



**HAL**  
open science

# Apprentissage à l'aide de techniques issues de la théorie de l'information pour la spectroscopie Raman comprimée

Timothée Alain Marcos Justel

► **To cite this version:**

Timothée Alain Marcos Justel. Apprentissage à l'aide de techniques issues de la théorie de l'information pour la spectroscopie Raman comprimée. Statistiques [stat]. Aix-marseille université, 2023. Français. NNT : 2023AIXM0371 . tel-04419536

**HAL Id: tel-04419536**

**<https://theses.hal.science/tel-04419536>**

Submitted on 26 Jan 2024

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

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



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License

# THÈSE DE DOCTORAT

Soutenue à Aix-Marseille Université  
le 31/10/2023 par

## Timothée Alain Marcos JUSTEL

Apprentissage à l'aide de techniques issues de la théorie de  
l'information pour la spectroscopie Raman comprimée

### Discipline

Physique et Sciences de la Matière

### Spécialité

Optique, Photonique et Traitement d'Image

### École doctorale

ED 352 Physique et Sciences de la Matière

### Laboratoire/Partenaires de recherche

Institut Fresnel

### Composition du jury

|   |                       |
|---|-----------------------|
| Pr. François GOUDAIL<br>Institut d'Optique Graduate<br>School | Rapporteur            |
| Pr. Olivier HAEBERLÉ<br>Université de Haute Alsace            | Rapporteur            |
| Pr. Mehdi ALOUINI<br>Université de Rennes 1                   | Président du jury     |
| Dr. Daniel DOLFI<br>Thales R&T                                | Examineur             |
| Dr. Corinne FOURNIER<br>Université Jean Monnet                | Examinatrice          |
| Pr. Antoine ROUEFF<br>Université de Toulon                    | Directeur de thèse    |
| Dr. Frédéric GALLAND<br>CNRS                                  | Co-directeur de thèse |



# Affidavit

Je soussigné, Timothée Alain Marcos Justel, déclare par la présente que le travail présenté dans ce manuscrit est mon propre travail, réalisé sous la direction scientifique d'Antoine Roueff et de Frédéric Galland, dans le respect des principes d'honnêteté, d'intégrité et de responsabilité inhérents à la mission de recherche. Les travaux de recherche et la rédaction de ce manuscrit ont été réalisés dans le respect à la fois de la charte nationale de déontologie des métiers de la recherche et de la charte d'Aix-Marseille Université relative à la lutte contre le plagiat.

Ce travail n'a pas été précédemment soumis en France ou à l'étranger dans une version identique ou similaire à un organisme examinateur.

Fait à Marseille le 28/06/2023



Cette œuvre est mise à disposition selon les termes de la [Licence Creative Commons Attribution - Pas d'Utilisation Commerciale - Pas de Modification 4.0 International](https://creativecommons.org/licenses/by-nc-nd/4.0/).

# Liste de publications et participation aux conférences

## Liste des publications et/ou brevets réalisées dans le cadre du projet de thèse :

### Revues internationales avec comité de lecture :

- ▷ Timothée Justel, Frédéric Galland et Antoine Roueff. « Compressed Raman method combining classification and estimation of spectra with optimized binary filters ». Dans : *Opt. Lett.* 47.5 (mars 2022), p. 1101-1104.
- ▷ Timothée Justel, Frédéric Galland et Antoine Roueff. « Optimal trade-off filters for compressed Raman classification and spectrum reconstruction ». Dans : *J. Opt. Soc. Am. A* 40.6 (mai 2023), p. 1058-1065.

## Participation aux conférences et écoles d'été au cours de la période de thèse :

### Actes de colloques à comités de lectures :

- ▷ Timothée Justel, Frédéric Galland et Antoine Roueff. « Spectroscopie Raman comprimée pour la classification supervisée d'espèces : méthodologie pour la conception d'approches robustes. » Colloque GRETSI, Grenoble, France, 28 août au 1<sup>er</sup> septembre 2023.

### Communications à des congrès, symposiums :

- ▷ Timothée Justel, Frédéric Galland et Antoine Roueff. « Classification et reconstruction de spectres à partir de mesures Raman comprimées. » Journées Imagerie Optique Non Conventionnelle, Paris, France, 17-18 mars 2022.
- ▷ Timothée Justel, Frédéric Galland et Antoine Roueff. « Filtres à compromis optimaux pour la classification en spectroscopie Raman comprimée. » Journées Imagerie Optique Non Conventionnelle, Paris, France, 30-31 mars 2023.

### Séminaires, workshops :

- ▷ Timothée Justel, Frédéric Galland et Antoine Roueff. « A compressed Raman imaging approach for classification and reconstruction of spectra. » Workshop on new microscopies for cell biology, Paris, France, 13 décembre 2022.

# Résumé

La spectroscopie par diffusion Raman spontanée est une technique très répandue qui permet de remonter à la nature des liaisons chimiques présentes dans un échantillon sans nécessiter de marquage. Elle souffre de la faiblesse du rayonnement Raman, qui impose de longs temps d'acquisition lorsque les mesures sont effectuées avec un spectromètre conventionnel. Pour s'affranchir de cette limitation, des techniques dites comprimées ont été développées. Ces techniques consistent à n'effectuer que quelques mesures avec un petit nombre de détecteurs très sensibles et très rapides. Cette méthodologie est particulièrement efficace si l'on dispose de connaissances a priori sur la composition chimique de l'échantillon analysé. On peut en effet les utiliser pour déterminer les fréquences utiles à la réalisation de la tâche ciblée et filtrer le rayonnement pour ne mesurer que ces fréquences.

L'objectif de cette thèse est d'adapter ces techniques afin de pouvoir questionner la validité des a priori. Pour ce faire, nous proposons de répartir une diversité de jeux de filtres sur les pixels de la zone à imager. Dans le cas de pixels purs, cela permet non seulement de classer chaque pixel parmi des espèces chimiques de spectres connus mais aussi de reconstruire les spectres rencontrés en combinant des mesures réalisées sur plusieurs pixels. Dans le cas de pixels où apparaissent des mélanges, cela permet à la fois d'estimer en chaque pixel les concentrations d'espèces chimiques de spectres connus et de réduire la part d'anomalies indétectables. La méthodologie proposée est fondée sur l'emploi de bornes statistiques pour caractériser les performances des jeux de filtres en chacune des tâches considérées. Une approche fondée sur des compromis optimaux nous permet de construire la diversité de jeux de filtres nécessaire et d'obtenir un aperçu global des performances atteignables pour une configuration donnée du système d'acquisition, ce qui permet de l'optimiser.

Mots clés : Traitement statistique des données, Théorie de l'information, Spectroscopie Raman comprimée, Borne de Cramér-Rao, Borne de Bhattacharyya, Compromis optimaux.

# Abstract

Spontaneous Raman spectroscopy is a widely used label-free analysis technique enabling to determine the chemical bonds appearing in a given sample. This technique suffers from the low intensity of Raman radiation, which implies large exposure times when measurements are performed with a conventional spectrometer. Compressed techniques were developed in order to overcome this limitation. They consist in performing few measurements with a limited number of high-sensitivity high-speed detectors. This methodology is especially efficient when a priori knowledge about the composition of the analysed sample is available. In fact, it is possible to use this knowledge to determine which frequencies are useful to achieve a given task and filter the radiation accordingly.

In this thesis, we build on these techniques to enable questioning the validity of the a priori knowledge. Specifically, we propose to distribute a diversity of filter sets onto the pixels of the area to be analysed. In the case of pure pixels, it enables one to classify each pixel amongst a list of chemical species of known spectra and also to reconstruct encountered spectra by aggregating measures acquired on different pixels. In the case of pixels where multiple chemical species are mixed together, it enables one to estimate the concentrations of chemical species of known spectra but also to reduce the set of undetectable anomalies. The proposed methodology is based on using statistical bounds to characterise the performance of filter sets in each of the considered tasks. We make use of an approach based on optimal trade-offs to build the necessary diversity of filter sets and to obtain an overview of the performance reachable with a given configuration of the acquisition device, thereby offering ways to optimise it.

Keywords: Statistical data analysis, Information theory, Compressed Raman spectroscopy, Cramér-Rao bound, Bhattacharyya bound, Optimal trade-offs.

# Remerciements

Je remercie tout d'abord François Goudail et Olivier Haeberlé pour avoir accepté de rapporter ce travail. Je remercie également Daniel Dolfi et Corinne Fournier pour avoir accepté de faire partie de mon jury de thèse, et Mehdi Alouini, pour l'avoir présidé.

Ce manuscrit a été comparé avec raison à un galet, c'est-à-dire à un objet très dense et poli, dont la surface n'offre par conséquent que peu d'aspérités auxquelles se raccrocher. La qualité des rapports et des questions qui m'ont été posées lors de la soutenance témoigne donc du soin important que tous les membres du jury ont accordé à leur lecture de ce travail, ce pour quoi je les remercie vivement.

Je remercie également mes directeurs de thèse, Antoine Roueff et Frédéric Galland, qui m'ont guidé pendant ces trois ans. J'ai pu malgré la distance compter sur le suivi régulier d'Antoine. Ses questions, promptes à exposer les failles ou à souligner l'intérêt d'une idée, ont été d'une grande aide. Frédéric s'est investi énormément dans cette thèse et m'a accompagné au plus proche de mon travail de recherche, ce qui a été très précieux. Nous avons développé ensemble de nombreuses idées.

J'ai pu avoir avec l'un comme l'autre de nombreux débats scientifiques vifs, instructifs et fructueux, ce qui a rendu cette thèse à la fois extrêmement stimulante et enrichissante. À l'issue de ces débats, seules les idées les plus convaincantes étaient conservées, sans considération pour leur provenance. Je remercie donc Antoine et Frédéric de la confiance qu'ils m'ont accordée, ainsi que pour m'avoir fait l'honneur de me traiter petit à petit plus comme un collègue que comme un élève.

Je suis reconnaissant aux autres membres de l'équipe  $\varphi$ TI, Nicolas Bertaux, Emmanuel Chevallier, Philippe Réfrégier et Muriel Roche, pour leur bon accueil. Leur soutien scientifique et humain a permis à ce travail de se dérouler dans les meilleures conditions possibles, et j'ai eu grand plaisir à travailler à leurs côtés.

Je souhaite en particulier remercier Philippe. Ses idées, systématiquement pertinentes, m'ont souvent permis d'avancer plus vite. Le partage de son expérience et de sa vision de la science en général a également été extrêmement instructif.

Il m'aurait été impossible d'aller aussi loin sans le soutien de mes amis et pairs centraliens Chiheb Daaloul et Ali Shakil. J'ai partagé avec le premier non seulement un logement mais aussi de nombreuses discussions mathématiques à des horaires improbables devant le tableau minuscule dudit logement. Avec le second, j'ai affronté les trajets en bus vers le laboratoire et la machine à café dudit laboratoire. Je n'aurais pu avoir de meilleurs compagnons pour parcourir le chemin d'étudiant ingénieur et de thésard. Et la route est encore longue!

Ils n'ont bien sûr pas été les seuls à m'accompagner toutes ces années, et il me faut

aussi remercier les autres, amis et collègues du laboratoire, qui le méritent mais que je ne peux tous citer ici.

Enfin, merci aux membres de ma famille d'avoir été là quand il le fallait, toujours fidèles à eux-mêmes. Parmi eux, il me faut remercier en particulier mes parents. C'est parce qu'ils m'ont montré la force du travail et appris que je devais étudier pour moi et pour personne d'autre que j'ai pu arriver jusqu'ici. Je leur en serai toujours reconnaissant.

# Table des matières

|   |           |
|---|-----------|
| Affidavit   | i         |
| Liste de publications et participation aux conférences  | ii        |
| Résumé  | iii       |
| Abstract  | iv        |
| Remerciements   | v         |
| Introduction  | 1         |
| <b>1 Méthodes comprimées en spectroscopie Raman</b>   | <b>5</b>  |
| 1.1 Spectroscopie par diffusion Raman spontanée . . . . .                                       | 6         |
| 1.1.1 Introduction à la diffusion Raman spontanée . . . . .                                     | 6         |
| 1.1.2 Dispositif de mesure du spectre Raman complet et limitations . . . . .                    | 7         |
| 1.2 Méthodes comprimées en spectroscopie Raman . . . . .  | 8         |
| 1.2.1 Méthodes non supervisées . . . . .  | 8         |
| 1.2.2 Méthodes supervisées . . . . .  | 10        |
| 1.3 Classification et estimation des concentrations en spectroscopie Raman comprimée . . . . .  | 11        |
| 1.3.1 Dispositif de mesure comprimé et modélisation . . . . .                                   | 12        |
| 1.3.2 Estimation des concentrations . . . . .   | 14        |
| 1.3.3 Classification des mesures parmi les espèces de référence . . . . .                       | 19        |
| 1.4 Conclusion et positionnement de la thèse . . . . .  | 23        |
| <b>2 Combiner classification et reconstruction de spectres en spectroscopie Raman comprimée</b> | <b>25</b> |
| 2.1 Modalités d'acquisition considérées . . . . .   | 26        |
| 2.1.1 Modélisation de l'échantillon . . . . .   | 26        |
| 2.1.2 Type de dispositif et stratégie d'acquisition . . . . .                                   | 26        |
| 2.2 Manque de robustesse des approches existantes . . . . .                                     | 27        |
| 2.3 Réaliser à la fois classification et reconstruction . . . . .                               | 30        |
| 2.3.1 Effectuer une diversité de mesures comprimées . . . . .                                   | 30        |
| 2.3.2 Stratégie d'acquisition permettant la reconstruction . . . . .                            | 30        |
| 2.3.3 Caractéristiques des filtres et reconstruction . . . . .                                  | 31        |
| 2.3.4 Filtres complémentaires et maîtrise du temps d'acquisition . . . . .                      | 33        |
| 2.3.5 Reconstruction du spectre observé . . . . .   | 34        |
| 2.4 Problème d'optimisation des filtres . . . . .   | 35        |
| 2.4.1 Critères à optimiser . . . . .  | 36        |

|          |   |           |
|----------|---|-----------|
| 2.4.2    | Optimisation et sélection d'une diversité de jeux de filtres . . . . .                                      | 37        |
| 2.5      | Approche fondée sur les compromis optimaux . . . . .  | 41        |
| 2.5.1    | Principe des filtres à compromis optimaux . . . . .   | 42        |
| 2.5.2    | Recherche des filtres à compromis optimaux . . . . .  | 42        |
| 2.5.3    | Algorithme de recherche des filtres à compromis optimaux . . . . .  | 44        |
| 2.5.4    | Mise en pratique de l'approche à compromis optimaux . . . . .   | 46        |
| 2.6      | Étude des performances du système d'acquisition . . . . .   | 51        |
| 2.6.1    | Influence de la forme des spectres des espèces de référence . . . . .                                       | 51        |
| 2.6.2    | Influence du nombre d'espèces de référence . . . . .  | 53        |
| 2.6.3    | Influence du nombre de jeux de filtres . . . . .  | 54        |
| 2.6.4    | Influence du nombre de photons à détecter . . . . .   | 55        |
| 2.6.5    | Influence du nombre de filtres . . . . .  | 56        |
| 2.6.6    | Influence du nombre de passages par pixel . . . . .   | 57        |
| 2.7      | Détection et caractérisation d'une anomalie . . . . .   | 61        |
| 2.7.1    | Méthodologie proposée et détection d'anomalie . . . . .   | 61        |
| 2.7.2    | Retour sur le premier exemple . . . . .   | 61        |
| 2.8      | Conclusion, limites et ouverture . . . . .  | 65        |
| <b>3</b> | <b>Combiner estimation des concentrations et tests d'inadéquation en spectroscopie Raman comprimée</b>      | <b>67</b> |
| 3.1      | Mélanges et stratégie d'acquisition . . . . .   | 68        |
| 3.1.1    | Rayonnement d'un mélange linéaire . . . . .   | 68        |
| 3.1.2    | Dispositif d'acquisition considéré . . . . .  | 68        |
| 3.2      | Détection des inadéquations . . . . .   | 69        |
| 3.2.1    | Ambiguïté à propos de l'origine du rayonnement . . . . .  | 69        |
| 3.2.2    | Conditions à vérifier en chaque pixel . . . . .   | 70        |
| 3.2.3    | Diversité de jeux de filtres et détection d'anomalies . . . . .   | 72        |
| 3.2.4    | Test d'inadéquation . . . . .   | 73        |
| 3.3      | Filtres à compromis optimaux entre estimation des concentrations et reconstruction de spectres . . . . .    | 75        |
| 3.3.1    | Critères à optimiser . . . . .  | 76        |
| 3.3.2    | Algorithme de recherche des compromis optimaux . . . . .  | 78        |
| 3.4      | Exemple d'application . . . . .   | 78        |
| 3.4.1    | Optimisation des jeux de filtres . . . . .  | 78        |
| 3.4.2    | Performances en détection des jeux de filtres sélectionnés . . . . .  | 80        |
| 3.4.3    | Estimation robuste des concentrations . . . . .   | 81        |
| 3.5      | Conclusion et ouverture . . . . .   | 85        |
|          | <b>Conclusion et perspectives</b>   | <b>87</b> |
|          | <b>Annexes</b>  | <b>91</b> |
| A        | Borne de Cramér-Rao de la reconstruction d'un spectre dans le cas d'une distribution multinomiale . . . . . | 92        |
| A.1      | Calcul de l'information de Fisher . . . . .   | 92        |
| A.2      | Inversibilité de la matrice d'information de Fisher . . . . .   | 93        |
| A.3      | Formule de Sherman et calcul de la matrice de Cramér-Rao . . . . .  | 96        |
| B        | Borne de Cramér-Rao de la reconstruction du spectre dans le cas d'un bruit de Poisson . . . . .             | 99        |
| B.1      | Calcul de l'information de Fisher . . . . .   | 99        |
| B.2      | Inversibilité de la matrice d'information de Fisher . . . . .   | 99        |

|   |     |  |     |
|---|-----|--|-----|
|   | B.3 | Formule de Sherman et calcul de la matrice de Cramér-Rao . . . . .       | 100 |
| C |     | Borne de Bhattacharyya . . . . .   | 102 |
|   | C.1 | Borne de Bhattacharyya à plusieurs jeux de filtres . . . . .             | 102 |
|   | C.2 | Borne de Bhattacharyya généralisée à plusieurs jeux de filtres . . . . . | 102 |
| D |     | Modification aléatoire de jeux de filtres . . . . .                      | 103 |
|   | D.1 | Modifications possibles en une itération . . . . .                       | 103 |
|   | D.2 | Tirage de la localisation de la modification . . . . .                   | 103 |
| E |     | Convexité de la courbe des compromis optimaux . . . . .                  | 104 |
| F |     | Conditions nécessaires à détecter les anomalies . . . . .                | 105 |
|   | F.1 | Notations spécifiques à la section . . . . .                             | 105 |
|   | F.2 | Décomposition des spectres normés . . . . .                              | 105 |
|   | F.3 | Démonstration de la proposition 3 . . . . .                              | 106 |
|   | F.4 | Démonstration de la proposition 4 . . . . .                              | 109 |

**Bibliographie**

**112**



# Introduction

La spectroscopie par diffusion Raman spontanée est très utilisée pour l'analyse chimique. Cette technique possède certaines caractéristiques très intéressantes pour l'expérimentateur. Premièrement, elle ne nécessite pas de marquage car la plupart des matériaux diffusent naturellement un rayonnement Raman lorsque soumis à un rayonnement monochromatique ultra-violet, visible ou proche infrarouge. Ainsi, ce type de spectroscopie ne nécessite pas ou peu de préparation des échantillons à analyser. Il est de plus possible de choisir des longueurs d'ondes d'excitation qui ne sont pas ou peu absorbées par l'eau, contrairement par exemple à la spectroscopie infrarouge. Cela permet de travailler facilement avec des échantillons biologiques ou en milieu aqueux. L'analyse du rayonnement Raman permet de remonter à des informations précises à propos de la composition de l'échantillon analysé, notamment la nature des liaisons chimiques présentes. Chaque composé chimique est ainsi identifiable par son spectre Raman.

Malgré ces avantages, son utilisation est limitée par la faiblesse de la diffusion Raman spontanée. Cette diffusion produit en effet un rayonnement de très faible intensité comparativement au rayonnement incident. De ce fait, le rayonnement Raman peut être noyé dans la fluorescence de l'échantillon souvent suscitée dans le même temps par l'excitation laser et dont l'intensité est plus forte. De plus, la faiblesse de ce signal rend la mesure de son spectre difficile. Dans un spectromètre conventionnel, le rayonnement à analyser est dispersé par un composant diffractif sur une barrette de détecteurs. Avec un rayonnement Raman, la faible quantité de signal reçue par chaque élément de la barrette après dispersion couplée avec le bruit électronique de cette barrette peut suffire à rendre les mesures inexploitable. Mesurer un spectre Raman avec un spectromètre conventionnel nécessite donc de longs temps de pose. Ces temps de pose sont incompatibles avec certaines applications, telle que le suivi de réactions chimiques. Plus généralement, il est difficile d'imager des échantillons dynamiques. De plus, une irradiation prolongée peut endommager les échantillons sensibles à la lumière, comme certains échantillons biologiques. On peut ajouter à cela qu'effectuer la mesure de la totalité du spectre du rayonnement en chaque pixel d'une image produit une quantité massive de données, ce qui demande une grande capacité de stockage.

Pour résoudre ces problèmes, des approches de spectroscopie Raman comprimée ont été proposées à partir des années 1990. Ces approches consistent à réduire le nombre de mesures effectuées dans le domaine spatial ou spectral, ce qui peut permettre de n'utiliser que quelques détecteurs, que l'on peut alors choisir plus sensibles et plus rapides qu'une barrette de détecteurs standard. Il devient alors possible de travailler avec des temps de mesure très courts ou à faible flux de photons. La quantité de données acquises est bien plus limitée qu'en spectroscopie conventionnelle, même si le traitement de ces données n'est pas forcément plus léger, dépendamment de la méthode utilisée.

Cette méthodologie est particulièrement efficace si l'on dispose d'a priori sur la composition chimique de l'échantillon analysé, notamment si on connaît les spectres des espèces chimiques le composant. En effet, il est alors possible d'utiliser ces a priori pour sélectionner uniquement les bandes spectrales les plus pertinentes pour la tâche considérée. Le rayonnement est alors

filtré et on ne fait l'acquisition que des combinaisons de fréquences d'intérêt. Ces méthodes comprimées supervisées permettent de réduire encore plus les temps de mesure, étendant le champ des applications possibles de la spectroscopie Raman. La sélection des fréquences à mesurer est équivalente à une compression, qui est en général avec pertes. Il faut donc sélectionner avec soin la manière dont cette compression est effectuée, sans quoi la perte de performance résultant réduirait à néant l'intérêt de l'approche comprimée. C'est ce problème que traitent les méthodes de la littérature.

Ces méthodes proposent en général d'utiliser systématiquement un même jeu de filtres pour effectuer les mesures, spécialisé au préalable pour la tâche et les a priori considérés. Comme nous le montrerons dans ce manuscrit, cette stratégie implique d'être aveugle à la plupart des écarts au modèle. Or, il est particulièrement important de les détecter puisque leur présence peut induire des erreurs dans les résultats de l'analyse des mesures. Un écart à la composition supposée peut également signifier une pollution de l'échantillon analysé par un composé imprévu, ce qui constitue un problème qu'il faut détecter et si possible caractériser.

## Objectifs et plan de la thèse

L'objectif de cette thèse est de développer des techniques permettant de réaliser de manière robuste la classification des mesures de spectroscopie Raman comprimée parmi une liste d'espèces chimiques de spectres connus ou l'estimation des concentrations d'espèces chimiques de spectres connus. Autrement dit, l'objectif est de construire des filtres permettant à la fois de réaliser la tâche considérée avec de bonnes performances si l'a priori sur la composition chimique de l'échantillon est vérifié, mais aussi donner la possibilité de détecter les écarts au modèle et de remonter à leur origine lorsque cela est possible.

Cette thèse se structure en trois chapitres. Dans le chapitre 1, nous décrivons l'état de l'art des méthodes de spectroscopie Raman comprimée supervisées et non supervisées. Nous décrivons ensuite succinctement les principes physiques de la diffusion Raman spontanée, ce qu'est un spectre Raman et comment le mesurer. Nous présentons également la modélisation statistique du système d'acquisition proposée dans la littérature, qui est celle que nous adoptons. Enfin, nous revenons en détails sur les techniques supervisées permettant la classification et l'estimation des concentrations à partir de mesures de spectroscopie Raman comprimée. Ces techniques constituent les fondements de ce travail de thèse.

Dans le chapitre 2, nous nous intéressons au problème de la classification des mesures comprimées parmi des espèces chimiques de spectres connus. Nous commençons par questionner la robustesse des méthodes de la littérature. Afin de résoudre le problème de robustesse décelé, nous proposons de construire des filtres permettant de reconstruire les spectres rencontrés à partir de mesures comprimées réalisées sur plusieurs pixels, ce qui donne la possibilité de détecter et d'identifier les anomalies sans mesure supplémentaire. Nous utilisons des bornes statistiques pour caractériser de manière rigoureuse les performances des filtres en classification et en reconstruction. Ces deux tâches étant antagonistes, une partie du chapitre s'attache à comment construire des filtres permettant de les concilier. Enfin, nous employons la méthodologie proposée pour réaliser une analyse complète des performances du système d'acquisition en fonction des différents paramètres.

Dans le chapitre 3, nous nous intéressons à comment estimer de manière robuste les concentrations d'espèces chimiques de spectres connus à partir de mesures comprimées. Les techniques développées dans le chapitre précédent pour la détection d'anomalies ne sont pas adaptées à des échantillons constitués de mélanges et il faut traiter ce cas séparément. Ce problème est beaucoup plus difficile car on ne fait aucune hypothèse sur les concentrations. Nous commençons par rechercher les conditions sous lesquelles il est possible de détecter les

anomalies rencontrées. Cette étude aboutit à des contraintes sur la forme des filtres qu'il est nécessaire d'appliquer pour limiter le nombre d'anomalies indétectables. Nous étendons alors les méthodes développées dans le second chapitre pour construire des filtres permettant de concilier estimation des concentrations et détection des anomalies.

Enfin, nous concluons en discutant des résultats principaux et des perspectives à ce travail.

Plusieurs annexes complètent ce manuscrit. Les annexes A, B et C traitent du calcul des bornes statistiques utilisées pour caractériser les performances des jeux de filtres employés, et peuvent être ignorées en première lecture. L'annexe D fournit certains détails techniques sur les procédures d'optimisation développées dans ce manuscrit, et il peut être intéressant de la consulter au moment où celles-ci sont développées. L'annexe E démontre certaines propriétés concernant les filtres construits dans le second chapitre. Enfin, l'annexe F fournit les preuves des propositions du troisième chapitre qui concernent la détection d'anomalies.



# Chapitre 1

## Méthodes comprimées en spectroscopie Raman

### Sommaire

---

|   |           |
|---|-----------|
| <b>1.1 Spectroscopie par diffusion Raman spontanée . . . . .</b>                                      | <b>6</b>  |
| 1.1.1 Introduction à la diffusion Raman spontanée . . . . .   | 6         |
| 1.1.2 Dispositif de mesure du spectre Raman complet et limitations . . .                              | 7         |
| <b>1.2 Méthodes comprimées en spectroscopie Raman . . . . .</b>                                       | <b>8</b>  |
| 1.2.1 Méthodes non supervisées . . . . .  | 8         |
| 1.2.2 Méthodes supervisées . . . . .  | 10        |
| <b>1.3 Classification et estimation des concentrations en spectroscopie Raman comprimée . . . . .</b> | <b>11</b> |
| 1.3.1 Dispositif de mesure comprimé et modélisation . . . . .   | 12        |
| 1.3.2 Estimation des concentrations . . . . .   | 14        |
| 1.3.3 Classification des mesures parmi les espèces de référence . . . . .                             | 19        |
| <b>1.4 Conclusion et positionnement de la thèse . . . . .</b>   | <b>23</b> |

---

## 1.1 Spectroscopie par diffusion Raman spontanée

### 1.1.1 Introduction à la diffusion Raman spontanée

Le contexte de cette thèse est la spectroscopie par diffusion Raman spontanée. Cette technique d'analyse se fonde sur l'effet Raman, un phénomène de diffusion inélastique des photons prédit par Smekal (1923) et constaté expérimentalement pour la première fois par Raman (1928) ainsi que par Landsberg et Mandelstam (1928) de manière indépendante.

Cette thèse ne porte pas sur la physique associée à la spectroscopie Raman, aussi nous ne présentons ici qu'une description très succincte de ce qu'est l'effet Raman. Nous renvoyons par exemple le lecteur vers Schrader (1995) pour un traitement plus détaillé ou vers Rigneault et Berto (2018) pour un tutoriel complet sur le sujet.

Lorsqu'un échantillon est excité par un laser, il diffuse une partie de ce rayonnement. La diffusion majoritaire est élastique, c'est-à-dire que les photons diffusés ont en général la même énergie que les photons incidents. Cette diffusion est appelée la diffusion *Rayleigh*. Son intensité peut atteindre un millième de l'intensité incidente. Il se produit également une diffusion inélastique, dont l'intensité est environ mille fois moins importante. Lorsqu'une partie de l'énergie des photons incidents est transférée aux molécules rencontrées, les photons diffusés sont d'énergie plus faible que les photons incidents. Ces photons sont appelés photons *Stokes*. À l'inverse, il est possible qu'une partie de l'énergie des molécules rencontrées soit transférée aux photons diffusés, qui sont alors de plus forte énergie que les photons incidents. Ces photons sont appelés photons *anti-Stokes*. La proportion de photons anti-Stokes dans les photons diffusés de manière inélastique croît avec la température, les photons Stokes étant largement majoritaires à température ambiante.

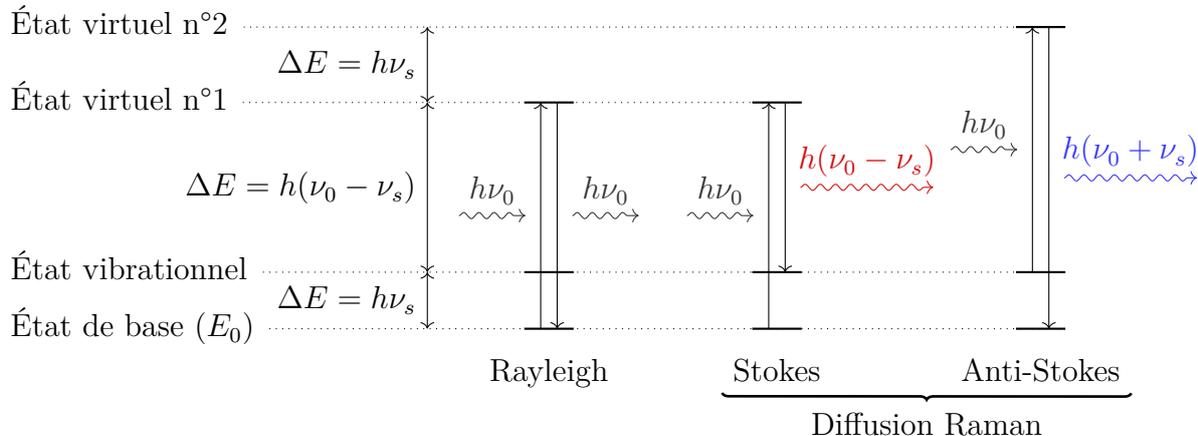


FIGURE 1.1 – Principe de l'effet Raman.

On modélise souvent ces interactions comme présenté dans la figure 1.1. Si la molécule éclairée est dans son état de base d'énergie  $E_0$  une collision avec un photon d'énergie  $h\nu_0$  l'élève momentanément dans un état virtuel situé à  $E_0 + h\nu_0$ . Elle se désexcite en général en émettant un photon de même énergie. Il arrive cependant que la molécule ne se désexcite pas entièrement, redescendant plutôt dans un état vibrationnel situé à  $E_0 + h\nu_s$ . Le photon émis lors de la désexcitation est alors d'énergie  $h(\nu_0 - \nu_s)$ . Si la molécule est déjà dans un état vibrationnel localisé à  $E_0 + h\nu_s$  lors de la collision avec un photon d'énergie  $h\nu_0$  alors cette collision la fait passer dans un état situé à  $E_0 + h(\nu_0 + \nu_s)$ . Si elle redescend directement dans son état de base après la collision alors le photon diffusé est d'énergie  $h(\nu_0 + \nu_s)$ .

On appelle *rayonnement Raman* le rayonnement issu de la diffusion inélastique. Le spectre de ce rayonnement est un spectre de raies, et on le représente souvent en traçant les décalages

en fréquence (ou de manière équivalente en nombre d'onde) par rapport à la longueur d'onde incidente, exprimés en  $\text{cm}^{-1}$ . Ces décalages en fréquence étant caractéristiques des liaisons chimiques présentes dans l'échantillon analysé, étudier le spectre Raman permet de caractériser la composition chimique des échantillons observés.

**Atouts de l'analyse du rayonnement Raman.** L'effet Raman apparaît sans nécessité de marquage. Il n'est donc pas nécessaire d'intervenir sur les échantillons à analyser, ce qui serait réhibitoire pour certaines applications. De plus, il est possible de choisir des longueurs d'onde d'excitation dans le spectre du visible, et donc peu absorbées par l'eau. Cela permet ainsi de travailler facilement avec des échantillons contenant de l'eau, et en particulier avec des échantillons d'origine biologique. Cette caractéristique explique sa complémentarité avec la spectroscopie infrarouge, qui permet de collecter des informations assez similaires à propos de la composition chimique d'un échantillon mais dont le rayonnement excitateur est absorbé en quasi-totalité par l'eau.

**Obstacles à l'analyse du rayonnement Raman.** La faible intensité du rayonnement Raman rend son analyse difficile. La section efficace de la diffusion Raman spontanée est de l'ordre de  $10^{-30} \text{ cm}^2$ . Cela peut être interprété de la manière suivante : si on considère que la surface des molécules illuminées est de l'ordre de  $10^{-24} \text{ cm}^2$  alors il y a une probabilité de  $10^{-30}/10^{-24} = 10^{-6}$  qu'un photon incident à une molécule soit diffusé de manière inélastique. De plus, l'excitation de l'échantillon par le laser entraîne souvent une fluorescence. Celle-ci est de forte intensité par rapport au rayonnement Raman et peut noyer sa contribution.

### 1.1.2 Dispositif de mesure du spectre Raman complet et limitations

En spectroscopie par diffusion Raman spontanée, l'échantillon à analyser est excité avec un laser monochromatique. L'approche la plus couramment utilisée pour analyser le rayonnement Raman est de le séparer du faisceau excitateur et du rayonnement Rayleigh grâce à un filtre en fréquence et de l'envoyer sur un composant diffractif, généralement un réseau de diffraction. Ce dernier transpose les fréquences lumineuses dans le domaine spatial, et le spectre entier est typiquement imagé sur une barrette de détecteurs. Le schéma d'un tel dispositif apparaît en figure 1.2.

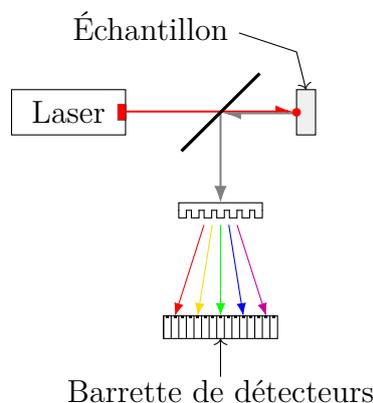


FIGURE 1.2 – Principe d'un spectromètre Raman conventionnel.

Lorsque le rayonnement diffracté est dispersé sur une barrette de détecteurs, chaque composant de cette barrette reçoit un très faible flux de photon, et le signal ainsi détecté peut être noyé dans le bruit électronique de la barrette. Pour pallier ce problème, il est nécessaire

d'effectuer de longues acquisitions. Ainsi, il est en général impossible de passer moins d'une milliseconde sur un point donné avec un spectromètre Raman conventionnel (voir Scotté (2020, section 3.2.2)). Cela rend ce dispositif incapable d'imager des échantillons présentant des dynamiques rapides, par exemple des spécimens vivants.

Pour augmenter la quantité de signal Raman produit par l'échantillon analysé, il est possible de changer de processus d'excitation ou de modifier l'échantillon directement. Plusieurs alternatives à la spectroscopie Raman conventionnelle basées sur ces principes existent. On peut évoquer la spectroscopie Raman stimulée (SRS, voir Ploetz, Laimgruber et al. (2007) et Freudiger, Min et al. (2008)), la spectroscopie Raman anti-Stokes cohérente (CARS, voir Zumbusch, Holtom et al. (1999) et Volkmer, Cheng et al. (2001)) ou la spectroscopie Raman par exaltation de surface (SERS, voir Félidj, Aubard et al. (2003)). Dans les deux premiers cas, la puissance du rayonnement diffusé est accrue en utilisant un ou plusieurs faisceaux pompe. Dans le dernier cas, elle est augmentée en introduisant dans l'échantillon des nanoparticules métalliques dans le but de produire une résonance. Ces dispositifs offrent des possibilités intéressantes mais sont toutefois coûteux et complexes à mettre en œuvre.

## 1.2 Méthodes comprimées en spectroscopie Raman

L'objectif de cette section est de donner un état de l'art des méthodes comprimées en spectroscopie Raman. Nous séparons ces méthodes en deux catégories. La première est celle des méthodes non supervisées, qui n'utilisent pas ou peu d'a priori et qui visent en général à reconstruire les spectres observés. La seconde est celle des méthodes supervisées, qui utilisent un fort a priori sur la composition de l'échantillon analysé pour effectuer la tâche considérée sans avoir besoin d'avoir accès à la totalité du spectre.

Comme expliqué dans la section précédente, la faible intensité du spectre Raman rend son analyse difficile. En effet, le bruit de lecture d'une barrette de détecteurs standard peut suffire à rendre les mesures inexploitable, en particulier lorsque le flux de photons est très faible. Pour s'affranchir de cette limitation il est possible d'utiliser des détecteurs à comptage de photons, plus sensibles et plus rapides. Ces détecteurs ne comportent cependant en général qu'un unique canal et leur volume ainsi que leur coût ne permettent pas d'envisager d'en employer plus de quelques unités à la fois. Il est alors impossible de mesurer un spectre entier avec une bonne résolution en une seule fois et il faut adapter la stratégie d'acquisition et d'analyse des mesures.

### 1.2.1 Méthodes non supervisées

Parmi les premiers travaux à proposer une alternative au spectromètre conventionnel, Wagner, B. W. Smith et al. (1995) développent un concept de spectromètre employant une matrice de micro-miroirs pour filtrer le rayonnement analysé afin de faire l'acquisition des fréquences une à une sur un tube photomultiplicateur. Ce détecteur très sensible peut permettre d'améliorer la qualité du spectre reconstruit par rapport au cas où une barrette de détecteurs standard est utilisée. La perte d'énergie est toutefois importante puisque toutes les fréquences du rayonnement sauf une sont bloquées à chaque mesure. Le gain par rapport à un spectromètre conventionnel n'est donc pas toujours évident.

Pour résoudre ce problème, DeVerse, Hammaker et al. (2000) proposent d'utiliser un dispositif similaire pour mesurer les projections du spectre observé sur les vecteurs de la base de Hadamard. Chacun des filtres correspondant laisse passer la moitié des fréquences, ce qui permet d'augmenter grandement la quantité de photons détectés à chaque mesure. Lorsque les projections sur tous les vecteurs de la base ont été mesurées, il est possible de reconstruire

le spectre observé en inversant le système. Cette technique permet d'améliorer grandement le rapport signal-à-bruit de la reconstruction. Elle requiert cependant de réaliser autant de mesures que de composantes du spectre et il est difficile d'envisager de l'appliquer en chaque pixel d'une image, car cela serait très chronophage.

Les travaux sur l'acquisition comprimée, notamment E. Candès, J. Romberg et al. (2006), E. J. Candès, J. K. Romberg et al. (2006) et E. J. Candès et Tao (2006) permettent de dépasser cette limitation. Ces travaux montrent qu'il est possible de reconstruire un signal parcimonieux ou approximativement parcimonieux avec un nombre de mesures inférieur à la taille de ce signal. Plus précisément, ils montrent qu'une reconstruction exacte ou avec une erreur proportionnelle au bruit est possible avec un nombre de mesures de l'ordre du nombre de coefficients non nuls dans le signal à reconstruire. La matrice de mesure, c'est-à-dire la matrice représentant la projection du signal observé sur les détecteurs, doit cependant posséder une propriété d'isométrie restreinte (RIP, Restricted Isometry Property). S'il est très difficile de vérifier qu'une matrice donnée possède cette propriété, ces travaux démontrent également que c'est presque sûrement le cas des matrices de mesure aléatoires. Ces résultats sont particulièrement intéressants pour la spectroscopie Raman où il est très avantageux de limiter le nombre de mesures, qui peuvent nécessiter un long temps de pose.

Il est fréquent que l'information spatiale contenue dans une image soit fortement redondante, du fait d'une certaine homogénéité ou de variations lentes de son contenu. Il est alors possible de n'en mesurer qu'une partie. Ainsi, Galvis-Carreño, Mejia-Melgarejo et al. (2014) proposent de sous-échantillonner spatialement une image (hyperspectrale) Raman et de la compléter avec des algorithmes de reconstruction comprimée. Dans Pavillon et N. I. Smith (2016) les auteurs implémentent cette technique en programmant un spectromètre pour ne parcourir que certains pixels de la zone à imager. En chacun de ces pixels ils utilisent une barrette de détecteurs pour faire l'acquisition de toutes les fréquences. L'image complète est reconstruite grâce à une minimisation de la variation totale, de façon à imposer une contrainte sur l'évolution spatiale des spectres. La redondance semble compenser le manque de signal mais l'utilisation d'une barrette limite fondamentalement la rapidité de cette approche du fait de la faible fréquence de lecture. Suivant des principes similaires, Thompson, Bixler et al. (2017) proposent de faire l'acquisition d'une image entière, filtrée spatialement par un masque aléatoire, sur un spectromètre possédant une matrice de détecteurs (la seconde dimension spatiale étant alors mélangée à la dimension spectrale). Le cube hyperspectral complet est reconstruit avec des algorithmes basés sur la minimisation de la variation totale. Ces algorithmes nécessitent néanmoins une très grande puissance de calcul.

Puisque le spectre Raman est un spectre de raies, et donc parcimonieux, il est également possible de sous-échantillonner les mesures en fréquence. Se basant sur ce principe, Berto, Scotté et al. (2017) proposent de ne mesurer en chaque pixel d'une image que les projections du spectre observé sur des sous-ensembles aléatoires des vecteurs de la base de Hadamard. De cette manière, il est possible de profiter des avantages du multiplexage en plus de la compression. Le spectre total est reconstruit par une minimisation de la norme  $\ell_1$ .

Il est également possible de combiner les deux approches, comme dans Soldevila, Dong et al. (2019). Dans ces travaux, les auteurs proposent à la fois de ne scanner que certains pixels, répartis de manière aléatoire dans l'image, et de ne mesurer en chacun de ces pixels qu'un nombre réduit de projections aléatoires du spectre observé. Ils utilisent ensuite la théorie de la factorisation de matrices non négatives pour reconstruire le cube hyperspectral.

### 1.2.2 Méthodes supervisées

Il n'est pas systématiquement nécessaire de mesurer ou même de reconstruire la totalité du spectre du rayonnement observé. Si l'on dispose d'informations a priori à propos des espèces en présence dans l'échantillon analysé, il devient possible d'optimiser les mesures pour ne collecter que l'information utile à la tâche que l'on veut réaliser. Le nombre de mesures peut être drastiquement réduit dans certains cas. Ce principe est celui sur lequel les méthodes comprimées supervisées sont fondées.

Un travail pionnier sur les méthodes comprimées supervisées en spectroscopie Raman est celui de Nelson, Aust et al. (1998). Dans cet article, les auteurs proposent d'utiliser des techniques d'analyse multivariée pour construire une paire de filtres optiques permettant de sélectionner des combinaisons de fréquences caractéristiques du spectre d'un composé chimique. Plus précisément, ces filtres permettent de projeter le spectre observé sur les détecteurs selon la première direction principale du spectre de l'espèce étudiée (les parties positives et négatives de ce vecteur correspondant chacune à un des filtres de la paire). Ils montrent sur des données existantes que cela permet de suivre la concentration de l'espèce étudiée dans un mélange en fonction du temps, en ne réalisant ainsi que deux mesures à chaque instant de l'expérience. Une implémentation de cette approche sur un système réel est présentée par O. Soyemi, Eastwood et al. (2001), qui introduisent le terme MOE (Multivariate Optical Element) pour désigner les filtres construits avec cette méthode, appelée MOC (Multivariate Optical Computing). O. O. Soyemi, Haibach et al. (2002) développent de manière plus approfondie les algorithmes permettant de construire de tels filtres

Les travaux de Uzunbajakava, Peinder et al. (2006) poussent plus loin ce concept en concevant un spectromètre Raman capable d'appliquer des éléments optiques multivariés de manière dynamique avec un modulateur à cristaux liquides, là où les travaux précédents utilisaient des filtres à base de couches minces, par nature statiques. Les filtres ainsi construits permettent de suivre les concentrations de plusieurs composés chimiques dans un mélange, avec une paire de filtres spécifique à chaque composé. La méthode est également implémentée par Davis, Hemphill et al. (2011) qui combinent une méthode non supervisée pour reconstruire les spectres de référence puis utilisent des filtres optiques multivariés pour faire de la détection ou de l'estimation des concentrations.

Dans le même cadre, Quyen, Jouan et al. (2008) et Quyen, Silva et al. (2008) proposent de réaliser un spectromètre intégrant une matrice de micro-miroirs, utilisée pour implémenter des filtres binaires pour faire la sélection des fréquences. La construction de ces filtres est plus simple que dans les précédents travaux puisqu'il s'agit simplement de sélectionner les pics discriminants des espèces considérées. Love (2009) étend ce principe en proposant d'utiliser un filtrage adapté. Z. J. Smith, Strombom et al. (2011) proposent d'appliquer des filtres permettant de retrouver les concentrations directement à partir des mesures, sans traitement. Corcoran (2018) propose d'utiliser des combinaisons de fonctions de Walsh-Hadamard adaptées aux spectres des espèces de référence.

Dans les travaux mentionnés jusqu'à présent, beaucoup de questions restaient ouvertes concernant le choix des filtres. Wilcox, Buzzard et al. (2012) développent une nouvelle manière de construire des filtres adaptés à l'estimation des concentrations. Les auteurs proposent d'optimiser numériquement le filtrage afin de minimiser l'erreur quadratique moyenne effectuée en estimant les concentrations des espèces de référence avec un certain estimateur. Ils emploient ensuite les concentrations estimées en chaque pixel de l'image pour classer ces pixels parmi les espèces de référence. Ils montrent que cela peut être fait avec des temps de mesure très courts. Wilcox, Buzzard et al. (2013) se basent sur le même principe et proposent d'utiliser des méthodes de résolution multivariée de courbes (MCR, Multivariate Curve Resolution) pour obtenir les spectres de référence à partir de mesures d'un mélange, ce

qui permet de traiter le cas où il n'est pas possible d'obtenir des composés purs sur lesquels mesurer directement ces spectres.

Réfrégier, Scotté et al. (2018) proposent d'utiliser la borne de Cramér-Rao pour optimiser les filtres pour l'estimation des concentrations, ce qui permet également de caractériser de manière rigoureuse les performances atteignables avec ce type d'approche et de les comparer avec celles des approches conventionnelles. Cette méthode a été implémentée sur un système réel par Scotté, Aguiar et al. (2018). Une autre contribution de Réfrégier, Scotté et al. (2018) est la proposition d'utiliser des filtres orthogonaux appliqués en parallèle pour gagner en temps d'acquisition. Ce concept a également été proposé par Rehrauer, Dinh et al. (2018).

Un problème souvent rencontré dans l'application de ces approches est la présence d'un fond de fluorescence de forte intensité, qui s'ajoute au rayonnement Raman. Afin de mitiger ce problème Rehrauer, Mankani et al. (2015) proposent de modéliser ce fond comme un polynôme de degré fixé et de faire l'estimation des coefficients de ce polynôme en estimant les concentrations des pseudo-spectres correspondant aux différents monômes grâce à l'approche de Wilcox, Buzzard et al. (2012).

Toutes les approches présentées jusqu'ici supposent d'utiliser autant de filtres que d'espèces de référence. Réfrégier et Galland (2019) démontrent que ce n'est pas nécessaire dans le cas où l'objectif n'est plus d'estimer les concentrations en chaque pixel mais de classer ces pixels parmi les espèces de référence. Le nombre de mesures peut alors être réduit. Les auteurs proposent également de construire des filtres minimisant la borne de Bhattacharyya, qui est une borne supérieure à la probabilité d'erreur de classification. Se basant sur principe, Réfrégier, Chevallier et al. (2019) proposent de réaliser l'acquisition à nombre de photons détectés fixé, ce qui permet de garantir un niveau d'erreur indépendant de l'intensité du rayonnement.

Il est possible d'améliorer encore le dispositif d'acquisition. Ainsi, dans Scotté, Sivankutty et al. (2020) l'acquisition de la zone à imagée est effectuée ligne par ligne avec un encodage spatio-spectral. Plus précisément, le rayonnement provenant de chaque ligne de l'image est intégré sur un détecteur ponctuel après l'application d'un masque spatial correspondant à un vecteur de la base de Hadamard et un masque spectral optimisé pour l'estimation des concentrations comme dans Réfrégier, Scotté et al. (2018). L'application d'un masque spatio-spectral est rendue possible par la nature 2D de la matrice de micro-miroirs, jusque-là jamais exploitée. Lorsque tous les vecteurs de la base de Hadamard ont été mesurés il est possible de reconstruire les concentrations associées à chaque pixel de la ligne correspondant. On bénéficie ainsi de la quantité de signal du multiplexage en plus de la spécificité des filtres.

### 1.3 Classification et estimation des concentrations en spectroscopie Raman comprimée

Comme nous avons pu le voir dans la section précédente, adopter une approche comprimée peut être particulièrement intéressant dans le cadre de la spectroscopie Raman, notamment lorsque l'on dispose de forts a priori sur les espèces présentes dans l'échantillon analysé. Cette approche peut en effet permettre de réduire significativement le nombre de mesures réalisées, et donc le temps d'acquisition.

Dans cette section, nous étudions plus en détails les systèmes Raman comprimés et décrivons de manière plus approfondie certaines des méthodes présentées dans la section précédente permettant de classer des mesures comprimées parmi des espèces chimiques de spectres connus ou d'estimer à partir de mesures comprimées les concentrations d'espèces chimiques de spectres connus dans un mélange. Ces méthodes constituent les fondements de

cette thèse.

L'idée principale de ces méthodes est de ne mesurer que certaines bandes spectrales sélectionnées en fonction des hypothèses effectuées pour leur pertinence vis-à-vis de la tâche considérée. Cette sélection est mise en œuvre grâce à un filtrage en fréquence réalisé en amont de quelques détecteurs ponctuels, vers lesquels ne sont redirigées que certaines combinaisons de fréquences d'intérêt.

Afin de visualiser ce concept et de mettre en avant la différence avec la spectroscopie conventionnelle, la figure 1.3 compare le schéma d'un montage conventionnel (à gauche) à celui d'un montage comprimé appliquant séquentiellement au rayonnement plusieurs filtres fréquentiels différents (à droite).

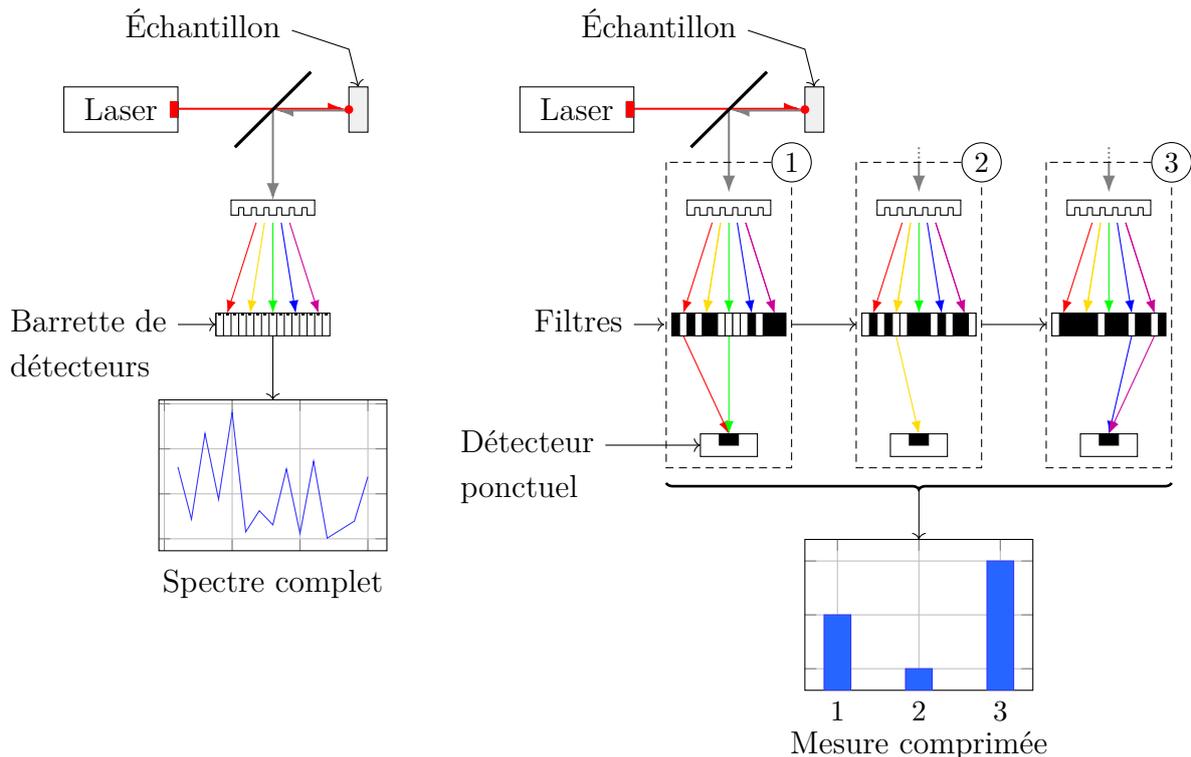


FIGURE 1.3 – Comparaison entre spectroscopie conventionnelle et spectroscopie comprimée.

Si le résultat de l'acquisition réalisée avec un montage conventionnel est directement lié au spectre du rayonnement, celui d'une acquisition réalisée avec un montage comprimé en fréquence n'est constitué que des projections du spectre observé sur les filtres appliqués. L'un des enjeux des méthodes comprimées est donc d'optimiser les filtres de sorte à ce que le signal soit projeté sur les détecteurs de manière à conserver l'information en fréquence pertinente pour la tâche à réaliser.

### 1.3.1 Dispositif de mesure comprimé et modélisation

Nous présentons dans cette sous-section la modélisation du processus d'acquisition qui est utilisée dans ce manuscrit. Cette modélisation correspond à celle qui est en général adoptée dans la littérature. Nous décrivons dans un premier temps le rayonnement Raman comme un ensemble de processus stochastiques. Dans un second temps, nous présentons le dispositif d'acquisition lui-même, et plus précisément l'effet du filtrage sur le rayonnement. Nous donnons alors la distribution des observations effectuées sur les différents détecteurs dans le cas où le temps de mesure est fixé.

Comme proposé par Wilcox, Buzzard et al. (2012, 2013), Réfrégier, Scotté et al. (2018), Réfrégier, Chevallier et al. (2019) et Réfrégier et Galland (2019), la diffusion du rayonnement Raman par l'échantillon analysé peut être modélisée par des processus de Poisson. Si on décompose le spectre d'un rayonnement en  $K$  bandes spectrales d'intérêt disjointes, le nombre de photons émis dans chacune de ces bandes spectrales suit une loi de Poisson de flux  $\phi(k)$ . Cela signifie que pour une durée d'émission  $\tau$  la probabilité d'observer  $n(k)$  photons dans la  $k^{\text{ème}}$  bande spectrale est

$$\mathcal{P}(n(k)) = \frac{(\tau\phi(k))^{n(k)}}{n(k)!} \exp(-\tau\phi(k)), \quad (1.1)$$

où  $n(k)!$  désigne la factorielle de  $n(k)$ . Le flux  $\phi(k)$  est le nombre moyen de photons diffusés par unité de temps dans la  $k^{\text{ème}}$  bande spectrale. La distribution du nombre de photons diffusés pendant un temps  $\tau$  est entièrement déterminée par la connaissance de ces flux.

L'intensité de ce rayonnement dépend en particulier de la quantité de matière dans le champ de vue. Autrement dit, la nature du composé chimique ne fixe que la densité spectrale du rayonnement et non son intensité. Dans ce manuscrit nous appelons *spectre* d'une espèce sa densité spectrale. Un tel spectre identifie de manière unique l'espèce chimique y étant associée. Pour refléter cette décomposition, il est possible d'écrire les flux sous la forme du produit d'une intensité  $\gamma$  et des composantes d'un spectre normalisé  $\mathbf{s} = {}^t(s(1), \dots, s(K))$  où  ${}^t \cdot$  est la transposition, c'est-à-dire  $\phi(k) = \gamma s(k)$  avec  $\sum_{k=1}^K s(k) = 1$  ou de manière équivalente  $s(k) = \phi(k) / \sum_{k=1}^K \phi(k)$  et  $\gamma = \sum_{k=1}^K \phi(k)$ .

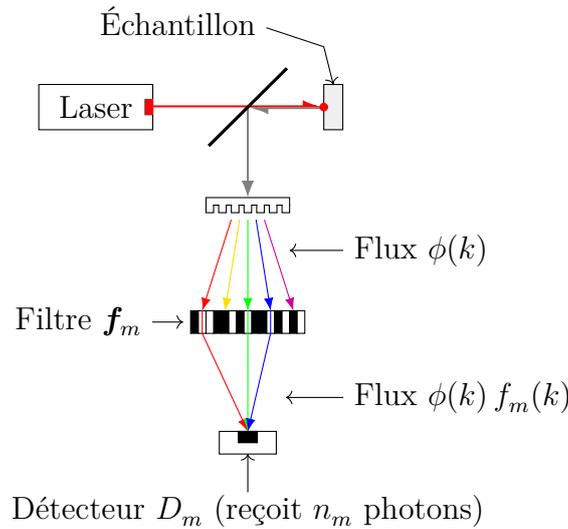


FIGURE 1.4 – Modélisation du filtrage effectué par un spectromètre Raman comprimé.

Le principe de la spectroscopie Raman comprimée est de filtrer le rayonnement provenant de l'échantillon afin de ne sélectionner que les combinaisons de fréquences pertinentes pour l'analyse. Ainsi, à chaque mesure, ce rayonnement est filtré par un jeu de  $P$  filtres que l'on note  $\mathbf{F} = (\mathbf{f}_1, \dots, \mathbf{f}_P)$ , avec  $\mathbf{f}_m = {}^t(f_m(1), \dots, f_m(K))$  le  $m^{\text{ème}}$  filtre et  $f_m(k) \in [0, 1]$  la transparence de ce filtre pour la  $k^{\text{ème}}$  bande spectrale. Dans les méthodes de la littérature les filtres employés sont souvent binaires, et on a alors  $f_m(k) \in \{0, 1\}$ . Dans tous les cas, le résultat du filtrage d'un rayonnement de flux  $\phi(k)$  par un filtre  $\mathbf{f}_m$  est un rayonnement de flux  $f_m(k) \phi(k)$  dans la  $k^{\text{ème}}$  bande spectrale.

Comme présenté en figure 1.4 chaque filtre  $\mathbf{f}_m$  est associé à un détecteur noté  $D_m$ . Le nombre de photons reçus par ce détecteur est noté  $n_m$ . Si les filtres sont appliqués en parallèle,

il est nécessaire d'utiliser autant de détecteurs que de filtres. Dans le cas contraire, le même détecteur peut être utilisé séquentiellement.

Les détecteurs utilisés dans le dispositif comprimé doivent être limités uniquement par le bruit de photons. Si ça n'est pas le cas, il n'est pas intéressant d'utiliser la spectroscopie comprimée, qui se base sur le fait de tirer avantage de détecteurs de meilleure qualité que les barrettes de détecteurs standard pour gagner en rapidité par rapport à un montage conventionnel. Par ailleurs, cette hypothèse est raisonnable si on utilise des détecteurs à comptage de photons. Cela a par exemple été vérifié expérimentalement pour une photodiode à avalanche dans Scotté (2020, section 2.3).

Le flux  $\phi_m$  vu par le détecteur  $D_m$  correspond à la somme des flux des différents intervalles en fréquence après filtrage par  $\mathbf{f}_m$ , soit

$$\phi_m = \sum_{k=1}^K f_m(k)\phi(k) = \gamma \sum_{k=1}^K f_m(k)s(k) = \gamma \mathbf{f}_m \cdot \mathbf{s}. \quad (1.2)$$

Comme le nombre de photons reçus dans chaque bande fréquentielle suit une loi de Poisson (voir équation (1.1)), le nombre  $n_m$  de photons reçus sur le détecteur  $D_m$  en sortie du filtre  $\mathbf{f}_m$  suit également une loi de Poisson, mais cette fois de flux  $\phi_m$ . Formellement,

$$\mathcal{P}(n_m) = \frac{(\tau_m \phi_m)^{n_m}}{n_m!} \exp(-\tau_m \phi_m).$$

où  $\tau_m$  est le temps d'acquisition lorsque le filtre  $\mathbf{f}_m$  est appliqué. Le nombre moyen de photons reçus par le détecteur  $D_m$  est alors

$$\mathbb{E}[n_m] = \tau_m \phi_m = \tau_m \gamma \mathbf{f}_m \cdot \mathbf{s}. \quad (1.3)$$

Soit  $\mathbf{n} = {}^t(n_1, \dots, n_P)$  le vecteur des nombres de photons détectés en sortie de chacun des filtres. Les composantes de ce vecteur étant indépendantes, la distribution de ce vecteur est simplement le produit de celles de ses composantes, c'est-à-dire

$$\mathcal{P}(\mathbf{n}) = \prod_{m=1}^P \frac{(\tau_m \phi_m)^{n_m}}{n_m!} \exp(-\tau_m \phi_m). \quad (1.4)$$

Un tel filtrage est équivalent à une compression, puisque seules  $P$  mesures  $n_1, \dots, n_P$  d'un rayonnement caractérisé par  $K$  coefficients  $\phi(1), \dots, \phi(K)$  sont effectuées. À l'exception du cas où l'information contenue dans les  $\phi(k)$  se situe dans un sous-espace de dimension inférieure ou égale à  $P$  de  $\mathbb{R}^K$ , cette compression ne peut se faire qu'avec perte. L'objectif des méthodes comprimées est alors de construire le meilleur filtrage pour une tâche donnée, c'est-à-dire celui permettant de préserver le plus d'informations utiles à sa réalisation et ainsi de la réaliser avec les meilleures performances possibles.

### 1.3.2 Estimation des concentrations

L'estimation des concentrations des espèces chimiques d'un mélange est l'une des premières tâches pour lesquelles des méthodes de spectroscopie Raman comprimée ont été développées (voir Nelson, Aust et al. (1998) et O. Soyemi, Eastwood et al. (2001)). Dans ce cadre, l'échantillon analysé est constitué de mélanges linéaires d'espèces chimiques dont on connaît les spectres et dont on veut déterminer les concentrations. Cet a priori peut paraître fort mais il existe de nombreuses situations où l'on dispose de la liste des espèces chimiques

pouvant composer l'échantillon étudié. On peut notamment penser au cas où on réalise le contrôle de la qualité d'un échantillon de composition normalement connue en vérifiant que les concentrations des différents composants de l'échantillon sont bien conformes aux spécifications.

Formellement, l'échantillon analysé est constitué de mélanges linéaires de  $M$  espèces connues dont on note les spectres normalisés  $\mathbf{s}_1, \dots, \mathbf{s}_M$ . Ces spectres, que l'on appelle *spectres de référence* dans ce manuscrit, peuvent par exemple être extraits de bases de données existantes ou mesurés sur des composés purs. Il a également été proposé d'utiliser des méthodes de résolution multivariée de courbes afin d'obtenir ces spectres dans le cas où les composés purs ne sont pas disponibles, comme ce qui est fait par Wilcox, Buzzard et al. (2013). Dans tous les cas, la mesure ou la reconstruction des spectres de référence complets est effectuée a priori, c'est-à-dire avant l'acquisition.

Les flux du rayonnement diffusé par un mélange sont alors les combinaisons linéaires des spectres de référence dont les poids sont les concentrations  $\mathbf{c} = {}^t(c_1, \dots, c_M)$  en chacune des espèces (voir par exemple Wilcox, Buzzard et al. (2012)), soit

$$\phi(k) = \sum_{u=1}^M c_u s_u(k). \quad (1.5)$$

On peut noter que ce que l'on appelle ici concentration ne correspond pas à la concentration chimique d'une espèce, mais est un produit de la concentration chimique par un facteur liant cette concentration et le nombre moyen de photons Raman diffusés par unité de temps dans les conditions de l'expérience. Par souci de simplicité, nous laissons ce problème de côté et continuons d'appeler ces grandeurs des concentrations dans toute la suite.

L'objectif est alors d'estimer les concentrations  $\mathbf{c}$  à partir des mesures  $\mathbf{n} = {}^t(n_1, \dots, n_P)$  réalisées par les détecteurs après filtrage par un jeu de  $P$  filtres  $\mathbf{F} = (\mathbf{f}_1, \dots, \mathbf{f}_P)$ . En injectant les flux de l'équation (1.5) dans l'équation (1.2) il est aisé de voir que le flux observé par le  $m^{\text{ème}}$  détecteur s'écrit

$$\phi_m = \sum_{k=1}^K f_m(k) \phi(k) = \sum_{k=1}^K f_m(k) \sum_{u=1}^M c_u s_u(k) = \mathbf{f}_m \cdot \sum_{u=1}^M c_u \mathbf{s}_u. \quad (1.6)$$

En posant  $\boldsymbol{\phi} = {}^t(\phi_1, \dots, \phi_P)$  le vecteur des flux reçus par les détecteurs et  $\mathbf{S} = (\mathbf{s}_1, \dots, \mathbf{s}_M)$  la matrice des spectres de référence, on obtient

$$\boldsymbol{\phi} = {}^t\mathbf{F} \mathbf{S} \mathbf{c}. \quad (1.7)$$

Les flux observés sont donc déterminés par les filtres, les spectres et les concentrations. Les filtres et les spectres étant connus, il est possible sous certaines conditions de remonter aux concentrations à partir des flux. L'observateur ne dispose pas d'un accès direct à ces flux et il est donc nécessaire de les estimer à partir des nombres de photons mesurés en sortie de chacun des filtres.

Un des estimateurs les plus simples est l'estimateur des moments. En effet, lorsque le temps de mesure par filtre est fixé à l'avance, la distribution du vecteur  $\mathbf{n}$  est celle de l'équation (1.4) où sont injectés les flux de l'équation (1.6). Un estimateur non biaisé des flux est alors simplement le rapport entre le nombre de photons détectés et le temps de mesure,

$$\hat{\phi}_m = \frac{n_m}{\tau_m}.$$

On peut introduire  $\mathbf{T}$  la matrice diagonale dont les composantes diagonales sont les temps d'acquisition, c'est-à-dire  $[\mathbf{T}]_{m,n} = \tau_m \delta_{m,n}$ , avec  $\delta_{m,n}$  le delta de Kronecker. On a alors

$$\hat{\phi} = \mathbf{T}^{-1} \mathbf{n}. \quad (1.8)$$

Lorsque la matrice  ${}^t\mathbf{FS}$  est inversible à gauche, l'estimateur des moments du vecteur des concentrations est donc

$$\hat{\mathbf{c}} = ({}^t\mathbf{FS})^+ \hat{\phi} = ({}^t\mathbf{FS})^+ \mathbf{T}^{-1} \mathbf{n}, \quad (1.9)$$

avec  $\mathbf{A}^+ = ({}^t\mathbf{AA})^{-1} {}^t\mathbf{A}$  l'inverse à gauche de  $\mathbf{A}$ , qui vérifie  $\mathbf{A}^+ \mathbf{A} = \mathbf{A}^+ \mathbf{A}^{-1}$  lorsque  $\mathbf{A}$  est carrée. Lorsque les mesures sont effectuées avec au moins autant de filtres que d'espèces de référence cet estimateur est sans biais et avec une variance finie (voir Wilcox, Buzzard et al. (2012) ou Réfrégier, Scotté et al. (2018)).

Pour obtenir de bonnes performances en estimation, Wilcox, Buzzard et al. (2012) proposent d'optimiser conjointement l'estimateur employé et le système d'acquisition, en particulier la forme des filtres et le temps d'acquisition par filtre. Dans ce travail, cette question est posée sous la forme d'un problème d'optimisation consistant à rechercher les matrices  $\mathbf{F}$ ,  $\mathbf{T}$  et la matrice  $\mathbf{E}$  correspondant à l'estimateur linéaire des concentrations  $\hat{\mathbf{c}} = \mathbf{E} \mathbf{T}^{-1} \mathbf{n}$  qui minimisent l'erreur quadratique moyenne (EQM) entre les concentrations estimées et les concentrations vraies. En utilisant cette formulation, les auteurs calculent les temps de mesures optimaux  $\tau_{m,\text{opt}}$  au sens de l'EQM. Pour un estimateur linéaire fixé et des concentrations données,

$$\tau_{m,\text{opt}} = \frac{\sqrt{\phi_m \sum_{l=1}^P [\mathbf{E}]_{\ell,m}^2}}{\sum_{n=1}^P \sqrt{\phi_n \sum_{l=1}^P [\mathbf{E}]_{\ell,n}^2}} \quad (1.10)$$

Par ailleurs, l'estimateur optimal au sens de l'EQM pour des temps de mesure  $\mathbf{T}$  fixés et des concentrations données est l'estimateur des moindres carrés pondérés par la matrice de covariance de  $\mathbf{T}^{-1} \mathbf{n}$ , soit la matrice diagonale dont les composantes diagonales sont  $\phi_m / \tau_m$ . Ces matrices dépendent en général des concentrations vraies et dans ce cas leur intérêt est limité. On peut néanmoins noter que l'estimateur optimal au sens de l'EQM lorsque  $P = M$  est celui de l'équation (1.9), soit  $\mathbf{E} = {}^t\mathbf{FS}$ . Cette matrice ne dépend alors pas de  $\mathbf{T}$  et il est par conséquent possible de calculer les temps de mesure optimaux  $\tau_{m,\text{opt}}$  à partir de l'équation (1.10). Toujours dans le cas où  $P = M$  les auteurs utilisent des routines standard d'optimisation pour trouver les filtres minimisant l'erreur quadratique moyenne. Le choix d'utiliser le nombre minimal de filtres nécessaire à réaliser l'estimation des concentrations est motivé par des raisons liées à l'implémentation du spectromètre, l'augmentation du nombre de filtres pouvant aboutir à une complexification du système d'acquisition.

Rehrauer, Dinh et al. (2018) proposent d'implémenter cette méthode avec des paires de filtres binaires, orthogonaux et complémentaires. L'orthogonalité des filtres permet de les appliquer en parallèle et ainsi de réaliser un gain de temps important lorsque l'on dispose de deux détecteurs. La définition de filtres complémentaires est celle qui suit.

**Définition 1** (Filtres complémentaires). On dit que  $P$  filtres binaires et orthogonaux  $\mathbf{f}_1, \dots, \mathbf{f}_P$  de taille  $K$  sont *complémentaires* si

$$\forall k \in \llbracket 1, K \rrbracket \exists ! m_1 \in \llbracket 1, P \rrbracket, \mathbf{f}_{m_1}(k) = 1$$

Avec ces filtres, aucune bande spectrale n'est entièrement filtrée. Autrement dit, tout photon atteignant les filtres sera redirigé sur un détecteur, ce qui permet de maximiser la quantité d'énergie intégrée sur les détecteurs.

Le schéma du montage proposé, basé sur ce principe et utilisant une matrice de micro-miroirs pour implémenter les filtres, est présenté en figure 1.5 ci-dessous. Le composant diffractif projette les différentes fréquences sur cette matrice de micro-miroirs, dont chaque colonne est associée à une bande spectrale. Ces miroirs sont contrôlés individuellement par voie électronique et peuvent prendre deux positions, renvoyant dans l'une la lumière vers le premier détecteur et dans l'autre la lumière vers le second. Ils jouent ainsi le rôle d'une paire de filtres nécessairement binaires, orthogonaux et complémentaires.

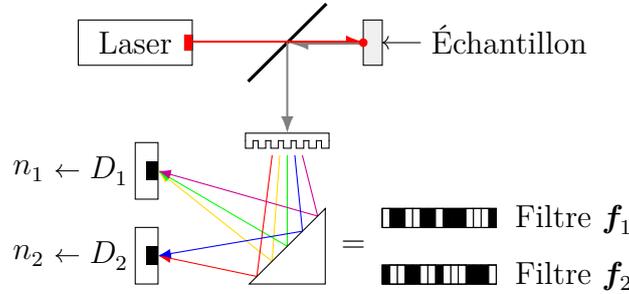


FIGURE 1.5 – Spectromètre Raman comprimé employant une paire de filtres binaires et orthogonaux appliqués en parallèle.

Ce travail montre qu'il est possible de gagner en rapidité en travaillant sur l'implémentation du système comprimé. Il existe d'autres travaux sur ce sujet, comme ceux de Scotté, Sivankutty et al. (2020) qui proposent de faire l'acquisition ligne par ligne de l'image d'un échantillon grâce à un encodage spatio-spectral. Ce type d'approche est toujours un domaine de recherche actif. Nous laissons toutefois ces développements de côté pour nous concentrer sur le problème d'estimation des concentrations lui-même.

### Estimation et borne de Cramér-Rao

Une approche plus générale, complémentaire à celle de Wilcox, Buzzard et al. (2012, 2013) a été proposée par Réfrégier, Scotté et al. (2018). Dans cet article, les auteurs proposent d'évaluer et d'optimiser les performances du système d'acquisition grâce à une méthodologie fondée sur la borne de Cramér-Rao. Cette approche permet de caractériser les performances du système comprimé et de les comparer aux performances des systèmes conventionnels indépendamment du choix de l'estimateur des concentrations.

Soient  $\mathbf{n} = {}^t(n_1, \dots, n_p)$  des réalisations de variables aléatoires dont la loi dépend des paramètres  $\mathbf{c} = {}^t(c_1, \dots, c_M)$ . Notons  $\mathbf{I}_c$  la matrice d'information de Fisher. Cette matrice est définie comme  $[\mathbf{I}_c]_{p,q} = -\mathbb{E}[\partial_{c_p} \partial_{c_q} \ell(\mathbf{n} | \mathbf{c})]$  où  $\ell(\mathbf{n} | \mathbf{c})$  est la log-vraisemblance. La matrice de Cramér-Rao  $\mathbf{\Upsilon}_c$  est l'inverse de la matrice d'information de Fisher. Comme expliqué dans Garthwaite, Jolliffe et al. (2002, section 2.4)  $\mathbf{\Upsilon}_c$  est sous certaines conditions de régularité sur la distribution de  $\mathbf{n}$  une borne inférieure sur la variance de tout estimateur non biaisé de  $\mathbf{c}$ . Ces conditions sont vérifiées par la loi de l'équation (1.4). Formellement, si  $\mathbf{\Gamma}$  est la matrice de covariance d'un tel estimateur on a

$$\mathbf{\Gamma} \succeq \mathbf{\Upsilon}_c \quad (1.11)$$

ce qui signifie que la matrice (symétrique)  $\mathbf{\Gamma} - \mathbf{\Upsilon}_c$  a des valeurs propres positives. En particulier,

$$\sigma_{c_p}^2 = [\mathbf{\Gamma}]_{p,p} \geq [\mathbf{\Upsilon}_c]_{p,p}$$

où  $\sigma_{c_p}^2$  est la variance d'estimation de  $c_p$ . Pour la loi de l'équation (1.4), la matrice d'information

de Fisher s'écrit

$$[\mathbf{I}_c]_{p,q} = \sum_{m=1}^P \tau_m \frac{(\mathbf{f}_m \cdot \mathbf{s}_p)(\mathbf{f}_m \cdot \mathbf{s}_q)}{\phi_m} = \sum_{m=1}^P \tau_m \frac{(\mathbf{f}_m \cdot \mathbf{s}_p)(\mathbf{f}_m \cdot \mathbf{s}_q)}{\mathbf{f}_m \cdot \sum_{u=1}^M c_u \mathbf{s}_u}. \quad (1.12)$$

La matrice  $\mathbf{\Upsilon}_c = \mathbf{I}_c^{-1}$  borne la variance de tout estimateur non biaisé des concentrations au sens de (1.11). Elle permet d'obtenir une précision de référence indépendante de l'estimateur utilisé. Son expression peut permettre d'identifier les paramètres du système ayant intrinsèquement un impact sur la précision d'estimation. En l'occurrence, on peut remarquer que l'information de Fisher évolue linéairement en fonction du temps de mesure. Par conséquent, multiplier le temps d'acquisition par un facteur donné revient à diviser l'écart-type minimal d'estimation des  $c_p$  par la racine carrée de ce facteur. De plus, puisque cette borne dépend de la forme des filtres, il est possible de l'utiliser pour optimiser le filtrage.

Il n'est pas certain dans le cas général qu'existe un estimateur non biaisé et *efficace*, c'est-à-dire dont la variance est égale à la borne de Cramér-Rao. Néanmoins, il est toujours possible de comparer la variance des estimateurs disponibles à cette précision de référence. On peut par exemple décider de cette manière s'il est souhaitable de chercher un autre estimateur que ceux dont on dispose déjà ou si le gain potentiel ne justifie pas le coût de cette recherche. En l'occurrence, Réfrégier, Scotté et al. (2018) ont montré que dans le cas où le nombre de filtres  $P$  est égal au nombre d'espèces  $M$  l'estimateur de l'équation (1.9) est non seulement sans biais mais également efficace.

Il est possible d'optimiser les performances d'estimation des concentrations en choisissant des filtres et des temps d'acquisition minimisant un critère dérivant de la borne de Cramér-Rao. Réfrégier, Scotté et al. (2018) proposent de construire des filtres binaires et orthogonaux en minimisant la trace de la borne de Cramér-Rao  $\text{tr } \mathbf{\Upsilon}_c$  à l'aide d'un algorithme d'optimisation consistant à réaliser de multiples modifications aléatoires successives de jeux de filtres binaires et orthogonaux, conservées uniquement si elles conduisent à une réduction de  $\text{tr } \mathbf{\Upsilon}_c$ . Puisque cette borne dépend des concentrations vraies, qui sont inconnues, les auteurs proposent de la calculer pour des concentrations uniformes (soit  $c_p = 1/M \forall p$ ) et montrent que lorsque  $P = M$  ce choix est équivalent à optimiser la moyenne de la trace de la borne de Cramér-Rao sur l'ensemble des concentrations positives se sommant à 1.

Par ailleurs, un apport important de Réfrégier, Scotté et al. (2018) est la comparaison entre un système Raman comprimé et un système Raman conventionnel, où toutes les fréquences sont mesurées simultanément. Pour que la comparaison soit pertinente, il est nécessaire de l'effectuer à temps de mesure total égal et que les détecteurs utilisés soient tous identiques, limités par un bruit de photons. Dans ce cas, mesurer la totalité des coefficients d'un spectre Raman de taille  $K$  est équivalent à effectuer une mesure avec un dispositif comprimé dont les  $K$  filtres appliqués en parallèle forment la base canonique de  $\mathbb{R}^K$ . Autrement dit, un dispositif conventionnel idéal est un cas limite d'un dispositif comprimé. En comparant les bornes de Cramér-Rao des deux systèmes, l'article montre qu'il est possible d'obtenir des performances d'estimation des concentrations proches de celles d'un système conventionnel idéal avec un système comprimé effectuant des mesures avec  $P = M$  filtres binaires et orthogonaux, mais uniquement si ceux-ci sont appliqués en parallèle.

Ces travaux montrent qu'il est possible d'optimiser les filtres d'un dispositif comprimé de sorte à ce que les performances d'estimation des concentrations soient proches de celles d'un système conventionnel. Autrement dit, les avantages en termes de vitesse d'acquisition et de coût de la spectroscopie Raman comprimée sont ici accessibles à faible perte de performance. Cette méthode a été implémentée sur un système réel par Scotté, Aguiar et al. (2018) et Grand, Scotté et al. (2022), mais avec des mesures effectuées séquentiellement et non en

parallèle car l'implémentation de plus de deux filtres en parallèle est difficile. Cela montre toutefois l'intérêt de la méthode pour l'expérimentation.

### 1.3.3 Classification des mesures parmi les espèces de référence

Une autre tâche pour laquelle des méthodes de spectroscopie Raman comprimée ont été développées est la classification d'une observation parmi une liste d'espèces de référence de spectres connus (voir Wilcox, Buzzard et al. (2012), Réfrégier, Chevallier et al. (2019) et Réfrégier et Galland (2019)). Dans ce cadre, une seule espèce chimique de la liste des espèces de référence est observée en chaque point de l'échantillon analysé. L'objectif de l'acquisition est alors de classer les mesures comprimées effectuées parmi les espèces de référence. Réaliser une telle classification est par exemple utile pour trier des échantillons de matières connues ou distinguer les zones saines des zones malades dans un échantillon donné.

Une possibilité est de réaliser la classification à partir d'une estimation de concentrations. Wilcox, Buzzard et al. (2012) suggèrent ainsi de supposer d'abord que le rayonnement observé est un mélange des espèces de référence, d'en estimer les concentrations comme présenté dans la sous-section précédente, puis de classer le résultat à l'aide de techniques telles que l'analyse discriminante linéaire ou quadratique. Employer cette méthodologie implique d'utiliser au moins autant de filtres que d'espèces de référence puisque cela est nécessaire pour estimer les concentrations.

Il est avantageux d'aborder ce problème de classification de manière plus directe. Ici, le rayonnement observé provient de la diffusion Raman du laser par une des  $M$  espèces de référence. Les spectres normalisés de ces espèces sont notés comme auparavant  $\mathbf{s}_1, \dots, \mathbf{s}_M$ . Le flux de photons diffusé dans la  $k^{\text{ème}}$  bande spectrale lorsque l'espèce de référence  $u$  est observée est alors le produit de l'intensité  $\gamma_u$  par la  $k^{\text{ème}}$  composante de  $\mathbf{s}_u$  soit

$$\phi_u(k) = \gamma_u \mathbf{s}_u(k).$$

Il s'agit donc de déterminer l'indice  $u$  de l'espèce observée à partir des mesures. Comme précédemment, ce rayonnement est filtré par un jeu de  $P$  filtres  $\mathbf{F} = (\mathbf{f}_1, \dots, \mathbf{f}_P)$  et le flux reçu par le détecteur  $D_m$  lorsque l'espèce numérotée  $u$  est observée s'écrit

$$\phi_{m,u} = \gamma_u \mathbf{f}_m \cdot \mathbf{s}_u.$$

En injectant ces flux dans l'équation (1.4) on obtient la loi du vecteur de mesures  $\mathbf{n}$  pour une espèce, des filtres et une intensité fixés.

#### Classifieur optimal et borne de Bhattacharyya

Réfrégier et Galland (2019) utilisent la théorie de la décision pour résoudre ce problème de classification dans le cas où  $M = 2$ . Le classifieur optimal des mesures  $\mathbf{n}$  au sens de la probabilité d'erreur de classification est celui du maximum de vraisemblance

$$\hat{u}_{\text{opt}}(\mathbf{n}) = \arg \max_u \mathcal{P}(\mathbf{n} | u) \mathcal{P}(u) \quad (1.13)$$

où  $\mathcal{P}(\mathbf{n} | u)$  est la loi des mesures lorsque la classe  $u$  est observée et  $\mathcal{P}(u)$  est la probabilité d'observer cette classe, avec  $\sum_{u=1}^M \mathcal{P}(u) = 1$  (voir Duda, Hart et al. (2012, p. 27)). La probabilité d'erreur  $\mathcal{P}(\epsilon)$  de ce classifieur est

$$\mathcal{P}(\epsilon) = 1 - \int_{\Omega} \max_u \mathcal{P}(\mathbf{n} | u) \mathcal{P}(u),$$

où  $\Omega$  est l'espace des échantillons. Il est en général difficile d'évaluer  $\mathcal{P}(\epsilon)$  analytiquement et il faut donc recourir au calcul numérique, ce qui peut être très coûteux en temps de calcul. Dans le cas d'un couple de classes  $(u, v)$  on dispose d'une borne supérieure de cette probabilité d'erreur, la borne de Bhattacharyya (voir Duda, Hart et al. (2012, p. 47)). Cette borne, que l'on note  $\mathcal{B}_{u,v}$ , est définie comme

$$\mathcal{B}_{u,v} = \sqrt{\mathcal{P}(u)\mathcal{P}(v)} \int_{\Omega} \sqrt{\mathcal{P}(\mathbf{n}|u)\mathcal{P}(\mathbf{n}|v)}$$

vérifie

$$\mathcal{P}(\epsilon) \leq \mathcal{B}_{u,v}.$$

La borne de Bhattacharyya est souvent plus facile à calculer et à interpréter que la probabilité d'erreur. Puisqu'il s'agit d'une borne supérieure, elle permet de contrôler les performances du classifieur, au sens où la minimiser signifie minimiser la probabilité d'erreur maximale.

Ces résultats peuvent être appliqués au cas de la classification de mesures comprimées parmi un couple d'espèces de référence  $(u, v)$ . Le classifieur optimal s'écrit alors

$$\hat{u}_{\text{opt}}(\mathbf{n}) = \arg \max_{i \in \{u,v\}} \sum_{m=1}^P (n_m \log(\tau_m \phi_{m,i}) - \tau_m \phi_{m,i}) + \log(\mathcal{P}(i)),$$

Les probabilités  $\mathcal{P}(u)$  et  $\mathcal{P}(v)$  correspondent ici à un a priori sur la probabilité d'observation de chaque espèce de référence dans l'échantillon analysé, qui est directement liée à leur proportion dans cet échantillon. Sans a priori particulier une possibilité est de choisir un prior uniforme, ce qui revient à ne favoriser aucune classe. On a alors simplement

$$\hat{u}_{\text{opt}}(\mathbf{n}) = \arg \max_{i \in \{u,v\}} \sum_{m=1}^P (n_m \log(\tau_m \phi_{m,i}) - \tau_m \phi_{m,i}).$$

La probabilité d'erreur de ce classifieur est bornée par la borne de Bhattacharyya

$$\mathcal{B}_{u,v} = \frac{1}{2} \exp(-\beta_{u,v}),$$

avec

$$\beta_{u,v} = \frac{1}{2} \sum_{m=1}^P \left[ \sqrt{\tau_m \phi_{m,u}} - \sqrt{\tau_m \phi_{m,v}} \right]^2 = \sum_{m=1}^P \left[ \sqrt{\tau_m \gamma_u \mathbf{f}_m \cdot \mathbf{s}_u} - \sqrt{\tau_m \gamma_v \mathbf{f}_m \cdot \mathbf{s}_v} \right]^2$$

la distance de Bhattacharyya. On peut remarquer que la borne de Bhattacharyya décroît exponentiellement lorsque le temps d'acquisition augmente. Elle croît en revanche exponentiellement lorsque l'intensité du rayonnement diminue. Nous reviendrons sur ce point en fin de cette sous-section. Conformément à l'intuition, il va être difficile de distinguer deux espèces chimiques produisant des rayonnements proches en spectre et en intensité. Il apparaît utile de construire un filtrage projetant les rayonnements des différentes espèces de référence de la manière la plus discriminante possible, c'est-à-dire produisant les flux les plus caractéristiques possibles sur les détecteurs.

Le calcul de ces flux et donc l'application du classifieur ainsi que le calcul de la borne requièrent la connaissance de l'intensité  $\gamma$  du rayonnement observé, que l'on a supposée inconnue. Il est toutefois possible de substituer à cette quantité son estimée au sens du maximum de vraisemblance  $\hat{\gamma}(\mathbf{n}) = \sum_{m=1}^P n_m$  ce qui conduit au classifieur du maximum de

vraisemblance généralisé

$$\hat{u}_{\text{gen}}(\mathbf{n}) = \arg \max_u \sum_{m=1}^P \left( n_m \log(\tau_m \hat{\phi}_{m,u}(\mathbf{n})) - \tau_m \hat{\phi}_{m,u}(\mathbf{n}) \right), \quad (1.14)$$

avec  $\hat{\phi}_{m,u}(\mathbf{n}) = \hat{\gamma}(\mathbf{n}) \mathbf{f}_m \cdot \mathbf{s}_u$ . Puisqu'il ne s'agit pas du classifieur optimal sa probabilité d'erreur n'est pas bornée par la borne de Bhattacharyya. Comme montré par Réfrégier et Galland (2019) elle est cependant bornée par une quantité qui en dérive.

Pour construire un filtrage permettant d'obtenir de bonnes performances en classification une solution est donc de minimiser la borne de Bhattacharyya  $\mathcal{B}_{u,v}$  ou sa variante. Tirant avantage de la rapidité avec laquelle il est possible d'évaluer ces quantités, Réfrégier et Galland (2019) proposent un algorithme d'optimisation basé sur un grand nombre de modifications aléatoires de jeux de filtres binaires et orthogonaux, acceptées uniquement si permettant de diminuer la borne choisie. Ils montrent que les filtres obtenus de cette manière permettent de classer les mesures comprimées parmi les espèces de référence avec de bonnes performances.

### Acquisition à nombre de photons détectés fixés

Lorsque le temps de mesure par filtre est fixé à l'avance, on a vu que les performances de classification dépendent de l'intensité du rayonnement diffusé. Il est dans certains cas possible de prévoir à l'avance, voire de contrôler l'intensité de ce rayonnement, par exemple en ajustant la puissance du rayonnement excitateur. Dans le cas contraire il est en revanche impossible de garantir un niveau de performance donné.

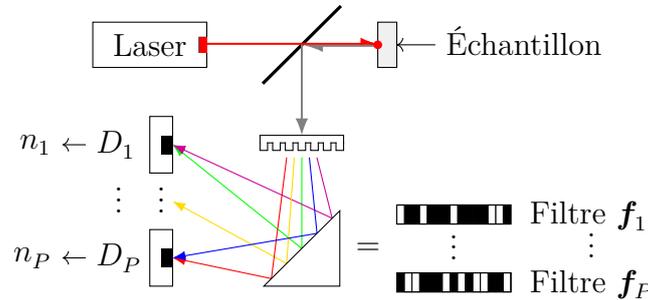


FIGURE 1.6 – Spectromètre Raman comprimé employant  $P$  filtres binaires et orthogonaux appliqués en parallèle.

Une solution à ce problème est présentée par Réfrégier, Chevallier et al. (2019). Dans cet article, les auteurs considèrent le cas d'un dispositif comprimé filtrant le rayonnement avec  $P$  filtres binaires et orthogonaux appliqués en parallèle (voir figure 1.6). L'objectif de l'acquisition est de classer les mesures parmi  $M$  espèces de référence de spectres  $\mathbf{s}_1, \dots, \mathbf{s}_M$  connus, et dont les probabilités d'apparition sont égales. On introduit la probabilité  $p_m$  qu'un photon détecté le soit par le  $m^{\text{ème}}$  détecteur. Cette probabilité correspond à la proportion moyenne de photons détectés l'ayant été par  $D_m$ . Pour l'espèce de spectre  $\mathbf{s}_u$ ,

$$p_{m,u} = \frac{\mathbb{E}[n_m]}{\sum_{\ell=1}^P \mathbb{E}[n_\ell]} = \frac{\phi_{m,u}}{\sum_{\ell=1}^P \phi_{\ell,u}} = \frac{\mathbf{f}_m \cdot \mathbf{s}_u}{\sum_{\ell=1}^P \mathbf{f}_\ell \cdot \mathbf{s}_u}, \quad (1.15)$$

avec  $\sum_{m=1}^P p_{m,u} = 1$ , ce qui n'a un sens que lorsque les filtres sont bel et bien appliqués en parallèle. L'article propose de changer la condition d'arrêt statique de l'acquisition en une condition dynamique, afin de s'adapter à l'intensité réelle du rayonnement et de maintenir

des performances constantes. Plus précisément, plutôt que d'arrêter l'acquisition après une durée fixée à l'avance, il est possible d'arrêter les mesures uniquement après avoir détecté une quantité de photons  $N$  fixée à l'avance, soit lorsque

$$\sum_{m=1}^P n_m = N. \quad (1.16)$$

La distribution des mesures  $\mathbf{n}$  conditionnellement à l'observation de l'espèce  $u$  est alors une loi multinomiale de paramètres  $p_{m,u}$  c'est-à-dire

$$\mathcal{P}(\mathbf{n} | u) = N! \prod_{m=1}^P \frac{(p_{m,u})^{n_m}}{n_m!}. \quad (1.17)$$

Cette loi est indépendante de l'intensité du rayonnement. Puisque les paramètres des lois correspondant aux différentes espèces de référence sont connus on peut utiliser directement le classifieur optimal au sens de la probabilité d'erreur de classification,

$$\hat{u}_{\text{opt}}(\mathbf{n}) = \arg \max_u \mathcal{P}(\mathbf{n} | u) = \arg \max_u \sum_{m=1}^P n_m \log(p_{m,u}). \quad (1.18)$$

L'utilisation de ce classifieur n'implique pas d'estimer les concentrations des  $M$  espèces de référence. Rien n'empêche donc d'utiliser moins de  $M$  filtres. Cela peut permettre de réduire significativement la complexité du dispositif d'acquisition, notamment s'il y a un grand nombre d'espèces de références.

Comme précédemment, dans le cas d'un couple  $(u, v)$  la borne de Bhattacharyya  $\mathcal{B}_{u,v}$  est une borne supérieure de la probabilité d'erreur de ce classifieur. Dans le cas d'une loi multinomiale, cette borne s'écrit

$$\mathcal{B}_{u,v} = \frac{1}{2} \left[ \sum_{m=1}^P \sqrt{p_{m,u} p_{m,v}} \right]^N = \frac{1}{2} [\mathbf{q}_u \cdot \mathbf{q}_v]^N, \quad (1.19)$$

avec  $\mathbf{q}_u$  le vecteur colonne tel que  $[\mathbf{q}_u]_m = (p_{m,u})^{1/2}$ . Ce formalisme est étendu au cas de la classification des mesures parmi  $M > 2$  espèces de référence équiprobables par Réfrégier, Chevallier et al. (2019), qui démontrent que la probabilité d'erreur du classifieur de l'équation (1.18) est bornée par une borne de Bhattacharyya généralisée

$$\mathcal{B} = \frac{1}{M} \sum_u \sum_{v>u} [\mathbf{q}_u \cdot \mathbf{q}_v]^N. \quad (1.20)$$

Les bornes  $\mathcal{B}_{u,v}$  et  $\mathcal{B}$  ne dépendent pas de l'intensité du rayonnement observé. La difficulté du problème de classification dépend ici uniquement du nombre total de photons détectés  $N$  et du produit scalaire entre les vecteurs  $\mathbf{q}_u$ . Il s'agit donc cette fois de construire un filtrage projetant les spectres de manière à ce que les probabilités de détection par filtre soient caractéristiques à chaque espèce de référence. En outre, les performances de classification augmentent avec  $N$ . On dispose donc d'un levier que l'on peut ajuster afin d'atteindre un certain niveau de performances.

## 1.4 Conclusion et positionnement de la thèse

La classification de mesures effectuées par un spectromètre Raman comprimé et l'estimation des concentrations d'un mélange à partir de ce même type de mesures ont été abordées de multiples manières dans la littérature. Leur traitement est effectué sous des hypothèses fortes, notamment la connaissance a priori des espèces composant l'échantillon analysé, et en particulier de leurs spectres. Cela permet d'optimiser les filtres appliqués en fonction de ces hypothèses et de la tâche à réaliser, et ainsi de remonter à l'information recherchée à partir d'un faible nombre de photons et avec seulement quelques mesures. Il est également possible dans ce cadre de quantifier les performances des estimateurs et des classifieurs utilisés.

La stratégie d'acquisition et l'analyse des mesures sont construites pour des espèces de référence données. Dans le cas où l'a priori sur les espèces en présence n'est pas vérifié, les résultats produits par l'analyse peuvent par conséquent être de moins bonne qualité que prévu, voire complètement erronés, et il est crucial de pouvoir le détecter.

Cette thèse vise à concevoir des approches de spectroscopie Raman comprimée robustes. Plus précisément, il s'agit de permettre à l'utilisateur de réaliser la tâche considérée avec les performances souhaitées lorsque les hypothèses effectuées sont vérifiées, mais également de détecter les écarts à ces hypothèses. Le défi est donc de concevoir un système à la fois spécialisé pour des espèces chimiques données mais également capable de détecter la présence d'espèces imprévues. Ce type d'avancée est nécessaire à étendre l'emploi de la spectroscopie Raman comprimée à des applications critiques, pour lesquelles il est essentiel de quantifier la confiance qu'il est véritablement possible d'avoir en les résultats de l'analyse.



# Chapitre 2

## Combiner classification et reconstruction de spectres en spectroscopie Raman comprimée

### Sommaire

---

|            |  |           |
|------------|--|-----------|
| <b>2.1</b> | <b>Modalités d'acquisition considérées</b>                   | <b>26</b> |
| 2.1.1      | Modélisation de l'échantillon                                | 26        |
| 2.1.2      | Type de dispositif et stratégie d'acquisition                | 26        |
| <b>2.2</b> | <b>Manque de robustesse des approches existantes</b>         | <b>27</b> |
| <b>2.3</b> | <b>Réaliser à la fois classification et reconstruction</b>   | <b>30</b> |
| 2.3.1      | Effectuer une diversité de mesures comprimées                | 30        |
| 2.3.2      | Stratégie d'acquisition permettant la reconstruction         | 30        |
| 2.3.3      | Caractéristiques des filtres et reconstruction               | 31        |
| 2.3.4      | Filtres complémentaires et maîtrise du temps d'acquisition   | 33        |
| 2.3.5      | Reconstruction du spectre observé                            | 34        |
| <b>2.4</b> | <b>Problème d'optimisation des filtres</b>                   | <b>35</b> |
| 2.4.1      | Critères à optimiser   | 36        |
| 2.4.2      | Optimisation et sélection d'une diversité de jeux de filtres | 37        |
| <b>2.5</b> | <b>Approche fondée sur les compromis optimaux</b>            | <b>41</b> |
| 2.5.1      | Principe des filtres à compromis optimaux                    | 42        |
| 2.5.2      | Recherche des filtres à compromis optimaux                   | 42        |
| 2.5.3      | Algorithme de recherche des filtres à compromis optimaux     | 44        |
| 2.5.4      | Mise en pratique de l'approche à compromis optimaux          | 46        |
| <b>2.6</b> | <b>Étude des performances du système d'acquisition</b>       | <b>51</b> |
| 2.6.1      | Influence de la forme des spectres des espèces de référence  | 51        |
| 2.6.2      | Influence du nombre d'espèces de référence                   | 53        |
| 2.6.3      | Influence du nombre de jeux de filtres                       | 54        |
| 2.6.4      | Influence du nombre de photons à détecter                    | 55        |
| 2.6.5      | Influence du nombre de filtres                               | 56        |
| 2.6.6      | Influence du nombre de passages par pixel                    | 57        |
| <b>2.7</b> | <b>Détection et caractérisation d'une anomalie</b>           | <b>61</b> |
| 2.7.1      | Méthodologie proposée et détection d'anomalie                | 61        |
| 2.7.2      | Retour sur le premier exemple                                | 61        |
| <b>2.8</b> | <b>Conclusion, limites et ouverture</b>                      | <b>65</b> |

---

Dans ce chapitre, on s'intéresse à la classification de mesures de spectroscopie Raman comprimée parmi des espèces chimiques de spectres connus. Ce problème apparaît par exemple dans un contexte où il est nécessaire de trier des échantillons de matières connues ou de séparer un champ de vue en différentes régions de compositions connues. Ce sujet a été abordé de plusieurs manières dans la littérature, notamment avec une approche fondée sur la théorie de la décision qui permet d'évaluer les performances des filtres du spectromètre comprimé et d'optimiser ce filtrage afin de minimiser la probabilité d'erreur de classification. Cette méthode a été rappelée dans la section 1.3.3 du chapitre précédent.

Cependant, comme nous le montrons dans la suite, les approches de classification comprimée proposées dans le cadre de la spectroscopie Raman ont un problème de robustesse car elles peuvent conduire à des résultats erronés lorsque des composés imprévus sont observés. Ce problème est d'autant plus sérieux que l'utilisateur n'a en général pas la possibilité de caractériser ou même de détecter les anomalies rencontrées. Cette faille rend l'emploi tel quel de ces méthodes difficilement envisageable pour des applications critiques, par exemple lorsque l'objectif est de contrôler la composition des échantillons analysés.

Nous verrons dans ce chapitre qu'il est possible d'adapter cette approche de manière à ce qu'il soit possible de questionner le modèle à partir des mesures au prix d'une perte limitée en performance de classification. Nous illustrons la méthodologie proposée dans un cas particulier où l'échantillon simulé est pollué par un composant totalement inconnu que nous parvenons néanmoins à détecter et caractériser sans mesure supplémentaire.

## 2.1 Modalités d'acquisition considérées

### 2.1.1 Modélisation de l'échantillon

On s'intéresse dans ce chapitre à la classification de mesures comprimées parmi une liste d'espèces chimiques de référence. On considère donc que le rayonnement observé a été diffusé par une seule espèce chimique à la fois et on souhaite déterminer quelle est l'espèce présente en chaque pixel de l'échantillon analysé. Cette situation correspond par exemple à l'observation de morceaux de différentes matières ou d'agrégats de grains de matière de taille plus importante que la résolution spatiale du microscope.

Si l'échantillon est composé de  $M$  espèces chimiques de spectres  $\mathbf{s}_1, \dots, \mathbf{s}_M$  alors une unique espèce est observée en chaque pixel et le spectre  $\mathbf{s}$  du rayonnement observé est tel que

$$\mathbf{s} \in \{\mathbf{s}_1, \dots, \mathbf{s}_M\},$$

ce qui correspond exactement à la situation de la section 1.3.3. On rappelle que nous supposons toujours les spectres normalisés et positifs, c'est-à-dire  $\sum_{k=1}^K s(k) = 1$  et  $s(k) \geq 0 \forall k$ .

### 2.1.2 Type de dispositif et stratégie d'acquisition

Comme expliqué dans la section 1.3.3 et démontré par Réfrégier, Chevallier et al. (2019) il n'est pas nécessaire d'employer autant de filtres que d'espèces chimiques de référence pour réaliser la classification. De plus, lorsque les filtres sont appliqués en parallèle et que l'acquisition est arrêtée après la détection d'un nombre de photons fixé à l'avance, la probabilité d'erreur de classification est indépendante de l'intensité du rayonnement.

Ce chapitre se place dans ce cadre et le dispositif d'acquisition considéré est un spectromètre Raman comprimé filtrant le rayonnement provenant de l'échantillon avec  $P$  filtres binaires et orthogonaux appliqués en parallèle. Ce dispositif est représenté en figure 1.6. Puisqu'il est en

pratique difficile d'appliquer plus de deux filtres en parallèle, les expériences de ce chapitre sont en général réalisées avec  $P = 2$ .

Avec cette configuration la loi des mesures est la loi multinomiale donnée à l'équation (1.17), que l'on rappelle ici. Soient  $\mathbf{n} = {}^t(n_1, \dots, n_P)$  les mesures réalisées avec un jeu de  $P$  filtres binaires et orthogonaux  $\mathbf{F} = (\mathbf{f}_1, \dots, \mathbf{f}_P)$ . Si l'espèce  $u$  de spectre  $\mathbf{s}_u$  est observée et si l'acquisition est arrêtée après la détection de  $N$  photons alors on a

$$\mathcal{P}(\mathbf{n} | u) = N! \prod_{m=1}^P \frac{(p_{m,u})^{n_m}}{n_m!} \quad (2.1)$$

où

$$p_{m,u} = \frac{\mathbf{f}_m \cdot \mathbf{s}_u}{\sum_{n=1}^P \mathbf{f}_n \cdot \mathbf{s}_u} \quad (2.2)$$

est la proportion moyenne de photons détectés l'étant par le détecteur  $D_m$ .

Bien entendu, la méthodologie présentée dans ce chapitre est généralisable à d'autres types de dispositifs, notamment à ceux où les filtres sont appliqués séquentiellement ainsi qu'à ceux où c'est le temps d'acquisition qui est fixé plutôt que le nombre de photons détectés.

## 2.2 Manque de robustesse des approches existantes

Comme présenté dans le chapitre précédent, les méthodes de spectroscopie Raman comprimée consistent à ne faire l'acquisition que de certaines bandes spectrales d'intérêt sur un petit nombre de détecteurs plus rapides et moins bruités que les barrettes de détecteurs utilisées dans un spectromètre Raman conventionnel. Une telle compression risque de dégrader significativement les performances du système, à moins d'optimiser les filtres employés.

Dans les méthodes existantes comme celles de Wilcox, Buzzard et al. (2012, 2013), Réfrégier, Chevallier et al. (2019) et Réfrégier et Galland (2019) l'optimisation consiste à spécialiser le filtrage pour la tâche et l'a priori considérés. Dans le cadre de la classification, il est notamment proposé par Réfrégier, Chevallier et al. (2019) de rechercher les filtres minimisant la borne de Bhattacharyya associée aux espèces de référence. Nous proposons d'illustrer les résultats obtenus avec cette approche sur un exemple simulé.

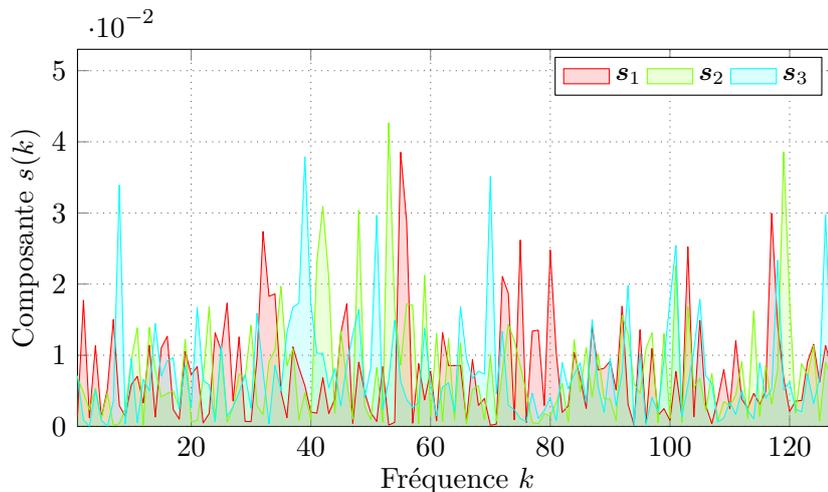


FIGURE 2.1 – Spectres des espèces chimiques de référence.

Dans cet exemple, on suppose que l'échantillon analysé est composé des  $M = 3$  espèces

de référence dont les spectres  $\mathbf{s}_1, \mathbf{s}_2, \mathbf{s}_3$  apparaissent en figure 2.1. Ces spectres possèdent  $K = 128$  composantes et ont été tirés uniformément et indépendamment sur le simplexe de  $\mathbb{R}^K$ , c'est-à-dire sur l'espace des vecteurs normalisés et positifs à  $K$  composantes. Le système d'acquisition considéré comporte  $P = 2$  filtres binaires, orthogonaux et complémentaires appliqués en parallèle. L'acquisition est arrêtée après la détection de  $N = 250$  photons.

On optimise un jeu de  $P = 2$  filtres pour la classification des mesures parmi les espèces de référence. Pour ce faire, on utilise l'algorithme de Réfrégier, Chevallier et al. (2019). À l'issue de cette optimisation on obtient une paire de filtres dont la borne de Bhattacharyya vaut  $\mathcal{B} = 1.43 \times 10^{-4}$ . Cela signifie que le classifieur optimal commettra en moyenne moins de deux erreurs sur  $10^4$  classifications en classant des mesures de rayonnements diffusés par les espèces 1 à 3 effectuées à travers ces filtres.

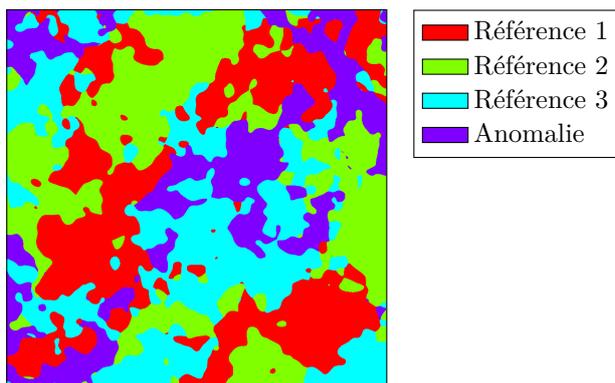


FIGURE 2.2 – Répartition des espèces chimiques dans l'échantillon.

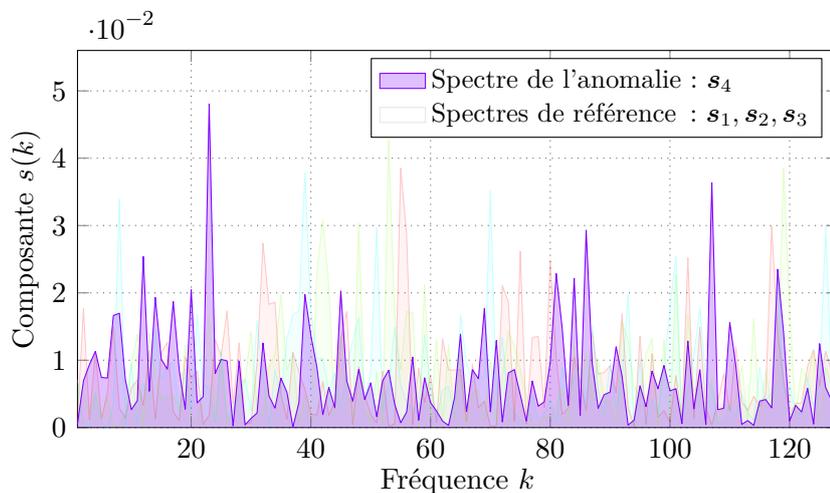


FIGURE 2.3 – Spectre de l'espèce chimique imprévue.

Afin de tester les performances des filtres optimisés, on simule l'acquisition d'un champ de vue carré de 1000 pixels de côté. Sur ce champ de vue, qui apparaît sur la figure 2.2, nous avons ajouté une quatrième espèce imprévue. La répartition des différentes espèces a été générée aléatoirement à l'aide d'un bruit de Perlin, bruit utilisé pour générer des textures pseudo-naturelles dans le contexte de la synthèse d'images par ordinateur et présenté dans Perlin (1985). Le spectre de l'anomalie a été généré par un tirage uniforme sur le simplexe et apparaît en figure 2.3. On suppose ignorer la présence de l'espèce anormale ainsi que son spectre. Cette situation correspond à celle où on contrôle la qualité d'un échantillon de composition normalement connue mais que celui-ci est pollué par un composé inconnu.

On simule l'acquisition de chaque pixel du champ de vue avec le jeu de filtres optimisé. Les mesures simulées sont classées parmi les espèces 1 à 3 avec le classifieur de l'équation (1.18), qui est optimal au sens de la probabilité d'erreur de classification. Le résultat de cette analyse apparaît en figure 2.4.

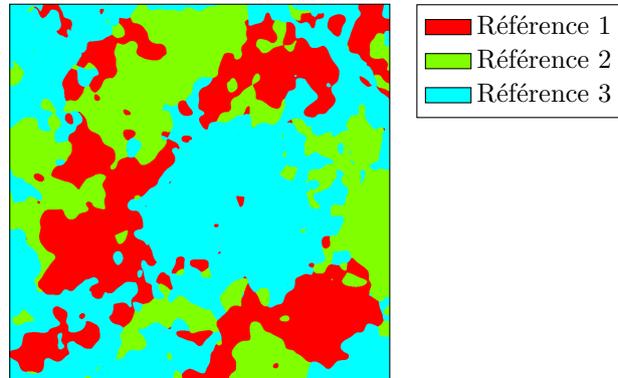


FIGURE 2.4 – Classification parmi les espèces de référence de mesures du champ de vue de la figure 2.2 réalisées avec un même jeu de  $P = 2$  filtres et  $N = 250$  photons détectés par pixel.

Aucune anomalie n'apparaît sur ces résultats car les pixels où la quatrième espèce est présente ont été systématiquement attribués à l'espèce 3. Ce résultat est causé par la proximité des projections des spectres  $\mathbf{s}_3$  et  $\mathbf{s}_4$  par les filtres que l'on a utilisés, c'est-à-dire des couples de probabilités  $(p_{1,3}, p_{2,3})$  et  $(p_{1,4}, p_{2,4})$ . Ces projections sont représentées sur la figure 2.5 avec celles correspondant aux autres espèces. On rappelle que la somme des projections correspondant à une même espèce vaut 1.

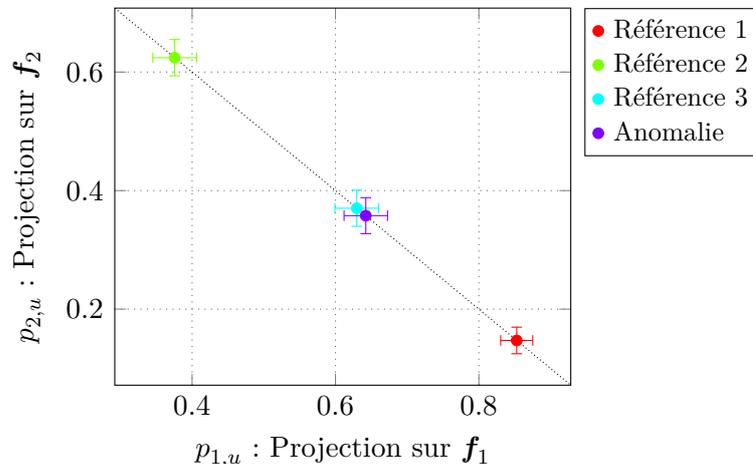


FIGURE 2.5 – Projections des spectres des différentes espèces par les filtres.

Pour visualiser la difficulté de distinguer  $\mathbf{s}_4$  de  $\mathbf{s}_3$  nous avons représenté sur la figure les barres d'erreur correspondant aux écart-types des  $\hat{p}_m = n_m/N$  pour  $N = 250$ . Formellement, la borne de Bhattacharyya associée à ce problème de classification vaut  $\mathcal{B}_{3,4} = 0.49$ . Cela signifie que même en connaissant  $\mathbf{s}_4$  et en utilisant le classifieur optimal pour séparer les espèces 3 et 4, la probabilité d'erreur est équivalente à celle obtenue en décidant de manière aléatoire entre ces deux espèces. Quoi qu'il en soit, le spectre de l'espèce 4 est supposé inconnu, ce qui empêche d'utiliser ce classifieur.

Une solution pour résoudre ce problème de robustesse serait de pouvoir reconstruire les spectres des espèces rencontrées dans l'échantillon. Comparer les spectres reconstruits aux

spectres de références permettrait en effet de détecter s'il y a un problème et d'en identifier l'origine le cas échéant.

Pour obtenir ces spectres, une possibilité est de faire des mesures sur certains pixels de l'échantillon avec un spectromètre Raman conventionnel. Cependant, outre le temps nécessaire à de telles mesures, il est parfois difficile voire impossible de retrouver le même champ de vue plusieurs fois. C'est par exemple le cas si l'objet analysé est en mouvement, ou simplement si l'échantillon ne peut être analysé à nouveau. L'objectif de la section suivante est donc de mettre en œuvre la reconstruction d'un spectre sans mesure supplémentaire.

## 2.3 Réaliser à la fois classification et reconstruction

### 2.3.1 Effectuer une diversité de mesures comprimées

Les approches existantes de la classification en spectroscopie Raman comprimée consistent à effectuer les mesures avec un unique jeu de  $P \ll K$  filtres construit pour maximiser les performances en classification. Comme nous l'avons illustré dans la section précédente, appliquer cette stratégie a un inconvénient majeur : elle permet très difficilement de détecter les écarts au modèle et ne permet pas de déterminer l'origine d'une anomalie. Nous proposons donc d'effectuer les mesures avec une diversité de jeux de filtres.

Ce principe a déjà été utilisé pour reconstruire des spectres Raman à partir de mesures comprimées. Il est par exemple possible de mesurer des projections aléatoires puis d'utiliser des hypothèses de parcimonie pour mettre en œuvre la reconstruction, comme dans Berto, Scotté et al. (2017). Une autre possibilité est de mesurer une base complète de l'espace des spectres afin d'effectuer la reconstruction directement. C'est ce qui est fait par DeVerse, Hammaker et al. (2000) avec une base de Hadamard.

Dans notre cas, l'objectif principal de l'acquisition est la classification des mesures parmi les espèces de référence en chaque pixel. Nous verrons dans la suite qu'employer des filtres aléatoires ou provenant d'une base arbitraire n'est pas adapté à cet objectif. En revanche, nous verrons qu'il existe la plupart du temps un large ensemble de jeux de filtres permettant de classer les mesures parmi les espèces de référence avec une faible probabilité d'erreur. Il paraît donc possible d'avoir les moyens de reconstruire les spectres rencontrés tout en conservant de bonnes performances en classification.

Appliquer une diversité de jeux de filtres implique de pouvoir modifier le filtrage appliqué par le spectromètre. De nombreux travaux évoqués dans la section 1.2 emploient une matrice de micro-miroirs pour implémenter les filtres du spectromètre comprimé. Ce système, contrôlé électroniquement, permet d'appliquer aisément toute paire de filtres binaires, orthogonaux, et complémentaires au rayonnement observé. Il est donc possible d'appliquer une diversité de jeux de filtres en faisant commuter les miroirs de la matrice chaque fois que cela est nécessaire. La fréquence de commutation d'un tel dispositif est de l'ordre de 20 kHz, d'où un temps mort de 50  $\mu$ s à chaque commutation (voir Scotté (2020, section 2.3.1)). Ce délai est comparable au temps d'acquisition d'un pixel, qui est en général de quelques dizaines ou centaines de microsecondes. On peut donc appliquer la diversité de jeux de filtres désirée sans augmenter exagérément le temps d'acquisition.

### 2.3.2 Stratégie d'acquisition permettant la reconstruction

Pour appliquer une diversité de jeux de filtres tout en conservant un nombre de mesures et un temps d'acquisition limités nous proposons d'assigner des jeux de filtres différents aux pixels de la zone à analyser. Autrement dit, on continue à n'appliquer qu'un jeu de filtres par

pixel mais on autorise à changer de jeu de filtres à chaque pixel. Le jeu de filtres appliqué au pixel  $w$  est noté  $\mathbf{F}^{(w)}$ . Ce dispositif est schématisé en figure 2.6.

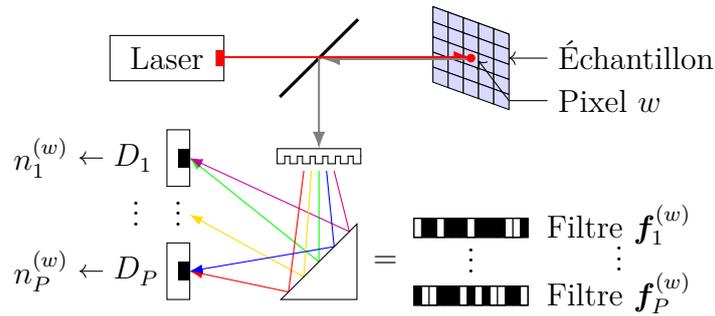


FIGURE 2.6 – Spectromètre Raman comprimé employant  $P$  filtres binaire et orthogonaux appliqués en parallèle et pouvant changer de jeu de filtres à chaque pixel.

Il faut alors utiliser des mesures réalisées sur plusieurs pixels si on veut effectuer une reconstruction de spectre. Si les pixels sélectionnés sont de composition chimique homogène, on obtient à l'issue de la reconstruction une approximation du spectre de l'espèce chimique correspondant. Dans le cas contraire, le spectre reconstruit est plus difficilement interprétable. Il est dans tous les cas nécessaire de s'assurer d'effectuer une diversité suffisante de mesures, sans quoi il est de toute façon impossible d'effectuer la reconstruction. Le choix des pixels à combiner peut se faire après les mesures, contrairement à la construction des filtres qui est effectuée avant l'acquisition. L'enjeu de la méthode est donc de construire les meilleurs filtres possibles, dans un sens qu'il faut préciser.

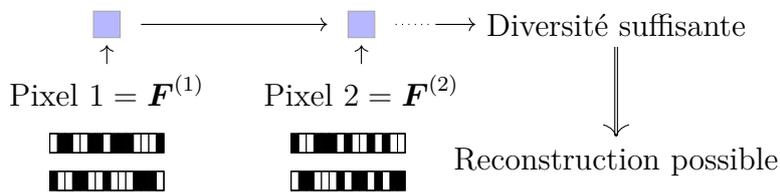


FIGURE 2.7 – Combiner plusieurs pixels pour réaliser la reconstruction.

L'optimisation du parcours de la zone imagée, c'est-à-dire de l'ordre dans lequel les pixels sont visités, sort du cadre de ce manuscrit. Soulignons toutefois que le nombre de jeux de filtres différents nécessaire à la reconstruction est en général largement inférieur au nombre de pixels de l'image. Par conséquent, une manière de réaliser l'acquisition rapidement est de visiter tous les pixels assignés à un même jeu de filtres avant de passer au jeu de filtres suivant. Cette méthode est avantageuse par rapport au cas où l'on ferait commuter le système à chaque pixel puisque les dispositifs d'acquisition permettent en général de parcourir rapidement le champ de vue, par exemple en utilisant des miroirs galvanométriques.

### 2.3.3 Caractéristiques des filtres et reconstruction

La stratégie que nous proposons consiste à effectuer des mesures avec une diversité de jeux de filtres, et la question est de savoir comment construire ces filtres pour réaliser au mieux à la fois la classification et la reconstruction.

Il est en particulier nécessaire d'établir un critère permettant d'évaluer les performances d'un ensemble de jeux de filtres en reconstruction. Pour ce faire, nous proposons d'employer la borne de Cramér-Rao. Cette borne, bien connue en statistiques, s'est révélée utile à la

quantification des performances des systèmes de traitement de l'information en estimation de paramètres dans des domaines très divers, depuis la théorie de la compression des impulsions radar de Woodward (1953) jusqu'aux systèmes de spectroscopie modernes comme par exemple dans Réfrégier, Scotté et al. (2018).

Le calcul de la borne de Cramér-Rao nécessite d'inverser la matrice d'information de Fisher. L'étude de l'inversibilité de cette matrice permet de préciser le nombre de jeux de filtres nécessaires à la reconstruction.

### Performances des filtres en reconstruction

On établit ici l'expression de la borne de Cramér-Rao de la reconstruction d'un spectre à partir de mesures comprimées effectuées sur plusieurs pixels de même composition chimique.

Soient  $\mathbf{n}^{(1)}, \dots, \mathbf{n}^{(W)}$  où on note  $\mathbf{n}^{(w)} = \mathbf{t}(n_1^{(w)}, \dots, n_P^{(w)})$  les mesures comprimées réalisées en appliquant les jeux de  $P$  filtres orthogonaux  $\mathbf{F}^{(1)}, \dots, \mathbf{F}^{(W)}$  où  $\mathbf{F}^{(w)} = (\mathbf{f}_1^{(w)}, \dots, \mathbf{f}_P^{(w)})$  sur  $W$  pixels où l'échantillon analysé est composé d'une espèce ayant un spectre  $\mathbf{s}$ . Pour manipuler plus aisément ces quantités, on introduit le vecteur  $\mathbf{n}$  qui est la concaténation des vecteurs comptant le nombre de photons détectés en chaque pixel, c'est-à-dire la transposée de  $(\mathbf{t}\mathbf{n}^{(1)}, \dots, \mathbf{t}\mathbf{n}^{(W)})$ . L'acquisition est arrêtée en chaque pixel  $w$  après la détection de  $N$  photons et on a donc  $\sum_{m=1}^N n_m^{(w)} = N \forall w$ . La distribution des mesures est donc indépendante de l'intensité du rayonnement. Les mesures effectuées sur différents pixels étant indépendantes,

$$\mathcal{P}(\mathbf{n} | \mathbf{s}) = \prod_{w=1}^W \mathcal{P}(\mathbf{n}^{(w)} | \mathbf{s}) = \prod_{w=1}^W N! \prod_{m=1}^P \frac{[p_m^{(w)}]^{n_m^{(w)}}}{n_m^{(w)}!}, \quad (2.3)$$

où  $p_m^{(w)}$  est la proportion moyenne de photons détectés par le  $m^{\text{ème}}$  détecteur lorsque le jeu de filtres  $\mathbf{F}^{(w)}$  est employé, et dont l'expression est donnée dans l'équation (1.15).

La matrice de Cramér-Rao est définie comme  $\mathbf{\Upsilon} = \mathbf{I}^{-1}$  où  $\mathbf{I}$  est la matrice d'information de Fisher, dont nous donnons l'expression dans l'équation (2.4). On rappelle que si la distribution des mesures est suffisamment régulière alors cette matrice est une borne inférieure de la variance de tout estimateur non biaisé (voir Garthwaite, Jolliffe et al. (2002, section 2.4)). Ces conditions de régularité sont vérifiées par la loi de l'équation (2.3). Plus formellement, si  $\mathbf{\Gamma}$  est la matrice de covariance d'un estimateur non biaisé alors

$$\mathbf{\Gamma} \succeq \mathbf{\Upsilon},$$

ce qui signifie que la matrice  $\mathbf{\Gamma} - \mathbf{\Upsilon}$  a des valeurs propres positives. En particulier,

$$\sigma_{s^{(k)}}^2 = [\mathbf{\Gamma}]_{k,k} \geq [\mathbf{\Upsilon}]_{k,k}$$

où  $\sigma_{s^{(k)}}^2$  est la variance d'estimation de  $s(k)$ . La borne de Cramér-Rao dépend de la forme des filtres et il est donc possible de l'utiliser pour optimiser le filtrage. C'est ce que nous faisons dans les sections 2.4 et 2.5.

L'indépendance de cette borne à l'estimateur employé peut permettre de dégager les paramètres clés du système d'acquisition qu'il est possible d'ajuster pour améliorer les performances en reconstruction. Elle permet également à l'étude effectuée dans la suite de ce chapitre d'être indépendante du choix de cet estimateur et de sa mise en oeuvre.

La matrice d'information de Fisher  $\mathbf{I}$  associée à l'estimation de  $\mathbf{s}$  est définie comme

$$[\mathbf{I}]_{k,l} = -\mathbb{E} [\partial_{s^{(k)}} \partial_{s^{(l)}} \ell(\mathbf{n} | \mathbf{s})]$$

où  $\ell(\mathbf{n} | \mathbf{s}) = \log \mathcal{P}(\mathbf{n} | \mathbf{s})$  est la log-vraisemblance. On rappelle que  $\sum_{k=1}^K s(k) = 1$  et qu'il n'y a donc que  $K - 1$  paramètres indépendants à estimer.  $\mathbf{I}$  est donc de taille  $(K - 1) \times (K - 1)$ . Comme démontré dans l'annexe A.1,

$$[\mathbf{I}]_{k,l} = \sum_{w=1}^W \frac{N}{\sum_{n=1}^P \mathbf{f}_n^{(w)} \cdot \mathbf{s}} \left[ \sum_{m=1}^P \frac{\delta f_m^{(w)}(k) \delta f_m^{(w)}(l)}{\mathbf{f}_m^{(w)} \cdot \mathbf{s}} - \frac{\sum_{n=1}^P \delta f_n^{(w)}(k) \sum_{n=1}^P \delta f_n^{(w)}(l)}{\sum_{n=1}^P \mathbf{f}_n^{(w)} \cdot \mathbf{s}} \right] \quad (2.4)$$

avec

$$\delta f_m^{(w)}(k) = f_m^{(w)}(k) - f_m^{(w)}(K) \quad (2.5)$$

### Diversité de filtres nécessaire à la reconstruction

Outre le calcul de la borne de Cramér-Rao, étudier la matrice d'information de Fisher permet de préciser les conditions sous lesquelles les mesures effectuées contiennent suffisamment d'information à propos du spectre observé pour qu'il soit possible de reconstruire chacune de ses composantes sans introduire d'hypothèse supplémentaire. En particulier, un résultat intéressant est qu'il n'existe en général pas d'estimateur non biaisé à variance finie si la matrice d'information de Fisher est singulière (voir Stoica et Marzetta (2001)). On étudie l'inversibilité de  $\mathbf{I}$  en annexe A.1, où on montre les résultats suivants.

**Proposition 1** (Diversité de filtres nécessaire à la reconstruction). Soient  $\mathbf{n}$  des mesures d'un rayonnement de spectre  $\mathbf{s} \in \mathbb{R}^K$  réalisées avec les jeux de filtres  $\mathbf{F}^{(1)}, \dots, \mathbf{F}^{(W)}$  à nombre de photons détectés fixé. Soit également  $\delta \mathbf{F}^{(w)} = (\delta \mathbf{f}_1^{(w)}, \dots, \delta \mathbf{f}_P^{(w)})$  avec  $[\delta \mathbf{f}_m^{(w)}]_k = \delta f_m^{(w)}(k)$  défini en (2.5). Si la matrice d'information de Fisher  $\mathbf{I}$  donnée par (2.4) est inversible alors la famille  $\{\delta \mathbf{F}^{(1)}, \dots, \delta \mathbf{F}^{(W)}\}$  engendre  $\mathbb{R}^{K-1}$ . Lorsque les filtres employés sont complémentaires au sens de la définition 1 la réciproque est vraie.

**Proposition 2** (Nombre de jeux de filtres nécessaires à la reconstruction). Soit  $\mathbf{n}$  des mesures d'un rayonnement de spectre  $\mathbf{s} \in \mathbb{R}^K$  réalisées avec  $W$  jeux de  $P$  filtres à nombre de photons détectés fixé. Si la matrice d'information de Fisher  $\mathbf{I}$  donnée par (2.4) est inversible alors

$$W \geq W_{\min} = \left\lceil \frac{K-1}{P-1} \right\rceil.$$

On peut remarquer qu'il n'est pas suffisant que les jeux de filtres forment une base de  $\mathbb{R}^K$  pour avoir la possibilité de reconstruire le spectre observé. Ainsi, diviser la base canonique en  $W = \lceil K/2 \rceil$  paires de filtres ne permet pas de vérifier la condition de la proposition 2 et donc d'obtenir l'inversibilité de la matrice d'information de Fisher, hormis bien sûr lorsque  $K = 2$ . Les mesures se sommant à  $N$  en chaque pixel on ne dispose en effet alors pas de suffisamment de mesures indépendantes pour estimer les  $K - 1$  coefficients indépendants du spectre observé. Il faut pour cela en faire environ deux fois plus.

### 2.3.4 Filtres complémentaires et maîtrise du temps d'acquisition

Le principal avantage de la spectroscopie Raman comprimée par rapport à la spectroscopie Raman conventionnelle est une réduction importante du temps de mesure. Il faut donc quantifier et contrôler ce temps autant que faire se peut afin de s'assurer que ce gain soit bien réalisé. Il est impossible de contrôler parfaitement le temps d'acquisition lorsque la condition d'arrêt des mesures est la détection d'un certain nombre de photons, mais on peut limiter sa valeur moyenne pour une intensité du rayonnement donnée. On note  $\tau_N$  le temps nécessaire

à l'acquisition de  $N$  photons en moyenne et  $\gamma$  l'intensité du rayonnement. L'équation (1.3) donne alors

$$\tau_N = \frac{N}{\sum_{m=1}^P \phi_m} = \frac{N}{\gamma \sum_{m=1}^P \mathbf{f}_m \cdot \mathbf{s}}. \quad (2.6)$$

Lorsque le spectre observé est connu, c'est-à-dire lorsque l'a priori à propos de la composition de l'échantillon analysé est vérifié, il est possible d'agir sur les filtres afin de limiter le temps d'acquisition pour une intensité donnée. C'est ce qui est implicitement fait par Réfrégier, Chevallier et al. (2019) en imposant une borne supérieure sur  $\tau_N$  lors de l'optimisation des filtres. Cela n'est plus possible si les spectres rencontrés ne sont pas les spectres de référence, comme on l'envisage ici.

Une solution à ce problème est d'utiliser des jeux de filtres complémentaires, qui sont une généralisation des paires de filtres orthogonaux complémentaires dont l'utilisation a été proposée par Réfrégier, Scotté et al. (2018) et Rehrauer, Dinh et al. (2018). La définition de filtres complémentaires est la définition 1 de la section 1.3.2

Si  $\mathbf{F} = (\mathbf{f}_1, \dots, \mathbf{f}_P)$  est un jeu de filtres complémentaires alors  $\sum_{m=1}^P \mathbf{f}_m = \mathbf{1}$  ou de manière équivalente  $\sum_{m=1}^P \mathbf{f}_m(k) = 1 \forall k \in \llbracket 1, K \rrbracket$ . Cela signifie qu'un tel filtrage ne coupe aucune bande spectrale de l'ensemble analysé. Ainsi, tout photon atteignant les filtres sera redirigé vers un détecteur. Par conséquent, pour tout spectre  $\mathbf{s}$

$$\sum_{m=1}^P \mathbf{f}_m \cdot \mathbf{s} = \left( \sum_{m=1}^P \mathbf{f}_m \right) \cdot \mathbf{s} = \mathbf{1} \cdot \mathbf{s} = 1 \quad (2.7)$$

et l'équation (2.6) donne

$$\tau_N = \frac{N}{\gamma}.$$

Autrement dit, utiliser des filtres complémentaires donne la certitude d'obtenir le temps d'acquisition minimal pour une intensité et un nombre de photons détectés donnés, et ce peu importe le spectre observé. C'est également le seul type de filtres pour lequel ce temps ne diverge que lorsque l'intensité du rayonnement est nulle.

Une autre motivation à utiliser des jeux de filtres complémentaires se trouve dans la mise en œuvre du filtrage. Utiliser des jeux de filtres non complémentaires implique de pouvoir rejeter certains photons. Cela suppose souvent de mettre en place une  $(P + 1)$ <sup>ème</sup> voie parallèle dans le spectromètre, où est placée une poubelle optique. Si c'est le cas, il est plus intéressant d'y installer plutôt un  $(P + 1)$ <sup>ème</sup> détecteur, créant ainsi un jeu de  $P + 1$  filtres complémentaires. Plus spécifiquement, les paires de filtres qu'il est possible d'appliquer en parallèle en utilisant une matrice de micro-miroirs sont nécessairement complémentaires, chaque miroir de la matrice ne pouvant être orienté que dans deux directions différentes.

Pour ces raisons nous utilisons dans les expériences présentées dans ce chapitre des filtres complémentaires. Le développement théorique reste quant à lui le plus général possible.

### 2.3.5 Reconstruction du spectre observé

Lorsque la reconstruction est possible, la question du choix de l'estimateur du spectre observé se pose. De nombreuses possibilités existent pour construire des estimateurs ayant de bonnes propriétés, et la littérature est abondante sur le sujet (voir par exemple Garthwaite, Jolliffe et al. (2002) ou Cox et Hinkley (1974)).

Soient  $\mathbf{n}^{(1)}, \dots, \mathbf{n}^{(W)}$  des mesures effectuées en appliquant les jeux de filtres  $\mathbf{F}^{(1)}, \dots, \mathbf{F}^{(W)}$  à un rayonnement de spectre  $\mathbf{s}$ . L'acquisition est arrêtée après la détection de  $N$  photons.

En injectant  $\mathbf{F}^{(w)}$  et  $\mathbf{s}$  dans l'équation (1.15) on a

$$p_m^{(w)} = \frac{\mathbf{f}_m^{(w)} \cdot \mathbf{s}}{\sum_{\ell=1}^P \mathbf{f}_\ell^{(w)} \cdot \mathbf{s}} \quad (2.8)$$

où  $p_m^{(w)}$  est la proportion moyenne de photons détectés par le  $m^{\text{ème}}$  détecteur au pixel  $w$ . Si les filtres utilisés sont complémentaires alors on obtient en utilisant l'équation (2.7)

$$p_m^{(w)} = \mathbf{f}_m^{(w)} \cdot \mathbf{s}.$$

Soit  $\mathbf{F} = (\mathbf{F}^{(1)}, \dots, \mathbf{F}^{(W)})$  la matrice des jeux de filtres et  $\mathbf{p}$  le vecteur obtenu en concaténant les  $\mathbf{p}^{(w)}$  c'est-à-dire la transposée de  $({}^t\mathbf{p}^{(1)}, \dots, {}^t\mathbf{p}^{(W)})$ . Si les filtres sont complémentaires,

$$\mathbf{p} = {}^t\mathbf{F} \mathbf{s}. \quad (2.9)$$

L'estimateur des moments de  $p_m^{(w)}$  est simplement

$$\hat{p}_m^{(w)} = \frac{n_m^{(w)}}{N}.$$

Cet estimateur est sans biais et efficace. Rappelons ici que  $N$  est fixé a priori pour garantir une certaine probabilité d'erreur. En pratique, il peut arriver que le nombre total de photons mesuré ne soit pas exactement  $N$ . Dans ce cas, la distribution des mesures est toujours une distribution multinomiale et les développements précédents restent valides, mais il faut remplacer  $N$  par le nombre total de photons réellement mesurés.

Lorsque  ${}^t\mathbf{F}$  est inversible à gauche, on peut estimer le spectre observé avec

$$\hat{\mathbf{s}} = {}^t\mathbf{F}^+ \hat{\mathbf{p}} \quad (2.10)$$

où  ${}^t\mathbf{F}^+ = (\mathbf{F} {}^t\mathbf{F})^{-1} \mathbf{F}$  est la pseudo-inverse de  ${}^t\mathbf{F}$  et  $\hat{\mathbf{p}}$  le vecteur obtenu en concaténant les  $\hat{p}_m^{(w)}$  c'est-à-dire la transposée de  $({}^t\hat{\mathbf{p}}^{(1)}, \dots, {}^t\hat{\mathbf{p}}^{(W)})$ . Il est aisé de montrer que cet estimateur est sans biais, mais il n'est pas efficace. Sa variance est néanmoins proche de la borne de Cramér-Rao dans les exemples que nous présentons dans ce manuscrit, comme dans celui de la figure 2.10. Ces propriétés et sa simplicité de calcul numérique en font un estimateur de choix bien que les spectres reconstruits ne soient pas nécessairement positifs.

## 2.4 Problème d'optimisation des filtres

Dans la section précédente, nous avons proposé de répartir une diversité de jeux de filtres sur la zone à imager afin d'être non seulement capable de classer chaque pixel parmi les espèces de référence mais également de reconstruire les spectres rencontrés en combinant des mesures effectuées sur plusieurs pixels.

Vouloir réaliser ces deux tâches antagonistes à partir des mêmes mesures rend l'optimisation du filtrage délicate. Nous posons dans cette section le problème d'optimisation associé à la construction de filtres permettant de réaliser à la fois classification et reconstruction avec de bonnes performances, et proposons une méthode naïve pour atteindre cet objectif.

### 2.4.1 Critères à optimiser

#### Critère de performance en classification

L'objectif principal de l'acquisition est de réaliser la classification des mesures parmi les espèces de référence avec une faible probabilité d'erreur. Comme nous l'avons vu dans la section 1.3.3, il est possible de quantifier les performances en classification d'un jeu de filtres  $\mathbf{F}^{(w)}$  avec sa borne de Bhattacharyya  $\mathcal{B}^{(w)}$ . Cette borne majore la probabilité d'erreur du classifieur optimal de l'équation (1.18).

Soient  $\mathbf{s}_1, \dots, \mathbf{s}_M$  les spectres des  $M$  espèces de référence et  $N$  le nombre de photons à détecter en chaque pixel. En injectant ces grandeurs dans l'équation (1.20), on a

$$\mathcal{B}^{(w)} = \frac{1}{M} \sum_u \sum_{v>u} \left[ \mathbf{q}_u^{(w)} \cdot \mathbf{q}_v^{(w)} \right]^N$$

où  $[\mathbf{q}_u^{(w)}]_m = (p_{m,u}^{(w)})^{1/2}$  avec  $p_{m,u}^{(w)}$  défini comme dans (2.8).

Pour maximiser les performances en classification d'un ensemble de  $W$  jeux de filtres, une possibilité est de minimiser la moyenne des bornes de Bhattacharyya, c'est-à-dire  $\bar{\mathcal{B}}$  telle que

$$\bar{\mathcal{B}} = \frac{1}{W} \sum_{w=1}^W \mathcal{B}^{(w)}.$$

Une autre possibilité est de minimiser  $\max_w \mathcal{B}^{(w)}$  mais n'avons pas constaté de différence notable entre ces deux critères pour l'optimisation des filtres.

#### Critère de performance en reconstruction

Comme détaillé dans la section 2.3.3, il est intéressant de minimiser la borne de Cramér-Rao pour être capable de reconstruire les spectres rencontrés avec de bonnes performances. En particulier, il est possible d'utiliser la trace de la matrice de Cramér-Rao  $\mathbf{\Upsilon}$  comme critère de performance en reconstruction, c'est-à-dire

$$\text{tr } \mathbf{\Upsilon} = \sum_{k=1}^{K-1} [\mathbf{\Upsilon}]_{k,k}.$$

Minimiser ce critère est équivalent à minimiser la somme des erreurs quadratiques moyennes de tout estimateur sans biais et efficace du spectre observé, soit  $\sum_{k=1}^{K-1} \mathbb{E} [(s(k) - \hat{s}(k))^2]$  où  $\hat{s}(k)$  est l'estimation de  $s(k)$ .

L'objectif de la reconstruction est de tester si l'espèce observée correspond à l'une des espèces chimiques de référence. Dans ce cadre, le spectre observé n'est pas toujours l'un des spectres de référence. On considère donc que le spectre à reconstruire est inconnu a priori. Or,  $\text{tr } \mathbf{\Upsilon}$  en dépend. Pour résoudre ce problème, nous avons empiriquement choisi de calculer  $\mathbf{\Upsilon}$  pour le spectre plat, qui est tel que  $s(k) = 1/K \forall k$ . D'autres solutions sont envisageables, notamment si l'on dispose d'informations sur les spectres qu'il est possible d'observer. Sauf précision contraire  $\mathbf{\Upsilon}$  désigne dans la suite de ce chapitre la matrice de Cramér-Rao calculée pour le spectre plat.

## 2.4.2 Optimisation et sélection d'une diversité de jeux de filtres

Nous montrons dans cette section qu'il existe une diversité de jeux de filtres ayant de bonnes performances en classification, comme nous l'avons affirmé précédemment. Partant de ce constat, nous présentons un premier algorithme d'optimisation des filtres. Cet algorithme consiste à optimiser un grand nombre de jeux de filtres pour la classification puis à en sélectionner  $W$  en maximisant les performances en reconstruction. Nous illustrons sur un exemple que cette méthode naïve permet de reconstruire les spectres rencontrés tout en limitant la perte de performance en classification.

### Diversité de jeux de filtres ayant de bonnes performances en classification

Pour construire une diversité de jeux de filtres permettant de classer les mesures parmi des espèces de référence données avec de bonnes performances, nous adoptons une approche itérative. Plus précisément, nous appliquons l'algorithme de Réfrégier, Chevallier et al. (2019) à un grand nombre d'initialisations aléatoires, ce qui revient à la procédure qui suit.

**Algorithme.** On tire tout d'abord uniformément  $V$  jeux de  $P$  filtres binaires et orthogonaux. Pour chacun de ces jeux de filtres, la procédure suivante est appliquée. À chaque itération le jeu de filtres courant est modifié aléatoirement comme décrit dans l'annexe D. Si cette modification conduit à une borne de Bhattacharyya plus importante qu'auparavant, elle est rejetée et il est interdit à l'algorithme de la tirer de nouveau. Dans le cas contraire, elle est acceptée et la liste des modifications interdites est remise à zéro. Cette procédure est appliquée jusqu'à ce qu'il n'existe plus de modification autorisée, ce qui arrive lorsque l'algorithme a exploré tous les jeux de filtres atteignables en une itération à partir du jeu de filtres courant et qu'aucun d'entre eux ne conduit à une meilleure borne de Bhattacharyya. Chaque jeu de filtres obtenu de la sorte minimise  $\mathcal{B}$  localement.

**Mise en œuvre sur un exemple.** Afin d'illustrer la méthode proposée, nous présentons le résultat de l'optimisation de  $V = 10^4$  jeux de  $P = 2$  filtres complémentaires pour la classification parmi les trois espèces dont les spectres apparaissent en figure 2.1. On considère que l'acquisition est arrêtée après la détection de  $N = 250$  photons. La figure 2.8 présente la distribution de  $\mathcal{B}^{(w)}$  pendant l'optimisation : à l'initialisation, cinq pas avant convergence et après convergence.

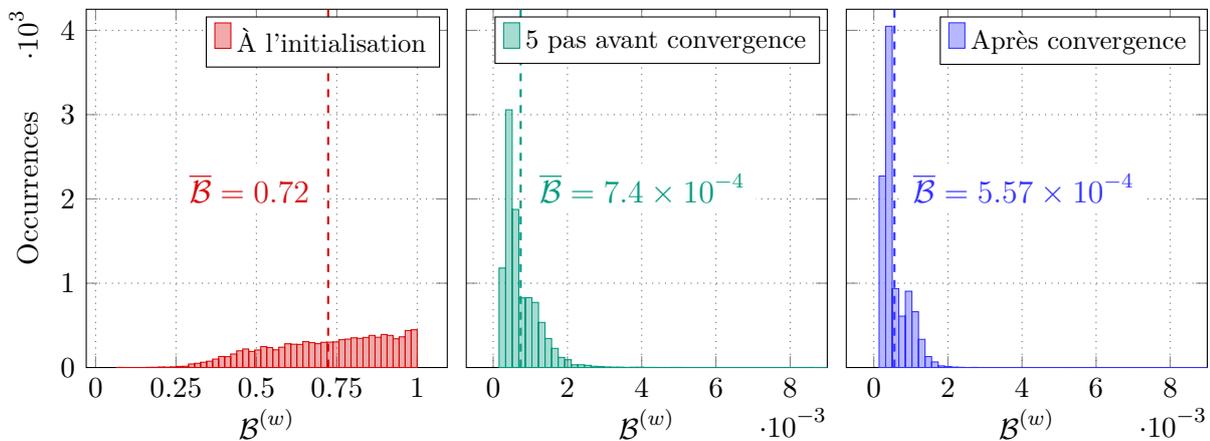


FIGURE 2.8 – Exemple d'évolution de la distribution de  $\mathcal{B}^{(w)}$  lors de l'optimisation.

On peut constater sur ces graphiques que l'optimisation a permis de réduire la borne de Bhattacharyya de plusieurs ordre de grandeur. Ainsi,  $\mathcal{B}^{(w)}$  est comprise entre  $6.46 \times 10^{-2}$  et 1 avec une moyenne  $\bar{\mathcal{B}} = 0.72$  pour les jeux de filtres initiaux, et elle est comprise entre  $1.43 \times 10^{-4}$  et  $2.77 \times 10^{-3}$  avec  $\bar{\mathcal{B}} = 5.57 \times 10^{-4}$  après convergence. On peut également remarquer que l'effet du retour cinq pas avant convergence sur les performances de classification est limité. Pour les jeux de filtres correspondant, la borne de Bhattacharyya est comprise entre  $1.58 \times 10^{-4}$  et  $8.63 \times 10^{-3}$  avec  $\bar{\mathcal{B}} = 7.4 \times 10^{-4}$ .

### Sélection des jeux de filtres ayant les meilleures performances en reconstruction

Puisque l'on dispose d'un ensemble constitué d'une diversité de jeux de filtres ayant de bonnes performances en classification, une idée qui vient naturellement pour concilier classification et reconstruction est de sélectionner les  $W$  jeux de filtres de cet ensemble permettant d'effectuer la reconstruction avec les meilleures performances possibles.

L'approche que nous avons choisie consiste à sélectionner les jeux de filtres un à un en comparant à chaque itération tous les jeux de filtres de l'ensemble pour ne retenir que le meilleur. Au début de la sélection, l'objectif est d'inverser la matrice d'information de Fisher le plus rapidement possible afin de pouvoir reconstruire les spectres rencontrés à partir d'un minimum de mesures. Lorsque cette matrice est inversible, la sélection minimise la trace de la borne de Cramér-Rao calculée pour un spectre plat.

Il arrive que les jeux de filtres optimisés pour la classification ne présentent pas une diversité suffisante à permettre de reconstruire les spectres rencontrés, ce qui signifie que l'optimisation a convergé vers des jeux de filtres trop proches. Si l'on choisit  $V \gg W_{\min}$ , où  $W_{\min}$  est le nombre minimum de jeux de filtres nécessaires à la reconstruction (voir proposition 2), il est en revanche probable que la diversité de jeux de filtres de l'ensemble initial ait été suffisante à pouvoir inverser la matrice d'information de Fisher. Une possibilité est alors de remonter le cours de l'optimisation pour se rapprocher de cette diversité initiale sans trop diminuer les performances en classification. On peut en effet supposer que les chemins parcourus des différentes initialisations aux filtres obtenus après convergence sont au moins partiellement disjoints. Nous avons constaté qu'ajouter les  $V$  jeux de filtres obtenus en remontant cinq pas avant la convergence (comme dans la figure 2.8) aux  $V$  jeux de filtres obtenus après la convergence permet dans la majeure partie des cas d'obtenir la diversité nécessaire à la reconstruction. On obtient ainsi à un ensemble de  $2 \times V$  jeux de filtres.

**Algorithme.** À l'initialisation, aucun jeu de filtres n'est sélectionné. Une itération de l'algorithme consiste à tester l'ajout de chaque jeu de filtres de l'ensemble initial à la sélection pour n'en choisir finalement qu'un seul. Lorsque la matrice d'information de Fisher est singulière, l'algorithme sélectionne le jeu de filtres permettant d'accroître le plus son rang. Ce seul critère ne suffit parfois pas à distinguer un jeu de filtres en particulier, par exemple lors des premières itérations où il est aisé d'ajouter de la diversité à la sélection. Quand il y a ambiguïté, le jeu de filtres sélectionné est celui permettant de minimiser la trace de  $\mathbf{I}^+$  la pseudo-inverse de  $\mathbf{I}$ .

Lorsque la matrice d'information de Fisher est inversible, l'algorithme sélectionne à chaque itération le jeu de filtres qui permet de réduire le plus  $\text{tr } \mathbf{Y}$  la trace de la matrice de Cramér-Rao. Si deux jeux de filtres aboutissent à la même  $\text{tr } \mathbf{Y}$ , celui ayant la plus petite borne de Bhattacharyya est sélectionné, ce qui est équivalent à minimiser  $\bar{\mathcal{B}}$ .

On laisse la possibilité à l'algorithme de sélectionner le même jeu de filtres plusieurs fois. Si cela ne permet évidemment pas de gagner en diversité, rien n'indique que cela ne peut être avantageux une fois l'inversibilité de  $\mathbf{I}$  atteinte.

**Implémentation.** Le nombre d'itérations de l'algorithme est en général limité puisque  $W$  est la plupart du temps de l'ordre de  $10^2$ , mais il est souvent nécessaire que  $V \sim 10^3 - 10^4$  pour obtenir une diversité suffisante. Il est donc important de pouvoir tester l'ajout d'un jeu de filtres à la sélection rapidement.

Le calcul et l'inversion de la matrice d'information de Fisher  $\mathbf{I}$  sont les étapes les plus chronophages de la sélection. La matrice  $\mathbf{I}$  s'écrivant comme une somme sur les jeux de filtres sélectionnés (voir équation (2.4)), il est possible pour la calculer d'ajouter simplement le terme correspondant au jeu de filtres testé à la matrice d'information de Fisher courante. Le temps de calcul de  $\mathbf{I}$  est alors négligeable par rapport à celui de son inversion, et le temps nécessaire à la sélection est environ proportionnel au nombre total d'inversions à effectuer, soit  $W \times V$ . Puisque les calculs de ce critère pour les différents jeux de filtres testés à chaque itération sont indépendants les uns des autres, il est possible de les paralléliser.

**Mise en œuvre sur un exemple.** Pour illustrer le fonctionnement de la méthode, on sélectionne  $W = 254$  jeux de filtres parmi l'ensemble de  $2 \times V = 2 \times 10^4$  jeux de filtres que nous avons construit précédemment. Choisir  $W = 254 = 2 \times W_{\min}$  permet de montrer l'évolution des critères de performance avant et après l'inversibilité de la matrice d'information de Fisher. La sélection des jeux de filtres a nécessité environ 20 minutes sur une machine possédant 20 processeurs.

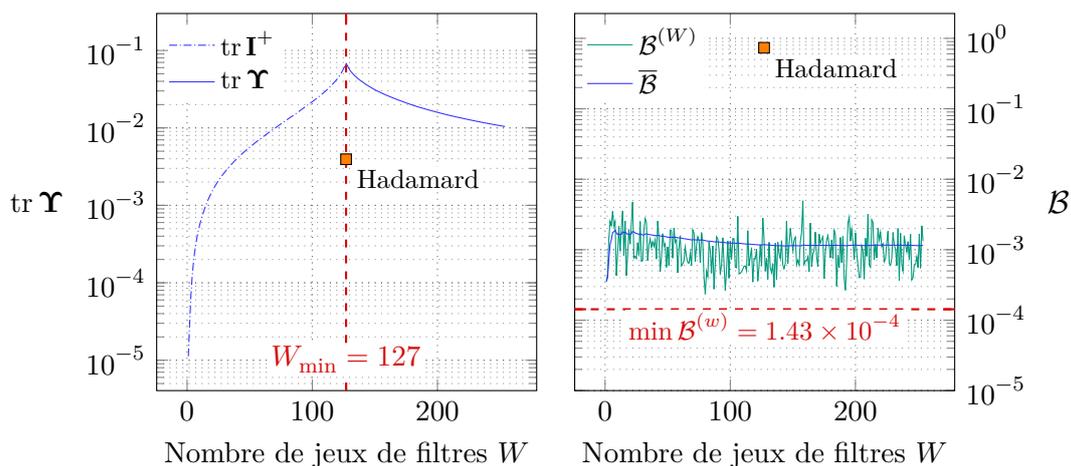


FIGURE 2.9 – Exemple d'évolution des critères de performance lors de la sélection.

Sur la gauche de la figure 2.9 apparaît la trace de la pseudo-inverse ou de l'inverse de la matrice d'information de Fisher en fonction du nombre de jeux de filtres sélectionnés. Jusqu'à  $W = W_{\min} = 127$  la matrice d'information de Fisher est singulière et les jeux de filtres sélectionnés sont ceux qui minimisent la trace de sa pseudo-inverse parmi ceux qui permettent d'accroître le plus son rang. La trace de  $\mathbf{I}^+$  croît au fur et à mesure des itérations car le nombre de paramètres qu'il est possible de reconstruire à partir des mesures est de plus en plus important. Lorsque  $W \geq W_{\min}$  la matrice d'information de Fisher est inversible et il est alors possible de minimiser la trace de la matrice de Cramér-Rao. Entre  $W = 127$  et  $W = 254$ ,  $\text{tr } \mathbf{Y}$  passe de  $6.86 \times 10^{-2}$  à  $1.04 \times 10^{-2}$ .

Sur la droite de la figure 2.9 apparaissent à la fois la borne de Bhattacharyya moyenne  $\bar{\mathcal{B}}$  à l'issue de l'itération  $W$  et la borne de Bhattacharyya  $\mathcal{B}^{(W)}$  du jeu de filtres sélectionné lors de cette itération. À la fin de la sélection  $\bar{\mathcal{B}} = 1.15 \times 10^{-3}$ , ce qui est environ 9 fois plus élevé que la borne de Bhattacharyya du jeu de filtres ayant les meilleures performances en classification, qui vaut  $1.43 \times 10^{-4}$ . On ne peut cependant pas réaliser la reconstruction en utilisant uniquement ce jeu de filtres.

Les performances des 127 premiers jeux de filtres sélectionnés peuvent être comparées à celles de la base de Hadamard. Cet ensemble de  $K - 1$  paires de filtres binaires, orthogonaux et complémentaires est souvent utilisé dans la littérature pour la reconstruction de spectres. Il apparaît sous la forme d'un carré orange sur la figure 2.9. L'utiliser permet d'obtenir de meilleures performances en reconstruction puisque dans ce cas  $\text{tr } \Upsilon = 3.94 \times 10^{-3}$  alors qu'avec les 127 premiers jeux de filtres sélectionnés  $\text{tr } \Upsilon = 6.86 \times 10^{-2}$ . On perd cependant largement en performances de classification, passant de  $\bar{\mathcal{B}} = 1.17 \times 10^{-3}$  avec nos jeux de filtres à  $\bar{\mathcal{B}} = 0.73$  avec ceux de la base de Hadamard, ce qui est inexploitable.

### Test des performances des jeux de filtres construits

Ici, l'objectif est de tester les performances en classification et en reconstruction des  $W = 254$  jeux de filtres construits. Pour tester leurs performances en classification, nous simulons pour chacun d'entre eux  $10^5$  mesures distribuées selon la loi multinomiale de l'équation (1.17) pour chaque spectre de référence  $\mathbf{s}_p$ . Nous calculons alors la probabilité d'erreur de classification empirique, que nous comparons avec la borne de Bhattacharyya. Cette comparaison apparaît en bas de la figure 2.10. Dans cet exemple, la probabilité d'erreur empirique est toujours inférieure à  $\mathcal{B}^{(w)}/2$ .

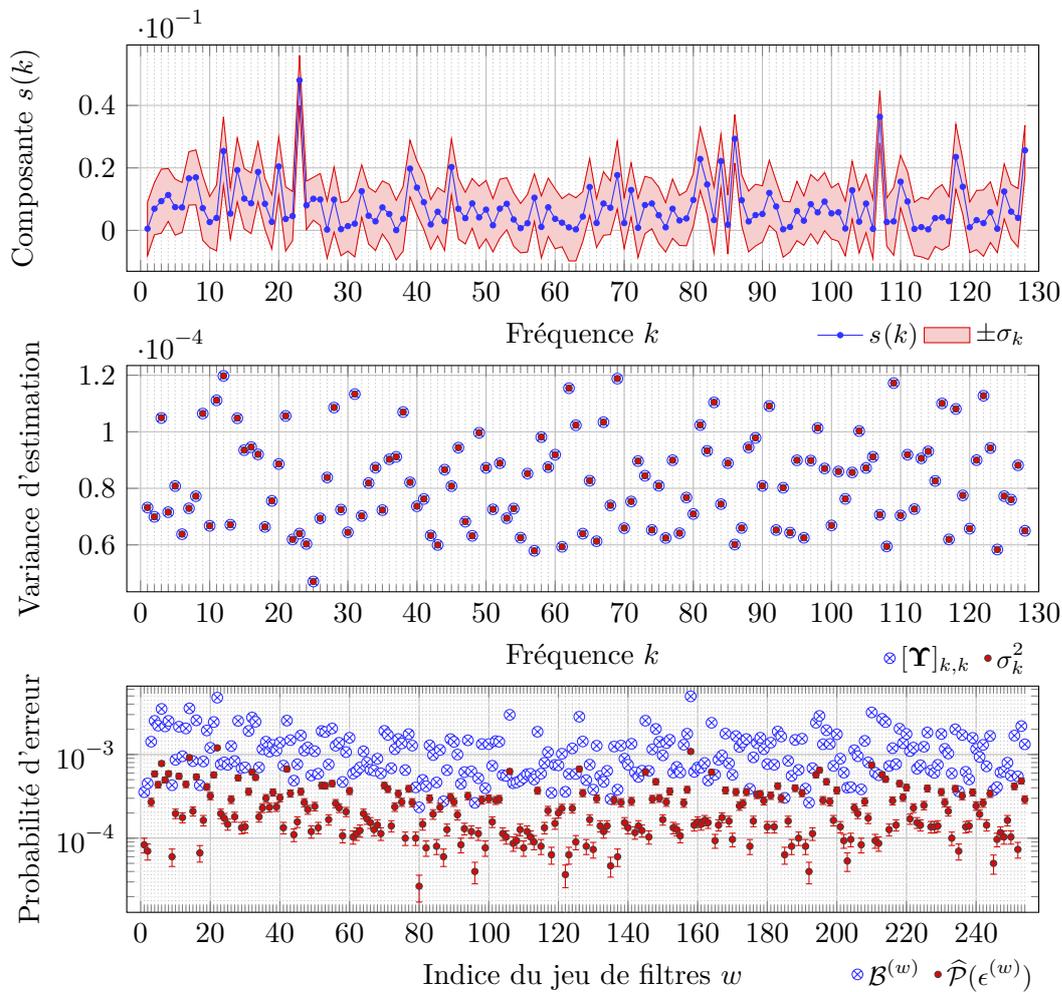


FIGURE 2.10 – Performances des jeux de filtres construits. En haut : spectre vrai et écart-type de l'estimateur de la pseudo-inverse. Au milieu : borne de Cramér-Rao et variance de l'estimateur de la pseudo-inverse. En bas : borne de Bhattacharyya et probabilité d'erreur empirique.

Pour visualiser les performances en reconstruction qu'il est possible d'obtenir avec les jeux de filtres, nous avons calculé la borne de Cramér-Rao et la covariance de l'estimateur de la pseudo-inverse de l'équation (2.10) pour le spectre  $\mathbf{s}$  apparaissant en haut de la figure 2.10. On peut constater que l'estimateur de la pseudo-inverse est presque efficace sur cet exemple. Les performances en reconstruction obtenues avec les filtres construits et l'estimateur de la pseudo-inverse ne permettent ici que d'estimer l'allure générale du spectre. S'il est nécessaire que la reconstruction soit plus précise, une possibilité est d'utiliser plus de pixels pour la réaliser, ce qui se fera nécessairement au détriment de la taille de la plus petite zone sur laquelle il est possible de reconstruire le spectre rencontré. On peut également choisir d'augmenter  $N$ , ce qui améliorera les performances globales mais allongera l'acquisition.

### Limitations de la méthode

L'approche décrite dans cette section comporte une large part d'empirisme, par exemple dans le choix d'ajouter aux jeux de filtres obtenus après convergence ceux correspondant à une des itérations antérieures ou dans celui de minimiser la trace de la pseudo-inverse de la matrice d'information de Fisher lorsque celle-ci n'est pas inversible.

De plus, bien que cette approche produise des filtres réalisant un certain compromis entre classification et reconstruction, il est impossible de garantir qu'il n'existe pas de filtres plus performants en ces deux tâches. En particulier, réaliser d'abord l'optimisation des performances en classification d'un ensemble de jeux de filtres puis en sélectionner un sous-ensemble sans effectuer de modification est probablement sous-optimal. En outre, il n'y a aucun contrôle sur les performances relatives en classification et en reconstruction alors que d'autres compromis peuvent être souhaitables. L'utilisateur peut en effet vouloir favoriser une tâche plutôt que l'autre, par exemple en donnant un poids plus important aux performances de reconstruction au détriment de la probabilité d'erreur de classification, ou inversement.

Il paraît donc utile de développer une nouvelle méthodologie afin de mieux traiter ce problème. Dans ce but, nous proposons dans la prochaine section une approche plus générale de l'optimisation du filtrage basée sur l'emploi de filtres à compromis optimaux.

## 2.5 Approche fondée sur les compromis optimaux

Nous avons développé une méthodologie consistant à utiliser des jeux de filtres permettant non seulement de classer chaque pixel parmi les espèces prévues par le modèle mais aussi de reconstruire les spectres rencontrés en combinant les mesures effectuées sur plusieurs pixels. Cela permet à l'analyse de gagner en robustesse en donnant notamment la possibilité de questionner les a priori utilisés pour réaliser la classification. Chacune de ces tâches peut être associée à un critère permettant de mesurer ses performances, borne de Bhattacharyya pour la classification et borne de Cramér-Rao pour la reconstruction.

Les jeux de filtres performants pour l'une ou l'autre de ces tâches ont des caractéristiques différentes. En particulier, si une diversité de jeux de filtres est nécessaire à la reconstruction il est plus avantageux de spécialiser les filtres pour réaliser la classification. Il paraît donc vain de chercher à construire des jeux de filtres maximisant simultanément les performances des deux tâches et un objectif plus raisonnable est plutôt de chercher à réaliser un compromis entre performances de classification et de reconstruction. Cette recherche de compromis est au cœur de la méthodologie que nous présentons dans cette section, qui est celle que nous avons publiée dans Justel, Galland et al. (2023a).

### 2.5.1 Principe des filtres à compromis optimaux

Il existe de nombreux compromis possibles entre performances de classification et de reconstruction, au sens où chaque ensemble de jeux de filtres aura des performances différentes et que la performance relative en chacune des deux tâches sera elle aussi variable en fonction de l'ensemble choisi. Puisqu'il n'existe pas d'optimum unique, il s'agit de savoir s'il est malgré tout possible de distinguer certains ensembles de jeux de filtres des autres. Une manière plus simple d'aborder ce problème est de se poser plutôt la question inverse : Dans quel cas est-il possible de préférer de manière objective un ensemble de jeux de filtres à un autre ?

Soient deux ensembles de jeux de filtres dont les performances sont évaluées pour le même problème de classification et de reconstruction. Si l'un de ces deux ensembles permet de mieux classer les mesures parmi les espèces de référence que l'autre mais que ce dernier permet de mieux reconstruire les spectres rencontrés alors il est impossible de les hiérarchiser objectivement : chacun pourrait être préféré à l'autre dépendamment des besoins de l'utilisateur. Si en revanche l'un de ces ensembles de jeux de filtres permet à la fois de faire mieux en classification et en reconstruction que l'autre, alors ce dernier est objectivement plus mauvais puisqu'il n'existe aucune situation qui pourrait conduire à le préférer. C'est dans ce raisonnement que se trouve les fondements de la notion suivante, utilisée par Réfrégier (1990, 1991) et Figue et Réfrégier (1993) pour concevoir des filtres optiques pour la reconnaissance de formes.

**Définition 2** (Filtres à compromis optimaux). Un ensemble de jeux de filtres est dit à compromis optimaux pour un problème donné s'il n'existe pas d'ensemble de jeux de filtres ayant de meilleures performances dans toutes les tâches considérées au regard des critères évalués. En l'occurrence, si  $\mathcal{F}_{\text{opt}}$  est constitué des ensembles de jeux de filtres à compromis optimaux et que  $\mathbf{F} = (\mathbf{F}^{(1)}, \dots, \mathbf{F}^{(W)})$ ,

$$\mathbf{F} \in \mathcal{F}_{\text{opt}} \Leftrightarrow \forall \mathbf{F}' \begin{cases} \bar{\mathcal{B}}(\mathbf{F}') \geq \bar{\mathcal{B}}(\mathbf{F}) \\ \text{ou} \\ \text{tr } \Upsilon(\mathbf{F}') \geq \text{tr } \Upsilon(\mathbf{F}). \end{cases} \quad (2.11)$$

Par définition, tout ensemble de jeux de filtres non optimal en ce sens possède au moins une contrepartie plus performante dans tous les domaines et ne doit donc pas être utilisé si cela ne dépend que de ces critères. L'utilisateur devrait donc choisir l'ensemble de jeux de filtres correspondant le mieux à ses besoins uniquement parmi les ensembles de jeux de filtres à compromis optimaux.

### 2.5.2 Recherche des filtres à compromis optimaux

La recherche des filtres à compromis optimaux est un problème d'optimisation multi-critères dans un espace discret de grande dimension. Ce problème est difficile du fait de la non-convexité des critères au regard des filtres et du fait que ceux-ci soient binaires. Il existe  $P^{K \times W} / W!$  ensembles de  $W$  jeux de  $P$  filtres binaires, orthogonaux et complémentaires de taille  $K$ . Au vu des valeurs usuelles de  $W$ ,  $P$  et  $K$  il est inenvisageable de tester l'optimalité d'un ensemble de jeux de filtres, puisque cela supposerait de comparer ses performances à celles de tous les autres ensembles de jeux de filtres, ce qui représente un très grand nombre d'opérations. Par exemple, il existe  $2^{16256} / 128! \approx 10^{4680}$  ensembles de  $W = 127$  paires de filtres binaires orthogonaux et complémentaires de taille  $K = 128$ , là où un processeur standard réalise  $\sim 10^9$  opérations par seconde.

Une solution à ce problème est de guider la recherche en essayant de minimiser une combinaison convexe des différents critères considérés, c'est-à-dire un critère  $\mathcal{C}(\mathbf{F}, \lambda)$  tel que

$$\mathcal{C}(\mathbf{F}, \lambda) = (1 - \lambda) \operatorname{tr} \mathbf{Y}(\mathbf{F}) + \lambda \bar{\mathcal{B}}(\mathbf{F}), \quad \lambda \in [0, 1]. \quad (2.12)$$

Comme démontré par Réfrégier (1990), minimiser ce critère aboutit effectivement à des ensembles de jeux de filtres à compromis optimaux. En faisant varier le poids  $\lambda$  de cette combinaison convexe on trouve des filtres réalisant différents compromis entre les tâches considérées. On peut ainsi tracer ce que l'on appelle abusivement la *courbe des compromis optimaux*, qui est la courbe dans l'espace des critères ( $\operatorname{tr} \mathbf{Y}, \bar{\mathcal{B}}$ ) correspondant aux ensembles de jeux de filtres minimisant  $\mathcal{C}(\mathbf{F}, \lambda)$  pour  $\lambda \in [0, 1]$ . Puisqu'il s'agit d'un espace discret, cette courbe est discontinue. Nous démontrons en annexe E qu'elle est également convexe. En conséquence, il n'est possible en minimisant le critère  $\mathcal{C}(\mathbf{F}, \lambda)$  que d'atteindre les filtres optimaux situés sur l'enveloppe convexe de l'espace des critères atteignables.

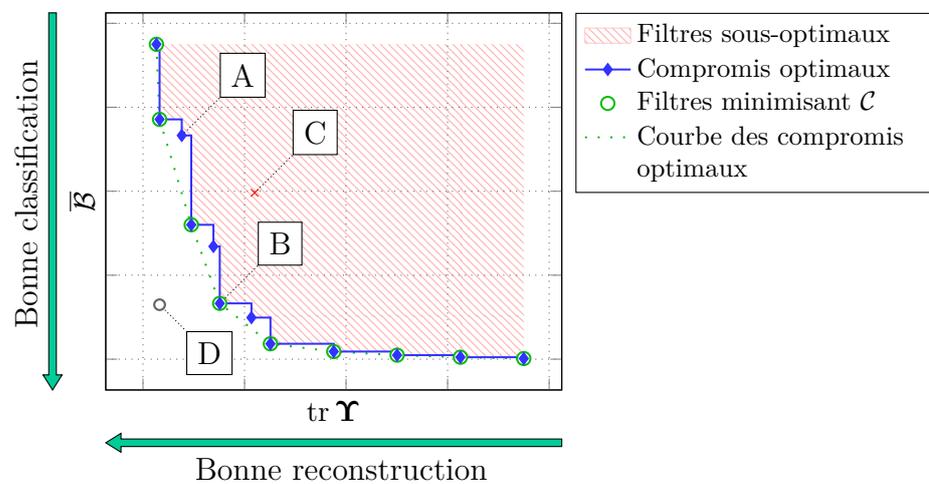


FIGURE 2.11 – Principe des filtres à compromis optimaux

Pour visualiser les concepts introduits, nous les avons schématisés en figure 2.11. L'espace des critères ( $\operatorname{tr} \mathbf{Y}, \bar{\mathcal{B}}$ ) est  $(\mathbb{R}_+)^2$ . Dans cet espace, seuls les points correspondant à des ensembles de jeux de filtres sont accessibles. Puisque le nombre d'ensembles de jeux de filtres binaires d'une taille donnée est fini, l'espace des critères atteignables est lui-même discret et restreint à une certaine zone du plan. Cette zone correspond à l'union de la courbe bleue et de la zone hachurée en rouge sur le schéma. En particulier, il est impossible d'atteindre le point D.

Sur la figure 2.11 on a représenté en bleu le bord inférieur gauche de l'espace des critères atteignables, c'est-à-dire l'ensemble des points de cet espace pour lesquels il est impossible de trouver d'autres points à la fois plus bas et plus à gauche sans le quitter. Améliorer un ensemble de jeux de filtres à la fois en classification et en reconstruction implique justement de trouver un autre ensemble de jeux de filtres localisé plus bas et plus à gauche, et les points de ce bord forment donc la totalité des ensembles de jeux de filtres à compromis optimaux. Tout point situé plus haut ou plus à droite est sous-optimal car il est possible de trouver un point du bord ayant de meilleures valeurs des deux critères. C'est le cas du point C, qui est inférieur en classification et en reconstruction au point B. À l'inverse, il est nécessaire pour se déplacer sur le bord d'accepter une perte en une tâche pour gagner en l'autre. Par exemple, passer du point A au point B permet de diminuer  $\bar{\mathcal{B}}$  au prix d'une augmentation de  $\operatorname{tr} \mathbf{Y}$ .

La forme du bord de l'espace des critères atteignables est quelconque contrairement à la courbe des compromis optimaux qui est convexe. En minimisant  $\mathcal{C}(\mathbf{F}, \lambda)$  on n'atteint donc que les points situés sur l'enveloppe convexe de ce bord. Par exemple, l'ensemble de jeux de

filtres à compromis optimaux correspondant au point A ne minimise  $\mathcal{C}(\mathbf{F}, \lambda)$  pour aucun  $\lambda \in [0, 1]$  et A n'est donc pas sur ce que nous appelons la courbe des compromis optimaux. En conséquence, il sera impossible de trouver ce point avec la méthode proposée.

### 2.5.3 Algorithme de recherche des filtres à compromis optimaux

Il s'agit ici de trouver les ensembles de jeux de filtres minimisant les combinaisons convexes  $\mathcal{C}(\mathbf{F}, \lambda)$  des critères considérés. Comme évoqué précédemment, l'espace des filtres binaires et orthogonaux est un espace discret de grande dimension. Le parcourir entièrement à la recherche de solutions n'est donc pas envisageable.

Nous proposons un algorithme permettant de trouver des minima locaux de  $\mathcal{C}(\mathbf{F}, \lambda)$  grâce à une marche aléatoire dans l'espace des filtres, ce qui permet d'estimer la courbe des compromis optimaux en faisant varier  $\lambda$ . La procédure utilisée est la suivante :

1. Choisir une valeur de  $\lambda$  pour laquelle optimiser  $\mathcal{C}(\mathbf{F}, \lambda)$ .
2. Construire les ensembles de jeux de filtres servant d'initialisations.
3. Modifier aléatoirement les filtres aussi longtemps que cela permet de diminuer  $\mathcal{C}(\mathbf{F}, \lambda)$ .
4. Sélectionner les filtres associés à la plus petite valeur de ce critère.
5. Recommencer la procédure avec d'autres valeurs de  $\lambda$  jusqu'à ce que la courbe des compromis optimaux soit échantillonnée avec la précision désirée.

Dans la suite de cette section, nous revenons plus en détails sur :

- ▷ Comment choisir les valeurs de  $\lambda$  successives pour obtenir un bon échantillonnage de la courbe des compromis optimaux.
- ▷ Quels sont les ensembles de jeux de filtres pouvant servir d'initialisation.
- ▷ Quelle est la procédure de modification aléatoire des ensembles de jeux de filtres.
- ▷ Comment calculer rapidement  $\mathcal{C}(\mathbf{F}, \lambda)$  au fil de la marche aléatoire.

#### Choix des valeurs de $\lambda$

Pour tracer la courbe des compromis optimaux, il faut optimiser le critère  $\mathcal{C}(\mathbf{F}, \lambda)$  pour différentes valeurs de  $\lambda$ . L'objectif est d'obtenir une diversité de compromis et donc un échantillonnage suffisamment fin du bord de l'espace des critères atteignables, et il faut pour cela choisir des valeurs de  $\lambda$  adéquates. Sélectionner les valeurs de  $\lambda$  n'est pas aisé car il est difficile d'anticiper la position dans l'espace des critères des filtres minimisant  $\mathcal{C}(\mathbf{F}, \lambda)$ . Nous décrivons ici deux possibilités pour choisir les poids successifs.

- ▷ **Valeurs linéairement espacées.** La borne de Bhattacharyya moyenne et la trace de la borne de Cramér-Rao varient de plusieurs ordres de grandeur au cours de l'optimisation. En conséquence, l'optimisation de  $\mathcal{C}(\mathbf{F}, \lambda)$  pour des valeurs de  $\lambda$  ordonnées et linéairement espacées entre 0 et 1 conduit souvent à une courbe ayant des points concentrés dans les zones où les deux critères sont du même ordre de grandeur et épars ailleurs. Si l'on veut malgré tout utiliser cette méthode, une solution est de minimiser  $\mathcal{C}_{\log}(\mathbf{F}, \lambda) = (1 - \lambda) \log(\text{tr } \mathbf{Y}(\mathbf{F})) + \lambda \log(\overline{\mathbf{B}}(\mathbf{F}))$  où les logarithmes semblent mitiger les problèmes d'ordres de grandeur. La courbe est alors convexe dans  $(\log(\text{tr } \mathbf{Y}), \log(\overline{\mathbf{B}}))$ .
- ▷ **Dichotomie.** Il est également possible de choisir la valeur des poids dynamiquement. On peut par exemple utiliser l'heuristique suivante, dont l'objectif est d'ajouter des points à la courbe à l'endroit où l'échantillonnage est le plus lâche. Les deux premières valeurs du poids  $\lambda$  sont 0 et 1. Par la suite, les  $\lambda$  sélectionnés sont au milieu des valeurs

déjà optimisées pour lesquelles les filtres correspondent aux points voisins les plus éloignés dans l'espace des critères. Autrement dit, si  $\mathbf{F}_{\lambda_1}$  et  $\mathbf{F}_{\lambda_2}$  sont ces ensembles de jeux de filtres, alors le prochain poids à optimiser est  $\lambda = (\lambda_1 + \lambda_2)/2$ .

### Choix des initialisations

La construction des jeux de filtres servant à initialiser l'optimisation pour un  $\lambda$  donné peut se faire de plusieurs manières. Tout ensemble de jeux de filtres satisfaisant la contrainte d'inversibilité de la matrice d'information de Fisher est une initialisation admissible pour l'algorithme présenté. Nous décrivons ici quelques possibilités d'initialisation.

- ▷ **Filtres aléatoires.** Il est possible d'employer des ensembles de jeux de filtres tirés uniformément pour initialiser l'algorithme. Il suffit en général de quelques dizaines de tirages pour obtenir des ensembles de filtres dont la matrice d'information de Fisher est inversible. Nous avons observé que ces ensembles de filtres ont de mauvaises performances en classification, des performances en reconstruction variables et sont souvent éloignés de la courbe des compromis optimaux. Cette méthode permet néanmoins d'obtenir une diversité d'initialisations à faible coût.
- ▷ **Algorithme de la section 2.4.2.** Une autre possibilité est d'utiliser l'algorithme présenté dans la section précédente, qui correspond à celui de Justel, Galland et al. (2022). Même si on peut supposer qu'un ensemble de jeux de filtres construit avec cet algorithme n'est pas trop éloigné de la courbe, il reste néanmoins difficile de savoir à quel niveau il se situe.
- ▷ **Optimisation de  $\mathcal{B}^{(w)}$  et sélection sur  $\bar{\mathcal{B}}$ .** Il est également possible d'utiliser  $\bar{\mathcal{B}}$  plutôt que  $\text{tr } \mathbf{Y}$  et  $\text{tr } \mathbf{I}^+$  comme critère de sélection dans l'algorithme de la section 2.4.2, toujours en s'assurant de maximiser le rang de  $\mathbf{I}$  lorsque celle-ci n'est pas inversible. Nous avons constaté que ces filtres sont en général proche du minimiseur de  $\mathcal{C}(\mathbf{F}, \lambda = 1)$ . Ce type d'initialisation est donc en général localisé à l'extrémité droite de la courbe.
- ▷ **Base de Hadamard.** Un ensemble de paires de filtres binaires, orthogonaux, et complémentaires souvent utilisé dans la littérature pour la reconstruction de spectres est la base de Hadamard (voir DeVerse, Hammaker et al. (2000)). Ces filtres minimisent en général  $\mathcal{C}(\mathbf{F}, \lambda = 0)$  et sont donc situés à l'extrémité gauche de la courbe.
- ▷ **Ensemble de jeux de filtres déjà optimisé.** Il est possible de construire la courbe de manière itérative, c'est-à-dire en repartant d'un ensemble de jeux de filtres résultant de l'optimisation effectuée pour un  $\lambda$  donné pour construire les filtres pour un autre  $\lambda$ . Nous avons constaté que construire la courbe de proche en proche peut accélérer la convergence. En revanche, l'espace exploré est alors moins important.

Afin d'accroître les chances d'atteindre la véritable courbe des compromis optimaux, on répète la procédure en utilisant de multiples initialisations pour chaque  $\lambda$ , en ne gardant à chaque fois que le meilleur ensemble de jeux de filtres construit au cours de ces différents essais. Cela permet d'explorer une plus grande diversité d'ensembles de jeux de filtres.

### Modifications aléatoires.

Une fois une valeur de  $\lambda$  et une initialisation choisies, l'algorithme consiste à tester successivement de multiples modifications aléatoires de l'ensemble de jeux de filtres initial. Une modification n'est conservée que si elle permet de diminuer la valeur de  $\mathcal{C}(\mathbf{F}, \lambda)$ . Plus précisément, on applique la procédure qui suit.

À chaque itération, l'indice d'un jeu de filtres  $w$  et celui d'un intervalle fréquentiel  $k$  sont choisis aléatoirement et les coefficients  $f_1^{(w)}(k), \dots, f_P^{(w)}(k)$  correspondant sont modifiés

en respectant les contraintes sur la forme des filtres (orthogonalité, binarité, éventuelle complémentarité). Cette procédure est décrite dans l'annexe D. Si la modification aboutit à une matrice d'information de Fisher  $\mathbf{I}$  inversible et permet de réduire le critère  $\mathcal{C}(\mathbf{F}, \lambda)$  elle est acceptée, sinon elle est rejetée et il est interdit à l'algorithme de re-tester à nouveau les mêmes jeux de filtres. La règle portant sur l'inversibilité de  $\mathbf{I}$  est appliquée peu importe le poids pour lequel le critère est évalué. C'est en particulier le cas pour  $\lambda = 1$ , autrement dit même si seules les performances en classification sont optimisées. L'optimisation est arrêtée lorsque l'ensemble de jeux de filtres courant est un optimum local, c'est-à-dire que toutes les modifications possibles de cet ensemble qu'il est possible de réaliser en une itération entraînent une augmentation du critère  $\mathcal{C}(\mathbf{F}, \lambda)$ .

### Calcul rapide des critères

L'algorithme proposé est basé sur l'exploration d'un grand nombre d'ensembles de jeux de filtres différents. Il est parfois nécessaire de tester plusieurs millions d'ensembles de jeux de filtres avant de trouver un minimum local de  $\mathcal{C}(\mathbf{F}, \lambda)$ . Par conséquent, il est essentiel de pouvoir tester une modification donnée d'un ensemble de jeux de filtres très rapidement. Les étapes les plus chronophages de ce test sont le calcul de la borne de Bhattacharyya moyenne ainsi que le calcul et l'inversion de la matrice d'information de Fisher.

Comme évoqué dans la section 2.4.2, la borne de Bhattacharyya moyenne  $\bar{\mathcal{B}}$  comme la matrice d'information de Fisher  $\mathbf{I}$  s'écrivent comme des sommes sur les jeux de filtres employés (voir équation (2.4)). Puisqu'un unique jeu de filtres est modifié à chaque itération, il est possible pour calculer la nouvelle matrice  $\mathbf{I}$  de soustraire à la matrice  $\mathbf{I}$  courante le terme correspondant à la précédente version du jeu de filtres modifié et d'y ajouter celui correspondant à ce jeu de filtres après modification, et de même pour  $\bar{\mathcal{B}}$ . Même si les calculs de  $\bar{\mathcal{B}}$  et  $\mathbf{I}$  sont alors rapides, l'inversion de  $\mathbf{I}$  reste très coûteuse en temps de calcul.

La modification aléatoire d'un jeu de filtres effectuée par l'algorithme lors d'une itération est la plupart du temps équivalente à une perturbation de rang 4 ou 6 de la matrice d'information de Fisher. Il est donc possible d'utiliser la formule de Sherman-Morrison-Woodbury ou lemme d'inversion matriciel (voir Guttman (1946)) pour calculer la nouvelle matrice de Cramér-Rao à partir de la précédente en n'inversant qu'une matrice de taille  $4 \times 4$  ou  $6 \times 6$  au lieu d'inverser directement la matrice d'information de Fisher de taille  $(K - 1) \times (K - 1)$ . Une description détaillée de l'utilisation de ce lemme pour le calcul de  $\mathbf{\Upsilon}$  et des conditions sous lesquelles cela est possible est donnée en annexe A.3.

Le choix des initialisations de l'algorithme ainsi que celui des poids  $\lambda$  de la combinaison convexe successivement optimisés ont un fort impact sur la durée de l'optimisation comme sur la qualité des résultats obtenus. Nous proposons dans la suite deux procédures correspondant chacune à différents choix d'initialisations et de valeurs de  $\lambda$  et analysons sur un exemple à la fois les performances des filtres obtenus avec ces procédures et leur temps de calcul.

### 2.5.4 Mise en pratique de l'approche à compromis optimaux

Dans cette sous-section nous utilisons l'algorithme décrit précédemment sur un exemple afin d'illustrer son fonctionnement. Ici, l'objectif est de trouver des ensembles de  $W_{\min} = 127$  paires de filtres binaires, orthogonaux et complémentaires réalisant des compromis optimaux entre la classification parmi les trois espèces dont les spectres apparaissent dans la figure 2.1 et la reconstruction de spectres.

Dans un premier temps, nous proposons un choix d'initialisations et de valeurs de  $\lambda$  qui vise à explorer une région de l'espace des filtres aussi grande que possible, de manière à maximiser la confiance dans la courbe obtenue, mais qui entraîne de longs temps de calcul. La

procédure correspondant nous permet d’obtenir une référence à laquelle comparer une seconde procédure plus rapide. Dans cette seconde procédure, qui utilise d’autres initialisations et d’autres valeurs de  $\lambda$ , nous imposons une borne supérieure sur le nombre d’itérations afin de limiter le temps de calcul, et obtenons malgré cela une courbe assez proche de la première.

### Première procédure de construction de la courbe des compromis optimaux

Dans cette première procédure le choix des valeurs de  $\lambda$  est dynamique. Plus précisément, on utilise l’heuristique décrite dans la section précédente qui consiste à choisir la valeur de  $\lambda$  au centre de celles correspondant aux points voisins les plus éloignés au sens d’une certaine norme. Ici, cette norme est la norme euclidienne dans  $(\log(\text{tr } \Upsilon), \log(\overline{\mathcal{B}}))$ . Ce choix est empirique et vise à obtenir une courbe bien échantillonnée à plusieurs échelles.

Afin de maximiser la probabilité d’obtenir des ensembles de jeux de filtres proches de la courbe des compromis optimaux, il est nécessaire d’employer une diversité d’initialisations. L’optimisation réalisée à partir de chaque initialisation va permettre d’explorer une région de l’espace au moins partiellement disjointe des autres, agrandissant ainsi l’espace total exploré. Ici, seize initialisations sont employées pour chaque valeur de  $\lambda$  :

- ▷ Les filtres de la base de Hadamard.
- ▷ Un ensemble de jeux de filtres optimisé comme dans Justel, Galland et al. (2022) en employant  $\overline{\mathcal{B}}$  comme critère de sélection.
- ▷ Les deux ensembles de jeux de filtres optimisés par l’algorithme pour les deux valeurs de  $\lambda$  les plus proches de la valeur courante.
- ▷ Des ensembles de jeux de filtres tirés aléatoirement.

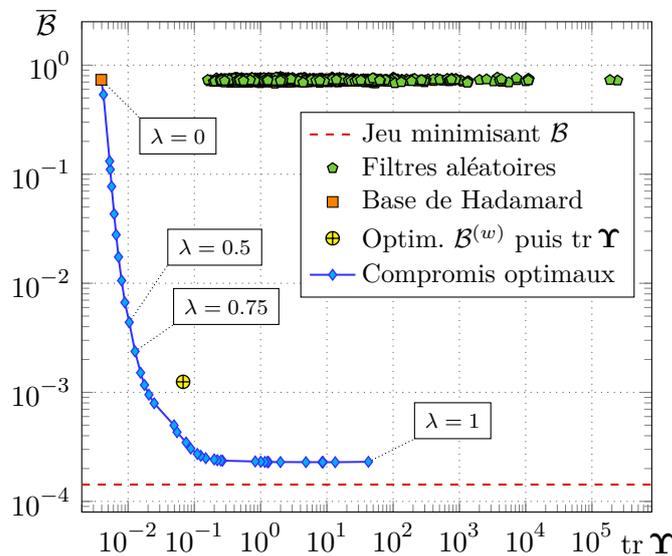


FIGURE 2.12 – Résultats de la recherche des filtres à compromis optimaux.

La figure 2.12 montre la courbe des compromis optimaux obtenue, à laquelle on compare les performances de plusieurs autres ensembles de jeux de filtres notables. Cette courbe comporte 41 points, chacun correspondant à une valeur de  $\lambda$ . En particulier, le point  $\lambda = 0$  est obtenu en maximisant uniquement les performances en reconstruction et le point  $\lambda = 1$  uniquement les performances en classification. L’algorithme proposé produit ici des ensembles de jeux de filtres réalisant une diversité de compromis entre classification parmi les espèces de référence et reconstruction des spectres rencontrés, et permet d’avoir un aperçu global des

performances possibles pour une configuration du système d'acquisition et des paramètres expérimentaux donnés. L'utilisateur peut alors choisir l'ensemble de jeux de filtres de la courbe qui correspond le mieux à ses besoins.

### Comparaison avec d'autres ensembles de jeux de filtres

On peut comparer les filtres construits par l'algorithme à d'autres ensembles de jeux de filtres. La ligne rouge tiretée correspond à la borne de Bhattacharyya associée à une paire de filtres complémentaires optimisée exclusivement pour la classification grâce à l'algorithme proposé par Réfrégier, Chevallier et al. (2019). Cette paire de filtres est celle permettant de réaliser la classification avec la probabilité d'erreur la plus faible. Il est néanmoins impossible de reconstruire les spectres rencontrés en n'utilisant que cette paire de filtres.

Chacun des  $10^3$  pentagones verts apparaissant sur la figure 2.12 correspond à un ensemble de paires de filtres tirées aléatoirement dont la matrice d'information de Fisher est inversible. On peut constater que ces ensembles de filtres ont une borne de Bhattacharyya supérieure à 0.68, ce qui implique un très mauvais contrôle sur la probabilité d'erreur. De plus, leurs performances en reconstruction sont très variables, allant de  $\text{tr } \mathbf{Y} = 0.16$  à  $\text{tr } \mathbf{Y} = 9.23 \times 10^5$ . Comme démontré par E. J. Candès, J. K. Romberg et al. (2006) et E. J. Candès et Tao (2006), la mesure de projections aléatoires du signal d'intérêt est particulièrement adaptée au cas où l'objectif de l'acquisition est de reconstruire des signaux à support à la fois inconnu et parcimonieux dans une certaine base. Il est alors nécessaire pour reconstruire un tel signal de disposer de plus de projections que de coefficients non nuls dans son support. Cependant, la situation analysée ici est différente, puisque nous n'avons fait aucune hypothèse sur la parcimonie des signaux observés et nous assurons d'effectuer suffisamment de mesures pour reconstruire tout signal normalisé de taille  $K$ . En l'occurrence, tous les ensembles de jeux de filtres correspondant à la partie gauche de la courbe des compromis optimaux sont meilleurs à la fois en classification et en reconstruction que les ensembles aléatoires.

Comme évoqué plus haut, les filtres dérivés de la base de Hadamard sont des paires de filtres binaires, orthogonaux, et complémentaires souvent utilisés pour la reconstruction de spectres. La trace de la borne de Cramér-Rao associée à ces filtres vaut  $\text{tr } \mathbf{Y} = 3.94 \times 10^{-3}$  et l'algorithme d'optimisation n'a pas trouvé d'ensemble de jeux de filtres plus performant en reconstruction. En revanche, ces jeux de filtres ont les pires performances en classification de la courbe des compromis optimaux, avec  $\bar{\mathcal{B}} = 0.73$  et ne sont donc adaptés qu'aux situations où la classification est secondaire.

L'ensemble de jeux de filtres construit en utilisant la méthode présentée dans la section 2.4.2 apparaît sur la figure sous la forme d'un disque jaune. Cet ensemble de jeux de filtres n'est pas situé sur la courbe des compromis optimaux. Il existe plusieurs ensembles de jeux de filtres à compromis optimaux ayant de meilleures performances à la fois en classification et en reconstruction. Ces ensembles lui sont donc toujours préférables.

### Forme de la courbe des compromis optimaux

On peut observer sur la figure 2.12 que la partie gauche de la courbe des compromis optimaux est presque verticale et que la partie basse est presque horizontale. Cela signifie qu'il est possible d'améliorer significativement les performances en l'une des tâches sans trop dégrader les performances en l'autre tâche si l'on s'écarte des extrémités de la courbe. La figure 2.13 présente la même courbe mais en faisant apparaître la différence entre ces extrémités et deux autres points.

Utiliser le point cerclé de rouge plutôt que les filtres de la base de Hadamard permet de diviser la borne de Bhattacharyya moyenne par 309.3, en ne multipliant la trace de la

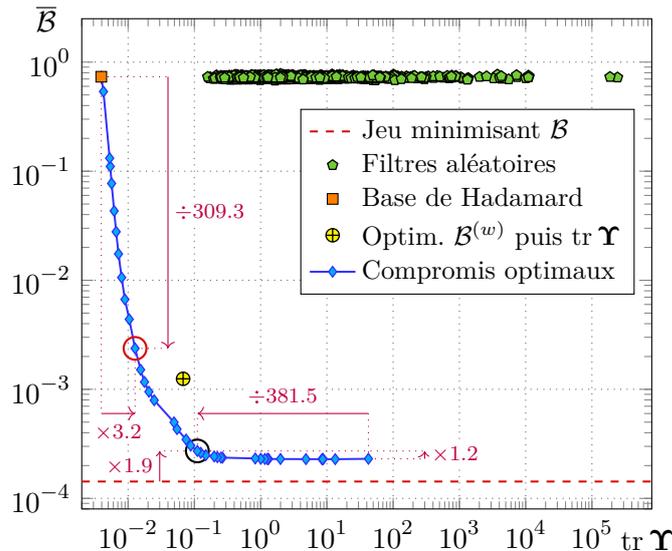


FIGURE 2.13 – Comparaison de différents points de la courbe des compromis optimaux.

matrice de Cramér-Rao que par 3.2. Autrement dit, il est possible de diviser par 309.3 la borne supérieure sur la probabilité d'erreur en ne multipliant l'écart-type de reconstruction que par un facteur  $\sqrt{3.2} \approx 1.8$ . De même, utiliser l'ensemble de jeux de filtres correspondant au point cerclé de noir plutôt que celui situé à l'extrême droite de la courbe permet de diviser l'écart-type d'estimation par  $\sqrt{381.5} \approx 19.5$  tout en ne multipliant  $\bar{\mathcal{B}}$  que par 1.2.

La borne de Bhattacharyya moyenne du point cerclé de noir n'est supérieure que d'un facteur 1.9 à la borne de Bhattacharyya minimale. Il est ainsi possible de réaliser à la fois classification parmi les espèces de référence et reconstruction des spectres rencontrés à faible coût en probabilité d'erreur de classification.

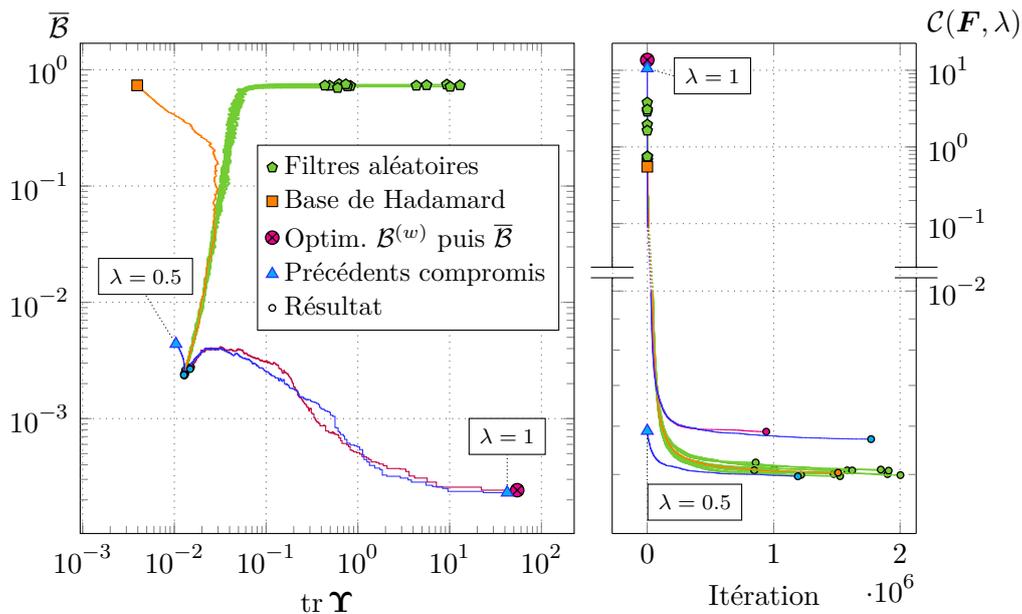
### Procédure rapide de calcul de la courbe des compromis optimaux

La courbe des compromis optimaux de la figure 2.12 a nécessité environ de 7 heures de calcul sur un serveur possédant 20 processeurs fonctionnant à une fréquence de 2.3GHz. Même si l'optimisation des jeux de filtres n'a lieu qu'une fois, avant l'acquisition, ce temps de calcul pourrait être un frein à l'adoption de la méthodologie proposée. Nous proposons ici une méthode visant à le réduire à quelques minutes sans trop dégrader les performances.

Afin de déterminer comment accélérer l'optimisation, il est nécessaire de regarder son déroulement en détails. La figure 2.14 montre les courbes dans l'espace  $(\text{tr } \mathbf{Y}, \bar{\mathcal{B}})$  (partie gauche) et les courbes en  $\mathcal{C}(\mathbf{F}, \lambda = 0.75)$  (partie droite) correspondant à l'optimisation des différentes initialisations lors de la recherche de filtres à compromis optimaux pour  $\lambda = 0.75$  effectuée lors du tracé de la courbe de la figure 2.12 .

Dans cet exemple, le nombre minimal d'itérations nécessaires à la convergence est supérieur à  $9 \times 10^5$ , d'où l'intérêt d'avoir une implémentation efficace. Une initialisation proche de la courbe des compromis optimaux ne garantit pas forcément que l'optimisation nécessite peu d'itérations. On peut en effet constater sur la partie droite de la figure 2.14 que la trajectoire correspondant au triangle bleu à  $\lambda = 0.5$  est de l'ordre de  $10^6$  itérations de longueur, ce qui est équivalent aux autres trajectoires. En revanche  $\mathcal{C}(\mathbf{F}, \lambda)$  ne décroît de manière significative que de 0 à  $\sim 10^5$  itérations, ce qui ne représente qu'environ 10% de la trajectoire.

Ces résultats suggèrent qu'il est possible de limiter fortement le nombre d'itérations sans perdre de trop en performances si les différentes initialisations sont proches des points de convergence. Pour ce faire, une possibilité est de sélectionner des valeurs de  $\lambda$  ordonnées et

FIGURE 2.14 – Recherche d'un minimum de  $\mathcal{C}(\mathbf{F}, \lambda = 0.75)$ .

de réaliser l'optimisation de proche en proche, c'est-à-dire en partant systématiquement de l'ensemble de jeux de filtres construit en optimisant  $\mathcal{C}(\mathbf{F}, \lambda)$  pour la précédente valeur de  $\lambda$ .

Réaliser l'optimisation en partant uniquement du dernier ensemble de jeux de filtres optimisé plutôt que d'initialisations multiples réduit l'espace exploré par l'algorithme. Le risque est de rester bloqué dans un minimum local. On peut voir sur la partie gauche de la figure 2.15 deux courbes des compromis optimaux tracées de manière itérative et en limitant à  $10^4$  le nombre d'itérations. La première part des jeux de filtres de la base de Hadamard et la seconde part des jeux de filtres optimisés et sélectionnés avec la borne de Bhattacharyya. Les  $\lambda$  optimisés sont répartis linéairement entre 0 et 1 avec un écart de  $2.5 \times 10^{-2}$  et le critère optimisé est  $\mathcal{C}_{\log}(\mathbf{F}, \lambda) = (1 - \lambda) \log(\text{tr } \Upsilon(\mathbf{F})) + \lambda \log(\overline{\mathcal{B}}(\mathbf{F}))$  plutôt que  $\mathcal{C}$  afin que les points optimisés soient répartis sur plusieurs échelles. Ces courbes ont été réalisées en moins de cinq minutes chacune.

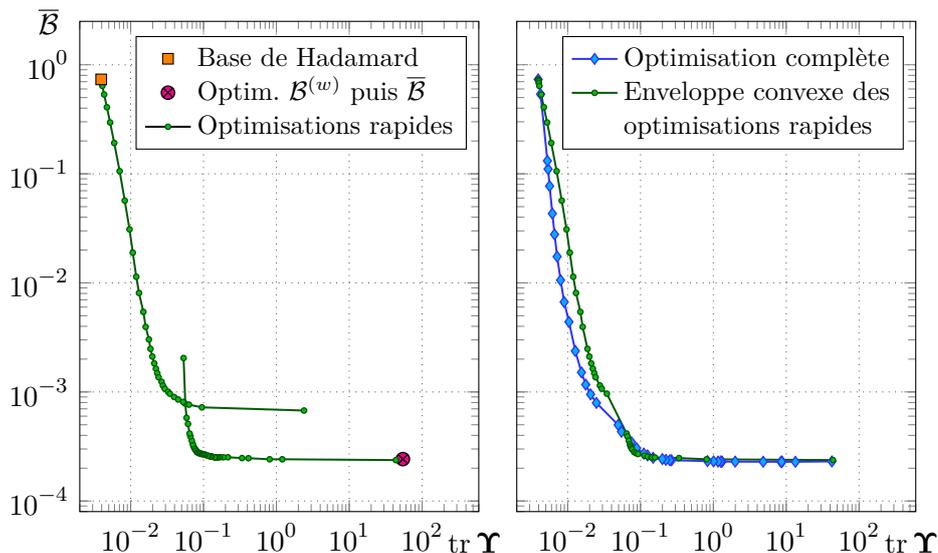


FIGURE 2.15 – Comparaison entre la procédure rapide et la procédure complète.

Chacune des deux courbes s'écarte assez largement de la courbe construite avec la première procédure, restant trop haut pour la courbe provenant des filtres de Hadamard ou trop à droite pour celle provenant de jeux optimisés et sélectionnés en  $\mathcal{B}$ . Les combiner permet de trouver une meilleure courbe. En particulier, leur enveloppe convexe dans l'espace des critères  $(\text{tr } \Upsilon, \bar{\mathcal{B}})$  est proche de la courbe construite en appliquant la procédure précédente, comme on peut le voir dans la partie droite de la figure 2.15. Le nombre d'itération maximal a été choisi arbitrairement, de manière à limiter le temps d'acquisition. Cette procédure rapide permet de calculer rapidement une première approximation de la courbe des compromis optimaux. En fonction des besoins de l'utilisateur il faudra ou non appliquer dans un second temps une autre procédure pour trouver de meilleurs filtres, mais ce n'est pas toujours nécessaire.

## 2.6 Étude des performances du système d'acquisition

Comme nous avons pu le voir dans la section précédente, tracer la courbe des compromis optimaux donne un aperçu des performances qu'il est possible d'obtenir pour des espèces de référence données avec une certaine configuration du système d'acquisition. Avec cette approche, l'utilisateur peut sélectionner parmi les filtres correspondant aux différents points de la courbe ceux permettant de satisfaire au mieux ses besoins. Ces ensembles de jeux de filtres sont donc autant de points de fonctionnement du système d'acquisition.

Cet outil permet également d'étudier l'évolution des performances en classification et en reconstruction en fonction des caractéristiques de l'expérience. On peut ainsi quantifier leur évolution en fonction de la nature des spectres des espèces chimiques censées composer l'échantillon (cf. sous-section 2.6.1) ainsi qu'avec leur nombre (cf. sous-section 2.6.2). Si la configuration du spectromètre évaluée ne permet pas d'avoir des performances suffisantes, l'approche proposée permet d'étudier l'effet des leviers dont on dispose pour ajuster les performances : nombre de pixels combinés pour reconstruire un spectre (cf. sous-section 2.6.3), nombre de photons détectés par pixel (cf. sous-section 2.6.4), nombre de filtres appliqués par pixel (cf. sous-section 2.6.5) et nombre de passages par pixel (cf. sous-section 2.6.6).

Nous présentons dans cette section une étude des performances du spectromètre Raman comprimé en fonction de ces différents paramètres. Cette étude, que nous effectuons avec l'approche à compromis optimaux, met en évidence l'influence de chacun de ces paramètres sur la qualité de l'analyse qu'il est possible d'effectuer à partir des mesures, démontrant l'intérêt de cette approche pour l'expérimentateur.

### 2.6.1 Influence de la forme des spectres des espèces de référence

Il est évident que la difficulté de la classification dépend de la forme des spectres de référence : plus ces spectres sont proches, plus il est difficile de les distinguer. La borne de Bhattacharyya permet de préciser la nature de cette dépendance. On rappelle ici son expression pour  $M$  spectres de référence  $\mathbf{s}_1, \dots, \mathbf{s}_M$ . Lorsque l'acquisition est réalisée à l'aide d'un jeu de filtres  $\mathbf{F}^{(w)}$  et arrêtée après la détection de  $N$  photons, cette borne s'écrit

$$\mathcal{B} = \frac{1}{M} \sum_u \sum_{v>u} \left[ \mathbf{q}_u^{(w)} \cdot \mathbf{q}_v^{(w)} \right]^N \quad (2.13)$$

où  $[\mathbf{q}_u^{(w)}]_m = (p_{m,u}^{(w)})^{\frac{1}{2}}$  est la racine carrée de la probabilité de détection d'un photon en sortie du filtre  $\mathbf{f}_m^{(w)}$  lorsque l'espèce de spectre  $\mathbf{s}_u$  est observée (voir (1.15) et (1.20)). Cette borne ne dépend des espèces de référence qu'à travers les termes  $\mathbf{q}_u^{(w)} \cdot \mathbf{q}_v^{(w)}$ . Il suffit donc pour

étudier l'influence de la forme des spectres de référence sur les performances de tracer les courbes des compromis optimaux pour des jeux de spectres corrélés à différents degrés.

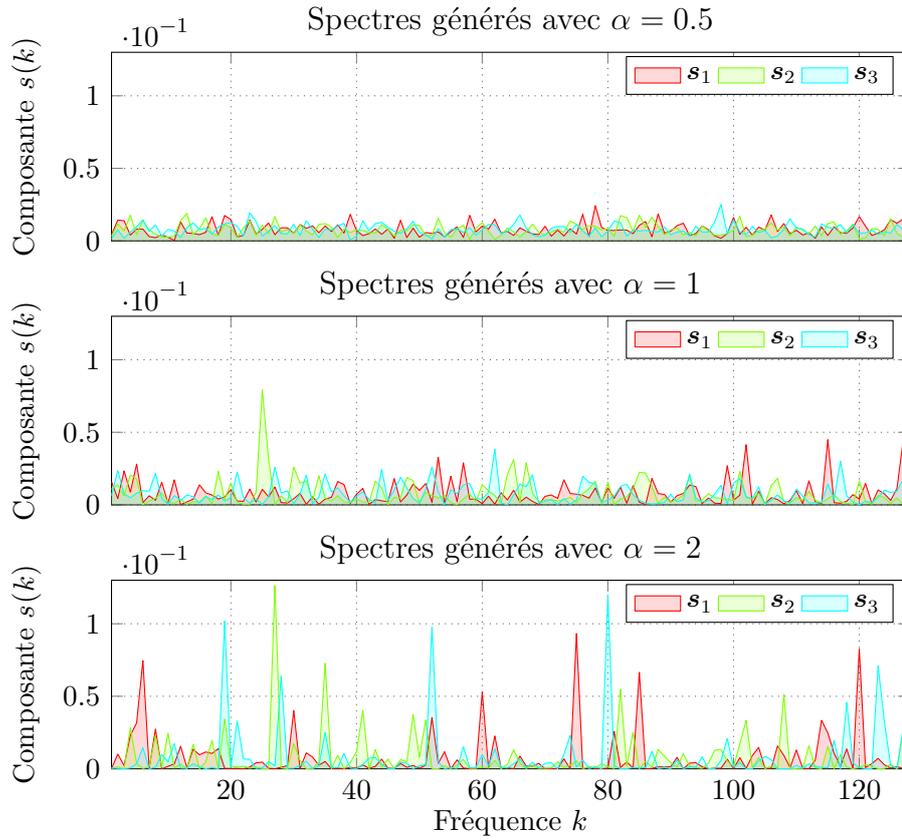


FIGURE 2.16 – Jeux de  $M = 3$  spectres générés avec différentes valeurs de  $\alpha$ .

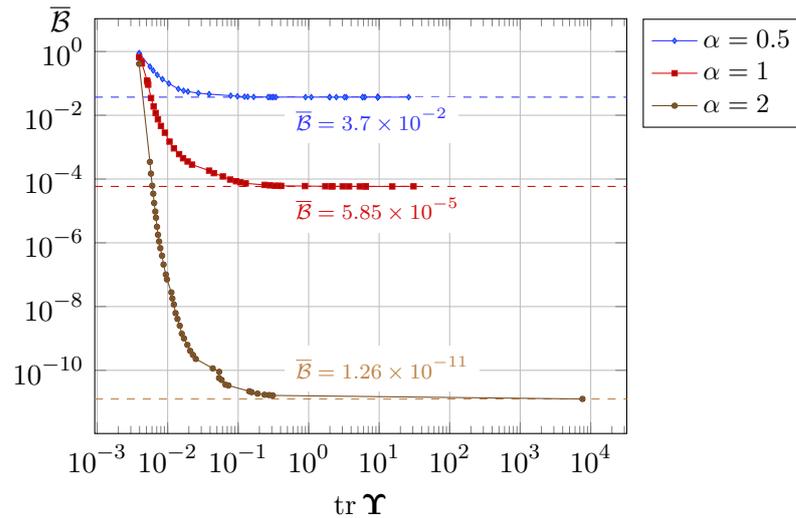
Comme évoqué par Réfrégier, Chevallier et al. (2019), il est possible d'obtenir des spectres plus ou moins corrélés en les générant comme des sommes normalisées de variables aléatoires indépendantes distribuées selon la loi exponentielle de moyenne 1 et élevées à une certaine puissance, c'est-à-dire comme

$$s(k) = \frac{(X_k)^\alpha}{\sum_{k=1}^K (X_k)^\alpha} \text{ où } X_k \sim \mathcal{E}(1) \quad (2.14)$$

où  $\alpha$  est un réel positif. Plus  $\alpha$  est grand et moins les spectres générés sont lisses. Comme démontré par Marsaglia (1961), le cas où  $\alpha = 1$  correspond à un tirage uniforme sur l'espace des spectres normalisés. La figure 2.16 montre des jeux de  $M = 3$  spectres générés pour  $\alpha \in \{0.5, 1, 2\}$ . Il paraît aisé de distinguer les spectres générés pour  $\alpha = 2$  les uns des autres, par exemple en se basant sur les pics caractéristiques de chaque espèce. À l'inverse, les spectres générés pour  $\alpha = 0.5$  sont assez plats et il est difficile de trouver des éléments discriminants.

La figure 2.17 présente les courbes des compromis optimaux correspondant à des ensembles de  $W = 127$  paires de filtres complémentaires optimisés à la fois pour la classification des mesures parmi les espèces ayant les spectres de la figure 2.16 et pour la reconstruction. Le nombre de photons détectés est fixé à  $N = 250$  par pixel.

Comme attendu, la probabilité d'erreur de classification parmi les espèces correspondant au jeu de spectres généré avec  $\alpha = 2$  est faible pour la plupart des points de fonctionnement. Plus précisément, la borne de Bhattacharyya moyenne est comprise entre  $1.26 \times 10^{-11}$  et  $3.44 \times 10^{-4}$  lorsque  $\lambda \neq 0$ . Les performances en classification diminuent avec  $\alpha$ . S'il est possible d'atteindre  $\bar{\mathcal{B}} = 5.85 \times 10^{-5}$  lorsque  $\alpha = 1$ , la borne de Bhattacharyya moyenne est


 FIGURE 2.17 – Courbes des compromis optimaux fonction de  $\alpha$ .

toujours supérieure à  $3.7 \times 10^{-2}$  pour  $\alpha = 0.5$ . Cette probabilité d'erreur de classification est élevée, et il paraît nécessaire de changer la configuration du spectromètre lorsque les spectres de référence sont de ce type. Il est par exemple possible d'accroître le nombre de photons détectés par pixel ou le nombre de passages par pixel (voir sous-sections 2.6.4 et 2.6.6).

## 2.6.2 Influence du nombre d'espèces de référence

Un autre paramètre clé pour la classification est le nombre d'espèces chimiques de référence. Conformément à l'intuition, plus le nombre d'espèces de référence est élevé et plus la classification est difficile. L'augmenter est en effet équivalent à augmenter le nombre de classes, ce qui accroît nécessairement la difficulté du problème de classification. Il faut donc éviter d'introduire plus d'espèces de références que nécessaire et se limiter à celles que l'on sait pouvoir apparaître dans l'échantillon.

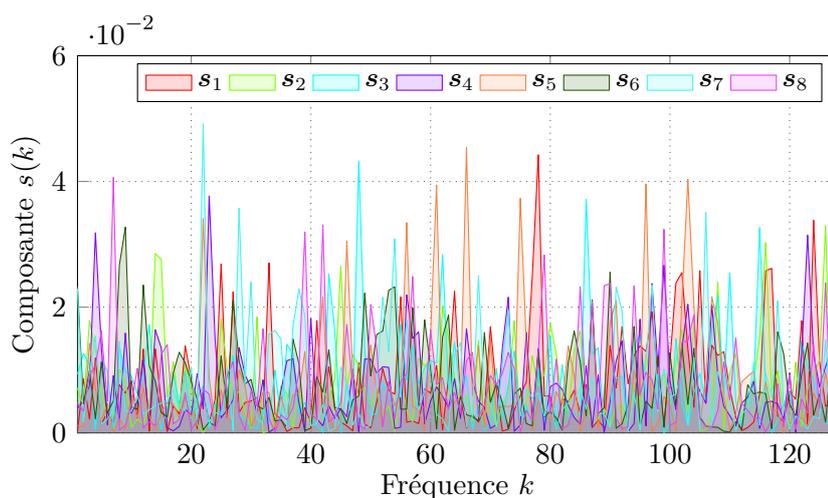


FIGURE 2.18 – Spectres de référence pour les courbes de la figure 2.19.

La figure 2.19 présente les courbes des compromis optimaux correspondant à des ensembles de  $W = 127$  paires de filtres complémentaires de dimension  $K = 128$ . Le nombre de photons détectés par pixel est fixé à  $N = 250$ . Pour chaque courbe, les spectres de référence sont

les  $M \in \{2, 4, 8\}$  premiers des huit spectres de la figure 2.18. Ces spectres ont été tirés uniformément et indépendamment sur l'ensemble des spectres normalisés.

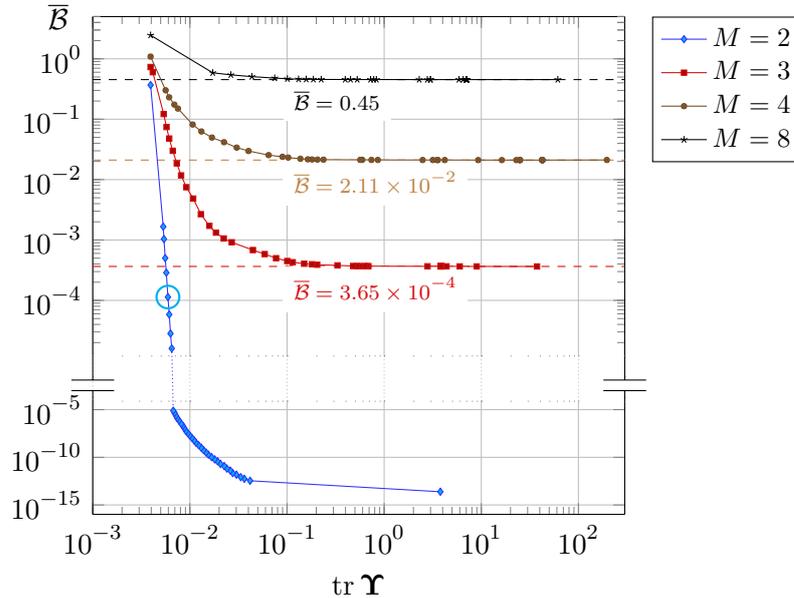


FIGURE 2.19 – Courbes des compromis optimaux en fonction de  $M$ .

Comme attendu, l'augmentation de  $M$  diminue significativement les performances en classification. Dans cet exemple, classer les mesures parmi les deux premières espèces de référence est facile et il est aisé de trouver un point de fonctionnement permettant d'avoir une faible probabilité d'erreur de classification. On peut par exemple choisir d'utiliser l'ensemble de jeux de filtres correspondant au point cerclé de bleu sur la figure, pour lequel  $\bar{\mathcal{B}} = 1.13 \times 10^{-4}$  et  $\text{tr } \Upsilon = 5.94 \times 10^{-3}$ . Dès que le nombre d'espèces atteint quatre, les performances en classification sont mauvaises pour tous les points de fonctionnement. En effet,  $\bar{\mathcal{B}}$  est supérieure à  $2.11 \times 10^{-2}$  lorsque  $M = 4$  et à  $0.45$  lorsque  $M = 8$ . Il semble donc nécessaire dans cet exemple de modifier les paramètres de l'acquisition pour  $M \geq 3$ .

### 2.6.3 Influence du nombre de jeux de filtres

Si les performances en reconstruction sont insuffisantes, une solution est d'augmenter le nombre de pixels combinés pour la réaliser. Il est possible de mettre en œuvre cette solution a posteriori, sans modifier les filtres appliqués. L'équation (2.4) montre en effet que la trace de la matrice de Cramér-Rao est divisée par  $Q$  si l'on combine  $W \times Q$  pixels plutôt que  $W$  et que chacun des  $W$  jeux de filtres de l'ensemble construit est assigné à  $Q$  de ces pixels. On peut donc considérer en première approximation que multiplier par  $Q$  le nombre de pixels utilisés pour reconstruire un spectre divise par  $Q$  la variance de la reconstruction. Cette stratégie revient à dupliquer certains jeux de filtres, ce qui est en général sous-optimal. Si on en constate la nécessité avant les mesures, il est plus avantageux de relancer l'optimisation des filtres en augmentant  $W$ , du moins jusqu'à un certain point.

Pour étudier cet effet, nous avons optimisé des ensembles de jeux de filtres de mêmes caractéristiques pour différents  $W$ . Le nombre de photons détectés par pixel est fixé à  $N = 250$  et le nombre d'espèces de référence est fixé à  $M = 4$ . Les spectres de ces espèces possèdent  $K = 128$  composantes et sont tirés uniformément sur le simplexe.

La partie gauche de la figure 2.20 montre les courbes des compromis optimaux correspondant à des ensembles de  $W \in \{W_{\min} = 127, 254, 381, 508, 635\}$  paires de filtres complémen-

taires. Sur la partie droite de la figure, on a tracé ces courbes en fonction de  $W/W_{\min} \operatorname{tr} \Upsilon$  afin de ne pas tenir compte de l'amélioration des performances en reconstruction due à l'augmentation du nombre de pixels combinés pour la réaliser, mais uniquement de celle obtenue grâce à l'augmentation de la diversité de jeux de filtres.

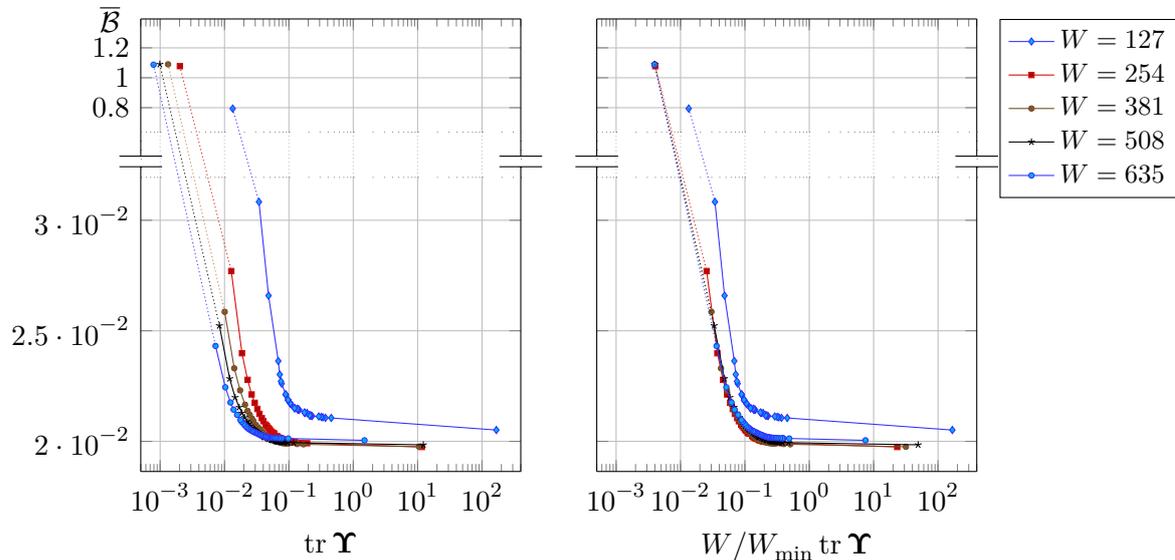


FIGURE 2.20 – Courbes des compromis optimaux fonction de  $W$ .

Comme prévu, les courbes montrent qu'augmenter  $W$  permet d'améliorer les performances en reconstruction sans impact significatif sur les performances de classification. Cette amélioration atteint néanmoins rapidement un plafond. En effet, seule la courbe à  $W = 127$  n'est pas confondue avec les autres courbes de la partie droite de la figure 2.20. Dans cet exemple,  $\operatorname{tr} \Upsilon$  atteint un régime asymptotique en  $1/W$  dès  $W = 254$ , variation qui est simplement due à l'augmentation du nombre de mesures et non à leur diversification.

#### 2.6.4 Influence du nombre de photons à détecter

Fixer le nombre de photons déclenchant l'arrêt de l'acquisition est équivalent à fixer la quantité d'énergie intégrée sur les détecteurs en chaque pixel. Augmenter cette quantité permet d'améliorer les performances, que ce soit en classification ou en reconstruction. Cependant, cela implique également d'allonger la durée de l'acquisition. En effet, le temps nécessaire à l'acquisition de  $N$  photons en moyenne est proportionnel à  $N$  (voir équation (2.6)). Il y a donc un compromis à trouver entre performances et durée de l'acquisition.

On peut préciser l'évolution des critères de performances choisis avec  $N$ . L'équation (2.4) montre que la borne de Cramér-Rao est inversement proportionnelle à  $N$ . Le résultat de l'optimisation d'un ensemble de jeux de filtres pour la reconstruction ne dépend donc pas de la valeur de  $N$  choisie. En revanche, il n'existe pas de relation linéaire entre  $N$  et la borne de Bhattacharyya (voir équation (2.13)). Ainsi, il n'est pas équivalent d'optimiser  $\mathcal{C}(\mathbf{F}, \lambda)$  pour deux valeurs de  $N$  différentes lorsque  $\lambda$  est strictement positif. Il est donc nécessaire d'estimer les courbes des compromis optimaux pour chaque valeur de  $N$  à étudier.

La figure 2.21 montre les courbes à compromis optimaux d'ensembles de  $W = 127$  paires de filtres complémentaires de taille  $K = 128$  optimisés à la fois pour la classification des mesures parmi  $M = 6$  espèces de référence et pour la reconstruction des spectres rencontrés. Les spectres des espèces de référence ont été tirés uniformément sur l'espace des spectres normalisés. Du fait de la valeur de  $M$  il s'agit d'un problème de classification difficile. Le

nombre de photons à détecter est compris entre 250 et 2500, ce qui permet d'analyser l'effet d'une multiplication par 10 du temps d'acquisition.

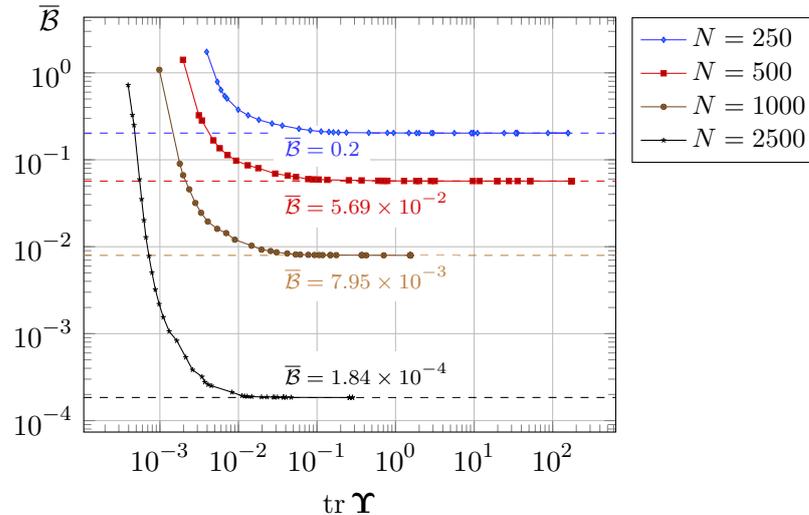


FIGURE 2.21 – Courbes des compromis optimaux en fonction de  $N$ .

Passer de  $N = 250$  à  $N = 2500$  photons détectés par pixel permet de réduire la valeur minimale de  $\bar{\mathcal{B}}$  atteignable de 0.2 à  $1.84 \times 10^{-4}$  soit une division par  $1.1 \times 10^3$ . Cela permet également de diviser la trace de la borne de Cramér-Rao minimale par 10, passant de  $\text{tr } \Upsilon = 3.94 \times 10^{-3}$  à  $\text{tr } \Upsilon = 3.94 \times 10^{-4}$ .

Dépendamment de l'intensité du rayonnement et des besoins de l'utilisateur, le temps de pose nécessaire à l'acquisition de plusieurs milliers de photons peut être rédhibitoire. Une possibilité est alors d'augmenter le nombre de filtres appliqués par pixel.

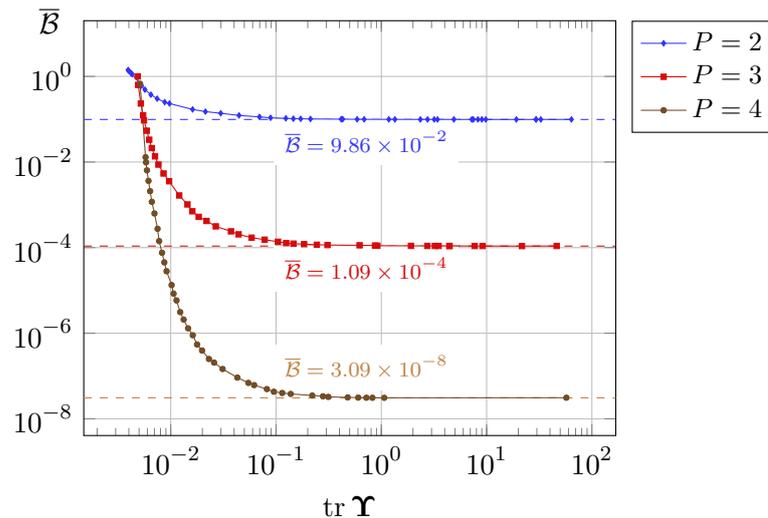
### 2.6.5 Influence du nombre de filtres

Lorsque les filtres sont appliqués en parallèle comme c'est le cas dans ce premier chapitre, il y a autant de détecteurs que de filtres et il paraît évident qu'augmenter le nombre de filtres améliore les performances en classifications. Cela permet également de mettre en œuvre la reconstruction sur de plus petites zones, car le nombre minimum  $W_{\min}$  de pixels à combiner pour reconstruire un spectre est inversement proportionnel à  $P - 1$  (voir proposition 2).

Pour tester cette affirmation nous comparons les performances d'ensembles de jeux de filtres construits avec différentes valeurs de  $P$  grâce aux courbes des compromis optimaux correspondant. Ces courbes sont tracées pour les mêmes espèces de référence, un même nombre de photons détectés par pixel et pour  $W_{\min}$  pixels. Plus précisément, la figure 2.22 montre les courbes des compromis optimaux d'ensembles de  $W \in \{127, 64, 43\}$  jeux de  $P \in \{2, 3, 4\}$  filtres complémentaires pour  $N = 250$  photons détectés par pixel et  $M = 5$  espèces de référence. Les spectres de ces espèces de référence sont tirés uniformément.

Augmenter le nombre de filtres appliqués en chaque pixel améliore considérablement les performances en classification. Là où la valeur minimale de  $\bar{\mathcal{B}}$  est de  $9.86 \times 10^{-2}$  lorsque  $P = 2$ , il est possible d'atteindre  $\bar{\mathcal{B}} = 3.09 \times 10^{-8}$  en utilisant des jeux de  $P = 4$  filtres. Les performances en reconstruction obtenues pour les différentes valeurs de  $P$  sont équivalentes, ce qui n'est pas étonnant puisque le nombre total de mesures indépendantes effectuées ne varie presque pas.

Il faut toutefois prendre en considération le fait qu'augmenter le nombre de filtres appliqués en parallèle accroît considérablement la complexité du système de mesure. Ainsi,

FIGURE 2.22 – Courbes des compromis optimaux fonction de  $P$ .

l'implémentation d'un spectromètre Raman comprimé possédant plus de deux filtres appliqués en parallèle est difficile. Une alternative pour améliorer les performances en classification sans complexifier le dispositif d'acquisition est d'appliquer séquentiellement plusieurs paires de filtres par pixel.

### 2.6.6 Influence du nombre de passages par pixel

Comme nous l'avons montré dans les sous-sections précédentes, réaliser l'acquisition en n'appliquant qu'une paire de filtres par pixel ne permet parfois pas d'obtenir de bonnes performances en classification, notamment dans le cas où les spectres de référence se ressemblent (cf. sous-section 2.6.1) ou lorsque le nombre d'espèces de référence est élevé (cf. sous-section 2.6.2). S'il est possible de relever le nombre de photons détectés déclenchant l'arrêt des mesures pour améliorer les performances (cf. sous-section 2.6.4) cela signifie également allonger l'acquisition, ce qui n'est parfois pas envisageable. Augmenter le nombre de filtres appliqués en parallèle permet de résoudre ce problème (cf. sous-section 2.6.5) mais cela est complexe à mettre en œuvre expérimentalement.

Une autre possibilité permettant d'améliorer la qualité de la classification tout en continuant à appliquer seulement deux filtres en parallèle est d'appliquer successivement plusieurs paires de filtres sur chaque pixel, ce qui revient à effectuer plusieurs passages par pixel. Dans cette sous-section, nous étudions l'impact de cette stratégie sur les performances en classification et en reconstruction. Nous montrons qu'elle permet effectivement de diminuer la probabilité d'erreur de classification tout en conservant le même temps d'acquisition si le nombre de passages par pixel n'est pas trop élevé.

Il faut préciser ici que l'optimisation des jeux de filtres est toujours effectuée avant les mesures. En d'autres termes, le système que nous proposons n'est pas dynamique, au sens où nous n'utilisons pas la connaissance acquise lors des premiers passages pour optimiser les filtres appliqués lors des passages suivants. Bien qu'il soit possible qu'un système dynamique permette d'atteindre de meilleures performances, le temps de calcul nécessaire à réaliser l'optimisation au cours de la prise de vue réduirait à néant le gain de temps réalisé grâce à l'emploi d'un système comprimé.

### Impact sur la classification

La totalité des mesures réalisées sur un même pixel peuvent être utilisées pour le classer. La qualité de la classification dépend alors de tous les jeux de filtres appliqués sur ce pixel. Il est donc nécessaire que les différents jeux de filtres appliqués sur un même pixel soient optimisés conjointement. Nous utilisons comme précédemment la borne de Bhattacharyya comme critère de performance en classification.

Soit  $L$  jeux de  $P$  filtres  $\mathbf{F}^{(w,1)}, \dots, \mathbf{F}^{(w,L)}$  appliqués sur un pixel  $w$  où apparaît l'une des  $M$  espèces de spectres  $\mathbf{s}_1, \dots, \mathbf{s}_M$  censées composer l'échantillon analysé. On démontre en annexe C que la probabilité d'erreur de classification des mesures réalisées en  $w$  avec le classifieur optimal de l'équation (1.18) vérifie

$$\mathcal{P}(\epsilon) \leq \mathcal{B}^{(w)} = \frac{1}{M} \sum_{u=1}^M \sum_{v>u} \prod_{\ell=1}^L 2 \mathcal{B}_{u,v}^{(w,\ell)}$$

où  $\mathcal{B}_{u,v}^{(w,\ell)}$  est la borne de Bhattacharyya associée à la classification de mesures réalisées à travers le jeu de filtres  $\mathbf{F}^{(w,\ell)}$  parmi les espèces  $u$  et  $v$  (voir équation (1.19)).

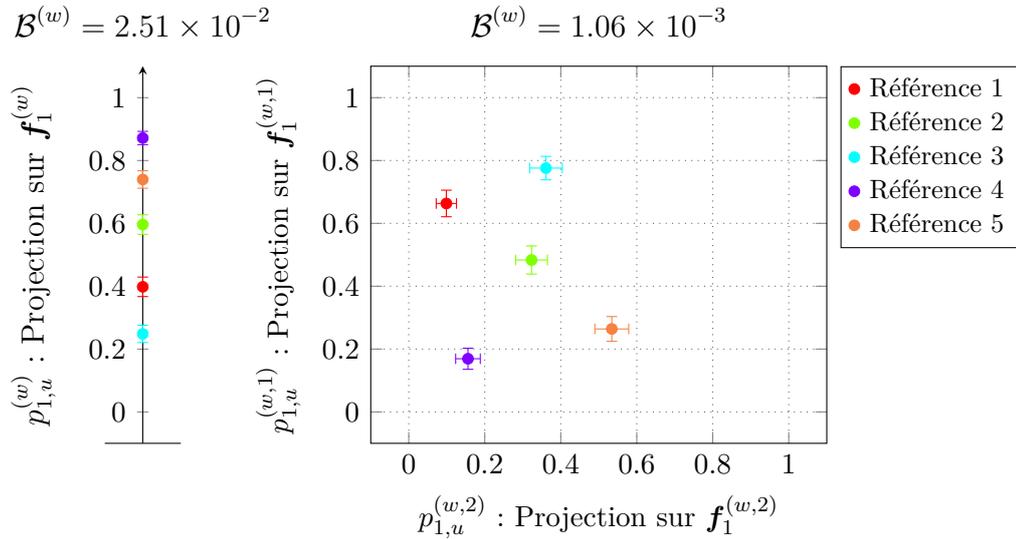


FIGURE 2.23 – Projections de  $M = 5$  spectres de référence sur différents filtres.

Pour illustrer l'intérêt d'appliquer plusieurs paires de filtres par pixel, on compare un cas où une unique paire de filtres  $\mathbf{F}^{(w)}$  est appliquée en un pixel  $w$  au cas où  $L = 2$  paires de filtres  $\mathbf{F}^{(w,1)}$  et  $\mathbf{F}^{(w,2)}$  y sont appliquées. Ces filtres ont été optimisés pour la classification des mesures parmi  $M = 5$  espèces de référence dont les spectres ont été tirés uniformément et indépendamment dans l'espace des spectres normalisés. Le nombre de photons détectés par pixel est fixé à 250, soit à 125 par paire de filtres lorsque  $L = 2$ .

La partie gauche de la figure 2.23 montre les projections des spectres de référence sur le filtre  $\mathbf{f}_1^{(w)}$  et la partie droite leurs projections sur les filtres  $\mathbf{f}_1^{(w,1)}$  et  $\mathbf{f}_1^{(w,2)}$ . On rappelle que  $\sum_{m=1}^P p_{m,u}^{(w)} = 1$ , ce qui permet de n'afficher que les projections des spectres sur le premier filtre de chaque paire sans perte d'information. On a également représenté sous la forme de barres d'erreur l'écart-type théorique des mesures normalisées correspondant à chaque filtre, c'est-à-dire des variables aléatoires  $n_1^{(w)}/N$  avec  $N = 250$  à gauche ainsi que  $n_1^{(w,1)}/N$  et  $n_1^{(w,2)}/N$  avec  $N = 125$  à droite, où  $n_m^{(w,\ell)}$  est la mesure correspondant au filtre  $\mathbf{f}_m^{(w,\ell)}$ .

L'écart entre les projections des différents spectres est globalement plus important lorsque

l'on emploie deux paires de filtres plutôt qu'une, même en prenant en compte les barres d'erreurs. On s'attend donc à ce que la probabilité d'erreur de classification soit moins élevée dans ce cas. Cette analyse empirique est confirmée par le calcul de la borne de Bhattacharyya associée à chacune des deux configurations. En effet, on a  $\mathcal{B}^{(w)} = 2.51 \times 10^{-2}$  quand  $L = 1$  et  $\mathcal{B}^{(w)} = 1.06 \times 10^{-3}$  pour  $L = 2$ .

### Impact sur la reconstruction

Les mesures effectuées avec différents jeux de filtres sont indépendantes. De ce fait, il est équivalent pour la reconstruction de combiner des mesures provenant d'un même pixel ou de plusieurs pixels de même composition, et on a simplement

$$[\mathbf{I}]_{k,l} = \sum_{w=1}^W [\mathbf{I}^{(w)}]_{k,l} = \sum_{w=1}^W \sum_{\ell=1}^L \frac{N}{\sum_{n=1}^P \mathbf{f}_n^{(w,\ell)} \cdot \mathbf{s}} \left[ \sum_{m=1}^P \frac{\delta f_m^{(w,\ell)}(k) \delta f_m^{(w,\ell)}(l)}{\mathbf{f}_m^{(w,\ell)} \cdot \mathbf{s}} - \frac{\sum_{n=1}^P \delta f_n^{(w,\ell)}(k) \sum_{n=1}^P \delta f_n^{(w,\ell)}(l)}{\sum_{n=1}^P \mathbf{f}_n^{(w,\ell)} \cdot \mathbf{s}} \right]$$

où la seule différence par rapport à l'équation (2.4) est l'ajout des termes correspondant aux différents jeux de filtres appliqués en chaque  $w$ .

Réaliser plusieurs passages par pixel permet de diminuer le nombre de pixels dont il est nécessaire de combiner les mesures pour pouvoir effectuer la reconstruction. Plus précisément, il faut combiner les mesures provenant de  $W_{\min}/L$  pixels au minimum lorsque l'on effectue  $L$  passages par pixel avec des jeux de filtres différents afin de pouvoir reconstruire le spectre correspondant.

### Impact sur le temps d'acquisition

Augmenter le nombre de jeux de filtres appliqués en chaque pixel n'est pas sans impact sur le temps d'acquisition, et il est nécessaire d'ajuster le nombre de photons détectés par jeu de filtres afin de pas allonger de trop l'acquisition.

On considère ici que les filtres utilisés sont complémentaires pour ne pas alourdir inutilement le formalisme. Avec cette hypothèse, le temps de mesure nécessaire à la détection de  $N$  photons en moyenne est  $\tau_N = N/\gamma$  où  $\gamma$  est l'intensité du rayonnement (voir équation (2.6)).

Une solution permettant de conserver le même temps de mesure par pixel peu importe  $L$  est de diviser  $N$  par  $L$  comme nous l'avons fait précédemment. Cependant, multiplier le nombre de passages par pixel accroît également le nombre de changements de filtres et de déplacements d'un pixel à l'autre, et donc le temps mort. On a ainsi en chaque pixel

$$\tau_{\text{total}} = L(\tau_N + \tau_{\text{mort}}) = L \left( \frac{N}{\gamma} + \tau_{\text{mort}} \right)$$

où  $\tau_{\text{total}}$  est le temps total d'acquisition et  $\tau_{\text{mort}}$  le temps mort par jeu de filtres. Pour garder un temps d'acquisition total fixé il est donc nécessaire que  $N$  vérifie

$$N(\tau_{\text{total}}, L) = \gamma \left( \frac{\tau_{\text{total}}}{L} - \tau_{\text{mort}} \right),$$

ce qui peut réduire significativement le nombre de photons dont il est possible d'attendre la détection. La figure 2.24 montre par exemple le nombre de photons  $N(\tau_{\text{total}}, L)$  en fonction de  $L$ , pour  $\tau_{\text{total}} \in \{50, 100, 250\} \mu\text{s}$ ,  $\tau_{\text{mort}} = 10 \mu\text{s}$  et  $\gamma = 1 \text{ photon}/\mu\text{s}$ .

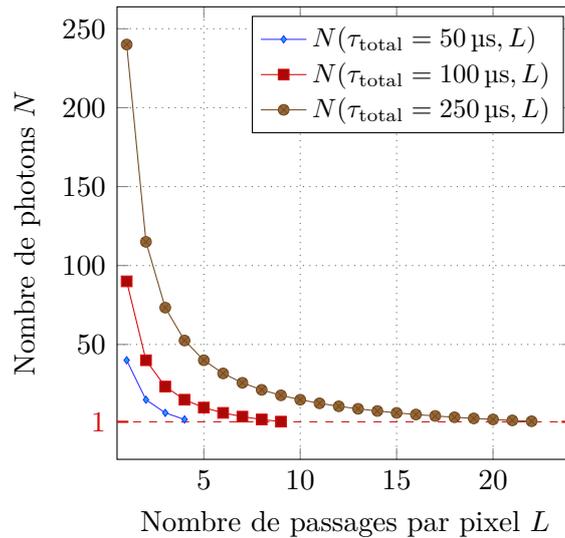


FIGURE 2.24 – Nombre  $N$  de photons à détecter à chaque passage en fonction du nombre  $L$  de passages par pixel pour conserver un temps d’acquisition total par pixel fixé à  $\tau_{\text{total}}$ .

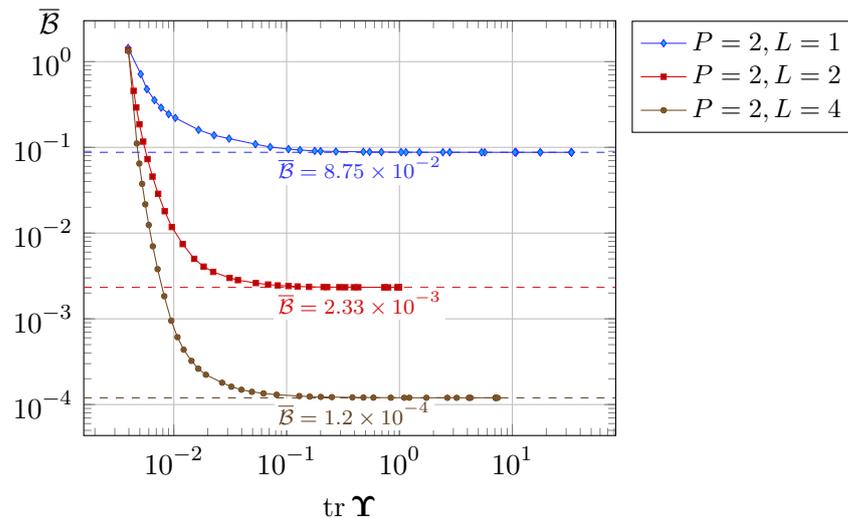
On constate qu’il est difficile ici d’envisager de réaliser plus d’une dizaine de passages par pixel si l’on veut conserver un temps d’acquisition total équivalent à 100 $\mu\text{s}$ . Ces résultats sont propres aux valeurs choisies, et il est nécessaire en pratique d’adapter cette étude au dispositif de mesure employé. L’analyse réalisée dans ce manuscrit se voulant générale, nous ne fixons pas  $\tau_{\text{mort}}$  où  $\gamma$  dans la suite. Nous nous sommes cependant restreints à quelques passages par pixel pour tenir compte de cette limitation.

### Courbe des compromis optimaux pour plusieurs passages par pixel

Afin de quantifier l’influence du nombre de passages par pixel sur les performances, nous avons tracé sur la figure 2.25 les courbes des compromis optimaux correspondant à des ensembles de  $W = 127 \times L$  jeux de filtres complémentaires de dimension  $K = 128$  pour  $L \in \{1, 2, 4\}$  et  $P = 2$ . Cela correspond à réaliser l’acquisition sur 127 pixels en faisant un, deux ou quatre passages par pixel. Les jeux de filtres sont divisés en blocs de taille  $L$  dont les membres sont assignés aux mêmes pixels et donc optimisés conjointement pour la classification. Les spectres des  $M = 5$  espèces de référence sont ceux utilisés pour l’expérience de la figure 2.23. Le nombre de photons détectés par jeu de filtres est fixé  $N = 250/L$  de sorte à ce que le nombre de photons détectés par pixel soit le même quel que soit  $L$ .

Dans cet exemple, augmenter le nombre de passages par pixel permet de diminuer significativement la probabilité d’erreur de classification. Plus précisément, la borne de Bhattacharyya minimale lorsque  $P = 2$  et  $L = 1$  est de  $\bar{\mathcal{B}} = 8.75 \times 10^{-2}$  lorsqu’elle est de  $\bar{\mathcal{B}} = 1.2 \times 10^{-4}$  pour  $P = 2$  et  $L = 4$ .

Ainsi, effectuer plusieurs passages par pixel permet de mieux traiter le problème de classification, ce qui donne la possibilité de réaliser des acquisitions courtes même lorsque le nombre d’espèces de référence est élevé ou lorsque les spectres de référence se ressemblent.

FIGURE 2.25 – Courbes des compromis optimaux en fonction de  $L$ .

## 2.7 Détection et caractérisation d'une anomalie

### 2.7.1 Méthodologie proposée et détection d'anomalie

Nous avons présenté dans ce chapitre une nouvelle approche permettant de réaliser à la fois la classification de chaque pixel parmi les espèces de référence et la reconstruction des spectres rencontrés en combinant des mesures comprimées réalisées sur plusieurs pixels.

Pouvoir reconstruire les spectres rencontrés offre la possibilité de tester les a priori sur les espèces en présence utilisés lors de la classification. Il est en effet possible de comparer les spectres reconstruits aux références, à l'instar de ce qui est possible sans filtrage. La reconstruction permet ainsi de tester la conformité des mesures au modèle. En cas de non-conformité il est possible d'identifier l'origine de l'anomalie détectée à partir de son spectre, ce qui permet à l'utilisateur de réagir de manière appropriée. La méthode proposée permet d'atteindre cet objectif sans mesure supplémentaire.

### 2.7.2 Retour sur le premier exemple

Il est intéressant de tester l'approche proposée sur l'exemple de la section 2.2 pour la comparer aux méthodes existantes. C'est ce que nous faisons dans cette sous-section.

On considère le même échantillon que dans la section 2.2. On rappelle que cet échantillon est censé être composé des  $M = 3$  espèces chimiques dont les spectres, qui possèdent  $K = 128$  composantes, apparaissent en figure 2.1. Les filtres sont donc optimisés en ne prenant en compte que ces spectres. Nous avons cependant ajouté une quatrième espèce dans l'échantillon. Le spectre de cette espèce apparaît en figure (2.3). On suppose ignorer a priori à la fois la présence et la nature de cette espèce anormale. Ce type de situation apparaît par exemple si l'échantillon analysé est pollué par un composé inconnu. Il est alors important de détecter et d'identifier ce composé.

La figure 2.26 montre la courbe des compromis optimaux pour des ensembles de  $W = 144$  paires de filtres binaires, orthogonaux et complémentaires. Cette valeur de  $W$  correspond au plus petit carré supérieur à  $W_{\min} = 127$ , ce qui sera utile dans la suite pour répartir les jeux de filtres sur l'image. Puisque l'on souhaite pouvoir réaliser à la fois la classification et la reconstruction avec de bonnes performances, nous choisissons d'utiliser l'ensemble de jeux de filtres cerclé de rouge sur la figure. Cet ensemble de jeux de filtres  $\mathbf{F} = (\mathbf{F}^{(1)}, \dots, \mathbf{F}^{(144)})$

correspond à l'optimisation de  $\mathcal{C}(\mathbf{F}, \lambda = 0.9375)$ . Sa borne de Bhattacharyya moyenne est  $\bar{\mathcal{B}} = 9.77 \times 10^{-4}$  et la trace de la borne de Cramér-Rao correspondante est  $\text{tr } \boldsymbol{\Upsilon} = 1.49 \times 10^{-2}$ .

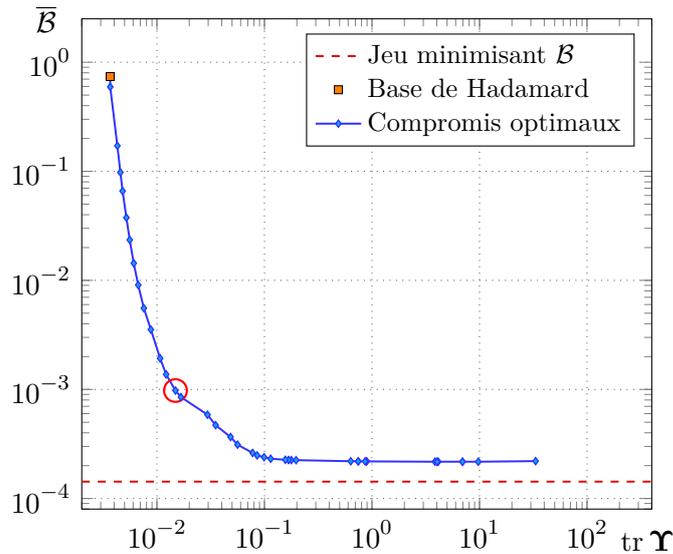


FIGURE 2.26 – Courbe des compromis optimaux pour l'exemple d'application.

On simule des mesures réalisées avec les jeux de filtres sélectionnés sur le champ de vue de la section 2.2. La composition de ce champ de vue est affichée en haut à gauche de la figure 2.28. Les 144 jeux de filtres  $y$  sont répartis de sorte à pouvoir reconstruire un spectre en combinant les pixels compris dans toute fenêtre carrée de 12 pixels de côté. Plus précisément, au pixel de coordonnées  $(i, j)$  est assigné le jeu de filtres  $\mathbf{F}^{(w(i,j))}$  avec  $w(i, j) = \text{mod}(i-1, \sqrt{W}) \times \sqrt{W} + \text{mod}(j-1, \sqrt{W}) + 1$  où  $\text{mod}(a, b)$  est le reste de la division euclidienne de  $a$  par  $b$ . Un exemple de ce type de répartition est affiché en figure 2.27.

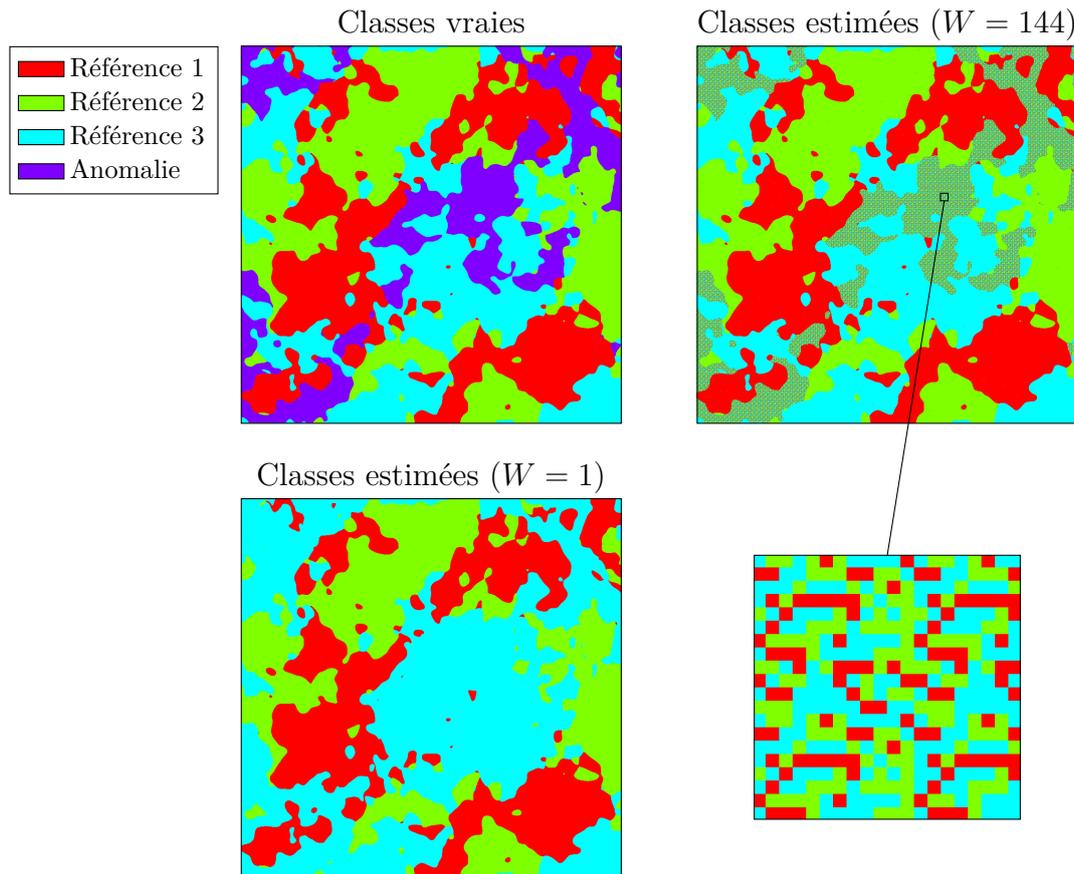
|                    |                    |                    |                    |                    |
|--------------------|--------------------|--------------------|--------------------|--------------------|
| $\mathbf{F}^{(1)}$ | $\mathbf{F}^{(2)}$ | $\mathbf{F}^{(3)}$ | $\mathbf{F}^{(1)}$ | $\mathbf{F}^{(2)}$ |
| $\mathbf{F}^{(4)}$ | $\mathbf{F}^{(5)}$ | $\mathbf{F}^{(6)}$ | $\mathbf{F}^{(4)}$ | $\mathbf{F}^{(5)}$ |
| $\mathbf{F}^{(7)}$ | $\mathbf{F}^{(8)}$ | $\mathbf{F}^{(9)}$ | $\mathbf{F}^{(7)}$ | $\mathbf{F}^{(8)}$ |
| $\mathbf{F}^{(1)}$ | $\mathbf{F}^{(2)}$ | $\mathbf{F}^{(3)}$ | $\mathbf{F}^{(1)}$ | $\mathbf{F}^{(2)}$ |
| $\mathbf{F}^{(4)}$ | $\mathbf{F}^{(5)}$ | $\mathbf{F}^{(6)}$ | $\mathbf{F}^{(4)}$ | $\mathbf{F}^{(5)}$ |

FIGURE 2.27 – Exemple de répartition de  $W = 9$  jeux de filtres.

La classification de ces mesures est effectuée avec le classifieur optimal de l'équation (1.18) connaissant les  $M = 3$  espèces de référence. Le résultat de cette classification apparaît en haut à droite de la figure 2.28. En bas à gauche de cette image apparaît également la classification qui a été réalisée dans la section 2.2 avec des mesures systématiquement effectuées avec le même jeu de filtres, soit pour  $W = 1$ .

Dans les deux cas la probabilité d'erreur de classification empirique dans les zones constituées d'une des trois espèces de référence est faible, en accord avec les bornes de Bhattacharyya correspondant. Des différences apparaissent dans les portions de l'image où l'espèce imprévue est présente.

Lorsque  $W = 1$  les pixels anormaux sont systématiquement attribués à l'espèce 3, ce qui rend l'anomalie impossible à détecter à partir des résultats de la classification. Il est de plus impossible de reconstruire les spectres rencontrés.

FIGURE 2.28 – Classification réalisée avec  $W = 1$  ou  $W = 144$  jeux de filtres.

Avec les  $W = 144$  jeux de filtres sélectionnés, la classification de ces pixels est très variable, comme on peut le constater en bas à droite de la figure 2.28. Cette variabilité est due à la diversité des jeux de filtres employés. Chacun d'entre eux est soumis à la même limitation que le jeu de filtres utilisé lorsque  $W = 1$ , c'est-à-dire que la projection du spectre de la quatrième espèce va être proche de l'une de celles des espèces de référence. Cependant, la diversité de jeux de filtres employés implique que l'espèce de référence la plus proche de la quatrième espèce change d'un pixel à l'autre. Pour cette raison, on voit apparaître des motifs dans la classification des zones correspondant à l'espèce 4. La régularité de ces motifs est due à la périodicité de la répartition des jeux de filtres dans l'image.

Les motifs visibles dans la classification des zones problématiques peuvent suffire à détecter la présence d'une espèce inconnue, au prix d'une hypothèse d'homogénéité sous-jacente. Il est en revanche probable que la détection pixel à pixel d'une anomalie ne puisse toujours pas se faire avec de bonnes performances.

Contrairement au cas où les mesures ont été systématiquement effectuées avec un même jeu de filtres, on dispose néanmoins de suffisamment de projections différentes pour avoir la possibilité de pousser plus loin l'analyse, en reconstruisant les spectres rencontrés. Ici, il est en particulier possible de reconstruire le spectre observé sur toute zone carrée de 12 pixels de côté. Une telle reconstruction peut être effectuée a posteriori sans mesure supplémentaire.

Par exemple, la reconstruction du spectre de la zone correspondant au rectangle noir apparaissant en haut du champ de vue sur la figure 2.29 est présentée à droite de cette même figure. La zone sélectionnée contient 15360 pixels. Le spectre reconstruit à partir des mesures effectuées dans cette zone correspond bien au spectre de la seconde espèce de référence, ce qui confirme qu'elle ne présente pas d'anomalie. Cela peut permettre de vérifier qu'il n'y a

pas de dérive des spectres de référence ou d'erreur de calibration de l'appareil.

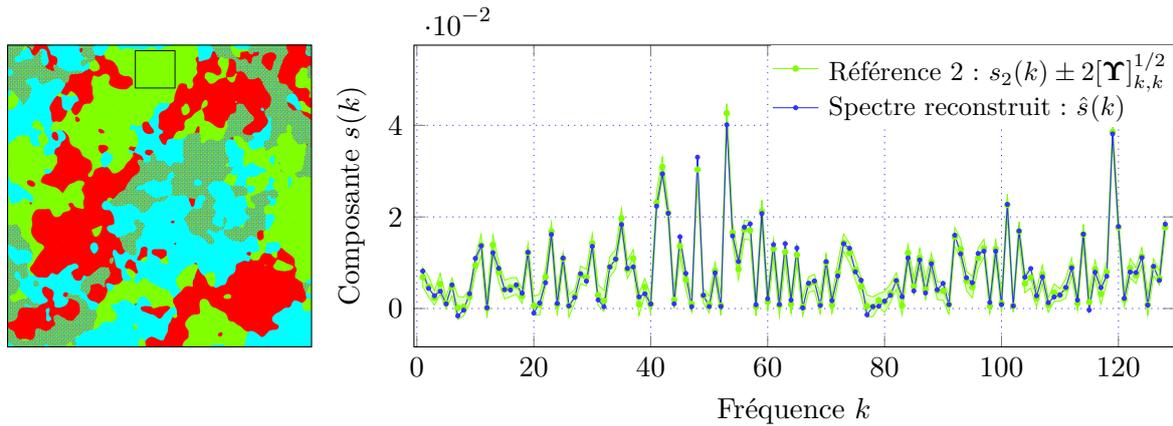


FIGURE 2.29 – Reconstruction de spectre et comparaison avec une référence.

Il est également possible de reconstruire le spectre observé sur une des zones où la classification semble incertaine. La partie droite de la figure 2.30 compare le spectre reconstruit à partir des mesures localisées dans le rectangle noir situé au centre du champ de vue apparaissant sur la partie gauche de cette figure au spectre de l'anomalie. La zone utilisée contient 4416 pixels. Cette fois, le spectre reconstruit diffère largement des spectres de référence, ce qui permet de confirmer la présence d'un problème. On peut également en déterminer l'origine, ce qui était impossible sans reconstruire ce spectre.

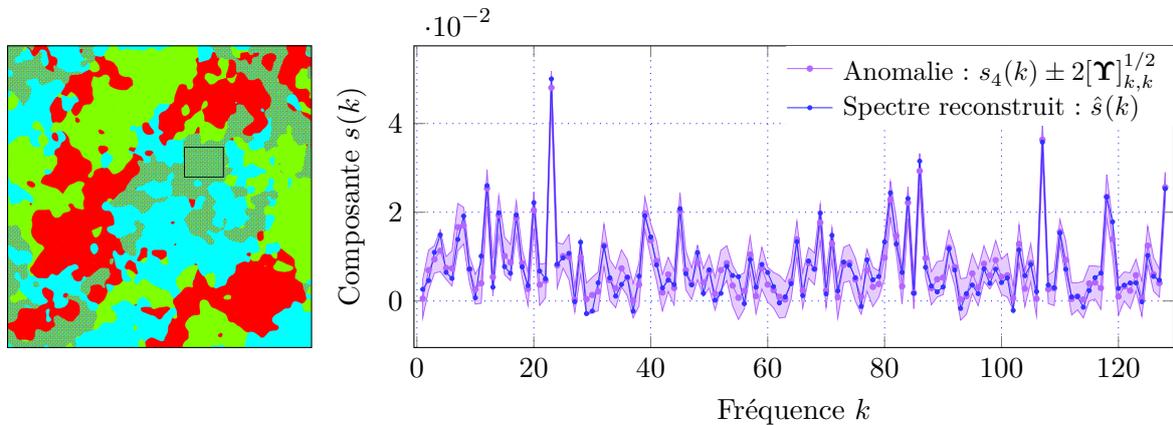


FIGURE 2.30 – Reconstruction du spectre de l'anomalie et comparaison avec le spectre vrai.

Dans cet exemple, la méthode proposée a permis de détecter et d'identifier sans mesure supplémentaire le composé inconnu par lequel l'échantillon analysé avait été pollué. Bien sûr, si les zones anormales ne sont pas de composition chimique homogène il n'est pas judicieux de reconstruire un spectre de la manière dont nous l'avons fait ici. Toutefois, rien n'empêche de mettre en place des méthodes de reconstruction plus poussées pour s'affranchir de cette limitation. Il est dans tous les cas nécessaire de s'assurer que la reconstruction puisse être effectuée correctement dans le cas favorable que nous avons traité ici, sans quoi il est peu probable qu'il existe une méthode qui parvienne à le faire avec de bonnes performances dans un cas plus difficile.

Une perspective à ce travail est de développer plus avant la détection d'anomalies. Ici, cette détection est effectuée visuellement, grâce à la présence de motifs dans la classification réalisée. Il est bien sûr possible de développer des tests d'inadéquation afin de réaliser la

détection de manière moins empirique. On peut alors utiliser le résultat de ces tests pour sélectionner les zones sur lesquelles il est intéressant de reconstruire un spectre, par exemple en couplant les résultats des tests d'inadéquation avec un algorithme de segmentation. Un exemple de méthodologie correspondant est présenté dans Justel, Galland et al. (2023b). Les concepts sur lesquels sont fondés à ce travail, notamment concernant les tests d'inadéquation, ont été dérivés des résultats présentés dans le chapitre suivant, aussi nous ne le présentons pas en détails ici.

## 2.8 Conclusion, limites et ouverture

Nous avons donné dans ce chapitre les outils nécessaires pour réaliser à la fois la classification de chaque pixel parmi les espèces de référence et la reconstruction des spectres rencontrés sur des zones homogènes. Une telle reconstruction donne une possibilité de détecter la présence et de déterminer la nature d'une espèce imprévue qui apparaîtrait dans l'échantillon.

Il est pour cela nécessaire d'utiliser une diversité de jeux de filtres. Le choix de ces jeux de filtres est crucial aux performances du système. Pour construire des filtres adaptés simultanément à la classification et à la reconstruction, il est possible d'utiliser l'approche à compromis optimaux. Cette méthode consiste à rechercher les ensembles de jeux de filtres qu'il n'est pas possible de surpasser à la fois en classification et en reconstruction. De tels ensembles de jeux de filtres peuvent être construits en minimisant une combinaison convexe des critères utilisés pour mesurer les performances en classification et en reconstruction de spectres, qui sont respectivement la borne de Bhattacharyya et la borne de Cramér-Rao.

L'approche à compromis optimaux permet non seulement de construire des filtres pertinents pour la classification et la reconstruction mais offre également un aperçu global des performances du système. Ceci nous a permis de réaliser une étude complète de l'influence des différents paramètres de l'acquisition sur les performances du système de mesure.

La limitation principale de cette méthodologie est celle de la classification. Puisque l'objectif de l'acquisition était de classer les pixels de la zone à imager parmi les espèces de référence, nous avons considéré que les pixels analysés étaient composés d'une unique espèce, ce qui implique que tout mélange soit considéré comme une anomalie.

Cette hypothèse limite le champ d'application de ce travail, car les méthodes développées sont inadaptées à des échantillons contenant des mélanges d'espèces. S'il existe dans un tel échantillon des zones de composition homogène, c'est-à-dire des zones où les concentrations des différentes espèces composant l'échantillon sont les mêmes, il est possible avec l'approche présentée dans ce chapitre de reconstruire le spectre du mélange correspondant et d'en déduire les concentrations. Il semble cependant plus approprié de traiter le cas des mélanges séparément. C'est que nous faisons dans le chapitre suivant, où nous explorons comment effectuer l'estimation des concentrations en chaque pixel avec de bonnes performances tout en étant capable de détecter les éventuelles anomalies.



# Chapitre 3

## Combiner estimation des concentrations et tests d'inadéquation en spectroscopie Raman comprimée

### Sommaire

---

|   |           |
|---|-----------|
| <b>3.1 Mélanges et stratégie d'acquisition . . . . .</b>  | <b>68</b> |
| 3.1.1 Rayonnement d'un mélange linéaire . . . . .   | 68        |
| 3.1.2 Dispositif d'acquisition considéré . . . . .  | 68        |
| <b>3.2 Détection des inadéquations . . . . .</b>  | <b>69</b> |
| 3.2.1 Ambiguïté à propos de l'origine du rayonnement . . . . .  | 69        |
| 3.2.2 Conditions à vérifier en chaque pixel . . . . .   | 70        |
| 3.2.3 Diversité de jeux de filtres et détection d'anomalies . . . . .   | 72        |
| 3.2.4 Test d'inadéquation . . . . .   | 73        |
| <b>3.3 Filtres à compromis optimaux entre estimation des concentrations et reconstruction de spectres . . . . .</b> | <b>75</b> |
| 3.3.1 Critères à optimiser . . . . .  | 76        |
| 3.3.2 Algorithme de recherche des compromis optimaux . . . . .  | 78        |
| <b>3.4 Exemple d'application . . . . .</b>  | <b>78</b> |
| 3.4.1 Optimisation des jeux de filtres . . . . .  | 78        |
| 3.4.2 Performances en détection des jeux de filtres sélectionnés . . . . .  | 80        |
| 3.4.3 Estimation robuste des concentrations . . . . .   | 81        |
| <b>3.5 Conclusion et ouverture . . . . .</b>  | <b>85</b> |

---

Dans le chapitre précédent, nous nous sommes intéressés à la classification de mesures de spectroscopie Raman comprimée parmi des espèces chimiques de référence de spectres connus. Dans ce cadre, le rayonnement observé en chaque pixel de l'échantillon analysé est censé ne provenir que d'une seule espèce chimique à la fois, et tout mélange est par conséquent considéré comme une anomalie. Cette approche n'est donc pas adaptée au cas d'échantillons composés de mélanges des espèces de référence.

L'objectif de ce chapitre est de déterminer comment réaliser une estimation robuste des concentrations d'un mélange d'espèces chimiques de référence de spectres connus, c'est-à-dire comment faire cette estimation avec de bonnes performances si l'a priori concernant les espèces en présence est vérifié tout en étant capable de détecter un écart à cette hypothèse.

Nous commençons par étudier dans quels cas peuvent être réalisés des tests permettant de vérifier si le mélange observé est en adéquation avec un mélange des espèces de référence. Cette étude conduit à établir plusieurs contraintes sur la forme et la diversité des jeux de filtres nécessaires à ne pas être systématiquement aveugle à une part non négligeable d'anomalies.

Dans la continuité du chapitre précédent, nous proposons ensuite d'étendre l'approche à compromis optimaux au cas des mélanges afin de construire une diversité de jeux de filtres vérifiant les contraintes établies pour la détection d'anomalies et permettant d'estimer au mieux les concentrations des espèces de référence en chaque pixel.

Cette méthodologie est illustrée sur un exemple simulé, où nous arrivons à détecter la présence d'une espèce imprévue mélangée avec les espèces de référence sans connaître son spectre et sans mesure supplémentaire. Nous concluons en discutant des limitations et perspectives de la méthode.

## 3.1 Mélanges et stratégie d'acquisition

### 3.1.1 Rayonnement d'un mélange linéaire

Le premier objectif de l'acquisition est d'estimer les concentrations des espèces chimiques de référence en chaque pixel. On considère donc que l'échantillon analysé est constitué de mélanges linéaires de ces  $M$  espèces dont les spectres  $\mathbf{s}_1, \dots, \mathbf{s}_M$  sont connus.

En chaque pixel  $w$  les flux  $\phi^{(w)}(k)$  du rayonnement diffusé dans les  $K$  bandes spectrales considérées sont par conséquent les combinaisons linéaires des coefficients  $s_u(k)$  dont les poids sont les concentrations  $c_1^{(w)}, \dots, c_M^{(w)}$  en chacune des espèces, c'est-à-dire

$$\phi^{(w)}(k) = \sum_{u=1}^M c_u^{(w)} s_u(k). \quad (3.1)$$

Ce cadre est celui de Wilcox, Buzzard et al. (2012, 2013), Réfrégier, Scotté et al. (2018) et Rehrauer, Dinh et al. (2018) rappelé dans la section 1.3.2.

### 3.1.2 Dispositif d'acquisition considéré

Comme évoqué dans la section 1.3.2 il faut appliquer au moins  $M$  filtres linéairement indépendants en chaque pixel afin de pouvoir réaliser l'estimation des concentrations des  $M$  espèces de référence, ce qui implique  $P \geq M$ . Cette contrainte n'était pas appliquée dans le chapitre précédent puisqu'elle n'est pas nécessaire à réaliser la classification des mesures parmi les espèces de référence.

Il est difficile d'implémenter expérimentalement un spectromètre Raman permettant d'appliquer plus de  $P = 2$  filtres en parallèle. Contrairement au chapitre précédent, nous

considérons par conséquent dans ce chapitre que les  $P$  filtres sont appliqués séquentiellement. Appliquer une contrainte sur le nombre total de photons détectés n'a alors plus de sens, et chacune des  $P$  mesures correspondant est donc réalisée à temps fixé.

Soient  $\mathbf{n}^{(w)} = {}^t(n_1^{(w)}, \dots, n_P^{(w)})$  les mesures réalisées au pixel  $w$  avec le jeu de filtres  $\mathbf{F}^{(w)} = (\mathbf{f}_1^{(w)}, \dots, \mathbf{f}_P^{(w)})$ . On pose  $\tau_m^{(w)}$  le temps de mesure associé au filtre  $\mathbf{f}_m^{(w)}$ . Afin d'alléger les notations, nous supposons dans la suite que tous les temps de mesures sont identiques, c'est-à-dire  $\tau_m^{(w)} = \tau$ . L'équation 1.4 devient alors

$$\mathcal{P}(\mathbf{n}^{(w)}) = \prod_{m=1}^P \frac{(\tau \phi_m^{(w)})^{n_m^{(w)}}}{n_m^{(w)}!} \exp(-\tau \phi_m^{(w)}) \quad (3.2)$$

avec

$$\phi_m^{(w)} = \mathbf{f}_m^{(w)} \cdot \sum_{u=1}^M c_u^{(w)} \mathbf{s}_u \quad (3.3)$$

le flux observé en sortie du filtre  $\mathbf{f}_m^{(w)}$ .

## 3.2 Détection des inadéquations

À l'instar de la classification, l'estimation des concentrations à partir de mesures comprimées s'effectue sous hypothèse de la connaissance a priori des espèces chimiques composant l'échantillon analysé et de leurs spectres. Ainsi, on considère lorsque l'on réalise cette estimation que l'échantillon est constitué uniquement de mélanges des espèces de référence.

Si cette hypothèse n'est pas vérifiée, les concentrations estimées peuvent être complètement erronées. Cela sera par exemple le cas si le mélange observé ne contient aucune des espèces de référence. Il est donc crucial de détecter les anomalies, c'est-à-dire les écarts à cette hypothèse.

Si cela ne pose pas de difficulté particulière lorsque l'on peut reconstruire le spectre du mélange observé en chaque pixel, comme c'est le cas avec un spectromètre conventionnel, cela est plus délicat dans le cadre de la spectroscopie comprimée. Dans cette section, nous établissons des contraintes qu'il est nécessaire d'imposer sur la forme des filtres appliqués en chaque pixel ainsi que sur la diversité des jeux de filtres afin de pouvoir détecter les anomalies.

### 3.2.1 Ambiguïté à propos de l'origine du rayonnement

#### Classe d'équivalence des spectres

Dès lors que l'on considère un rayonnement diffusé par un mélange d'espèces chimiques de concentrations inconnues, il existe une ambiguïté dans la manière de formuler ce mélange. Il est en effet possible de l'écrire avec plusieurs combinaisons de spectres et de concentrations.

Soient  $\mathbf{S} = (\mathbf{s}_1, \dots, \mathbf{s}_M)$  des spectres normalisés et  $\tilde{\mathbf{D}}$  une matrice de taille  $M \times \tilde{M}$  dont les colonnes se somment à 1 et dont les coefficients sont positifs. Le jeu de spectres  $\tilde{\mathbf{S}} = \mathbf{S} \tilde{\mathbf{D}}$  est constitué de  $\tilde{M}$  combinaisons convexes des spectres de  $\mathbf{S}$ . Les spectres de  $\tilde{\mathbf{S}}$  sont à la fois positifs et normalisés. Tout mélange de  $\tilde{\mathbf{S}}$  avec des concentrations positives peut s'exprimer sous la forme d'un mélange de  $\mathbf{S}$  avec des concentrations également positives. On a en effet

$$\tilde{\mathbf{S}} \tilde{\mathbf{c}}^{(w)} = \mathbf{S} \tilde{\mathbf{D}} \tilde{\mathbf{c}}^{(w)} = \mathbf{S} \mathbf{c}^{(w)}$$

avec  $\mathbf{c}^{(w)} = \tilde{\mathbf{D}} \tilde{\mathbf{c}}^{(w)} \geq 0$ . De ce fait, tout mélange physique de  $\tilde{\mathbf{S}}$  est indiscernable d'un mélange physique de  $\mathbf{S}$ . Ce que l'on appelle ici mélange physique est un mélange de spectres

normalisés et positifs en concentrations positives, qui pourrait par conséquent correspondre à un véritable mélange d'espèces chimiques.

Une telle classe d'équivalence apparaît même sans filtrage. On retrouve pour  $M = \tilde{M}$  l'indétermination classique des sources d'un mélange. Ces sources ne peuvent être estimées qu'à une permutation ou un changement d'échelle près, comme évoqué par exemple dans Comon et Jutten (2010, section 1.2).

Dans la suite de ce chapitre, on désigne par *ensemble d'espèces anormal* un ensemble d'espèces chimiques dont les spectres  $\tilde{\mathbf{S}}$  n'appartiennent pas à la classe d'équivalence des spectres de référence. La présence d'un de ces ensembles d'espèce anormal dans l'échantillon constitue une *anomalie* qu'il faut détecter.

### Ambiguïté supplémentaire introduite par le filtrage

En chaque pixel, les détecteurs comptent les photons diffusés par l'échantillon arrivant en sortie des filtres. Si le modèle est vérifié, c'est-à-dire si l'on observe bien un mélange des espèces de référence, la distribution de ces mesures est celle de l'équation (3.2). Cette distribution dépend des concentrations  $\mathbf{c}^{(w)}$  inconnues.

Il ne sera possible de détecter un écart au modèle que si la distribution des mesures du rayonnement correspondant à cet écart ne peut s'exprimer sous cette forme, et plus précisément que si les flux observés par les détecteurs ne peuvent pas s'écrire à partir des spectres de référence comme dans l'équation (3.3). Ainsi, un mélange d'espèces de spectres  $\tilde{\mathbf{S}}$  en concentrations  $\tilde{\mathbf{c}}^{(w)}$  sera indiscernable d'un mélange des espèces de référence *lorsqu'observé à travers*  $\mathbf{F}^{(w)}$  si et seulement si il existe des concentrations  $\mathbf{c}^{(w)}$  telles que

$${}^t\mathbf{F}^{(w)} \tilde{\mathbf{S}} \tilde{\mathbf{c}}^{(w)} = {}^t\mathbf{F}^{(w)} \mathbf{S} \mathbf{c}^{(w)}.$$

Le filtrage peut donc introduire une ambiguïté entre un mélange de spectres  $\mathbf{S}$  et un mélange de spectres  $\tilde{\mathbf{S}}$  qui ne sont pas dans la classe d'équivalence de  $\mathbf{S}$ . Comme nous le montrons dans la suite de cette section, il est nécessaire que les filtres possèdent certaines caractéristiques pour que ce type d'ambiguïté n'empêche pas de détecter une part significative des anomalies.

### 3.2.2 Conditions à vérifier en chaque pixel

L'objectif de l'acquisition est d'estimer les concentrations des  $M$  espèces chimiques de référence dans le mélange observé à chaque pixel. Puisque multiplier le nombre de filtres appliqués en chaque pixel signifie souvent augmenter le temps d'acquisition ou la complexité du système de mesure, il est fréquemment proposé d'utiliser le nombre minimal de filtres nécessaire à réaliser cette tâche, c'est-à-dire  $M$  filtres. C'est par exemple le cas dans Wilcox, Buzzard et al. (2012, 2013) et Réfrégier, Scotté et al. (2018) ainsi que dans la majeure partie des travaux portant sur l'estimation des concentrations présentés dans la section 1.2.2.

Cependant, ce choix implique d'être aveugle à une part non négligeable des anomalies. En effet, nous démontrons en annexe F.3 la proposition 3 ci-dessous.

**Proposition 3** (Diversité de filtres nécessaire à ne pas être aveugle). Soit un jeu de filtres  $\mathbf{F}^{(w)}$  permettant d'estimer les concentrations des  $M$  espèces de référence, donc tel que  ${}^t\mathbf{F}^{(w)} \mathbf{S}$  soit de rang plein. Afin de ne pas être aveugle à une part non négligeable des anomalies lorsque l'on effectue des mesures avec  $\mathbf{F}^{(w)}$  il est nécessaire et suffisant que

$$\text{rg}(\mathbf{F}^{(w)}) > M.$$

**Corollaire 3.1** (Nombre de filtres nécessaire à ne pas être aveugle). Si  $P$  est le nombre de filtres employés pour réaliser les mesures alors il est nécessaire que

$$P > M.$$

**Idée de preuve.** Une manière d'aborder le problème est de s'intéresser aux flux de photons sur les détecteurs. Si on observe le rayonnement diffusé par un mélange des espèces de référence à travers un jeu de filtres  $\mathbf{F}^{(w)}$  alors les flux  $\phi^{(w)}$  observés sur les détecteurs sont de la forme de l'équation (1.7), c'est-à-dire

$$\phi^{(w)} = {}^t\mathbf{F}^{(w)} \mathbf{S} \mathbf{c}^{(w)}.$$

Les flux observés se situent donc dans un sous-ensemble de  $\mathbb{R}^P$  engendré par les colonnes de la matrice  ${}^t\mathbf{F}^{(w)} \mathbf{S}$  et qui peut être parcouru en faisant varier les concentrations  $\mathbf{c}^{(w)}$ . On désigne ce sous-ensemble par *facette* dans la suite. Afin de réaliser l'estimation des concentrations des espèces de référence, il est nécessaire que  ${}^t\mathbf{F}^{(w)} \mathbf{S}$  soit de rang plein. Cette facette est donc de dimension  $M$ .

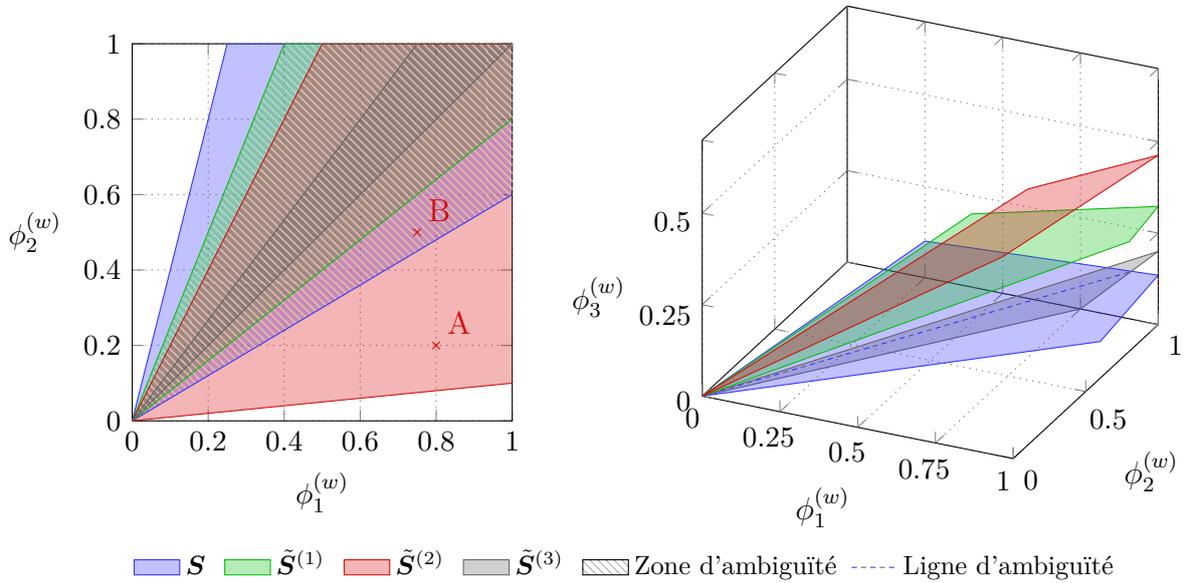


FIGURE 3.1 – Visualisation des facettes associées à des jeux de  $M = 2$  espèces chimiques lorsque  $P = 2$  (à gauche) ou  $P = 3$  (à droite) filtres sont appliqués pour effectuer les mesures.

Lorsque  $P = M$  la facette est de même dimension que l'espace dans laquelle elle existe. Il existe donc un recouvrement non négligeable entre la facette correspondant à  $\mathbf{S}$  et celles correspondant à la plupart des autres jeux de spectres. Cela est illustré sur la partie gauche de la figure 3.1. Puisque la facette correspondant au jeu de spectres  $\tilde{\mathbf{S}}^{(1)}$  est située à l'intérieur de celle de  $\mathbf{S}$ , aucun mélange des espèces correspondant à  $\tilde{\mathbf{S}}^{(1)}$  ne pourra être distingué d'un mélange des espèces de référence, et de même pour  $\tilde{\mathbf{S}}^{(3)}$ . Pour certains jeux de spectres, le recouvrement est partiel. Par exemple, si les flux observés sont ceux du point B alors il y a ambiguïté entre un mélange de  $\tilde{\mathbf{S}}^{(2)}$  et un mélange de  $\mathbf{S}$ , tandis qu'il est possible de détecter que le mélange de  $\tilde{\mathbf{S}}^{(2)}$  correspondant au point A n'est pas conforme au modèle.

Il est possible de mitiger ce problème en réalisant des mesures avec strictement plus de filtres linéairement indépendants que d'espèces de référence, ce qui implique  $P > M$ . La facette correspondant à  $\mathbf{S}$  étant toujours de dimension  $M$ , son recouvrement avec une autre

facette est de dimension inférieure ou égale à  $M$  et donc de mesure nulle dans l'espace des flux  $\mathbb{R}^P$ . Ainsi, cette facette est discernable de la plupart de celles des autres jeux de spectres, comme on peut le voir sur la partie droite de la figure 3.1.

De ce fait, si la contrainte de la proposition 3 est vérifiée alors la plupart des mélanges de la plupart des ensembles d'espèces chimiques anormaux sont potentiellement discernables de mélanges des espèces de référence. Autrement dit, bien qu'il soit toujours possible qu'un ensemble d'espèces anormal donné soit indiscernable des espèces de référence lorsqu'apparaissant dans l'échantillon analysé avec certaines concentrations, ces concentrations appartiennent en général à un ensemble de mesure nulle.

Le problème de robustesse n'est pas entièrement résolu puisque vérifier cette contrainte n'élimine pas totalement la possibilité qu'existent des ensembles d'espèces anormaux dont les mélanges sont *systématiquement* indiscernables de mélanges des espèces de référence, c'est-à-dire peu importe les concentrations. Comme nous allons le voir dans la section suivante, il faut pour cela employer une diversité de jeux de filtres.

### 3.2.3 Diversité de jeux de filtres et détection d'anomalies

Pour faire en sorte qu'en chaque pixel les seules espèces systématiquement indiscernables des espèces de référence soient celles dont les spectres sont dans la classe d'équivalence de  $\mathbf{S}$  il est nécessaire que  $P = K$ . Il n'y a en effet que dans ce cas qu'il est possible de reconstruire le spectre observé en chaque pixel pour tester s'il correspond à une combinaison linéaire des spectres de référence. Cela revient toutefois à utiliser un spectromètre conventionnel plutôt qu'un spectromètre comprimé.

Ici, on veut conserver  $P \ll K$  et le noyau de chaque jeu de filtres est donc forcément non nul. Cela signifie qu'il existe des anomalies qui seront systématiquement indiscernables des espèces de référence en chaque pixel. Vérifier les contraintes établies dans la proposition 3 permet de faire en sorte que cela ne soit pas le cas de la plupart des anomalies. Cela reste néanmoins vrai pour certaines d'entre elles, dépendamment du jeu de filtres.

L'enjeu est alors qu'aucun ensemble d'espèces anormal ne soit systématiquement indiscernable des espèces de référence pour tous les jeux de filtres employés. Autrement dit, il faut qu'il existe pour tout ensemble anormal au moins un jeu de filtres pour lequel il n'y a pas systématiquement ambiguïté avec les espèces de référence. Si cet objectif est atteint, tout ensemble anormal apparaissant sur plusieurs pixels est alors potentiellement détectable en combinant les mesures effectuées sur ces pixels avec suffisamment de jeux de filtres différents. Cela dépendra évidemment des concentrations des espèces correspondant, mais il n'est pas nécessaire que ces concentrations soient égales. Pour atteindre cet objectif, il est nécessaire d'employer une diversité de jeux de filtres. Plus précisément, on démontre la proposition suivante dans l'annexe F.4.

**Proposition 4.** Soit  $\mathbf{F} = (\mathbf{F}^{(1)}, \dots, \mathbf{F}^{(W)})$  un ensemble de jeux de filtres tel que  $\forall w \in \llbracket 1, W \rrbracket$   $\text{rg}(\mathbf{F}^{(w)}) > M$  et  ${}^t\mathbf{F}^{(w)}\mathbf{S}$  soit de rang plein. On fixe  $\mathbf{B}^{(\perp, w)}$  une base normée de  $\ker({}^t\mathbf{F}^{(w)})$  et  $\mathbf{B}^{(\perp, w)}$  une base normée du supplémentaire orthogonal de  $\text{vect}(\mathbf{S}, \mathbf{B}^{(\perp, w)})$ . Afin que pour tout  $\tilde{\mathbf{S}}$  anormal il existe au moins un  $w$  tel que l'ensemble des mélanges de  $\tilde{\mathbf{S}}$  indiscernables de mélanges de  $\mathbf{S}$  lorsqu'observés à travers  $\mathbf{F}^{(w)}$  soit négligeable, il est nécessaire que

$$\text{vect}(\mathbf{S}, \mathbf{B}^{(\perp, 1)}, \dots, \mathbf{B}^{(\perp, W)}) = \mathbb{R}^K$$

où  $\text{vect}(\mathbf{A})$  est l'espace vectoriel engendré par  $\mathbf{A}$ .

**Corollaire 4.1.** Si  $W$  jeux de  $P$  filtres sont utilisés alors il est nécessaire que

$$W \geq \left\lceil \frac{K - M}{P - M} \right\rceil.$$

**Idée de preuve.** Soit  $\mathbf{F}^{(w)}$  tel que  $\text{rg}(\mathbf{F}^{(w)}) > M$  et  ${}^t\mathbf{F}^{(w)}\mathbf{S}$  soit de rang plein. Il est possible d'écrire tout spectre comme une combinaison linéaire des spectres  $\mathbf{S}$ , de vecteurs de  $\mathbf{B}^{(\perp, w)}$  et de vecteurs de  $\mathbf{B}^{(\mathcal{L}, w)}$ . Lorsque  $\mathbf{F}^{(w)}$  est appliqué au rayonnement incident, la partie non filtrée de ce rayonnement est localisée dans  $\text{vect}(\mathbf{S}, \mathbf{B}^{(\mathcal{L}, w)})$ .

Si  $(\mathbf{S}, \mathbf{B}^{(\mathcal{L}, 1)}, \dots, \mathbf{B}^{(\mathcal{L}, W)})$  ne génère pas  $\mathbb{R}^K$  alors il est possible de compléter cette famille avec une base  $\mathbf{E}$  de son supplémentaire orthogonal pour former une famille génératrice de  $\mathbb{R}^K$ . Toute anomalie dont les spectres peuvent s'écrire comme une combinaison linéaire de spectres de  $\mathbf{S}$  et de vecteurs de  $\mathbf{E}$  apparaîtra en chaque pixel comme une combinaison des spectres de référence et sera donc indétectable. Il faut donc bien que  $\text{vect}(\mathbf{S}, \mathbf{B}^{(\mathcal{L}, 1)}, \dots, \mathbf{B}^{(\mathcal{L}, W)}) = \mathbb{R}^K$ .

On a  $\dim(\text{vect}(\mathbf{B}^{(\mathcal{L}, w)})) = \text{rg}(\mathbf{F}^{(w)}) - M \leq P - M$  et  $\dim(\text{vect}(\mathbf{S})) = M$ . Par conséquent  $\dim(\text{vect}(\mathbf{S}, \mathbf{B}^{(\mathcal{L}, 1)}, \dots, \mathbf{B}^{(\mathcal{L}, W)})) \leq W(P - M) + M$  et il faut que cette dimension soit égale à  $\dim(\mathbb{R}^K) = K$ , d'où le corollaire.

### 3.2.4 Test d'inadéquation

Dans le cadre de l'estimation des concentrations, on suppose que l'échantillon analysé est constitué de mélanges linéaires des espèces de référence, dont on connaît les spectres. L'hypothèse à tester pour vérifier la présence d'anomalie est donc simplement :

$H_0$  : Les spectres observés sont des combinaisons linéaires des spectres de référence.

Lorsque les contraintes des propositions 3 et 4 sont vérifiées, la plupart des mélanges comportant au moins une espèce anormale diffusent sur les détecteurs des flux de photons qui n'auraient pas pu être produits par un mélange physique des espèces de référence. Tester si  $H_0$  est vérifiée est donc presque équivalent à tester l'hypothèse qui suit :

$$H'_0 : \forall w \in \llbracket 1, W \rrbracket \exists \mathbf{c}^{(w)} \text{ tel que } \phi^{(w)} = {}^t\mathbf{F}^{(w)}\mathbf{S}\mathbf{c}^{(w)}.$$

Pour détecter les écarts à cette hypothèse nous proposons de mettre en place un test de rapport de vraisemblance, ou LRT (Likelihood Ratio Test). Ce test d'inadéquation classique et ses propriétés sont présentés par exemple dans Garthwaite, Jolliffe et al. (2002, section 4.6). On pose comme auparavant  $\mathbf{n}$  la transposée de  $({}^t\mathbf{n}^{(1)}, \dots, {}^t\mathbf{n}^{(W)})$ . Ici, la statistique du LRT effectué à partir de ces mesures pour décider entre  $H'_0$  et son complémentaire  $\overline{H'_0}$  s'écrit

$$T(\mathbf{n}) = -2 \log \frac{\mathcal{P}(\mathbf{n} | \mathbf{c}, \mathbf{S})}{\mathcal{P}(\mathbf{n} | \phi)} \quad (3.4)$$

où  $\hat{\mathbf{c}}$  est la concaténation des  $\mathbf{c}^{(w)}$  et  $\phi$  celle des  $\phi^{(w)}$ .

L'observateur n'a bien sûr pas accès à la valeur exacte des concentrations ou des flux, mais peut les estimer à partir des nombres de photons  $n_m^{(w)}$  détectés en chaque pixel et les injecter dans l'expression du test. Celui-ci devient alors un test de rapport de vraisemblance généralisé, ou GLRT (Generalized Likelihood Ratio Test). La statistique de ce test est

$$\hat{T}(\mathbf{n}) = -2 \log \frac{\max_{\mathbf{c}} \mathcal{P}(\mathbf{n} | \mathbf{c}, \mathbf{S})}{\max_{\phi} \mathcal{P}(\mathbf{n} | \phi)} = -2 \log \frac{\mathcal{P}(\mathbf{n} | \hat{\mathbf{c}}, \mathbf{S})}{\mathcal{P}(\mathbf{n} | \hat{\phi})} \quad (3.5)$$

où  $\hat{\mathbf{c}}$  est la concaténation des  $\hat{\mathbf{c}}^{(w)}$  et  $\hat{\phi}$  celle des  $\hat{\phi}^{(w)}$ , ces estimateurs étant ceux du maximum de vraisemblance. Nous abordons le calcul de ces estimateurs en fin de cette section. Puisque les différentes mesures sont indépendantes on a

$$\begin{aligned}\hat{T}(\mathbf{n}) &= -2 \log \prod_{w=1}^W \prod_{m=1}^P \frac{\mathcal{P}(n_m^{(w)} | \hat{\mathbf{c}}^{(w)}, \mathbf{S})}{\mathcal{P}(n_m^{(w)} | \hat{\phi}^{(w)})} \\ &= -2 \sum_{w=1}^W \sum_{m=1}^P \left[ n_m^{(w)} \log \left( \frac{{}^t \mathbf{f}_m^{(w)} \mathbf{S} \hat{\mathbf{c}}^{(w)}}{\hat{\phi}_m^{(w)}} \right) - \tau \left( {}^t \mathbf{f}_m^{(w)} \mathbf{S} \hat{\mathbf{c}}^{(w)} - \hat{\phi}_m^{(w)} \right) \right].\end{aligned}$$

Le test lui-même s'écrit

$$\hat{T}(\mathbf{n}) \underset{H_0'}{\overset{H_0'}{\gtrless}} \mu \quad (3.6)$$

où  $\mu$  est un seuil à choisir. Comme démontré dans Cox et Hinkley (1974, section 9.3), si l'hypothèse  $H_0'$  est vérifiée alors la statistique  $\hat{T}(\mathbf{n})$  est asymptotiquement distribuée selon une loi du  $\chi^2$  dont le nombre de degrés de liberté est égal à la différence entre le nombre de paramètres indépendants estimés sous  $H_0'$  et le nombre de paramètres indépendants estimés sous  $H_0$  injectés dans le calcul de la vraisemblance. Dans notre cas il y a  $W \times P$  flux estimés sous  $H_0'$  et  $W \times M$  concentrations estimées sous  $H_0$ , donc

$$\hat{T}(\mathbf{n}) \xrightarrow[n_m^{(w)} \rightarrow +\infty]{Loi} \chi_{W \times (P-M)}^2 \quad \text{sous } H_0'.$$

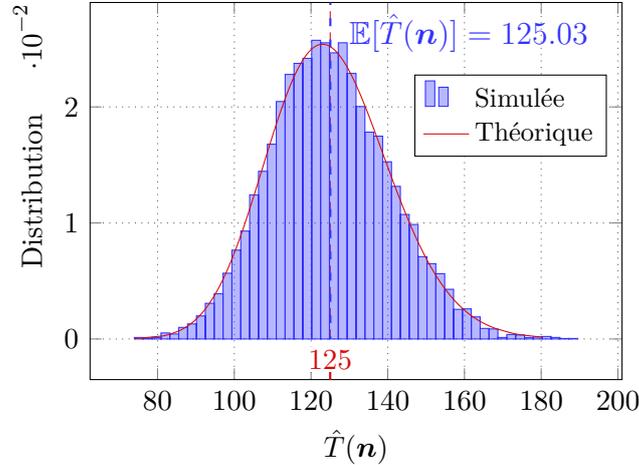


FIGURE 3.2 – Distribution du GLRT sous hypothèse  $H_0'$ .

Pour que ce résultat soit vérifié il est nécessaire que les estimateurs des paramètres soient ceux du maximum de vraisemblance. Comme nous le verrons en fin de cette section, l'estimateur du maximum de vraisemblance des concentrations sous hypothèse  $H_0'$  n'admet pas d'expression analytique lorsque  $P > M$ , et nous le calculons avec un algorithme itératif. Il ne s'agit que d'une distribution asymptotique, et donc d'une approximation de la véritable distribution de  $\hat{T}(\mathbf{n})$ . Pour vérifier la qualité de cette approximation, nous avons simulé des mesures de mélanges de  $M = 3$  espèces de référence dont les spectres ont été tirés uniformément sur le simplexe (c'est-à-dire comme dans l'équation (2.14) avec  $\alpha = 1$ ), avec  $W = 125$  jeux de  $P = 4$  filtres binaires optimisés pour l'estimation des concentrations de ces espèces. Les concentrations des espèces de référence ont été tirées uniformément entre 0 et 1.

Le temps d'acquisition étant fixé à  $10^3$ , cela correspond à la détection de 1500 photons par filtre en moyenne. Chacune des  $W \times P$  mesures a été simulée  $10^4$  fois, ce qui correspond à  $10^4$  valeurs de  $\hat{T}(\mathbf{n})$ . La figure 3.2 compare la distribution théorique de  $\hat{T}(\mathbf{n})$  à l'histogramme normalisé des valeurs de  $\hat{T}(\mathbf{n})$  obtenues en simulation. Sur cet exemple, la loi du  $\chi^2$  est une bonne approximation de la loi de  $\hat{T}(\mathbf{n})$ , et nous avons pu constater que cela reste le cas tant qu'au moins quelques centaines de photons par filtre sont détectés.

Cette approximation de la distribution du test sous hypothèse  $H'_0$  permet en particulier de choisir la probabilité de fausse alarme (PFA) en fixant la valeur de  $\mu$  appropriée, c'est-à-dire

$$\mu(\text{PFA}) = \mathcal{F}_{\chi^2}^{-1}(1 - \text{PFA}),$$

où  $\mathcal{F}_{\chi^2}^{-1}$  est l'inverse de fonction de répartition du  $\chi^2$  à  $W \times (P - M)$  degrés de liberté.

**Estimateur du maximum de vraisemblance des flux.** Il est possible d'estimer en chaque pixel les flux observés par chaque détecteur. On rappelle que le nombre de photons détectés par le détecteur  $D_m$  suit au pixel  $w$  une loi de Poisson dont la moyenne est le produit entre le flux  $\phi_m^{(w)}$  et le temps d'acquisition  $\tau$ , c'est-à-dire

$$\mathcal{P}(n_m^{(w)} | \phi^{(w)}) = \left( \tau \phi_m^{(w)} \right)^{n_m^{(w)}} \frac{\exp(-\tau \phi_m^{(w)})}{n_m^{(w)}!}.$$

Sans autre hypothèse, l'estimateur du maximum de vraisemblance des flux est simplement le rapport du nombre de photons détectés par le temps d'acquisition, c'est-à-dire

$$\hat{\phi}_m^{(w)} = \frac{n_m^{(w)}}{\tau}.$$

**Estimateur du maximum de vraisemblance des concentrations.** Lorsque les hypothèses à propos des espèces en présence sont vérifiées, les flux observés par les détecteurs peuvent s'exprimer à partir des spectres des espèces de référence et de leurs concentrations dans le mélange analysé, c'est-à-dire

$$\mathcal{P}(n_m^{(w)} | \mathbf{c}^{(w)}, \mathbf{S}) = \left( \tau \phi_m^{(w)} \right)^{n_m^{(w)}} \frac{\exp(-\tau \phi_m^{(w)})}{n_m^{(w)}!}.$$

avec

$$\phi_m^{(w)} = {}^t \mathbf{F}^{(w)} \mathbf{S} \mathbf{c}^{(w)}.$$

L'estimateur du maximum de vraisemblance des concentrations n'admet pas d'expression analytique lorsque  $P > M$ . Nous utilisons pour le calculer un algorithme de scoring de Fisher initialisé avec l'estimateur de la pseudo-inverse de l'équation (1.9).

### 3.3 Filtres à compromis optimaux entre estimation des concentrations et reconstruction de spectres

Dans la section précédente, nous avons montré qu'il est nécessaire que les jeux de filtres vérifient certaines contraintes afin de ne pas être aveugle à la plupart des anomalies. Ces contraintes, énoncées dans les propositions 3 et 4, portent sur le rang des jeux de filtres et sur leur diversité. Elles ne suffisent pas à désigner un ensemble de jeux de filtres à utiliser. Il

reste donc à déterminer comment construire ces filtres. Nous proposons pour cela d'utiliser l'approche à compromis optimaux présentée dans le chapitre précédent.

Le premier objectif de l'acquisition reste d'estimer les concentrations des espèces de référence en chaque pixel avec les meilleures performances possibles. Le premier critère à optimiser est donc évidemment un critère de performance en estimation des concentrations.

Comme précédemment dans le cas de la classification, optimiser les performances en estimation des concentrations ne conduit pas à une diversité de jeux de filtres, puisqu'il est optimal pour cette tâche d'utiliser toujours le même jeu de filtres, comme cela est par exemple fait dans Wilcox, Buzzard et al. (2012, 2013) et Réfrégier, Scotté et al. (2018).

Dans la continuité du chapitre précédent, nous proposons que le second critère à optimiser soit un critère de performance en reconstruction des spectres observés. Nous avons en effet vu qu'optimiser un tel critère permet d'accroître la diversité des jeux de filtres construits et que l'approche à compromis optimaux permet de le faire à faible coût.

On peut également remarquer qu'il est nécessaire pour reconstruire le spectre du rayonnement observé de projeter ce rayonnement sur les détecteurs dans des directions autres que celles des spectres de référence. Empiriquement, optimiser un critère de performance en reconstruction de spectres va donc dans le sens de la proposition 4, mais en prenant en compte le bruit. On peut donc penser que cela est intéressant pour la détection des anomalies.

D'autres critères sont bien sûr possibles, notamment si l'on dispose de plus d'informations à propos des anomalies à détecter ou des concentrations. Nous confirmons plus loin la pertinence de l'approche que nous proposons ici sur un exemple simulé. Dans cette section, nous détaillons les différents critères à optimiser ainsi que l'algorithme permettant de construire des filtres réalisant des compromis optimaux entre les critères choisis. Cette approche nous permet d'obtenir une diversité de jeux de filtres ayant des performances en estimation des concentrations proches de l'optimal.

### 3.3.1 Critères à optimiser

#### Critère de performance en estimation des concentrations

Comme rappelé dans la section 1.3.2, un critère pertinent pour évaluer les performances d'un jeu de filtres  $\mathbf{F}^{(w)}$  en estimation des concentrations est  $\text{tr } \mathbf{\Upsilon}_c^{(w)}$  la trace de la matrice de Cramér-Rao. Cette trace est une borne inférieure à la somme des erreurs quadratiques moyennes de tout estimateur non biaisé des concentrations

La matrice de Cramér-Rao  $\mathbf{\Upsilon}_c^{(w)}$  est l'inverse de la matrice d'information de Fisher  $\mathbf{I}_c^{(w)}$ . En notant  $\mathbf{s}_1, \dots, \mathbf{s}_M$  les spectres des espèces de référence,  $\mathbf{c}^{(w)}$  leurs concentrations et  $\tau$  le temps de mesure, l'équation (1.12) donne

$$[\mathbf{I}_c^{(w)}]_{p,q} = \sum_{m=1}^P \tau \frac{(\mathbf{f}_m^{(w)} \cdot \mathbf{s}_p)(\mathbf{f}_m^{(w)} \cdot \mathbf{s}_q)}{\mathbf{f}_m^{(w)} \cdot \sum_{u=1}^M c_u^{(w)} \mathbf{s}_u}.$$

Cette expression dépend des concentrations  $\mathbf{c}^{(w)}$  qui sont inconnues. Pour résoudre ce problème, une solution est de l'évaluer pour des proportions uniformes, soit  $c_u^{(w)} = 1/M$ . Réfrégier, Scotté et al. (2018) ont montré que lorsque  $P = M$  cette solution est équivalente à optimiser la moyenne de la borne de Cramér-Rao sur l'ensemble des concentrations normalisées.

Ce critère est spécifique à un jeu de filtres  $\mathbf{F}^{(w)}$  donné. Pour maximiser les performances d'un ensemble de  $W$  jeux de filtres en estimation des concentrations, une possibilité est de

minimiser la moyenne de la trace de la borne de Cramér-Rao sur cet ensemble, c'est-à-dire

$$\overline{\text{tr } \mathbf{Y}_c} = \frac{1}{W} \sum_{w=1}^W \text{tr } \mathbf{Y}_c^{(w)}.$$

### Critère de performance en reconstruction de spectres

Comme proposé dans le chapitre 2 nous employons la matrice de Cramér-Rao pour évaluer les performances d'un ensemble de jeux de filtres en reconstruction. Comme calculé en annexe B.1, la matrice d'information de Fisher associée à la reconstruction d'un spectre  $\mathbf{r}$  à partir de mesures de ce spectre réalisées à temps fixé avec les jeux de filtres  $\mathbf{F}^{(1)}, \dots, \mathbf{F}^{(W)}$  est la matrice  $\mathbf{I}$  telle que

$$[\mathbf{I}]_{k,l} = \sum_{w=1}^W \sum_{m=1}^P \tau \frac{f_m^{(w)}(k) f_m^{(w)}(l)}{\phi_m^{(w)}}$$

où

$$\phi_m^{(w)} = \mathbf{f}_m^{(w)} \cdot \mathbf{r}.$$

Notons que le spectre que nous considérons ici n'est pas normalisé. En effet, puisque les mesures sont réalisées à temps fixé les performances de reconstruction ne sont pas indépendantes de l'intensité du rayonnement observé. Pour plus de simplicité nous avons donc choisi d'inclure l'intensité dans le spectre plutôt que de traiter ce paramètre séparément.

Il est nécessaire d'employer une diversité de jeux de filtres afin de pouvoir inverser la matrice d'information de Fisher et donc utiliser ce critère. Plus précisément nous démontrons dans l'annexe B.2 la proposition qui suit.

**Proposition 5.** Soient des mesures d'un spectre possédant  $K$  composantes réalisées avec un ensemble de jeux de filtres  $\mathbf{F} = (\mathbf{F}^{(1)}, \dots, \mathbf{F}^{(W)})$  en arrêtant l'acquisition après un temps fixé à l'avance. Il est nécessaire et suffisant que

$$\text{rg}(\mathbf{F}) = K$$

pour que la matrice d'information de Fisher  $\mathbf{I}$  associée soit inversible.

**Corollaire 5.1.** Si  $W$  jeux de  $P$  filtres sont utilisés alors cela implique que

$$W \geq \left\lceil \frac{K}{P} \right\rceil.$$

Comme dans le chapitre précédent, nous utilisons comme critère de performance en reconstruction la trace de la borne de Cramér-Rao calculée pour le spectre plat, c'est-à-dire tel que  $r(k) = 1/K \forall k \in \llbracket 1, K \rrbracket$ . On note ce critère  $\text{tr } \mathbf{Y}$ .

L'hypothèse utilisée pour calculer ce critère est que les mesures sont effectuées sur des pixels de même composition. Sans hypothèse sur la variation spatiale des concentrations, il est incertain que l'on puisse trouver des pixels de ce type. Comme évoqué en introduction de cette section, il semble toutefois pertinent d'optimiser ce critère, non pas dans l'objectif de mettre en œuvre la reconstruction mais pour construire la diversité de jeux de filtres nécessaire à la détection. Nous vérifions la validité de cette approche dans la section suivante.

### 3.3.2 Algorithme de recherche des compromis optimaux

L'algorithme de recherche des compromis optimaux est le même que celui présenté dans la section 2.5.3, à deux exceptions près. La première est que l'on s'assure que les jeux de filtres soient de rang strictement supérieur à  $M$  afin de respecter la contrainte énoncée dans la proposition 3. La seconde est que le premier critère à optimiser est  $\overline{\text{tr } \Upsilon_c}$  plutôt que  $\overline{\mathcal{B}}$ . La combinaison convexe à minimiser est donc

$$\mathcal{C}(\mathbf{F}, \lambda) = (1 - \lambda) \text{tr } \Upsilon(\mathbf{F}) + \lambda \overline{\text{tr } \Upsilon_c(\mathbf{F})}.$$

On démontre dans l'annexe B.3 qu'il est comme auparavant possible d'utiliser le lemme d'inversion matriciel pour calculer rapidement la matrice de Cramér-Rao associée à la reconstruction de spectre au fil de la marche aléatoire réalisée par l'algorithme.

Les ensembles de jeux de filtres servant d'initialisations à l'algorithme sont également adaptés à ce nouveau cas, et le critère optimisé lors de leur construction est  $\text{tr } \Upsilon_c$  plutôt que  $\mathcal{B}$ . On s'assure en outre que  $\mathbf{I}_c$  soit inversible pour toutes les initialisations, en particulier pour celles constituées de jeux de filtres aléatoires.

Il serait très chronophage de tester à chaque itération si la contrainte correspondant à la proposition 4 est vérifiée. Toutefois, comme nous le verrons plus loin, cette contrainte est souvent vérifiée à l'issue de l'optimisation.

## 3.4 Exemple d'application

Afin d'illustrer les performances qu'il est possible d'obtenir avec cette approche, nous présentons ici sa mise en œuvre sur un exemple. On suppose que l'échantillon analysé est composé de mélanges de  $M = 3$  espèces chimiques de spectres connus. Les spectres de ces espèces ont été tirés uniformément sur le simplexe de  $\mathbb{R}^K$  (soit comme dans l'équation (2.14) avec  $\alpha = 1$ ) et apparaissent sur la figure 3.3.

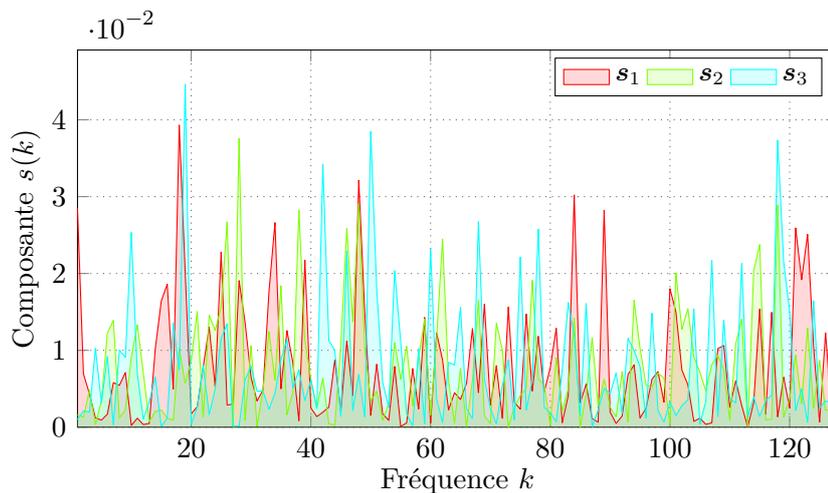


FIGURE 3.3 – Spectres de référence dont on veut estimer les concentrations.

### 3.4.1 Optimisation des jeux de filtres

On veut construire des jeux de filtres permettant de réaliser à la fois l'estimation des concentrations des espèces de référence en chaque pixel et la détection des anomalies. Suivant

l'approche proposée, on recherche des filtres à compromis optimaux entre estimation des concentrations et reconstruction de spectres. Comme évoqué dans la section précédente, il est nécessaire d'imposer des contraintes sur la forme des filtres afin de pouvoir détecter le plus d'anomalies possibles.

Pour ce placer dans le cadre de la proposition 3, il est nécessaire que  $\text{rg}(\mathbf{F}^{(w)}) > M \forall w$ . On choisit donc  $P = M + 1 = 4$ . Le corollaire de la proposition 4 indique qu'il faut  $W \geq \lceil (K - M)/(P - M) \rceil = 125$ . Cette condition est moins forte que celle de la proposition 5 qui impose seulement  $W \geq \lceil K/P \rceil = 32$ . Ici, nous avons choisi  $W = 144$ , ce qui vérifie ces deux conditions et permet d'effectuer un test d'inadéquation combinant des mesures effectuées avec tous les jeux de filtres construits sur toute fenêtre carrée de 12 pixels de côté si l'on répartit les jeux de filtres comme dans la figure 2.27.

La figure 3.4 montre la courbe des compromis optimaux correspondant, calculée avec un temps d'acquisition fixé à  $\tau = 1000$ . Ce choix arbitraire revient à faire l'acquisition de 1500 photons par pixel en moyenne si les concentrations des espèces de référence sont réparties uniformément entre 0 et 1. Sur cette courbe,  $\text{tr } \Upsilon \in [2.42 \times 10^{-4}, 1.61 \times 10^{-2}]$  et  $\overline{\text{tr } \Upsilon_c} \in [3.24 \times 10^{-3}, 9.67 \times 10^{-2}]$ .

Les points de la courbe des compromis optimaux apparaissant sous la forme de losanges verts vérifient la condition énoncée dans la proposition 4 et les points apparaissant sous la forme de triangles rouges ne la vérifient pas. Empiriquement, donner plus de poids à la reconstruction lors de l'optimisation favorise la diversité des jeux de filtres construits, et il n'est donc pas étonnant que les quelques points ne vérifiant pas cette condition soient proches de l'extrémité droite de la courbe.

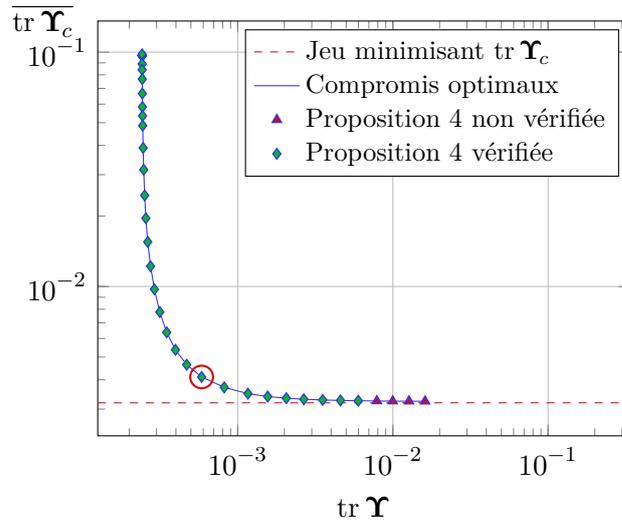


FIGURE 3.4 – Courbe des compromis optimaux entre estimation et reconstruction.

On peut une nouvelle fois remarquer qu'il est intéressant de choisir un point de fonctionnement proche du coude de la courbe des compromis optimaux plutôt qu'au niveau de ses extrémités. Cela permet en effet d'obtenir un gain de performances significatif en l'une des tâches tout en perdant peu en l'autre tâche par rapport aux points extrêmes.

La ligne en pointillés rouges représente la valeur minimale de  $\text{tr } \Upsilon_c$  de courbe, soit la meilleure borne de performance en estimation des concentrations qu'il est possible d'obtenir en un pixel donné en utilisant un des jeux de filtres construits par l'algorithme. Cette quantité vaut ici  $\text{tr } \Upsilon_c = 3.19 \times 10^{-3}$ . Ainsi, utiliser une diversité de jeux de filtres plutôt qu'un seul ne diminue pas significativement les performances en estimation des concentrations.

Pour les expériences de la suite de cette section, on choisit d'employer l'ensemble de jeux de filtres correspondant au point cerclé de rouge sur la figure 3.4. Cet ensemble de jeux de filtres situé au niveau du coude de la courbe des compromis optimaux vérifie les conditions des propositions 3 et 4 et permet d'obtenir de bonnes performances à la fois en estimation des concentrations et en reconstruction. Pour cet ensemble, on a en effet  $\text{tr } \overline{\Upsilon}_c = 4.11 \times 10^{-3}$  et  $\text{tr } \Upsilon = 5.86 \times 10^{-4}$ , ce qui ne correspond qu'à une multiplication de l'écart-type de l'estimation des concentrations par 1.13 et de l'écart-type de l'estimation de spectre par 1.56 par rapport aux points extrêmes de la courbe.

### 3.4.2 Performances en détection des jeux de filtres sélectionnés

Afin d'avoir un aperçu des performances en détection de l'ensemble de jeux de filtres sélectionné, nous avons estimé les probabilités de détection de différentes anomalies. Pour simuler des anomalies, nous avons généré aléatoirement 100 spectres  $\tilde{\mathbf{s}}_u$  distincts des spectres de référence. Chacun de ces spectres a été ajouté aux spectres de référence afin d'obtenir 100 jeux de spectres  $\tilde{\mathbf{S}}^{(u)} = (\mathbf{S}, \tilde{\mathbf{s}}_u)$  anormaux. Ces anomalies peuvent être très difficiles à détecter puisque les jeux de spectres correspondant contiennent la totalité des spectres de référence, auxquels n'a été ajouté qu'un seul spectre anormal.

Pour chaque  $\tilde{\mathbf{S}}^{(u)}$  nous avons tiré uniformément  $W$  vecteurs de concentrations et simulé les mesures  $\mathbf{n}$  des mélanges correspondant par les  $W$  jeux de filtres choisis. Nous avons ainsi simulé  $10^3$  réalisations de  $\mathbf{n}$  pour chaque  $\tilde{\mathbf{S}}^{(u)}$ . Pour chacune de ces réalisations il est possible d'effectuer le test d'inadéquation de l'équation (3.6). On peut alors estimer la probabilité de détection en fonction du seuil  $\mu$  du test. Nous avons fait de même avec  $10^3$  réalisations de  $\mathbf{n}$  avec un mélange des espèces de référence pour estimer la probabilité de fausse alarme.

Afin de montrer l'intérêt à utiliser une diversité de jeux de filtres, nous également réalisé cette expérience en simulant des mesures effectuées sur  $W$  pixels en appliquant le jeu de filtres minimisant  $\text{tr } \Upsilon_c$  représenté par une ligne en pointillés rouges sur la figure 3.4. Le rang de ce jeu de filtres est strictement supérieur à  $M$ , ce qui permet de se placer dans le cadre de la proposition 3.

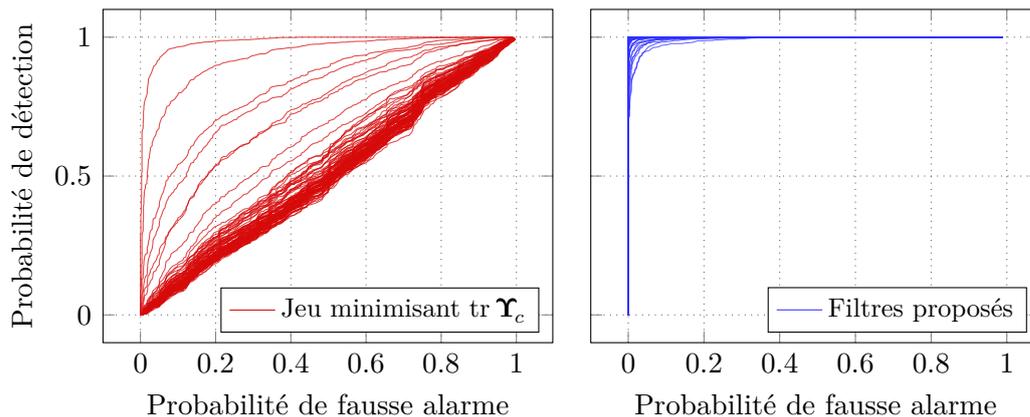


FIGURE 3.5 – Courbes ROC obtenues pour 100 spectres anormaux générés aléatoirement, lorsque l'on utilise un unique jeu de filtres (à gauche) où les filtres cerclés de rouge dans la figure 3.4 (à droite).

En faisant varier  $\mu$  on peut tracer les 100 courbes ROC (Receiver Operating Characteristic) correspondant aux différentes anomalies  $\tilde{\mathbf{S}}^{(u)}$  pour les deux ensembles de jeux de filtres. Ces courbes sont celles de l'évolution de la probabilité de détection en fonction de la probabilité de fausse alarme et apparaissent en figure 3.5. Elles montrent l'intérêt d'utiliser une diversité

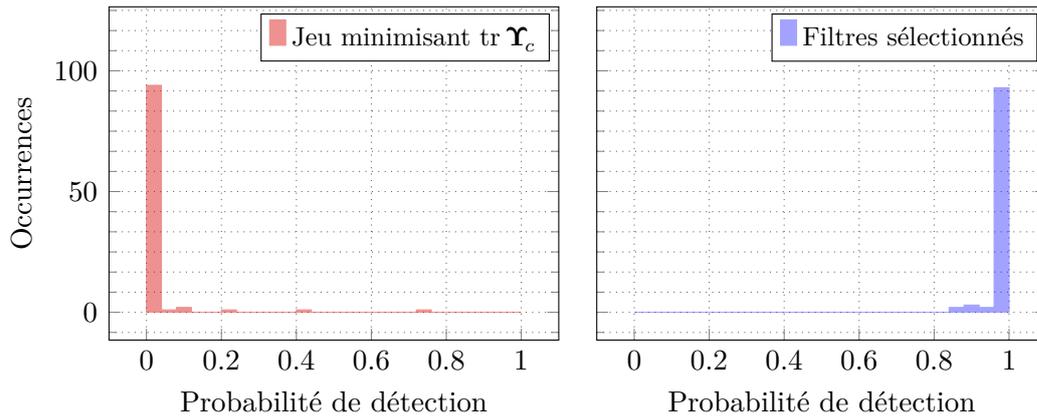


FIGURE 3.6 – Probabilités de détection pour une probabilité de fausse alarme de  $10^{-2}$ . Histogrammes pour 100 spectres anormaux générés aléatoirement, lorsque l'on utilise un unique jeu de filtres (à gauche) où les filtres cerclés de rouge dans la figure 3.4 (à droite).

de jeux de filtres : si ce n'est pas le cas (partie gauche) il existe un grand nombre d'anomalies pour lesquelles la probabilité de détection est équivalente à la probabilité de fausse alarme, alors qu'avec une diversité de jeux de filtres (partie droite) la probabilité de détection est très supérieure à la probabilité de fausse alarme.

Pour aller plus loin, la figure 3.6 montre les histogrammes des probabilités de détection des différentes anomalies quand la probabilité de fausse alarme est fixée à  $10^{-2}$ . On constate sur ces figures qu'employer une diversité de jeux de filtres permet d'améliorer les performances en détection. Plus précisément, la majeure partie des anomalies simulées ne peuvent être détectées avec de bonnes performances lorsque les mesures sont réalisées uniquement avec le jeu de filtres minimisant  $\text{tr } \Upsilon_c$ . Ces résultats peuvent être expliqués empiriquement par la spécialisation de ce jeu de filtres, qui focalise les mesures dans les directions correspondant aux spectres des espèces de référence et réduit par conséquent les chances de détecter une anomalie dont le spectre est quelconque.

Utiliser l'ensemble de jeux de filtres à compromis optimaux sélectionné permet donc de conjuguer estimation des concentrations et détection des anomalies.

### 3.4.3 Estimation robuste des concentrations

Nous illustrons à présent les performances de l'ensemble de jeux de filtres sélectionné sur une image simulée de l'échantillon pour lequel ces jeux de filtres ont été optimisés. On rappelle que cet échantillon est censé être composé des  $M = 3$  espèces de référence dont les spectres apparaissent en figure 3.3.

L'image simulée est un champ de vue carré de  $10^3$  pixels de côté. On y répartit les  $W = 144$  jeux de filtres choisis comme dans la figure 2.27, afin de disposer de la diversité de jeux de filtres nécessaire à se placer sous les conditions des propositions 3 et 4 dès que la zone testée comporte un carré de plus de  $\sqrt{W} = 12$  pixels de côté.

Sur ce champ de vue, nous avons ajouté une quatrième espèce imprévue en plus des espèces de référence. Celle-ci représente par exemple une pollution de l'échantillon analysé par un composé inconnu. Le spectre de cette espèce est celui de la figure 3.7. Les cartes des concentrations des différentes espèces sont présentées dans la figure 3.8. Ces cartes ont été générées aléatoirement avec une procédure employant un bruit de Perlin. L'espèce 4 n'est présente que sur environ 25% de l'image.

On simule alors les mesures réalisées avec le spectromètre Raman comprimé sur chaque

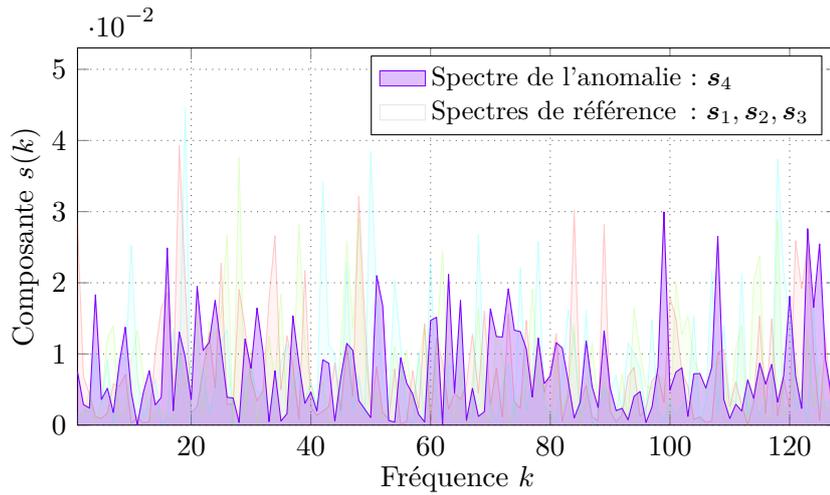


FIGURE 3.7 – Spectre de l'espèce imprévue.

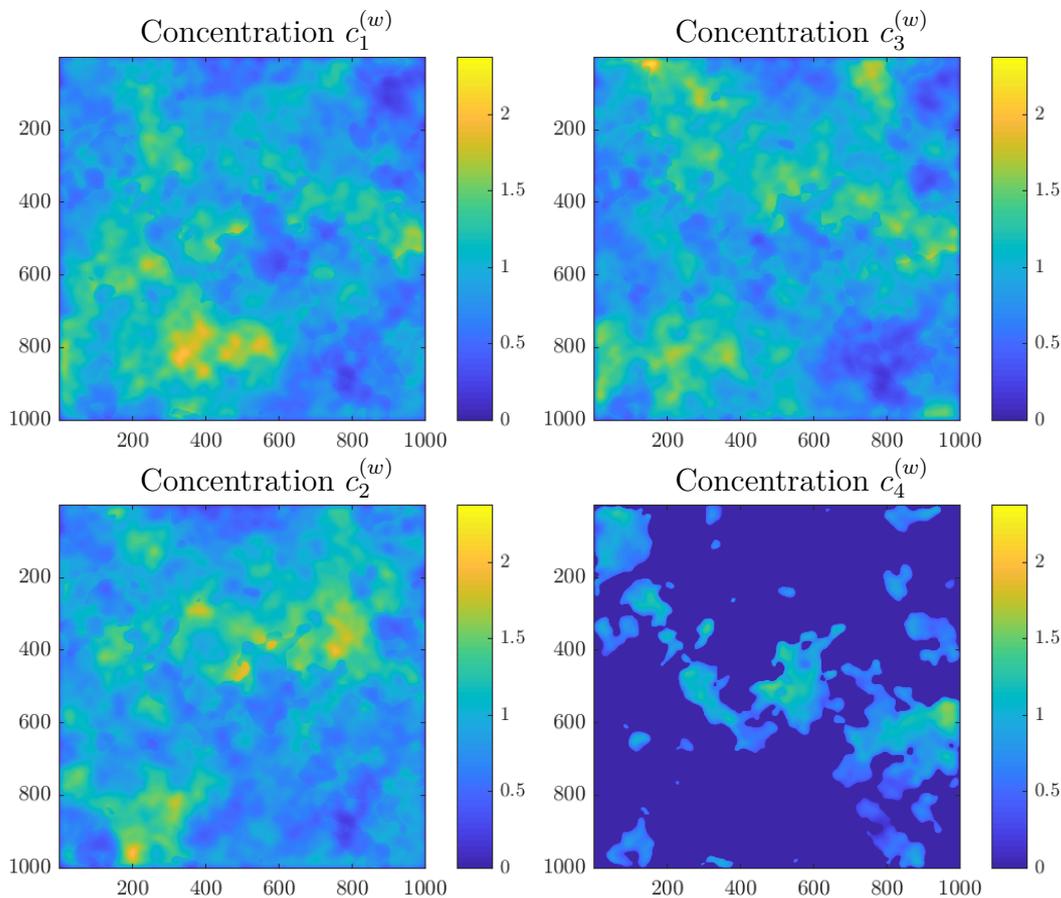


FIGURE 3.8 – Concentrations des 4 espèces composant l'échantillon simulé.

pixel du champ de vue avec le jeu de filtres  $y$  étant assigné, et on estime en chacun de ces pixels les concentrations des espèces de référence avec l'estimateur du maximum de vraisemblance  $\hat{\mathbf{c}}^{(w)}$  décrit dans la section précédente. Le résultat de cette estimation ainsi que la somme des erreurs quadratiques entre les concentrations estimées et les concentrations vraies en chaque pixel  $\|\mathbf{c}^{(w)} - \hat{\mathbf{c}}^{(w)}\|_2 = \sqrt{\sum_{u=1}^M (c_u^{(w)} - \hat{c}_u^{(w)})^2}$  sont affichés en figure 3.9. On affiche également en figure 3.10 l'histogramme de  $\sqrt{\text{tr } \Upsilon_c}$  (à gauche), celui des erreurs sur les zones en adéquation avec le modèle (au milieu) et celui des erreurs sur les zones anormales

(à droite). Sur les zones en adéquation avec le modèle, la somme des erreurs quadratiques est comprise entre  $3.77 \times 10^{-4}$  et 0.51, et vaut en moyenne  $9.96 \times 10^{-2}$ , ce qui est proche de l'erreur minimale théorique pouvant être calculée à partir de la borne de Cramér-Rao. L'erreur est plus importante sur les zones où l'espèce inconnue est présente, y étant comprise entre  $6.63 \times 10^{-3}$  et 1.11 avec une moyenne de 0.431. Cependant, il est difficile de voir les anomalies sur la carte des concentrations estimées de la figure 3.9.

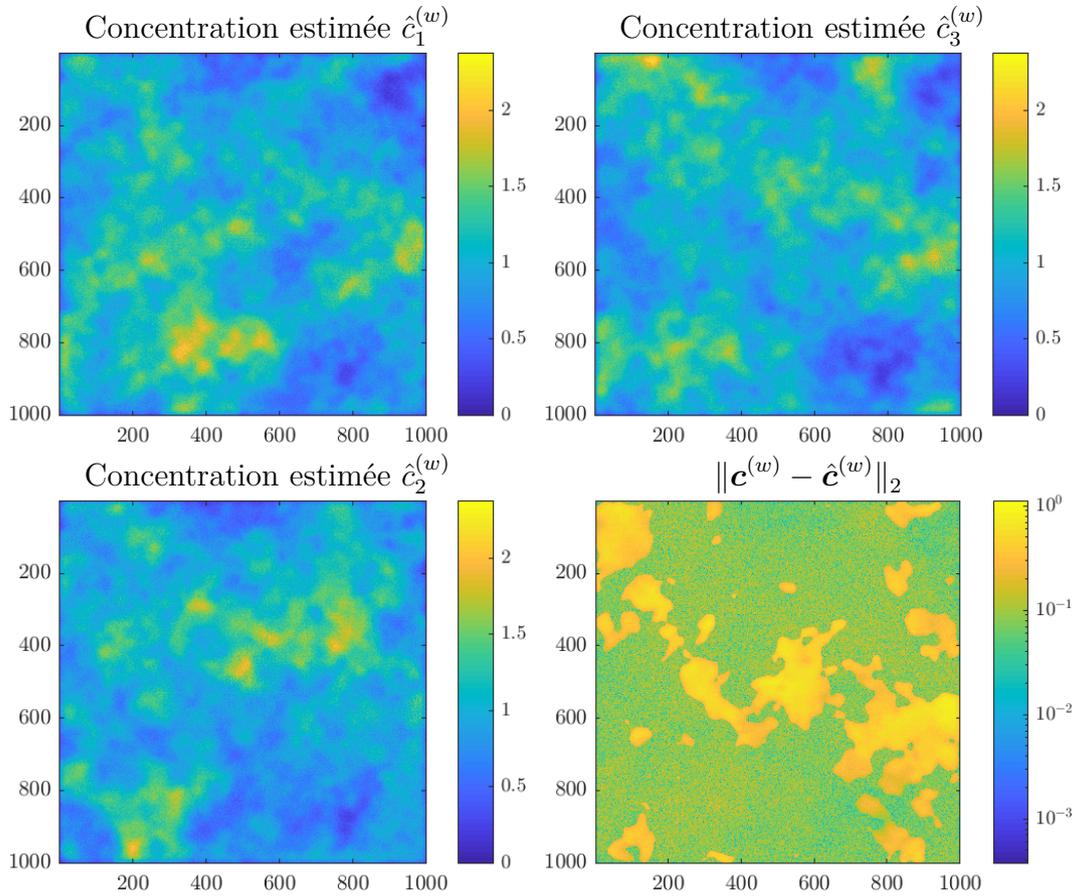


FIGURE 3.9 – Concentrations estimées et somme des erreurs quadratiques.

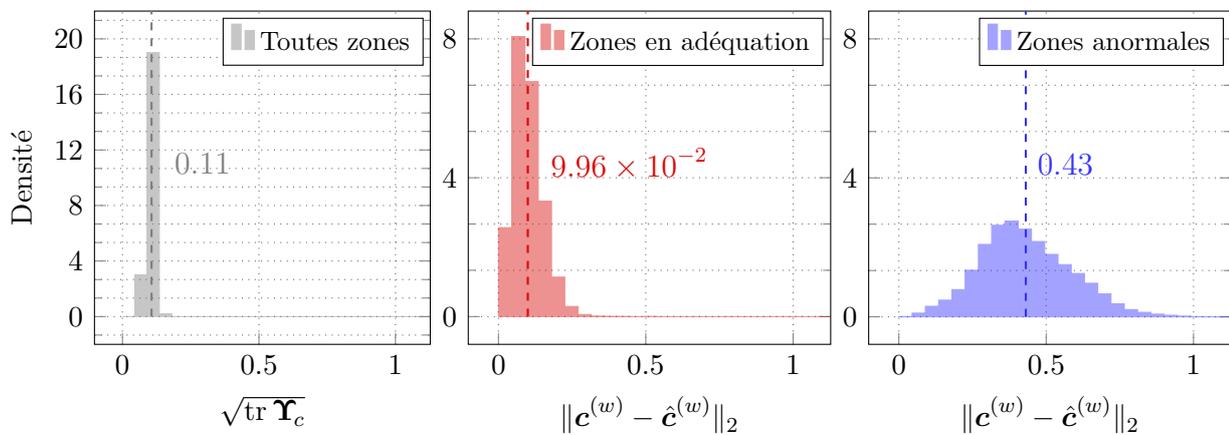


FIGURE 3.10 – Histogrammes de la racine de la trace de la borne de Cramér-Rao et de la somme des erreurs quadratiques des concentrations estimées.

Afin de détecter les anomalies, nous effectuons le test d'inadéquation proposé dans la

section 3.2.4 sur des fenêtres glissantes carrées de côté  $\sqrt{W} = 12$  pixels avec un seuil calculé pour que la probabilité de fausse alarme soit de  $10^{-2}$ . Le résultat du test est affecté au pixel central de la fenêtre. La figure 3.11 compare la carte des concentrations en espèce inconnue  $c_4^{(w)}$  (en haut à gauche), la carte des valeurs de  $\hat{T}(\mathbf{n})$  correspondant à chaque fenêtre (en haut à droite), le résultat du test en chaque pixel (en bas à gauche, rouge si le test a été déclenché, vert sinon) et l'évaluation de la qualité du test (en bas à droite, vert si en accord avec la réalité, rouge si l'anomalie n'a pas été détectée, bleu en cas de fausse alarme). Ici, la probabilité de détection empirique de l'anomalie est de 89.3% et la probabilité de fausse alarme empirique est de 1.8%. Cette probabilité de fausse alarme est plus élevée que la prédiction du fait des erreurs commises sur les bords des zones anormales.

À titre de comparaison, la probabilité de détection empirique obtenue sur cet échantillon en fixant un seuil correspondant à une probabilité de fausse alarme de  $10^{-2}$  et en réalisant les tests sur des fenêtres de taille  $12 \times 12$  en utilisant systématiquement le jeu de filtres minimisant  $\text{tr } \Upsilon_c$  est de 2% et la probabilité de fausse alarme empirique est de 1%.

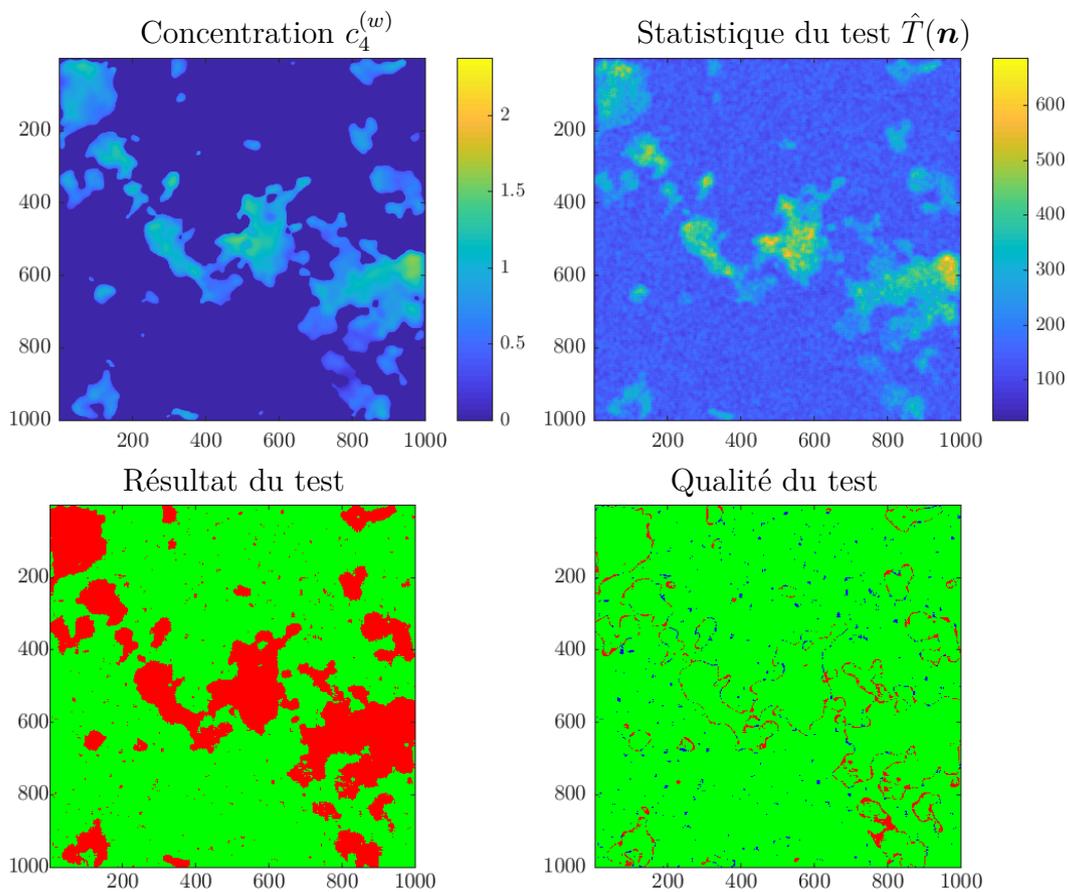
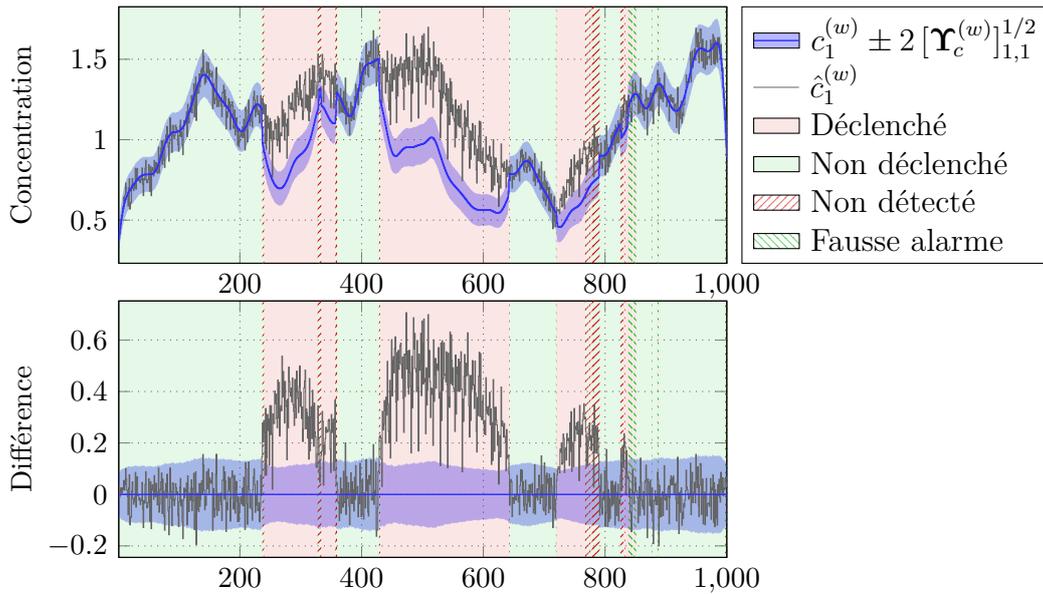


FIGURE 3.11 – Résultats du test d'inadéquation. Sur la carte en bas à gauche, les pixels apparaissent en rouge si le test a déclenché et en vert sinon. Sur la carte en bas à droite, les pixels apparaissent en vert si la bonne décision a été prise, en rouge en cas de non-détection et en bleu en cas de fausse alarme.

Pour mieux visualiser les performances en estimation des concentrations et en détection d'anomalies, la figure 3.12 compare le long de la 500<sup>ème</sup> ligne de l'image la concentration de la première espèce de référence  $c_1^{(w)}$  avec son estimation  $\hat{c}_1^{(w)}$ . La borne de Cramér-Rao affichée a été calculée à partir des concentrations vraies. Sur la partie basse de l'image, on affiche la différence entre l'estimation et la concentration vraie. Dans les zones en adéquation, les estimations  $\hat{c}_1^{(w)}$  sont pour la plupart dans l'intervalle  $c_1^{(w)} \pm 2 [\Upsilon_c^{(w)}]_{1,1}^{1/2}$ . En revanche,


 FIGURE 3.12 – Estimation des concentrations et détection sur la 500<sup>ème</sup> ligne de l'image.

on observe un fort biais sur  $\hat{c}_1^{(w)}$  dans les zones anormales. Dans ces zones l'estimation des concentrations est inexploitable, et il est essentiel de le détecter.

Nous avons également affiché sur cette figure les résultats du test d'inadéquation : le fond est coloré de rouge si le test a déclenché et de vert sinon. Les zones hachurées en rouge et en vert représentent respectivement les zones où l'anomalie n'a pas été détectée et celles où il y a une fausse alarme. Il est difficile de détecter l'anomalie simplement à partir de la courbe de  $\hat{c}_1^{(w)}$ , mais le test d'inadéquation permet de détecter la majeure partie des zones où le modèle est inexact, zones dans lesquelles l'estimation des concentrations en espèces de référence est erronée. Cette détection est effectuée sans faire appel à des hypothèses sur les concentrations.

Cette expérience illustre l'intérêt de l'approche proposée. En construisant des filtres à compromis optimaux entre estimation des concentrations et reconstruction de spectres vérifiant les contraintes nécessaires à se placer dans les conditions des propositions 3 et 4, il devient non seulement possible d'estimer les concentrations des espèces de référence mais également de détecter les anomalies et donc de déterminer quand avoir confiance en l'estimation.

### 3.5 Conclusion et ouverture

Dans ce chapitre, nous avons étudié comment estimer les concentrations des espèces de référence dans un échantillon de manière robuste, c'est-à-dire en étant capable de détecter la présence d'espèces chimiques imprévues qui fausseraient les concentrations estimées.

Nous avons étudié les contraintes qu'il est nécessaire d'imposer sur le filtrage pour avoir la possibilité de détecter les anomalies. Un résultat important est qu'il faut appliquer strictement plus de filtres linéairement indépendants que d'espèces de référence en chaque pixel afin de ne pas être aveugle à une part non négligeable des anomalies. Nous avons également montré qu'utiliser une diversité de jeux de filtres permet de limiter la part d'espèces anormales systématiquement indiscernables des espèces de référence, c'est-à-dire peu importe les concentrations avec lesquelles elles apparaissent dans l'échantillon analysé.

Pour construire une telle diversité de jeux de filtres, nous avons étendu l'approche à compromis optimaux proposée dans le chapitre précédent afin de construire des ensembles de

jeux de filtres réalisant des compromis entre estimation des concentrations des espèces de référence en chaque pixel et reconstruction de spectres.

En choisissant un ensemble de jeux de filtres à compromis optimaux vérifiant les contraintes établies, nous avons été capables de détecter sur un exemple la présence d'un composé inconnu mélangé avec les espèces de référence, permettant ainsi de déterminer dans quelles zones les résultats de l'estimation des concentrations étaient inexploitable.

Une perspective intéressante à ce travail serait de pouvoir reconstruire les spectres des espèces anormales à partir des mesures comprimées. Résoudre ce problème permettrait de déterminer l'origine des anomalies détectées, donnant à l'utilisateur la possibilité de pousser plus loin l'analyse du problème. Nous reviendrons sur cette perspective dans notre conclusion générale.

# Conclusion et perspectives

L'objectif de cette thèse était de développer des méthodes permettant de réaliser de manière robuste et à partir de peu de photons la classification de mesures de spectroscopie Raman comprimée parmi des espèces chimiques de spectres connus ou l'estimation des concentrations d'espèces chimiques de spectres connus dans un mélange à partir de ce même type de mesures. Réaliser ces tâches de manière robuste signifie ici pouvoir tester la validité des hypothèses effectuées lors de leur réalisation et déterminer l'origine du problème lorsqu'elles ne sont pas valides, si cela est possible.

Le développement de la méthodologie que nous avons proposée est d'abord passé par l'analyse de l'intérêt mais aussi des problèmes de robustesse des méthodes comprimées proposées dans la littérature. Ces problèmes ont été mis en évidence grâce à des simulations et leurs fondements ont été analysés de manière théorique.

Une fois ces problèmes exposés, l'enjeu était de déterminer comment construire un filtrage en fréquence permettant à la fois de réaliser la tâche considérée avec de bonnes performances lorsque les hypothèses sur les espèces chimiques composant l'échantillon analysé sont vérifiées mais également de pouvoir détecter et caractériser les anomalies rencontrées. La méthodologie que nous avons développée s'appuie d'abord sur des bornes statistiques pour caractériser les performances des filtres en ces deux tâches antagonistes puis sur une stratégie fondée sur des compromis optimaux pour construire des filtres binaires remplissant les objectifs établis.

Dans cette conclusion, nous revenons plus en détails sur les principaux résultats obtenus au cours de cette thèse et proposons plusieurs perspectives à ces travaux.

## Principaux résultats

Le premier problème que nous avons traité concerne la classification des mesures comprimées parmi des espèces chimiques de spectres connus, que nous appelons espèces de référence dans ce manuscrit. Le traitement de ce problème correspond au chapitre 2.

Afin de réaliser cette tâche avec les meilleures performances possibles, les approches de la littérature suggèrent d'appliquer un même jeu de filtres sur tous les pixels du champ de vue, ce jeu de filtres étant préalablement optimisé pour la classification des mesures parmi les espèces de référence. Il est notamment possible pour ce faire de minimiser la borne de Bhattacharyya, qui majore la probabilité d'erreur du classifieur optimal.

Comme nous l'avons mis en évidence en simulation, utiliser systématiquement un même jeu de filtres spécialisé de la sorte ne permet en général pas de détecter la présence d'une espèce chimique imprévue. Une telle limitation rend cette technique inadaptée au contrôle de la composition chimique d'un échantillon, au sens où un échantillon pollué apparaîtra malgré tout la plupart du temps à l'observateur comme étant conforme aux hypothèses effectuées sur les espèces en présence. Il paraît donc nécessaire de changer de stratégie d'acquisition.

Constatant l'existence d'une grande diversité de jeux de filtres ayant de bonnes performances en classification, nous avons proposé de changer de jeu de filtres en chaque pixel afin d'être à la fois capable de classer chaque pixel parmi les espèces de référence avec de

bonnes performances mais également de pouvoir reconstruire un spectre en combinant des mesures réalisées sur plusieurs pixels. Une telle reconstruction donne la possibilité de détecter la présence d'une anomalie rencontrée sur plusieurs pixels et de l'identifier sans effectuer de mesure supplémentaire.

Afin d'évaluer les performances d'un ensemble de jeux de filtres en reconstruction de spectres, nous avons proposé d'employer un critère fondé sur la borne de Cramér-Rao associée à ce problème d'inférence. Cette borne minore la variance de tout estimateur non biaisé du spectre, et la minimiser permet d'optimiser les performances des filtres en reconstruction. La classification et la reconstruction de spectres étant antagonistes, la question qui se posait était alors : comment construire un filtrage permettant de concilier ces deux tâches ?

Pour résoudre ce problème, nous avons proposé d'utiliser une approche dite à compromis optimaux. Cette approche permet de construire des ensembles de jeux de filtres optimaux au sens où il n'existe pas d'ensemble de jeux de filtres meilleur à la fois en classification et en reconstruction. Il est ainsi possible de construire une diversité de points de fonctionnement pour le système, chacun réalisant à un compromis différent entre les deux tâches. L'utilisateur peut choisir parmi ces points de fonctionnement celui qui correspond le mieux à ses besoins.

L'approche à compromis optimaux permet également d'obtenir un aperçu complet des performances qu'il est possible d'atteindre pour des espèces chimiques de référence et une configuration du système d'acquisition donnés. Cela nous a permis d'analyser l'influence de différents paramètres clés sur les performances, tels que le nombre d'espèces de référence, la nature de leurs spectres et le nombre de mesures effectuées en chaque pixel. Il ressort des exemples que nous avons traités qu'il existe dans de nombreuses situations des ensembles de jeux de filtres ayant des performances proches des meilleures performances atteignables à la fois en classification et en reconstruction de spectres. Cela signifie qu'il est en général possible d'implémenter notre approche avec une faible perte de performance en classification par rapport aux approches de la littérature.

Dans le chapitre 3, nous avons étudié le problème plus général de l'estimation des concentrations des espèces de référence dans un échantillon que l'on suppose être normalement composé uniquement de ces espèces. À l'instar de la classification, les méthodes de la littérature proposent d'utiliser un même jeu de filtres spécialisé en cette tâche sur tous les pixels de la zone à imager. Ce jeu de filtres peut notamment être optimisé en minimisant la trace de la borne de Cramér-Rao de l'estimation des concentrations.

La présence de multiples espèces chimiques ayant des concentrations inconnues à chaque pixel de l'échantillon rend la détection d'anomalies plus difficile que dans le cas traité dans le second chapitre. Nous avons montré qu'il est nécessaire d'imposer des contraintes sur la forme et sur la diversité des jeux de filtres afin d'avoir la possibilité de détecter le plus d'anomalies possibles. Il est en particulier nécessaire d'appliquer en chaque pixel strictement plus de filtres linéairement indépendants que d'espèces de référence, sans quoi il est impossible de détecter une part non négligeable des anomalies. Utiliser une diversité de jeux de filtres permet également de réduire la part d'anomalies systématiquement indétectables, c'est-à-dire peu importe les concentrations des espèces chimiques correspondant dans l'échantillon.

Dans la continuité du second chapitre, nous avons adapté l'approche à compromis optimaux afin de construire des ensembles de jeux de filtres ayant de bonnes performances en estimation des concentrations et possédant la diversité nécessaire à la détection des éventuelles anomalies.

Pour détecter les anomalies à partir des mesures réalisées avec les jeux de filtres ainsi construits, nous avons mis en place un test d'inadéquation. Ce test, dont on peut fixer la probabilité de fausse alarme, nous a permis de mettre en œuvre l'approche proposée sur un exemple simulé et d'en valider la pertinence.

## Perspectives

La perspective peut être la plus évidente à ce travail est l'implémentation de la méthodologie proposée sur un système réel. Cette implémentation nécessite d'adapter les dispositifs comprimés existants afin qu'ils soient capables d'appliquer une diversité de jeux de filtres sur la zone analysée. La mise en place d'un tel système est en cours dans l'équipe Mosaic de l'Institut Fresnel.

Nous avons brièvement évoqué dans la dernière partie du premier chapitre la possibilité de réaliser de multiples passages par pixel et avons souligné que cela permet d'améliorer significativement les performances. Il serait intéressant de pousser plus loin l'analyse de cette possibilité car elle pourrait permettre de compenser les limitations des spectromètres comprimés actuels, qui ne peuvent en général pas appliquer plus de deux filtres en parallèle.

Une hypothèse forte des approches comprimées sur lesquelles se base ce travail est la connaissance parfaite des spectres des espèces chimiques de référence. Si les spectres dont on dispose diffèrent des spectres véritables, les résultats de la classification ou de l'estimation des concentrations seront entachés d'erreurs. Il est possible que ces inadéquations puissent être détectées en appliquant la méthodologie proposée dans ce manuscrit. Toutefois, une meilleure solution serait de prendre en compte l'incertitude sur la forme des spectres de référence lors la classification des mesures ou de l'estimation des concentrations, mais aussi lors de l'optimisation des filtres. Une telle généralisation du modèle pourrait toutefois aboutir à une forte complexification de l'analyse théorique, dépendamment de l'incertitude modélisée.

Dans ce travail, nous avons montré qu'il est pertinent d'utiliser une diversité de jeux de filtres pour réaliser les mesures. Pour construire de tels jeux de filtres, nous avons optimisé les performances du système d'acquisition en reconstruction de spectres, en supposant que la reconstruction était réalisée à partir de mesures effectuées sur des pixels de composition homogène. Cette hypothèse d'homogénéité permet d'obtenir un critère de performance simple pour optimiser les filtres. En pratique, utiliser une diversité de jeux de filtres peut donner la possibilité de mettre en œuvre la reconstruction même lorsque l'on ne dispose pas de zones strictement homogènes. Il est par exemple possible de faire des hypothèses sur la variation spatiale de la composition chimique de l'échantillon pour régulariser le problème. Plus généralement, utiliser des approches parcimonieuses pourrait s'avérer particulièrement intéressant pour remonter a posteriori aux spectres contenus dans l'image. On pourrait également coupler la méthodologie proposée à des techniques de factorisation de matrices non négatives pour reconstruire les spectres des composés chimiques anormaux même dans le cas de mélanges. Employer ces techniques permettrait de donner à l'utilisateur la possibilité de remonter à l'origine des anomalies détectées dans une plus grande variété de situations.

Enfin, il serait particulièrement intéressant d'exporter les méthodes développées dans cette thèse vers d'autres applications que la spectroscopie Raman.



# Annexes

## Sommaire

---

|          |  |            |
|----------|--|------------|
| <b>A</b> | <b>Borne de Cramér-Rao de la reconstruction d'un spectre dans le cas d'une distribution multinomiale . . . . .</b> | <b>92</b>  |
| A.1      | Calcul de l'information de Fisher . . . . .  | 92         |
| A.2      | Inversibilité de la matrice d'information de Fisher . . . . .  | 93         |
| A.3      | Formule de Sherman et calcul de la matrice de Cramér-Rao . . . . .   | 96         |
| <b>B</b> | <b>Borne de Cramér-Rao de la reconstruction du spectre dans le cas d'un bruit de Poisson . . . . .</b>             | <b>99</b>  |
| B.1      | Calcul de l'information de Fisher . . . . .  | 99         |
| B.2      | Inversibilité de la matrice d'information de Fisher . . . . .  | 99         |
| B.3      | Formule de Sherman et calcul de la matrice de Cramér-Rao . . . . .   | 100        |
| <b>C</b> | <b>Borne de Bhattacharyya . . . . .</b>  | <b>102</b> |
| C.1      | Borne de Bhattacharyya à plusieurs jeux de filtres . . . . .   | 102        |
| C.2      | Borne de Bhattacharyya généralisée à plusieurs jeux de filtres . . . . .   | 102        |
| <b>D</b> | <b>Modification aléatoire de jeux de filtres . . . . .</b>   | <b>103</b> |
| D.1      | Modifications possibles en une itération . . . . .   | 103        |
| D.2      | Tirage de la localisation de la modification . . . . .   | 103        |
| <b>E</b> | <b>Convexité de la courbe des compromis optimaux . . . . .</b>   | <b>104</b> |
| <b>F</b> | <b>Conditions nécessaires à détecter les anomalies . . . . .</b>   | <b>105</b> |
| F.1      | Notations spécifiques à la section . . . . .   | 105        |
| F.2      | Décomposition des spectres normés . . . . .  | 105        |
| F.3      | Démonstration de la proposition 3 . . . . .  | 106        |
| F.4      | Démonstration de la proposition 4 . . . . .  | 109        |

---

# A Borne de Cramér-Rao de la reconstruction d'un spectre dans le cas d'une distribution multinomiale

## A.1 Calcul de l'information de Fisher

Soient  $\mathbf{n}^{(1)}, \dots, \mathbf{n}^{(W)}$  des mesures d'un rayonnement ayant un spectre normalisé  $\mathbf{s}$  effectuées avec les jeux de filtres  $\mathbf{F}^{(1)}, \dots, \mathbf{F}^{(W)}$  en arrêtant l'acquisition après la détection de  $N$  photons pour chaque jeu de filtres. On pose  $\mathbf{n}$  la transposée de  $(\mathbf{t}\mathbf{n}^{(1)}, \dots, \mathbf{t}\mathbf{n}^{(W)})$ . On veut calculer la borne de Cramér-Rao associée à l'estimation du spectre  $\mathbf{s}$  à partir de ces mesures. On note  $\ell(\mathbf{n} | \mathbf{s})$  la log-vraisemblance. Le vecteur  $\mathbf{n}$  ayant la loi de l'équation (2.3) on a

$$\ell(\mathbf{n} | \mathbf{s}) = \sum_{w=1}^W \sum_{m=1}^P \log N! - \log n_m^{(w)}! + n_m^{(w)} \log p_m^{(w)},$$

avec

$$p_m^{(w)} = \frac{\sum_{k=1}^K f_m^{(w)}(k) s(k)}{\sum_{n=1}^P \sum_{k=1}^K f_n^{(w)}(k) s(k)} = \frac{\mathbf{f}_m^{(w)} \cdot \mathbf{s}}{\sum_{n=1}^P \mathbf{f}_n^{(w)} \cdot \mathbf{s}}.$$

Puisque  $s(K) = 1 - \sum_{k=1}^{K-1} s(k)$  on a

$$\partial_{s(k)} \mathbf{f}_m^{(w)} \cdot \mathbf{s} = \partial_{s(k)} \sum_{k=1}^K f_m^{(w)}(k) s(k) = f_m^{(w)}(k) - f_m^{(w)}(K) = \delta f_m^{(w)}(k).$$

On en déduit

$$\begin{aligned} \partial_{s(k)} \ell(\mathbf{n} | \mathbf{s}) &= \sum_{w=1}^W \sum_{m=1}^P \partial_{s(k)} \left[ n_m^{(w)} \log p_m^{(w)} \right] \\ &= \sum_{w=1}^W \sum_{m=1}^P n_m^{(w)} \left[ \partial_{s(k)} \log \mathbf{f}_m^{(w)} \cdot \mathbf{s} - \partial_{s(k)} \log \sum_{n=1}^P \mathbf{f}_n^{(w)} \cdot \mathbf{s} \right] \\ &= \sum_{w=1}^W \sum_{m=1}^P n_m^{(w)} \partial_{s(k)} \log \mathbf{f}_m^{(w)} \cdot \mathbf{s} - \sum_{w=1}^W N \partial_{s(k)} \log \sum_{n=1}^P \mathbf{f}_n^{(w)} \cdot \mathbf{s} \\ &= \sum_{w=1}^W \sum_{m=1}^P n_m^{(w)} \frac{\delta f_m^{(w)}(k)}{\mathbf{f}_m^{(w)} \cdot \mathbf{s}} - N \sum_{w=1}^W \frac{\sum_{n=1}^P \delta f_n^{(w)}(k)}{\sum_{n=1}^P \mathbf{f}_n^{(w)} \cdot \mathbf{s}} \end{aligned}$$

puis

$$\partial_{s(k)} \partial_{s(l)} \ell(\mathbf{n} | \mathbf{s}) = - \sum_{w=1}^W \sum_{m=1}^P n_m^{(w)} \frac{\delta f_m^{(w)}(k) \delta f_m^{(w)}(l)}{(\mathbf{f}_m^{(w)} \cdot \mathbf{s})^2} + N \sum_{w=1}^W \frac{\sum_{n=1}^P \delta f_n^{(w)}(k) \sum_{n=1}^P \delta f_n^{(w)}(l)}{(\sum_{n=1}^P \mathbf{f}_n^{(w)} \cdot \mathbf{s})^2}$$

et

$$\begin{aligned} [\mathbf{I}]_{k,l} &= - \langle \partial_{s(k)} \partial_{s(l)} \ell(\mathbf{n} | \mathbf{s}) \rangle \\ &= \sum_{w=1}^W N \left[ \sum_{m=1}^P p_m^{(w)} \frac{\delta f_m^{(w)}(k) \delta f_m^{(w)}(l)}{(\mathbf{f}_m^{(w)} \cdot \mathbf{s})^2} - \frac{\sum_{n=1}^P \delta f_n^{(w)}(k) \sum_{n=1}^P \delta f_n^{(w)}(l)}{(\sum_{n=1}^P \mathbf{f}_n^{(w)} \cdot \mathbf{s})^2} \right] \end{aligned}$$

$$\begin{aligned}
&= \sum_{w=1}^W N \left[ \sum_{m=1}^P \frac{\mathbf{f}_m^{(w)} \cdot \mathbf{s}}{\sum_{n=1}^P \mathbf{f}_n^{(w)} \cdot \mathbf{s}} \frac{\delta f_m^{(w)}(k) \delta f_m^{(w)}(l)}{(\mathbf{f}_m^{(w)} \cdot \mathbf{s})^2} - \frac{\sum_{n=1}^P \delta f_n^{(w)}(k) \sum_{n=1}^P \delta f_n^{(w)}(l)}{(\sum_{n=1}^P \mathbf{f}_n^{(w)} \cdot \mathbf{s})^2} \right] \\
&= \sum_{w=1}^W \frac{N}{\sum_{n=1}^P \mathbf{f}_n^{(w)} \cdot \mathbf{s}} \left[ \sum_{m=1}^P \frac{\delta f_m^{(w)}(k) \delta f_m^{(w)}(l)}{\mathbf{f}_m^{(w)} \cdot \mathbf{s}} - \frac{\sum_{n=1}^P \delta f_n^{(w)}(k) \sum_{n=1}^P \delta f_n^{(w)}(l)}{\sum_{n=1}^P \mathbf{f}_n^{(w)} \cdot \mathbf{s}} \right]
\end{aligned}$$

où il est important de noter que le dernier terme de la partie droite de l'équation n'est pas dans le second niveau de sommation.

## A.2 Inversibilité de la matrice d'information de Fisher

On étudie dans cette sous-section les conditions sur les filtres sous lesquelles la matrice d'information de Fisher de la reconstruction de spectre dans le cas multinomial est inversible, et on démontre notamment les proposition 1 et 2 de la section 2.3.3. On commence pour cela par démontrer deux lemmes qui correspondent à l'écriture de la matrice d'information de Fisher sous deux formes différentes. Pour cela, on pose les notations suivantes :

| Grandeur                  | Domaine                           | Définition  |
|---------------------------|-----------------------------------|---|
| $\delta \mathbf{F}^{(w)}$ | $\{0, 1\}^{(K-1) \times P}$       | $[\delta \mathbf{F}^{(w)}]_{k,m} = \delta f_m^{(w)}(k)$     |
| $\delta \mathbf{F}$       | $\{-1, 0, 1\}^{(K-1) \times P W}$ | $(\delta \mathbf{F}^{(1)}, \dots, \delta \mathbf{F}^{(W)})$ |
| $g^{(w)}(k)$              | $\{0, 1\}$                        | $\sum_{m=1}^P f_m^{(w)}(k)$                                 |
| $\mathbf{g}^{(w)}$        | $\{0, 1\}^K$                      | $[\mathbf{g}^{(w)}]_k = g^{(w)}(k)$                         |
| $\mathbf{G}$              | $\{0, 1\}^{K \times W}$           | $(\mathbf{g}^{(1)}, \dots, \mathbf{g}^{(W)})$               |
| $\delta g^{(w)}(k)$       | $\{-1, 0, 1\}$                    | $g^{(w)}(k) - g^{(w)}(K)$                                   |
| $\delta \mathbf{g}^{(w)}$ | $\{-1, 0, 1\}^{(K-1)}$            | $[\delta \mathbf{g}^{(w)}]_k = \delta g^{(w)}(k)$           |
| $\delta \mathbf{G}$       | $\{-1, 0, 1\}^{(K-1) \times W}$   | $(\delta \mathbf{g}^{(1)}, \dots, \delta \mathbf{g}^{(W)})$ |

Avec ces notations on a

$$[\mathbf{I}]_{k,l} = \sum_{w=1}^W \frac{N}{\mathbf{g}^{(w)} \cdot \mathbf{s}} \left[ \sum_{m=1}^P \frac{\delta f_m^{(w)}(k) \delta f_m^{(w)}(l)}{\mathbf{f}_m^{(w)} \cdot \mathbf{s}} - \frac{\delta g^{(w)}(k) \delta g^{(w)}(l)}{\mathbf{g}^{(w)} \cdot \mathbf{s}} \right].$$

**Lemme 1** (Écriture de  $\mathbf{I}$  comme un produit de matrices). Soit  $\mathbf{Q}$  la matrice diagonale de côté  $PW$  dont les coefficients diagonaux sont

$$[\mathbf{Q}]_{i=j=P(w-1)+m} = \frac{N}{(\mathbf{f}_m^{(w)} \cdot \mathbf{s}) (\mathbf{g}^{(w)} \cdot \mathbf{s})}$$

et  $\mathbf{R}$  la matrice diagonale de côté  $W$  dont les coefficients diagonaux sont

$$[\mathbf{R}]_{i=j=w} = \frac{N}{(\mathbf{g}^{(w)} \cdot \mathbf{s})^2}$$

alors on a

$$\mathbf{I} = \delta \mathbf{F} \mathbf{Q} \delta \mathbf{F} - \delta \mathbf{G} \mathbf{R} \delta \mathbf{G} \tag{7}$$

soit

$$\mathbf{I} = (\delta \mathbf{F} \ \delta \mathbf{G}) \begin{pmatrix} \mathbf{Q} & (0) \\ (0) & -\mathbf{R} \end{pmatrix} \begin{pmatrix} \delta \mathbf{F} \\ \delta \mathbf{G} \end{pmatrix}. \tag{8}$$

Avec des filtres complémentaires,  $\delta g^{(w)}(k) = \sum_{m=1}^P f_m^{(w)}(k) - \sum_{m=1}^P f_m^{(w)}(K) = 0$  et on a

$$\mathbf{I} = \delta \mathbf{F} \mathbf{Q}^t \delta \mathbf{F}. \quad (9)$$

**Démonstration.** On développe le membre de droite de l'équation (7). On a

$$\begin{aligned} [\delta \mathbf{F} \mathbf{Q}]_{k, P(w-1)+m} &= \sum_i [\delta \mathbf{F}]_{k,i} [\mathbf{Q}]_{i, P(w-1)+m} \\ &= [\delta \mathbf{f}]_{k, P(w-1)+m} [\mathbf{Q}]_{P(w-1)+m, P(w-1)+m} \\ &= N \frac{\delta f_m^{(w)}(k)}{(\mathbf{f}_m^{(w)} \cdot \mathbf{s})(\mathbf{g}^{(w)} \cdot \mathbf{s})} \end{aligned}$$

soit

$$\begin{aligned} [\delta \mathbf{F} \mathbf{Q}^t \delta \mathbf{F}]_{k,l} &= \sum_i [\delta \mathbf{F} \mathbf{Q}]_{k,i} [\delta \mathbf{F}]_{i,l} \\ &= \sum_{w=1}^W \sum_{m=1}^P [\delta \mathbf{F} \mathbf{Q}]_{k, P(w-1)+m} [\delta \mathbf{F}]_{l, P(w-1)+m} \\ &= \sum_{w=1}^W \frac{N}{\mathbf{g}^{(w)} \cdot \mathbf{s}} \sum_{m=1}^P \frac{\delta f_m^{(w)}(k) \delta f_m^{(w)}(l)}{\mathbf{f}_m^{(w)} \cdot \mathbf{s}} \end{aligned}$$

et de plus

$$[\delta \mathbf{G} \mathbf{R}]_{k,w} = \sum_i [\delta \mathbf{G}]_{k,i} [\mathbf{R}]_{i,w} = [\delta \mathbf{F}]_{k,w} [\mathbf{R}]_{w,w} = N \frac{\delta g^{(w)}(k)}{(\mathbf{g}^{(w)} \cdot \mathbf{s})^2}$$

par conséquent

$$[\delta \mathbf{G} \mathbf{R}^t \delta \mathbf{G}]_{k,l} = \sum_{w=1}^W [\delta \mathbf{G} \mathbf{R}]_{k,w} [\delta \mathbf{G}]_{l,w} = \sum_{w=1}^W [\delta \mathbf{F} \mathbf{Q}]_{k, P(w-1)+m} [\delta \mathbf{F}]_{l, P(w-1)+m} = \sum_{w=1}^W N \frac{\delta g^{(w)}(k) \delta g^{(w)}(l)}{(\mathbf{g}^{(w)} \cdot \mathbf{s})^2}$$

ce qui permet de retrouver la formule (7), dont découlent (8) et (9). ■

**Lemme 2** (Seconde écriture de  $\mathbf{I}$  comme un produit de matrices). Soit  $\mathbf{H}$  la matrice de taille  $(K-1) \times (WP)$  telle que

$$[\mathbf{H}]_{k, P(w-1)+m} = \frac{N}{(\mathbf{f}_w \cdot \mathbf{s})} \left( \delta f_m^{(w)}(k) - (\mathbf{f}_m^{(w)} \cdot \mathbf{s}) \frac{\delta g^{(w)}(k)}{\mathbf{g}^{(w)} \cdot \mathbf{s}} \right) \quad (10)$$

et  $\mathbf{J}$  la matrice de taille  $(K-1) \times (WP)$  telle que

$$[\mathbf{J}]_{k, P(w-1)+m} = \frac{\delta f_m^{(w)}(k)}{\mathbf{f}_m^{(w)} \cdot \mathbf{s}} + \frac{\delta g^{(w)}(k)}{\mathbf{g}^{(w)} \cdot \mathbf{s}}$$

alors on a

$$\mathbf{I} = \mathbf{H}^t \mathbf{J}. \quad (11)$$

**Démonstration.**

$$\begin{aligned}
[\mathbf{H}^\dagger \mathbf{J}]_{k,l} &= \sum_{w=1}^W \sum_{m=1}^P [\mathbf{H}]_{k,P(w-1)+m} [\mathbf{J}]_{l,P(w-1)+m} \\
&= \sum_{w=1}^W \frac{N}{(\mathbf{g}^{(w)} \cdot \mathbf{s})} \sum_{m=1}^P \left( \delta f_m^{(w)}(k) - (\mathbf{f}_m^{(w)} \cdot \mathbf{s}) \frac{\delta g^{(w)}(k)}{\mathbf{g}^{(w)} \cdot \mathbf{s}} \right) \left( \frac{\delta f_m^{(w)}(l)}{\mathbf{f}_m^{(w)} \cdot \mathbf{s}} + \frac{\delta g^{(w)}(l)}{\mathbf{g}^{(w)} \cdot \mathbf{s}} \right) \\
&= \sum_{w=1}^W \frac{N}{(\mathbf{g}^{(w)} \cdot \mathbf{s})} \sum_{m=1}^P \left[ \frac{\delta f_m^{(w)}(k) \delta f_m^{(w)}(l)}{\mathbf{f}_m^{(w)} \cdot \mathbf{s}} - (\mathbf{f}_m^{(w)} \cdot \mathbf{s}) \frac{\delta g^{(w)}(k) \delta g^{(w)}(l)}{(\mathbf{g}^{(w)} \cdot \mathbf{s})^2} \right. \\
&\quad \left. + \delta f_m^{(w)}(k) \frac{\delta g^{(w)}(l)}{\mathbf{g}^{(w)} \cdot \mathbf{s}} - \delta f_m^{(w)}(l) \frac{\delta g^{(w)}(k)}{\mathbf{g}^{(w)} \cdot \mathbf{s}} \right] \\
&= \sum_{w=1}^W \frac{N}{(\mathbf{g}^{(w)} \cdot \mathbf{s})} \left[ \sum_{m=1}^P \frac{\delta f_m^{(w)}(k) \delta f_m^{(w)}(l)}{\mathbf{f}_m^{(w)} \cdot \mathbf{s}} - \sum_{m=1}^P (\mathbf{f}_m^{(w)} \cdot \mathbf{s}) \frac{\delta g^{(w)}(k) \delta g^{(w)}(l)}{(\mathbf{g}^{(w)} \cdot \mathbf{s})^2} \right. \\
&\quad \left. + \sum_{m=1}^P \delta f_m^{(w)}(k) \frac{\delta g^{(w)}(l)}{\mathbf{g}^{(w)} \cdot \mathbf{s}} - \sum_{m=1}^P \delta f_m^{(w)}(l) \frac{\delta g^{(w)}(k)}{\mathbf{g}^{(w)} \cdot \mathbf{s}} \right] \\
&= \sum_{w=1}^W \frac{N}{(\mathbf{g}^{(w)} \cdot \mathbf{s})} \left[ \sum_{m=1}^P \frac{\delta f_m^{(w)}(k) \delta f_m^{(w)}(l)}{\mathbf{f}_m^{(w)} \cdot \mathbf{s}} - \frac{\delta g^{(w)}(k) \delta g^{(w)}(l)}{\mathbf{g}^{(w)} \cdot \mathbf{s}} \right. \\
&\quad \left. + \sum_{m=1}^P \delta f_m^{(w)}(k) \frac{\delta g^{(w)}(l)}{\mathbf{g}^{(w)} \cdot \mathbf{s}} - \sum_{m=1}^P \delta f_m^{(w)}(l) \frac{\delta g^{(w)}(k)}{\mathbf{g}^{(w)} \cdot \mathbf{s}} \right] \\
&= [\mathbf{I}]_{k,l} + \sum_{w=1}^W \frac{N}{(\mathbf{g}^{(w)} \cdot \mathbf{s})} \left( \sum_{m=1}^P \delta f_m^{(w)}(k) \frac{\delta g^{(w)}(l)}{\mathbf{g}^{(w)} \cdot \mathbf{s}} - \sum_{m=1}^P \delta f_m^{(w)}(l) \frac{\delta g^{(w)}(k)}{\mathbf{g}^{(w)} \cdot \mathbf{s}} \right) \\
&= [\mathbf{I}]_{k,l} + \sum_{w=1}^W \frac{N}{(\mathbf{g}^{(w)} \cdot \mathbf{s})} \left( \frac{\delta g^{(w)}(k) \delta g^{(w)}(l)}{\mathbf{g}^{(w)} \cdot \mathbf{s}} - \frac{\delta g^{(w)}(l) \delta g^{(w)}(k)}{\mathbf{g}^{(w)} \cdot \mathbf{s}} \right) = [\mathbf{I}]_{k,l}.
\end{aligned}$$

■

**Démonstration de la proposition 1.** On peut facilement déduire du lemme 1 que

$$\mathbf{I} = (\delta \mathbf{F} \ \delta \mathbf{G}) \begin{pmatrix} \mathbf{Q} & (0) \\ (0) & -\mathbf{R} \end{pmatrix} \begin{pmatrix} {}^t \delta \mathbf{F} \\ {}^t \delta \mathbf{G} \end{pmatrix}.$$

Le rang d'un produit de matrices est inférieur ou égal au rang de chacun de ses termes. On a donc en particulier

$$\text{rg}(\mathbf{I}) \leq \text{rg}(\delta \mathbf{F} \ \delta \mathbf{G}).$$

Puisque  $\delta \mathbf{g}^{(w)} = \sum_{m=1}^P \delta \mathbf{f}_m^{(w)}$ , le rang de la matrice  $(\delta \mathbf{F} \ \delta \mathbf{G})$  est celui de  $\delta \mathbf{F}$ , et on a donc

$$\text{rg}(\mathbf{I}) \leq \text{rg}(\delta \mathbf{F})$$

ce qui signifie qu'il est nécessaire que  $\text{rg}(\delta \mathbf{F}) = K - 1$  pour que  $\mathbf{I}$  soit inversible.

Lorsque les filtres employés sont complémentaires, la matrice d'information de Fisher peut s'écrire sous la forme donnée par l'équation (9), c'est-à-dire

$$\mathbf{I} = \delta \mathbf{F} \mathbf{Q} {}^t \delta \mathbf{F},$$

avec  $[Q]_{i=j} = N/(\mathbf{f}_m^{(w)} \cdot \mathbf{s})$ . Puisque  $Q$  est diagonale et que ses coefficients diagonaux sont strictement positifs, on peut écrire de manière équivalente

$$\mathbf{I} = \mathbf{U}^t \mathbf{U}$$

avec  $\mathbf{U} = \delta \mathbf{F} \sqrt{Q}$ , et  $\mathbf{I}$  est la matrice de Gram des lignes de  $\mathbf{U}$ . Il est donc nécessaire et suffisant que  $\text{rg}(\mathbf{U}) = \text{rg}(\delta \mathbf{F}) = K - 1$  pour que la matrice d'information de Fisher soit inversible lorsque les filtres sont complémentaires. ■

**Démonstration de la proposition 2.** Le lemme 2 indique que peu importe la forme des filtres on peut écrire la matrice d'information de Fisher comme

$$\mathbf{I} = \mathbf{H}^t \mathbf{J}$$

avec  $\mathbf{H}$  la matrice de taille  $(K - 1) \times (WP)$  définie dans l'équation (10). Or, on a pour tout  $w$  et pour tout  $k$

$$\begin{aligned} \sum_{m=1}^P [\mathbf{H}]_{k, P(w-1)+m} &= \sum_{m=1}^P \frac{N}{(\mathbf{f}_w \cdot \mathbf{s})} \left( \delta f_m^{(w)}(k) - (\mathbf{f}_m^{(w)} \cdot \mathbf{s}) \frac{\delta g^{(w)}(k)}{\mathbf{g}^{(w)} \cdot \mathbf{s}} \right) \\ &= \frac{N}{(\mathbf{f}_w \cdot \mathbf{s})} \left( \sum_{m=1}^P \delta f_m^{(w)}(k) - \sum_{m=1}^P (\mathbf{f}_m^{(w)} \cdot \mathbf{s}) \frac{\delta g^{(w)}(k)}{\mathbf{g}^{(w)} \cdot \mathbf{s}} \right) \\ &= \frac{N}{(\mathbf{f}_w \cdot \mathbf{s})} \left( \delta g^{(w)}(k) - (\mathbf{g}^{(w)} \cdot \mathbf{s}) \frac{\delta g^{(w)}(k)}{\mathbf{g}^{(w)} \cdot \mathbf{s}} \right) \\ &= \frac{N}{(\mathbf{f}_w \cdot \mathbf{s})} \left( \delta g^{(w)}(k) - \delta g^{(w)}(k) \right) \\ &= 0 \end{aligned}$$

et la matrice  $\mathbf{H}$  ne comporte par conséquent au maximum que  $W(P - 1)$  colonnes linéairement indépendantes. On a en particulier

$$\text{rg}(\mathbf{I}) \leq \text{rg}(\mathbf{H}) \leq W(P - 1).$$

La matrice d'information de Fisher étant de taille  $K - 1$ , il est par conséquent nécessaire à son inversibilité que

$$W \geq \left\lceil \frac{K - 1}{P - 1} \right\rceil.$$

■

### A.3 Formule de Sherman et calcul de la matrice de Cramér-Rao

Les algorithmes présentés dans les sections 2.4.2 et 2.5.3 consistent à modifier aléatoirement des jeux de filtres un très grand nombre de fois, et demandent de calculer la trace de la borne de Cramér-Rao correspondant à la reconstruction d'un spectre à partir de mesures effectuées avec chaque ensemble de jeux de filtres résultant de ces modifications.

Nous montrons dans cette section que la modification effectuée à chaque itération de ces algorithmes ne résulte qu'en une perturbation de faible rang de la matrice d'information de Fisher. Il est par conséquent possible d'utiliser la formule de Sherman-Morrison-Woodbury pour calculer rapidement la matrice de Cramér-Rao.

Soit un ensemble de jeux de filtres  $\mathbf{F}^{(t)} = (\mathbf{F}^{(1,t)}, \dots, \mathbf{F}^{(W,t)})$ . Ces jeux de filtres sont modifiés selon la procédure décrite dans l'annexe D, à l'issue de laquelle on obtient  $\mathbf{F}^{(t)} = (\mathbf{F}^{(1,t+1)}, \dots, \mathbf{F}^{(W,t+1)})$ . Cet ensemble de jeux de filtres est identique au précédent, à l'exception des coefficients  $f_{m_0}^{(w_0,t+1)}(k_0)$  et  $f_{m_1}^{(w_0,t+1)}(k_0)$  qui diffèrent des coefficients  $f_{m_0}^{(w_0,t)}(k_0)$  et  $f_{m_1}^{(w_0,t)}(k_0)$ . On note  $\mathbf{I}^{(t)}$  et  $\mathbf{I}^{(t+1)}$  les matrices d'information de Fisher de  $\mathbf{F}^{(t)}$  et  $\mathbf{F}^{(t+1)}$  ainsi que  $\Upsilon^{(t)}$  et  $\Upsilon^{(t+1)}$  leurs matrices de Cramér-Rao. On note  $\Delta_t x^{(t)} = x^{(t+1)} - x^{(t)}$ . L'équation (2.4) donne

$$\Delta_t [\mathbf{I}^{(t)}]_{k,l} = [\mathbf{I}^{(t+1)}]_{k,l} - [\mathbf{I}^{(t)}]_{k,l} = \Delta_t \left[ \frac{N}{\mathbf{g}^{(w_0,t)} \cdot \mathbf{s}} \left( \sum_{n=1}^M \frac{\delta f_n^{(w_0,t)}(k) \delta f_n^{(w_0,t)}(l)}{\mathbf{f}_n^{(w_0,t)} \cdot \mathbf{s}} - \frac{\delta g^{(w_0,t)}(k) \delta g^{(w_0,t)}(l)}{\mathbf{g}^{(w_0,t)} \cdot \mathbf{s}} \right) \right].$$

On peut donc écrire

$$\mathbf{I}^{(t+1)} - \mathbf{I}^{(t)} = \mathbf{C} \mathbf{D}^t \mathbf{C}$$

avec la matrice  $\mathbf{C} = (\mathbf{C}^{(t+1)}, \mathbf{C}^{(t)})$  de taille  $(K-1) \times 2(P+1)$  telle que

$$\mathbf{C}^{(t)} = \left( \delta \mathbf{f}_1^{(w_0,t)}, \dots, \delta \mathbf{f}_P^{(w_0,t)}, \delta \mathbf{g}^{(w_0,t)} \right)$$

et  $\mathbf{D}$  la matrice diagonale de côté  $2(P+1)$  dont les coefficients diagonaux sont ceux du vecteur  $\mathbf{d} = (\mathbf{d}^{(t+1)}, -\mathbf{d}^{(t)})$  où  $\mathbf{d}^{(t)}$  est le vecteur ligne tel que

$$\mathbf{d}^{(t)} = \left( \frac{N}{(\mathbf{g}^{(w_0,t)} \cdot \mathbf{s})(\mathbf{f}_1^{(w_0,t)} \cdot \mathbf{s})}, \dots, \frac{N}{(\mathbf{g}^{(w_0,t)} \cdot \mathbf{s})(\mathbf{f}_P^{(w_0,t)} \cdot \mathbf{s})}, \frac{-N}{(\mathbf{g}^{(w_0,t)} \cdot \mathbf{s})^2} \right)$$

La formule de Sherman (voir Guttman (1946)) donne alors

$$\Upsilon^{(t+1)} = \Upsilon^{(t)} - \Upsilon^{(t)} \mathbf{C} \left( \mathbf{C} \Upsilon^{(t)} \mathbf{C} + \mathbf{D}^{-1} \right)^{-1} \mathbf{C} \Upsilon^{(t)} \quad (12)$$

et il n'est ainsi nécessaire que d'inverser une matrice carré de côté  $2(P+1)$  pour calculer la nouvelle matrice de Cramér-Rao. Pour calculer sa trace, on peut utiliser la formule suivante

$$\text{tr}(\Upsilon^{(t+1)}) = \text{tr}(\Upsilon^{(t)}) - \text{tr} \left[ \left( \mathbf{C} \Upsilon^{(t)} \mathbf{C} + \mathbf{D}^{-1} \right)^{-1} \left( \mathbf{C} \Upsilon^{(t)} \right) \left( \Upsilon^{(t)} \mathbf{C} \right) \right]$$

où on a permuté le produit de matrices, ce qui permet de réduire encore le temps de calcul car les produits de matrices successifs sont de moins grande dimension. De plus, si on a

$$\frac{N}{\mathbf{g}^{(w_0,t)} \cdot \mathbf{s}} = \frac{N}{\mathbf{g}^{(w_0,t+1)} \cdot \mathbf{s}} = \beta^{(w_0)} \quad (13)$$

alors on peut écrire la différence des matrices d'information de Fisher sous la même forme mais avec  $\mathbf{C}$  de taille  $K \times 6$  telle que

$$\mathbf{C} = \left( \delta \mathbf{f}_{m_0}^{(w_0,t+1)}, \delta \mathbf{f}_{m_1}^{(w_0,t+1)}, \delta \mathbf{f}_{m_0}^{(w_0,t)}, \delta \mathbf{f}_{m_1}^{(w_0,t)}, \delta \mathbf{g}^{(w_0,t+1)}, \delta \mathbf{g}^{(w_0,t)} \right)$$

et  $\mathbf{D}$  diagonale de taille  $6 \times 6$  dont les coefficients diagonaux sont ceux du vecteur

$$\mathbf{d} = \left( \frac{\beta^{(w_0)}}{\mathbf{f}_{m_0}^{(w_0, t+1)} \cdot \mathbf{s}}, \frac{\beta^{(w_0)}}{\mathbf{f}_{m_1}^{(w_0, t+1)} \cdot \mathbf{s}}, -\frac{\beta^{(w_0)}}{\mathbf{f}_{m_0}^{(w_0, t)} \cdot \mathbf{s}}, -\frac{\beta^{(w_0)}}{\mathbf{f}_{m_1}^{(w_0, t)} \cdot \mathbf{s}}, -\frac{\beta^{(w_0)}}{\mathbf{g}^{(w_0, t+1)} \cdot \mathbf{s}}, \frac{\beta^{(w_0)}}{\mathbf{g}^{(w_0, t)} \cdot \mathbf{s}} \right)$$

et il n'est alors nécessaire que d'inverser une matrice de taille  $6 \times 6$  au lieu de  $2(P+1) \times 2(P+1)$  pour calculer  $\Upsilon^{(t+1)}$ , ce qui est intéressant quand  $P > 2$ . Pour vérifier la condition de l'équation (13) il est nécessaire que  $\sum_{m=1}^P \mathbf{f}_m^{(w_0, t)} = \sum_{m=1}^P \mathbf{f}_m^{(w_0, t+1)}$  ce qui n'est le cas que si le vecteur  $(f_1^{(w_0, t+1)}(k_0), \dots, f_P^{(w_0, t+1)}(k_0))$  est une permutation de  $(f_1^{(w_0, t)}(k_0), \dots, f_P^{(w_0, t)}(k_0))$ . C'est en particulier toujours le cas lorsque les filtres sont complémentaires.

## B Borne de Cramér-Rao de la reconstruction du spectre dans le cas d'un bruit de Poisson

### B.1 Calcul de l'information de Fisher

Soient  $\mathbf{n}^{(1)}, \dots, \mathbf{n}^{(W)}$  où  $\mathbf{n}^{(w)} = (n_1^{(w)}, \dots, n_P^{(w)})$  des mesures d'un spectre  $\mathbf{r}$  effectuées avec les jeux de filtres  $\mathbf{F}^{(1)}, \dots, \mathbf{F}^{(W)}$  en arrêtant l'acquisition après un temps  $\tau$  pour chacun des filtres. On pose  $\mathbf{n}$  la transposée de  $(\mathbf{n}^{(1)}, \dots, \mathbf{n}^{(W)})$ . On veut calculer la borne de Cramér-Rao associée à l'estimation du spectre  $\mathbf{r}$  à partir de  $\mathbf{n}$ . Puisque les mesures sont indépendantes les unes des autres et que la distribution de  $\mathbf{n}^{(w)}$  est celle de l'équation 1.4, on a

$$\mathcal{P}(\mathbf{n}) = \prod_{w=1}^W \mathcal{P}(\mathbf{n}^{(w)}) = \prod_{w=1}^W \prod_{m=1}^P \frac{(\tau \phi_m^{(w)})^{n_m^{(w)}}}{n_m^{(w)}!} \exp(-\tau \phi_m^{(w)}).$$

où  $\phi_m^{(w)} = \mathbf{f}_m^{(w)} \cdot \mathbf{r}$ . Si on note  $\ell(\mathbf{n} | \mathbf{r})$  la log-vraisemblance associée aux mesures alors

$$\ell(\mathbf{n} | \mathbf{r}) = \sum_{w=1}^W \sum_{m=1}^P n_m^{(w)} \log(\tau \phi_m^{(w)}) - \log n_m^{(w)}! - \tau \phi_m^{(w)}.$$

Le spectre  $\mathbf{r}$  que nous considérons ici n'est pas normalisé. En effet, contrairement au cas où les mesures sont effectuées à nombre de photons fixé, les performances de reconstruction ne sont pas indépendantes de l'intensité du rayonnement observé lorsque les mesures sont réalisées à temps fixé comme c'est le cas ici. Pour plus de simplicité, nous avons choisi d'inclure l'intensité dans le spectre à reconstruire plutôt que de traiter ce paramètre séparément. Par conséquent,

$$\partial_{r(k)} \phi_m^{(w)} = \partial_{r(k)} \sum_{k=1}^K f_m^{(w)}(k) r(k) = f_m^{(w)}(k).$$

On en déduit

$$\begin{aligned} \partial_{r(k)} \ell(\mathbf{n} | \mathbf{r}) &= \sum_{w=1}^W \sum_{m=1}^P \partial_{r(k)} \left[ n_m^{(w)} \log(\tau \phi_m^{(w)}) - \tau \phi_m^{(w)} \right] \\ &= \sum_{w=1}^W \sum_{m=1}^P n_m^{(w)} \frac{f_m^{(w)}(k)}{\phi_m^{(w)}} + \tau f_m^{(w)}(k) \end{aligned}$$

puis

$$\partial_{r(k)} \partial_{r(l)} \ell(\mathbf{n} | \mathbf{r}) = - \sum_{w=1}^W \sum_{m=1}^P n_m^{(w)} \frac{f_m^{(w)}(k) f_m^{(w)}(l)}{\phi_m^{(w)2}}$$

et

$$[\mathbf{I}]_{k,l} = - \langle \partial_{r(k)} \partial_{r(l)} \ell(\mathbf{n} | \mathbf{r}) \rangle = \sum_{w=1}^W \sum_{m=1}^P \tau \frac{f_m^{(w)}(k) f_m^{(w)}(l)}{\phi_m^{(w)}} \quad (14)$$

### B.2 Inversibilité de la matrice d'information de Fisher

On étudie dans cette sous-section les conditions sur les filtres sous lesquelles la matrice d'information de Fisher de la reconstruction de spectre dans le cas de mesures réalisées à temps fixé est inversible. On démontre en particulier la proposition 5 de la section 3.3. Dans

ce but, on commence par écrire  $\mathbf{I}$  comme un produit de matrices.

**Lemme 3.** La matrice d'information de Fisher peut s'écrire comme

$$\mathbf{I} = \mathbf{F} \mathbf{Q}^\dagger \mathbf{F}$$

avec  $\mathbf{Q}$  une matrice diagonale de taille  $WP \times WP$  telle que  $[\mathbf{Q}]_{i=j=P(w-1)+m} = \tau/\phi_m^{(w)}$ .

**Démonstration.**

$$\begin{aligned} [\mathbf{F} \mathbf{Q}]_{k,P(w-1)+m} &= \sum_{v=1}^W \sum_{q=1}^P [\mathbf{F}]_{k,P(v-1)+q} [\mathbf{Q}]_{P(v-1)+q,P(w-1)+m} \\ &= \sum_{v=1}^W \sum_{q=1}^P f_q^{(v)}(k) \frac{\tau}{\phi_m^{(w)}} \delta_{v,w} \delta_{m,q} \\ &= f_m^{(w)}(k) \frac{\tau}{\phi_m^{(w)}} \end{aligned}$$

et

$$\begin{aligned} [\mathbf{F} \mathbf{Q}^\dagger \mathbf{F}]_{k,l} &= \sum_{w=1}^W \sum_{m=1}^P [\mathbf{F} \mathbf{Q}]_{k,P(w-1)+m} [\mathbf{F}]_{P(w-1)+m,l} \\ &= \sum_{w=1}^W \sum_{m=1}^P f_m^{(w)}(k) \frac{\tau}{\phi_m^{(w)}} f_m^{(w)}(l) \\ &= [\mathbf{I}]_{k,l}. \end{aligned}$$

■

**Démonstration de la proposition 5.** Le lemme 3 indique que matrice d'information de Fisher peut s'écrire

$$\mathbf{I} = \mathbf{F} \mathbf{Q}^\dagger \mathbf{F}.$$

avec  $\mathbf{Q}$  une matrice diagonale dont les coefficients diagonaux sont strictement positifs. On peut par conséquent écrire

$$\mathbf{I} = \mathbf{F} \sqrt{\mathbf{Q}}^\dagger (\mathbf{F} \sqrt{\mathbf{Q}})$$

ce qui signifie que  $\mathbf{I}$  est la matrice de Gram des lignes de  $\mathbf{F} \sqrt{\mathbf{Q}}$ . Puisque  $\sqrt{\mathbf{Q}}$  est diagonale et que ses coefficients sont strictement positifs alors il est nécessaire et suffisant que  $\mathbf{F}$  soit de rang plein pour que la matrice d'information de Fisher soit inversible. ■

### B.3 Formule de Sherman et calcul de la matrice de Cramér-Rao

Les algorithmes présentés dans les sections 2.4.2 et 2.5.3 consistent à modifier aléatoirement des jeux de filtres un très grand nombre de fois, et demandent de calculer la trace de la borne de Cramér-Rao de chaque ensemble de jeux de filtres résultant de ces modifications.

Comme dans le cas multinomial, nous montrons dans cette section que la modification effectuée à chaque itération de ces algorithmes ne résulte qu'en une perturbation de faible rang de la matrice d'information de Fisher calculée dans le cas d'un bruit de Poisson. Il est par conséquent possible d'utiliser la formule de Sherman-Morrison-Woodbury pour calculer rapidement la matrice de Cramér-Rao correspondant.

Soit un ensemble de jeux de filtres  $\mathbf{F}^{(t)} = (\mathbf{F}^{(1,t)}, \dots, \mathbf{F}^{(W,t)})$ . Ces jeux de filtres sont modifiés selon la procédure décrite dans l'annexe D, à l'issue de laquelle on obtient  $\mathbf{F}^{(t)} = (\mathbf{F}^{(1,t+1)}, \dots, \mathbf{F}^{(W,t+1)})$ . Cet ensemble de jeux de filtres est identique au précédent, à l'exception des coefficients  $f_{m_0}^{(w_0,t+1)}(k_0)$  et  $f_{m_1}^{(w_0,t+1)}(k_0)$  qui diffèrent des coefficients  $f_{m_0}^{(w_0,t)}(k_0)$  et  $f_{m_1}^{(w_0,t)}(k_0)$ . On note  $\mathbf{I}^{(t)}$  et  $\mathbf{I}^{(t+1)}$  les matrices d'information de Fisher de  $\mathbf{F}^{(t)}$  et  $\mathbf{F}^{(t+1)}$  ainsi que  $\Upsilon^{(t)}$  et  $\Upsilon^{(t+1)}$  leurs matrices de Cramér-Rao. On note  $\Delta_t x^{(t)} = x^{(t+1)} - x^{(t)}$ . L'équation (14) donne

$$\begin{aligned} \Delta_t [\mathbf{I}^{(t)}]_{k,l} &= [\mathbf{I}^{(t+1)}]_{k,l} - [\mathbf{I}^{(t)}]_{k,l} \\ &= \Delta_t \left[ \sum_{w=1}^W \sum_{m=1}^P \tau \frac{f_m^{(w)}(k) f_m^{(w)}(l)}{\phi_m^{(w)}} \right] \\ &= \Delta_t \left[ \tau \frac{f_{m_0}^{(w_0,t)}(k) f_{m_0}^{(w_0,t)}(l)}{\phi_{m_0}^{(w_0,t)}} + \tau \frac{f_{m_1}^{(w_0,t)}(k) f_{m_1}^{(w_0,t)}(l)}{\phi_{m_1}^{(w_0,t)}} \right] \end{aligned}$$

On peut donc écrire

$$\Delta_t \mathbf{I} = \mathbf{C} \mathbf{D}^t \mathbf{C}$$

avec la matrice  $\mathbf{C}$  de taille  $K \times 4$  telle que

$$\mathbf{C} = \left( \mathbf{f}_{m_0}^{(w_0,t+1)}, \mathbf{f}_{m_1}^{(w_0,t+1)}, \mathbf{f}_{m_0}^{(w_0,t)}, \mathbf{f}_{m_1}^{(w_0,t)} \right)$$

et  $\mathbf{D}$  la matrice diagonale de côté 4 dont la diagonale est le vecteur  $\mathbf{d}$  tel que

$$\mathbf{d} = \left( \frac{\tau}{\phi_{m_0}^{(w_0,t+1)}}, \frac{\tau}{\phi_{m_1}^{(w_0,t+1)}}, \frac{-\tau}{\phi_{m_0}^{(w_0,t)}}, \frac{-\tau}{\phi_{m_1}^{(w_0,t)}} \right).$$

La formule de Sherman permet alors d'exprimer  $\Upsilon^{(t+1)}$  en fonction de  $\Upsilon^{(t)}$  comme dans l'équation (12), ce qui permet de la calculer à partir de  $\Upsilon^{(t)}$  en n'inversant qu'une matrice de taille  $4 \times 4$  au lieu de  $K \times K$ .

## C Borne de Bhattacharyya

### C.1 Borne de Bhattacharyya à plusieurs jeux de filtres

Soient des mesures  $\mathbf{n}^{(w,1)}, \dots, \mathbf{n}^{(w,L)}$  réalisées avec  $L$  jeux de  $P$  filtres  $\mathbf{F}^{(w,1)}, \dots, \mathbf{F}^{(w,L)}$  sur un unique pixel  $w$  où apparaît l'une des  $M$  espèces de référence. On note les spectres de ces espèces  $\mathbf{s}_1, \dots, \mathbf{s}_M$ . Pour chaque jeu de filtres, l'acquisition est stoppée lorsque le nombre total de photons détectés atteint  $N$ . Puisque les mesures sont indépendantes, la borne de Bhattacharyya pour un couple d'espèces  $u$  et  $v$  a pour expression

$$\mathcal{B}_{u,v}^{(w)} = \frac{1}{2} \sum_{\mathbf{n}^{(w,1)}} \cdots \sum_{\mathbf{n}^{(w,L)}} \prod_{\ell=1}^L \sqrt{\mathcal{P}(\mathbf{n}^{(w,\ell)} | u) \mathcal{P}(\mathbf{n}^{(w,\ell)} | v)}.$$

De plus, comme démontré par Réfrégier, Chevallier et al. (2019),

$$\sum_{\mathbf{n}^{(w,\ell)}} \sqrt{\mathcal{P}(\mathbf{n}^{(w,\ell)} | u) \mathcal{P}(\mathbf{n}^{(w,\ell)} | v)} = \sum_{m=1}^P \sqrt{p_{m,u}^{(w,\ell)} p_{m,v}^{(w,\ell)}}^N = \left[ \mathbf{q}_u^{(w,\ell)} \cdot \mathbf{q}_v^{(w,\ell)} \right]^N.$$

où  $p_{m,u}^{(w,\ell)}$  est défini comme dans l'équation (2.8) pour le jeu de filtres  $\mathbf{F}^{(w,\ell)}$  et le spectre  $\mathbf{s}_u$  et  $\mathbf{q}_u^{(w,\ell)}$  est le vecteur colonne tel que  $[\mathbf{q}_u^{(w,\ell)}]_m = (p_{m,u}^{(w,\ell)})^{1/2}$ . Par conséquent,

$$\mathcal{B}_{u,v}^{(w)} = \frac{1}{2} \prod_{\ell=1}^L \left( \mathbf{q}_u^{(w,\ell)} \cdot \mathbf{q}_v^{(w,\ell)} \right)^N = \frac{1}{2} \prod_{\ell=1}^L 2 \mathcal{B}_{u,v}^{(w,\ell)}.$$

On obtient également une expression simple de la distance de Bhattacharyya :

$$\beta_{u,v}^{(w)} = -\log(2 \mathcal{B}_{u,v}^{(w)}) = \sum_{\ell=1}^L -\log(2 \mathcal{B}_{u,v}^{(w,\ell)}) = \sum_{\ell=1}^L \beta_{u,v}^{(w,\ell)}.$$

### C.2 Borne de Bhattacharyya généralisée à plusieurs jeux de filtres

D'après Réfrégier, Chevallier et al. (2019), la borne de Bhattacharyya généralisée pour  $M$  espèces équiprobables s'écrit :

$$\mathcal{B}^{(w)} = \frac{2}{M} \sum_u \sum_{v>u} \mathcal{B}_{u,v}^{(w)}.$$

Le raisonnement employé pour arriver à cette expression ne dépend pas de la distribution des mesures. La borne de Bhattacharyya généralisée pour plusieurs jeux de filtres par pixel s'écrit donc ici :

$$\mathcal{B}^{(w)} = \frac{1}{M} \sum_u \sum_{v>u} \prod_{\ell=1}^L 2 \mathcal{B}_{u,v}^{(w,\ell)}.$$

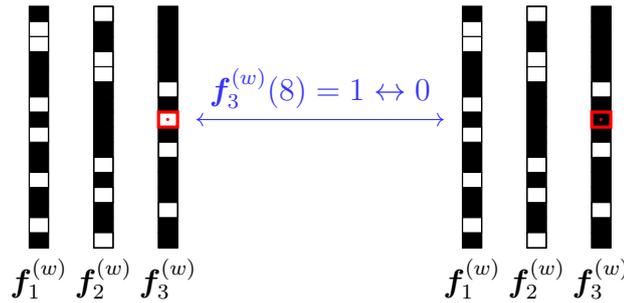
## D Modification aléatoire de jeux de filtres

Les algorithmes présentés dans les sections 2.4.2 et 2.5.3 consistent à réaliser un grand nombre de modifications aléatoires d'ensembles de jeux de filtres, et nous décrivons ici l'algorithme utilisé pour tirer ces modifications.

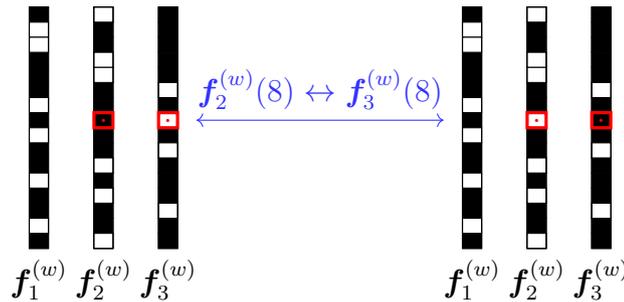
### D.1 Modifications possibles en une itération

Soit un jeu de  $P$  filtres binaires  $\mathbf{F}^{(w)} = (\mathbf{f}_1^{(w)}, \dots, \mathbf{f}_P^{(w)})$ . On n'autorise l'algorithme à effectuer en une itération que les changements suivants :

- ▷ Suppression du coefficient à 1 pour une fréquence, ou inversement.



- ▷ Échange du coefficient à 1 entre deux filtres pour une fréquence.



Il existe des restrictions supplémentaires pour certains types de filtres :

- ▷ Pour les jeux de filtres orthogonaux les modifications faisant apparaître plusieurs  $f_m^{(w)}(k)$  non nuls pour un même couple  $(k, w)$  sont interdites, car cela briserait l'orthogonalité.
- ▷ Pour les jeux de filtres complémentaires il est de surcroît interdit d'effectuer le premier type de modification, car cela briserait la complémentarité.

### D.2 Tirage de la localisation de la modification

Soit un ensemble de  $W$  jeux de  $P$  filtres  $\mathbf{F} = (\mathbf{F}^{(1)}, \dots, \mathbf{F}^{(W)})$ . Le tirage de la modification à appliquer à une itération donnée consiste en la procédure suivante :

- ▷ On commence par tirer l'indice  $w$  du jeu de filtres à modifier. Ce tirage est uniforme parmi l'ensemble des indices des jeux de filtres dont la modification est autorisée.
- ▷ On tire ensuite l'indice  $k$  de la fréquence à modifier dans le jeu de filtres  $\mathbf{F}^{(w)}$ . Cet indice est tiré uniformément parmi la liste des fréquences dont la modification est autorisée pour ce jeu de filtres.
- ▷ Il faut ensuite tirer la modification à appliquer au jeu de filtres  $w$  pour cette fréquence. Cette modification est tirée uniformément parmi les modifications autorisées pour ce jeu de filtres à cette fréquence, qui sont donc un sous-ensemble des modifications décrite dans la sous-section précédente.

## E Convexité de la courbe des compromis optimaux

La courbe des compromis optimaux correspond aux ensembles de jeux de filtres qui minimisent  $\mathcal{C}(\mathbf{F}, \lambda) = (1 - \lambda) \operatorname{tr} \mathbf{\Upsilon}(\mathbf{F}) + \lambda \overline{\mathcal{B}}(\mathbf{F})$  pour  $\lambda \in [0, 1]$ . Nous montrons ici que cette courbe est localisée dans l'enveloppe convexe de l'ensemble des critères  $(\operatorname{tr} \mathbf{\Upsilon}, \overline{\mathcal{B}})$  atteignables.

On note  $\mathbf{F}_\lambda = \arg \min \mathcal{C}(\mathbf{F}, \lambda)$ . Soient  $\mathbf{F}$  et  $\mathbf{F}'$  des ensembles de jeux de filtres tels que  $\operatorname{tr} \mathbf{\Upsilon}(\mathbf{F}) \leq \operatorname{tr} \mathbf{\Upsilon}(\mathbf{F}_\lambda) \leq \operatorname{tr} \mathbf{\Upsilon}(\mathbf{F}')$ . Du fait de cette inégalité, il existe  $\mu \in [0, 1]$  tel que

$$\operatorname{tr} \mathbf{\Upsilon}(\mathbf{F}_\lambda) = \mu \operatorname{tr} \mathbf{\Upsilon}(\mathbf{F}) + (1 - \mu) \operatorname{tr} \mathbf{\Upsilon}(\mathbf{F}'). \quad (15)$$

De plus, par définition

$$\mathcal{C}(\mathbf{F}_\lambda, \lambda) \leq \mathcal{C}(\mathbf{F}, \lambda) \text{ and } \mathcal{C}(\mathbf{F}_\lambda, \lambda) \leq \mathcal{C}(\mathbf{F}', \lambda)$$

donc

$$\mathcal{C}(\mathbf{F}_\lambda, \lambda) \leq \mu \mathcal{C}(\mathbf{F}, \lambda) + (1 - \mu) \mathcal{C}(\mathbf{F}', \lambda).$$

En utilisant l'équation (15) on a

$$\begin{aligned} \lambda \operatorname{tr} \mathbf{\Upsilon}(\mathbf{F}_\lambda) + (1 - \lambda) \overline{\mathcal{B}}(\mathbf{F}_\lambda) \\ \leq \lambda \operatorname{tr} \mathbf{\Upsilon}(\mathbf{F}_\lambda) + (1 - \lambda) (\mu \overline{\mathcal{B}}(\mathbf{F}) + (1 - \mu) \overline{\mathcal{B}}(\mathbf{F}')), \end{aligned}$$

par conséquent,

$$\overline{\mathcal{B}}(\mathbf{F}_\lambda) \leq \mu \overline{\mathcal{B}}(\mathbf{F}) + (1 - \mu) \overline{\mathcal{B}}(\mathbf{F}')$$

et la courbe des compromis optimaux est bien localisée dans l'enveloppe convexe de l'ensemble des critères atteignables dans l'espace  $(\operatorname{tr} \mathbf{\Upsilon}, \overline{\mathcal{B}})$ .

## F Conditions nécessaires à détecter les anomalies

### F.1 Notations spécifiques à la section

Dans cette section on note :

| Grandeur                        | Domaine                            | Définition   | Description   |
|---------------------------------|------------------------------------|--|---|
| $\mathbf{S}$                    | $\mathbb{R}^{K \times M}$          | $(\mathbf{s}_1, \dots, \mathbf{s}_M)$  | Les $M$ spectres de référence.                                    |
| $\mathbf{F}^{(w)}$              | $\mathbb{R}^{K \times P}$          | $(\mathbf{f}_1^{(w)}, \dots, \mathbf{f}_P^{(w)})$                                    | Un jeu de $P$ filtres.  |
| $d(w)$                          | $\mathbb{N}$                       | $\dim(\ker({}^t\mathbf{F}^{(w)}))$   | La dimension de $\ker({}^t\mathbf{F}^{(w)})$ .                    |
| $\mathbf{B}^{(\perp, w)}$       | $\mathbb{R}^{K \times d(w)}$       | $(\mathbf{b}_1^{(\perp, w)}, \dots, \mathbf{b}_{d(w)}^{(\perp, w)})$                 | Une base normée de $\ker({}^t\mathbf{F}^{(w)})$ .                 |
| $\mathbf{B}^{(\mathcal{L}, w)}$ | $\mathbb{R}^{K \times (K-M-d(w))}$ | $(\mathbf{b}_1^{(\mathcal{L}, w)}, \dots, \mathbf{b}_{K-M-d(w)}^{(\mathcal{L}, w)})$ | Une base normée de $\perp(\mathbf{S}, \mathbf{B}^{(\perp, w)})$ . |
| $\mathbf{F}$                    | $\mathbb{R}^{K \times (WP)}$       | $(\mathbf{F}^{(1)}, \dots, \mathbf{F}^{(W)})$  | La matrice des jeux de filtres.                                   |
| $D$                             | $\mathbb{N}$                       | $\dim(\ker({}^t\mathbf{F}))$   | La dimension de $\ker({}^t\mathbf{F})$ .                          |
| $\mathbf{B}^{(\perp)}$          | $\mathbb{R}^{K \times D}$          | $(\mathbf{b}_1^{(\perp)}, \dots, \mathbf{b}_D^{(\perp)})$                            | Une base normée de $\ker({}^t\mathbf{F})$ .                       |
| $\mathbf{B}^{(\mathcal{L})}$    | $\mathbb{R}^{K \times (K-M-D)}$    | $(\mathbf{b}_1^{(\mathcal{L})}, \dots, \mathbf{b}_{K-M-D}^{(\mathcal{L})})$          | Une base normée de $\perp(\mathbf{S}, \mathbf{B}^{(\perp)})$ .    |

On désigne par  $\perp\mathbf{A}$  le supplémentaire orthogonal à  $\text{vect}(\mathbf{A})$ . Notons également que l'on considère que les  $M$  spectres de référence sont linéairement indépendants. Si ce n'est pas le cas, le problème est mal posé. Il est toujours possible d'extraire de  $\mathbf{S}$  une famille libre, mais nous ne distinguons pas ce cas afin de ne pas alourdir inutilement le formalisme.

### F.2 Décomposition des spectres normés

Dans cette sous-section on établit le lemme 6 qui sera utile à la démonstration des propositions 3 et 4. Les lemmes 4 et 5 servent à établir le lemme 6.

**Lemme 4.** S'il est possible d'estimer les concentrations des espèces de référence de spectres  $\mathbf{S}$  à partir de mesures réalisées avec les filtres de  $\mathbf{F}^{(w)}$  alors  $(\mathbf{S}, \mathbf{B}^{(\perp, w)})$  est une famille libre.

**Démonstration.** S'il est possible d'estimer les concentrations des espèces de référence de spectres  $\mathbf{S}$  à partir de mesures réalisées avec  $\mathbf{F}^{(w)}$  alors l'application linéaire correspondant à la matrice  ${}^t\mathbf{F}^{(w)}\mathbf{S}$  est injective. En effet, si ce n'est pas le cas alors plusieurs vecteurs de concentrations  $\mathbf{c}^{(w)}$  génèrent les mêmes flux  $\phi = {}^t\mathbf{F}^{(w)}\mathbf{S}\mathbf{c}^{(w)}$  ce qui rend l'estimation des concentrations impossible. En particulier, aucun  $\mathbf{s}_u$  de  $\mathbf{S}$  ne doit se trouver dans  $\ker({}^t\mathbf{F}^{(w)})$ .

Montrons que  $(\mathbf{S}, \mathbf{B}^{(\perp, w)})$  est libre. Par l'absurde, si  $(\mathbf{S}, \mathbf{B}^{(\perp, w)})$  était liée alors il existerait  $\alpha = (\alpha_1, \dots, \alpha_M)$  et  $\beta = (\beta_2, \dots, \beta_{d(w)})$  deux vecteurs non tous nuls tels que

$$\mathbf{b}_1^{(\perp, w)} = \sum_{v=1}^M \alpha_v \mathbf{s}_v + \sum_{p=2}^{d(w)} \beta_p \mathbf{b}_p^{(\perp, w)}.$$

Or, par définition de  $\mathbf{B}^{(\perp, w)}$  on a  ${}^t\mathbf{F}^{(w)}\mathbf{b}_p^{(\perp, w)} = 0 \forall p$ , donc

$${}^t\mathbf{F}^{(w)}\mathbf{b}_1^{(\perp, w)} = {}^t\mathbf{F}^{(w)} \left( \sum_{v=1}^M \alpha_v \mathbf{s}_v + \sum_{p=2}^d \beta_p \mathbf{b}_p^{(\perp, w)} \right) = {}^t\mathbf{F}^{(w)} \sum_{v=1}^M \alpha_v \mathbf{s}_v = {}^t\mathbf{F}^{(w)}\mathbf{S}\alpha = 0.$$

Puisque l'application linéaire correspondant à  ${}^t\mathbf{F}^{(w)}\mathbf{S}$  est injective, cela implique  $\alpha_v = 0 \forall v$ . Il faudrait donc qu'il existe  $\boldsymbol{\beta}$  non nul tel que

$$\mathbf{b}_1^{(\perp)} = \sum_{p=2}^{d(w)} \beta_p \mathbf{b}_p^{(\perp, w)}$$

ce qui est impossible puisque  $\mathbf{B}^{(\perp, w)}$  est une famille libre en tant que base de  $\ker({}^t\mathbf{F}^{(w)})$ . ■

**Lemme 5.** Si  $d(w) = K - M$  alors  $(\mathbf{S}, \mathbf{B}^{(\perp, w)})$  est une base de  $\mathbb{R}^K$ . Si  $d(w) < K - M$ , il est possible de compléter  $(\mathbf{S}, \mathbf{B}^{(\perp, w)})$  en une base de  $\mathbb{R}^K$  avec une famille de  $K - M - d(w)$  vecteurs normés  $\mathbf{B}^{(\perp, w)}$  n'appartenant pas à  $\ker({}^t\mathbf{F}^{(w)})$  et indépendants de  $\mathbf{S}$ .

**Démonstration.** Le lemme 4 indique que  $(\mathbf{S}, \mathbf{B}^{(\perp, w)})$  est une famille libre. Son cardinal est  $M + d(w)$ . Si  $d(w) = K - M$  alors cette famille est une base de  $\mathbb{R}^K$ . Si  $d(w) < K - M$  le théorème de la base incomplète indique qu'il est possible de compléter  $(\mathbf{S}, \mathbf{B}^{(\perp, w)})$  en une base de  $\mathbb{R}^K$  avec  $K - M - d(w)$  vecteurs indépendants de  $\mathbf{S}$  et de  $\mathbf{B}^{(\perp, w)}$ . ■

**Lemme 6** (Décomposition des spectres adaptée à  $\mathbf{S}$  et à  $\mathbf{F}^{(w)}$ ). On peut écrire tout spectre  $\tilde{\mathbf{s}}$  normé sous la forme

$$\tilde{\mathbf{s}} = \left( \delta\mathbf{S}, \mathbf{s}_M, \delta\mathbf{B}^{(\perp, w)}, \delta\mathbf{B}^{(\perp, w)} \right) \tilde{\mathbf{d}}^{(w)}$$

où  $\tilde{\mathbf{d}}^{(w)}$  est un vecteur de taille  $K$  dont la  $M^{\text{ème}}$  composante vaut 1 et

$$\begin{aligned} \delta\mathbf{S} &= (\mathbf{s}_1 - \mathbf{s}_M, \dots, \mathbf{s}_{M-1} - \mathbf{s}_M) \\ \delta\mathbf{B}^{(\perp, w)} &= (\mathbf{b}_1^{(\perp, w)} - \mathbf{s}_M, \dots, \mathbf{b}_{d(w)}^{(\perp, w)} - \mathbf{s}_M) \\ \delta\mathbf{B}^{(\perp, w)} &= (\mathbf{b}_1^{(\perp, w)} - \mathbf{s}_M, \dots, \mathbf{b}_{K-M-d(w)}^{(\perp, w)} - \mathbf{s}_M). \end{aligned}$$

**Démonstration.** Le lemme 5 indique que la famille  $(\mathbf{S}, \mathbf{B}^{(\perp, w)}, \mathbf{B}^{(\perp, w)})$  est une base de  $\mathbb{R}^K$ . La sous-famille  $(\mathbf{s}_1, \dots, \mathbf{s}_{M-1}, \mathbf{B}^{(\perp, w)}, \mathbf{B}^{(\perp, w)})$  est donc libre. Retirer à tous les vecteurs d'une famille libre un même vecteur linéairement indépendant de tous ses membres résulte en une famille libre, donc  $(\delta\mathbf{S}, \delta\mathbf{B}^{(\perp, w)}, \delta\mathbf{B}^{(\perp, w)})$  est libre. Tous les vecteurs de cette famille se somment à 0, comme différence de deux vecteurs normés. La famille  $(\delta\mathbf{S}, \delta\mathbf{B}^{(\perp, w)}, \delta\mathbf{B}^{(\perp, w)})$  étant une famille libre de cardinal  $K - 1$  du sous-espace vectoriel de dimension  $K - 1$  de  $\mathbb{R}^K$   $\mathcal{X}_0 = \{\mathbf{x}_0 \in \mathbb{R}^K, \sum_{k=1}^K x(k) = 0\}$  alors c'est une base de  $\mathcal{X}_0$ .

Le plan  $\mathcal{X}_0$  est parallèle à  $\mathcal{X}_1 = \{\mathbf{x}_1 \in \mathbb{R}^K, \sum_{k=1}^K x(k) = 1\}$ . Le simplexe de  $\mathbb{R}^K$  est l'ensemble des vecteurs positifs de  $\mathcal{X}_1$ . Ainsi, tous les vecteurs du simplexe peuvent s'exprimer en ajoutant à un vecteur de  $\mathcal{X}_0$  un vecteur se sommant à 1 que l'on choisit égal à  $\mathbf{s}_M$ . Autrement dit

$$\forall \mathbf{x}_1 \in \mathcal{X}_0 \exists! \mathbf{x}_0 \mathbf{x}_1 = \mathbf{x}_0 + \mathbf{s}_M.$$

On peut vérifier simplement que pour tout  $\tilde{\mathbf{d}}^{(w)}$  tel que  $[\tilde{\mathbf{d}}^{(w)}]_M = 1$  on a  $\sum_{k=1}^K \tilde{s}(k) = 1$ . ■

### F.3 Démonstration de la proposition 3

On suppose dans cette section que  $\mathbf{F}^{(w)}$  vérifie les conditions nécessaires à pouvoir réaliser l'estimation des concentrations des espèces de références de spectre  $\mathbf{S}$ . Si  $d(w) > K - M$  alors  $\text{rg}({}^t\mathbf{F}) < M$  et il est impossible d'estimer les concentrations. On suppose donc que  $d(w) \leq K - M$  et on veut montrer qu'il est nécessaire et suffisant que  $d(w) < K - M$  afin de ne pas être aveugle à une part non négligeable d'anomalies.

**Démonstration.**

**(Condition nécessaire).** On suppose ici que  $d(w) = K - M \Leftrightarrow \text{rg}(\mathbf{F}^{(w)}) = M$  et on démontre qu'il existe dans ce cas un ensemble non négligeable de spectres systématiquement indiscernables des spectres de référence, c'est à dire peu importe les concentrations des espèces correspondant dans le mélange observé.

Puisque  $\text{rg}(\mathbf{F}^{(w)}) = M$ , alors  $\text{vect}(\mathbf{S}, \mathbf{B}^{(\perp, w)}) = \mathbb{R}^K$  et le lemme 6 indique que tout spectre  $\tilde{\mathbf{s}}_u$  normalisé peut s'écrire sous la forme

$$\tilde{\mathbf{s}}_u = \left( \delta \mathbf{S}, \mathbf{s}_M, \delta \mathbf{B}^{(\perp, w)} \right) \tilde{\mathbf{d}}_u^{(w)}$$

avec  $\tilde{\mathbf{d}}_u^{(w)}$  un vecteur de taille  $K$  dont la  $M^{\text{ème}}$  composante vaut 1. En posant

$$\tilde{\mathbf{d}}_u^{(w)} = \left( \tilde{\alpha}_{1,u}^{(w)}, \dots, \tilde{\alpha}_{M-1,u}^{(w)}, 1, \tilde{\beta}_{1,u}^{(w)}, \dots, \tilde{\beta}_{K-M,u}^{(w)} \right)$$

on a

$$\tilde{\mathbf{s}}_u = \sum_{v=1}^{M-1} \tilde{\alpha}_{v,u}^{(w)} \mathbf{s}_v + \sum_{p=1}^{K-M} \tilde{\beta}_{p,u}^{(w)} \mathbf{b}_p^{(\perp, w)} + \left( 1 - \sum_{v=1}^{M-1} \tilde{\alpha}_{v,u}^{(w)} - \sum_{p=1}^{K-M} \tilde{\beta}_{p,u}^{(w)} \right) \mathbf{s}_M.$$

Autrement dit, on peut écrire tout spectre  $\tilde{\mathbf{s}}_u$  comme une perturbation d'une combinaison linéaire des spectres de référence par des vecteurs  $\mathbf{B}^{(\perp, w)}$  de  $\ker({}^t \mathbf{F}^{(w)})$ . Pour tout jeu de spectres  $\tilde{\mathbf{S}}$  il existe donc une matrice  $\tilde{\mathbf{D}}^{(w)}$  possédant  $K$  lignes et dont la  $M^{\text{ème}}$  ligne est composée de 1 telle que

$$\tilde{\mathbf{S}} = \left( \delta \mathbf{S}, \mathbf{s}_M, \delta \mathbf{B}^{(\perp, w)} \right) \tilde{\mathbf{D}}^{(w)}.$$

Puisque les vecteurs de  $\mathbf{B}^{(\perp, w)}$  appartiennent à  $\ker({}^t \mathbf{F}^{(w)})$  on a pour tout  $\tilde{\mathbf{c}}^{(w)}$  de taille  $M$

$$\begin{aligned} {}^t \mathbf{F}^{(w)} \tilde{\mathbf{S}} \tilde{\mathbf{c}}^{(w)} &= {}^t \mathbf{F}^{(w)} \left( \delta \mathbf{S}, \mathbf{s}_M, \delta \mathbf{B}^{(\perp, w)} \right) \tilde{\mathbf{D}}^{(w)} \tilde{\mathbf{c}}^{(w)} \\ &= {}^t \mathbf{F}^{(w)} \left( \delta \mathbf{S}, \mathbf{s}_M, -\mathbf{s}_M, \dots, -\mathbf{s}_M \right) \tilde{\mathbf{D}}^{(w)} \tilde{\mathbf{c}}^{(w)} \\ &= {}^t \mathbf{F}^{(w)} \mathbf{S} \mathbf{c}^{(w)} \end{aligned}$$

avec  $\mathbf{c}^{(w)}$  le vecteur de taille  $M$  tel que

$$\begin{cases} c_{v \neq M}^{(w)} = \sum_u \tilde{c}_u^{(w)} \tilde{\alpha}_{v,u}^{(w)} \\ c_M^{(w)} = \sum_u \tilde{c}_u^{(w)} \left( 1 - \sum_{v=1}^{M-1} \tilde{\alpha}_{v,u}^{(w)} - \sum_{p=1}^{K-M} \tilde{\beta}_{p,u}^{(w)} \right). \end{cases}$$

Un mélange de  $\tilde{\mathbf{S}}$  en concentrations  $\tilde{\mathbf{c}}^{(w)}$  est donc équivalent à un mélange de  $\mathbf{S}$  en concentrations  $\mathbf{c}^{(w)}$  lorsque l'observation est réalisée à travers le jeu de filtres  $\mathbf{F}^{(w)}$ . Si  $\mathbf{c}^{(w)}$  est positif alors ces concentrations sont physiques et un tel mélange ne pourra être discerné d'un mélange des espèces de référence.

Pour que  $\mathbf{c}^{(w)}$  soit positif peu importe  $\tilde{\mathbf{c}}^{(w)}$  positif et donc que tout mélange de  $\tilde{\mathbf{S}}$  avec des concentrations physiques soit indiscernable d'un mélange de  $\mathbf{S}$  avec des concentrations physiques, il est nécessaire et suffisant que tout spectre  $\tilde{\mathbf{s}}_u$  de  $\tilde{\mathbf{S}}$  soit tel que

$$\begin{cases} \tilde{\alpha}_{1,u}^{(w)}, \dots, \tilde{\alpha}_{M-1,u}^{(w)} & \geq 0 \\ \sum_{v=1}^{M-1} \tilde{\alpha}_{v,u}^{(w)} + \sum_{p=1}^{K-M} \tilde{\beta}_{p,u}^{(w)} & \leq 1 \end{cases}$$

L'ensemble des vecteurs normés vérifiant ces contraintes est de mesure non nulle. C'est également le cas des vecteurs normés et positifs vérifiant ces contraintes, dans le cas général.

Il existe également des jeux de spectres  $\tilde{\mathbf{S}}$  tels que le vecteur  $\mathbf{c}^{(w)}$  n'est positif que pour certaines valeurs de  $\tilde{\mathbf{c}}^{(w)}$ . Les espèces correspondant ne seront discernables des espèces de référence que lorsque les concentrations du mélange observé ne permettent pas de vérifier la positivité de  $\mathbf{c}^{(w)}$ , ce qui peut représenter une fraction potentiellement significative des mélanges possibles.

**(Condition suffisante).** On suppose ici  $d(w) = \dim(\ker({}^t\mathbf{F}^{(w)})) < K - M \Leftrightarrow \text{rg}(\mathbf{F}^{(w)}) > M$ . On montre que l'ensemble des spectres systématiquement indiscernables des spectres de référence est alors négligeable.

Le lemme 6 indique que tout spectre  $\tilde{\mathbf{s}}_u$  normalisé peut s'écrire sous la forme

$$\tilde{\mathbf{s}}_u = \left( \delta\mathbf{S}, \mathbf{s}_M, \delta\mathbf{B}^{(\perp,w)}, \delta\mathbf{B}^{(\neq,w)} \right) \tilde{\mathbf{d}}_u^{(w)}$$

avec  $\tilde{\mathbf{d}}_u^{(w)}$  un vecteur de taille  $K$  dont la  $M^{\text{ème}}$  composante vaut 1. En posant

$$\tilde{\mathbf{d}}_u^{(w)} = \left( \tilde{\alpha}_{1,u}^{(w)}, \dots, \tilde{\alpha}_{M-1,u}^{(w)}, 1, \tilde{\beta}_{1,u}^{(w)}, \dots, \tilde{\beta}_{d(w),u}^{(w)}, \tilde{\gamma}_{1,u}^{(w)}, \dots, \tilde{\gamma}_{K-M-d(w),u}^{(w)} \right)$$

on a

$$\begin{aligned} \tilde{\mathbf{s}}_u = & \sum_{v=1}^{M-1} \tilde{\alpha}_{v,u}^{(w)} \mathbf{s}_v + \sum_{p=1}^{d(w)} \tilde{\beta}_{p,u}^{(w)} \mathbf{b}_p^{(\perp,w)} + \sum_{p=1}^{K-M-d(w)} \tilde{\gamma}_{p,u}^{(w)} \mathbf{b}_p^{(\neq,w)} \\ & + \left( 1 - \sum_{v=1}^{M-1} \tilde{\alpha}_{v,u}^{(w)} - \sum_{p=1}^{d(w)} \tilde{\beta}_{p,u}^{(w)} - \sum_{p=1}^{K-M-d(w)} \tilde{\gamma}_{p,u}^{(w)} \right) \mathbf{s}_M. \end{aligned}$$

Il existe donc pour tout jeu de spectres  $\tilde{\mathbf{S}}$  une matrice  $\tilde{\mathbf{D}}^{(w)}$  possédant  $K$  lignes et dont la  $M^{\text{ème}}$  ligne est composée de 1 telle que

$$\tilde{\mathbf{S}} = \left( \delta\mathbf{S}, \mathbf{s}_M, \delta\mathbf{B}^{(\perp,w)}, \delta\mathbf{B}^{(\neq,w)} \right) \tilde{\mathbf{D}}^{(w)}.$$

Puisque les vecteurs de  $\mathbf{B}^{(\perp,w)}$  appartiennent à  $\ker({}^t\mathbf{F}^{(w)})$  mais pas les vecteurs de  $\mathbf{B}^{(\neq,w)}$  on a pour tout  $\tilde{\mathbf{c}}^{(w)}$

$$\begin{aligned} {}^t\mathbf{F}^{(w)} \tilde{\mathbf{S}} \tilde{\mathbf{c}}^{(w)} &= {}^t\mathbf{F}^{(w)} \left( \delta\mathbf{S}, \mathbf{s}_M, \delta\mathbf{B}^{(\perp,w)}, \delta\mathbf{B}^{(\neq,w)} \right) \tilde{\mathbf{D}}^{(w)} \tilde{\mathbf{c}}^{(w)} \\ &= {}^t\mathbf{F}^{(w)} \mathbf{S} \mathbf{c}^{(w)} + {}^t\mathbf{F}^{(w)} \mathbf{B}^{(\neq,w)} \mathbf{c}^{(\neq,w)} \end{aligned}$$

avec  $\mathbf{c}^{(w)}$  de taille  $M$  tel que

$$\begin{cases} c_{v \neq M}^{(w)} = \sum_u \tilde{c}_u^{(w)} \tilde{\alpha}_{v,u}^{(w)} \\ c_M^{(w)} = \sum_u \tilde{c}_u^{(w)} \left( 1 - \sum_{v=1}^{M-1} \tilde{\alpha}_{v,u}^{(w)} - \sum_{p=1}^{d(w)} \tilde{\beta}_{p,u}^{(w)} - \sum_{p=1}^{K-M-d(w)} \tilde{\gamma}_{p,u}^{(w)} \right) \end{cases}$$

et  $\mathbf{c}^{(\neq,w)}$  de taille  $K - M - d(w)$  tel que

$$c_p^{(\neq,w)} = \sum_u \tilde{c}_u^{(w)} \tilde{\gamma}_{p,u}^{(w)}.$$

Pour que  $\mathbf{c}^{(w)}$  soit positif et  $\mathbf{c}^{(\perp,w)}$  soit nul peu importe  $\tilde{\mathbf{c}}^{(w)}$  positif et donc que tout mélange de  $\tilde{\mathbf{S}}$  en concentrations physiques soit indiscernable d'un mélange de  $\mathbf{S}$  en concentrations physiques, il est nécessaire et suffisant que tout spectre  $\tilde{\mathbf{s}}_u$  de  $\tilde{\mathbf{S}}$  soit tel que

$$\begin{cases} \tilde{\alpha}_{1,u}^{(w)}, \dots, \tilde{\alpha}_{M-1,u}^{(w)} & \geq 0 \\ \sum_{v=1}^{M-1} \tilde{\alpha}_{v,u}^{(w)} + \sum_{p=1}^{d(w)} \tilde{\beta}_{p,u}^{(w)} & \leq 1 \\ \tilde{\gamma}_{1,u}^{(w)}, \dots, \tilde{\gamma}_{K-M-d(w),u}^{(w)} & = 0 \end{cases}$$

L'ensemble des vecteurs normés et donc des spectres normés et positifs vérifiant ces contraintes est de dimension strictement inférieure à  $K$  du fait des contraintes sur les coordonnées  $\tilde{\gamma}_{p,u}^{(w)}$  et donc de mesure nulle dans le simplexe de  $\mathbb{R}^K$ . ■

#### F.4 Démonstration de la proposition 4

Soit  $\mathbf{F} = (\mathbf{F}^{(1)}, \dots, \mathbf{F}^{(W)})$  un ensemble de jeux de filtres tel que  $\forall w \in \llbracket 1, W \rrbracket \text{ rg}(\mathbf{F}^{(w)}) > M$  et que  $\mathbf{F}^{(w)}$  vérifie les conditions nécessaires à pouvoir réaliser l'estimation des concentrations des espèces de références de spectres  $\mathbf{S}$ .

**Démonstration de la proposition 4.** On suppose que la proposition 2 n'est pas vérifiée, c'est-à-dire que  $(\mathbf{S}, \mathbf{B}^{(\perp,1)}, \dots, \mathbf{B}^{(\perp,W)})$  ne génère pas  $\mathbb{R}^K$ . On montre qu'il est alors possible de trouver des jeux de spectres  $\tilde{\mathbf{S}}$  n'étant pas dans la classe d'équivalence de  $\mathbf{S}$  systématiquement indiscernables des spectres de référence.

Puisque tous les vecteurs de  $\mathbf{B}^{(\perp)}$  sont orthogonaux à  $\mathbf{S}$ , il existe une sous-famille  $\mathbf{B}'$  de  $(\mathbf{B}^{(\perp,1)}, \dots, \mathbf{B}^{(\perp,W)})$  telle que  $(\mathbf{S}, \mathbf{B}')$  soit libre. Le théorème de la base incomplète indique que l'on peut compléter  $(\mathbf{S}, \mathbf{B}')$  en une base de  $\mathbb{R}^K$  avec une base normée du supplémentaire orthogonal de  $(\mathbf{S}, \mathbf{B}')$ , que l'on note  $\mathbf{E}$  et dont le cardinal est noté  $\text{card}(\mathbf{E})$ .

Le lemme 6 indique que tout vecteur  $\mathbf{e}_p$  de  $\mathbf{E}$  peut se décomposer comme

$$\mathbf{e}_p = (\delta\mathbf{S}, \mathbf{s}_M, \delta\mathbf{B}^{(\perp,w)}, \delta\mathbf{B}^{(\perp,w)}) \mathbf{d}_p^{(w)}$$

avec  $\mathbf{d}_p^{(w)}$  un vecteur dont la  $M^{\text{ème}}$  composante est égale à 1. En posant

$$\mathbf{d}_p^{(w)} = \left( \alpha_{1,p}^{(w)}, \dots, \alpha_{M-1,p}^{(w)}, 1, \beta_{1,p}^{(w)}, \dots, \beta_{d(w),p}^{(w)}, \gamma_{1,p}^{(w)}, \dots, \gamma_{K-M-d(w),p}^{(w)} \right)$$

on a

$$\begin{aligned} \mathbf{e}_p = & \sum_{v=1}^{M-1} \alpha_{v,p}^{(w)} \mathbf{s}_v + \sum_{m=1}^{d(w)} \beta_{m,p}^{(w)} \mathbf{b}_m^{(\perp,w)} + \sum_{m=1}^{K-M-d(w)} \gamma_{m,p}^{(w)} \mathbf{b}_m^{(\perp,w)} \\ & + \left( 1 - \sum_{v=1}^{M-1} \alpha_{v,p}^{(w)} - \sum_{m=1}^{d(w)} \beta_{m,p}^{(w)} - \sum_{m=1}^{K-M-d(w)} \gamma_{m,p}^{(w)} \right) \mathbf{s}_M. \end{aligned}$$

Nous utilisons le lemme suivant, que nous démontrons plus loin.

**Lemme 7.** Pour tout vecteur  $\mathbf{e}_p$  on a  $\gamma_{m,p}^{(w)} = 0 \forall m, w$ .

On a donc

$$\mathbf{e}_p = \sum_{v=1}^{M-1} \alpha_{v,p}^{(w)} \mathbf{s}_v + \sum_{m=1}^{d(w)} \beta_{m,p}^{(w)} \mathbf{b}_m^{(\perp,w)} + \left( 1 - \sum_{v=1}^{M-1} \alpha_{v,p}^{(w)} - \sum_{m=1}^{d(w)} \beta_{m,p}^{(w)} \right) \mathbf{s}_M. \quad (16)$$

Puisque  $(\mathbf{S}, \mathbf{B}', \mathbf{E})$  est une base normée alors  $(\mathbf{S}, \mathbf{E})$  est libre et normée. En employant la même méthode que dans la preuve du lemme 6 on peut construire des spectres normés  $\tilde{\mathbf{s}}_u$  qui s'écrivent

$$\tilde{\mathbf{s}}_u = (\delta \mathbf{S}, \mathbf{s}_M, \delta \mathbf{E}) \tilde{\mathbf{d}}_u$$

avec

$$\delta \mathbf{E} = (\mathbf{e}_1 - \mathbf{s}_M, \dots, \mathbf{e}_{\text{card}(\mathbf{E})} - \mathbf{s}_M)$$

et  $\tilde{\mathbf{d}}_u$  un vecteur de taille  $K$  dont le  $M^{\text{ème}}$  coefficient est égal à 1. En posant

$$\tilde{\mathbf{d}}_u = \left( \tilde{\alpha}_{1,u}, \dots, \tilde{\alpha}_{M-1,u}, 1, \tilde{\beta}_{1,u}, \dots, \tilde{\beta}_{\text{card}(\mathbf{E}),u} \right)$$

on a

$$\tilde{\mathbf{s}}_u = \sum_{v=1}^{M-1} \tilde{\alpha}_{v,u} \mathbf{s}_v + \sum_{p=1}^{\text{card}(\mathbf{E})} \tilde{\beta}_{p,u} \mathbf{e}_p + \left( 1 - \sum_{v=1}^{M-1} \tilde{\alpha}_{v,u} - \sum_{p=1}^{\text{card}(\mathbf{E})} \tilde{\beta}_{p,u} \right) \mathbf{s}_M.$$

Pour chaque  $w$  on peut donc décomposer un tel spectre sous la forme

$$\begin{aligned} \tilde{\mathbf{s}}_u &= \sum_{v=1}^{M-1} \tilde{\alpha}_{v,u} \mathbf{s}_v \\ &+ \sum_{p=1}^{\text{card}(\mathbf{E})} \tilde{\beta}_{p,u} \left[ \sum_{v=1}^{M-1} \alpha_{v,p}^{(w)} \mathbf{s}_v + \sum_{m=1}^{d(w)} \beta_{m,p}^{(w)} \mathbf{b}_m^{(\perp,w)} + \left( 1 - \sum_{v=1}^{M-1} \alpha_{v,p}^{(w)} - \sum_{m=1}^{d(w)} \beta_{m,p}^{(w)} \right) \mathbf{s}_M \right] \\ &+ \left( 1 - \sum_{v=1}^{M-1} \tilde{\alpha}_{v,u} - \sum_{p=1}^{\text{card}(\mathbf{E})} \tilde{\beta}_{p,u} \right) \mathbf{s}_M \\ &= \sum_{v=1}^{M-1} \mathbf{s}_v \left[ \tilde{\alpha}_{v,u} + \sum_{p=1}^{\text{card}(\mathbf{E})} \tilde{\beta}_{p,u} \alpha_{v,p}^{(w)} \right] \\ &+ \mathbf{s}_M \left[ 1 - \sum_{v=1}^{M-1} \tilde{\alpha}_{v,u} - \sum_{p=1}^{\text{card}(\mathbf{E})} \tilde{\beta}_{p,u} + \sum_{p=1}^{\text{card}(\mathbf{E})} \tilde{\beta}_{p,u} \left( 1 - \sum_{v=1}^{M-1} \alpha_{v,p}^{(w)} - \sum_{m=1}^{d(w)} \beta_{m,p}^{(w)} \right) \right] \\ &+ \sum_{p=1}^{\text{card}(\mathbf{E})} \tilde{\beta}_{p,u} \sum_{m=1}^{d(w)} \beta_{m,p}^{(w)} \mathbf{b}_m^{(\perp,w)}. \end{aligned}$$

On peut ainsi construire un jeu de spectres  $\tilde{\mathbf{S}}$  composé de vecteurs de cette forme. Pour un tel jeu de spectres il existe une matrice  $\tilde{\mathbf{D}}$  possédant  $M + \text{card}(\mathbf{E})$  lignes telle que

$$\tilde{\mathbf{S}} = (\delta \mathbf{S}, \mathbf{s}_M, \delta \mathbf{E}) \tilde{\mathbf{D}}.$$

Puisque les vecteurs  $\mathbf{e}_p$  s'écrivent comme dans l'équation (16),  $\delta \mathbf{E}$  est composé de combinaisons

linéaires de spectres de  $\mathbf{S}$  et de vecteurs de  $\ker({}^t\mathbf{F}^{(w)})$ . On a donc pour tout  $\tilde{\mathbf{c}}^{(w)}$

$${}^t\mathbf{F}^{(w)} \tilde{\mathbf{S}} \tilde{\mathbf{c}}^{(w)} = {}^t\mathbf{F}^{(w)} (\delta\mathbf{S}, \mathbf{s}_M, \delta\mathbf{E}) \tilde{\mathbf{D}} \tilde{\mathbf{c}}^{(w)} = {}^t\mathbf{F}^{(w)} \mathbf{S} \mathbf{c}^{(w)}$$

avec  $\mathbf{c}^{(w)}$  tel que

$$\begin{cases} c_{v \neq M}^{(w)} = \sum_u \tilde{c}_u^{(w)} \left( \tilde{\alpha}_{v,u} + \sum_{p=1}^{\text{card}(\mathbf{E})} \tilde{\beta}_{p,u} \alpha_{v,p}^{(w)} \right) \\ c_M^{(w)} = \sum_u \tilde{c}_u^{(w)} \left[ 1 - \sum_{v=1}^{M-1} \tilde{\alpha}_{v,u} - \sum_{p=1}^{\text{card}(\mathbf{E})} \tilde{\beta}_{p,u} \left( \sum_{v=1}^{M-1} \alpha_{v,p}^{(w)} + \sum_{p=1}^{d(w)} \beta_{m,p}^{(w)} \right) \right]. \end{cases}$$

Si on a pour tout  $u$

$$\begin{cases} \tilde{\alpha}_{v,u} + \sum_{p=1}^{\text{card}(\mathbf{E})} \tilde{\beta}_{p,u} \alpha_{v,p}^{(w)} & \geq 0 \quad \forall v \in \llbracket 1, M-1 \rrbracket \\ 1 - \sum_{v=1}^{M-1} \tilde{\alpha}_{v,u} - \sum_{p=1}^{\text{card}(\mathbf{E})} \tilde{\beta}_{p,u} \left( \sum_{v=1}^{M-1} \alpha_{v,p}^{(w)} + \sum_{m=1}^{d(w)} \beta_{m,p}^{(w)} \right) & \geq 0 \end{cases}$$

alors tout mélange des espèces correspondant à  $\tilde{\mathbf{S}}$  en concentrations physiques est indiscernable d'un mélange des espèces de référence en concentrations physiques lorsqu'observé à travers tout jeu de filtres de  $\mathbf{F}$ .

Puisque  $\mathbf{E}$  est dans le supplémentaire orthogonal de  $(\mathbf{S}, \mathbf{B}^{(\neq)})$  alors  $\text{vect}(\mathbf{S}, \mathbf{E}) \not\subset \text{vect}(\mathbf{S})$  et donc ces spectres ne sont pas nécessairement inclus dans la classe d'équivalence de  $\mathbf{S}$ .

On a ainsi montré que si l'on n'impose pas la contrainte de la proposition 4 alors on peut construire des anomalies systématiquement indiscernables des espèces de référence.  $\blacksquare$

**Démonstration du lemme 7.** On a

$$\begin{aligned} \mathbf{e}_p = & \sum_{v=1}^{M-1} \alpha_{v,p}^{(w)} \mathbf{s}_v + \sum_{m=1}^{d(w)} \beta_{m,p}^{(w)} \mathbf{b}_m^{(\perp,w)} + \sum_{m=1}^{K-M-d(w)} \gamma_{m,p}^{(w)} \mathbf{b}_m^{(\neq,w)} \\ & + \left( 1 - \sum_{v=1}^{M-1} \alpha_{v,p}^{(w)} - \sum_{m=1}^{d(w)} \beta_{m,p}^{(w)} - \sum_{m=1}^{K-M-d(w)} \gamma_{m,p}^{(w)} \right) \mathbf{s}_M \end{aligned}$$

et donc

$$\begin{aligned} \mathbf{e}_p - \sum_{v=1}^{M-1} \alpha_{v,p}^{(w)} \mathbf{s}_v - \sum_{m=1}^{d(w)} \beta_{m,p}^{(w)} \mathbf{b}_m^{(\perp,w)} - \left( 1 - \sum_{v=1}^{M-1} \alpha_{v,p}^{(w)} - \sum_{m=1}^{d(w)} \beta_{m,p}^{(w)} - \sum_{m=1}^{K-M-d(w)} \gamma_{m,p}^{(w)} \right) \mathbf{s}_M \\ = \sum_{m=1}^{K-M-d(w)} \gamma_{m,p}^{(w)} \mathbf{b}_m^{(\neq,w)}. \end{aligned}$$

Or, les vecteurs de  $\mathbf{B}^{(\neq,w)}$  sont dans le supplémentaire orthogonal de  $(\mathbf{S}, \mathbf{B}^{(\perp,w)})$ . De plus,  $\mathbf{e}_p$  est orthogonal aux vecteurs de  $\mathbf{B}^{(\perp,w)}$ . Les deux membres de cette équation sont donc orthogonaux et il faut nécessairement

$$\sum_{m=1}^{K-M-d(w)} \gamma_{m,p}^{(w)} \mathbf{b}_m^{(\neq,w)} = 0.$$

Puisque  $\mathbf{B}^{(\neq,w)}$  est une famille libre cela implique  $\gamma_{m,p}^{(w)} = 0 \quad \forall m, p, w$ .  $\blacksquare$



# Bibliographie

- Berto, P., C. Scotté, F. Galland, H. Rigneault et H. B. de Aguiar. « Programmable single-pixel-based broadband stimulated Raman scattering ». Dans : *Opt. Lett.* 42.9 (mai 2017), p. 1696-1699.
- Candès, E., J. Romberg et T. Tao. « Robust uncertainty principles : exact signal reconstruction from highly incomplete frequency information ». Dans : *IEEE Transactions on Information Theory* 52.2 (2006), p. 489-509.
- Candès, E. J., J. K. Romberg et T. Tao. « Stable signal recovery from incomplete and inaccurate measurements ». Dans : *Communications on Pure and Applied Mathematics* 59.8 (2006), p. 1207-1223.
- Candès, E. J. et T. Tao. « Near-Optimal Signal Recovery From Random Projections : Universal Encoding Strategies ? ». Dans : *IEEE Transactions on Information Theory* 52.12 (2006), p. 5406-5425.
- Comon, P. et C. Jutten. *Handbook of Blind Source Separation, Independent Component Analysis and Applications*. Academic Press (Elsevier), fév. 2010.
- Corcoran, T. C. « Compressive Detection of Highly Overlapped Spectra Using Walsh–Hadamard-Based Filter Functions ». Dans : *Applied Spectroscopy* 72.3 (2018), p. 392-403.
- Cox, D. R. et D. V. Hinkley. *Theoretical Statistics*. Chapman & Hall, 1974.
- Davis, B. M., A. J. Hemphill, D. Cebeci Maltaş, M. A. Zipper, P. Wang et D. Ben-Amotz. « Multivariate Hyperspectral Raman Imaging Using Compressive Detection ». Dans : *Analytical Chemistry* 83.13 (2011), p. 5086-5092.
- DeVerse, R. A., R. M. Hammaker et W. G. Fateley. « Realization of the Hadamard Multiplex Advantage Using a Programmable Optical Mask in a Dispersive Flat-Field Near-Infrared Spectrometer ». Dans : *Appl. Spectrosc.* 54.12 (déc. 2000), p. 1751-1758.
- Duda, R., P. Hart et D. Stork. *Pattern Classification*. Wiley, 2012.
- Félidj, N., J. Aubard, G. Lévi, J. R. Krenn, A. Hohenau, G. Schider, A. Leitner et F. R. Aussenegg. « Optimized surface-enhanced Raman scattering on gold nanoparticle arrays ». Dans : *Applied Physics Letters* 82.18 (avr. 2003), p. 3095-3097.
- Figue, J. et P. Réfrégier. « Optimality of trade-off filters ». Dans : *Appl. Opt.* 32.11 (avr. 1993), p. 1933-1935.
- Freudiger, C. W., W. Min, B. G. Saar, S. Lu, G. R. Holtom, C. He, J. C. Tsai, J. X. Kang et X. S. Xie. « Label-Free Biomedical Imaging with High Sensitivity by Stimulated Raman Scattering Microscopy ». Dans : *Science* 322.5909 (2008), p. 1857-1861.
- Galvis-Carreño, D., Y. Mejia-Melgarejo et H. Arguello. « Efficient reconstruction of Raman spectroscopy imaging based on compressive sensing ». Dans : *DYNA* 81 (déc. 2014), p. 116-124.
- Garthwaite, P. H., I. T. Jolliffe et B. Jones. *Statistical Inference*. 2<sup>e</sup> éd. Wiley, 2002.
- Grand, C., C. Scotté et H. Rigneault. « Fast Compressive Raman Imaging of Polymorph Molecules and Excipients in Pharmaceutical Tablets ». Dans : *Analytical Chemistry* 94.48 (2022), p. 16632-16637.

- Guttman, L. « Enlargement Methods for Computing the Inverse Matrix ». Dans : *The Annals of Mathematical Statistics* 17.3 (1946), p. 336-343.
- Justel, T., F. Galland et A. Roueff. « Compressed Raman method combining classification and estimation of spectra with optimized binary filters ». Dans : *Opt. Lett.* 47.5 (mars 2022), p. 1101-1104.
- « Optimal trade-off filters for compressed Raman classification and spectrum reconstruction ». Dans : *J. Opt. Soc. Am. A* 40.6 (mai 2023), p. 1058-1065.
- « Spectroscopie Raman comprimée pour la classification supervisée d'espèces : méthodologie pour la conception d'approches robustes. » 2023. Colloque GRETSI, Grenoble, France, 28 août au premier septembre 2023.
- Landsberg, G. et L. Mandelstam. « New phenomenon in scattering of light (preliminary report) ». Dans : *Journal of the Russian Physico-Chemical Society, Physics Section* 5.60 (1928), p. 335.
- Love, S. P. « Programmable matched filter and Hadamard transform hyperspectral imagers based on micromirror arrays ». Dans : *Emerging Digital Micromirror Device Based Systems and Applications*. T. 7210. SPIE. 2009, p. 55-69.
- Marsaglia, G. *Uniform distributions over a simplex*. Rapp. tech. Boeing Scientific Research Labs Seattle WA, 1961.
- Nelson, M. P., J. F. Aust, J. A. Dobrowolski, P. G. Verly et M. L. Myrick. « Multivariate Optical Computation for Predictive Spectroscopy ». Dans : *Analytical Chemistry* 70.1 (1998), p. 73-82.
- Pavillon, N. et N. I. Smith. « Compressed sensing laser scanning microscopy ». Dans : *Opt. Express* 24.26 (déc. 2016), p. 30038-30052.
- Perlin, K. « An Image Synthesizer ». Dans : *SIGGRAPH Comput. Graph.* 19.3 (juill. 1985), p. 287-296.
- Ploetz, E., S. Laimgruber, S. Berner, W. Zinth et P. Gilch. « Femtosecond stimulated Raman microscopy ». Dans : *Applied Physics B* 87 (2007), p. 389-393.
- Quyen, N. T., M. D. Jouan, N. Q. Dao, E. D. Silva et D. A. Phuong. « New Raman Spectrometer Using a Digital Micromirror Device and a Photomultiplier Tube Detector for Rapid On-Line Industrial Analysis. Part II : Choice of Analytical Methods ». Dans : *Appl. Spectrosc.* 62.3 (mars 2008), p. 279-284.
- Quyen, N. T., E. D. Silva, N. Q. Dao et M. D. Jouan. « New Raman Spectrometer Using a Digital Micromirror Device and a Photomultiplier Tube Detector for Rapid On-Line Industrial Analysis. Part I : Description of the Prototype and Preliminary Results ». Dans : *Appl. Spectrosc.* 62.3 (mars 2008), p. 273-278.
- Raman, C. V. « A new radiation ». Dans : *Indian J. Phys* 2 (1928), p. 387-398.
- Réfrégier, P. « Filter design for optical pattern recognition : multicriteria optimization approach ». Dans : *Opt. Lett.* 15.15 (août 1990), p. 854-856.
- « Optimal trade-off filters for noise robustness, sharpness of the correlation peak, and Horner efficiency ». Dans : *Opt. Lett.* 16.11 (juin 1991), p. 829-831.
- Réfrégier, P., E. Chevallier et F. Galland. « Compressed Raman classification method with upper-bounded error probability ». Dans : *Opt. Lett.* 44.23 (déc. 2019), p. 5836-5839.
- Réfrégier, P. et F. Galland. « Bhattacharyya bound for Raman spectrum classification with a couple of binary filters ». Dans : *Opt. Lett.* 44.9 (mai 2019), p. 2228-2231.
- Réfrégier, P., C. Scotté, H. B. de Aguiar, H. Rigneault et F. Galland. « Precision of proportion estimation with binary compressed Raman spectrum ». Dans : *J. Opt. Soc. Am. A* 35.1 (jan. 2018), p. 125-134.

- Rehrauer, O. G., V. C. Dinh, B. R. Mankani, G. T. Buzzard, B. J. Lucier et D. Ben-Amotz. « Binary Complementary Filters for Compressive Raman Spectroscopy ». Dans : *Applied Spectroscopy* 72.1 (2018), p. 69-78.
- Rehrauer, O. G., B. R. Mankani, G. T. Buzzard, B. J. Lucier et D. Ben-Amotz. « Fluorescence modeling for optimized-binary compressive detection Raman spectroscopy ». Dans : *Opt. Express* 23.18 (sept. 2015), p. 23935-23951.
- Rigneault, H. et P. Berto. « Tutorial : Coherent Raman light matter interaction processes ». Dans : *APL Photonics* 3.9 (juill. 2018), p. 091101.
- Schrader, B. *Infrared and Raman spectroscopy : methods and applications*. John Wiley & Sons, 1995.
- Scotté, C. « Spontaneous compressive Raman technology : developments and applications ». 2020AIXM0174. Thèse de doct. 2020.
- Scotté, C., H. B. de Aguiar, D. Marguet, E. M. Green, P. Bouzy, S. Vergnole, C. P. Winlove, N. Stone et H. Rigneault. « Assessment of compressive Raman versus hyperspectral Raman for microcalcification chemical imaging ». Dans : *Analytical chemistry* 90.12 (2018), p. 7197-7203.
- Scotté, C., S. Sivankutty, R. A. Bartels et H. Rigneault. « Line-scan compressive Raman imaging with spatio-spectral encoding ». Dans : *Opt. Lett.* 45.19 (oct. 2020), p. 5567-5570.
- Smekal, A. « Zur Quantentheorie der Dispersion ». Dans : *Naturwissenschaften* 11.43 (oct. 1923), p. 873-875.
- Smith, Z. J., S. Strombom et S. Wachsmann-Hogiu. « Multivariate optical computing using a digital micromirror device for fluorescence and Raman spectroscopy ». Dans : *Opt. Express* 19.18 (août 2011), p. 16950-16962.
- Soldevila, F., J. Dong, E. Tajahuerce, S. Gigan et H. B. de Aguiar. « Fast compressive Raman bio-imaging via matrix completion ». Dans : *Optica* 6.3 (mars 2019), p. 341-346.
- Soyemi, O., D. Eastwood, L. Zhang, H. Li, J. Karunamuni, P. Gemperline, R. A. Synowicki et M. L. Myrick. « Design and Testing of a Multivariate Optical Element : The First Demonstration of Multivariate Optical Computing for Predictive Spectroscopy ». Dans : *Analytical Chemistry* 73.6 (2001), p. 1069-1079.
- Soyemi, O. O., F. G. Haibach, P. J. Gemperline et M. L. Myrick. « Nonlinear Optimization Algorithm for Multivariate Optical Element Design ». Dans : *Appl. Spectrosc.* 56.4 (avr. 2002), p. 477-487.
- Stoica, P. et T. L. Marzetta. « Parameter estimation problems with singular information matrices ». Dans : *IEEE Transactions on Signal Processing* 49.1 (2001), p. 87-90.
- Thompson, J. V., J. N. Bixler, B. H. Hokr, G. D. Noojin, M. O. Scully et V. V. Yakovlev. « Single-shot chemical detection and identification with compressed hyperspectral Raman imaging ». Dans : *Opt. Lett.* 42.11 (juin 2017), p. 2169-2172.
- Uzunbajakava, N., P. de Peinder, G. W. 't Hoof et A. T. M. van Gogh. « Low-Cost Spectroscopy with a Variable Multivariate Optical Element ». Dans : *Analytical Chemistry* 78.20 (2006), p. 7302-7308.
- Volkmer, A., J.-X. Cheng et X. Sunney Xie. « Vibrational Imaging with High Sensitivity via Epidetected Coherent Anti-Stokes Raman Scattering Microscopy ». Dans : *Phys. Rev. Lett.* 87 (2 juin 2001), p. 023901.
- Wagner, E. P., B. W. Smith, S. Madden, J. D. Winefordner et M. Mignardi. « Construction and Evaluation of a Visible Spectrometer Using Digital Micromirror Spatial Light Modulation ». Dans : *Appl. Spectrosc.* 49.11 (nov. 1995), p. 1715-1719.
- Wilcox, D. S., G. T. Buzzard, B. J. Lucier, O. G. Rehrauer, P. Wang et D. Ben-Amotz. « Digital compressive chemical quantitation and hyperspectral imaging ». Dans : *Analyst* 138 (17 2013), p. 4982-4990.

- Wilcox, D. S., G. T. Buzzard, B. J. Lucier, P. Wang et D. Ben-Amotz. « Photon level chemical classification using digital compressive detection ». Dans : *Analytica Chimica Acta* 755 (2012), p. 17-27.
- Woodward, P. « Theory of radar information ». Dans : *Transactions of the IRE Professional Group on Information Theory* 1.1 (1953), p. 108-113.
- Zumbusch, A., G. R. Holtom et X. S. Xie. « Three-Dimensional Vibrational Imaging by Coherent Anti-Stokes Raman Scattering ». Dans : *Phys. Rev. Lett.* 82 (20 mai 1999), p. 4142-4145.