le *maximum posterior mode* (MPM) qui permettent respectivement d'obtenir de faibles valeurs des taux d'erreurs et de fausses alarmes. En outre, nous pouvons tirer une mesure de confiance, ou incertitude, de l'estimation au sens MPM, en comparant les probabilités des deux classes en chaque site. La modélisation par champ de Markov couple est encore à l'étude, notamment en la comparant avec des champs de Markov cachés.

Champs de Markov triplets

Les champs de Markov triplets sont un autre modèle proposé par Pieczynski et coll. [211]. Aux données \mathbf{y} et aux inconnues \mathbf{x} est ajoutée une variable supplémentaire \mathbf{v} de telle sorte que le triplet $\mathbf{t} = (\mathbf{y}, \mathbf{x}, \mathbf{v})$ soit un champ de Markov. Nous nous sommes intéressés à la segmentation d'images présentant des textures orientées. Il existe peu de travaux fondés sur la modélisation markovienne des structures orientées. À notre connaissance, les travaux les plus adaptés sont ceux de Dass [58]. L'intérêt des champs de Markov triplets est d'estimer conjointement les classes et les orientations. Ainsi, \mathbf{y} est l'image observée, \mathbf{x} regroupe les classes des pixels et \mathbf{v} regroupe les orientations sous-jacentes. En plus de l'estimation des classes et des orientations, il est possible d'évaluer l'incertitude associée aux segmentations à condition d'utiliser l'estimateur du MPM. Un exemple de segmentation est donné figure 1.5 : le champ de Markov triplet fourni un résultat plus « propre » que celui obtenu par un modèle classique de champ de Markov caché.

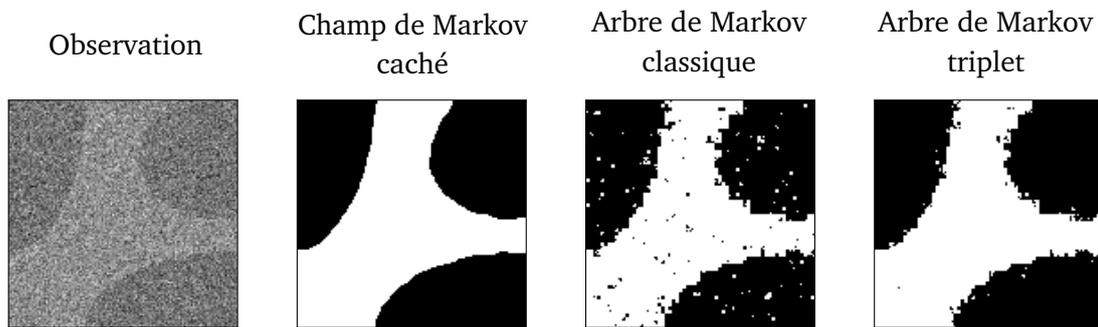


FIGURE 1.6 – Segmentation d’une image synthétique avec un rapport signal-à-bruit de -8 dB. La meilleure segmentation est bien sûr obtenue avec un champ de Markov caché, mais dans le cadre des arbres de Markov, le triplet permet de minimiser l’effet de bloc observé avec un arbre classique.

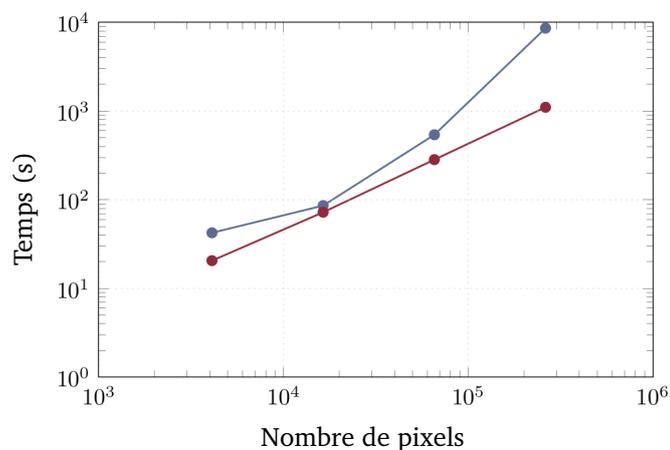


FIGURE 1.7 – Temps de calculs moyens (sur 10 réalisations) pour la segmentation avec un champ de Markov caché (bleu) et un arbre de Markov triplet (rouge).

Arbres de Markov triplets

Les champs de Markov ont l’avantage d’apporter une grande richesse de modélisation, mais ils dépendent très souvent d’algorithmes stochastiques, ce qui implique un temps de calcul conséquent. Les arbres de Markov apportent un compromis entre un modèle polyvalent et le temps de calcul [131]. Ils définissent une hiérarchie de résolution sur laquelle porte l’aspect markovien : chaque pixel de l’image est un « enfant », et chaque carré de quatre enfants est regroupé en un « parent ». De même, les carrés de quatre parents sont regroupés et ainsi de suite jusqu’à un parent unique appelé « racine ». Les arbres de Markov ont cependant comme principal défaut d’introduire un effet de « blocs » dans l’estimation. Plusieurs types de modèles ont été proposés pour compenser cet effet [166, 172]. Les arbres de Markov, tout comme les modèles par chaînes ou par champs, ont été enrichis par l’introduction de modèles triplets [210] en ajoutant, en plus du processus observé y et du processus de classe x , un troisième processus v dans le but de modéliser des phénomènes complexes. Ainsi, nous avons proposé un modèle d’arbre de Markov triplet dans lequel le processus auxiliaire permet de redéfinir les transitions parent/enfant, en fonction de la classe des voisins du parent. Cette formulation facilite la modélisation de structures homogènes à grande échelle, et évite les effets de blocs, comme on peut l’observer figure 1.6. Même si la qualité de la segmentation reste en deçà d’une segmentation avec un champ de Markov caché, elle a l’avantage d’être plus rapide (figure 1.7).

Application aux données MUSE

Parmi les quatre techniques développées durant la thèse de Jean-Baptiste Courbot, les deux plus pertinentes pour être appliquées aux données de MUSE sont le champ de Markov couple et l'arbre de Markov triplet. Ces deux techniques ont donc été utilisées pour une observation de champs profonds, c'est-à-dire une région du ciel contenant des objets très lointains de l'Univers jeune. Cette région du ciel est le *Hubble Ultra Deep Field* qui a été observé la première fois en 2003 par le télescope spatial Hubble [13, 18].

La localisation spatiale et spectrale des émetteurs de Lyman-alpha a été obtenue indépendamment par des experts. Chaque objet détecté est catalogué et identifié par un numéro [110]. Au total, 242 émetteurs de Lyman-alpha ont été traités ; la figure 1.8 en présente trois.

Plusieurs commentaires peuvent être tirés de l'inspection visuelle de ces résultats. Tout d'abord, la méthode basée sur le champ de Markov couple semble produire des résultats satisfaisants (l'objet 6298 par exemple). Néanmoins, ce modèle ne permet pas la détection à très faible rapport signal-à-bruit, comme l'illustre l'objet 148, bien que la carte d'incertitude associée puisse aider à repérer des objets très peu lumineux, comme dans le cas de l'objet 2069. La détection par arbre de Markov triplet est une alternative pertinente à faible rapport signal-à-bruit : l'objet 148 est bien détecté par cette méthode.

Le traitement des 242 objets montre que la méthode basée sur le champ de Markov couple permet

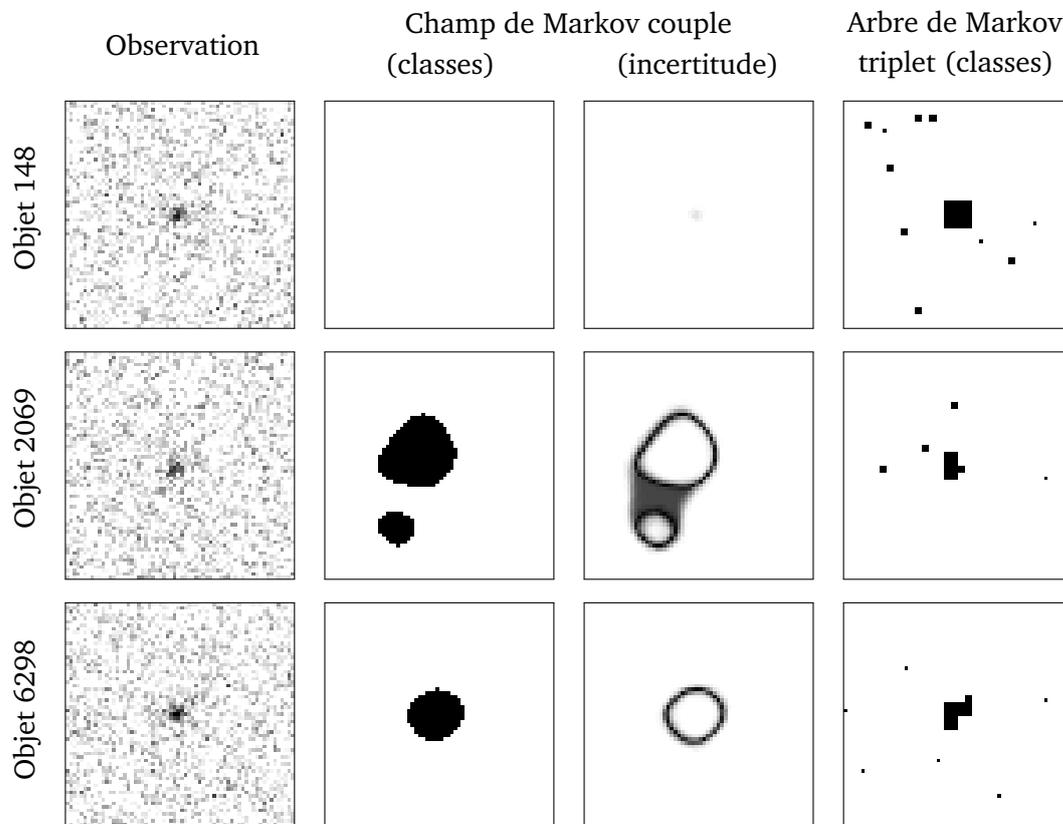


FIGURE 1.8 – Segmentation de trois objets avec un champ de Markov couple (estimateur du MPM) et un arbre de Markov triplet. L'observation est représentée en inverse vidéo. L'incertitude associée à la méthode par champ de Markov couple s'étale du blanc (très certain) au noir (très incertain).

une détection de 59,5 % des objets considérés. La méthode par arbre de Markov triplet complète ce résultat en détectant 16,5 % d'objets supplémentaires. Finalement, la combinaison de ces deux modèles produit un taux de détection de 76 %.

1.4 Publications personnelles

Thèse de doctorat

[T-1] V. MAZET : *Développement de méthodes de traitement de signaux spectroscopiques : estimation de la ligne de base et du spectre de raies*. Thèse de doctorat, Université Henri Poincaré, Nancy 1, 2005.

Chapitres dans des ouvrages

[O-1] C. COLLET, B. PERRET et V. MAZET : *Multivariate Image Processing*, chapitre Panoramic Integral-Field Spectrograph : Ultraspectral Data to Understand the History of the Universe, pages 437–450. ISTE Ltd and John Wiley & Sons, 2010.

[O-2] V. MAZET, F. FLITTI et C. COLLET : *Multivariate Image Processing*, chapitre Detection and Tracking of Emission Rays in Radioastronomy, pages 169–200. ISTE Ltd and John Wiley & Sons, 2010.

Articles dans des revues internationales avec comité de lecture

[RI-1] S. AWALI, L. POISSON, B. SOEP, M.-A. GAVEAU, M. BRIANT, C. POTHIER, J.-M. MESTDAGH, M. Ben El Hadj RHOUMA, M. HOCHLAF, V. MAZET et S. FAISAN : Time resolved observation of the solvation dynamics of a Rydberg excited molecule deposited on an Argon cluster — I : DABCO* at short times. *Physical Chemistry Chemical Physics*, 2014.

[RI-2] J.-B. COURBOT, V. MAZET, E. MONFRINI et C. COLLET : Extended faint source detection in astronomical hyperspectral images. *Signal Processing*, 135:274–283, juin 2017.

[RI-3] J.-B. COURBOT, V. MAZET, E. MONFRINI et C. COLLET : Pairwise Markov fields for segmentation in astronomical hyperspectral images. *Signal Processing*, pages 41–48, 2019.

[RI-4] J.-B. COURBOT, E. MONFRINI, V. MAZET et C. COLLET : Oriented triplet Markov fields. *Pattern Recognition Letters*, 103:16–22, février 2018.

[RI-5] A. LIETARD, G. PIANI, M. BRIANT, M.-A. GAVEAU, S. FAISAN, V. MAZET, B. SOEP, J.-M. MESTDAGH et L. POISSON : Self-trapping relaxation decay investigated by time-resolved photoelectron spectroscopy. *Physical Chemistry Chemical Physics*, 16, juin 2018.

[RI-6] A. MASSON, L. POISSON, M.-A. GAVEAU, B. SOEP, J.-M. MESTDAGH, V. MAZET et F. SPIEGELMAN : Dynamics of highly excited Barium atoms deposited on large Argon clusters. I. General trends. *The Journal of Chemical Physics*, 133(054307), 2010.

[RI-7] V. MAZET : Joint Bayesian decomposition of a spectroscopic signal sequence. *IEEE Signal Processing Letters*, 18(3):181–184, mars 2011.

[RI-8] V. MAZET, D. BRIE et C. CAIRONI : Sparse spike train deconvolution using the Hunt filter and a thresholding method. *IEEE Signal Processing Letters*, 11(5):486–489, mai 2004.

- [RI-9] V. MAZET, C. CARTERET, D. BRIE, J. IDIER et B. HUMBERT : Background removal from spectra by designing and minimising a non-quadratic cost function. *Chemometrics and Intelligent Laboratory Systems*, 76(2):121–133, avril 2005.
- [RI-10] V. MAZET, S. FAISAN, S. AWALI, M.-A. GAVEAU et L. POISSON : Unsupervised joint decomposition of a spectroscopic signal sequence. *Signal Processing*, 109:193–205, 2015.
- [RI-11] H. MORTADA, V. MAZET, C. SOUSSEN, C. COLLET et L. POISSON : Parameterized source separation for delayed spectroscopic signals. *Signal Processing*, 158:48–60, 2019.
- [RI-12] T.-T. NGO, V. MAZET, C. COLLET et P. de FRAIPONT : Shape-based building detection in visible band images using shadow information. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 10(3):920–932, mars 2017.
- [RI-13] B. PERRET, V. MAZET, C. COLLET et É. SLEZAK : Hierarchical multispectral galaxy decomposition using a MCMC algorithm with multiple temperature simulated annealing. *Pattern Recognition*, 44(6):1328–1342, juin 2011.

Articles dans des revues nationales avec comité de lecture

- [RN-1] V. MAZET, S. FAISAN, A. MASSON, M.-A. GAVEAU, L. POISSON et J.-M. MESTDAGH : Approche bayésienne pour la décomposition conjointe d’une séquence de spectres de photo-électrons. *Traitement du signal*, 30:9–34, 2013.
- [RN-2] T.T. NGO, C. COLLET et V. MAZET : Détection simultanée de l’ombre et la végétation sur des images aériennes couleur en haute résolution. *Traitement du Signal*, 32:311–333, 2015.

Articles dans des conférences internationales

- [CI-1] J.-B. COURBOT, V. MAZET, E. MONFRINI et C. COLLET : Detection of faint extended sources in hyperspectral data and application to HDF-S MUSE observations. *In ICASSP*, Shanghai, Chine, 2016.
- [CI-2] J.-B. COURBOT, E. MONFRINI, V. MAZET et C. COLLET : Oriented triplet Markov fields for hyperspectral image segmentation. *In WHISPERS*, Los Angeles, États-Unis, 2016.
- [CI-3] J.-B. COURBOT, E. MONFRINI, V. MAZET et C. COLLET : Triplet Markov trees for image segmentation. *In IEEE Statistical Signal Processing Workshop*, Freiburg, Allemagne, 2018.
- [CI-4] V. MAZET, D. BRIE et J. IDIER : Baseline spectrum estimation using half-quadratic minimization. *In EUSIPCO*, Vienne, Autriche, 2004.
- [CI-5] V. MAZET, D. BRIE et J. IDIER : Simulation of positive normal variables using several proposal distributions. *In IEEE Workshop Statistical Signal Processing*, Bordeaux, France, 2005.
- [CI-6] V. MAZET, D. BRIE et J. IDIER : Decomposition of a chemical spectrum using a marked point process and a constant dimension model. *In MaxEnt*, Paris, France, 2006.
- [CI-7] V. MAZET, C. COLLET et B. VOLLMER : Decomposition and classification of spectral lines in astronomical radio data cubes. *In SCIA*, Oslo, Norvège, 2009.
- [CI-8] V. MAZET, S. FAISAN, A. MASSON, M.-A. GAVEAU et L. POISSON : Unsupervised joint Bayesian decomposition of a sequence of photoelectron spectra. *In Whispers*, Lisbonne, Portugal, 2011.
- [CI-9] V. MAZET, S. FAISAN, A. MASSON, M.-A. GAVEAU, L. POISSON et J.-M. MESTDAGH : Joint Bayesian decomposition of a spectroscopic signal sequence with RJMCMC. *In IEEE Workshop on Statistical Signal Processing*, Ann Arbor, États-Unis, 2012.

- [CI-10] H. MORTADA, V. MAZET, C. SOUSSEN et C. COLLET : Separation of delayed parameterized sources. *In EUSIPCO*, Kos, Grèce, 2017.
- [CI-11] H. MORTADA, V. MAZET, C. SOUSSEN et C. COLLET : Spectroscopic decomposition of astronomical multispectral images using B-splines. *In Whispers*, Amsterdam, Pays-Bas, 2018.
- [CI-12] T.-T. NGO, C. COLLET et V. MAZET : MRF and Dempster–Shafer theory for simultaneous shadow/vegetation detection on high resolution aerial color images. *In ICIP*, Paris, France, 2014.
- [CI-13] T.-T. NGO, C. COLLET et V. MAZET : Automatic rectangular building detection from VHR aerial imagery using shadow and image segmentation. *In ICIP*, Québec, Canada, 2015.
- [CI-14] B. PERRET, V. MAZET, C. COLLET et É. SLEZAK : Galaxy decomposition in multispectral images using Markov chain Monte Carlo algorithms. *In SCIA*, Oslo, Norvège, 2009.
- [CI-15] M. PETREMAND, M. LOUYS, C. COLLET, V. MAZET, A. JALOBÉANU et F. SALZENSTEIN : New Bayesian fusion scheme and visualization tool for astronomical hyperspectral data cubes. *In Astronomical and Data Analysis*, Monastir, Tunisie, 2010.
- [CI-16] D. SERRE, E. VILLENEUVE, H. CARFANTAN, V. MAZET, S. BOURGUIGNON, A. JARNO et L. JO-LISSAINT : Modeling the spatial PSF at the VLT focal plane for MUSE data deconvolution purpose. *In SPIE Symposium on Astronomical Telescopes and Instrumentation*, San Diego, États-Unis, 2010.

Articles dans des conférences nationales

- [CN-1] J.B. COURBOT, E. MONFRINI, V. MAZET et C. COLLET : Arbres de Markov triplets pour la segmentation d'image. *In GRETSI*, Juan-les-Pins, France, 2017.
- [CN-2] V. MAZET : Simulations pédagogiques en traitement du signal avec JavaScript. *In GRETSI*, Juan-les-Pins, France, 2017.
- [CN-3] V. MAZET, D. BRIE et C. CAIRONI : Déconvolution impulsionnelle par filtre de Hunt et seuillage. *In GRETSI*, Paris, France, 2003.
- [CN-4] V. MAZET, D. BRIE et J. IDIER : Simuler une distribution normale à support positif à partir de plusieurs lois candidates. *In GRETSI*, Louvain-la-Neuve, Belgique, 2005.
- [CN-5] V. MAZET, C. COLLET et B. VOLLMER : Décomposition et classification de composantes spectrales dans des cubes de données radio-astronomiques. *In GRETSI*, Dijon, France, 2009.
- [CN-6] V. MAZET, S. FAISAN, A. MASSON, M.-A. GAVEAU et L. POISSON : Approche bayésienne pour la décomposition conjointe d'une séquence de spectres de photoélectrons. *In GRETSI*, Bordeaux, France, 2011.
- [CN-7] V. MAZET, S. FAISAN, L. POISSON, M.-A. GAVEAU et J.-M. MESTDAGH : Décomposition d'une séquence de spectres avec modèle markovien et algorithme RJMCMC à deux variables de dimension. *In GRETSI*, Brest, France, 2013.
- [CN-8] V. MAZET, J. IDIER et D. BRIE : Déconvolution impulsionnelle positive myope. *In GRETSI*, Louvain-la-Neuve, Belgique, 2005.
- [CN-9] V. MAZET, J. IDIER, D. BRIE, B. HUMBERT et C. CARTERET : Estimation de l'arrière-plan de spectres par différentes méthodes dérivées des moindres carrés. *In Chimométrie*, Paris, France, 2003.
- [CN-10] V. MAZET, C. SOUSSEN et E.-H. DJERMOUNE : Décomposition de spectres en motifs paramétriques par approximation parcimonieuse. *In GRETSI*, Brest, France, 2013.

- [CN-11] H. MORTADA, V. MAZET, C. SOUSSEN et C. COLLET : Séparation de sources retardées et paramétriques. *In GRETSI*, Juan-les-Pins, France, 2017.
- [CN-12] T.-T. NGO, C. COLLET et V. MAZET : Détection simultanée de l'ombre et la végétation sur des images aériennes couleur en haute résolution. *In RFIA*, Rouen, France, 2014.
- [CN-13] B. PERRET, V. MAZET, C. COLLET et É. SLEZAK : Décomposition d'images multispectrales de galaxies au moyen d'algorithmes de Monte Carlo par chaînes de Markov. *In GRETSI*, Dijon, France, 2009.
- [CN-14] E. VILLENEUVE, H. CARFANTAN, A. JARNO, D. SERRE, V. MAZET et S. BOURGUIGNON : Modélisation et estimation de la PSF d'un instrument hyperspectral au sol pour l'astrophysique. *In GRETSI*, Bordeaux, France, 2011.

Chapitre 2

Décomposition d'un signal spectroscopique en raies

Le problème étudié dans ce chapitre concerne la décomposition d'un signal unidimensionnel $\mathbf{y} \in \mathbb{R}^N$ en une somme de motifs paramétriques $\varphi \in \mathbb{R}^N$ dont on cherche à estimer leur position, leurs caractéristiques (par exemple leur intensité, leur forme, etc.) et éventuellement leur nombre K (sans se restreindre à une valeur maximale de K). Ce problème s'apparente au problème d'ajustement de courbe, appelé communément *curve fitting* ou *peak fitting*. Le modèle, sous l'hypothèse d'un bruit $\mathbf{b} \in \mathbb{R}^N$ additif, est le suivant :

$$\mathbf{y} = \sum_{k=1}^K a_k \varphi(c_k, w_k) + \mathbf{b} \quad (2.1)$$

où K est le nombre de motifs et a_k, c_k, w_k représentent respectivement l'amplitude, le centre et les paramètres de forme du motif k . Notons que dans plupart des applications l'amplitude est un paramètre linéaire, ce qui permet de l'écrire en facteur de φ . De plus, les paramètres de formes peuvent être nombreux ; par simplicité nous considérons qu'il n'y en a qu'un par motif, ce qui explique la notation scalaire de w_k .

La décomposition d'un signal unidimensionnel en motifs se rencontre dans des disciplines très diverses. Malgré cela, on retrouve les mêmes idées, les mêmes modèles et les mêmes méthodes dans toutes ces disciplines, sans pour autant que ces similarités soient explicitées dans les publications. En spectroscopie, on s'intéresse à la détection des raies d'émission ou d'absorption : les motifs φ sont généralement supposés à support borné, unimodaux et symétriques. Plusieurs modèles sont utilisés (gaussienne, lorentzienne, fonctions de Pearson VII, pseudo-fonction de Voigt, etc.) en fonction du type de spectroscopie et de la précision souhaitée. Ce problème est considéré depuis plusieurs décennies et pour tout type de spectroscopies : spectroscopie de photoélectrons [146], de masse [8, 45], des pertes d'énergie (EELS : *electron energy loss*) [78], infrarouge [180, 148], gamma [97, 249], optique [30], Raman [140, 148, 264] ou RMN (résonance magnétique nucléaire) [45, 167, 168, 226]. En sismique-réflexion, on cherche à retrouver dans un signal des impulsions caractéristiques des milieux géologiques. Les travaux en sismique-réflexion sont déjà anciens puisqu'on trouve des références dès les années 1980 avec Kormylo et Mendel [126] ou Oldenburg, Levy et Stinson [193]. Depuis, la sismique-réflexion a continué à intéresser de nombreux chercheurs [67, 103, 109, 114, 119, 190, 216, 224]. L'estimation des raies (*line spectra estimation*) dans des spectres de Fourier peut également être traitée comme la décomposition d'un signal en motifs [14]. Enfin, d'autres études modélisent les motifs φ par des sinusoides ou des exponentielles complexes [7, 29, 143, 241].

Le modèle décrit par l'équation 2.1 convient également à des signaux multidimensionnels, comme

des images. C'est un modèle similaire qui est utilisé pour détecter des galaxies lointaines en imagerie astronomique [164] ou des objets particuliers en télédétection [129, 196].

Dans ce chapitre, nous prendrons l'exemple de la décomposition d'un signal spectroscopique¹ en raies d'émission, ces dernières étant donc les motifs recherchés. Sans perdre en généralité, on considère le modèle répandu de motifs gaussiens [117] :

$$\forall n \in \{1, \dots, N\}, \quad \varphi_n(c_k, w_k) = \exp\left(-\frac{(\lambda_n - c_k)^2}{2w_k^2}\right)$$

où λ_n est la valeur de la variable d'évolution (typiquement la longueur d'onde) à l'échantillon $n \in \{1, \dots, N\}$. Le paramètre de forme w_k est égal à la largeur à mi-hauteur de la raie divisée par $2\sqrt{2\ln(2)}$.

Ce chapitre fait le point sur le problème de la décomposition d'un signal spectroscopique en raies, en mettant clairement en évidence les difficultés du problème et en donnant un état de l'art des méthodes de résolution.

2.1 Difficultés de la décomposition en raies

Si des méthodes simples échouent à obtenir une décomposition correcte d'un signal spectroscopique, c'est en premier lieu parce que le problème est mal posé. En effet, la forme des raies et le niveau de bruit rendent l'identification des raies délicate. En outre, le critère n'est pas du tout convexe et les raies peuvent se superposer et être en nombre inconnu : ces deux points sont détaillés ci-après.

Échantillonnage irrégulier

Les signaux spectroscopiques ne sont pas toujours échantillonnés régulièrement. Je pensais initialement qu'une raie contenant peu d'échantillons pouvait être moins bien estimée qu'une raie très bien échantillonnée, puisque la plupart des méthodes cherchent à minimiser l'énergie du résidu entre les observations et le signal reconstruit, ou toute quantité équivalente. Par conséquent, une raie de petite amplitude mais bien échantillonnée pourrait-elle plus influencer l'estimation qu'une raie de grande amplitude mais peu échantillonnée ? Autrement dit, un échantillonnage irrégulier peut-il poser une difficulté pour l'estimation ?

En fait, il n'en est rien, du moins dans une certaine mesure.

La figure 2.1 représente une même raie échantillonnée avec deux pas différents. Ces deux signaux ont même puissance $\|\mathbf{y}\|^2/N$, égale à $35,609 \cdot 10^{-3}$. Cette expérience peut être conduite pour plusieurs valeurs du pas d'échantillonnage Δ : la figure 2.2 représente l'évolution de l'énergie en fonction de Δ pour une raie de paramètres $a^* = 1$, $c^* = 100$ et $w^* = 4$, dans le cas non bruité. On observe que lorsque le pas d'échantillonnage est inférieur à 6, l'énergie reste à peu près constante. On peut reproduire cette figure pour d'autres paramètres de raies : il apparaît que la valeur maximale du pas d'échantillonnage permettant de conserver une énergie à peu près constante est compris entre 1 et 1,5 fois la largeur w de la raie (les pas d'échantillonnage utilisés dans la figure 2.1 correspondent à ces deux valeurs). Tant que cette condition est satisfaite, l'échantillonnage n'a donc pas d'influence sur la qualité de l'estimation.

1. J'utiliserai le terme « signal spectroscopique » plutôt que « spectre » qui peut être confondu avec la représentation obtenue après transformée de Fourier.

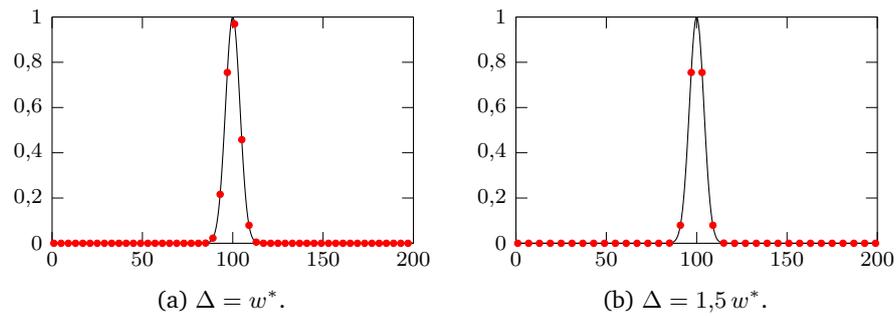


FIGURE 2.1 – Exemples de signaux spectroscopiques échantillonnés à différents pas Δ (ici, $w^* = 4$). Les échantillons sont représentés par des points sur le signal continu.

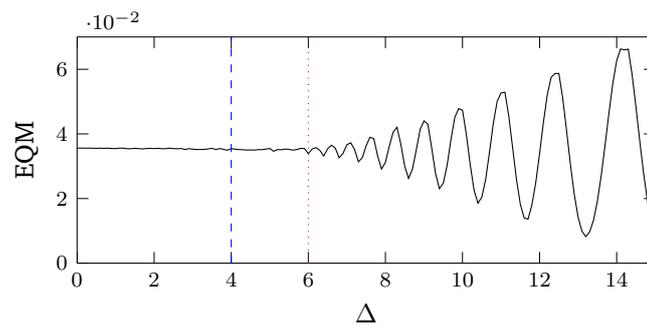


FIGURE 2.2 – Évolution de l'EQM en fonction du pas d'échantillonnage. Les lignes verticales correspondent à 1 (tirets) et 1,5 (pointillés) fois la largeur $w^* = 4$.

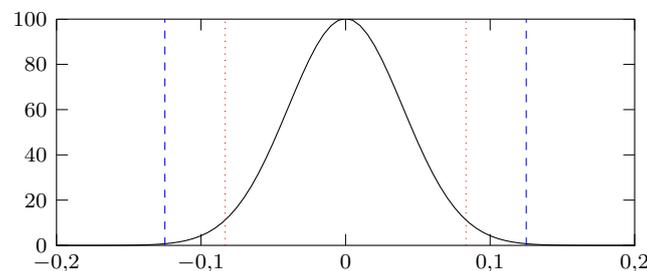
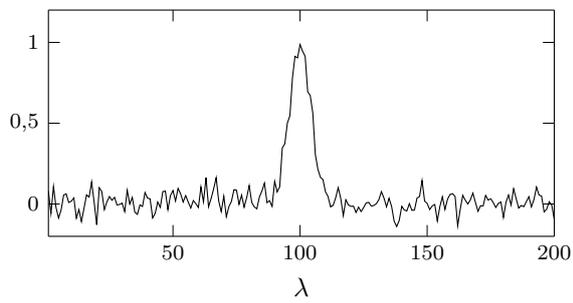


FIGURE 2.3 – Module de la transformée de Fourier discrète d'une gaussienne. Les deux lignes verticales représentent les fréquences maximales correspondant à 1 (tirets) et 1,5 (pointillés) fois la largeur w^* .

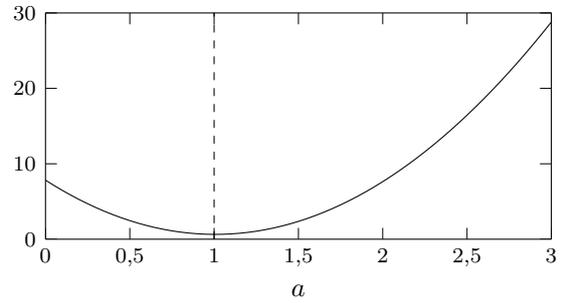
Cette conclusion s'explique très bien en observant le spectre d'une gaussienne : le module de la transformée de Fourier discrète d'une gaussienne échantillonnée très finement est représenté figure 2.3. Les fréquences correspondant à des pas d'échantillonnage de 1 et 1,5 fois la largeur w^* de la raie sont également représentés. On remarque qu'un pas égal à w^* permet de conserver la quasi totalité de l'énergie du spectre, contrairement à un pas égal à 1,5 fois w^* .

Optimisation du critère

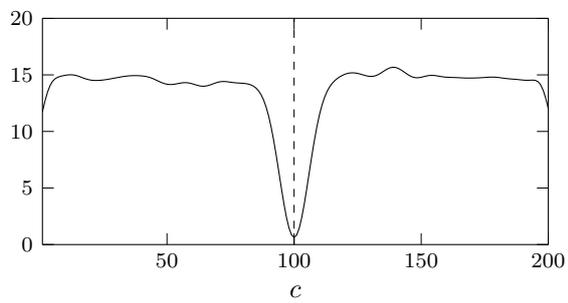
En supposant le nombre de raies K connu, l'estimation des paramètres c , a et w au sens du maximum de vraisemblance sous hypothèse de bruit b gaussien centré revient à minimiser le critère



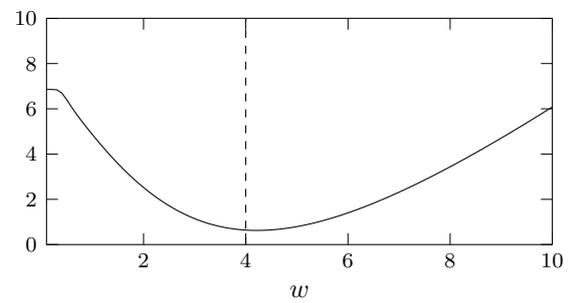
(a) signal spectroscopique synthétique.



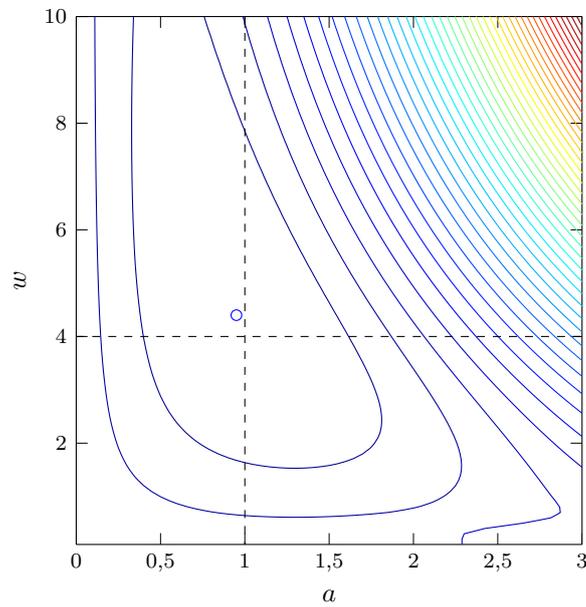
(b) $\mathcal{C}(a, c^*, w^*)$.



(c) $\mathcal{C}(a^*, c, w^*)$.



(d) $\mathcal{C}(a^*, c^*, w)$.



(e) $\mathcal{C}(a, c^*, w)$.

FIGURE 2.4 – Évolution du critère $\mathcal{C}(a, c, w)$ (2.2) appliqué au signal spectroscopique (a). Les lignes en tirets représentent la valeur réelle du paramètre considéré. Sur le critère 2D, le minimum du critère est indiqué par \circ .

des moindres carrés :

$$\hat{\mathbf{a}}, \hat{\mathbf{c}}, \hat{\mathbf{w}} = \arg \max_{\mathbf{a}, \mathbf{c}, \mathbf{w}} \mathcal{C}(\mathbf{a}, \mathbf{c}, \mathbf{w}) \quad \text{où} \quad \mathcal{C}(\mathbf{a}, \mathbf{c}, \mathbf{w}) = \left\| \mathbf{y} - \sum_{k=1}^K a_k \varphi(c_k, w_k) \right\|^2. \quad (2.2)$$

Le modèle étant linéaire en \mathbf{a} , l'estimation des amplitudes est très facile et il existe une expression explicite de l'estimateur; ce n'est pas le cas pour les centres \mathbf{c} et les largeurs \mathbf{w} .

Aussi, pour appréhender le comportement du critère \mathcal{C} , nous nous plaçons dans le cas simple d'un signal spectroscopique contenant une unique raie ($K = 1$). Remarquons que ce cas est très semblable au cas d'une raie isolée des autres raies d'un même signal spectroscopique. Ce signal spectroscopique est représenté figure 2.4a, il est de taille $N = 200$, de RSB 10 dB et contient une unique raie de paramètres

$$a^* = 1, \quad c^* = 100, \quad w^* = 4.$$

Les figures 2.4b, 2.4c et 2.4d représentent respectivement $\mathcal{C}(a, c^*, w^*)$, $\mathcal{C}(a^*, c, w^*)$ et $\mathcal{C}(a^*, c^*, w)$, c'est-à-dire le critère représenté selon un paramètre lorsque les autres sont fixés à leur valeur réelle. Il apparaît que le minimum du critère correspond bien aux valeurs réelles a^* , c^* , w^* . Par ailleurs, il est quadratique en a comme prévu et relativement convexe en w (ce qui semble indiquer qu'une simple minimisation suffit pour trouver la solution). Concernant l'évolution du critère par rapport à c , on remarque qu'il est convexe au voisinage de c^* . Donc, en partant d'une initialisation dans cette zone, la convergence vers le minimum global est facile. Sur l'exemple illustré figure 2.4c, on observe qu'une initialisation dans l'intervalle $[c^* - 15, c^* + 15]$ conviendrait; cet intervalle correspond à environ $[c^* - 4w^*, c^* + 4w^*]$. En revanche, on peut identifier deux difficultés quant à la minimisation lorsque l'initialisation s'éloigne de la solution :

- la première et plus importante difficulté est la présence de zones relativement plates, sans gradient réel. En conséquence, une optimisation dans ces zones est difficile car il n'existe pas d'information pour guider l'estimateur vers la bonne solution. On notera également la présence de plusieurs minima locaux dus au bruit;
- l'autre difficulté est liée aux effets de bords : le critère présente de forts minima locaux aux extrémités du signal. Cela s'explique par le fait que positionner une raie aux extrémités du signal spectroscopique est moins coûteux que dans les zones plates puisque la raie n'existe qu'en partie. Cette difficulté peut toutefois être levée facilement, par exemple en contraignant les raies à être positionnées loin des bords ou bien en calculant l'erreur en dehors de la plage de mesure et en considérant que le signal spectroscopique y est nul. En pratique, cette difficulté reste extrêmement rare.

Les conclusions de cette section ne sont valides que dans le cas d'une raie seule ou isolée des autres raies d'un signal spectroscopique. Nous discutons dans la section suivante le cas de raies superposées.

Superposition des raies

Comme nous l'avons vu dans la section 2.1, l'estimation de l'amplitude a ou de la largeur w d'une raie isolée est simple si l'on connaît son centre. Mais cette affirmation n'est vraie que parce que la raie considérée n'est pas superposée à d'autres raies.

En guise de contre-exemple à cette apparente facilité de minimisation, nous avons considéré un signal spectroscopique peu bruité (RSB = 20 dB) contenant trois raies de paramètres

$$\mathbf{a}^* = \begin{bmatrix} 1 & 1 & 1 \end{bmatrix}, \quad \mathbf{c}^* = \begin{bmatrix} 95 & 100 & 105 \end{bmatrix}, \quad \mathbf{w}^* = \begin{bmatrix} 8 & 8 & 8 \end{bmatrix}.$$

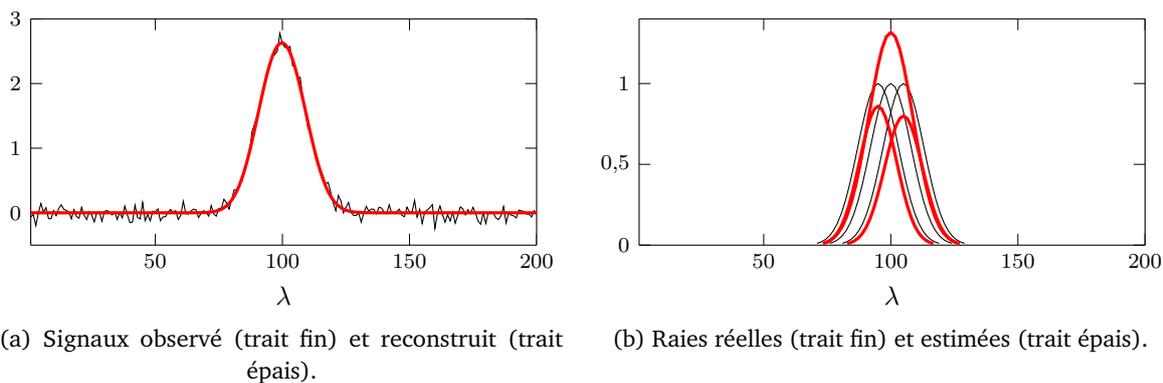


FIGURE 2.5 – Difficulté de minimisation du critère $\mathcal{C}(\mathbf{a}, \mathbf{c}^*, \mathbf{w})$ par rapport à \mathbf{a} et \mathbf{w} .

Le nombre de raies, leurs centres c et la variance du bruit sont connus. l'estimation ne porte donc que sur \mathbf{a} et \mathbf{w} qui sont initialisés aux valeurs de la vérité terrain \mathbf{a}^* et \mathbf{w}^* . Un échantillonneur de Gibbs [222] a permis de générer des échantillons suivant la vraisemblance, et les inconnues \mathbf{a} et \mathbf{w} sont estimées au sens de l'espérance a posteriori en moyennant les derniers échantillons générés. Les résultats obtenus sont représentés figure 2.5. L'estimation est très bonne puisque la valeur du critère avec ces valeurs (0,884) est meilleure que la valeur du critère avec les valeurs réelles des paramètres (0,932), ce qui indique que l'échantillonneur de Gibbs a convergé correctement. Pourtant, bien que le signal spectroscopique reconstruit soit très proche du signal spectroscopique observé, les raies estimées sont très différentes des véritables raies. Cet exemple montre que l'estimation des amplitudes et des largeurs peut être délicate en cas de superposition des raies.

Un cas limite de cet exemple est celui d'un signal spectroscopique de deux raies de centres et de largeurs identiques. Les raies étant gaussiennes, leur somme est également une gaussienne : on ne peut donc pas distinguer les deux raies. Heureusement, ce cas extrême n'a aucune raison d'exister en pratique.

Allen et McMeeking [6] abordent le problème de l'identifiabilité d'une ou plusieurs raies superposées en un ou plusieurs motifs. Ils montrent différents cas mais ne proposent pas de critère permettant de considérer que deux raies sont ou ne sont pas superposées. En optique, plusieurs critères existent pour considérer que deux sources ponctuelles sont résolues ou non. Par exemple, le critère de Rayleigh est vérifié si l'écart entre les deux sources est supérieur ou égal à la largeur à mi-hauteur de la réponse instrumentale. Le critère de Sparrow quant à lui est vérifié s'il existe un creux dans le signal entre les centres des deux raies [236, 238]. Ces critères se traduisent, dans le cas de deux gaussiennes de même largeur w situées en c_1 et c_2 et en l'absence de bruit (cf. figure 2.6) par :

$$\text{Critère de Rayleigh :} \quad |c_1 - c_2| \geq w \times 2\sqrt{2 \ln(2)}, \quad (2.3)$$

$$\text{Critère de Sparrow :} \quad |c_1 - c_2| \geq w \times 1,256\sqrt{2 \ln(2)}. \quad (2.4)$$

Des mesures de distances spectrales peuvent également être appliquées, comme par exemple les distances euclidiennes ou de Manhattan, Tchebychev ou Canberra [61]. McLean et ses co-auteurs [162] ont proposé un test de Neyman-Pearson qui donne la probabilité maximale de détection de deux raies superposées pour une probabilité de fausse alarme donnée. Il est donc illusoire de tenter de séparer deux raies lorsqu'elles sont trop proches.

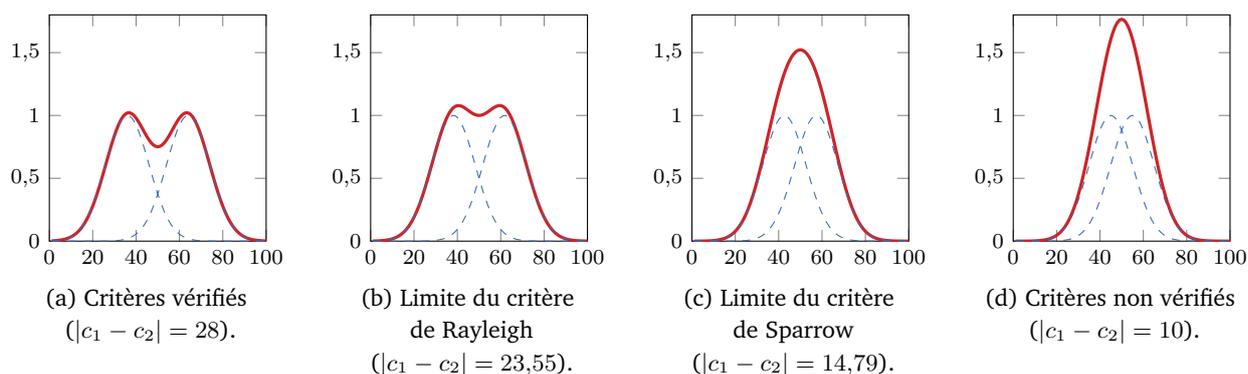


FIGURE 2.6 – Illustration des critères de Rayleigh et Sparrow avec deux raies gaussiennes de largeur $w = 10$.

Nombre inconnu de motifs

Très souvent, le nombre de raies n'est pas connu lors de la décomposition d'un signal spectroscopique, et plus généralement lors de la décomposition en motifs élémentaires. Le nombre d'inconnues est donc lui-même inconnu, ce qui augmente la difficulté du problème : l'espace des solutions est en effet plus grand, et il est délicat de comparer deux estimations si elles ne sont pas de même dimension. En 1978, Allen et McMeeking [6] notaient déjà à juste titre : « *Clearly the number of component bands used for analysis of a curve envelope can be increased to any arbitrary number; conversely, satisfactory fits may be obtained with fewer curves than are present in reality* ». Cette idée correspond au principe de simplicité énoncé par Guillaume d'Ockham.

Parmi les approches pour estimer le nombre de raies, on peut citer par exemple la représentation d'espace échelle de Goshtasby et O'Neill [92], l'utilisation des critères AIC (*Akaike information criterion*) [4], BIC (*Bayesian information criterion*) [230], GIC (*generalized information criteria*) [125] ou MDL (*Minimum description length*) [220]. Certains algorithmes sont également capables de considérer le nombre de motifs comme une variable supplémentaire, comme par exemple les algorithmes RJMCMC [94], *multiple birth and death* [64] ou *multiple birth and cut* [85]. Nous proposons d'ailleurs une méthode utilisant l'algorithme RJMCMC dans la section 3.3.

2.2 État de l'art

Vue d'ensemble

Les premières méthodes de décomposition d'un signal spectroscopique datent du début des années 1980. La décomposition y est considérée comme un problème de moindres carrés non linéaires :

$$\hat{\mathbf{a}}, \hat{\mathbf{c}}, \hat{\mathbf{w}} = \arg \max_{\mathbf{a}, \mathbf{c}, \mathbf{w}} \left\| \mathbf{y} - \sum_{k=1}^K a_k \varphi(c_k, w_k) \right\|^2. \quad (2.5)$$

L'estimation des paramètres est effectuée à l'aide de méthodes d'optimisation locale (algorithme du gradient, de recherche directe, de Newton-Raphson, du simplexe, des moindres carrés généralisés...), voir par exemple [134, 193] et l'état de l'art donné dans [6]. Ces méthodes, qui sont parfois empiriques, fonctionnent correctement dans des cas simples (peu de bruit, raies non superposées, nombre de raies connu...) où lorsque l'initialisation est déjà très proche de la solution. En revanche,

dès que ces conditions ne sont plus valides, elles sont incapables de donner une bonne estimation et restent coincées dans un minimum local ; ce problème était déjà soulevé en 1978 [6].

La méthode de Caruana *et al.* [37] repose sur le fait que le logarithme d'une gaussienne est une parabole :

$$\log \varphi_n(c, w) = -\frac{1}{2w^2} \lambda_n^2 + \frac{c}{w^2} \lambda_n - \frac{c^2}{2w^2}$$

Ainsi, le passage au logarithme sur le support de chaque raie permet de linéariser le problème et d'estimer les coefficients de la parabole (et donc les paramètres de la raie) par moindres carrés. Cette méthode est toutefois délicate à mettre en œuvre car elle nécessite non seulement de connaître le nombre de raies mais également leur support pour que la transformation permettent d'obtenir une parabole. Elle est également très sensible au bruit [98].

Plus tard, Goshtasby et O'Neill [92] effectuent une représentation d'espace-échelle des données pour déterminer le nombre de raies et une initialisation de leurs paramètres. Cette représentation est obtenue en appliquant un lissage plus ou moins fort sur les données, puis en déterminant les passages à zéro de la dérivée seconde du signal obtenu. L'algorithme de Marquardt peut ensuite être utilisé pour affiner les paramètres obtenus. Une méthode très proche a été proposée par Makarewicz et coll. [66, 141].

D'autres approches ont été ponctuellement proposées ; citons par exemple l'utilisation d'ondelettes couplée à une approche multi-échelle [8], ou la considération que les observations correspondent à une réalisation d'un mélange de gaussiennes [78].

Mais la grande majorité des méthodes de décomposition en motifs sont fondées sur l'équation (2.1) et modélisent les observations \mathbf{y} comme un signal impulsionnel \mathbf{x} dont chaque impulsion est convoluée par un motif. Dans le cas où le motif est identique pour chaque impulsion, la décomposition en motifs est équivalente à une déconvolution impulsionnelle. Le problème étant mal posé, on cherchera donc à bénéficier de la connaissance que le signal d'intérêt \mathbf{x} est impulsionnel. Ainsi, les modèles statistiques et les représentations parcimonieuses ont été largement utilisés.

Modèles statistiques

Plusieurs modèles statistiques ont été proposés, avec des techniques d'optimisation diverses : méthode heuristique [128], algorithme de Viterbi [103], algorithme EM [17, 167, 168], méthode *split Bregman* [140], décomposition du signal en sous-bandes [66], ou méthode bayésienne variationnelle [14].

Un modèle statistique très répandu est le processus Bernoulli-gaussien qui permet de modéliser facilement le caractère impulsionnel du signal \mathbf{x} . Kormylo et Mendel [126] ont été les premiers à proposer ce modèle, qui a depuis été très utilisé [14, 28, 44, 93, 100, 190, 148, 224]. Dans un processus Bernoulli-gaussien, chaque échantillon $x_n \in \mathbb{R}^N$ est représenté par le couple (q_n, a_n) . La variable q_n représente l'occurrence d'une raie à l'échantillon n et est modélisé par une loi de Bernoulli :

$$\forall n, \quad p(q_n) = \begin{cases} 1 - \lambda & \text{si } q_n = 0, \\ \lambda & \text{si } q_n = 1 \end{cases} \quad (2.6)$$

où $\lambda \in [0, 1]$ est un hyperparamètre contrôlant la parcimonie du signal. La variable a_n représente l'amplitude de l'échantillon n et est modélisé par une distribution gaussienne :

$$\forall n, \quad a_n \sim \begin{cases} \delta_0 & \text{si } q_n = 0, \\ \mathcal{N}(0, \sigma^2) & \text{si } q_n = 1 \end{cases} \quad (2.7)$$

où δ_0 est la distribution de Dirac et σ est un second hyperparamètre contrôlant l'écart-type des amplitudes. La distribution gaussienne peut être restreindre à \mathbb{R}^+ si l'on cherche seulement des raies d'amplitude positive. Une autre définition consiste à modéliser a_n par une gaussienne, quelle que soit la valeur de q_n [126], mais alors cela peut poser le problème de l'estimation de a_n lorsqu'il n'y a pas de raies en n : la valeur de a_n n'apparaissant pas dans la vraisemblance (puisqu'elle est annulée par $q_n = 0$), son estimation est donc indépendante des données. Kormylo et Mendel [126] ont proposé l'algorithme optimisation SMLR (*single most likely replacement*) qui a été réutilisé plusieurs fois depuis [109, 115, 237]. Les algorithmes EM [93] et SEM [190] sont également utilisés.

Le cadre bayésien est capable de considérer le processus Bernoulli-gaussien. Plusieurs travaux ont alors utilisé des algorithmes MCMC : l'échantillonneur de Gibbs ou sa variante PCGS [28, 44, 116, 216, 224, 148] et l'algorithme *Exchange Monte Carlo*, connu également sous le nom de *Parallel Tempering* [180, 247]. L'intérêt d'un a priori Bernoulli-gaussien permet de rester dans un espace de recherche à dimension constante, même si le nombre de raies est inconnu.

D'autres travaux, tout en utilisant le cadre bayésien, considèrent directement l'équation (2.1) avec un nombre de motifs K inconnu. En traitement d'images, les modèles par processus ponctuels marqués [65] utilisent cette formulation. La méthode MCMC par excellence est l'algorithme RJMCMC de Green [94] et est utilisé dans de nombreux travaux [78, 97, 226]. Le nombre de raies est estimé au sens du maximum a posteriori marginal, mais l'estimateur du signal impulsionnel est malheureusement rarement explicité dans les publications alors que c'est une étape indispensable et qui, dans le cas de la décomposition en motifs élémentaires, fait apparaître des problèmes de permutation d'indices (*label switching*) rendant l'estimation du signal très délicate. Sans utiliser de méthodes MCMC, les auteurs se tournent alors vers des algorithmes à dimension constante qui testent toutes les valeurs possibles de K [171].

Représentations parcimonieuses

Les méthodes d'approximation parcimonieuse reposent sur un terme de régularisation, exprimé sous la forme d'une norme favorisant la parcimonie.

La pseudo-norme ℓ_0 est la plus appropriée pour cela [31, 237], mais elle n'est ni convexe, ni continue, et aboutit à un problème NP-difficile [181]. L'optimisation est alors effectuée à l'aide de méthodes combinatoires, comme par exemple l'optimisation entière (*mixed-integer programming*) [31], SBR [237] ou FNNOMP [264].

Beaucoup d'auteurs se tournent alors vers l'alternative convexe de la pseudo-norme ℓ_0 , à savoir la norme ℓ_1 [30, 40, 45, 119, 134, 140]. La régularisation issue de cette norme est d'ailleurs parfois adaptée pour la rendre différentiable [119], l'utiliser sur un espace particulier [40] ou en y ajoutant un terme d'entropie [45].

Notons enfin les travaux de Repetti et coll. [218] qui utilisent une norme mixte ℓ_1/ℓ_2 et un algorithme proximal.

La figure 2.7 présente le comportement de ces normes dans un espace bidimensionnel ; l'énergie d'un a priori Bernoulli-gaussien est également représentée pour comparaison (on s'aperçoit qu'elle est très proche de la pseudo-norme ℓ_0). On y voit que les valeurs nulles des coordonnées sont privilégiées (ce sont elles qui donnent des valeurs faibles aux normes), favorisant ainsi la parcimonie.

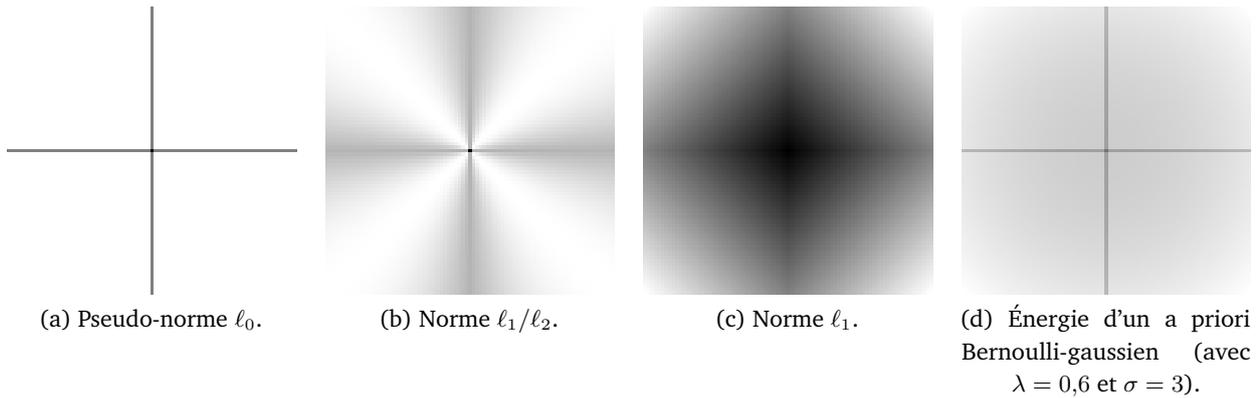


FIGURE 2.7 – Représentation de différentes normes en deux dimensions. Le niveau de gris est proportionnel à la valeur de la norme.

2.3 Méthodes développées

Durant ma thèse [148], j'ai proposé deux méthodes de décomposition d'un signal spectroscopique. Par la suite, j'ai étudié l'intérêt des algorithmes RJMCMC et HMC (*Hamiltonian Monte Carlo*, parfois appelé *hybrid Monte Carlo*) sur ce même problème.

La première méthode développée durant ma thèse supposait les raies de même forme : le problème se ramenait à un problème de déconvolution impulsionnelle myope. Le signal impulsionnel était modélisé par une loi a priori Bernoulli-gaussienne et estimé au sens de l'espérance a posteriori à partir de réalisations générées par un échantillonneur de Gibbs. L'originalité principale du travail concernait l'estimation : contrairement aux travaux similaires de l'époque [44, 224], l'estimation fournie utilisait le couple occurrences–amplitudes (q, a) , et pas seulement le vecteur des amplitudes. Cela permettait une analyse plus fine du résultat et une meilleure détection des raies d'amplitude faible.

La seconde méthode tenait compte d'une variabilité dans les formes des raies et permettait également de considérer les raies en dehors de la grille d'échantillonnage. Une modélisation Bernoulli-gaussienne du vecteur d'intérêt était également utilisée, avec l'originalité que le nombre de sites dans ce vecteur était fixé à $K_{\max} > K$: cela permettait de conserver un espace de dimension constante, et donc d'éviter le recours (fastidieux) à des algorithmes trans-dimensionnels, comme le RJMCMC. Le problème concernait alors l'estimation puisqu'il fallait gérer le problème de permutation d'indices (*label switching*), ce type de problème est assez difficile à résoudre correctement et reste encore aujourd'hui un problème ouvert. Plusieurs auteurs ont déjà proposé des solutions [41, 42, 111, 144, 161, 219, 239, 240], parfois en considérant ce problème comme un problème inverse, ce qui a pour conséquence d'ajouter une couche algorithmique à une méthode de simulation qui peut déjà être lourde.

Par la suite, nous avons utilisé, avec Sylvain Faisan, l'algorithme RJMCMC pour décomposer une séquence de plusieurs signaux (cf. chapitre 3) [153]. Il s'est avéré que dans ce cadre, l'algorithme RJMCMC a été difficile à implémenter et peinait à explorer convenablement l'espace des solutions, de sorte qu'il a fallu parfois le « forcer » à choisir une direction d'exploration (un peu comme les travaux de [208]), rendant ainsi caduque le caractère automatisé de cet algorithme, qui en faisant a priori la force.

C'est la raison pour laquelle nous avons exploré l'utilisation de l'algorithme HMC lors du stage de M2 de Argheesh Bhanot. Cet algorithme, proposé en particulier pour résoudre la difficulté qu'a un

algorithme à marche aléatoire d'explorer un espace de grande dimension, consiste à proposer des candidats en faisant appel à la dynamique hamiltonienne [182]. Or, comme on l'a vu section 2.1, le critère présente de nombreuses zones plates, où l'algorithme HMC n'est pas capable de choisir une bonne direction de recherche. Cette conclusion est en accord avec les études de la littérature : l'algorithme HMC échoue si la distribution est à queue lourde [203]. Il s'avère d'ailleurs qu'il n'y a pas vraiment d'algorithme pour travailler en grande dimension avec des queues lourdes ; on peut penser qu'un algorithme comme l'algorithme BPS (*Bouncy Particle Sampler*) [24] pourrait fonctionner².

2.4 Conclusion

Ce chapitre fait le point sur les difficultés et l'état de l'art de la décomposition d'un seul signal unidimensionnel. Le problème de base étant à présent établi, les chapitres suivants exposent les développements effectués pour des données plus complexes.

Un échantillonnage irrégulier n'a pas de grande influence à condition bien sûr qu'il permettent d'échantillonner convenablement la forme des raies. La première difficulté du problème est bien entendu la présence de bruit dans les données. Nous avons vu également que dans le cas d'une raie isolée, estimer son amplitude et sa largeur n'est pas difficile. En revanche, l'estimation de son centre est facile seulement si l'initialisation est suffisamment proche de la solution, autrement la présence de zones sans réel gradient dans le critère rend l'estimation du centre difficile. De plus, l'estimation de tous les paramètres devient difficile dans le cas de raies superposées. À cela s'ajoute le fait que le nombre de raies n'est pas toujours connu.

De nombreuses méthodes de décomposition ont été proposées depuis quarante ans. Il s'en dégagent deux grandes familles : les méthodes statistiques (notamment bayésiennes) et les méthodes d'approximation parcimonieuse. La raison est que ces deux cadres de travail sont capables d'introduire facilement et naturellement le fait que le signal sous-jacent à la décomposition est impulsionnel. Cela dit, le problème reste encore ouvert car il n'existe toujours pas de méthode qui réponde convenablement aux difficultés du problème dans un temps de calcul acceptable.

2. Source : communication personnelle avec M. Pereyra.

Chapitre 3

Décomposition conjointe d'une séquence de signaux spectroscopiques

Une extension du problème traité dans le chapitre précédent concerne la décomposition de *plusieurs* signaux spectroscopiques : nous nous intéressons désormais à la décomposition de S signaux $\mathbf{y}_s \in \mathbb{R}^N$ où $s \in \mathbb{S}$. L'ensemble \mathbb{S} représente l'espace dans lequel sont disposés les signaux \mathbf{y}_s :

- si les signaux sont acquis selon un échantillonnage temporel régulier, on a $\mathbb{S} = \{1, \dots, S\}$. Un exemple d'application correspondant à ce cas est l'acquisition à différents instants d'un spectre électromagnétique ;
- si les signaux sont acquis dans l'espace, alors $\mathbb{S} \subset \mathbb{N}^2$. Une image hyperspectrale correspond à ce type de données : chaque pixel de l'image est un spectre \mathbf{y}_s où s représente l'indice du pixel correspondant.

Chaque signal \mathbf{y}_s est modélisé de la même manière que dans le chapitre précédent (équation (2.1)) et regroupe K_s raies d'amplitudes a_{ks} , de centre c_{ks} et de paramètres de forme w_{ks} :

$$\forall s \in \mathbb{S}, \quad \mathbf{y}_s = \sum_{k=1}^{K_s} a_{ks} \varphi(c_{ks}, w_{ks}) + \mathbf{b}_s. \quad (3.1)$$

Le nombre de raies K_s peut être considéré constant et égal à une valeur K pour tout s . Dans ce cas, il est possible d'autoriser les amplitudes a_{ks} à être nulles, entraînant alors un nombre variable de raies par signal. Autrement, le nombre de raies K_s est variable et doit alors être estimé. Par ailleurs, nous supposons que les paramètres de même indice k présentent une évolution douce à travers l'espace \mathbb{S} . Autrement dit, les paramètres de deux signaux acquis à des instants proches ou correspondant à deux pixels voisins ont des valeurs proches. Cette hypothèse est justifiée d'un point de vue des applications considérées (cf. section 1.1). La figure 3.1 donne un exemple de séquence de plusieurs signaux spectroscopiques : on y voit clairement l'évolution douce des paramètres que l'on cherchera à reproduire. Cette hypothèse apporte également une connaissance a priori très forte qui permettra d'améliorer le résultat, comme il est montré section 3.2. La difficulté réside alors dans la modélisation de cette hypothèse et l'optimisation du problème résultant puisque les liens entre paramètres structurent fortement l'espace des solutions et accroissent la difficulté de son exploration.

À ma connaissance, ce problème n'a jamais été traité tel quel dans la littérature même s'il existe des similarités avec d'autres problèmes (cf. section 3.1). Une première approche pour résoudre le problème de la décomposition d'une séquence de signaux spectroscopiques est d'oublier l'aspect

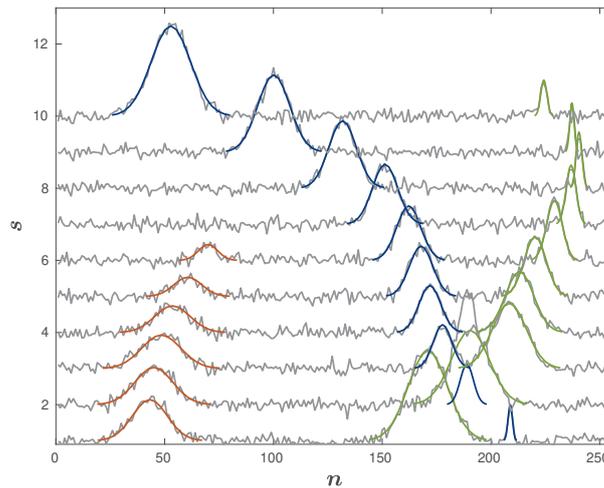


FIGURE 3.1 – Exemple d’une séquence unidimensionnelle de dix signaux ($\mathbb{S} = \{1, \dots, 10\}$). Chaque couleur regroupe les raies qui suivent la même évolution douce. Remarquez que le nombre de raies par spectre n’est pas constant.

conjoint. Ainsi, j’ai proposé en 2009 [151] une approche en deux étapes : les signaux sont dans un premier temps décomposés indépendamment les uns des autres, puis un post-traitement associe les raies obtenues en différents sites s en fonction de la similarité de leurs paramètres. Ce fonctionnement séquentiel, qui ne tient pas compte de l’hypothèse d’évolution douce des paramètres, est peu efficace pour deux raisons. D’une part il n’est pas possible d’améliorer une décomposition déjà obtenue lors de la première étape, d’autre part les décompositions étant indépendantes les unes des autres, on ne profite pas de l’a priori d’évolution douce. Nous montrons en section 3.2 que cela génère des erreurs d’estimation et qu’une décomposition conjointe est préférable. En 2011 [149] nous avons donc proposé une première méthode de décomposition conjointe afin d’éviter les écueils précédents. Elle a été établie dans un cadre bayésien et l’optimisation a été effectuée avec des algorithmes MCMC ; la méthode est détaillée section 3.3. Nous montrons que cette décomposition conjointe est meilleure que la décomposition séquentielle. En 2015 [153], nous avons introduit dans la méthode précédente la possibilité d’avoir un nombre de raies variables entre les spectres. Ce besoin était surtout nécessaire en imagerie hyperspectrale astronomique où les signaux n’ont pas tous le même nombre de raies. Il s’avère que cette méthode est lente et difficile à utiliser. Aussi, après une première étude en 2013 [160], nous avons proposé en 2019 [177] une alternative plus rapide et plus simple en s’inspirant des méthodes d’approximation parcimonieuse (section 3.4). Au final, ce travail a fait l’objet de quatorze publications [10, 138, 149, 151, 153, 154, 156, 157, 158, 160, 174, 175, 176, 177].

Ce travail a permis d’apporter les contributions suivantes.

- Le problème de la décomposition en motifs d’un signal où les motifs présentent une lente évolution a été défini.
- Un modèle de décomposition conjointe associé à ce problème a été proposé. Ce modèle a été formulé dans le cadre bayésien puis sous forme d’approximation parcimonieuse en définissant un terme additionnel représentant une variation totale.
- Un algorithme RJMCMC capable d’évoluer dans un espace à deux dimensions a été utilisé dans le cadre bayésien.
- Un algorithme inspiré des méthodes d’approximation parcimonieuse a été proposé, et il permet d’avoir une structuration des vecteurs inconnus plus souple que les méthodes d’approximation jointe, simultanée ou structurée.

3.1 État de l'art

Bien que la décomposition conjointe d'une séquence de signaux spectroscopiques n'ait pas, à notre connaissance, été traitée dans la littérature, nous pouvons lister des problèmes similaires qui nous ont parfois inspirés. Certains de ces travaux sont dans la discipline du traitement du signal multi-capteur (*multisensor signal processing*).

La séparation de sources [48] ou le démélange spectral [121] sont des problèmes proches du nôtre, dans le sens où les données regroupent plusieurs signaux ordonnés dans un espace particulier et modélisés comme une combinaison (souvent linéaire) de motifs élémentaires. Il existe en revanche deux différences principales avec notre problème. En premier lieu, les motifs élémentaires sont le plus souvent des signaux qui ne sont pas paramétrisables ou qui sont à support très large. Mais la différence majeure provient du fait que ces motifs ne présentent pas d'évolution particulière à travers les données (hormis pour l'amplitude des motifs). Or, si cette hypothèse existe, elle doit être prise en compte pour un traitement efficace. Gobinet et coll. [91] ont proposé plusieurs prétraitements afin que les méthodes traditionnelles de séparation de sources puissent être utilisées pour la décomposition d'une séquence de signaux spectroscopiques. En particulier, ils ont proposé une méthode d'alignement des raies et d'homogénéisation de leur largeur. Or, cette méthode nécessite une connaissance a priori sur la localisation des raies. Il existe des alternatives [248] qui ne nécessitent pas d'a priori mais elles sont supervisées et n'ont pas donné de résultats satisfaisants.

Le problème peut également être vu comme un problème d'approximation parcimonieuse dont le dictionnaire est construit à partir de motifs paramétriques [30, 227, 250, 263]. Plus précisément, l'approximation parcimonieuse jointe décompose plusieurs signaux en utilisant le même dictionnaire. Les méthodes de parcimonie simultanée résolvent le problème dit MMV (*multiple measurement vector*) [251] dans lequel les signaux partagent le même support. Cette hypothèse est très restrictive et ne convient pas au cas que l'on considère, dans lequel les raies n'ont pas les mêmes centres. D'autres approches, comme l'approximation parcimonieuse récursive [255] détermine le vecteur parcimonieux d'un signal à partir des vecteur des signaux précédents : la décomposition se fait de proche en proche, mais il n'est pas possible de revenir sur une décomposition antérieure. Les techniques de parcimonie sociale ou structurée [11, 127, 233] supposent une structure particulière des différents supports. Mais cette structure doit être définie a priori, ce qui n'est pas adapté à notre cas. Sans imposer de connaissance particulière sur le support de la solution, la modélisation de la réponse impulsionnelle à travers les signaux peut aider à considérer une évolution douce : c'est le cas dans [30] où la réponse instrumentale spatiale de l'instrument est prise en compte dans la décomposition des spectres d'une image hyperspectrale astronomique. Toutefois, dans Belmerhnia et coll. [20], le terme favorisant la parcimonie (une norme $\ell_{1,1}$) est appliqué sur les différences finies de la matrice des amplitudes, se rapprochant ainsi des solutions que nous présentons ci-après.

La séparation de sources retardée (ou anéchoïque) est un cas particulier de la séparation de source convolutive dans laquelle les filtres sont des impulsions de Dirac. Il est ainsi possible de tenir compte des retards que peuvent subir les sources [101, 169, 189, 194, 260, 265]. À notre connaissance, les approches développées ne considèrent pas de sources paramétriques ; elles ne sont donc pas capables de modéliser l'évolution douce des paramètres de forme w . En outre, la grande majorité des méthodes fait l'hypothèse que les sources sont indépendantes, décorréelées ou différentes dans un domaine particulier. Ces méthodes utilisent ainsi des stratégies de linéarisation [43, 112], d'analyse en fréquence [178, 198] ou en temps-fréquence [189, 194]. Des propositions ont été faites pour relaxer les hypothèses d'indépendance et de décorrélation, comme dans [215]. Des algorithmes de *dynamic time warping* ont été utilisés pour contrebalancer le retard des sources [101, 106], après les avoir détectées dans un premier temps. Dans [178], la méthode NMF (*non-negative matrix*

factorization) a été étendue pour considérer les retards. Cependant, nos tests montrent que cette approche fonctionne mal en cas de sources très corrélées [177].

La déconvolution multi-canaux est l'extension à plusieurs signaux du problème de sismique-réflexion présenté dans le chapitre 2. Là encore, la modélisation par processus Bernoulli-gaussien est utilisée [103, 109, 216]. Mais dans ces travaux, seuls les centres et les amplitudes des motifs sont estimés : les motifs sont tous de même forme et les amplitudes sont restreintes à être sur la grille.

Plus proche de l'une de nos applications, Villeneuve et coll. [258] ont proposé une approche bayésienne pour effectuer la déconvolution de cartes de vitesses de galaxies. Leur méthode est toutefois limitée au cas où il n'y a qu'une seule raie ($K = 1$), et l'évolution douce de cette raie n'est pas prise en compte. Bouché et coll. [25] utilisent un modèle d'ellipse tridimensionnelle sur la galaxie : les auteurs ne recherchent donc qu'une seule raie avec un modèle de cinématique très contraint. À la différence de ces travaux, nous considérons plusieurs raies en leur imposant une évolution douce : l'évolution autorisée est donc plus souple que dans [25] mais n'est pas totalement libre comme dans [258].

En temps-fréquence, l'estimation des fréquences de signaux modulés en fréquence [266, 267] s'apparente à la décomposition en motifs d'une séquence de signaux. Cependant, les fréquences instantanées des signaux sont contraintes à évoluer de manière assez définies (sinusoïdes ou constantes par morceaux).

Le suivi de cible (*target tracking*) consiste à estimer l'état d'un objet (en particulier sa position) à partir de mesures bruitées [16, 56, 108, 259]. Dans la plupart des travaux, il existe deux différences principales avec notre problème. La première différence est que les données que nous considérons ne sont pas des mesures bruitées des positions des raies, mais des mesures construites à partir des positions. Pour appliquer à notre contexte des méthodes de suivi de cible, il nous faudrait donc détecter les raies dans les spectres avant de pouvoir les suivre. La seconde différence est que le suivi de cible fonctionne souvent dans un cadre récursif, c'est-à-dire que les données doivent être traitées au fur et à mesure, alors que nous avons à disposition toutes les données, autorisant ainsi d'effectuer une décomposition conjointe des spectres. Par ailleurs, les méthodes actuelles de suivi de cibles ne sont pas toujours aptes à gérer l'interaction entre plusieurs cibles.

D'autres techniques, plus rares, ont été proposées dans différents domaines où l'objectif est également de suivre des objets. Ainsi, une approche issue du traitement d'images a été proposée pour des signaux de sismique-réflexion [225]. On peut voir également des similarités avec le problème de l'extraction de vaisseaux sanguins ou de fibres nerveuses en imagerie médicale [122, 200]. Mais dans nos applications, il n'est pas permis aux raies de se diviser ou de fusionner, contrairement aux structures anatomiques.

3.2 Intérêt d'une décomposition conjointe

Cette section a pour objectif de montrer l'intérêt d'une décomposition conjointe par rapport à une décomposition séquentielle ; elle reprend les modèles présentés dans [149]. Dans la continuité du chapitre précédent, le problème est traité dans le cadre bayésien puisqu'il donne accès à des méthodes efficaces qui pourront en outre modéliser facilement la dépendance entre signaux. Deux modèles sont donc étudiés : le premier pour une décomposition séquentielle (les signaux sont décomposés séparément les uns des autres), le second pour une décomposition conjointe. Dans ces deux modèles, les données sont modélisées selon l'équation (3.1). Dans les deux cas, une méthode MCMC couplée à un schéma de recuit simulé permet ensuite de déterminer l'estimateur MAP.

Modèle bayésien pour la décomposition séquentielle

Les centres peuvent se situer partout dans l'intervalle $[1, N]$:

$$c_{ks} \sim \mathbb{I}_{[1, N]}(c_{ks}). \quad (3.2)$$

Les amplitudes sont positives et on privilégie les valeurs faibles :

$$a_{ks} \sim \mathcal{N}^+(0, \sigma_a^2) \quad (3.3)$$

où la notation $\mathcal{N}^+(0, \sigma_a^2)$ représente une loi normale de moyenne nulle et variance σ_a^2 tronquée sur \mathbb{R}^+ :

$$p(a_{ks}) = \frac{2}{\sqrt{2\pi\sigma_a^2}} \exp\left(-\frac{a_{s,k}^2}{2\sigma_a^2}\right) \mathbb{I}_{\mathbb{R}^+}(a_{ks}). \quad (3.4)$$

Enfin, les paramètres de forme sont positifs :

$$w_{ks} \sim \mathbb{I}_{\mathbb{R}^+}(w_{ks}). \quad (3.5)$$

Le bruit est supposé blanc, gaussien et de variance r_b :

$$b_{s,n} \sim \mathcal{N}(0, r_b). \quad (3.6)$$

Nous choisissons une loi inverse gamma comme distribution de la variance du bruit :

$$r_b \sim \mathcal{IG}(\alpha_b, \beta_b) \quad (3.7)$$

dont les paramètres sont très petits ($\alpha_b = \beta_b \ll 1$) de sorte à s'approcher de l'a priori classique pour les variances : la loi de Jeffreys [118, 221].

Modèle bayésien pour la décomposition conjointe

Dans le premier modèle de décomposition conjointe que nous avons proposé [149] un quatrième paramètre était attribué à chaque raie k du signal s : il s'agissait de la classe $z_{ks} \in \mathbb{N}$ à laquelle appartenait la raie. Ainsi, si deux raies k et l issues de deux signaux s et t voisins sont de même classe ($z_{ks} = z_{lt}$) alors leurs centres, amplitudes et largeurs ont respectivement des valeurs très proches ($c_{ks} \approx c_{lt}$, $a_{ks} \approx a_{lt}$, $w_{ks} \approx w_{lt}$).

Par commodité, nous noterons dans la suite \mathbf{c}^l les paramètres $\{c_{ks}\}$ tels que $z_{ks} = l$ pour tout s . Puisque les centres des raies de même classe l évoluent lentement dans \mathbb{Z} , un a priori de douceur est choisi pour régulariser la solution, et nous imposons toujours aux centres d'être dans l'intervalle d'observation $[1, N]$. La loi a priori s'exprime donc [86] :

$$p(\mathbf{c}^l | \mathbf{z}) \propto \exp\left(-\frac{1}{2r_c} \|\mathbf{D}\mathbf{c}^l\|^2\right) \mathbb{I}_{[1, N]}(\mathbf{c}^l) \quad (3.8)$$

où r_c est un hyperparamètre contrôlant la force de l'a priori (il est fixé) et \mathbf{D} est l'opérateur des différences finies (dérivées discrètes). Par exemple, une dérivée à l'ordre 2 favorise une évolution linéaire :

$$(\mathbf{D}\mathbf{c}^l)_s = c_{s^-}^l - 2c_s^l + c_{s^+}^l \quad (3.9)$$

où s^- , s et s^+ représentent trois sites alignés et consécutifs dans \mathbb{S} . Par ailleurs, à s fixé les centres c_{ks} sont supposés indépendants. Ainsi, la loi a priori sur \mathbf{c} s'écrit :

$$p(\mathbf{c} | \mathbf{z}) \propto \exp\left(-\frac{1}{2r_c} \sum_{l=1}^K \|\mathbf{D}\mathbf{c}^l\|^2\right) \mathbb{I}_{[1, N]}(\mathbf{c}). \quad (3.10)$$

Les amplitudes et les largeurs sont également supposées suivre une évolution douce, et donc de la même manière que pour les centres, leurs lois a priori s'écrivent :

$$p(\mathbf{a}|\mathbf{z}) \propto \exp\left(-\frac{1}{2r_a} \sum_{l=1}^K \|\mathbf{D}\mathbf{a}^l\|^2\right) \mathbb{I}_{\mathbb{R}^+}(\mathbf{a}), \quad (3.11)$$

$$p(\mathbf{w}|\mathbf{z}) \propto \exp\left(-\frac{1}{2r_w} \sum_{l=1}^K \|\mathbf{D}\mathbf{w}^l\|^2\right) \mathbb{I}_{\mathbb{R}^+}(\mathbf{w}). \quad (3.12)$$

Le quatrième paramètre à modéliser est la classe des raies, pour laquelle il n'y a, a priori, aucune information disponible. Les classes sont donc distribuées suivant une loi uniforme :

$$\mathbf{z}_s \sim \mathcal{U}_{\mathcal{S}_K} \quad (3.13)$$

où \mathcal{S}_K est l'ensemble des permutations sur $\{1, \dots, K\}$.

Pour terminer, les distributions du bruit et de son hyperparamètre sont identiques au modèle pour la décomposition séquentielle.

Optimisation des deux modèles avec l'échantillonneur de Gibbs

Les lois a posteriori des deux modèles sont donnés dans [149]. Leurs expressions étant complexes, nous choisissons de les simuler avec l'échantillonneur de Gibbs (elles sont par ailleurs intégrables puisque les lois a priori dont elles sont issues sont elles-mêmes intégrables). Plus précisément, les centres, largeurs et classes sont simulés avec un algorithme de Metropolis-Hastings et les amplitudes et le bruit peuvent être simulés directement. Il est connu que l'échantillonneur de Gibbs peut rester coincer dans un minimum local [222]. C'est pourquoi il est intégré dans un schéma de recuit simulé afin de trouver l'estimateur du MAP (une description du recuit simulé est donnée section 4.3).

Comparaison des deux modèles sur données synthétiques

Les deux modèles sont testés sur les données synthétiques présentées figure 3.2 ; c'est une séquence unidimensionnelle ($\mathbb{S} = \{1, \dots, S\}$) de $S = 10$ spectres de $N = 256$ échantillons contenant chacun $K = 3$ raies gaussiennes. Le rapport signal-à-bruit est d'approximativement 15 dB. Les températures initiale et finale du recuit simulé sont égales à $T_i = 50$ et $T_f = 1$. Les hyperparamètres de douceur sont fixés à

$$r_c \approx 39, \quad r_a \approx 4 \cdot 10^{-3}, \quad r_w \approx 0,4.$$

Ces valeurs correspondent à une fraction de la dynamique des paramètres, la justification est détaillée dans [149]. Enfin, le nombre d'itérations des deux méthodes est fixé de sorte à avoir le même temps de calcul pour les deux modèles (environ 97 s), soit 5700 itérations pour la décomposition séquentielle et 5000 itérations pour la décomposition conjointe.

Les estimations au sens du MAP sont représentées figure 3.3. L'estimation obtenue avec la méthode conjointe est clairement meilleure que celle obtenue avec la méthode séquentielle : en effet, les paramètres sont bien mieux estimés et en plus la méthode conjointe est capable de classer correctement les raies, ce qui permet de suivre leur évolution à travers la séquence.

Il serait malgré tout possible d'obtenir une classification des raies à la suite d'une décomposition séquentielle à l'aide d'un post-traitement, comme nous l'avons proposé dans [151], mais la mauvaise estimation initiale conduirait forcément à une classification erronée.

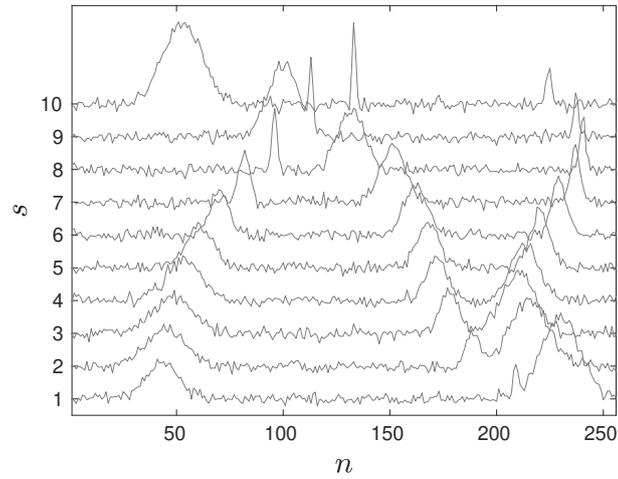


FIGURE 3.2 – Séquence simulée de $S = 10$ spectres de $N = 256$ échantillons et de $K = 3$ raies.

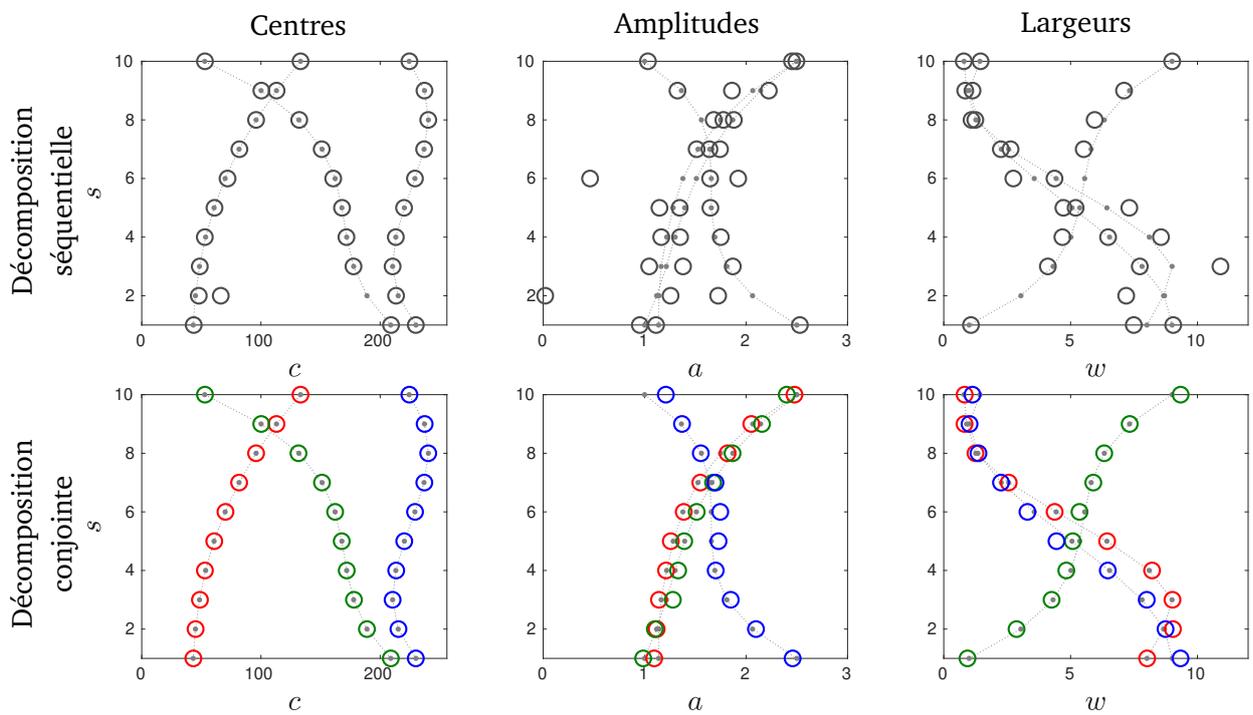


FIGURE 3.3 – Paramètres estimés (cercles) sur la séquence de la figure 3.2 avec les méthodes de décomposition séquentielle et conjointe. Les lignes en pointillés et les points représentent la vérité terrain. Les couleurs des cercles correspondent aux classes des raies estimées.

Pour affiner cette conclusion, nous avons mesuré l'EQM des estimations obtenues avec les deux modèles sur 20 séquences contenant $K = 4$ raies. La décomposition séquentielle obtient une EQM de $7.20 \cdot 10^{-2}$ alors que la décomposition conjointe obtient une EQM de seulement $1.76 \cdot 10^{-2}$. Ce résultat peut s'expliquer par le fait que la décomposition séquentielle peut manquer certaines raies. Au contraire, l'a priori de douceur intégré dans la décomposition conjointe favorise la recherche de raies dans les zones de l'espace des solutions en contenant probablement. En d'autres termes, la décomposition conjointe de tous les spectres permet d'utiliser l'estimation obtenue en un spectre pour améliorer l'estimation en un spectre contigu.

3.3 Approche bayésienne

Le modèle proposé dans la section précédente pour effectuer une décomposition conjointe est simple et performant dans le cas où le nombre de raies par spectre est connu et reste constant. Cette limite peut être levée en utilisant l'algorithme RJMCMC [94]. Bien qu'il existe d'autres algorithmes trans-dimensionnels (voir [95] pour une revue approfondie) l'algorithme RJMCMC est le plus ancien, le plus utilisé et le plus polyvalent.

L'extension du modèle et l'implémentation de l'algorithme RJMCMC a été menée en collaboration avec Sylvain Faisan, maître de conférence dans mon équipe. Le modèle que nous avons proposé est présenté ci-dessous et détaillé dans [153].

Modèle bayésien à dimension variable

Dans le modèle à dimension fixe présenté section 3.2, les raies sont associées à des classes z de sorte que chaque spectre s contienne une et une seule raie de chaque classe. La variable K représente à la fois le nombre de raies par spectre et le nombre de classes.

Si le nombre de raies peut varier d'un spectre à l'autre, cela signifie que les classes ne sont pas forcément toutes présentes dans chaque spectre. Une classe est composée d'une raie par spectre, avec la contrainte que toutes les raies d'une même classe soit présentes dans des spectres contigus. Nous greffons donc au modèle à dimension fixe la notion de début et de longueur de classe. Ainsi, chaque classe $k \in \{1, \dots, K\}$ débute au spectre d_k et est de longueur l_k . Dit autrement, les raies de la trace k sont situées dans les spectres d_k à $d_k + l_k - 1$: la classe k est un ensemble de $l_k \leq S$ raies similaires à travers les données. On note $M = \sum_{k=1}^K l_k$ le nombre total de raies. Ainsi, les objets élémentaires du modèle (les M raies) sont regroupées dans des objets plus gros (les K classes). Pour fixer les idées, la figure 3.7 représente une séquence synthétique avec les valeurs d_k et l_k pour chaque classe.

Le modèle s'écrit alors :

$$\forall s \in \mathbb{S}, \quad \mathbf{y}_s = \sum_{k=1}^K \sum_{m=1}^{l_k} a_{km} \varphi(c_{km}, w_{km}) \delta_{d_k+m-1, s} + \mathbf{b}_s. \quad (3.14)$$

Le symbole de Kronecker $\delta_{d_k+m-1, s}$ est égal à 1 si et seulement si $d_k + m - 1 = s$: il permet de coder la présence ou l'absence de la raie m de la classe k dans le spectre s .

Ce modèle est à comparer avec l'équation (3.1) : si toutes les classes sont présentes à chaque spectre, alors $d_k = 1$ et $l_k = S$ quel que soit k . Le modèle devient alors :

$$\forall s \in \mathbb{S}, \quad \mathbf{y}_s = \sum_{k=1}^K \sum_{m=1}^S a_{km} \varphi(c_{km}, w_{km}) \delta_{m, s} + \mathbf{b}_s \quad (3.15)$$

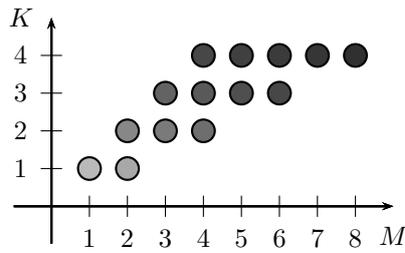


FIGURE 3.4 – Valeurs admissibles de M et K dans le cas où $S = 2$ et $K_{\max} = 4$. Chaque disque représente un espace de dimension (M, K) , et son niveau de gris indique sa probabilité : plus le disque est clair, plus grande est la probabilité $p(K, \mathbf{d}, \mathbf{l})$ (ici : $\xi = 0,9$).

et peut donc se simplifier en l'équation (3.1).

Nous pouvons maintenant définir les loi a priori des nouvelles variables, en l'occurrence K , \mathbf{d} et \mathbf{l} . Nous souhaitons que les données soient décrites avec le plus petit nombre de raies et de classes, aussi l'a priori $p(K, \mathbf{d}, \mathbf{l})$ doit favoriser les faibles valeurs de K et M . Nous avons donc choisi d'exprimer la loi jointe de ces variables en fonction de K et M :

$$p(K, \mathbf{d}, \mathbf{l}) \propto \xi^{M+SK} \mathbb{1}_{\mathbb{X}}(K, \mathbf{d}, \mathbf{l}) \quad (3.16)$$

où ξ est fixé à une valeur dans $]0,1[$ et \mathbb{X} est l'ensemble des valeurs possibles du triplet $(K, \mathbf{d}, \mathbf{l})$:

- le nombre K de classes est borné à une valeur maximale K_{\max} de sorte que le nombre de configurations soit fini et donc que $p(K, \mathbf{d}, \mathbf{l})$ soit intégrable ;
- les éléments de \mathbf{l} sont à valeurs dans $\{1, \dots, S\}$ (et donc $M \in \{K, \dots, SK\}$) ;
- les éléments de \mathbf{d} sont à valeurs dans $\{1, \dots, S - l_k + 1\}$.

Il est facile de voir que pour un nombre M de raies fixé, cet a priori privilégie les valeurs de K faibles. En conséquence, cet a priori favorise les classes les plus longues possibles. Du fait de ces contraintes, les variables K et M sont très corrélées. L'ensemble des valeurs possibles pour K et M est schématisé figure 3.4. Chaque couple (K, M) dans cette figure définit un espace de dimension constante. Avec cette modélisation, les modèles pour lesquels $M + SK$ sont identiques ont la même probabilité. Cette modélisation est toutefois inhabituelle car en général, les lois a priori sont définies indépendamment sur chaque variable. Or, cette démarche n'est pas possible car elle se confronte notamment à l'explicitation de constantes de normalisation qui, du fait de l'aspect combinatoire de cette modélisation, est trop difficile à utiliser.

Les autres lois a priori sont similaires à celles du modèle à dimension constante (section 3.2). En particulier, les lois a priori sur les paramètres des raies (équations (3.10)–(3.12)) sont redéfinies en tenant compte que les raies ne sont pas présentes partout. Le modèle est représenté par le graphe acyclique orienté figure 3.5.

Optimisation avec l'algorithme RJMCMC

L'algorithme RJMCMC a été choisi en raison de la non-linéarité du modèle et du nombre inconnu de variables. Nous résumons ci-dessous les mouvements implémentés et les moyens mis en œuvre pour accélérer la convergence de l'algorithme.

- Les traditionnels mouvements de naissance et de mort existent pour les raies, mais également pour les classes. On introduit également des combinaisons de ces mouvements (cf. figure 3.6). En comptant la simulation de l'a posteriori à dimension constante, nous aboutissons à huit mouvements :

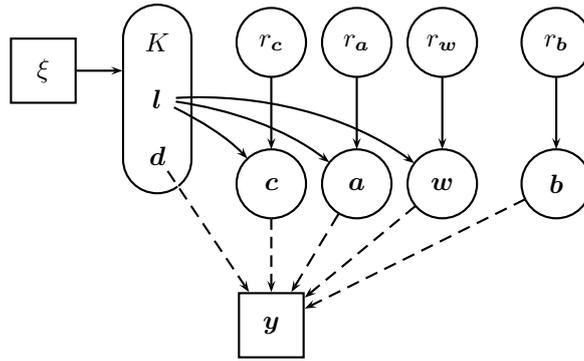


FIGURE 3.5 – Graphe acyclique orienté du modèle bayésien à dimension variable. Les carrés représentent les variables connues ou fixées, les cercles représentent les variable aléatoires inconnues. Les flèches indiquent les dépendances entre variables ; un trait continu correspond à une dépendance stochastique alors que des pointillés correspondent à une dépendance déterministe.

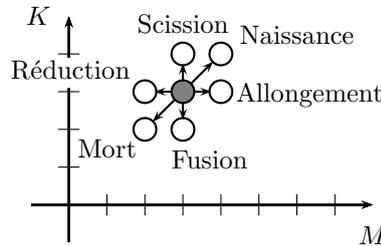


FIGURE 3.6 – Mouvements trans-dimensionnels dans l'espace (M, K) à partir du disque gris.

- une naissance ajoute une nouvelle classe contenant une unique raie ;
- une mort supprime une raie à l'extrémité d'une classe ne contenant qu'une raie ;
- une scission divise une classe en deux classes contiguës ;
- une fusion regroupe deux classes contiguës en une seule ;
- un allongement agrandit une classe en y ajoutant une raie à l'une des extrémités ;
- une réduction supprime une raie située à l'extrémité d'une classe ;
- une ré-affectation échange certaines raies d'une classe avec les raies d'une autre classe ;
- enfin, la mise à jour est un échantillonneur de Gibbs qui simule la loi a posteriori sans en changer la dimension.

Les six premiers mouvements sont trans-dimensionnels et permettent de faire varier seulement M , ou seulement K , ou les deux. Les mouvements proposés profitent ainsi de la représentation bidimensionnelle de l'espace (M, K) . C'est une particularité de notre modèle qui, à notre connaissance, n'a jamais été considérée dans le contexte de l'algorithme RJMCMC.

- Les probabilités d'effectuer ces mouvements ne sont pas égales et varient en fonction des mouvements effectués. Par exemple, une mise à jour a plus de chance d'être proposée après un mouvement trans-dimensionnel.
- L'utilisation de lois candidates mixtes permet d'augmenter le rapport d'acceptation des propositions. Les lois sont mixtes car elles permettent de générer des propositions suivant une loi uniforme ou suivant une loi construite à partir des données (ou du modèle).
- Une estimation au sens du maximum a posteriori est obtenue en choisissant parmi les échantillons générés celui qui maximise la loi a posteriori. Nous évitons ainsi de conserver en

mémoire la totalité de la chaîne générée et nous nous prévenons des problèmes de réordonnement d'indices (*label switching*) [111, 223, 240].

- L'algorithme n'est pas appliqué directement sur toutes les données, mais nous utilisons une approche multi-résolution : seuls quelques signaux sont d'abord décomposés et le résultat obtenu sert d'initialisation à l'étape suivante où plus de signaux sont traités. Cette procédure est possible grâce à l'hypothèse d'évolution douce des motifs.

Comparaison avec le modèle à dimension constante sur données synthétiques

Dans un premier temps, nous comparons le modèle présenté avec celui de la section 3.2. Ces deux modèles sont établis dans le cadre bayésien, mais le premier n'est pas capable de gérer le nombre inconnu de variables : K est connu et constant, ce qui signifie que le nombre de raies ne varie pas d'un spectre à l'autre. Les deux méthodes sont comparées sur une séquence regroupant $S = 60$ spectres de $N = 256$ échantillons. Le bruit est un bruit additif blanc gaussien de telle sorte à avoir un PSNR de 30 dB :

$$\text{PSNR} = 10 \log_{10} \left[\left(r_b \frac{1}{S} \sum_{s=1}^S \frac{1}{\max(\mathbf{y}_s)^2} \right)^{-1} \right]. \quad (3.17)$$

où r_b est la variance du bruit. La figure 3.7 représente quelques uns de ces spectres. On peut y voir des comportements délicats à traiter, tels que des raies très proches et un croisement. La méthode à dimension constante a été exécutée avec $K = 5$ pendant environ 4 h 16 alors que la méthode capable de gérer un nombre variable de raies a fonctionné durant environ 30 minutes. Le résultat obtenu est présenté figure 3.8. Avec la première méthode, certaines véritables raies sont estimées avec deux raies (comme par exemple au niveau de la zone grise). Cela a pour conséquence de créer des échanges de raies entre classes. En outre, le croisement en $s = 14$ entre les classes 1 et 2, où les vraies raies ont même centre et même largeur, n'est pas correctement estimé : les raies zigzaguent sur plusieurs signaux. Au contraire, la nouvelle approche permet une bien meilleure estimation des paramètres des raies, ainsi que du nombre de raies et de classes.

Résultat sur des données réelles de photoélectrons

La spectroscopie de photoélectrons résolue en temps [242] est une méthode permettant d'étudier le processus de relaxation de l'énergie qui apparaît après absorption d'un photon par une molécule ou un atome (ou un assemblage des deux). La détection de cette dynamique repose sur l'ionisation du système étudié, éjectant ainsi un électron, appelé alors photoélectron. L'énergie cinétique de ce dernier porte une information sur l'état du système au moment de son ionisation : à énergie de photon incident constante, l'électron est émis avec d'autant moins d'énergie cinétique que le système a relaxé son énergie électronique en énergie thermique. Un spectre de photoélectrons est la distribution des électrons en fonction de leur énergie. L'acquisition d'un spectre de photoélectrons en fonction du temps permet donc de suivre l'évolution du phénomène de relaxation [90] : on aboutit alors à une séquence temporelle de spectres de photo-électrons, comme celle présentée figure 3.9. Les centres c des raies correspondent aux énergies auxquelles la photo-émission apparaît. Les amplitudes a et les paramètres de forme w sont liés à l'aire de chaque raie qui correspond au nombre relatif de photo-électrons émis à l'énergie considérée. L'objectif de l'étude est de déterminer comment l'énergie et l'aire des raies évoluent dans le temps (donc à travers la séquence) indiquant ainsi les modifications électroniques que subit le système étudié. Comme d'autres travaux en dynamique réactionnelle [213, 214], le nombre et la position des raies pour chaque spectre ne sont pas connus

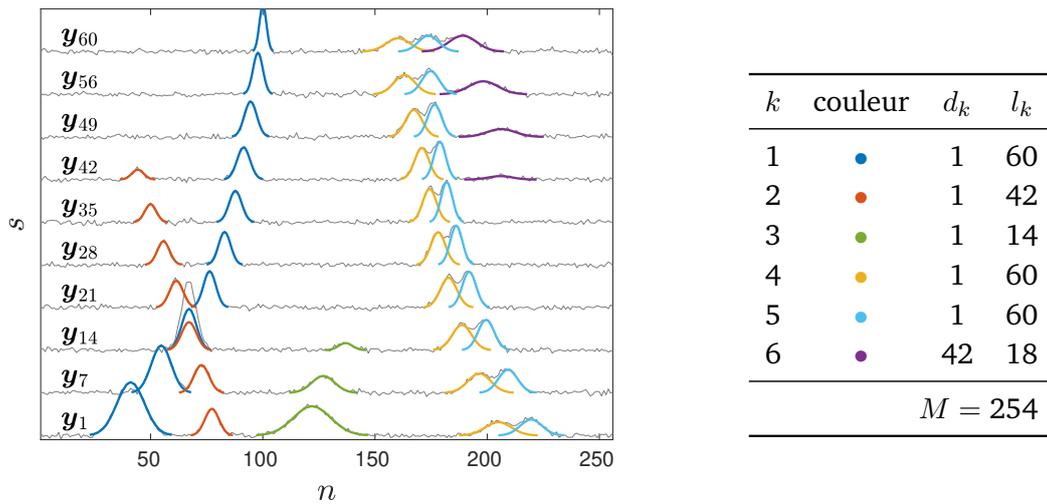


FIGURE 3.7 – Séquence simulée de $S = 60$ spectres de $N = 256$ échantillons. Seuls dix spectres de la séquence sont représentés (en gris). Les raies en couleurs correspondent aux raies réelles. Le tableau donne pour chaque classe k les débuts d_k et longueur l_k .

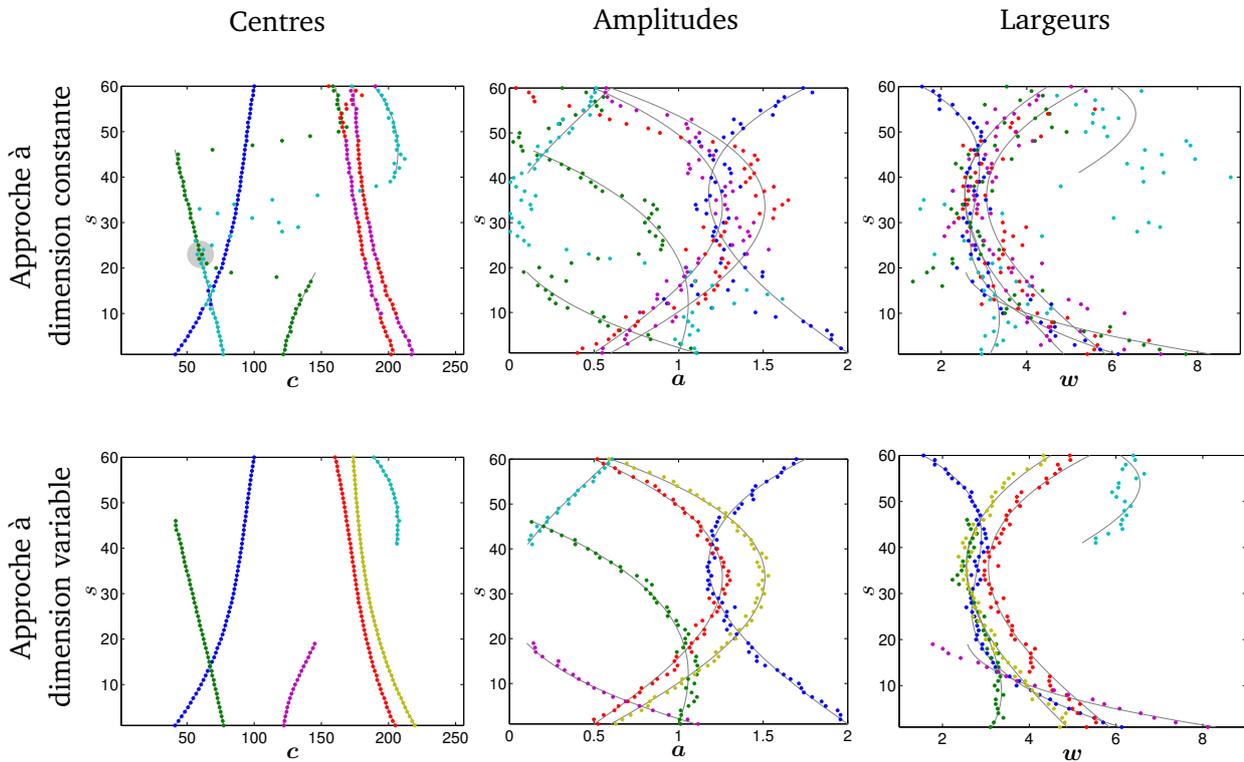


FIGURE 3.8 – Paramètres estimés sur la séquence de la figure 3.7 avec les approches à dimension constante et variable. Les trajectoires grises représentent la vérité terrain. Les couleurs des points correspondent aux classes des raies estimées.

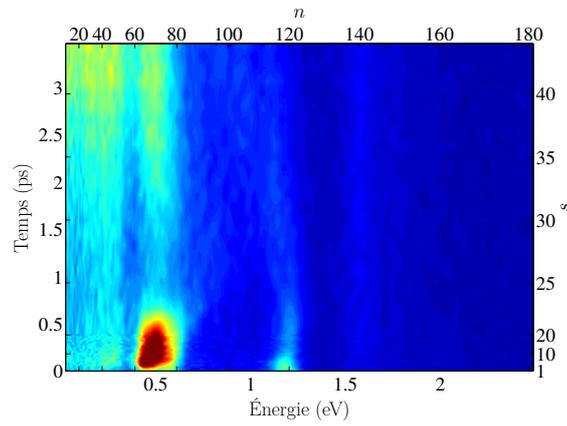
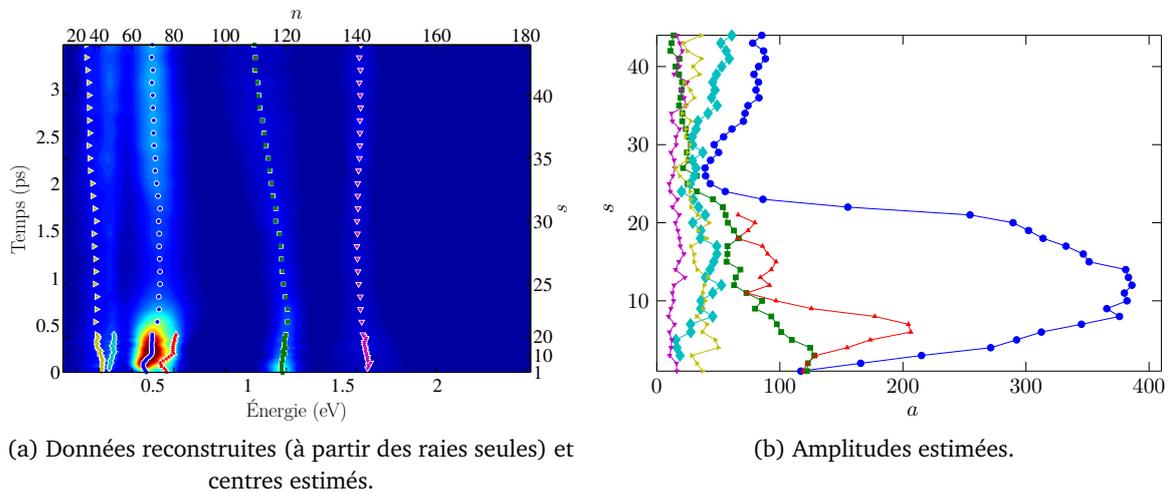


FIGURE 3.9 – Séquence de 44 spectres de photoélectrons. Les spectres de photoélectrons ne sont pas représentés comme des signaux, mais comme une image dont la couleur indique l’amplitude des spectres.



(a) Données reconstruites (à partir des raies seules) et centres estimés.

(b) Amplitudes estimées.

FIGURE 3.10 – Résultats sur la séquence de la figure 3.9.

et ils peuvent varier d’un spectre à l’autre. Par ailleurs, les raies peuvent se superposer et le rapport signal-sur-bruit peut être mauvais.

La séquence considérée ici correspond à la relaxation d’un atome de baryum [146]. Cette séquence regroupe $S = 44$ spectres (d’une durée totale de 3,47 ps), chacun de longueur $N = 181$ échantillons (de 0,02 eV à 2,52 eV). Cette séquence présente, en plus des raies, une ligne de base qui se traduit par des amplitudes fortes aux énergies faibles. La ligne de base a été modélisée comme une exponentielle complexe $\alpha_s \exp(-\lambda_n/\beta_s)$ et s’ajoute à l’équation 3.1. Les paramètres α_s et β_s de la ligne de base sont traités de la même manière que les paramètres des raies : leurs lois a priori sont similaires et l’algorithme MCMC les génère de la même façon.

La méthode proposée fournit une estimation très satisfaisante (figure 3.10), puisque les raies évoluent lentement (contrairement à ce qu’on obtiendrait avec une approche séquentielle [146, Fig. 4]). Par ailleurs, certaines zones sont estimées par deux raies qui se superposent (aux alentours de 0,25 eV et 0,5 eV dans les premiers temps). Ces résultats ont été validés par les experts et confirment les observations qualitatives données dans [146].

3.4 Approche par représentation parcimonieuse

Les approches bayésiennes décrites précédemment ont l'avantage d'offrir une modélisation simple et efficace, mais l'utilisation de méthodes MCMC implique un temps de calcul très long avec, en pratique, une mauvaise exploration de l'espace des solutions. En outre, le modèle que nous avons développé ne permet pas de gérer le cas où l'espace \mathbb{S} des signaux est bidimensionnel. En effet, comme l'équation (3.14) définit un début et une longueur pour chaque classe, elle n'est pas adaptée au cas où les classes s'étendent dans un espace bidimensionnel.

Pour ces raisons, nous avons souhaité explorer l'utilisation des représentations parcimonieuses. Justement, les signaux que nous considérons sont intrinsèquement parcimonieux puisqu'ils sont constitués d'un faible nombre de raies. En plus, de nombreux algorithmes d'optimisation ont été développés ces dernières années et on peut supposer qu'ils sont beaucoup plus rapides que des algorithmes MCMC. La méthode présentée a été développée dans le cadre de la thèse d'Hassan Mortada [173, 177]. Elle est présentée ici sous un éclairage différent.

Modèle

Par rapport aux modèles précédents, le modèle développé ici diffère sur deux points : nous supposons que le nombre K de raies est connu et que les paramètres de formes ne varient pas au sein d'une classe (donc $w_{ks} = w_k$)¹. Le cas où K n'est pas constant est possible et sera évoqué dans la suite. Le modèle des signaux s'écrit donc :

$$\forall s \in \mathbb{S}, \quad \mathbf{y}_s = \sum_{k=1}^K a_{ks} \varphi(c_{ks}, w_k) + \mathbf{b}_s. \quad (3.18)$$

Sous hypothèse de bruit blanc gaussien, le terme d'attache aux données s'écrit :

$$E(\mathbf{a}, \mathbf{c}, \mathbf{w}) = \sum_{s \in \mathbb{S}} \left\| \mathbf{y}_s - \sum_{k=1}^K a_{ks} \varphi(c_{ks}, w_k) \right\|^2. \quad (3.19)$$

L'hypothèse d'évolution douce des paramètres est prise en compte en ajoutant au terme (3.19) un terme de régularisation R . Nous nous sommes restreint à ne considérer que l'évolution des centres (l'évolution des paramètres est inutile puisqu'ils sont supposés constants pour une même classe). Pour favoriser une évolution douce des centres, nous faisons le choix classique de minimiser l'écart entre deux centres consécutifs : le terme de régularisation R utilise donc la variation totale. On retrouve la même idée que dans la modélisation bayésienne (3.10). Dans le cas où l'espace des signaux \mathbb{S} est unidimensionnel, par exemple $\mathbb{S} = \{1, \dots, S\}$, alors

$$R(\mathbf{c}) = \sum_{s=2}^S \sum_{k=1}^K (c_{ks} - c_{k(s-1)})^2 \quad (3.20)$$

et dans le cas d'un espace \mathbb{S} bidimensionnel qui peut être défini en toute généralité $\mathbb{S} = \{1, \dots, \Xi\} \times \{1, \dots, Z\}$ alors

$$R(\mathbf{c}) = \sum_{\xi=1}^{\Xi} \sum_{\zeta=2}^Z \sum_{k=1}^K (c_{k\xi\zeta} - c_{k\xi(\zeta-1)})^2 + \sum_{\xi=2}^{\Xi} \sum_{\zeta=1}^Z \sum_{k=1}^K (c_{k\xi\zeta} - c_{k(\xi-1)\zeta})^2. \quad (3.21)$$

1. Une extension de la méthode présentée qui est capable de considérer des paramètres de forme variables a été abordée dans la thèse d'Hassan Mortada [173, annexe B]. Mais cette extension reste encore une pré-étude qui mérite d'être éprouvée.

Pour simplifier l'exposé, nous n'utiliserons dans la suite que l'expression (3.20) ; l'extension au cas où \mathbb{S} est à plusieurs dimensions est directe.

Le problème est donc la minimisation de la somme pondérée du terme d'attache aux données et du terme de régularisation :

$$(\hat{\mathbf{a}}, \hat{\mathbf{c}}, \hat{\mathbf{w}}) = \arg \min_{\mathbf{a} \geq 0, \mathbf{c}, \mathbf{w}} F(\mathbf{a}, \mathbf{c}, \mathbf{w}) \quad (3.22)$$

où $F(\mathbf{a}, \mathbf{c}, \mathbf{w}) = E(\mathbf{a}, \mathbf{c}, \mathbf{w}) + \tau R(\mathbf{c})$ et τ est un paramètre de régularisation à régler manuellement. On notera la présence d'une contrainte de non-négativité des amplitudes. La minimisation est effectuée en plusieurs étapes, détaillées ci-après.

Minimisation de $F(\mathbf{a}, \mathbf{c}, \mathbf{w})$ par rapport à \mathbf{a} , \mathbf{c} et \mathbf{w}

L'algorithme 3.1 donne un aperçu général de la minimisation du critère par rapport à \mathbf{a} , \mathbf{c} et \mathbf{w} . Nous choisissons d'utiliser un schéma de moindres carrés alternés (ALS : *alternating least squares*) pour effectuer la minimisation de (3.22). Ce schéma consiste à minimiser le critère en fonction de certaines variables en fixant les autres, et vice-versa. Bien qu'il n'existe pas de preuve de convergence vers le minimum global, cette méthode est très utilisée en séparation de sources retardées où, par exemple, les sources, les retards et les amplitudes sont estimées alternativement [106, 178, 189]. Nous considérons donc la convergence atteinte lorsque la descente relative du critère entre deux itérations successives devient inférieure à un seuil. La minimisation en fonction de \mathbf{w} est un problème de moindres carrés non linéaires et peut être optimisée avec un algorithme classique comme l'algorithme de Levenberg-Marquardt. En revanche, la minimisation en fonction de \mathbf{a} et \mathbf{c} est plus difficile, elle est détaillée ci-après.

- 1 Initialisation : $i = 0$, $\mathbf{w}^{(0)}$ généré aléatoirement
- 2 Répéter
- 3 $i = i + 1$
- 4 $\mathbf{a}^{(i)}, \mathbf{c}^{(i)} = \arg \min_{\mathbf{a} \geq 0, \mathbf{c}} F(\mathbf{a}, \mathbf{c}, \mathbf{w}^{(i-1)})$ (cf. algorithme 3.2)
- 5 $\mathbf{w}^{(i)} = \arg \min_{\mathbf{w}} E(\mathbf{a}^{(i)}, \mathbf{c}^{(i)}, \mathbf{w})$ (Levenberg-Marquardt)
- 6 Jusqu'à convergence
- 7 Solution : $\hat{\mathbf{a}} = \mathbf{a}^{(i)}$, $\hat{\mathbf{c}} = \mathbf{c}^{(i)}$, $\hat{\mathbf{w}} = \mathbf{w}^{(i)}$

ALGORITHME 3.1 – Algorithme ALS pour minimiser $F(\mathbf{a}, \mathbf{c}, \mathbf{w})$ en fonction de \mathbf{a} , \mathbf{c} et \mathbf{w} .

Minimisation de $F(\mathbf{a}, \mathbf{c}, \mathbf{w})$ par rapport à \mathbf{a} et \mathbf{c}

Comme on l'a dit plus haut, les signaux considérés sont intrinsèquement parcimonieux. Aussi, pour minimiser le critère $F(\mathbf{a}, \mathbf{c}, \mathbf{w})$ par rapport à \mathbf{a} et \mathbf{c} à \mathbf{w} fixé, nous nous inspirons des algorithmes d'approximation parcimonieuse. En particulier, les algorithmes gloutons fonctionnent très bien lorsque le niveau de parcimonie K est faible et que les atomes du dictionnaire sont très corrélés. De plus, ils sont suffisamment simples pour être facilement adaptés aux représentations parcimonieuses avec structure. C'est la raison pour laquelle nous avons choisi de nous inspirer de l'algorithme NN-OMP (*non-negative orthogonal matching pursuit*) [35] : l'algorithme 3.2 suit les mêmes étapes que l'algorithme proposé dans [35] :

1. balayage : le minimum du critère $\epsilon(k)$ est calculé pour toutes les raies qui ne sont pas déjà sélectionnées (ligne 4) ;
2. mise à jour du support : les lignes 5–7 cherchent la meilleure classe k et l'ajoute au support \mathcal{S} . Les centres c sont également mis à jour ;
3. mise à jour de la solution : la ligne 8 calcule les amplitudes \mathbf{a} en tenant compte du support \mathcal{S} et de la contrainte de non-négativité ;
4. règle de convergence : contrairement à [35], la règle de convergence consiste à comparer la valeur du critère à deux itérations successives : si le critère remonte, alors la dernière itération n'est finalement pas utilisée pour l'estimation. La raison de cette règle de convergence est que l'ajout d'une nouvelle classe peut faire remonter le critère F en raison de la présence du terme de régularisation.

Notons que l'étape de mise à jour du résidu, présente dans [35], n'est pas utile ici puisque le résidu n'est pas explicitement utilisé.

```

1 Initialisation :  $i = 0, \mathbf{c}^{(0)} = \mathbf{0}, \mathcal{S}^{(0)} = \emptyset$ 
2 Répéter
3    $i = i + 1$ 
4   Calculer pour tout  $k \notin \mathcal{S}^{(i-1)}$  (cf. algorithme 3.3) :
           
$$\epsilon(k) = \min_{\mathbf{a}_k \geq 0, c_k} F(\mathbf{a}^{(i-1)}, \mathbf{c}^{(i-1)}, \mathbf{w})$$

5   Choisir  $\tilde{k}$  tel que  $\epsilon(\tilde{k}) \leq \epsilon(k)$  pour tout  $k$ 
6   Mettre à jour le support :  $\mathcal{S}^{(i)} = \mathcal{S}^{(i-1)} \cup \{\tilde{k}\}$ 
7   Mettre à jour les centres :  $\mathbf{c}^{(i)} = \mathbf{c}^{(i-1)}$  sauf  $c_{\tilde{k}}^{(i)}$  qui est la valeur de  $c_k$  qui a permis de
   calculer  $\epsilon(\tilde{k})$ 
8   Mettre à jour les amplitudes (expression explicite) :
           
$$\mathbf{a}^{(i)} = \min_{\mathbf{a}} E(\mathbf{a}, \mathbf{c}^{(i)}, \hat{\mathbf{w}}) \quad \text{tel que} \quad \text{Support}\{\mathbf{a}^{(i)}\} = \mathcal{S}^{(i)} \quad \text{et} \quad \mathbf{a}^{(i)} \geq 0$$

9 Jusqu'à  $F(\mathbf{a}^{(i)}, \mathbf{c}^{(i)}, \mathbf{w}) > F(\mathbf{a}^{(i-1)}, \mathbf{c}^{(i-1)}, \mathbf{w})$ 
10 Solution :  $\hat{\mathbf{a}} = \mathbf{a}^{(i-1)}, \hat{\mathbf{c}} = \mathbf{c}^{(i-1)}$ 

```

ALGORITHME 3.2 – Algorithme glouton pour minimiser $F(\mathbf{a}, \mathbf{c}, \mathbf{w})$ en fonction de \mathbf{a} et \mathbf{c} .

Du fait du caractère parcimonieux, l'algorithme peut s'écrire sous une autre forme, en mettant en évidence un dictionnaire : c'est ce que nous avons présenté dans [177]. Le dictionnaire s'écrit alors comme l'ensemble des raies φ décalées. Le vecteur parcimonieux a également une structure spécifique afin de prendre en compte le fait que la classe k n'est présente qu'une fois dans chaque signal.

Minimisation de $F(\mathbf{a}, \mathbf{c}, \mathbf{w})$ par rapport à \mathbf{a}_k et c_k

Le problème à considérer est de calculer, pour k fixé :

$$\epsilon(k) = \min_{\mathbf{a}_k \geq 0, c_k} F(\mathbf{a}, \mathbf{c}, \mathbf{w}). \quad (3.23)$$

où les trois vecteurs \mathbf{a} , \mathbf{c} et \mathbf{w} sont connus, à l'exception des éléments de \mathbf{a}_k et \mathbf{c}_k :

$$\widehat{\mathbf{a}}_k, \widehat{\mathbf{c}}_k = \arg \min_{\mathbf{a}_k \geq 0, \mathbf{c}_k} F(\mathbf{a}, \mathbf{c}, \mathbf{w}) \quad (3.24)$$

$$= \arg \min_{\mathbf{a}_k \geq 0, \mathbf{c}_k} \sum_{s \in \mathbb{S}} \left\| \mathbf{r}_s^k - a_{ks} \boldsymbol{\varphi}(c_{ks}, w_k) \right\|^2 + \tau \sum_{s=2}^S (c_{ks} - c_{k(s-1)})^2 \quad (3.25)$$

où \mathbf{r}_s^k est le résidu défini comme

$$\mathbf{r}_s^k = \mathbf{y}_s - \sum_{j \neq k} a_{js} \boldsymbol{\varphi}(c_{js}, w_j). \quad (3.26)$$

La minimisation de (3.25) par rapport à \mathbf{a}_k et pour \mathbf{c}_k fixé possède une expression explicite :

$$\forall s, \quad \widehat{a}_{ks} = \left[\mathbf{r}_s^{kT} \boldsymbol{\varphi}(c_{ks}, w_k) \right]_+ \quad (3.27)$$

où la notation $[u]_+$ correspond à $[u]_+ = \max(u, 0)$. En intégrant cette expression dans (3.25), on obtient :

$$\widehat{\mathbf{c}}_k = \arg \max_{\mathbf{c}_k} \sum_{s \in \mathbb{S}} \left[\mathbf{r}_s^{kT} \boldsymbol{\varphi}(c_{ks}, w_k) \right]_+^2 - \tau \sum_{s=2}^S (c_{ks} - c_{k(s-1)})^2 \quad (3.28)$$

Pour résoudre ce problème combinatoire, nous nous sommes inspirés de l'algorithme ICM (*iterated conditional modes*) [23], généralement utilisé en traitement d'images. Bien qu'il converge vers un optimum local, les résultats sont satisfaisants dans nos applications. Ainsi, en chaque site $s \in \mathbb{S}$, le centre c_{ks} est estimé comme celui qui minimise :

$$\widehat{c}_{ks} = \arg \max_{c_{ks}} \left[\mathbf{r}_s^{kT} \boldsymbol{\varphi}(c_{ks}, w_k) \right]_+^2 - \tau (c_{ks} - c_{k(s-1)})^2 - \tau (c_{k(s+1)} - c_{ks})^2 \quad (3.29)$$

Bien entendu, lorsqu'on considère un site au bord du domaine \mathbb{S} , alors l'un des termes de régularisation n'apparaît pas dans l'expression précédente. Ainsi, tous les sites sont visités de proche en proche. Une fois que tout le domaine \mathbb{S} est visité, l'algorithme recommence tant que le nombre d'éléments de $\widehat{\mathbf{c}}_k$ ayant été modifié depuis l'itération précédente reste trop grand. L'algorithme est initialisé en commençant par le site ayant le terme $\left[\mathbf{r}_s^{kT} \boldsymbol{\varphi}(c_{ks}, w_k) \right]_+^2$ le plus important. L'algorithme 3.3 résume ces différentes étapes.

```

1 Initialisation :  $\mathbf{c}_k^{(0)} = \emptyset$ 
2 Répéter
3    $i = i + 1$ 
4    $\mathbf{c}_k^{(i)} = \mathbf{c}_k^{(i-1)}$ 
5   Choisir le site  $s$ 
6     — à la première itération : celui qui maximise  $\left[ \mathbf{r}_s^{kT} \boldsymbol{\varphi}(c_{ks}^{(i)}, w_k) \right]_+^2$ 
7     — aux itérations suivantes : un site au hasard
6   Pour chaque site  $s$ , en les prenant de proche en proche :
7      $c_{ks}^{(i)} = \arg \max_{c_{ks}} \left[ \mathbf{r}_s^{kT} \boldsymbol{\varphi}(c_{ks}^{(i)}, w_k) \right]_+^2 - \tau (c_{ks} - c_{k(s-1)}^{(i)})^2 - \tau (c_{k(s+1)}^{(i)} - c_{ks})^2$ 
8   Fin
9 Jusqu'à ce que  $\mathbf{c}_k^{(i)}$  soit presque le même que  $\mathbf{c}_k^{(i-1)}$ 
10 Solution :  $\widehat{\mathbf{c}}_k = \mathbf{c}_k^{(i)}$ 

```

ALGORITHME 3.3 – Algorithme pour minimiser $F(\mathbf{a}, \mathbf{c}, \mathbf{w})$ en fonction de \mathbf{c}_k .

Résultats sur des données simulées

La méthode est évaluée sur la séquence représentée figure 3.11. Elle contient $S = 40$ spectres de $N = 200$ échantillons et $K = 4$ classes de raies gaussiennes. Les données présentent des particularités comme des sources proches, des croisements et un nombre non constant de raies (les classes ne sont pas toujours présentes).

L'approche par représentation parcimonieuse est comparée sur une séquence simulée avec les méthodes de Mørup [178], Castella [38] et la méthode bayésienne de la section 3.3. Les méthodes de Mørup et Castella ont été choisies car ce sont des méthodes de séparation de sources retardées ou convolutives, qui sont capables d'estimer les raies (les sources), mais n'utilisent pas l'information que les raies sont paramétrées. Ainsi, bien que la reconstruction soit satisfaisante (figure 3.12), les raies estimées peuvent être très différentes des raies réelles. Les centres et amplitudes estimés sont également non satisfaisants : les estimations sont très bruitées et ne correspondent pas à la vérité terrain. En revanche, les deux méthodes développées dans ce chapitre fournissent des résultats très satisfaisants et acceptables : les reconstructions et les valeurs des paramètres estimés sont très satisfaisantes (c'est également confirmé par des mesures d'erreurs entre les estimations et la vérité terrain [177]). Les deux principales différences entre les deux méthodes présentées dans ce chapitre sont d'une part la gestion du nombre variable de raies pour l'approche bayésienne (ce qui n'est pas le cas de l'approche par représentation parcimonieuse) et d'autre part le temps de calcul (44,2 s pour l'approche bayésienne contre 0,5 s pour l'approche par représentation parcimonieuse).

Résultats sur des données réelles de photoélectrons

L'approche par représentation parcimonieuse est appliquée sur les données de photoélectron déjà traitées ci-avant. Les paramètres de la ligne de base sont estimés grâce à la méthode de Levenberg-Marquardt dans une étape supplémentaire de l'algorithme 3.1. Les résultats obtenus avec cette nouvelle approche sont présentés figure 3.13 : ils sont équivalents à ceux obtenus avec l'approche bayésienne décrite dans la section 3.3, hormis en ce qui concerne le nombre de raies dont le nombre n'est pas, avec la version actuelle de l'approche par représentation parcimonieuse, estimé. Toutefois, un post-traitement permettrait d'éliminer les raies d'amplitudes trop faibles. On peut également imaginer une variante des algorithmes proposés de manière à ce que les raies ajoutées au fur et à mesure ne le soit qu'à condition qu'elles permettent une descente significative du critère. En outre, l'approche par représentation parcimonieuse remplit l'objectif initial puisque cette approche est environ 560 fois plus rapide que la méthode bayésienne (7,5 secondes contre 70 minutes) ! Cette nouvelle méthode est également bien plus simple à mettre en œuvre, ce qui remplit un deuxième objectif.

Résultats sur des données réelles de cinématique de galaxies

Un autre intérêt de l'approche par représentation parcimonieuse est sa capacité à traiter des données réparties dans un espace de dimension supérieure ou égale à 1. Ainsi, la décomposition des spectres d'une image multispectrale est possible. La méthode est donc appliquée à une observation réelle de la galaxie NGC 4254 dans les bandes radio [209] (figure 3.14) Plus précisément, la galaxie est observée à des longueurs d'onde proches de 21 cm, qui correspond au rayonnement d'atomes d'hydrogène neutre présents en grande quantité dans le gaz galactique. En raison de la rotation à très grande vitesse de la galaxie, la longueur d'onde d'émission de la raie peut se décaler en longueur d'onde : c'est l'effet Doppler. La mesure de ce décalage indique la vitesse de déplacement du gaz.

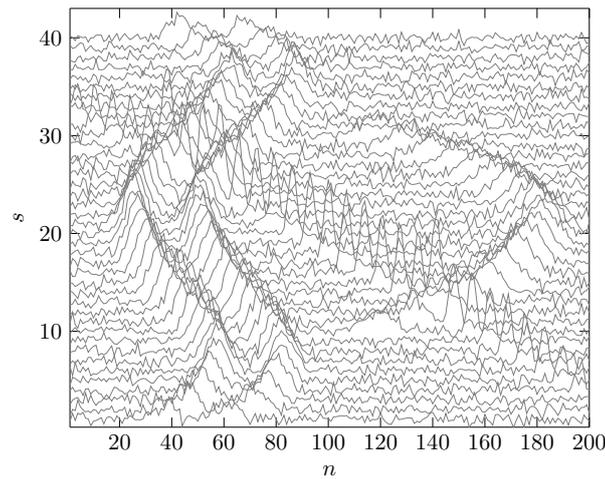


FIGURE 3.11 – Séquence simulée de $S = 40$ spectres de $N = 200$ échantillons.

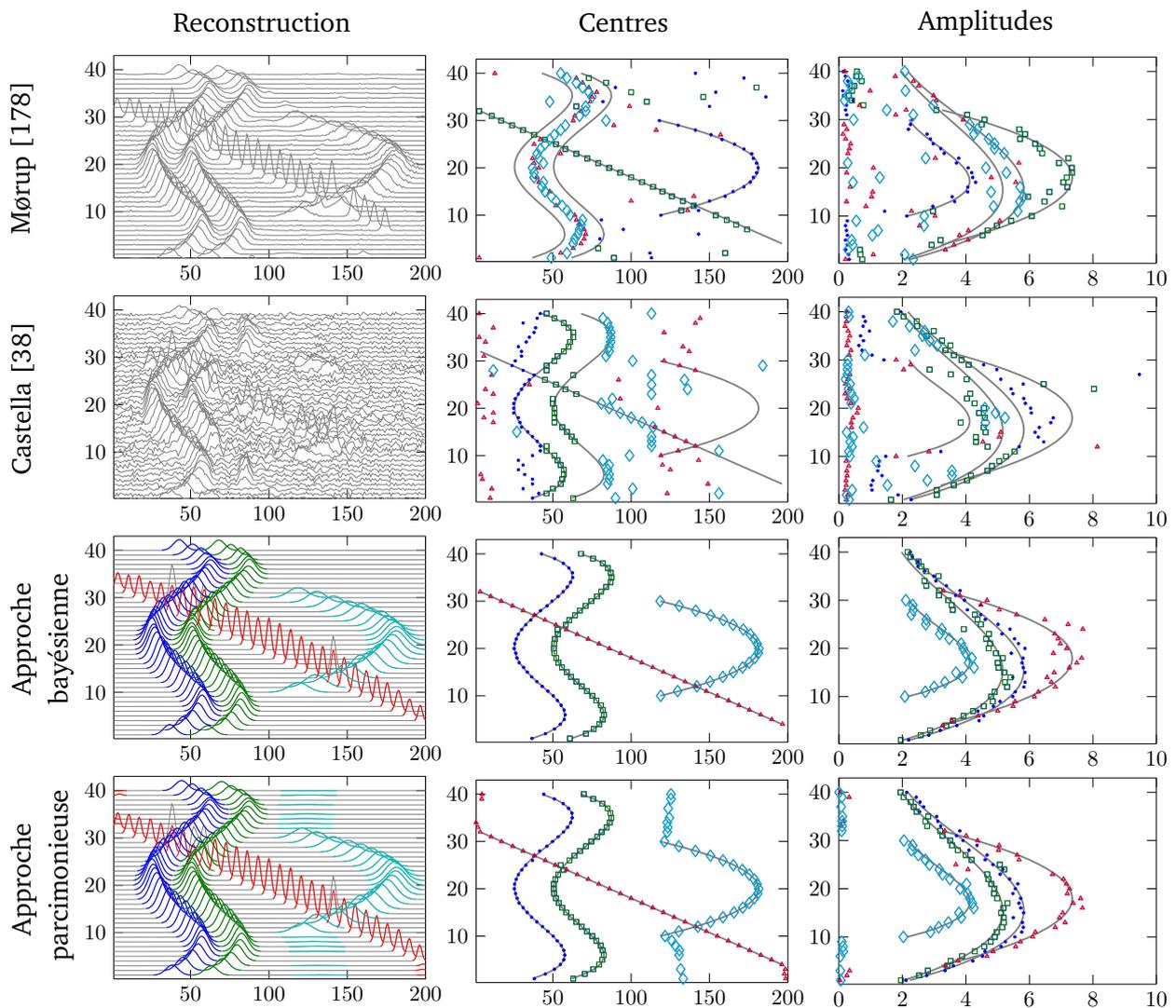


FIGURE 3.12 – Résultats sur la séquence de la figure 3.11 avec quatre approches. Les colonnes représentent respectivement les données reconstruites (et, le cas échéant, les raies estimées) et les estimations des centres et des amplitudes. Les figures sont graduées en fonction des échantillons (en abscisses) et des signaux (en ordonnées). Les trajectoires grises des deux colonnes de droites correspondent aux centres et amplitudes réels. Les couleurs et les marques correspondent aux classes.

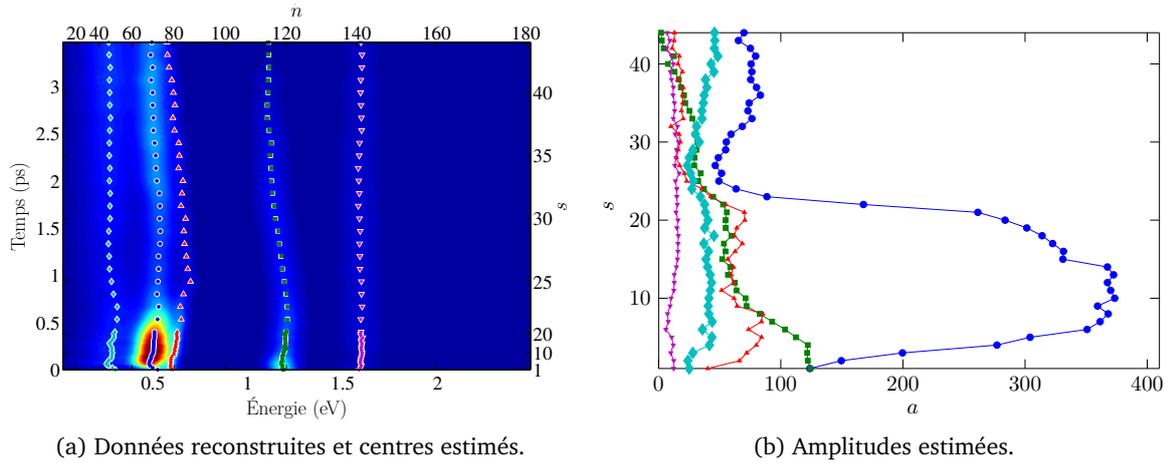


FIGURE 3.13 – Résultats sur la séquence de la figure 3.9.

Pour schématiser, la galaxie est en rotation et donc les déplacements mesurés à travers l'image seront différents d'un pixel à l'autre : les paramètres des raies évoluent à travers la séquence. Cette mesure de déplacement permet de représenter la cinématique des différentes structures composant la galaxie, et finalement de pouvoir comprendre l'histoire et l'évolution de la galaxie. Pour un état de l'art des méthodes existantes, voir par exemple [19].

La séquence considérée regroupe 140×140 signaux. Le nombre de classes est fixé à $K = 3$. Dans ces données, il est important de gérer un nombre variable de raies. Pour ce faire, l'algorithme 3.3 est modifié de sorte à définir un sous-ensemble de \mathbb{S} qui correspond aux sites s pour lesquels la corrélation $\left[\mathbf{r}_s^{kT} \boldsymbol{\varphi}(c_{ks}, w_k) \right]_+$ est supérieure à un seuil préalablement défini. Le parcours des sites est effectué grâce à un algorithme de croissance de région. L'algorithme 3.2 est également adapté pour ne prendre en compte que les sites de ce sous-ensemble. Les détails de cette implémentation sont disponibles dans [173, annexe A].

La figure 3.14 donne la reconstruction obtenue et le résidu, et la figure 3.15 donne les trois classes estimées. La première classe représente la principale structure cinématique de la galaxie : deux bras spiraux (que l'on distingue sur les amplitudes estimées) tournant autour du centre de la galaxie. La position estimée, qui est équivalente au décalage spectral de la raie à 21 cm, montre clairement que le gaz tourne autour d'un axe situé au niveau des pixels représentés en jaune. La seconde classe représente une structure très atténuée en amplitude et dont la cinématique est constante. Pour obtenir cette classe, il a fallu initialiser manuellement l'algorithme 3.3 en précisant comme pixel initial un pixel situé dans cette zone. Enfin, la troisième classe représente une structure synthétique qui a été ajoutée aux données. Cette structure a une géométrie circulaire, une intensité décroissante depuis un pixel central et une cinématique constante. Ces caractéristiques sont effectivement retrouvées par notre approche. Le temps de calcul pour traiter ces 140^2 signaux est encore une fois très rapide : 271 s.

3.5 Conclusion

Ce chapitre regroupe les différentes recherches que j'ai effectuées concernant la décomposition d'une séquence de signaux spectroscopiques en raies d'émission. Deux types de données ont été considérés et elles correspondent à des séquences unidimensionnelle ou bidimensionnelle. De par la nature des données, les raies évoluent entre les différents signaux, elles peuvent se croiser et sont très corrélées. Heureusement, la modélisation paramétrique des raies permet de surmonter ces

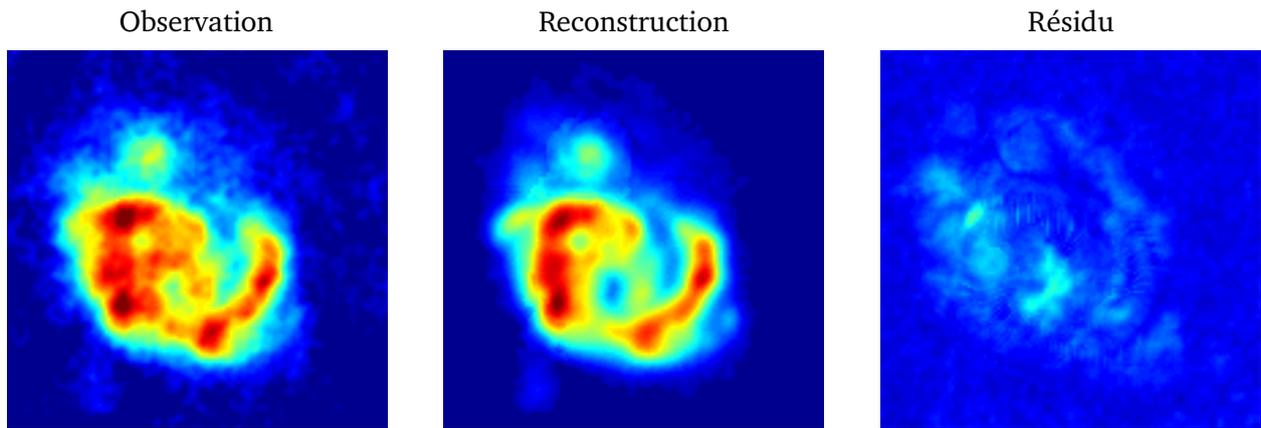


FIGURE 3.14 – Image multispectrale de la galaxie NGC 4254. Ces représentations sont appelées « images blanches » et correspondent aux moyennes selon les longueurs d’onde. Les données représentent une séquence de 140×140 signaux. Les niveaux de couleurs sont identiques sur les trois images.

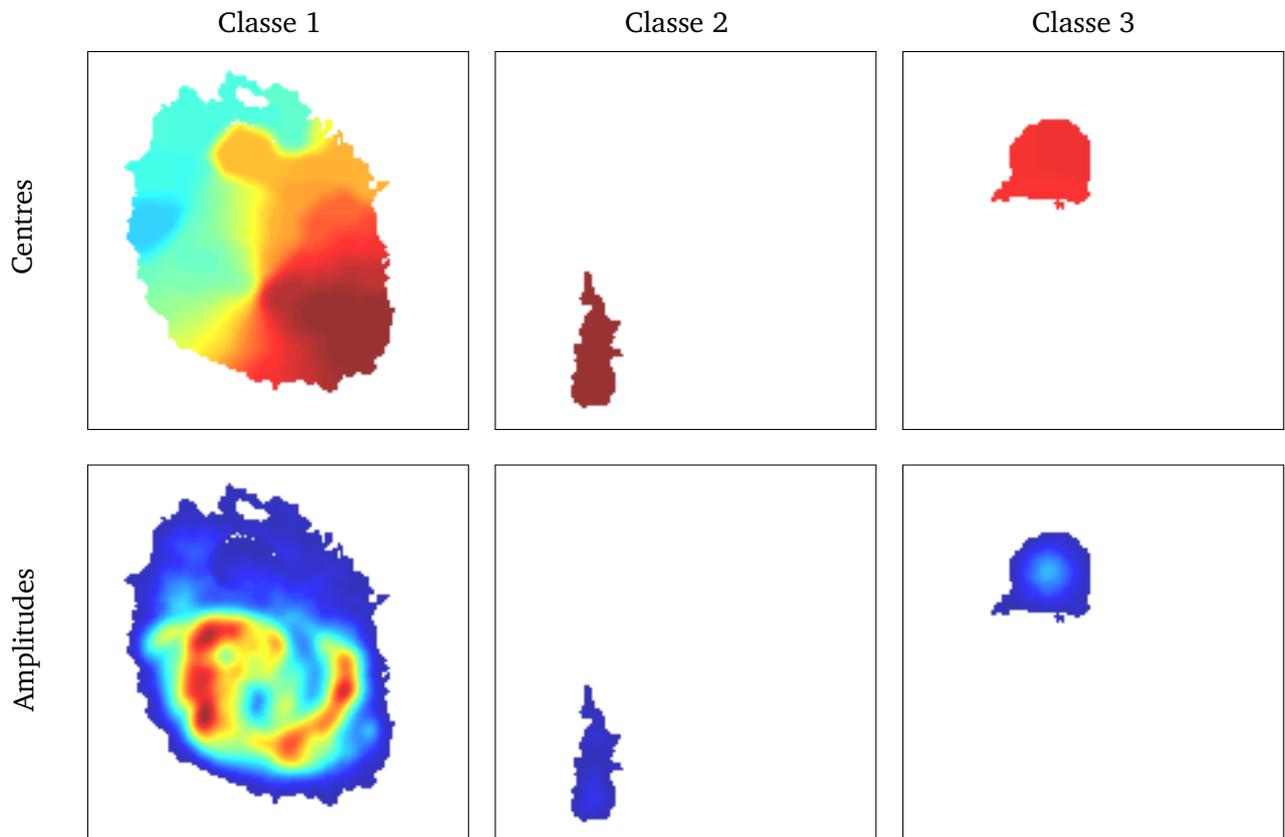


FIGURE 3.15 – Paramètres des trois classes estimées à partir des données de la figure 3.14. Les niveaux de couleurs sont identiques sur les trois images.

difficultés.

Dans un premier temps, nous avons montré qu'une décomposition conjointe est préférable à une décomposition séquentielle car cela permet non seulement d'obtenir une estimation plus réaliste mais également de guider la décomposition des signaux voisins. Pour cela, une première approche bayésienne a été mise en œuvre avec un échantillonneur de Gibbs. Ce modèle a ensuite été étendu pour gérer un nombre variables de raies et de classes ; nous avons alors choisi l'algorithme RJMCMC pour réaliser l'échantillonnage de l'a posteriori. Mais cette approche et notre implémentation pâ-tissent de plusieurs limitations : modélisation et implémentation laborieuses, mauvaise exploration de l'espace des solutions, temps de calcul très long et extension au cas de séquences bidimensionnelles très difficile. Par conséquent, une troisième approche utilisant les techniques d'approximation parcimonieuse a été mise en place. Elle donne des résultats similaires à l'approche bayésienne mais est beaucoup plus simple et rapide. Cette approche est toutefois heuristique sous certains aspects, et bien qu'elle fournisse des solutions très satisfaisantes, il n'y a pas de preuve de convergence théorique et il y a le risque de convergence vers un minimum local.

Les perspectives de ce travail constituent une partie de mon projet de recherche, détaillé dans le chapitre 5. Brièvement, il s'agit d'explorer les capacités des techniques d'approximation parcimonieuse à permettre une sélection de modèle et à effectuer un apprentissage de dictionnaire paramétré.

Chapitre 4

Décomposition d'une image astronomique en structures physiques

Contrairement aux deux chapitres précédents, les signaux sont désormais supposés bidimensionnels : le problème correspond donc à la décomposition d'une image multibande $\mathbf{y} \in \mathbb{R}^{M \times N \times B}$ en motifs paramétriques $\varphi \in \mathbb{R}^{M \times N \times B}$. En présence de bruit additif $\mathbf{b} \in \mathbb{R}^{M \times N \times B}$, le modèle s'écrit :

$$\mathbf{y} = \sum_{k=1}^K a_k \varphi(\mathbf{c}_k, \mathbf{w}_k) + \mathbf{b} \quad (4.1)$$

où les K motifs sont d'amplitude a_k , de centre \mathbf{c}_k et de paramètres de forme \mathbf{w}_k .

Ce problème a été étudié lors du stage de M2 de Fatma Abdelmoulah en 2009 [1] et de la thèse de Benjamin Perret en 2010 [204]. Il a fait l'objet de quatre publications [47, 206, 205, 207]. Dans ce chapitre, nous présentons l'approche développée lors de la thèse de Benjamin Perret.

4.1 Problématique

Les galaxies sont habituellement classées en suivant la séquence de Hubble [107] ou l'une de ses alternatives [60, 253] (figure 4.1). Or, ces deux séquences ont été établies en utilisant uniquement

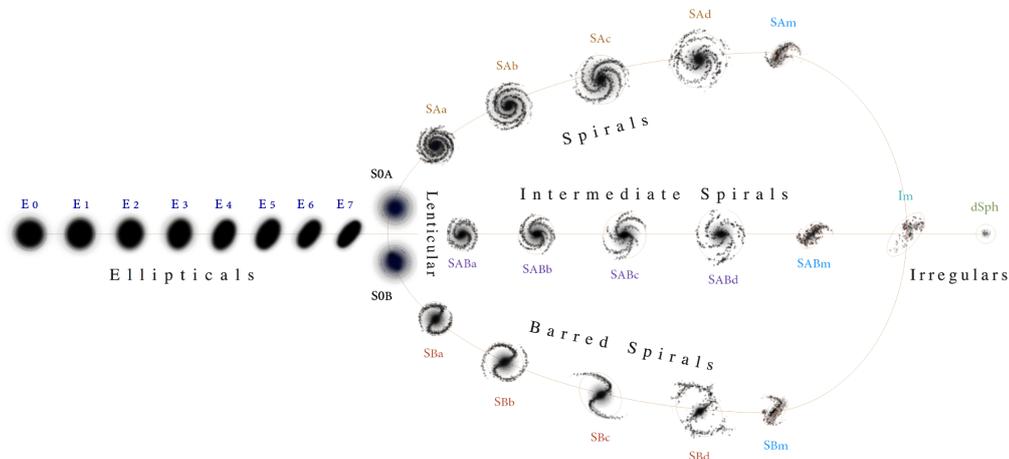


FIGURE 4.1 – Classification de Vaucouleurs (d'après Ciccolella et De Leo, *Hubble–de Vaucouleurs Galaxy Morphology Diagram*, Wikipedia).

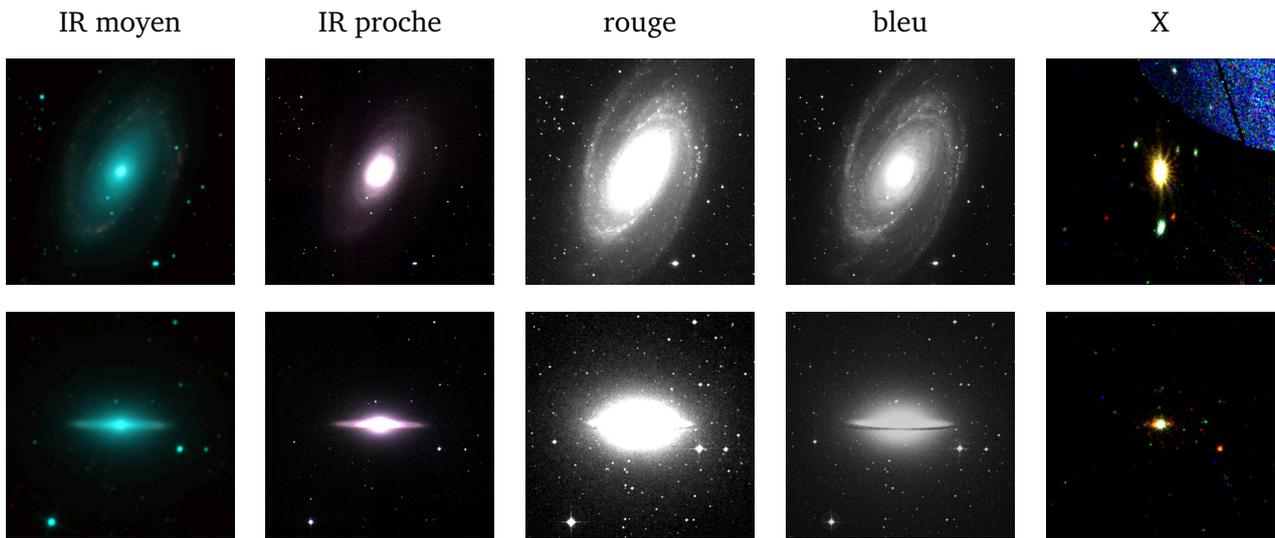


FIGURE 4.2 – Aspect des galaxies M81 (en haut) et M104 (en bas) observées à différentes longueurs d’onde.

les longueurs d’onde du visible, alors que les galaxies peuvent être d’aspect très différent dans d’autres longueurs d’ondes comme l’illustre la figure 4.2. Nous avons donc exploré la possibilité de modéliser les galaxies avec l’originalité de bénéficier d’observations multibandes couvrant un maximum de longueurs d’ondes, l’objectif étant d’aboutir à une classification fondée sur plus de données.

Une galaxie est composée de plusieurs structures physiques : un bulbe, un disque, des bras spiraux, une barre stellaire, des anneaux, et dans le cas d’une observation sur la tranche, d’une barre de poussière (toutes ces structures n’existent pas forcément dans chaque galaxie). L’observation est donc la somme des intensités de ces structures, qui correspondent, par rapport à l’équation (4.1), à des motifs φ tous différents. Par ailleurs, l’observation est entachée par la réponse impulsionnelle h de l’instrumentation. Le modèle est détaillé section 4.2.

Ce travail a permis d’apporter les contributions suivantes.

- Au moment de la thèse de Benjamin Perret, la plupart des méthodes de l’état de l’art fournissaient une décomposition en structures mathématiques et non physiques ; citons par exemple l’utilisation de shapelets [120, 145] ou la classification à partir de caractéristiques géométriques de l’image [245]. Il n’existait que trois algorithmes disponibles pour réaliser des décompositions physiques : Gim2D [235], Budda [59] et Galfit [201], mais ces algorithmes n’étaient pas capables de gérer les galaxies spirales ou ne fonctionnaient qu’avec des images monobandes. L’équation (4.1) a l’avantage de pouvoir effectuer une modélisation physique fine de la galaxie afin de fournir aux astronomes un résultat qu’ils puissent interpréter. Par ailleurs, la méthode proposée est capable de traiter plusieurs bandes (de l’infrarouge à l’ultraviolet) plutôt que les seules bandes de la lumière visible.
- Les structures choisies pour modéliser les galaxies sont des extensions au cas multibande de modèles existants. Le modèle des bras spiraux est quant à lui totalement original et permet de représenter relativement bien les observations réelles. Ce modèle est détaillé en section 4.2.
- La présence de zones très lumineuses (appelées régions HII) dans les bras de la galaxie peut perturber la bonne estimation des bras. Plutôt que de complexifier le modèle des bras, nous avons préféré le compléter par un filtrage des observations, les paramètres du filtre faisant partie des paramètres du modèle. Ce filtre n’est pas un pré-traitement car il est mis à jour en même temps que le modèle des structures, il permet par ailleurs de conserver naturellement

l'information de localisation des zones filtrées. Nous avons ainsi exploré l'utilisation conjointe des approches statistiques et de la morphologie mathématique, qui sont par ailleurs deux des thèmes traités dans notre équipe. Il existe très peu de liens entre ces deux disciplines et la littérature à ce sujet est rare (voir par exemple les travaux de Forbes et Raftery [82]). Cette contribution est détaillée section 4.2.

- L'optimisation de l'équation (4.1) n'est pas simple. C'est pour cela que nous avons choisi de travailler avec une hiérarchie de modèles : des modèles de plus en plus fins sont tour à tour estimés, en initialisant chacun d'eux avec l'estimation du modèle précédent. Cette approche permet de réduire le temps de calcul et d'améliorer la robustesse de la méthode par rapport aux conditions initiales, elle est détaillée en section 4.3.
- L'optimisation est effectuée à l'aide d'un échantillonneur de Gibbs couplé à un schéma de recuit simulé. Une contribution de ce travail est d'appliquer un schéma de température différent pour chaque bloc d'inconnues. Ce schéma est particulièrement adapté aux problèmes d'optimisation dont les variables sont de natures variées (ici, il s'agit de positions, longueurs, brillances, angles. . .). Il permet de mieux répartir l'effort de simulation en se concentrant sur les paramètres les plus difficiles à estimer. Cette contribution est détaillée section 4.3.

4.2 Modèle

La plupart des motifs que nous proposons pour modéliser les structures de la galaxie sont inspirés des modèles classiques de la littérature ; nous les avons adaptés aux données, notamment pour prendre en compte l'aspect multibande. Seul le motif modélisant les bras spiraux est complètement nouveau. Il a d'ailleurs l'originalité de ne pas être entièrement génératif : il est composé d'un modèle paramétrique traditionnel et d'un filtre paramétrique dont le rôle est de modifier les observations afin d'en supprimer certains aspects que nous ne souhaitons pas modéliser et qui pourraient perturber le processus d'estimation.

Nous résumons à présent les modèles de forme et de brillance des principales structures. Notons que les galaxies n'étant pas toujours observées de face, ces modèles doivent être projetés sur le plan d'observation pour représenter le plus fidèlement possible les conditions d'observation.

Modèle des structures autres que les bras spiraux

La forme des structures est modélisée par une ellipse généralisée [59, 201]. Une ellipse généralisée de centre (x_0, y_0) , d'angle α , d'axe majeur r , d'élongation e et de déformation c est formée par l'ensemble des points $(x, y) \in \mathbb{R}^2$ vérifiant :

$$r^{c+2} = \left| (x - x_0) \cos \alpha - (y - y_0) \sin \alpha \right|^{c+2} + e^{c+2} \left| (x - x_0) \sin \alpha + (y - y_0) \cos \alpha \right|^{c+2} \quad (4.2)$$

La figure 4.3 donne des exemples d'ellipses généralisées.

La brillance des structures est modélisée par une loi de Sérsic [243] :

$$I(r) = I_0 \exp \left[-k \left(\left(\frac{r}{R} \right)^{\frac{1}{n}} - 1 \right) \right] \quad (4.3)$$

avec I_0 la brillance au rayon effectif R , n l'indice de Sérsic (si $n = \frac{1}{2}$ on obtient une gaussienne), et k une fonction auxiliaire telle que $\Gamma(2n) = 2\gamma(2n, k)$ où Γ et γ sont respectivement la fonction Gamma et la fonction Gamma incomplète. La figure 4.4 donne des exemples de lois de Sérsic : on s'aperçoit que lorsque n devient grand, la loi devient très piquée.

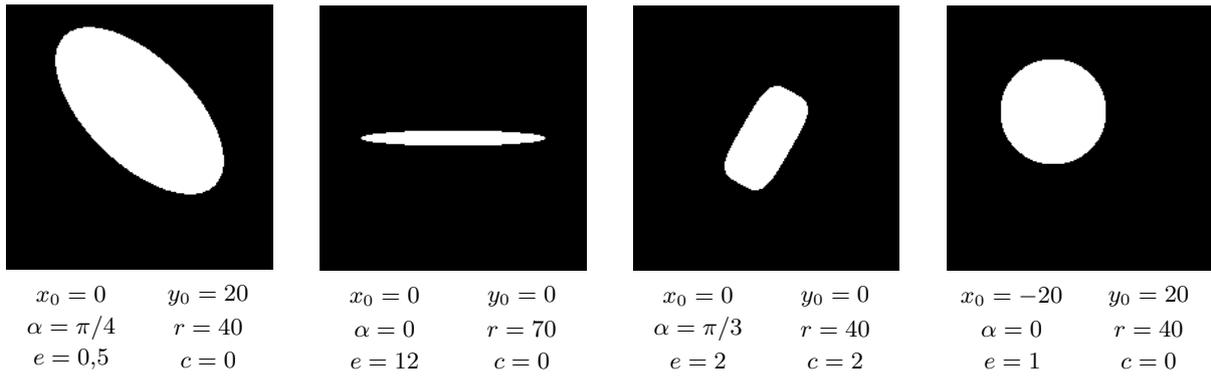


FIGURE 4.3 – Exemples de motifs dont la forme est une ellipse généralisée. L'origine est au centre des images.

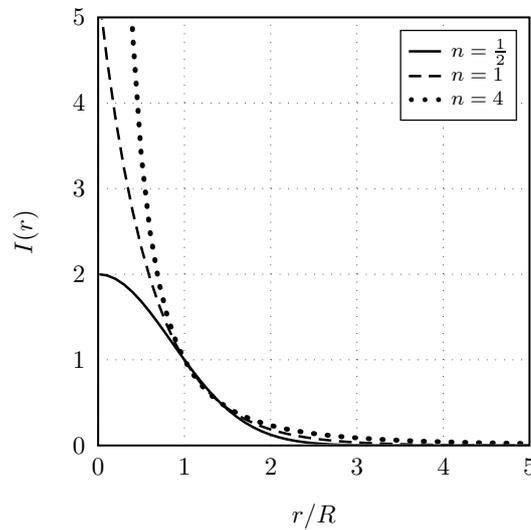


FIGURE 4.4 – Loi de Sérsic pour différentes valeurs de n .

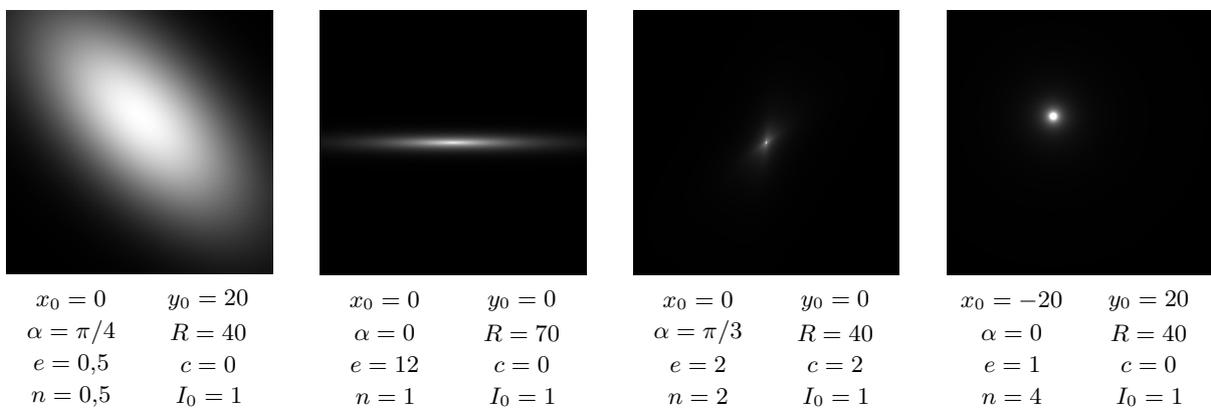


FIGURE 4.5 – Exemples de simulations obtenues en combinant l'aspect géométrique d'une ellipse généralisée avec une distribution lumineuse de Sérsic. Afin de visualiser suffisamment les structures, les intensités sont saturées.

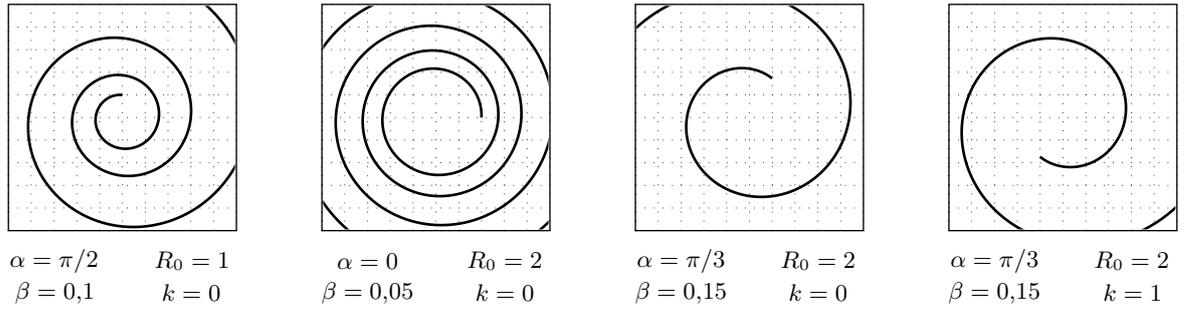


FIGURE 4.6 – Exemples de spirales logarithmiques. L'origine est au centre des graphes.

La combinaison du descripteur de forme (l'ellipse généralisée) avec le descripteur de brillance (la loi de Sérsic) permet de générer des images dont le rendu est très proche des observations réelles (figure 4.5). Suivant le type de structure à modéliser, les paramètres sont contraints à être dans des intervalles bien définis (cf. [204, tableau 2.4]).

Modèle des bras spiraux

Il existe peu de modèles paramétriques des bras spiraux, et les seuls disponibles (par exemple [102, 170, 202]) ne sont pas satisfaisants. C'est pourquoi nous avons proposé un modèle original, constitué de deux parties : un modèle paramétrique de l'aspect général des deux bras et un filtre adaptatif pour isoler les régions HII qui sont très lumineuses. Les bras spiraux peuvent avoir une structure très complexe, mais la plupart du temps on observe uniquement deux bras : le modèle proposé ne permet donc de représenter que deux bras.

Modèle paramétrique La forme des bras spiraux est supposé suivre une double spirale logarithmique [9] qui regroupent les points $(\rho, \phi) \in \mathbb{R}^+ \times \mathbb{R}$ du plan :

$$\rho = R_0 \exp(|\beta|\theta) \quad \phi = \beta\theta + \alpha + k\pi \quad (4.4)$$

où $\theta \in \mathbb{R}^+$ permet de se déplacer le long de la spirale, R_0 est le rayon de départ, β est l'angle d'ouverture (dont le signe code le sens de rotation de la spirale), α est l'angle de position et $k \in \{0,1\}$ est le numéro du bras. La figure 4.6 donne quelques exemples de spirales logarithmiques, mettant notamment en évidence l'influence des différents paramètres.

Les bras d'une galaxie possèdent une certaine épaisseur et la spirale logarithmique en modélise la crête. La brillance des bras suit une loi de Sérsic qui décroît en fonction de la distance à la crête, mais également en fonction de la distance au centre de la galaxie (les bras sont de moins en moins lumineux lorsqu'on s'éloigne du centre de la galaxie). En pratique, la simulation des bras spiraux suit les trois étapes suivantes :

1. la crête des bras est représentée à l'aide de deux spirales ;
2. une transformée de distance est calculée sur tout le plan selon la distance à la crête et la distance au centre ;
3. une loi de Sérsic est appliquée à la carte de distance obtenue à l'étape précédente.

Un exemple de simulation de bras est donné figure 4.7

Filtre adaptatif Certaines galaxies font apparaître des zones très lumineuses dans les longueurs d'onde courtes : ce sont les régions HII. Il serait inutile de modéliser ces structures avec un modèle

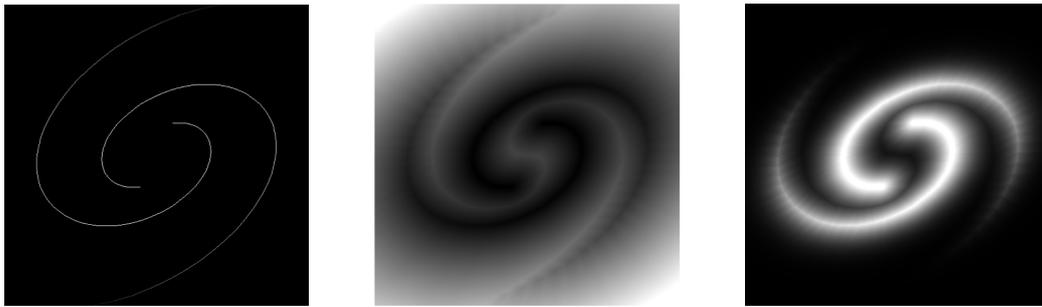


FIGURE 4.7 – Processus de génération des bras spiraux.
De gauche à droite : crête des bras, carte de distance et structure finale.

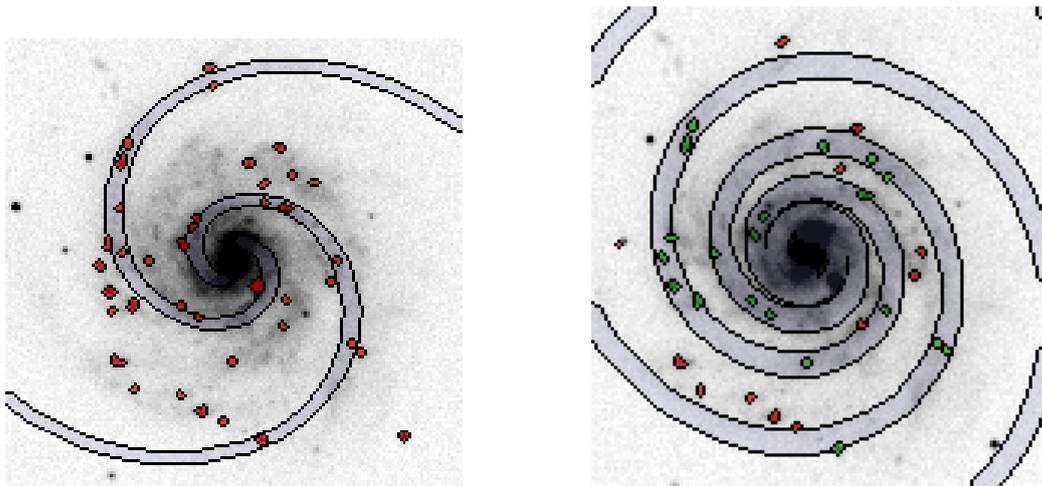


FIGURE 4.8 – Intérêt du filtrage adaptatif sur l'estimation de la spirale (galaxie PGC 35538) : sans prise en compte des régions HII (à gauche) et avec (à droite). Les zones rouges et vertes correspondent aux régions HII détectées.

paramétrique car ni la forme, ni le profil de brillance de ces régions ne contiennent d'information intéressante pour la classification des galaxies. Par ailleurs, il serait trop complexe de vouloir les modéliser. C'est pourquoi notre idée a été d'appliquer un filtrage f des observations pour isoler les régions HII. Les paramètres θ de ce filtre dépendent du modèle paramétrique, qui est lui-même estimé à partir de l'observation filtrée : le processus est donc itératif et c'est la raison pour laquelle le filtrage n'est pas qu'un simple prétraitement.

Ainsi, les régions HII identifiées sont répertoriées dans une carte qui permet de filtrer les observations y , de sorte que les observations filtrées $f(y, \theta)$ n'aient plus de régions HII qui risqueraient de déstabiliser l'estimation du modèle paramétrique. L'algorithme permettant d'obtenir cette carte fait appel aux arbres de composantes connexes et sort du contexte de ce document : le lecteur intéressé pourra donc consulter le chapitre 3 de la thèse de Benjamin Perret [204].

La figure 4.8 représente la galaxie PGC 35538 sur laquelle sont indiquées les régions HII détectées. Sans les prendre en compte, l'estimation de la spirale est très mauvaise, alors qu'elle est bien meilleure en tenant compte des régions HII détectées.

Combinaison des structures

La combinaison de plusieurs structures suivant les modèles présentés ci-avant permet donc de synthétiser des observations de galaxies (figure 4.9). Une observation multispectrale est obtenue

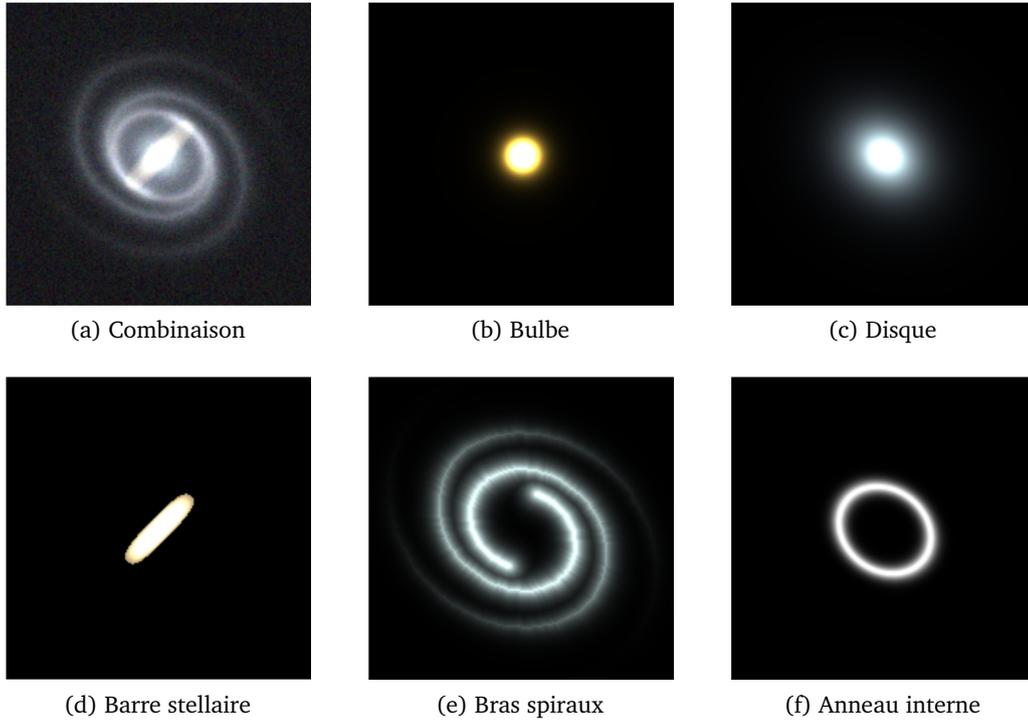


FIGURE 4.9 – Simulation d’une galaxie comprenant quatre structures ellipsoïdales et une structure spirale. Cette simulation pourrait correspondre à une galaxie de type SBbR.

bulbe	disque	halo	barre	anneau	bras
$5 + 3B$	$4 + 3B$	$4 + 3B$	$6 + 3B$	$3B$	$5 + 3B$

TABLE 4.1 – Nombre de paramètres pour chaque structure (B est le nombre de bandes).

en associant plusieurs observations monochromes où les descripteurs de forme des structures sont identiques entre les observations. Pour fixer les idées, le nombre de paramètres pour chaque structure est donné dans la table 4.1.

Modèle bayésien

En tenant compte de la modélisation précédente, de la réponse impulsionnelle et du filtrage adaptatif, l’équation d’observation (4.1) s’écrit finalement¹ :

$$f(\mathbf{y}, \boldsymbol{\theta}) = \mathbf{h} * \sum_{k=1}^K \mathbf{a}_k \varphi_k(\mathbf{c}, \mathbf{w}_k) + \mathbf{b} \quad (4.5)$$

où $\boldsymbol{\theta}$ regroupe les paramètres du modèle utilisé par le filtre adaptatif f et chaque motif φ_k modélise une structure particulière avec sa propre morphologie et son propre modèle d’intensité et a pour paramètres :

- un vecteur \mathbf{a}_k regroupant les intensités de la structure dans chaque bande et qui peut être nul si la structure n’existe pas dans la galaxie ;

1. En réalité, le modèle proposé dans le cadre de la thèse de Benjamin Perret est un peu plus complexe car il permet d’ajouter la présence d’une barre de poussière ; celle-ci intervient dans l’équation (4.5) sous la forme d’un motif additionnel qui est en facteur de la somme, permettant ainsi de décroître l’intensité des autres motifs. Nous n’avons pas détaillé cette structure particulière pour simplifier la présentation du modèle et parce que la barre de poussière n’apparaît que dans le cas de galaxies vues par la tranche (ce qui arrive assez rarement).

- le centre c de la structure qui est la même quel que soit k (les structures ont toutes le même centre, en l'occurrence le centre de la galaxie) ;
- des paramètres de forme w_k .

Le bruit b a plusieurs origines mais peut être approximé par un bruit gaussien, non corrélé et de moyenne nulle. Sa matrice de variance-covariance peut être assez facilement estimée [22] et est donc supposée connue. Les lois a priori des autres paramètres sont des lois uniformes sur les domaines respectifs de chaque variable ; il n'est pas apparu nécessaire de choisir des lois plus « informatives ». Ces domaines ont été fixés de manière à modéliser correctement les différentes galaxies existantes. Par exemple, l'indice de Sérsic n du bulbe est compris entre 1,5 et 8 alors que celui du disque est égal à 1. Par ailleurs, la déformation c du bulbe et du disque est nulle (la structure est une pure ellipse) tandis que celle de la barre stellaire est comprise entre 0,5 et 2 (la barre est presque rectangulaire). La liste exhaustive des contraintes est donnée dans la table 2.4 de la thèse de Benjamin Perret [204, p. 56].

4.3 Optimisation

Le modèle proposé pour décrire les données est complexe car il est constitué de plusieurs structures différentes, chacune définie à partir de plusieurs paramètres contraints, et à cause de la non linéarité de la plupart d'entre eux la distribution a posteriori n'a pas une expression simple.

Ces raisons nous ont donc conduit à utiliser un échantillonneur de Gibbs [89] dans lequel les blocs de variables sont simulés soit directement (pour les brillances dont la loi a posteriori conditionnelle est une gaussienne tronquée), soit avec un algorithme de Metropolis-Hastings (pour tous les autres paramètres dont la loi a posteriori conditionnelle n'est pas simulable directement). Ce dernier est, plus exactement, un algorithme de Metropolis-Hastings à marche aléatoire adaptatif, c'est-à-dire que la loi candidate de l'algorithme est une gaussienne centrée sur la valeur courante du paramètre à générer (marche aléatoire) et d'écart-type adapté au taux d'acceptation : si le taux d'acceptation est trop élevé, l'écart-type est augmenté et inversement.

La mesure de la convergence d'un algorithme MCMC n'est pas aisée. Il existe une littérature riche sur le sujet mais la plupart des méthodes publiées nécessitent la génération de plusieurs chaînes, ce qui est une solution trop coûteuse en temps de calcul pour l'application visée [33, 34, 39]. En s'inspirant de [80, 113], nous choisissons d'évaluer la convergence de la moyenne de la chaîne en construisant un intervalle de confiance valide asymptotiquement. La grande dimensionnalité de la chaîne rend néanmoins la construction d'un intervalle de confiance complexe et nous avons choisi de n'étudier que l'évolution de l'EQM en fonction des itérations. En notant σ l'écart-type de l'EQM sur les S dernières itérations, la convergence est supposée atteinte si :

$$t \frac{\sigma}{\sqrt{S}} < \varepsilon \quad (4.6)$$

où t est le quantile de Student et ε est la demi-largeur de l'intervalle de confiance souhaité. En pratique, nous avons fixé $t = 0,9$, $\varepsilon = 0,05$ et $S = 25$.

Par ailleurs, plusieurs techniques détaillées ci-après permettent de réduire le nombre d'itérations nécessaires pour obtenir une estimation satisfaisante :

- une hiérarchie de modèles pour estimer des modèles de précision croissante ;
- le regroupement au sein d'un même bloc de l'échantillonneur de Gibbs de paramètres corrélés ;
- un schéma de recuit simulé pour faciliter la convergence vers le maximum global ;

- une décroissance de température dans chaque bloc adaptée à la complexité des lois conditionnelles correspondantes.

Hiérarchie de modèles

Une hiérarchie de modèles consiste à décrire les observations par une succession de modèles de complexité croissante. L'estimation obtenue pour chaque modèle sert d'initialisation au modèle de complexité supérieure. Deux raisons justifient l'intérêt de cette procédure : d'une part, un modèle grossier utilise peu de paramètres et est donc plus rapide qu'un modèle précis ; d'autre part, l'initialisation de chaque modèle étant basée sur l'estimation du modèle antérieur, on peut la considérer comme étant de bonne qualité ce qui a pour effet de réduire le nombre d'itérations. L'idée d'effectuer un traitement grossier pour l'affiner peu à peu se retrouve également dans les représentations multi-échelle, dans lesquelles on commence par effectuer un traitement sur une image très fortement sous-échantillonnée pour terminer par l'image initiale, en passant par plusieurs échelles intermédiaires.

La hiérarchie des modèles est définie par l'importance relative des structures en termes de flux lumineux. Par exemple, il est connu que les bras spiraux ont une influence relativement faible sur les paramètres du bulbe et du disque [170] alors que la présence d'une barre stellaire peut avoir un impact important sur les paramètres du bulbe [84] : on choisira donc d'introduire la barre avant les bras spiraux. Il se peut néanmoins que, pour une galaxie particulière, l'ordre des modèles ne soit pas idéal.

Ainsi, la première étape consiste à ajuster un bulbe et un disque sur l'image blanche (moyenne des bandes). En fonction des intensités relatives de ces deux motifs, on peut décider si la galaxie n'a besoin que de ces deux motifs (dans ce cas, la décomposition est terminée) ou s'il faut en rajouter. La présence d'une barre stellaire, de bras spiraux ou d'un anneau est ensuite testée. Ces tests sont basés par exemple sur l'élongation du bulbe, une analyse de Fourier ou la couleur bleutée du disque. Une dernière estimation est donc effectuée sur les données multispectrales en tenant compte d'un bulbe, d'un disque et des structures dont la présence est validée.

Regroupement des paramètres corrélés par blocs

Le principe de l'échantillonneur de Gibbs est de structurer l'ensemble des inconnues du problème en plusieurs blocs de variables, pour chacun desquels la loi a posteriori conditionnelle est simulée. Ces blocs peuvent être scalaires (chaque variable est simulée seule) ou vectoriels. Afin d'améliorer l'exploration de la loi cible, les paramètres fortement corrélés entre eux sont regroupés [222]. C'est le cas par exemple des paramètres de la fonction de Sérsic qui sont donc échantillonnés conjointement. Au contraire, les paramètres de forme étant faiblement corrélés, ils sont échantillonnés indépendamment.

Schéma de recuit simulé

L'échantillonneur de Gibbs est couplé à un schéma de recuit simulé (cette technique d'optimisation a été proposée indépendamment par Kirkpatrick, Gelatt et Vecchi [123] puis par Černý [256]). Le principe de ce schéma d'optimisation empirique est de commencer par simuler une loi (à peu près) uniforme afin d'explorer facilement tout l'espace des solutions, puis de déformer cette loi pour qu'elle se rapproche petit à petit de la loi a posteriori $p(\theta|\mathbf{y})$. En pratique, la loi échantillonnée est

$p(\theta|\mathbf{y})^{1/T}$ où T est appelée « température » et décroît d'une valeur très importante jusqu'à 1. On peut même exagérer la déformation en déformant la loi a posteriori vers une distribution de Dirac centrée sur le maximum de la loi a posteriori (ce qui revient à $T = 0$).

Il est prouvé que la température doit suivre un schéma de décroissance logarithmique pour assurer la convergence presque sûre vers l'optimum [87, 99]. En d'autres termes, la température $T^{(i)}$ à l'itération i doit être supérieure ou égale à $c/\log(1+i)$. Les frères Geman [87] ont donné une condition suffisante sur c , puis Hajek [99] a précisé deux ans plus tard ce choix en donnant une condition nécessaire et suffisante. Cependant, une décroissance logarithmique est en pratique trop longue, et nous choisissons (comme la grande majorité des utilisateurs) une décroissance géométrique qui consiste à réduire la température d'une petite fraction (la température est multipliée par un facteur généralement compris entre 0,85 et 0,99).

Décroissance de température par palier propre à chaque bloc

Le schéma de décroissance de température adopté est une décroissance adaptative par paliers [105, 130] : la température est maintenue constante pendant un certain nombre d'itérations (le palier), à l'issue desquelles la moyenne des probabilités des échantillons générés est calculée. Si cette moyenne a augmenté par rapport au palier précédent, c'est que le système est encore en train de converger vers le maximum : la température est donc conservée. Au contraire, si la moyenne a diminué, il est alors temps de diminuer la température pour affiner cette convergence.

Dans un schéma de recuit simulé classique, tous les blocs échantillonnés ont la même température. Dans notre cas, certaines variables (telles que le centre de la galaxie) sont suffisamment simples à simuler et peuvent donc converger rapidement. D'autres variables, au contraire, sont très corrélées ou ont une loi a posteriori conditionnelle complexe. C'est pourquoi nous avons fait le choix de laisser chaque bloc avoir sa propre décroissance de température ; la rapidité de décroissance étant alors inversement proportionnelle à la complexité de la loi a posteriori conditionnelle.

Les variables qui ont réussi à converger rapidement, et donc pour lesquelles la température du bloc correspondant a rapidement atteint la valeur 1, pourront alors ne plus être aussi souvent échantillonnées que les autres, d'où un gain de temps. Nous pensons que cette méthode est équivalente à celle consistant à échantillonner certaines variables moins souvent que d'autres [133].

4.4 Estimation

Pour chaque paramètre θ du modèle, une estimation au sens du MAP (maximum a posteriori) est calculée sur J itérations en sélectionnant l'échantillon qui maximise la loi a posteriori du modèle². On peut également obtenir une « barre d'erreur », indiquant la confiance que l'on peut avoir en l'estimation, en calculant l'estimateur OBM (*overlapped batch means*) [5, 165] pour chaque paramètre θ :

$$\hat{V}_\theta = \frac{JW}{(J-W)(J-W+1)} \sum_{i=W}^J (\bar{\theta}_i - \bar{\theta})^2 \quad (4.7)$$

2. D'autres estimateurs auraient pu être utilisés sur certains paramètres, tel que l'estimateur de l'espérance a posteriori. Cependant, l'estimateur du MAP est approprié pour des paramètres tels que le sens de rotation de bras spiraux, qui peuvent prendre des valeurs négatives ou positives.

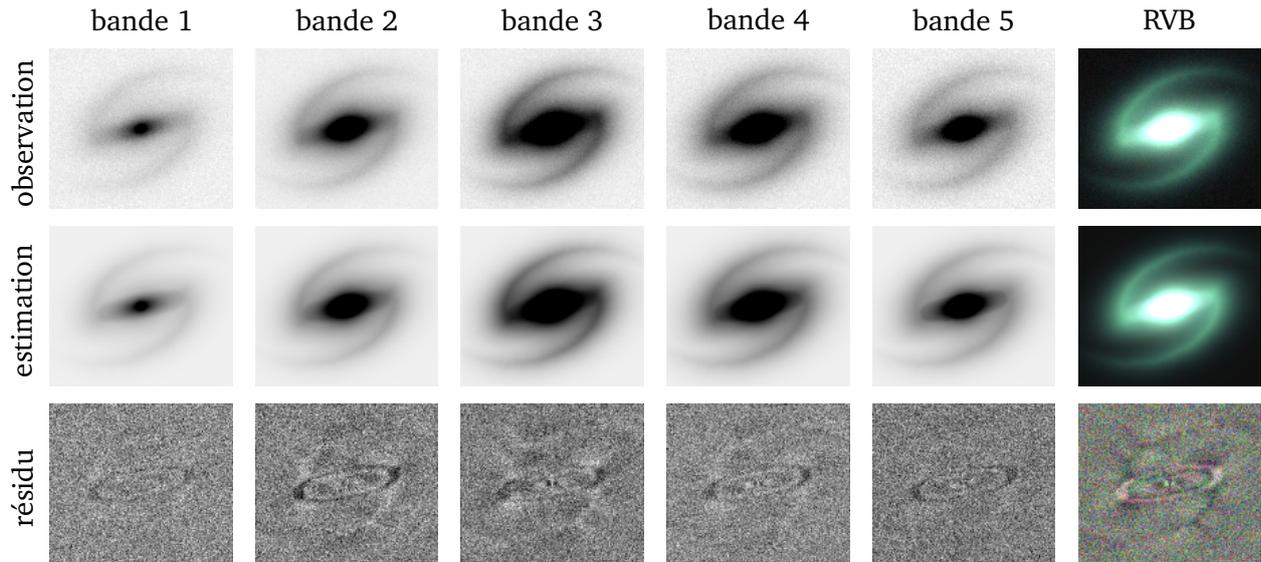


FIGURE 4.10 – Décomposition d’une galaxie spirale barrée synthétique. La dernière colonne (RVB) est une composition colorée à partir des trois bandes du visible. Les images sont affichées avec la même dynamique, excepté pour les résidus.

où $\bar{\theta}$ est la moyenne des échantillons générés θ_i et $\bar{\theta}_i$ est la moyenne des échantillons calculée sur une fenêtre glissante de W échantillons :

$$\bar{\theta} = \frac{1}{J} \sum_{k=1}^J \theta_k \quad \text{et} \quad \bar{\theta}_i = \frac{1}{W} \sum_{k=i-W+1}^i \theta_k. \quad (4.8)$$

Cet estimateur est plus robuste que l’estimateur classique lorsque la corrélation entre les échantillons est élevée et par ailleurs il possède de bonnes propriétés de convergence et une variance faible [5].

Or, le recuit simulé ne garantit pas l’homogénéité de la chaîne de Markov générée, et donc les estimateurs précédents ne peuvent pas être correctement déterminés. C’est pourquoi la température de tous les blocs est fixée à $T = 1$ sur quelques dizaines d’itérations afin de générer les J échantillons finaux qui permettront de calculer les estimateurs.

4.5 Résultats

Résultats sur données simulées

La figure 4.10 présente les résultats de la décomposition d’une galaxie synthétique à partir d’une image à cinq bandes. Visuellement, l’estimation est très correcte même s’il reste une structure ellipsoïdale dans le résidu. Des tests sur plusieurs galaxies simulées ont montré que les paramètres de forme tels que le centre, l’angle ou l’élongation sont estimés avec une très bonne précision, alors que les paramètres de la fonction de Sérsic sont moins bien estimés. Ces imprécisions proviennent majoritairement du fait que l’estimation porte sur un très faible nombre de pixels puisque le bulbe ou la barre stellaire sont de petites structures.

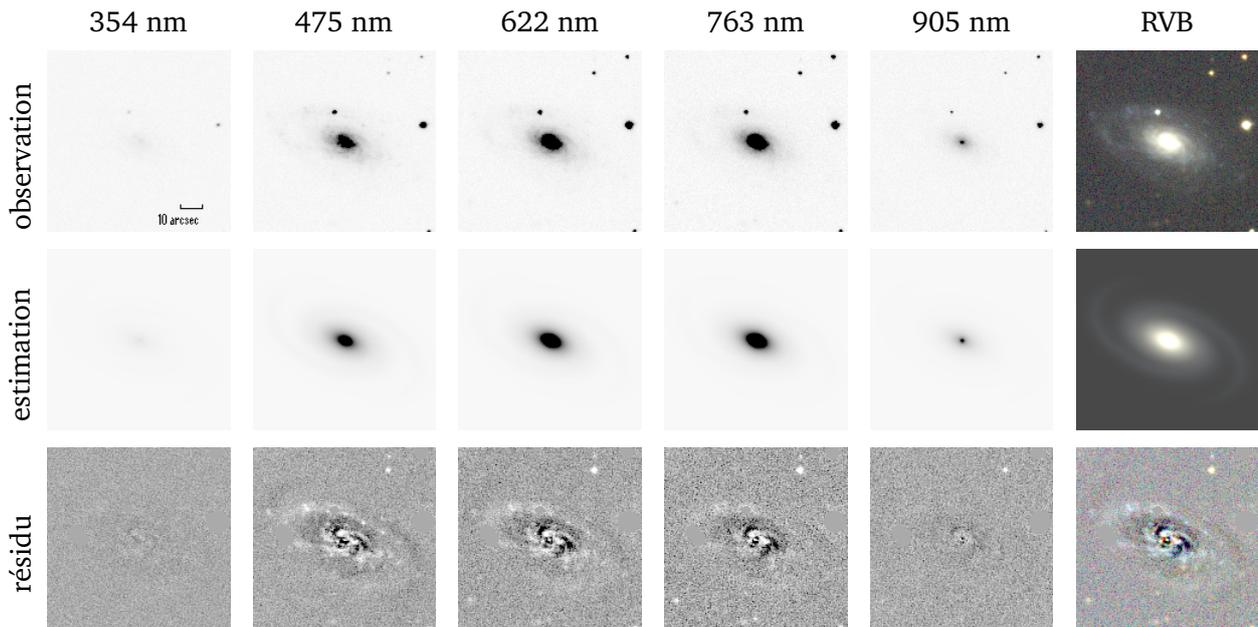


FIGURE 4.11 – Décomposition de la galaxie PGC 23443. La dernière colonne (RVB) est une composition colorée à partir des trois bandes du visible. Les images sont affichées avec la même dynamique, excepté pour les résidus.

Résultats sur données réelles

La méthode proposée a été appliquée sur près de 1500 galaxies issues de la base CFIGI [15]. Ces galaxies ont été observées par le télescope du *Sloan Digital Sky Survey* dans cinq bandes, de l'infrarouge à l'ultraviolet. Une interpolation a permis d'obtenir des images de taille 256×256 pixels. Seules les galaxies de résolution et de rapport signal-à-bruit suffisants ont été sélectionnées. Elles ont par ailleurs été annotées par des experts afin de disposer de leur classification suivant la séquence de Hubble. En moyenne, une image est traitée en deux bonnes heures avec une implémentation Java.

À titre d'illustration, la galaxie PGC 23443 et sa décomposition sont représentées figure 4.11.

La méthode proposée a été comparée avec le logiciel Galfit [201]. Même s'il ne fonctionne qu'avec des images monobandes, il utilise à peu près le même modèle que notre méthode. L'algorithme d'optimisation est déterministe et l'initialisation, manuelle, requiert une certaine précision. Les tests ont porté sur la décomposition bulbe–disque de cinq galaxies et les résultats obtenus sont donnés table 4.2. Les deux méthodes donnent des estimations très proches, ce qui conforte notre méthode.

Galaxie	EQM		Magnitude du disque		Magnitude du bulbe	
	Nous	Galfit	Nous	Galfit	Nous	Galfit
PGC 2182	2,53	2,69	12,6	12,55	15,65	15,35
PGC 2600	122,64	122,82	11,69	11,60	15,17	15,20
PGC 3563	8,48	9,65	13,13	13,13	12,24	12,24
PGC 3584	7,25	7,27	12,18	12,18	14,73	14,74
PGC 5055	1,56	1,59	14,92	14,95	13,24	13,24

TABLE 4.2 – Comparaison entre notre méthode et Galfit. L'EQM est définie [204, équation (2.49)].

Nous avons également procédé à une évaluation de la méthode proposée sur près de 1500 galaxies :

les estimations des paramètres du modèle fournies par notre méthode ont été comparées avec les attributs donnés par les experts. Or, les estimations sont quantitatives (par exemple : flux, taille de la barre, ouverture des bras. . .) alors que les attributs donnés par les experts sont qualitatifs (par exemple : « petit », « moyen », « long ». . .). Aussi, nous avons établi, en collaboration avec notre expert (Éric Slezak), le lien entre les attributs qualitatifs et les valeurs numériques de notre modèle. Par exemple, un barre stellaire « petite » correspond à une taille comprise entre 10 et 20. Il en ressort que la méthode proposée produit globalement de bonnes estimations des caractéristiques étudiées. Les difficultés rencontrées sur certaines caractéristiques sont principalement liées au fait que le filtrage adaptatif des objets n'est pas assez puissant et que le modèle de spirale logarithmique est trop contraint pour les bras spiraux (qui peuvent être plus que deux, non symétriques, etc.).

4.6 Conclusion

Ce chapitre traite de la décomposition conjointe de plusieurs images. À la différence du chapitre précédent, les signaux et les motifs sont bidimensionnels, mais il n'y a pas d'évolution douce des paramètres des motifs (les paramètres sont soit constants parmi les différents signaux, soit indépendants). Un modèle bayésien et un schéma hiérarchique d'optimisation a été proposé. Ce schéma combine un algorithme MCMC et un filtre adaptatif utilisant un arbre de composantes connexes. Des dispositifs ont été mis en place pour accélérer le temps de calcul.

Le fait probablement le plus intéressant dans ce travail a été le couplage entre un modèle bayésien paramétrique et le filtre adaptatif construit à partir de techniques de morphologie mathématique. Cela revient à une modélisation du problème reposant notamment sur l'usage d'un opérateur non linéaire. Cet opérateur ne s'exprime pas mathématiquement (sous forme d'équations), mais informatiquement (sous forme d'algorithme). On peut imaginer que le couplage de ces outils issus de disciplines différentes peut avoir un grand intérêt dans certaines applications.

Chapitre 5

Bilan et perspectives

5.1 Bilan général

Mes recherches ont traité d'aspects méthodologiques (comme par exemple la prise en compte de l'évolution des motifs dans la décomposition conjointe) et d'applications diverses (en astronomie, en physique, en télédétection). Pour mener à bien ces recherches, je me suis appuyé sur l'expertise de plusieurs collègues de différents laboratoires et de différentes disciplines. J'ai participé à divers projets et j'en ai également porté deux qui m'ont permis, avec mes collègues et notamment ceux du thème IPSEO, d'explorer nos questions de recherche. Depuis fin 2018, je suis également co-porteur d'un projet industriel avec une entreprise locale, dont le but est la détection sans microscope de micro-organismes dans des images biologiques. Les projets SPECTRODEC et DSIM, dont j'étais porteur, ont regroupé une douzaine de chercheurs et ont financé deux thèses et plusieurs stages de master.

Ce mémoire d'habilitation s'est concentré sur la décomposition de signaux ou d'images en motifs paramétriques dont les paramètres évoluent doucement à travers les données. Les principales contributions sont détaillées ci-dessous.

- Nous avons montré qu'une décomposition conjointe est meilleure qu'une décomposition séquentielle. C'est un résultat qui est était attendu et qui s'explique par le fait que l'aspect joint aide la décomposition en chaque site à profiter de la décomposition des voisins. Elle favorise également l'obtention d'un résultat ayant un comportement physiquement significatif.
- Nous avons proposé un modèle bayésien original adapté au problème, notamment à travers l'introduction de deux variables de dimensions que sont le nombre de motifs et le nombre de groupes. Ce modèle est optimisé à l'aide d'un algorithme RJMCMC ayant des mouvements adaptés.
- Nous avons développé une méthode fondée sur les méthodes d'approximation parcimonieuse pour effectuer la décomposition conjointe de signaux. Cette méthode s'apparente aux méthodes de parcimonie structurée, mais utilise une structure plus souple que ce qui est proposé dans la littérature. Elle est par ailleurs plus simple et plus rapide que l'approche bayésienne et fournit des résultats comparables, même si elle ne dispose pas des mêmes garanties de convergence.
- Nous avons combiné une approche bayésienne de décomposition avec une approche issue de la morphologie mathématique qui permet d'identifier des zones particulière d'une image. Cette combinaison n'est pas qu'une suite de traitements : les deux approches fonctionnent de pair, chacune utilisant le résultat de l'autre pour améliorer l'estimation.

5.2 Perspectives

La représentation parcimonieuse de signaux est un sujet très ancien mais qui reste d'actualité. Ainsi, la déconvolution impulsionnelle, après avoir été un sujet de recherche riche dans les années 1980 et 1990, est à nouveau considérée à l'heure actuelle dans le cadre des représentations parcimonieuses. Elle peut être traitée dans un cadre bayésien [14, 27, 26, 32] ou en modélisant le vecteur impulsionnel par une somme d'impulsions de Dirac [70, 71, 163]. Certains auteurs établissent également un lien entre les méthodes d'approximation parcimonieuse et la modélisation bayésienne avec un a priori Bernoulli-gaussien [32, 68, 69, 104].

Après avoir traité le problème de la déconvolution impulsionnelle durant ma thèse, je me suis intéressé au problème plus général de décomposition en motifs. Pour cela, je me suis tourné vers les représentations parcimonieuses et les algorithmes associés. La plupart des algorithmes s'avèrent rapides, simples et possédant de bonnes propriétés de convergence. Mais, ils ne sont pas optimisés pour traiter certaines situations, comme par exemple l'utilisation de dictionnaires paramétriques ou la décomposition simultanée de plusieurs signaux¹.

Mon projet de recherche regroupe trois axes qui s'attachent à l'étude et à l'adaptation de ces algorithmes à de nouvelles situations. Dans la mesure du possible, les développements devront suivre les spécificités du traitement du signal, à savoir des méthodes ayant de bonnes garanties mathématiques, des algorithmes rapides et peu consommateurs de ressources, et des résultats qui soient interprétables physiquement [79].

Comment utiliser les représentations parcimonieuses pour la sélection de modèle ?

En traitement du signal et des images, la sélection de modèle regroupe des problèmes pour lesquels le nombre de composantes ou la dimension de la variable d'intérêt sont inconnus et doivent être estimés. En d'autres termes, le nombre d'inconnues est lui-même inconnu. On peut citer comme exemples la décomposition d'un signal en motifs, mais également la restauration de signaux impulsionnels [97], l'estimation de signaux continus par morceaux [132], la détection d'objets [164, 196], la séparation de sources [229], l'internet des objets [139], etc.

Une première façon de résoudre ce problème est de fixer la dimension du problème à une valeur connue M : c'est le cas du modèle Bernoulli-gaussien et des représentations parcimonieuses. Le signal est modélisé par la somme de motifs potentiellement présents dans les données. L'inconnue du problème est un vecteur regroupant les M amplitudes des motifs potentiels ; les motifs absents des données étant d'amplitude nulle. L'intérêt de ce modèle est qu'il est à dimension constante, ce qui permet l'utilisation de méthodes simples, telles que les algorithmes d'approximation parcimonieuse (OMP, BP...), l'échantillonneur de Gibbs [44] ou d'autres algorithmes spécifiques (comme SMLR [126]). Mais ce modèle souffre de deux inconvénients principaux. D'une part, les motifs sont contraints à rester sur la grille d'échantillonnage du vecteur, et d'autre part la taille de ce vecteur peut être très grande, augmentant alors le nombre de variables à estimer. Des approches bayésiennes non paramétriques ont récemment été introduites [57, 88, 96, 179, 195]. Leur optimisation est effectuée à l'aide d'algorithmes MCMC ou d'approches variationnelles.

Une deuxième façon de résoudre le problème est de travailler à dimension variable avec des algorithmes capables d'estimer le nombre d'inconnues, et donc d'explorer un espace de dimension inconnue. Le signal est modélisé par la somme des seuls motifs présents dans les données : le vecteur

1. Il est question ici des cas où ces signaux ont une décomposition proche mais différente ; le cas où les signaux utilisent les mêmes composantes a déjà été l'objet de nombreuses études [11, 127, 233, 251, 255].

des amplitudes ne contient plus de valeurs nulles mais il est de taille K inconnue. En plus des amplitudes, il devient indispensable d'estimer les paramètres des motifs (par exemple, leur position). Ainsi, on peut utiliser des algorithmes à dimension constante en testant toutes les tailles possibles, puis utiliser un critère comme AIC (*Akaike information criterion* [4]), BIC (*Bayesian information criterion* [230]) ou MDL (*Minimum description length* [220]) pour identifier la solution. Seulement, cette méthode est très gourmande en temps de calcul car il faut tester toutes les dimensions possibles de l'espace. De plus, un grand nombre de données sont générées mais inutilisées pour l'estimation. Les algorithmes de Monte Carlo par chaînes de Markov transdimensionnels tels que RJMCMC (*Reversible jump Monte Carlo Markov chain*) [94] sont capables de sauter d'un espace de solution à un autre, et évitent ainsi d'avoir à exécuter plusieurs fois l'algorithme. Mais ces algorithmes sont très lents en temps de calcul s'ils ne sont pas parallélisés et complexes à mettre en œuvre (il faut calculer des jacobiens, définir des sauts efficaces, choisir des lois candidates appropriées...).

Je souhaite donc explorer une troisième voie, que l'on peut considérer à mi-chemin entre les deux précédentes. L'idée est de se placer à dimension constante pour bénéficier d'algorithmes simples. Cependant, le vecteur des amplitudes est de taille L beaucoup plus petite que M pour avoir moins d'inconnues à estimer. La taille de ce vecteur doit tout de même être aussi grande que le plus grand nombre de motifs attendu. On a donc : $K < L \ll M$. Afin d'avoir une bonne précision sur les motifs, il est indispensable d'estimer leurs paramètres. En résumé, les trois modèles nécessitent d'estimer respectivement M , $2K$ et $2L$ paramètres. En fonction des rapports entre ces trois variables et le temps nécessaire à l'estimation des différentes variables, il est envisageable que le modèle à explorer soit plus rapide et plus précis que l'approche à dimension constante tout en étant plus simple que l'approche à dimension variable.

Cette modélisation peut être vue comme un processus ponctuel marqué où l'une des marques est l'occurrence du motif. D'ailleurs, Stephens [240] remarquait que les motifs en nombre variable dans un problème transdimensionnel pouvaient être vus comme un processus ponctuel marqué.

J'avais déjà étudié une ébauche de ce type de modèle juste après ma thèse [150]. En 2017, Baidu et coll. [14] ont proposé un modèle dans lequel les amplitudes sont modélisées par un a priori Bernoulli-gaussien et les paramètres des motifs sont estimés hors-grille. À la différence de l'idée proposée, le nombre potentiel de motifs est égal à M et l'optimisation est effectuée à l'aide d'une approche variationnelle. Mais considérer un nombre de motifs connu et juste un peu plus grand que la vérité a déjà été abordé en statistiques pour l'estimation d'un mélange de lois [142]. L'application au traitement du signal reste donc à faire. Enfin, Carfantan, Bourguignon et Boudineau ont développé récemment un modèle de représentation parcimonieuse dans lequel les centres des motifs sont estimés (bien qu'ils soit contraints à rester dans le voisinage de la grille) [26]. L'utilisation de méthodes d'approximation parcimonieuse semble une alternative simple et rapide. Il faudrait alors se tourner vers les méthodes hors-grille pour estimer les paramètres des motifs.

Le premier risque de cette approche est que le bon fonctionnement des algorithmes parcimonieux est bien étudié et garanti lorsque la parcimonie est grande, ce qui n'est pas le cas avec le modèle proposé où le vecteur des amplitudes contient très peu de termes nuls. Le second risque est le problème de *label switching* qui peut rendre complexe l'estimation.

Ce projet est plus méthodologique que les deux suivants et aucune application spécifique n'est visée pour l'instant. C'est une recherche « amont » qui peut s'avérer fructueuse car elle peut aboutir à de nouvelles pistes de recherche sans être contrainte par des objectifs applicatifs trop spécifiques.

Comment estimer un dictionnaire paramétré ?

Dans beaucoup d'applications, les lois de la physique et les connaissances a priori autorisent une modélisation paramétrique des atomes du dictionnaire D . Par exemple, les raies d'émission de spectres électromagnétiques sont généralement modélisées par des gaussiennes [117], les galaxies lointaines observées par un télescope peuvent être vues comme des ellipses [164], et les objets observés sur une image de télédétection peuvent être représentés par des formes géométriques [129, 196]. Ainsi, les M atomes d'un dictionnaire D s'expriment par une fonction paramétrique φ :

$$D = \left(\varphi(\theta_1) \quad \dots \quad \varphi(\theta_k) \quad \dots \quad \varphi(\theta_M) \right) \quad (5.1)$$

où θ_k regroupe les paramètres associés à l'atome k . Le cas de dictionnaires paramétrés est très peu étudié dans la littérature. En général, le dictionnaire est soit parfaitement connu (c'est par exemple un ensemble de spectres mesurés en laboratoire) soit presque totalement inconnu (on fait alors appel à des méthodes d'apprentissage de dictionnaire [74]). Mais si on a la chance de connaître la fonction paramétrique φ , comment estimer conjointement ses paramètres et l'estimation du vecteur parcimonieux x ?

La solution la plus simple est d'utiliser un très grand nombre de valeurs de θ pour construire le dictionnaire. Dit autrement, l'espace dans lequel se trouve θ est échantillonné très finement, et chaque point de cet espace correspond à une colonne de la matrice D : il en résulte un dictionnaire de très grande taille. Le dictionnaire étant paramétrique, il n'est pas nécessaire de le stocker en mémoire puisqu'il suffit de générer les composantes à chaque fois qu'on en a besoin. Mais d'autres complications apparaissent :

- le temps de calcul augmente avec le nombre de composantes (car il faut tester toutes ces composantes ou manipuler algébriquement cette grande matrice),
- les valeurs des paramètres sont contraintes à la grille d'échantillonnage (donc pour augmenter la précision de l'estimation, il faut augmenter le nombre de composantes à tester),
- le conditionnement du dictionnaire diminue avec la finesse de l'échantillonnage (ce qui peut limiter le bon fonctionnement de certains algorithmes).

Aussi, la communauté scientifique s'est récemment intéressée aux méthodes dites continues (*continuous sparse modelling*, *continuous dictionary*) ou hors grille (*off-the-grid*). En particulier, les approches hors grille peuvent s'appuyer sur les récentes études en super-résolution [36, 77, 257].

L'approche que nous avons proposé durant la thèse de Hassan Mortada [173] est d'alterner l'optimisation en x avec l'optimisation en θ . La mise au point d'un tel schéma et son étude de convergence méritent toutefois d'être plus explorées que nous l'avons fait.

Les premiers travaux dans la littérature ont considéré le cas particulier où les seuls paramètres à estimer sont les décalages des atomes. Ces méthodes reposent sur l'interpolation de la grille d'échantillonnage autour des composantes détectées dans un premier temps à l'aide d'approches ℓ_2 - ℓ_0 [83] ou ℓ_2 - ℓ_1 [73, 124]. Il serait intéressant d'étudier comment ces méthodes peuvent s'étendre au cas général où les paramètres du dictionnaire ne représentent pas seulement les positions des composantes. Cela dit, de premiers tests effectués durant la thèse de Hassan Mortada ont semblé montrer que l'utilisation d'un dictionnaire très finement échantillonné était en fait plus rapide (et plus simple) que les techniques d'interpolation, pour la même qualité d'estimation. La question qui se pose alors est : quelles sont les conditions dans lesquelles un algorithme d'interpolation sera plus efficace ou plus rapide que l'utilisation d'un dictionnaire finement échantillonné ?

Plus tard ont été introduits des modèles « perturbés » [46, 81, 244]. L'idée est de remplacer le dictionnaire par la matrice $D + \Delta$ où Δ est une déviation à estimer. Les algorithmes propo-

sés minimisent un critère $\ell_2\text{-}\ell_0$ en estimant conjointement le vecteur parcimonieux x et cette déviation Δ .

Enfin, le problème a également été considéré comme un problème de déconvolution impulsionnelle (*sparse spikes deconvolution*) : le produit Dx est remplacé par la convolution $\Phi\mu$ où μ est une mesure associée à une somme pondérée de masses de Dirac [70, 71, 163]. Duval et Peyré [70, 71] aboutissent notamment à un problème d'optimisation convexe contraint qui est optimisé grâce à l'algorithme de Frank–Wolfe.

Les futures recherches devront donc étudier comment ces approches peuvent gérer les dictionnaires dont les paramètres sont de différents types (pas seulement des décalages). Les applications naturelles concernent la décomposition de spectres (X, gamma, RMN, de photoélectrons. . .), qui sont par nature des signaux impulsionnels, et dont les formes et positions exactes des raies peuvent être connues a priori mais sans grande précision. L'utilisation de dictionnaires dont les raies s'adapteraient avec l'optimisation aurait alors un grand intérêt.

Comment intégrer une connaissance a priori sur l'évolution des motifs ?

Dans le cadre du problème de décomposition conjointe d'une séquence de signaux, nous avons pris en compte l'hypothèse d'évolution douce des motifs grâce à une régularisation de Tikhonov, autant dans le cas bayésien qu'avec une représentation parcimonieuse².

Dans le contexte des représentations parcimonieuses, les hypothèses d'évolution qu'autorisent les méthodes de la littérature sont plus contraintes : les paramètres ne peuvent tout simplement pas évoluer [251] ou alors leur évolution est imposée [127, 233].

Mais pour certaines applications, l'évolution des composantes mérite d'être modélisée plus finement. Ainsi, dans le cas de la cinématique d'un gaz dans le milieu interstellaire ou dans une galaxie, on peut introduire une information a priori sur la façon dont évolue ce gaz dans les observations. En outre, la réponse instrumentale de l'instrument peut nécessiter une certaine régularisation spatiale sur les paramètres, celle-ci pouvant être isotrope (dans la cas d'une convolution) ou non (si la réponse instrumentale est variable dans le champ). En spectroscopie, on peut prendre en compte l'échantillonnage irrégulier auquel sont acquis les spectres.

Ainsi, le troisième axe de mon projet de recherche correspond à proposer des régularisations qui tiennent compte de la physique du problème et les algorithmes associés. Cela aboutira à des méthodes d'approximation parcimonieuse pour des mesures multiples (*multiple measurements*) qui permettent de modéliser plus de liberté dans les évolutions des paramètres que les méthodes de l'état de l'art. On peut ainsi espérer qu'une modélisation plus fine du problème donnera des résultats plus réalistes et donc plus satisfaisants d'un point de vue physique.

Plusieurs termes sont donc à prendre dans l'optimisation : l'adéquation aux données bien sûr, la parcimonie de la solution, les termes de régularisation, une éventuelle non-négativité, la somme à 1, etc. La première étape est donc d'identifier avec l'aide de l'expert applicatif quelles sont les connaissances sur l'expérimentation pour définir le critère à optimiser. La difficulté majeure vient ensuite, lors de la définition de l'algorithme d'optimisation adapté à ce critère. L'idée est de partir d'un algorithme existant et de le modifier en fonction des termes de régularisation. Cependant, il faut veiller à garantir de bonnes propriétés de convergences. Le critère pourra aussi être défini en parallèle de l'algorithme, afin d'assurer une bonne adéquation entre les aspects modèle et

2. Dans le cas bayésien, il s'agissait plus exactement d'un champ de Markov gaussien appliqué sur la dérivée du vecteur des paramètres, ce qui s'interprète comme une régularisation de Tikhonov.

optimisation.

Dans le cadre de ce projet de recherche, j'envisage de collaborer avec François Orioux (L2S) et Marc-Antoine Miville-Deschenes (Institut d'astrophysique spatiale, CEA) qui s'intéressent à la physique du milieu interstellaire. Cet espace est rempli de gaz en mouvement : l'analyse de ces mouvements se fait donc, comme pour la cinématique des galaxies, en analysant le déplacement des raies.

Bibliographie

- [1] F. ABDELMOULAH : Nouveaux modèles pour la détection et la caractérisation d'objets en imagerie multivariée. Mémoire de master, Université de Strasbourg, 2009.
- [2] R. ACHANTA, A. SHAJI, K. SMITH, A. LUCCHI, P. FUA et S. SÜSSTRUNK : SLIC superpixels compared to state-of-the-art superpixel methods. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 34(11):2274–2282, 2012.
- [3] K.R.M. ADELIN, M. CHEN, X. BRIOTTET, S.K. PANG et N. PAPANASTASIOU : Shadow detection in very high spatial resolution aerial images : A comparative study. *ISPRS Journal of Photogrammetry and Remote Sensing*, 80:21–38, 2013.
- [4] H. AKAIKE : A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19(6):716–723, 1974.
- [5] C. ALEXOPOULOS, D. GOLDSMAN et J.R. WILSON : Overlapping batch means : Something more for nothing? *In Winter Simulation Conference*, 2011.
- [6] G.C. ALLEN et R.F. MCMEEKING : Deconvolution of spectra by least-squares fitting. *Analytica Chimica Acta*, 103:73–108, 1978.
- [7] C. ANDRIEU et A. DOUCET : Joint Bayesian model selection and estimation of noisy sinusoids via reversible jump MCMC. *IEEE Transactions on Signal Processing*, 47(10):2667–2676, 1999.
- [8] A. ANTONIADIS, J. BIGOT et S. LAMBERT-LACROIX : Peaks detection and alignment for mass spectrometry data. *Journal de la Société Française de Statistique*, 151(1):17–37, 2010.
- [9] E. ATHANASSOULA : The spiral structure of galaxies. *Physics Reports*, 114:319–403, 1984.
- [10] S. AWALI, L. POISSON, B. SOEP, M.-A. GAVEAU, M. BRIANT, C. POTHIER, J.-M. MESTDAGH, M. Ben El Hadj RHOUMA, M. HOCHLAF, V. MAZET et S. FAISAN : Time resolved observation of the solvation dynamics of a Rydberg excited molecule deposited on an argon cluster — I : DABCO* at short times. *Physical Chemistry Chemical Physics*, 2014.
- [11] F. BACH, R. JENATTON, J. MAIRAL et G. OBOZINSKI : Structured sparsity through convex optimization. *Statistical Science*, 27(4):450–468, 2012.
- [12] R. BACHER, F. CHATELAIN et O. MICHEL : Global error control procedure for spatially structured targets. *In EUSIPCO*, 2017.
- [13] R. BACON, S. CONSEIL, D. MARY, J. BRINCHMANN, M. SHEPHERD, M. AKHLAGHI, P. WEILBACHER, L. PIQUERAS, L. WISOTZKI, D. LAGATTUTA, B. EPINAT, A. GUEROU, H. INAMI, S. CANTALUPO, C. CLASTRES, J.-B. COURBOT, T. CONTINI, J. ROCHARD, M. MASEDA, R. BOUWENS, N. BOUCHÉ, W. KOLLATSCHNY, J. SCHAYE, R. ANNA MARINO, R. PELLO, C. HERENZ, B. GUIDERDONI et M. CAROLLO : The MUSE Hubble ultra deep field survey : I. Survey description, data reduction and source detection. *Astronomy & Astrophysics*, 608, 2017.
- [14] M.-A. BADIU, T.L. HANSEN et B.H. FLEURY : Variational Bayesian inference of line spectra. *IEEE Transactions on Signal Processing*, 65(9):2247–2261, 2017.

- [15] A. BAILLARD, E. BERTIN, Y. MELLIER, H.J. MCCRACKEN, T. GÉRAUD, R. PELLÓ, F. LEBORGNE et P. FOUQUÉ : Project EFIGI : Automatic classification of galaxies. In C. GABRIEL, C. ARVISET, D. PONZ et S. ENRIQUE, éditeurs : *Astronomical Data Analysis Software and Systems XV*, volume 351. Astronomical Society of the Pacific, 2006.
- [16] Y. BAR-SHALOM, F. DAUM et J. HUANG : The probabilistic data association filter. *IEEE Control Systems Magazine*, 29(6):82–100, 2009.
- [17] J.P. BAUDRY et G. CELEUX : EM for mixtures. initialization requires special care. *Statistics and Computing*, 25:713–726, 2015.
- [18] S.V.W. BECKWITH, M. STIAVELLI, A.M. KOEKEMOER, J.A.R. CALDWELL, H.C. FERGUSON, R. HOOK, R.A. LUCAS, L.E. BERGERON, M. CORBIN, S. JOGEE, N. PANAGIA, M. ROBERTO, P. ROYLE, R.S. SOMERVILLE et M. SOSEY : The Hubble ultra deep field. *The Astronomical Journal*, 132(5):1729–1755, 2006.
- [19] G. BEKIARIS, K. GLAZEBROOK, C.J. FLUKE et R. ABRAHAM : Kinematic modelling of disc galaxies using graphics processing units. *Monthly Notices of the Royal Astronomical Society*, 455:754–784, 2016.
- [20] L. BELMERHIA, E.-H. DJERMOUNE, D. BRIE et C. CARTERET : A regularized sparse approximation method for hyperspectral image classification. In *IEEE Statistical Signal Processing Workshop*, 2016.
- [21] L. BENTABET et J. MAODONG : A combined Markovian and Dirichlet sub-mixture modeling for evidence assignment : Application to image fusion. *Pattern Recognition Letters*, 29(13):1775–1783, 2008.
- [22] E. BERTIN et S. ARNOUITS : SExtractor : Software for source extraction. *Astronomy and Astrophysics*, 117:393–404, 1996.
- [23] J. BESAG : On the statistical analysis of dirty pictures. *Journal of the Royal Statistical Society : Series B*, 48(3):259–302, 1986.
- [24] A. BOUCHARD-CÔTÉ, S.J. VOLLMER et A. DOUCET : The bouncy particle sampler : A nonreversible rejection-free Markov chain Monte Carlo method. *Journal of the American Statistical Association*, 113(522):855–867, 2018.
- [25] N. BOUCHÉ, H. CARFANTAN, I. SCHROETTER, L. MICHEL-DANSAC et T. CONTINI : GalPaK3D : a Bayesian parametric tool for extracting morpho-kinematics from 3D data. *The Astronomical Journal*, 150(3), 2015.
- [26] M. BOUDINEAU : *Vers la résolution « optimale » de problèmes inverses non linéaires parcimonieux grâce à l'exploitation de variables binaires sur dictionnaires continus : applications en astrophysique*. Thèse de doctorat, Université de Toulouse, 2019.
- [27] M. BOUDINEAU, H. CARFANTAN, S. BOURGUIGNON et M. BAZOT : Sampling schemes and parameter estimation for nonlinear Bernoulli-Gaussian sparse models. In *SSP*, 2016.
- [28] S. BOURGUIGNON et H. CARFANTAN : Bernoulli-Gaussian spectral analysis of unevenly spaced astrophysical data. In *SSP*, pages 811–816, Bordeaux, France, 2005.
- [29] S. BOURGUIGNON, H. CARFANTAN et T. BÖHM : Sparspec : a new method for fitting multiple sinusoids with irregularly sampled data. *Astronomy & Astrophysics*, 462:379–387, 2007.
- [30] S. BOURGUIGNON, D. MARY et É. SLEZAK : Restoration of astrophysical spectra with sparsity constraints : Models and algorithms. *IEEE Journal of Selected Topics in Signal Processing*, 5(5):1002–1013, 2011.
- [31] S. BOURGUIGNON, J. NININ, H. CARFANTAN et M. MONGEAU : Exact sparse approximation problems via mixed-integer programming : Formulations and computational performance. *IEEE Transactions on Signal Processing*, 64(6):1405–1419, 2016.

- [32] S. BOURGUIGNON, C. SOUSSEN, H. CARFANTAN et J. IDIER : Sparse deconvolution : Comparison of statistical and deterministic approaches. *In SSP*, 2011.
- [33] S.P. BROOKS, P. GIUDICI et A. PHILIPPE : Nonparametric convergence assessment for MCMC model selection. *Journal of Computational & Graphical Statistics*, 12(1):1–22, 2003.
- [34] S.P. BROOKS et G.O. ROBERTS : Assessing convergence of Markov chain Monte Carlo algorithms. *Statistics and Computing*, 8:319–335, 1997.
- [35] A.M. BRUCKSTEIN, M. ELAD et M. ZIBULEVSKY : On the uniqueness of nonnegative sparse solutions to underdetermined systems of equations. *IEEE Transactions on Information Theory*, 54(11):4813–4820, 2008.
- [36] E.J. CANDÈS et C. FERNANDEZ-GRANDA : Towards a mathematical theory of super-resolution. *Communications on Pure and Applied Mathematics*, 67:906–956, 2014.
- [37] R.A. CARUANA, R.B. SEARLE, T. HELLER et S.I. SHUPACK : Fast algorithm for the resolution of spectra. *Analytical Chemistry*, 58:1162–1167, 1986.
- [38] M. CASTELLA, J.-C. PESQUET et A.P. PETROPULU : A family of frequency-and time-domain contrasts for blind separation of convolutive mixtures of temporally dependent signals. *IEEE Transactions on Signal Processing*, 53(1):107–120, 2005.
- [39] J.M. CASTELLOE et D.L. ZIMMERMAN : Convergence assessment for reversible jump MCMC samplers. Rapport technique, University of Iowa, 2002.
- [40] P. CATALA, V. DUVAL et G. PEYRÉ : A low-rank approach to off-the-grid sparse deconvolution. *In Journal of Physics : Conference Series*, volume 904, page 012015, 2017.
- [41] G. CELEUX : Bayesian inference for mixtures : the label switching problem. *In COMPSTAT*, pages 227–232, 1998.
- [42] G. CELEUX, M. HURN et C.P. ROBERT : Computational and inferential difficulties with mixture posterior distributions. *Journal of the American Statistical Association*, 95(451):957–970, 2000.
- [43] G. CHABRIEL et J. BARRÈRE : An instantaneous formulation of mixtures for blind separation of propagating waves. *IEEE Transactions on Signal Processing*, 54(1):49–58, 2006.
- [44] Q. CHENG, R. CHEN et T.-H. LI : Simultaneous wavelet estimation and deconvolution of reflection seismic signals. *IEEE Geoscience and Remote Sensing Society*, 34(2), 1996.
- [45] A. CHERNI, E. CHOUZENOUX et M.A. DELSUC : PALMA, an improved algorithm for DOSY signal processing. *Analyst*, 142(5):772–779, 2017.
- [46] É. CHOUZENOUX, J.-C. PESQUET et A. FLORESCU : A multi-parameter optimization approach for complex continuous sparse modelling. *In International Conference on Digital Signal Processing*, 2014.
- [47] C. COLLET, B. PERRET et V. MAZET : *Multivariate Image Processing*, chapitre Panoramic Integral-Field Spectrograph : Ultraspectral Data to Understand the History of the Universe. John Wiley and Sons, 2009.
- [48] P. COMON et C. JUTTEN, éditeurs. *Séparation de sources*. Traité IC2. Lavoisier, 2007.
- [49] J.-B. COURBOT, V. MAZET, E. MONFRINI et C. COLLET : Detection of faint extended sources in hyperspectral data and application to HDF-S MUSE observations. *In ICASSP*, pages 1891–1895, 2016.
- [50] J.-B. COURBOT, V. MAZET, E. MONFRINI et C. COLLET : Extended faint source detection in astronomical hyperspectral images. *Signal Processing*, 135:274–283, 2017.

- [51] J.-B. COURBOT, V. MAZET, E. MONFRINI et C. COLLET : Pairwise Markov fields for segmentation in astronomical hyperspectral images. *Signal Processing*, À paraître.
- [52] J.-B. COURBOT, E. MONFRINI, V. MAZET et C. COLLET : Oriented triplet Markov field for hyperspectral image segmentation. *In Whispers*, 2016.
- [53] J.-B. COURBOT, E. MONFRINI, V. MAZET et C. COLLET : Arbres de Markov triplets pour la segmentation d'images. *In GRETSI*, 2017.
- [54] J.-B. COURBOT, E. MONFRINI, V. MAZET et C. COLLET : Oriented triplet Markov fields. *Pattern Recognition Letters*, 103(1):16–22, 2018.
- [55] J.-B. COURBOT, E. MONFRINI, V. MAZET et C. COLLET : Triplet Markov trees for image segmentation. *In SSP*, 2018.
- [56] P. CRACIUN : *Stochastic geometry for automatic object detection and tracking in remotely sensed image sequences*. Thèse de doctorat, Université de Nice–Sophia Antipolis, 2015.
- [57] H.P. DANG et P. CHAINAIS : Towards dictionaries of optimal size : A Bayesian non parametric approach. *Journal of Signal Processing Systems*, 90(2):221–232, 2018.
- [58] S.C. DASS : Markov random field models for directional field and singularity extraction in fingerprint images. *IEEE Transactions on Image Processing*, 13(10):1358–1367, 2004.
- [59] R.E. DE SOUZA, D.A. GADOTTI et S. DOS ANJOS : Budda : A new two-dimensional bulge/disk decomposition code for detailed structural analysis of galaxies. *The Astrophysical Journal Supplement Series*, 153:411–427, 2004.
- [60] G. de VAUCOULEURS : Classification and morphology of external galaxies. *Handbuch der Physik*, 53:275–310, 1959.
- [61] H. DEBORAH, N. RICHARD et J.Y. HARDEBERG : A comprehensive evaluation of spectral distance functions and metrics for hyperspectral image processing. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 8(6):3224–3234, 2015.
- [62] A.P. DEMPSTER : Upper and lower probabilities induced by a multivalued mapping. *The Annals of Mathematical Statistics*, 38(2):325–339, 1967.
- [63] T. DENÇEUX : Conjunctive and disjunctive combination of belief functions induced by nondistinct bodies of evidence. *Artificial Intelligence*, 172(2–3):234–264, 2008.
- [64] X. DESCOMBES, R. MINLOS et E. ZHIZHINA : Object extraction using a stochastic birth-and-death dynamics in continuum. *Journal of Mathematical Imaging and Vision*, 33:347–359, 2009.
- [65] X. DESCOMBES et J. ZERUBIA : Marked point process in image analysis. *IEEE Signal Processing Magazine*, 19(5):77–84, 2002.
- [66] E.-H. DJERMOUNE, M. TOMCZAK et D. BRIE : NMR data analysis : A time-domain parametric approach using adaptive subband decomposition. *Oil & Gas Science and Technology*, 69(2):229–244, 2014.
- [67] C. DOSSAL et S. MALLAT : Sparse spike deconvolution with minimum scale. *In SPARS*, 2005.
- [68] A. DRÉMEAU, C. HERZET et L. DAUDET : Soft Bayesian pursuit algorithm for sparse representations. *In SSP*, 2011.
- [69] A. DRÉMEAU, C. HERZET et L. DAUDET : Boltzmann machine and mean-field approximation for structured sparse decompositions. *IEEE Transactions on Signal Processing*, 60(7):3425–3438, 2012.
- [70] V. DUVAL et G. PEYRÉ : Sparse regularization on thin grids I : the LASSO. *Inverse Problems*, 33(5), 2017.

- [71] V. DUVAL et G. PEYRÉ : Sparse spikes super-resolution on thin grids II : the continuous basis pursuit. *Inverse Problems*, 33(095008), 2017.
- [72] O. ECHES, J.A. BENEDIKTSSON, N. DOBIGEON et J.-Y. TURNERET : Adaptive Markov random fields for joint unmixing and segmentation of hyperspectral images. *IEEE Transactions on Image Processing*, 22(1):5–16, 2013.
- [73] C. EKANADHAM, D. TRANCHINA et E.P. SIMONCELLI : Recovery of sparse translation-invariant signals with continuous basis pursuit. *IEEE Transactions on Signal Processing*, 59(10):4735–4744, 2011.
- [74] M. ELAD et M. AHARON : Image denoising via sparse and redundant representations over learned dictionaries. *IEEE Transactions on Image Processing*, 15(12):3736–3745, 2006.
- [75] T. FAWCETT : An introduction to ROC analysis. *Pattern Recognition Letters*, 27:861–874, 2006.
- [76] J. FEMIANI, E. LI, A. RAZDAN et P. WONKA : Shadow-based rooftop segmentation in visible band images. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 8(5):2063–2077, 2015.
- [77] C. FERNANDEZ-GRANDA : Super-resolution of point sources via convex programming. *Information and Inference : A Journal of the IMA*, 5(3):251–303, 2016.
- [78] R. FISCHER et V. DOSE : Analysis of mixtures in physical spectra. *Monographs of Official Statistics : Bayesian Methods With Applications to Science, Policy, and Official Statistics*, pages 145–154, 2005. Il semble qu’une autre référence existe, bien qu’elle soit difficile à trouver : ISBA 2000 (6th world meeting of the International Society for Bayesian Analysis), Heraklion, Grèce. C’est en tout cas la citation que je donne dans ma thèse, qui date de 2005.
- [79] P. FLANDRIN : Le traitement du signal, au cœur de la science et de notre vie quotidienne. The Conversation [Accédé le 27/05/2019], 2017.
- [80] J.M. FLEGAL : *Monte Carlo Standard Errors in Markov Chain Monte Carlo*. Thèse de doctorat, University of Minnesota, 2008.
- [81] A. FLORESCU, É. CHOUZENOUX, J.-C. PESQUET et S. CIOCHINA : A constrained optimization approach for complex sparse perturbed models. *In SPARS*, 2013.
- [82] F. FORBES et A.E. RAFTERY : Bayesian morphology : Fast unsupervised Bayesian image analysis. *Journal of the American Statistical Association*, 94(446):555–568, 1999.
- [83] K. FYHN, M.F. DUARTE et S.H. JENSEN : Compressive parameter estimation for sparse translation-invariant signals using polar interpolation. *IEEE Transactions on Signal Processing*, 63(4):870–881, 2015.
- [84] D.A. GADOTTI : Image decomposition of barred galaxies and agn hosts. *Monthly Notices of the Royal Astronomical Society*, 384(1):420–439, 2008.
- [85] A. GAMAL-ELDIN, X. DESCOMBES, G. CHARPIAT et J. ZERUBIA : Multiple birth and cut algorithm for multiple object detection. *Journal of Multimedia Processing and Technologies*, 2012.
- [86] D. GEMAN et C. YANG : Nonlinear image recovery with half-quadratic regularization. *IEEE Transactions on Image Processing*, 4(7):932–946, 1995.
- [87] S. GEMAN et D. GEMAN : Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6(6):721–741, 1984.
- [88] S.J. GERSHMAN et D.M. BLEI : A tutorial on Bayesian nonparametric models. *Journal of Mathematical Psychology*, 56(1):1–12, 2012.
- [89] W.R. GILKS, S. RICHARDSON et D. SPIEGELHALTER, éditeurs. *Markov Chain Monte Carlo in Practice*. Chapman & Hall/CRC, 1996.

- [90] A. GLOAGUEN, J.-M. MESTDAGH, L. POISSON, F. LEPETIT, J.-P. VISTICOT, B. SOEP, M. COROIU, A. EPPINK et D.H. PARKER : Experimental evidence for ultrafast electronic relaxation in molecules, mediated by diffuse states. *Journal of the American Chemical Society*, 127(47): 16529–16534, 2005.
- [91] C. GOBINET, V. VRABIE, M. MANFAIT et O. PIOT : Preprocessing methods of Raman spectra for source extraction on biomedical samples : Application on paraffin-embedded skin biopsies. *IEEE Transactions on Biomedical Engineering*, 56(5):1371–1382, 2009.
- [92] A. GOSHTASBY et W.D. O’NEILL : Curve fitting by a sum of Gaussians. *CVGIP : Graphical Models and Image Processing*, 56:281–288, 1994.
- [93] Y. GOUSSARD : Blind deconvolution of sparse spike trains using stochastic optimization. In *ICASSP*, pages 593–596, 1992.
- [94] P.J. GREEN : Reversible jump Markov chain Monte Carlo computation and Bayesian model determination. *Biometrika*, 82:711–732, 1995.
- [95] P.J. GREEN : *Highly Structured Stochastic Systems*, chapitre Trans-dimensional Markov chain Monte Carlo. Oxford University Press, 2003.
- [96] T.L. GRIFFITHS et Z. GHAHRAMANI : The Indian buffet process : An introduction and review. *Journal of Machine Learning Research*, 12:1185–1224, 2011.
- [97] S. GULAM RAZUL, W.J. FITZGERALD et C. ANDRIEU : Bayesian model selection and parameter estimation of nuclear emission spectra using RJMCMC. *Nuclear Instruments and Methods in Physics Research Section A*, 497(2):492–510, 2003.
- [98] H. GUO : A simple algorithm for fitting a Gaussian function. *IEEE Signal Processing Magazine*, 28(5):134–137, 2011.
- [99] B. HAJEK : Cooling schedules for optimal annealing. *Mathematics of operations research*, 13:311–329, 1988.
- [100] T.L. HANSEN, B.H. FLEURY et B.D. RAO : Superfast line spectral estimation. *IEEE Transactions on Signal Processing*, 66(10):2511–2526, 2018.
- [101] R.A. HARSHMAN, S. HONG et M.E. LUNDY : Shifted factor analysis – part I : Models and properties. *Journal of Chemometrics*, 17:363–378, 2003.
- [102] J. HE et X. YANG : Distributions of constant directional derivatives and a model of spiral patterns. *Astrophysics and Space Science*, 302:7–15, 2006.
- [103] A. HEIMER et I. COHEN : Multichannel seismic deconvolution using Markov–Bernoulli random-field modeling. *IEEE Transactions on Geoscience and Remote Sensing*, 47(7):2047–2058, 2009.
- [104] C. HERZET et A. DRÉMEAU : Bayesian pursuit algorithms. arxiv.org/abs/1401.7538, 2014.
- [105] K.H. HOFFMANN, D. WÜRTZ, D. DE GROOT et M. HANF : Concepts in optimizing simulated annealing schedules : An adaptive approach for parallel and vector machines. In M. GRAUER et D.B. PRESSMAR, éditeurs : *Parallel Computing and Mathematical Optimization. Lecture Notes in Economics and Mathematical Systems*, volume 367. Springer, 1991.
- [106] S. HONG et R.A. HARSHMAN : Shifted factor analysis – Part II : Algorithms. *Journal of Chemometrics*, 17(7):379–388, 2003.
- [107] E.P. HUBBLE : Extragalactic nebulae. *Astrophysical Journal*, 64:321–369, 1926.
- [108] C. HUE, J.-P. LE CADRE et P. PÉREZ : Tracking multiple objects with particle filtering. Rapport technique, Inria, 2000.
- [109] J. IDIER et Y. GOUSSARD : Multichannel seismic deconvolution. *IEEE Transactions on Geoscience and Remote Sensing*, 31(5):961–979, 1993.

- [110] H. INAMI, R. BACON, J. BRINCHMANN, J. RICHARD, T. CONTINI, S. CONSEIL, S. HAMER, M. AKHLAGHI, N. BOUCHÉ, B. CLÉMENT, G. DESPREZ, A. B. DRAKE, T. HASHIMOTO, F. LECLERCQ, M. MASEDA, L. MICHEL-DANSAC, M. PAALVAST, L. TRESSE, E. VENTOU, W. KOLLATSCHNY, L. A. BOOGAARD, H. FINLEY, R. A. MARINO, J. SCHAYE et L. WISOTZKI : The MUSE Hubble ultra deep field survey II. spectroscopic redshifts and comparisons to color selections of high-redshift galaxies. *Astronomy & Astrophysics*, 608, 2017.
- [111] A. JASRA, C.C. HOLMES et D.A. STEPHENS : Markov chain Monte Carlo methods and the label switching problem in Bayesian mixture modelling. *Statistical Science*, 20:50–67, 2005.
- [112] N. JIANG et D. FARINA : Covariance and time-scale methods for blind separation of delayed sources. *IEEE Transactions on Biomedical Engineering*, 58(3):550–556, 2011.
- [113] G.L. JONES, M. HARAN, B.S. CAFFO et R. NEATH : Fixed-width output analysis for Markov chain Monte Carlo. *Journal of the American Statistical Association*, 101:1537–1547, 2006.
- [114] K.F. KAARESEN : Evaluation and applications of the iterated window maximization method for sparse deconvolution. *IEEE Transactions on Signal Processing*, 46(3):609–624, 1998.
- [115] G. KAIL, F. HLAWATSCH et C. NOVAK : SMLR-type blind deconvolution of sparse pulse sequences under a minimum temporal distance constraint. *IEEE Transactions on Signal Processing*, 63(18):4838–4853, 2015.
- [116] G. KAIL, J.-Y. TOURNERET, F. HLAWATSCH et N. DOBIGEON : Blind deconvolution of sparse pulse sequences under a minimum distance constraint : a partially collapsed Gibbs sampler method. *IEEE Transactions on Signal Processing*, 60(6):2727–2743, 2012.
- [117] H.G. KAPER, D.W. SMITS, U. SCHWARZ, K. TAKAKUBO et H. VAN WOERDEN : Computer analysis of observed distributions into Gaussian components. *Bulletin of the Astronomical Institutes of the Netherlands*, 18:465–487, 1966.
- [118] E. KASS et L. WASSERMAN : The selection of prior distributions by formal rules. *Journal of the American Statistical Association*, 91(435):1343–1370, 1996.
- [119] N. KAZEMI et M.D. SACCHI : Sparse multichannel blind deconvolution. *Geophysics*, 79(4):v143–v152, 2014.
- [120] B.C. KELLY et T.A. MCKAY : Morphological classification of galaxies by shapelet decomposition in the Sloan digital sky survey. *Astronomical Journal*, 127:625–645, 2004.
- [121] N. KESHAVA et J.F. MUSTARD : Spectral unmixing. *IEEE Signal Processing Magazine*, 19(1):44–57, 2002.
- [122] C. KIRBAS et F. QUEK : A review of vessel extraction techniques and algorithms. *ACM Computing Surveys (CSUR)*, 36(2):81–121, 2004.
- [123] S. KIRKPATRICK, C.D. GELATT et M.P. VECCHI : Optimization by simulated annealing. *Science*, 220(4598), 1983.
- [124] K.C. KNUDSON, J. YATES, A. HUK et J.W. PILLOW : Inferring sparse representations of continuous signals with continuous orthogonal matching pursuit. *In NIPS*, 2014.
- [125] S. KONISHI et G. KITAGAWA : Generalised information criteria in model selection. *Biometrika*, 83(4):875–890, 1996.
- [126] J. KORMYLO et J. MENDEL : Maximum likelihood detection and estimation of Bernoulli–Gaussian processes. *IEEE Transactions on Information Theory*, 28(3):482–488, 1982.
- [127] M. KOWALSKI, K. SIEDENBURG et M. DÖRFLER : Social sparsity! Neighborhood systems enrich structured shrinkage operators. *IEEE Transactions on Signal Processing*, 61(10):2498–2511, 2013.

- [128] H. KWAKERNAAK : Estimation of pulse heights and arrival times. *Automatica*, 16:367–377, 1980.
- [129] C. LACOSTE, X. DESCOMBES et J. ZERUBIA : Point processes for unsupervised line network extraction in remote sensing. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(10):1568–1579, 2005.
- [130] C. LACOSTE, X. DESCOMBES et J. ZERUBIA : Unsupervised line network extraction in remote sensing using a polyline process. *Pattern Recognition*, 43:1631–1641, 2010.
- [131] J.-M. LAFERTÉ, P. PÉREZ et F. HEITZ : Discrete Markov image modeling and inference on the quadtree. *IEEE Transactions on Image Processing*, 9(3):390–404, 2000.
- [132] M. LAVIELLE et É. LEBARBIER : An application of MCMC methods for the multiple change-points problem. *Signal Processing*, 81:39–53, 2001.
- [133] R.A. LEVINE et G. CASELLA : Optimizing random scan Gibbs samplers. *Journal of Multivariate Analysis*, 97:2071–2100, 2006.
- [134] S. LEVY et P. FULLAGAR : Reconstruction of a sparse spike train from a portion of its spectrum and application to high resolution deconvolution. *Geophysics*, 46(9):1235–1243, 1981.
- [135] E. LI, J. FEMIANI, S. XU, X. ZHANG et P. WONKA : Robust rooftop extraction from visible band images using higher order CRF. *IEEE Transactions on Geoscience and Remote Sensing*, 53(8):4483–4495, 2015.
- [136] J. LI, J.M. BLOUCAS-DIAS et A. PLAZA : Spectral–spatial hyperspectral image segmentation using subspace multinomial logistic regression and Markov random fields. *IEEE Transactions on Geoscience and Remote Sensing*, 50(3):809–823, 2012.
- [137] W. LI, S. PRASAD et J.E. FOWLER : Hyperspectral image classification using Gaussian mixture models and Markov random fields. *IEEE Geoscience and Remote Sensing Letters*, 11(1):153–157, 2014.
- [138] A. LIETARD, G. PIANI, M. BRIANT, M.-A. GAVEAU, S. FAISAN, V. MAZET, B. SOEP, J.-M. MEST-DAGH et L. POISSON : Self-trapping relaxation decay investigated by time-resolved photoelectron spectroscopy. *Physical Chemistry Chemical Physics*, 2018.
- [139] S. LING et T. STROHMER : Blind deconvolution meets blind demixing : Algorithms and performance bounds. *IEEE Transactions on Information Theory*, 63(7):4497–4520, 2017.
- [140] H. LIU, S. LIU, Z. ZHANG, J. SUN et J. SHU : Adaptive total variation-based spectral deconvolution with the split Bregman method. *Applied Optics*, 53:8240–8248, 2014.
- [141] H.D. MAKAREWICZ, M. PARENTE et J.L. BISHOP : Deconvolution of VNIR spectra using modified Gaussian modeling (MGM) with automatic parameter initialization (API) applied to CRISM. *In Whispers*, 2009.
- [142] G. MALSINER-WALLI, S. FRÜHWIRTH-SCHNATTER et B. GRÜN : Model-based clustering based on sparse finite Gaussian mixtures. *Statistics and Computing*, 26(1–2):303–324, 2016.
- [143] B. MAMANDIPOOR, D. RAMASAMY et U. MADHOW : Newtonized orthogonal matching pursuit : Frequency estimation over the continuum. *IEEE Transactions on Signal Processing*, 64(19):5066–5081, 2016.
- [144] J.M. MARIN, K. MENGERSEN et C.P. ROBERT : Bayesian modelling and inference on mixtures of distributions. *Handbook of statistics*, 25:459–507, 2005.
- [145] R. MASSEY, A. REFREGIER, C.J. CONSELICE, J. DAVID et J. BACON : Image simulation with shapelets. *Monthly Notices of the Royal Astronomical Society*, 348:214–226, 2004.

- [146] A. MASSON, L. POISSON, M.A. GAVEAU, B. SOEP, J.M. MESTDAGH, V. MAZET et F. SPIEGELMAN : Dynamics of highly excited barium atoms deposited on large argon clusters. I. General trends. *The Journal of chemical physics*, 133(5), 2010.
- [147] B.W. MATTHEWS : Comparison of the predicted and observed secondary structure of T4 phage lysozyme. *Biochimica et Biophysica Acta — Protein Structure*, 405(2):442–451, 1975.
- [148] V. MAZET : *Développement de méthodes de traitement de signaux spectroscopiques : estimation de la ligne de base et du spectre de raies*. Thèse de doctorat, Université Henri Poincaré, Nancy I, 2005.
- [149] V. MAZET : Joint Bayesian decomposition of a spectroscopic signal sequence. *IEEE Signal Processing Letters*, 18(3):181–184, mars 2011.
- [150] V. MAZET, D. BRIE et J. IDIER : Decomposition of a chemical spectrum using a marked point process and a constant dimension model. *In MaxEnt*, 2006.
- [151] V. MAZET, C. COLLET et B. VOLLMER : Decomposition and classification of spectral lines in astronomical radio data cubes. *In SCIA*, Oslo, Norvège, 2009.
- [152] V. MAZET, C. COLLET et B. VOLLMER : Décomposition et classification de composantes spectrales dans des cubes de données radio-astronomiques. *In GRETSI*, Dijon, France, 2009.
- [153] V. MAZET, S. FAISAN, S. AWALI, M.-A. GAVEAU et L. POISSON : Unsupervised joint decomposition of a spectroscopic signal sequence. *Signal Processing*, 109:193–205, 2015.
- [154] V. MAZET, S. FAISAN, A. MASSON, M.-A. GAVEAU et L. POISSON : Approche bayésienne pour la décomposition conjointe d’une séquence de spectres de photoélectrons. *In GRETSI*, Bordeaux, France, 2011.
- [155] V. MAZET, S. FAISAN, A. MASSON, M.-A. GAVEAU et L. POISSON : Unsupervised joint Bayesian decomposition of a sequence of photoelectron spectra. *In WHISPERS*, Lisbonne, Portugal, 2011.
- [156] V. MAZET, S. FAISAN, A. MASSON, M.-A. GAVEAU, L. POISSON et J.-M. MESTDAGH : Joint Bayesian decomposition of a spectroscopic signal sequence with RJMCMC. *In IEEE Workshop on Statistical Signal Processing*, Ann Arbor, États-Unis, 2012.
- [157] V. MAZET, S. FAISAN, A. MASSON, M.-A. GAVEAU, L. POISSON et J.-M. MESTDAGH : Approche bayésienne pour la décomposition conjointe d’une séquence de spectres de photo-électrons. *Traitement du signal*, 30(1):9–34, 2013.
- [158] V. MAZET, S. FAISAN, L. POISSON, M.-A. GAVEAU et J.-M. MESTDAGH : Décomposition d’une séquence de spectres avec modèle markovien et algorithme RJMCMC à deux variables de dimension. *In GRETSI*, Brest, France, 2013.
- [159] V. MAZET, F. FLITTI et C. COLLET : *Multivariate Image Processing*, chapitre Detection and Tracking of Emission Rays in Radioastronomy. John Wiley and Sons, 2009.
- [160] V. MAZET, C. SOUSSEN et E.-H. DJERMOUNE : Décomposition de spectres en motifs paramétriques par approximation parcimonieuse. *In GRETSI*, Brest, France, 2013.
- [161] G. MCLACHLAN et D. PEEL : *Finite Mixture Models*. Wiley, 2000.
- [162] C. MCLEAN, M. PAULEY et J.H. MANTON : Limitations of decision based pile-up correction algorithms. *In IEEE Statistical Signal Processing Workshop*, pages 693–697, 2018.
- [163] S. MEIGNEN, Q. LEGROS, Y. ALTMANN et S. MCLAUGHLIN : A novel algorithm for the identification of dirac impulses from filtered noisy measurements. *Signal Processing*, 162:268–281, 2019.
- [164] C. MEILLIER : *Détection de sources quasi-ponctuelles dans des champs de données massifs*. Thèse de doctorat, Université Grenoble Alpes, 2015.

- [165] M.S. MEKETON et B.W. SCHMEISER : Overlapping batch means : Something for nothing? *In Winter Simulation Conference*, 1984.
- [166] M. MIGNOTTE, C. COLLET, P. PEREZ et P. BOUTHEMY : Sonar image segmentation using an unsupervised hierarchical MRF model. *IEEE Transactions on Image Processing*, 9(7):1216–1231, 2000.
- [167] M.I. MILLER, S.C. CHEN, D.A. KUEFLER et D.A. D’AVIGNON : Maximum likelihood and the EM algorithm for 2D NMR spectroscopy. *Journal of Magnetic Resonance, Series A*, 104(3):247–257, 1993.
- [168] M.I. MILLER et A. GREENE : Maximum-likelihood estimation for nuclear magnetic resonance spectroscopy. *Journal of Magnetic Resonance*, 83(3):525–548, 1989.
- [169] S. MIRZAEI, H. VAN HAMME et Y. NOROUZI : Blind audio source counting and separation of anechoic mixtures using the multichannel complex NMF framework. *Signal Processing*, 115:27–37, 2015.
- [170] A. MISIRIOTIS, N.D. KYLAFIS, J. PAPAMASTORAKIS et E.M. XILOURIS : Is the exponential distribution a good approximation of dusty galactic disks? *Astronomy and Astrophysics*, 353:117–123, 2000.
- [171] A. MOHAMMAD-DJAFARI : Une méthode bayésienne pour la localisation et la séparation de sources de formes connues. *In GRETSI*, 1997.
- [172] E. MONFRINI et W. PIECZYNSKI : Estimation de mélanges généralisés dans les arbres de Markov cachés, application à la segmentation des images de cartons d’orgue de Barbarie. *Traitement du Signal*, 22(2):135–147, 2005.
- [173] H. MORTADA : *Separation of parameterized and delayed sources : application to spectroscopic and multispectral data*. Thèse de doctorat, Université de Strasbourg, 2018.
- [174] H. MORTADA, V. MAZET, C. SOUSSEN et C. COLLET : Separation of delayed parameterized sources. *In EUSIPCO*, Kos, Grèce, 2017.
- [175] H. MORTADA, V. MAZET, C. SOUSSEN et C. COLLET : Séparation de sources retardées et paramétriques. *In GRETSI*, 2017.
- [176] H. MORTADA, V. MAZET, C. SOUSSEN et C. COLLET : Spectroscopic decomposition of astronomical multispectral images using B-splines. *In Whispers*, Amsterdam, Pays-Bas, 2018.
- [177] H. MORTADA, V. MAZET, C. SOUSSEN, C. COLLET et L. POISSON : Parameterized source separation for delayed spectroscopic signals. *Signal Processing*, 158:48–60, 2019.
- [178] M. MØRUP, K.H. MADSEN et L.K. HANSEN : Shifted non-negative matrix factorization. *In IEEE Workshop on Machine Learning for Signal Processing*, pages 139–144, 2007.
- [179] P. MÜLLER et R. MITRA : Bayesian nonparametric inference – why and how. *Bayesian Analysis*, 8(2):269–302, 2013.
- [180] K. NAGATA, S. SUGITA et M. OKADA : Bayesian spectral deconvolution with the exchange Monte Carlo method. *Neural Networks*, 28:82–89, 2012.
- [181] B.K. NATARAJAN : Sparse approximate solutions to linear systems. *SIAM Journal on Computing*, 24(2):227–234, 1995.
- [182] R. NEAL : *Handbook of Markov Chain Monte Carlo*, chapitre MCMC using Hamiltonian dynamics. Chapman & Hall/CRC, 2013.
- [183] T.T. NGO : *Shadow/Vegetation and building detection from single optical remote sensing image*. Thèse de doctorat, Université de Strasbourg, 2015.

- [184] T.T. NGO, C. COLLET et V. MAZET : Détection simultanée de l'ombre et la végétation sur des images aériennes couleur en haute résolution. *In RFIA*, 2014.
- [185] T.T. NGO, C. COLLET et V. MAZET : MRF and Dempster–Shafer theory for simultaneous shadow/vegetation detection on high resolution aerial color images. *In ICIP*, 2014.
- [186] T.T. NGO, C. COLLET et V. MAZET : Automatic rectangular building detection from VHR aerial imagery using shadow and image segmentation. *In ICIP*, 2015.
- [187] T.T. NGO, C. COLLET et V. MAZET : Détection simultanée de l'ombre et la végétation sur des images aériennes couleur en haute résolution. *Traitement du Signal*, 32:311–333, 2015.
- [188] T.T. NGO, V. MAZET, C. COLLET et P. DE FRAIPONT : Shape-based building detection in visible band images using shadow information. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 10(3):920–932, 2017.
- [189] D. NION, B. VANDEWOESTYNE, S. VANAUVERBEKE, K. VAN DEN ABEELE, H. DE GERSEM et L. DE LATHAUWER : A time-frequency technique for blind separation and localization of pure delayed sources. *In International Conference on Latent Variable Analysis and Signal Separation*, 2010.
- [190] B. NSIRI, J.M. BOUCHER et T. CHONAVEL : Multichannel blind deconvolution application to marine seismic. *In Oceans 2003*, volume 5, pages P2761–P2766, San Diego, Californie, 2003.
- [191] A.O. OK : Automated detection of buildings from single VHR multispectral images using shadow information and graph cuts. *ISPRS Journal of Photogrammetry and Remote Sensing*, 86:21–40, 2013.
- [192] A.O. OK, C. SENARAS et B. YUKSEL : Automated detection of arbitrarily shaped buildings in complex environments from monocular VHR optical satellite imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 51(3):1701–1717, 2013.
- [193] D.W. OLDENBURG, S. LEVY et K.J. STINSON : Inversion of band-limited reflection seismograms : theory and practice. *Proceedings of the IEEE*, 74(3):487–497, 1986.
- [194] L. OMLOR et M.A. GIESE : Anechoic blind source separation using Wigner marginals. *Journal of Machine Learning Research*, 12:1111–1148, 2011.
- [195] P. ORBANZ et J.M. BUHMANN : Nonparametric Bayesian image segmentation. *International Journal of Computer Vision*, 77(1–3):25–45, 2008.
- [196] M. ORTNER, X. DESCOMBES et J. ZERUBIA : A marked point process of rectangles and segments for automatic analysis of digital elevation models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 30(1):105–119, 2008.
- [197] N. OTSU : A threshold selection method from gray-level histograms. *IEEE Transactions on Systems, Man, and Cybernetics*, 9(1):62–66, 1979.
- [198] W.S.B. OUEDRAOGO, B. NICOLAS, B. OUDOMPHENG, J.I. MARS et C. JUTTEN : A frequency method for blind separation of an anechoic mixture. *In EUSIPCO*, 2014.
- [199] S. PARIS, R.F.R. SULEIMAN, D. MARY et A. FERRARI : Constrained likelihood ratios for detecting sparse signals in highly noisy 3D data. *In ICASSP*, pages 3947–3951, 2013.
- [200] N. PASSAT, C. RONSE, J. BARUTHIO, J.-P. ARMSPACH, C. MAILLOT et C. JAHN : Region-growing segmentation of brain vessels : an atlas-based automatic approach. *Journal of Magnetic Resonance Imaging*, 21(6):715–25, 2005.
- [201] C.Y. PENG, L.C. HO, C.D. IMPEY et H.-W. RIX : Detailed structural decomposition of galaxy images. *The Astronomical Journal*, 124(1):266–293, 2002.
- [202] C.Y. PENG, L.C. HO, C.D. IMPEY et H.-W. RIX : Detailed decomposition of galaxy images. II. beyond axisymmetric models. *The Astronomical Journal*, 139(6), 2010.

- [203] M. PEREYRA, P. SCHNITER, É. CHOUZENOUX, J.-C. PESQUET, J.-Y. TOURNERET, A.O. HERO et S. MCLAUGHLIN : A survey of stochastic simulation and optimization methods in signal processing. *IEEE Journal of Selected Topics in Signal Processing*, 10(2):224–241, 2016.
- [204] B. PERRET : *Caractérisation multibande de galaxies par hiérarchie de modèles et arbres de composantes connexes*. Thèse de doctorat, Université de Strasbourg, 2010.
- [205] B. PERRET, V. MAZET, C. COLLET et E. SLEZAK : Décomposition d’images multispectrales de galaxies au moyen d’algorithmes de Monte Carlo par chaînes de Markov. In *GRETSI*, 2009.
- [206] B. PERRET, V. MAZET, C. COLLET et E. SLEZAK : Galaxy decomposition in multispectral images using Markov chain Monte Carlo algorithms. In *Scandinavian Conference on Image Analysis*, 2009.
- [207] B. PERRET, V. MAZET, C. COLLET et E. SLEZAK : Hierarchical multispectral galaxy decomposition using a MCMC algorithm with multiple temperature simulated annealing. *Pattern Recognition*, 44(6):1328–1342, 2011.
- [208] G. PERRIN : *Étude du couvert forestier par processus ponctuels marqués*. Thèse de doctorat, École Centrale de Paris, 2006.
- [209] B. PHOOKUN et S.N. Vogel L.G. MUNDY : NGC 4254 : A spiral galaxy with an $m = 1$ mode and infalling gas. *The Astrophysical Journal*, 418:113–122, 1993.
- [210] W. PIECZYNSKI : Arbres de Markov triplet et fusion de Dempster–Shafer. *Comptes Rendus Mathématique*, 336(10):869–872, 2003.
- [211] W. PIECZYNSKI, D. BENBOUDJEMA et P. LANCHANTIN : Statistical image segmentation using triplet Markov fields. In *SPIE International Symposium on Remote Sensing*, 2002.
- [212] W. PIECZYNSKI et A.-N. TEBBACHE : Pairwise Markov random fields and segmentation of textured images. *Machine Graphics and Vision*, 2000.
- [213] L. POISSON, R. MAKSIMENKA, B. SOEP, J.-M. MESTDAGH, D.H. PARKER, M. NSANGOU et M. HOCHLAF : Unusual quantum interference in the S1 state of DABCO and observation of intramolecular vibrational redistribution. *Journal of Physical Chemistry A*, 114:3313–3319, 2010.
- [214] L. POISSON, J.-M. MESTDAGH, B. SOEP, A. GONZALEZ et M. CHERGUI : Direct observation of microscopic solvation at the surface of clusters by ultrafast photoelectron imaging. *Journal of Physical Chemistry A*, 112(39):9200–9210, 2008.
- [215] M. PUIGT et Y. DEVILLE : Time–frequency ratio-based blind separation methods for attenuated and time-delayed sources. *Mechanical Systems and Signal Processing*, 19(6):1348–1379, 2005.
- [216] I. RAM, I. COHEN et S. RAZ : Multichannel deconvolution of seismic signals using statistical MCMC methods. *IEEE Transactions on Signal Processing*, 58(5):2757–2770, 2010.
- [217] G. RELIER, X. DESCOMBES, F. FALZON et J. ZERUBIA : Texture feature analysis using a Gauss–Markov model in hyperspectral image classification. *IEEE Transactions on Geoscience and Remote Sensing*, 42(7):1543–1551, 2004.
- [218] A. REPETTI, M.Q. PHAM, L. DUVAL, É. CHOUZENOUX et J.-C. PESQUET : Euclid in a taxicab : Sparse blind deconvolution with smoothed ℓ_1/ℓ_2 regularization. *IEEE Signal Processing Letters*, 22(5):539–543, 2015.
- [219] S. RICHARDSON et P.J. GREEN : On Bayesian analysis of mixtures with an unknown number of components. *Journal of the Royal Statistical Society : Series B*, 59(4):731–792, 1997.
- [220] J. RISSANEN : Modeling by shortest data description. *Automatica*, 14(5):465–658, 1978.
- [221] C.P. ROBERT : *The Bayesian Choice*. Springer, 2 édition, 2001.

- [222] C.P. ROBERT et G. CASELLA : *Monte Carlo Statistical Methods*. Springer-Verlag, 2^e édition, 2004.
- [223] C.E. RODRÍGUEZ et S.G. WALKER : Label switching in Bayesian mixture models : Deterministic relabeling strategies. *Journal of Computational and Graphical Statistics*, 2014.
- [224] O. ROSEC, J.-M. BOUCHER, B. NSIRI et T. CHONAVEL : Blind marine seismic deconvolution using statistical MCMC methods. *IEEE Journal of Oceanic Engineering*, 28(3), 2003.
- [225] A. ROUEFF, J. CHANUSSOT, J.I. MARS et M.-Q. NGUYEN : Unsupervised separation of seismic waves using the watershed algorithm on time-scale images. *Geophysical Prospecting*, 52:287–300, 2004.
- [226] D.V. RUBTSOV et J.L. GRIFFIN : Time-domain Bayesian detection and estimation of noisy damped sinusoidal signals applied to NMR spectroscopy. *Journal of Magnetic Resonance*, 188:367–379, 2007.
- [227] S. SAHNOUN, E.-H. DJERMOUNE, D. BRIE et P. COMON : A simultaneous sparse approximation method for multidimensional harmonic retrieval. *Signal Processing*, 131:36–48, 2017.
- [228] E. SALVADOR, A. CAVALLARO et T. EBRAHIMI : Shadow identification and classification using invariant color models. In *ICASSP*, pages 1545–1548, 2001.
- [229] H. SAWADA, S. WINTER, R. MUKAI, S. ARAKI et S. MAKINO : Estimating the number of sources for frequency-domain blind source separation. In *International Conference on Independent Component Analysis and Signal Separation*, 2004.
- [230] G. SCHWARZ : Estimating the dimension of a model. *Annals of Statistics*, 6(2):461–464, 1978.
- [231] S.M. SCHWEIZER et J.M. MOURA : Hyperspectral imagery : clutter adaptation in anomaly detection. *IEEE Transactions on Information Theory*, 2000.
- [232] G. SHAFER : *A Mathematical Theory of Evidence*. Princeton University Press, 1976.
- [233] N. SHERVASHIDZE et F. BACH : Learning the structure for structured sparsity. *IEEE Transactions on Signal Processing*, 63(18):4894–4902, 2015.
- [234] N. SHORTER et T. KASPARIS : Automatic vegetation identification and building detection from a single nadir aerial image. *Remote Sensing*, 1:731–757, 2009.
- [235] L. SIMARD, C.N.A. WILLMER, N.P. VOGT, V.L. SARAJEDINI, A.C. PHILLIPS, B.J. WEINER, D.C. KOO, M. IM, G.D. ILLINGWORTH et S.M. FABER : The deep groth strip survey. II. Hubble space telescope structural parameters of galaxies in the groth strip. *The Astrophysical Journal Supplement Series*, 142(1), 2002.
- [236] R.C. SMITH : Resolution in the derivative recording of absorption spectra. *Review of Scientific Instruments*, 34:296–297, 1963.
- [237] C. SOUSSEN, J. IDIER, D. BRIE et J. DUAN : From Bernoulli–Gaussian deconvolution to sparse signal restoration. *IEEE Transactions on Signal Processing*, 59(10):4572–4584, 2011.
- [238] C.M. SPARROW : On spectroscopic resolving power. *The Astrophysical Journal*, 44:76, 1916.
- [239] M. SPERRIN, T. JAKI et E. WIT : Probabilistic relabelling strategies for the label switching problem in Bayesian mixture models. *Statistics and Computing*, 20:357–366, 2010.
- [240] M. STEPHENS : Dealing with label-switching in mixture models. *Journal of the Royal Statistical Society, Series B*, 62:795–809, 2000.
- [241] P. STOICA, R.L. MOSES, B. FRIEDLANDER et T. SODERSTROM : Maximum likelihood estimation of the parameters of multiple sinusoids from noisy measurements. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 37(3):378–392, 1989.

- [242] A. STOLOW, A.E. BRAGG et D.M. NEUMARK : Femtosecond time-resolved photoelectron spectroscopy. *Chemical Reviews*, 104(4):1719–1758, 2004.
- [243] J.L. SÉRSIC : *Atlas de galaxies australes*. Observatorio Astronomico, Universidad Nacional de Cordoba, 1968.
- [244] O. TEKE, A.C. GURBUZ et O. ARIKAN : Perturbed orthogonal matching pursuit. *IEEE Transactions on Signal Processing*, 61(24):6220–6231, 2013.
- [245] M. THONNAT : Toward an automatic classification of galaxies. In *The World of Galaxies*, 1989.
- [246] J. TIAN, J. SUN et Y. TANG : Tricolor attenuation model for shadow detection. *IEEE Transactions on Image Processing*, 18(10):2355–2363, 2009.
- [247] S. TOKUDA, K. NAGATA et M. OKADA : Simultaneous estimation of noise variance and number of peaks in Bayesian spectral deconvolution. *Journal of the Physical Society of Japan*, 86(024001), 2017.
- [248] G. TOMASI, F. VAN DEN BERG et C. ANDERSSON : Correlation optimized warping and dynamic time warping as preprocessing methods for chromatographic data. *Journal of Chemometrics*, 18:231–241, 2004.
- [249] T. TRIGANO, I. GILDIN et Y. SEPULCRE : Pileup correction algorithm using an iterated sparse reconstruction method. *IEEE Signal Processing Letters*, 22(9):1392–1395, 2015.
- [250] T. TRIGANO, I. SHEVTSOV et D. LUENGO : Cosa : An accelerated ISTA algorithm for dictionaries based on translated waveforms. *Signal Processing*, 139:131–135, 2017.
- [251] J.A. TROPP, A.C. GILBERT et M.J. STRAUSS : Algorithms for simultaneous sparse approximation. part I : Greedy pursuit. *Signal Processing*, 86:572–588, 2006.
- [252] V.J.D. TSAI : A comparative study on shadow compensation of color aerial images in invariant color models. *IEEE Transactions on Geoscience and Remote Sensing*, 44(6):1661–1671, 2006.
- [253] S. van der BERGH : Luminosity classifications of dwarf galaxies. *Astronomical Journal*, 71:922–926, 1966.
- [254] P. VANNOORENBERGHE, O. COLOT et D. DE BRUCQ : Color image segmentation using Dempster-Shafer’s theory. In *ICIP*, 1999.
- [255] N. VASWANI et J. ZHAN : Recursive recovery of sparse signal sequences from compressive measurements : A review. *IEEE Transactions on Signal Processing*, 64(13):3523–3549, 2016.
- [256] V. ČERNÝ : Thermodynamical approach to the traveling salesman problem : An efficient simulation algorithm. *Journal of Optimization Theory and Applications*, 45:41–51, 1985.
- [257] M. VETTERLI, P. MARZILIANO et T. BLU : Sampling signals with finite rate of innovation. *IEEE Transactions on Signal Processing*, 50(6):1417–1428, 2002.
- [258] E. VILLENEUVE et H. CARFANTAN : Nonlinear deconvolution of hyperspectral data with MCMC for studying the kinematics of galaxies. *IEEE Transactions on Image Processing*, 23(10):4322–4335, 2014.
- [259] B.-N. VO et W.-K. MA : The Gaussian mixture probability hypothesis density filter. *IEEE Transactions on Signal Processing*, 54(11):4091–4104, 2006.
- [260] V.D. VRABIE, J.I. MARS et J.-L. LACOUME : Modified singular value decomposition by means of independent component analysis. *Signal Processing*, 84:645–652, 2004.
- [261] D.M. WOEBBECKE, G.E. MEYER, K. VON BARGEN et D.A. MORTENSEN : Color indices for weed identification under various soil, residue, and lighting conditions. *Transactions of the ASAE*, 38(1):259–269, 1995.

- [262] J. XIA, J. CHANUSSOT, P. DU et X. HE : Spectral–spatial classification for hyperspectral data using rotation forests with local feature extraction and Markov random fields. *IEEE Transactions on Geoscience and Remote Sensing*, 53(5):2532–2546, 2015.
- [263] M. YAGHOOBI, L. DAUDET et M.E. DAVIES : Parametric dictionary design for sparse coding. *IEEE Transactions on Signal Processing*, 57(12):4800–4810, 2009.
- [264] M. YAGHOOBI, D. WU, R.J. CLEWES et M.E. DAVIES : Fast sparse Raman spectral unmixing for chemical fingerprinting and quantification. In *Optics and Photonics for Counterterrorism, Crime Fighting, and Defence XII*, volume 9995, page 99950E, 2016.
- [265] Ö. YILMAZ et S. RICKARD : Blind separation of speech mixtures via time-frequency masking. *IEEE Transactions on Signal Processing*, 52(7):1830–1847, 2004.
- [266] H. ZHANG, G. BI, L. YANG, S. GULAM RAZUL et C.-M. SAMSON SEE : IF extraction of multicomponent radar signals based on time-frequency gradient image. In *IEEE Radar Conference*, 2014.
- [267] H. ZHANG, G. BI, L. ZHAO, S. GULAM RAZUL et C.-M. SAMSON SEE : Time-varying filtering and separation of nonstationary FM signals in strong noise environments. In *ICASSP*, 2014.

Mémoire d'habilitation à diriger des recherches

Inférence bayésienne et représentations parcimonieuses pour la décomposition en motifs

Vincent MAZET

Résumé

Ce mémoire résume les principaux travaux de recherche effectués depuis ma thèse de doctorat. Le problème considéré est la décomposition conjointe de plusieurs signaux en motifs paramétrés, avec l'originalité que les paramètres de ces motifs évoluent lentement à travers les signaux. Ce problème se retrouve en spectroscopie de photoélectrons, où les données sont une séquence temporelle de signaux spectroscopiques et les motifs sont modélisés par des gaussiennes. Ce problème se retrouve également en imagerie multispectrale astronomique, où chaque pixel des images est un signal spectroscopique où les motifs sont les raies d'émission. Dans un premier temps, l'approche bayésienne permet de modéliser le problème, et l'optimisation est effectuée à l'aide d'un échantillonneur de Gibbs ou de l'algorithme RJMCMC, selon que le nombre de motifs est connu ou non. Dans un deuxième temps, nous explorons l'utilisation d'une modélisation parcimonieuse qui donne des résultats tout aussi satisfaisants mais bien plus rapidement et avec le bénéfice d'être plus simple à mettre en œuvre.

Mots clés Déconvolution impulsienne, séparation de sources, algorithmes MCMC, échantillonneur de Gibbs, algorithme RJMCMC, approximation parcimonieuse, spectroscopie de photoélectrons, imagerie astronomique, cinématique des galaxies.

Abstract

This thesis resumes the main works studied since my Ph.D. We consider the problem of the joint decomposition of several signals in parameterized patterns, with the originality that the pattern parameters evolve slowly through the signals. This problem originates from photoelectron spectroscopy, where data are a temporal sequence of spectroscopic signals whose patterns are modelled by Gaussians. Also, this problem is found in multispectral imaging in astronomy, where each pixel in the image is a spectroscopic signal whose patterns are the emission lines. On the one hand, a Bayesian approach is used to model the problem, and the optimization is done using a Gibbs sampler or the RJMCMC algorithm, whether the pattern number is known. On the other hand, we study the use of sparse approximation which yields satisfactory results while being faster and simpler to implement.

Keywords Sparse spike deconvolution, source separation, MCMC algorithms, Gibbs sampler, RJMCMC algorithm, sparse approximation, photoelectron spectroscopy, astronomical image processing, galaxy kinematics.