UFR sciences et techniques mathématiques informatique automatique
École doctorale IAEM Lorraine
DFD automatique et production automatisée

THÈSE

présentée pour l'obtention du

Doctorat de l'Université Henri Poincaré, Nancy 1
Spécialité automatique, traitement du signal et génie informatique

par

Vincent Mazet

Développement de méthodes de traitement de signaux spectroscopiques : estimation de la ligne de base et du spectre de raies

soutenue publiquement le 1er décembre 2005

Composition du jury :

Rapporteurs :	Thierry Chonavel	Professeur à l'ENST Bretagne, dép. Signal & Communications, Brest
	Guy Demoment	Professeur à l'Université de Paris Sud, L2S, Orsay
Examinateurs :	Olivier Cappé	Chargé de recherche CNRS, LTCI, Paris
	Jérôme Idier	Chargé de recherche CNRS, IRCCyN, Nantes
	Bernard Humbert	Professeur à l'Université Henri Poincaré, LCPME, Nancy
	David Brie	Professeur à l'Université Henri Poincaré, CRAN, Nancy
Invité :	Cédric Carteret	Maître de conférences à l'Université Henri Poincaré, LCPME, Nancy

Centre de recherche en automatique de Nancy
CRAN – CNRS UMR 7039

Remerciements

Cette thèse a débuté au rayon fruits et légumes d'un supermarché nancéien lorsque les professeurs Humbert et Brie, respectivement président de mon jury et directeur de thèse, mais également anciens camarades de l'université, se sont rencontrés, la tête pleine de soucis. Le premier cherchait à améliorer l'analyse de ses spectres, le second cherchait un maracudja bien mûr. Quelques mois plus tard, le petit Vincent arrivait.

Vincent commença sa thèse en engrangeant du bois pour son patron, mais la maturité et la confiance en soi lui fire refuser de repeindre l'intérieur de sa villa quelques années plus tard.

Bien que la couverture de ce présent document ne fasse apparaître que son nom, ce travail de thèse n'aurait pu voir le jour sans les conseils et le soutien d'une myriade d'autres personnes qui auraient presque mérité, tout comme lui, d'apparaître sur la première page de ce présent mémoire. Ces pages ont donc pour but de réparer cette omission.

Aussi, que soient remerciés monsieur le professeur Francis Lepage, ancien directeur du CRAN, pour m'avoir permis de commencer cette thèse, et monsieur le professeur Alain Richard, directeur actuel, pour m'avoir permis de la terminer.

Je témoigne également toute ma gratitude à messieurs les professeurs Guy Demoment et Thierry Chonavel pour avoir accepté le lourd travail qui est de rapporter une thèse.

Je souhaite également remercier monsieur Olivier Cappé pour avoir accepter d'examiner mon travail.

Je remercie très chaleureusement monsieur Jérôme Idier, sans qui ma thèse aurait probablement été très différente. J'ai passé un excellent séjour à Nantes au début de ma thèse, tant sur le plan personnel que professionnel. Les discussions qui ont suivi m'ont été très précieuses et sont à l'origine de paragraphes entiers de la thèse... J'espère que notre collaboration continuera. Je le remercie également pour avoir accepté d'être examinateur de ma thèse.

Merci également à monsieur le professeur Bernard Humbert, tout d'abord pour avoir été à l'origine d'un travail qui a occupé ces trois années et pour aider à le réaliser, mais également pour avoir examiné ma thèse et présidé mon jury.

Merci enfin à monsieur Cédric Carteret pour toutes ses remarques, son amitié, et avant tout sa disponibilité.

Je souhaite également remercier, d'un point de vue professionnel, messieurs Christian Heinrich, Manuel Davy, El Hadi Djermoune, Saïd Moussaoui et Stéphane Thil pour leurs conseils qui m'ont débloqué de plusieurs situations dans lesquelles je m'étais embourbé.

Je n'oublie évidemment pas tous les membres du couloir auprès de qui j'ai passé tant de moments drôlement bien ! En vrac, voici le générique :

Merci donc à Miko, le Champollion du Vichnout, le pharaon de la géographie, le chanteur aux cordes vocales dorées (ou plutôt ambrées !). Ces quatre années de collocation ont été vraiment super, tant que tu n'abusais pas du déo. J'espère que notre amitié continuera encore longtemps, vers ce lointain finistère aux longues plages de silence (et au delà)...

Merci aussi à mon frère de thèse Saïd, grâce à qui je ne me sens plus coupable de passer à l'orange !

Merci à Fartass, l'homme qui renversait son verre d'alcool plus vite que son ombre. Grâce à lui, j'ai pu m'exercer à la pratique de la muscu et des jeux vidéos.

Merci à Ludo, qui m'a fait découvrir la musique (notamment cette magnifique chanteuse japonaise à la voix si envoûtante) et le cinéma (ce film de superhéros aux physiques de Steve Urkel est devenu ma référence). Je te souhaite plein de bonnes choses avec Lena et sus aux Covenants !

Merci à Bounoît, grâce à qui j'ai passé pas mal de soirées sympas, et surtout un week-end extraordinaire au pays des volcans. C'est d'ailleurs un peu grâce à lui que je suis maintenant un playboy au regard intense.

Merci à Mag, pour sa présence féminine si importante, et à Hadi.

Merci aux frères siamois Stéphane et JPP pour leur bonne humeur.

Merci à Ericus pour les parties de badminton, où j'ai pu exercé toute ma créativité sportive.

Merci à Manu, en qui je suis mis en exergue, de les potins, et pas besoin de acheté le *Voici*, même si c'est pas loin de atchez moi le marchand del journal et merci pour, les dou tonnes de ta thèse, que les correctiòns j'ai corrigé en bonne heure. Pas merci par contre pour tes conversations téléphoniques en espagnol à 450 dB pendant que de l'autre côté de la cloison, je rédigeais ! Néanmoins, louée soit ta piña colada.

Merci au Pr. Fredouille qui, tel un John Travolta endiablé, enflammait les *dancefloors* de Nancy de ses déhanchements envoûtant et envahissait le couloir du labo de ses éclats de rire extravagants !

Merci à Hicham, qui a bien mérité ses surnoms de Jamel Debbouze de l'IUP (ça, c'est l'effet humour du terroir), de marteau d'Agadir ou encore de pilier de l'Antika.

Merci à Philou pour sa jovialité. Si un jour j'ai une super maison pour pas cher, ce sera grâce à tes conseils ! Je tiens également à remercier Fabienne qui nous a nourri pendant presque tout un été et à Toby, leur premier enfant.

Merci à Zitoune, l'ingénieur informaticien qui aime les ordinateurs, Windows 98 et tout le tralala...

Merci à Sab, la plus merveilleuse secrétaire du monde (je t'ai d'ailleurs piqué une phrase, j'espère que tu ne m'en voudras pas). Ô Sabine, sans toi, nous croulerions sous cet océan de paperasse administrative !

Je n'oublie pas non plus Thierry, qui m'a fait entrer dans la recherche, Fateh, Sinuhé et Brian, Jean-Marc, Christophe, Éric et Christian les trois mousquetaires-CNAM, Jean-Marie (mon maître), Yann et les autres...

Mais les amis du labo n'ont pas été les seules personnes qui ont participé à ces trois années. Merci donc à Adeline et Émiliche, accoucheuses de soirées sympas et de week-ends agréables. Merci à Damien, dorénavant mon DJ attitré, pour m'avoir fourni quelques MP3 qui ont rythmé ma rédaction. Merci à ma famille de m'avoir soutenu aveuglément, sans comprendre un traître mot de mon travail. Et enfin merci à ma chérie d'amour pour m'avoir permis de penser à autre chose qu'au boulot.

Je tiens enfin à remercier chaleureusement tous mes sponsors, sans qui ces trois longues années de recherches passionnées et assidues auraient été un enfer : Nutella, le restaurant universitaire de Vandœuvre et sa sympathique et souriante serveuse, le Caméo, la Parenthèse et l'Association, mes CD, Funky Georgy, *Friends*, *Un gars une fille* et *Kaamelott*, l'équipe du Café des Anges et surtout celle du New Mustang, les autochtones de Louvain-la-Neuve, etc.

Mais il me semble que j'oublie quelqu'un... Bong sang mais c'est bien sûr ! *Last but not least* donc, merci à toi, David, grand professeur qui m'a montré le chemin vers la Connaissance et la Vérité, toi qui m'a guidé de ta lumière ! Ah patron ! C'que vous êtes fort ! Non, sérieux, merci pour tout, notamment de m'avoir fait rencontrer d'autres chercheurs, de m'avoir appris à écrire français et d'avoir été aussi... disons... embêtant sur la rédaction (c'est pas pour dire... mais le patron quand même... quel type !).

Et maintenant, en route pour de nouvelles aventures !

Table des matières

Abréviations et notations

Abréviations

EAP	Espérance a posteriori
EM	*Expectation-maximisation* [27]
EQM	Erreur quadratique moyenne, définie dans le cadre de l'estimation de la ligne de base (chapitre 2) par : $$\mathrm{EQM} = \frac{1}{N}\sum_{n=1}^{N}(\mathbf{z}_n - \widehat{\mathbf{z}}_n)^2$$
ICM	*Iterative conditional mode* [5]
i.i.d.	Indépendants et identiquement distribués
MA	*Moving Average*
MAP	Maximum a posteriori
MCMC	Méthodes de Monte Carlo par chaînes de Markov (*Monte Carlo Markov chain*) (voir l'annexe A, page 123)
MMAP	Maximum a posteriori marginal
MPM	*Maximum posterior mode* (voir page 70)
RJMCMC	Méthode de Monte Carlo par chaînes de Markov à saut réversible (*reversible jump Monte Carlo Markov chain*) [46, 91]
RSB	Rapport signal à bruit : – dans le cadre de l'estimation de la ligne de base (chapitre 2), le RSB est défini comme le rapport des énergies de la ligne de base $\mathbf{z}$ sur le spectre bruité $\mathbf{w}$. $$\mathrm{RSB} = 10\log\left(\frac{\sum_{n=1}^{N}\mathbf{z}_n^2}{\sum_{n=1}^{N}\mathbf{w}_n^2}\right).$$ – dans le cas de l'estimation des raies du spectre (chapitres 3 et 4), le RSB est défini comme : $$\mathrm{RSB} = 10\log\left(\frac{\sum_{n=1}^{N}\mathbf{v}_n^2}{\sum_{n=1}^{N}\mathbf{b}_n^2}\right).$$
SAEM	*Stochastic approximation EM* [13]
SEM	*Stochastic EM* [12]
u.a.	Unité arbitraire (utilisée généralement pour l'axe des intensités des spectres)

Notations

Afin d'alléger les notations, nous ne distinguerons pas dans cette thèse les variables aléatoires de leur réalisation.

La liste des variables représentant les différents signaux peut être trouvée dans la section 1.7, et en particulier sur la figure 1.11 page 24.

$\sim$	Distribué selon
$\star$	Produit de convolution
$\propto$	Proportionnel à
$\widehat{\cdot}$	Estimation
$\cdot^T$	Transposée
$\mathbf{0}$	Matrice nulle
$\mathbb{1}_E$	Fonction indicatrice, vaut un sur l'espace E et zéro ailleurs
δ_n	Impulsion numérique ou impulsion de Dirac centrée en n
$\mathbb{E}_{x\|\theta}[f(x)]$	Espérance de $f(x)$ suivant x conditionnellement à θ
i	Itération particulière ($i \in \{1, \ldots, I\}$)
I	Nombre d'itérations
I_0	Longueur de la période de transition de la chaîne de Markov
$\mathbf{I}$	Matrice identité
n	Variable discrète du nombre d'onde ($n \in \{1, \ldots, N\}$)
N	Longueur des signaux
$\mathbb{N}$	Entiers naturels
$p(\cdot)$	Densité de probabilité
$\mathbb{R}$	Entiers réels
t	Variable continue du nombre d'onde ($t \in [1, N]$)
$\mathrm{Var}_{x\|\theta}[f(x)]$	Variance de $f(x)$ suivant x conditionnellement à θ
$\mathbf{x}_n$	n^{e} élément du vecteur $\mathbf{x}$
$\mathbf{x}_{-n}$	Vecteur $\mathbf{x}$ sans l'élément $\mathbf{x}_n$
$\mathbf{X}_n$	n^{e} colonne de la matrice $\mathbf{X}$
$\mathbf{X}_{i,j}$	Élement (i, j) de la matrice $\mathbf{X}$

Introduction

Cette thèse s'inscrit dans le cadre d'une collaboration entre le Centre de recherche en automatique de Nancy (CRAN, UHP–INPL–CNRS UMR 7039) et le Laboratoire de chimie-physique et microbiologie pour l'environnement (LCPME, UHP–CNRS UMR 7564).

L'objectif de la collaboration est de développer des méthodes numériques pour faciliter l'analyse de spectres infrarouge et Raman. Cette collaboration a donc des applications en chimiométrie, qui est la science de l'acquisition, de la validation et du traitement des données dans le domaine de la chimie analytique [4]. Les spectroscopies infrarouge et Raman, bien que faisant appel à des phénomènes différents, sont souvent exploitées ensemble pour leur aspect complémentaire. L'analyse des spectres consiste à retrouver les nombres d'onde, les intensités et les aires des raies ; cela permet d'identifier les molécules présentes dans l'échantillon analysé, ainsi que leur dosage et leur configuration structurale. Or, l'analyse de ces signaux est très souvent perturbée par le fait que les raies ne sont pas infiniment fines et peuvent donc se chevaucher lorsqu'elles sont trop proches. Par conséquent, l'analyse est plus délicate. Il arrive également que des phénomènes de fluorescence, de lumière extérieure, etc. créent une *ligne de base* qui ajoute au spectre une courbe dont l'intensité peut être parfois très importante. Cette ligne de base rend alors la recherche des raies plus difficile puisqu'elle est inconnue et qu'elle introduit une variation non linéaire de l'intensité du spectre.

Un deuxième aspect de la collaboration intervient lorsque l'expérience est menée en faisant varier un paramètre physique (pression, pH, température, ...). Pour chaque valeur du paramètre, un spectre est enregistré ; celui-ci revient en fait à une combinaison linéaire des spectres élémentaires de chaque élément de l'échantillon analysé. Le but est de retrouver ces différents spectres ainsi que les coefficients de mélange. Le problème est traité dans un cadre de séparation de sources dans la thèse de mon collègue et néanmoins ami Saïd Moussaoui [85].

Cette thèse a donc pour objectif de résoudre deux problèmes.

Le premier consiste en l'élimination de la ligne de base. En général, on suppose que la ligne de base évolue plus lentement que les raies du spectre. C'est d'ailleurs la seule connaissance que l'on ait car il n'existe pas de ligne de base unique : deux utilisateurs peuvent en trouver deux différentes dans le même spectre. En fait, cela dépend du signal que l'on considère comme étant d'intérêt ; un signal d'intérêt pour une personne peut être considéré comme un simple bruit pour une autre. C'est la raison pour laquelle la méthode d'estimation doit être supervisée, en permettant l'ajustement de paramètres afin d'estimer la ligne de base attendue. Dès lors, elle doit être également rapide, car l'ajustement de paramètres nécessite de procéder à des essais successifs jusqu'à obtenir un résultat qui satisfasse l'utilisateur. Enfin, comme cette méthode est sensée être appliquée sur des signaux issus de plusieurs techniques de spectroscopie dont les caractéristiques de forme des raies sont différentes, elle ne doit pas être fondée sur une modélisation précise des raies.

Le deuxième problème traité dans cette thèse est celui de l'estimation des raies du spectre. C'est typiquement un problème inverse mal posé. Aussi, il est indispensable de restreindre l'espace des solutions, par exemple en apportant un maximum d'informations sur les variables recherchées. Comme les spectres infrarouge et Raman sont des spectres optiques, une première information importante est de considérer que les raies sont toutes positives. En outre, il est communément admis que ces raies ont des formes connues (gaussiennes pour les spectres infrarouge et lorentziennes pour les spectres Raman) : cette information doit également être introduite dans la méthode. Enfin, et contrairement à

l'estimation de la ligne de base, il est souhaitable d'adopter une méthode non supervisée car l'estimation des raies ne dépend pas de l'utilisateur.

Cette thèse couvre un champ assez large des méthodologies du traitement du signal puisque nous proposons d'une part une approche déterministe pour l'estimation de la ligne de base (en minimisant une fonction-coût grâce à un algorithme d'optimisation déterministe), d'autre part une approche probabiliste pour l'estimation du spectre de raies (grâce à l'inférence bayésienne et des algorithmes d'optimisation stochastiques).

Les premières méthodes de régularisation de problèmes inverses furent déterministes : elle consistaient soit à contrôler la dimension de la solution, soit à ajouter au critère d'adéquation aux données un terme de régularisation. Malheureusement, hormis pour des cas simples, il n'est pas aisé de minimiser le critère composite. Parallèlement, l'inférence bayésienne est une méthodologie cohérente et homogène pour coder toute l'information connue sur les quantités d'intérêt, fût-elle qualitative. Toutefois, le fait d'accumuler un grand nombre d'informations a priori complique l'expression du critère et il faut alors souvent faire appel à des techniques d'optimisation stochastiques mais qui sont lentes à converger. Bien que ces deux approches paraissent différentes, il existe un lien entre elles, puisqu'il suffit de considérer une fonction strictement monotone du critère composite pour obtenir une loi a posteriori. En général, cette fonction est une exponentielle décroissante, ce qui convient bien aux modèles linéaires dont le bruit est i.i.d. et gaussien. Par ailleurs, l'estimateur du maximum a posteriori correspond au minimiseur du critère composite. Mais l'approche déterministe ne possède pas tous les outils de l'approche bayésienne, comme par exemple la possibilité de déterminer les coefficients de régularisation et les hyperparamètres du critère, ou d'utiliser des méthodes d'optimisation stochastiques.

Plusieurs raisons expliquent cette volonté de travailler dès le départ avec deux méthodologies différentes du traitement du signal. Tout d'abord, cela a permis d'étudier un large champ de techniques et d'en être plus familier. De plus, l'intérêt de traiter les deux problèmes séparément et non conjointement est motivé par le fait que, parfois, seule une correction de ligne de base est nécessaire si la résolution du spectre est suffisante pour l'analyse. En outre, une approche déterministe pour l'estimation de la ligne de base permet d'obtenir une méthode rapide, sans pour autant dégrader le résultat, tandis qu'une approche probabiliste pour l'estimation du spectre de raies est indispensable pour utiliser toute l'information existante car le nombre d'inconnues est grand par rapport au nombre de données.

Par ailleurs, nous avons souhaité traiter ces deux problèmes dans leur globabilité, c'est-à-dire considérer la modélisation, l'optimisation, l'estimation et finalement l'implémentation des méthodes proposées. En effet, la qualité d'une méthode dépend non seulement de son aspect théorique, mais également de sa mise en œuvre.

Ce document est organisé en quatre chapitres.

Le chapitre 1 est tout d'abord consacré aux motivations de ce travail de thèse : nous justifions l'intérêt de développer des méthodes numériques pour l'analyse de signaux spectroscopiques. Ensuite, le principe de la spectroscopie est présenté d'un point de vue général puis en s'attardant aux spectroscopies infrarouge et Raman. Enfin, nous considérons la modélisation du système ; ce modèle constituera par la suite le point de départ du développement des méthodes et servira également à créer des signaux simulés sur lesquels seront testées les approches proposées.

Le chapitre 2 traite du problème de la correction de la ligne de base. Comme nous l'avons dit, l'estimation de la ligne de base n'est pas une fin en soi, mais est nécessaire pour obtenir une estimation correcte des paramètres du spectre. Nous proposons une méthode déterministe en considérant la modélisation répandue qui suppose que la ligne de base peut être estimée par un polynôme d'ordre faible. Ce polynôme est obtenu en minimisant une fonction-coût choisie de manière à affecter un coût faible aux raies du signal, ce qui revient à diminuer leur influence. Deux fonctions-coûts sont étudiées (fonction de Huber et parabole tronquée) et nous montrons que, dans le cas de spectres optiques où les raies sont toutes positives, les versions asymétriques des fonctions-coûts donnent les meilleurs résultats. La minimisation des fonctions-coûts est réalisée grâce à l'algorithme de minimisation semi-quadratique

LEGEND : cela permet d'obtenir une méthode simple et rapide sans avoir à considérer un quelconque modèle sur les raies et le bruit du spectre. Le choix des paramètres de la méthode est également discuté, et une étude du comportement de l'algorithme lorsque les paramètres du spectre varient est donnée. La méthode est finalement validée sur des spectres infrarouge et Raman expérimentaux.

Le chapitre 3 considère, après correction de la ligne de base, l'estimation des raies du spectre. Grâce à deux hypothèses simplificatrices, le problème est traité dans le cadre désormais classique de la déconvolution impulsionnelle myope non supervisée. En outre, comme le nombre d'inconnues est grand par rapport au nombre de données, nous choisissons de nous placer dans un cadre bayésien afin d'apporter le maximum de connaissance a priori. Plus précisément, nous proposons une approche hiérarchique, c'est-à-dire que les paramètres des lois a priori sont eux-mêmes des variables du problème pour lesquelles une modélisation probabiliste est nécessaire. Aussi, le spectre est modélisé par un processus Bernoulli-gaussien (à support positif) convolué avec une réponse impulsionnelle de forme connue mais de paramètres de forme à estimer. À cet égard, la simulation d'une loi normale tronquée (ou à support positif) est indispensable : l'annexe C propose une méthode originale qui utilise un algorithme d'acceptation-rejet mixte. Les méthodes MCMC sont parfaitement justifiées pour l'optimisation de la loi a posteriori ; en particulier, l'échantillonneur de Gibbs s'avère naturel et facile à implémenter du fait du caractère hiérarchique de la modélisation. Finalement, un nouvel estimateur du signal impulsionnel permet d'obtenir un résultat plus pertinent que les estimateurs classiques (comme par exemple l'estimateur de l'espérance a posteriori). Une étude sur des spectres simulés est ensuite présentée. En définitive, cette méthode s'avère relativement efficace. Toutefois, les hypothèses simplificatrices impliquent deux inconvénients majeurs qui sont le fait que les raies estimées sont situées forcément en des nombres d'onde discrets et qu'elles sont toutes identiques (aux nombre d'onde et intensité près).

Le chapitre 4 propose, pour pallier les inconvénients de la méthode précédente, de modéliser le spectre par la somme de plusieurs motifs de forme connue, mais dont le nombre est inconnu : le problème devient alors un problème de décomposition en motifs élémentaires. Ce type de modèle est très difficile à résoudre mais l'avènement récent des méthodes MCMC a permis de proposer de nouvelles approches. Comme le nombre de motifs peut varier, l'utilisation d'une méthode MCMC classique est impossible. Ainsi, une grande majorité des travaux propose d'utiliser des algorithmes qui permettent de se déplacer entre des espaces de dimensions différentes ; l'algorithme RJMCMC (*Reversible Jump MCMC*) est le plus souvent proposé. Nous préférons utiliser une approche originale qui s'inspire de l'idée du processus Bernoulli-gaussien : on suppose qu'il existe un nombre maximal de motifs, qui peuvent ou non apparaître dans le spectre. Par conséquent, la dimension du problème est fixe et permet d'utiliser l'échantillonneur de Gibbs. Dès lors, cette approche est une alternative intéressante à la déconvolution impulsionnelle. Elle est en outre plus rapide que la précédente car il y a moins de variables à estimer. Cependant, elle doit faire face à un problème commun à toutes les méthodes MCMC appliquées à la décomposition en motifs élémentaires qui est le problème de permutation d'indices : une revue des méthodes existantes est présentée, puis un nouvel algorithme est proposé. Des simulations sont ensuite effectuées sur des spectres où les raies sont identiques puis différentes. Enfin, la méthode est appliquée sur les spectres expérimentaux du chapitre 2 dont la ligne de base a été corrigée. Par rapport à la déconvolution impulsionnelle myope, cette méthode a l'avantage de pouvoir considérer des raies de formes différentes qui de plus peuvent être situées hors des nombres d'onde discrets, mais également d'avoir moins de variables à estimer, ce qui réduit l'espace des solutions.

Chapitre 1

Principe et modélisation des spectres infrarouge et Raman

1.1 Introduction

La spectroscopie optique est la science qui produit, mesure et analyse les spectres électromagnétiques caractérisant la matière : elle consiste à obtenir des informations sur un échantillon de matière à partir de son interaction avec un rayonnement électromagnétique. C'est une méthode non destructive dont les applications typiques sont l'identification chimique des substances, leur analyse quantitative et la détermination de leur structure moléculaire [32].

La spectroscopie optique regroupe les spectroscopies infrarouge, Raman, UV-visible, ou de fluorescence. Dans cette thèse, nous considérons seulement le cas des spectroscopies infrarouge et Raman.

En spectroscopie infrarouge et Raman, chaque photon du rayonnement électromagnétique incident est, en fonction de sa longueur d'onde (c'est-à-dire de son énergie), à l'origine des vibrations des molécules : on parle alors de spectroscopie vibrationnelle. Ces vibrations sont caractéristiques des liaisons chimiques entre les atomes de la molécule et se traduisent sur le spectre par l'apparition ou la disparition de plusieurs longueurs d'ondes. Un spectre est donc caractéristique des molécules présentent dans l'échantillon. En d'autres termes, c'est une « image » des molécules de l'échantillon.

Le but de cette thèse est de développer des méthodes numériques afin d'améliorer l'analyse du spectre. Ainsi, nous présentons dans la section 1.2 les motivations qui ont conduit à ce travail de recherche : quel est l'intérêt de développer des méthodes numériques pour estimer la ligne de base et le spectre de raies ?

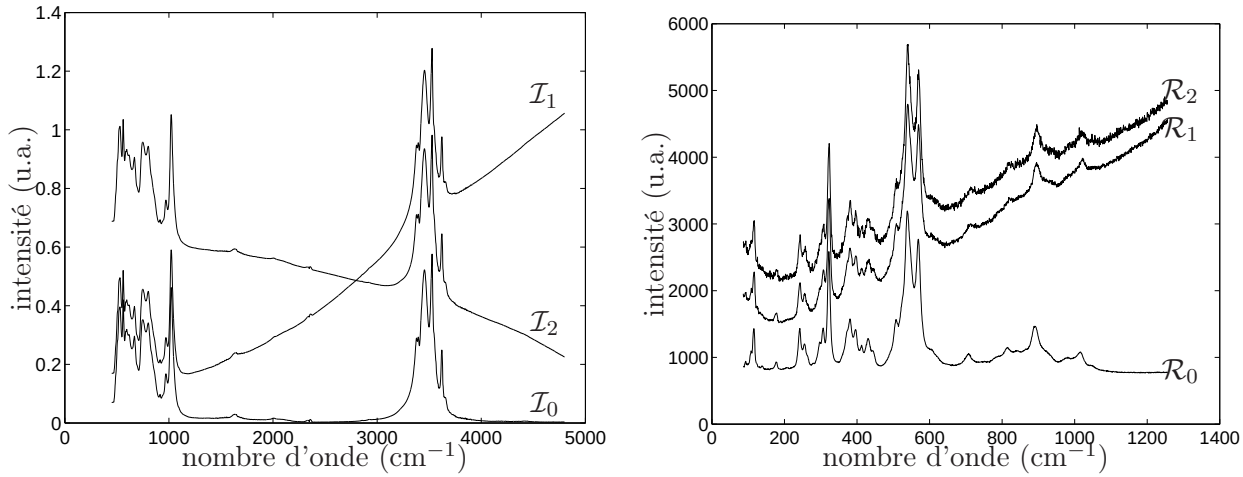
Les principes de bases de la spectroscopie vibrationnelle sont ensuite présentés dans la section 1.3 : nous étudions les propriétés du photon, puis montrons que les vibrations des molécules peuvent être décomposées en mouvements simples (appelés modes normaux de vibration) caractéristiques des liaisons chimiques. Nous détaillons les principales caractéristiques des spectroscopies infrarouge et Raman dans les sections 1.4 et 1.5. Enfin, nous montrons dans la section 1.6 qu'il est très souvent possible qu'un signal parasite (la ligne de base) apparaisse dans le spectre. Ce signal gêne l'analyse du spectre et doit donc être supprimé : c'est l'objectif du chapitre 2.

Les sections précédentes ont pour but non seulement de présenter le problème, mais également d'introduire le modèle choisi pour le système. Ce modèle servira à poser les hypothèses des méthodes proposées dans cette thèse, mais également à fournir un simulateur grâce auquel ces méthodes pourront être testées et comparées. Ce modèle est présenté dans la section 1.7.

1.2 Motivations

L'objectif de cette thèse est de développer des méthodes numériques pour faciliter l'analyse des spectres infrarouge et Raman. Les méthodes développées dans cette thèse seront validées sur des

spectres expérimentaux de gibbsite fournis par le LCPME (figure 1.1).



(a) Spectres infrarouge de la gibbsite.

(b) Spectres Raman de la gibbsite.

FIG. 1.1 – Spectres infrarouge et Raman de la gibbsite.

La gibbsite ($Al(OH)_3$) a été préparée par oxydation de poudre d'aluminium dans une solution d'hydroxyde de sodium (voir [62, 88] pour des détails sur la préparation).

Les spectres infrarouge ont été obtenus grâce à un spectromètre infrarouge à transformée de Fourier (Perkin Elmer system 2000) par transmission à travers des pastilles de gibbsite. Ces pastilles de 15 mm de diamètre ont été préparées par dilution de gibbsite dans un milieu non-absorbant (KBr à 0,1 % en poids de gibbsite) sous une pression de 10 MPa ; les spectres ont été enregistrés sur une bande de 450 à 4800 cm^{-1}. Le temps d'acquisition a été de une minute par spectre. En pratique, on retire au spectre infrarouge de l'échantillon un signal de référence correspondant au spectre d'une pastille pure de KBr. Le spectre est alors présenté en unité d'absorbance : $A = -\log_{10}(I/I_0)$ où I est l'intensité du spectre de l'échantillon et I_0 celle du spectre de référence. Aussi, trois spectres de référence on été obtenus avec des pastilles de KBr plus ou moins pures, ce qui a permis d'obtenir trois spectres infrarouges composés du même spectre d'absorption mais avec des lignes de base différentes. Ces trois spectres sont dénotés $\mathcal{I}_0$, $\mathcal{I}_1$ et $\mathcal{I}_2$ et sont présentés sur la figure 1.1(a).

Les spectres Raman ont été enregistrés grâce à un spectromètre monochromateur triple-soustractif (Jobin Yvon T64000) équipé d'un microscope confocal. Le détecteur est une caméra CCD refroidie par azote liquide. Les spectres Raman ont été obtenus en excitant l'échantillon de gibbsite avec un laser de longueur d'onde 514,3 nm. Le signal mesuré correspond au signal rétrodiffusé, c'est-à-dire au rayon lumineux qui fait un angle de 0° avec le rayon incident, dispersé sur un réseau de 1800 traits par millimètre afin d'obtenir une résolution d'échantillonnage de 0,6 cm^{-1}. Les spectres Raman ont été obtenus en mesurant le même échantillon pendant des durées différentes, afin d'obtenir différentes amplitudes du bruit. Trois spectres ont été enregistrés : un échantillon de gibbsite pure et deux échantillons de gibbsite pollués. Ces échantillons pollués ont été obtenus en laissant la poudre de gibbsite dans une atmosphère de fumée de cigarette, ils ont donc (à peu près) les mêmes signaux Raman mais des lignes de base de fluorescence différentes. Ces trois spectres sont dénotés $\mathcal{R}_0$, $\mathcal{R}_1$ et $\mathcal{R}_2$ et sont présentés sur la figure 1.1(b).

Les spectroscopies infrarouge et Raman sont intéressantes dans le sens où ce sont des techniques complémentaires. En effet, en considérant une gamme de longueur d'onde donnée, les raies absentes du spectre infrarouge apparaissent généralement dans le spectre Raman et vice-versa. C'est la raison pour laquelle ces deux techniques de spectroscopie sont souvent utilisées conjointement pour analyser des échantillons de matière.

L'analyse d'un signal spectroscopique permet d'obtenir trois types de renseignements sur l'échantillon analysé :

- tout d'abord, elle permet de réaliser l'*identification* des molécules qui composent l'échantillon. Chaque molécule possède sa propre « empreinte spectrale », c'est-à-dire qu'à chaque molécule correspond un spectre particulier qui permet de la différencier d'une autre molécule ;
- ensuite, elle permet d'effectuer un *dosage* de chaque molécule de l'échantillon : en effet, les intensités des spectres permettent de déterminer les rapports en quantité entre chaque molécule ;
- enfin, elle permet d'obtenir une *interprétation structurale* qui consiste à caractériser l'organisation des molécules entre elles. Cette information est notamment obtenue en mesurant les largeurs à mi-hauteur de chaque raie (un échantillon amorphe aura des raies plus larges que le même échantillon sous forme cristalline).

Pour obtenir ces trois informations, il est nécessaire de pouvoir caractériser les nombres d'onde, intensités et largeur à mi-hauteur de chaque raie. Or, en fonction du bruit, de la résolution du spectre et de la largeur des raies, l'analyse du spectre peut être difficile, en particulier parce que les raies sont susceptibles de se chevaucher. Aussi, l'objectif principal de ce travail consiste à développer une méthode qui permette de séparer les raies, ce qui revient à augmenter la résolution du spectre.

Cependant, l'estimation des raies peut être perturbée lorsque le spectre fait apparaître une *ligne de base*. Ce signal est dû à des phénomènes intrinsèques ou étrangers à l'échantillon qui ajoutent au spectre un « fond » irrégulier (voir section 1.6). Dès lors, pour estimer correctement les raies du spectre, il s'avère indispensable d'éliminer la ligne de base du spectre[1]. Il existe bien sûr des procédés expérimentaux, mais ils ne sont pas toujours complètement efficaces. C'est pourquoi l'estimation de la ligne de base est le second objectif de ce travail.

L'approche optimale consiste à estimer la ligne de base conjointement avec l'estimation des paramètres du spectre [48]. À cet égard, l'approche bayésienne, couplée à des techniques d'optimisation stochastiques, s'avère particulièrement bien adaptée. Néanmoins, nous proposons ici d'estimer la ligne de base séparément à l'aide d'une méthode déterministe. En effet, il n'est pas toujours nécessaire de procéder à une estimation des raies du spectre et, parfois, seule la correction de la ligne de base est intéressante. En outre, la méthode proposée dans cette thèse est très rapide et permet à l'utilisateur de fixer ses propres paramètres, car, comme nous le verrons dans le chapitre 2, la ligne de base recherchée peut être différente d'une personne à l'autre. Une approche bayésienne peut également laisser des paramètres libres, mais, du fait de l'utilisation de méthodes MCMC, cette approche serait beaucoup trop longue si les valeurs de ces paramètres sont fixées par essais/erreurs.

Pourquoi ne pas utiliser une base de données des modes normaux de vibration ?

Nous allons montrer dans les sections suivantes que chaque raie du spectre infrarouge ou Raman est la conséquence d'une vibration particulière d'une molécule de l'échantillon étudié : le nombre d'onde de la raie est liée à la fréquence de la vibration. Ainsi, nous pourrions créer une base de données des fréquences de vibrations, ce qui permettrait de retrouver les molécules d'un échantillon inconnu à partir de son spectre. Malheureusement, cela n'est pas possible.

La raison la plus évidente est que la base de données devrait contenir toutes les fréquences de toutes les molécules existantes. Elle serait alors de taille rédhibitoire pour permettre non seulement de la stocker, mais également de la consulter en un temps raisonnable.

En outre, pour connaître parfaitement la fréquence de chaque vibration, il faut évidemment être capable de la calculer. Or, il n'existe pas de modèle analytique représentant parfaitement le modèle anharmonique qui régit les vibrations des atomes dans les molécules. De nombreuses approximations mathématiques ont été proposées (les plus courantes sont données dans [4]), mais aucune n'est assez précise pour pouvoir être fournir des résultats exacts.

Néanmoins, il existe des tables expérimentales, mais il est difficile de les utiliser car une molécule

[1] « *a proper estimate of the area of peaks can only be achieved through a good approximation of the background* » [48].

peut avoir un spectre légèrement différent en fonction de son environnement. Par ailleurs, l'étude des vibrations nous indique si une raie peut apparaître dans le spectre, mais ne donnent pas son intensité, qui peut être nulle ou très faible [54, p. 123]. Ainsi, si un spectre contient des raies indétectables, il sera difficile, voire impossible, d'identifier la substance en comparant le spectre à la base de données.

De toute façon, pour pouvoir comparer le spectre mesuré avec une quelconque table, il est indispensable d'éliminer la ligne de base et, parfois, d'améliorer la résolution du spectre, motivant l'utilisation des méthodes proposées ici.

1.3 Bases de la spectroscopie vibrationnelle

1.3.1 Photon et spectre électromagnétique

Le photon est la particule élémentaire du rayonnement électromagnétique. L'énergie E (en joule) d'un photon est associée à sa fréquence ν (en hertz) selon la relation :

$$E = h\nu, \tag{1.1}$$

où h est la constante de Planck, égale à $6,626\,1 \times 10^{-34}$ J·s. La relation classique entre vitesse, temps et longueur permet de lier la fréquence du photon à sa longueur d'onde λ (en centimètres) :

$$\nu = c/\lambda,$$

où c est la vitesse de la lumière dans le milieu considéré (dans la majorité des cas, on considère la vitesse de la lumière dans le vide, égale à $2,997\,924\,580 \times 10^{10}$ cm·s^{-1}). Très souvent, on utilise plutôt le nombre d'onde $\overline{\nu}$ (en cm^{-1}) qui correspond à l'inverse de la longueur d'onde.

Le photon peut être considéré comme une onde électromagnétique constituée d'un champ électrique $\mathbf{E}$ et d'un champ magnétique $\mathbf{B}$, perpendiculaires entre eux et de modules :

$$E = E_0 \sin(2\pi\nu t - \phi),$$
$$B = B_0 \sin(2\pi\nu t - \phi),$$

où E_0 et B_0 sont les amplitudes des champs et ϕ une constante de déphasage. Ainsi, les champs oscillent en phase à une pulsation de $2\pi\nu$. En général, l'interaction du photon avec la matière se fait par la composante électrique.

La figure 1.2 présente le spectre électromagnétique, des rayons gamma (de faible longueur d'onde, très énergétiques) aux ondes radios (de grande longueur d'onde, peu énergétiques)[2]. On distingue habituellement l'infrarouge proche (P) qui utilise le rayonnement compris entre 4000 et 12 000 cm^{-1}, moyen (M) entre 400 et 4000 cm^{-1} ou lointain (L) entre 10 et 400 cm^{-1}.

1.3.2 Modes normaux de vibration

Une molécule de N atomes a $3N$ degrés de liberté. En effet, la molécule entière possède trois mouvements de translation selon les trois dimensions et trois mouvements de rotation autour de son centre de gravité. En considérant alors l'analogie classique avec la mécanique, on peut modéliser la molécule par un ensemble de masses représentant les atomes et de ressorts représentant les liaisons entre atomes (on peut considérer que les mouvements entre atomes sont des mouvements de compression ou d'extension). Un choc sur cette molécule sera la conséquence de mouvements complexes entre les atomes (mouvement de Lissajous). On peut montrer que ces mouvements sont la combinaison de $3N-6$ degrés de liberté pour lesquels les atomes vibrent selon des mouvements harmoniques simples qui sont

[2] L'énergie est représentée en électronvolt (1 eV $\approx 1,602 \times 10^{-19}$ J).

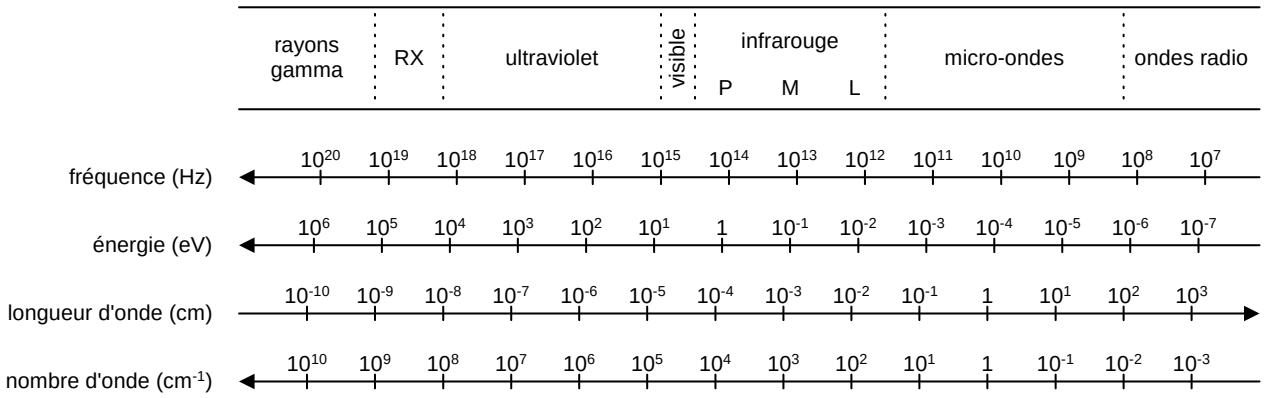


FIG. 1.2 – Spectre électromagnétique.

appelé *modes normaux de vibration* [23]. Toutefois, si la molécule est linéaire, elle ne possède que deux mouvements de rotation, et, par conséquent, $3N - 5$ modes normaux de vibration.

Pour fixer les idées, considérons la molécule de CO_2. Cette molécule étant linéaire, elle possède $3N - 5$ modes normaux de vibration, soit quatre modes normaux de vibration représentés figure 1.3.

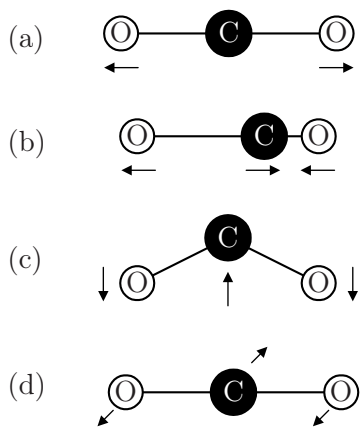


FIG. 1.3 – Les quatre modes normaux de vibration du dioxyde de carbone [7, 23].

Les quatres modes normaux de vibration de la molécule de CO_2 sont :
- une élongation symétrique, où les deux atomes d'hydrogène vibrent symétriquement de part et d'autre de l'atome de carbone ;
- une élongation anti-symétrique, où l'un des atomes d'hydrogène s'approche de l'atome de carbone quand l'autre s'en éloigne ;
- une déformation de l'angle entre les deux liaisons C–O dans le plan de la feuille ;
- une déformation de l'angle entre les deux liaisons C–O dans le plan perpendiculaire à la feuille.

L'élongation symétrique vibre à une fréquence de 1337 cm^{-1} et l'élongation anti-symétrique à une fréquence de 2349 cm^{-1}. Les deux derniers modes de vibration sont en fait identiques et vibrent à la même fréquence de 667 cm^{-1} : ils ne peuvent donc pas être différenciés ; on les appelle *modes dégénérés*.

1.3.3 Énergie vibrationnelle des molécules

On peut distinguer quatre types d'énergie moléculaire [23] :
- l'énergie de translation, qui est l'énergie fournie ou nécessaire à la molécule pour se mouvoir selon les trois dimensions ;
- l'énergie de rotation (rayonnement micro-ondes) qui permet à la molécule de tourner autour de son centre de gravité ;

– l'énergie vibrationnelle (rayonnement infrarouge) qui est l'énergie résultante des vibrations des liaisons entre atomes ;
– l'énergie électronique (rayonnement visible ou ultraviolet) qui correspond à l'énergie des électrons dans les orbitales moléculaires.

Considérons à présent l'énergie vibrationnelle des molécules, conséquence des vibrations des liaisons inter-atomiques et responsable des phénomènes exploités en spectroscopie infrarouge et Raman.

En utilisant à nouveau l'analogie avec la mécanique, une molécule diatomique est modélisée par deux masses reliées par un ressort : elle se comporte donc comme un oscillateur dont la force de rappel f est proportionnelle à la différence entre la distance r entre les deux atomes et la position d'équilibre r_0 :

$$f = -k\Delta r,$$

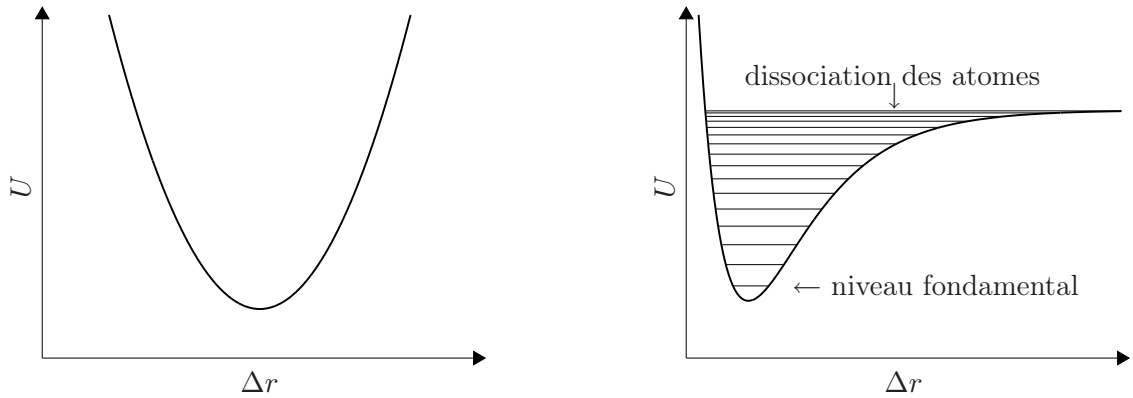
où $\Delta r = r - r_0$ et k est la constante de rappel. L'énergie potentielle de vibration U est régie par l'équation :

$$f = -\frac{dU}{d(\Delta r)}$$

qui donne après intégration :

$$U = \frac{k}{2}(\Delta r)^2 + c,$$

où c est une constante. L'énergie potentielle U est donc une parabole fonction de la distance entre les atomes et minimale en r_0 (figure 1.4(a)).



(a) Oscillateur harmonique. (b) Oscillateur anharmonique.

FIG. 1.4 – Représentation schématique de l'énergie vibrationnelle de liaison entre deux atomes.

En utilisant la seconde loi du mouvement de Newton, on obtient [4] :

$$\Delta r = r_0 \cos(2\pi\nu t + \phi)$$

où ν est la fréquence de vibration et ϕ est la phase qui dépend des conditions initiales. Le mouvement décrit par cette équation est un exemple de mouvement harmonique simple.

Cependant, le modèle de l'oscillateur harmonique est trop simpliste car il reste limité aux faibles déplacements des atomes. En particulier, il suppose que les liaisons sont parfaitement élastiques. Or, ces liaisons peuvent se briser lorsque l'amplitude des vibrations est trop importante.

Un modèle plus complexe, et donc mieux adapté à la réalité, est celui de l'oscillateur anharmonique dont l'énergie potentielle est représentée figure 1.4(b) [4, 23].

Cependant, l'énergie de vibration de la liaison ne peut pas prendre toutes les valeurs de la courbe. En effet, la mécanique quantique restreint l'énergie potentielle de la molécule à des valeurs discrètes

appelés *niveaux d'énergie*. Nous ne détaillerons pas ici l'aspect théorique ; pour plus de détails, le lecteur pourra se référer à [4, 23]. Les niveaux d'énergie pour l'oscillateur anharmonique sont représentés par les lignes horizontales de la figure 1.4(b).

Le niveau fondamental correspond au niveau de plus basse énergie, où la molécule est au repos. Tous les atomes y sont statiques et à leur position d'équilibre. Cet état n'existe que lorsque la température correspond au zéro absolu. La valeur limite des niveaux d'énergie correspond à la dissociation des deux atomes.

La transition d'un niveau de grande énergie à un niveau d'énergie plus faible s'accompagne de l'émission d'un photon dont l'énergie (la fréquence) correspond à la différence d'énergie entre les deux niveaux. Inversement, une liaison qui passe d'un état de faible énergie à un état d'énergie plus grande absorbe un photon dont l'énergie (la fréquence) correspond à la différence d'énergie entre les deux niveaux. Ce phénomène est à l'origine de la spectroscopie vibrationnelle qui utilise les radiations électromagnétiques pour exciter les molécules en les faisant passer d'un niveau d'énergie à un autre. En vibrant les uns par rapport aux autres, les atomes déforment les nuages électroniques, formés par les électrons gravitant autour des noyaux. Ces déformations sont à l'origine du spectre infrarouge et du spectre Raman ; les sections suivantes en détaillent le principe.

1.4 Spectroscopie infrarouge

1.4.1 Moment dipolaire

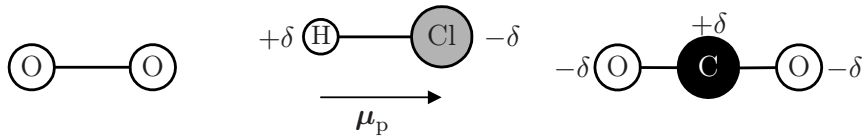
Un atome seul est électriquement neutre. Par exemple, l'atome de carbone possède six charges négatives (les électrons) et six charges positives (les protons). Or, dans les molécules, les atomes ne sont pas réellement neutres. En effet, les nuages électroniques autour des noyaux peuvent être déformés en fonction des atomes voisins. Tout se passe comme si les atomes étaient des ions dans la molécule : on parle d'électronégativité et on note $+\delta$ si l'atome est chargé positivement, $-\delta$ dans le cas contraire. Ainsi, une molécule, bien qu'électriquement neutre, possède tout de même des charges positives et négatives dues à l'hétérogénéité du nuage électronique : il y a donc des dipôles au sein de la plupart des molécules. Le *moment dipolaire permanent* d'une molécule mesure l'asymétrie de répartition de ces charges. C'est une quantité vectorielle.

Lorsqu'une charge positive $+q$ et une charge négative $-q$ sont séparées par une distance $\boldsymbol{d}$, le moment dipolaire est égal à :

$$\boldsymbol{\mu}_\mathrm{p} = q\boldsymbol{d},$$

où q est en coulombs, $\boldsymbol{r}$ en mètres et $\boldsymbol{\mu}_\mathrm{p}$ en debyes ($1\ \mathrm{D} \approx 3,335\,64 \times 10^{-30}$ C·m).

La figure 1.5 représente les molécules de dioxygène, d'acide chlorhydrique et de dioxyde de carbone, ainsi que leurs charges et le moment dipolaire correspondant. La molécule de dioxygène ayant deux



(a) (b) (c)

Fig. 1.5 – Moment dipolaire des molécules de dioxygène (a), d'acide chlorhydrique (b) et de dioxyde de carbone (c).

atomes de même électronégativité, elle ne possède pas de moment dipolaire. Au contraire, la molécule d'acide chlorhydrique a deux atomes d'électronégativités différentes, et donc possède un moment dipolaire. Enfin, compte tenu de la disposition linéaire de ses atomes, la molécule de dioxyde de carbone n'a pas de moment dipolaire au repos, bien qu'elle ait des atomes d'électronégativités différentes.

1.4.2 Principe de la spectroscopie infrarouge

C'est la variation du moment dipolaire permanent d'une molécule, due aux vibrations des liaisons, qui est à l'origine de son spectre infrarouge.

Les molécules vibrent naturellement suivant tous leurs modes de vibration, mais avec des amplitudes très faibles. Or, comme nous l'avons vu dans la section 1.3.1, le photon possède une composante électrique sinusoïdale. Si la fréquence du photon correspond à la fréquence de vibration d'un des modes normaux de la molécule, alors la molécule va entrer en résonance et vibrer avec une amplitude très grande. En d'autres termes, un photon dont l'énergie est égale à l'énergie nécessaire à la molécule pour passer d'un état de basse énergie à un état excité sera absorbé et son énergie sera transformée en énergie de vibration. La figure 1.6 schématise ce phénomène : seul le photon dont l'énergie $h\nu_2$ est égale à l'énergie de transition $E_2 - E_1$ est absorbé. Par conséquent, le photon absorbé fera défaut dans

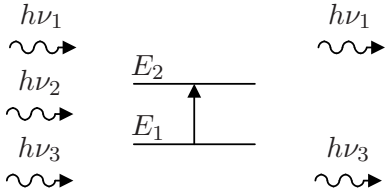


FIG. 1.6 – Absorption infrarouge.

le rayonnement transmis : le spectre infrarouge est donc un spectre d'absorption.

L'absorption de certains photons incidents fait apparaître dans le spectre infrarouge de la molécule des raies correspondant aux photons qui n'ont pas été transmis. Cette absorption est caractéristique des liaisons entre atomes du produit considéré puisque chaque mode normal de vibration correspond à un mouvement unique de la molécule : il existe donc une correspondance directe entre la fréquence de radiation absorbée et la structure de la molécule.

Par exemple, la molécule d'acide chlorhydrique ayant deux atomes, elle n'a donc qu'un seul mode de vibration ($3N - 5 = 1$) qui est l'élongation symétrique. Lors de cette vibration, la distance $\boldsymbol{d}$ entre les atomes varie, impliquant une variation du moment dipolaire : cette vibration est alors active en infrarouge.

Pour le dioxygène, la molécule a également un seul mode de vibration mais, comme ces atomes sont de même électronégativité, le moment dipolaire permanent ne varie pas : cette vibration n'est donc pas active en infrarouge.

Dans le cas de la molécule de dioxyde de carbone, l'élongation symétrique ne crée pas de variation de moment dipolaire, alors que l'élongation anti-symétrique en implique une : ce mode sera donc visible sur le spectre infrarouge, au même titre que les déformations d'angle.

1.4.3 Instrumentation

Un spectromètre souvent utilisé pour obtenir le spectre infrarouge d'un échantillon est l'interféromètre de Michelson dont le schéma de principe est représenté figure 1.7 (le schéma est simplifié au maximum afin de ne faire apparaître que les éléments essentiels). La source de lumière est une source polychromatique dont le faisceau de lumière est séparé en deux grâce à la séparatrice. Chaque faisceau est alors renvoyé via les miroirs sur l'échantillon. L'un des deux miroirs peut se déplacer, permettant ainsi d'obtenir un interférogramme, c'est-à-dire un signal représentant les franges d'interférences obtenues grâce à la différence de chemin optique du faisceau de lumière. Au centre de l'interférogramme est placé un détecteur (en général thermique) qui mesure l'intensité du signal en fonction du déplacement du miroir : on obtient donc un signal qui est échantillonné implicitement par le procédé mécanique.

Ce signal est alors numérisé, puis une transformée de Fourier est calculée pour obtenir le spectre infrarouge. Auparavant, le signal est multiplié par une fonction d'*apodisation*, ce qui a pour effet de diminuer les lobes des sinus cardinaux dus au fenêtrage rectangulaire, mais qui a pour conséquence

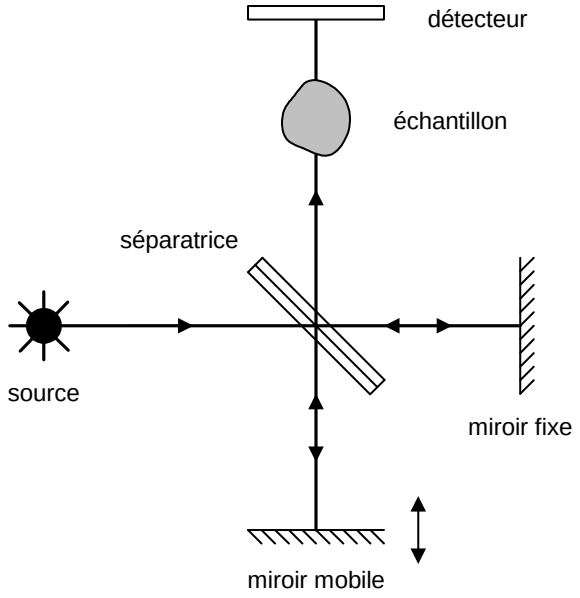


FIG. 1.7 – Schéma de principe de la spectroscopie infrarouge.

d'élargir les raies du signal. Malgré tout, cette technique permet d'obtenir un signal dont le rapport signal-à-bruit est très grand.

1.5 Spectroscopie Raman

1.5.1 Polarisabilité

Considérons dans un premier temps un simple atome, constitué de sa charge positive (les protons du noyau) et de sa charge négative (le nuage d'électrons). Cet atome n'a donc pas de moment dipolaire permanent. En revanche, s'il est placé dans un champ électrique $\mathbf{E}$ (par exemple entre les deux plaques d'un condensateur), les protons seront attirés par la borne négative et les électrons par la borne positive.

L'atome polarisé ainsi créé aura donc un *moment dipolaire induit* $\boldsymbol{\mu}_\mathrm{i}$ de même direction que $\mathbf{E}$:

$$\boldsymbol{\mu}_\mathrm{i} = \boldsymbol{\alpha}\mathbf{E}.$$

$\boldsymbol{\alpha}$ est appelé la polarisabilité de l'atome (en $\mathrm{C{\cdot}V^{-1}{\cdot}m^2}$). C'est un tenseur qui peut être s'écrire sous la forme d'une matrice symétrique :

$$\boldsymbol{\alpha} = \begin{pmatrix} \alpha_{xx} & \alpha_{xy} & \alpha_{xz} \\ \alpha_{yx} & \alpha_{yy} & \alpha_{yz} \\ \alpha_{zx} & \alpha_{zy} & \alpha_{zz} \end{pmatrix}.$$

Le tenseur de polarisabilité peut être représenté graphiquement sous forme d'une ellipsoïde [23, 54, 71].

De façon analogue à un atome, une molécule aura un moment dipolaire induit lorsqu'elle sera placée dans un champ électrique.

1.5.2 Principe de la spectroscopie Raman

La spectroscopie Raman est basée sur le phénomène de diffusion inélastique de la lumière, qui a été mis en évidence par Sir Chandrasekhara Venkata Raman (1888–1970, prix Nobel de physique 1930) en 1928.

Comme nous l'avons vu, les modes actifs en spectroscopie infrarouge sont les modes pour lesquels il y a une variation du moment dipolaire. Dans le cas de la spectroscopie Raman, un mode est actif lorsqu'il y a une variation de la polarisabilité au cours de la vibration. Cette variation de polarisabilité intervient lorsqu'il y a compression ou étirement du nuage électronique durant la vibration de la molécule.

Par exemple, l'élongation symétrique du dioxyde de carbone est active en Raman car la polarisabilité de la molécule change (figure 1.8). Cette vibration symétrique se situe à 1340 cm^{-1}, faisant ainsi

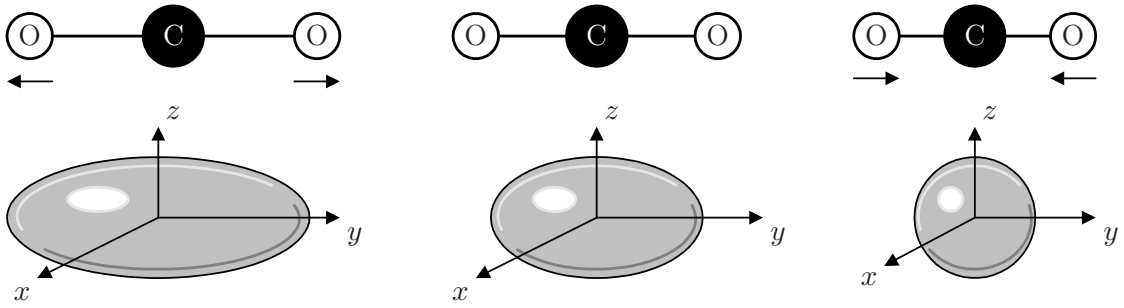


FIG. 1.8 – Polarisabilité de la molécule de dioxyde de carbone dans le cas de l'élongation symétrique.

apparaître une raie à 1340 cm^{-1} dans le spectre.

Comme en spectroscopie infrarouge, le champ électrique du rayonnement électromagnétique peut faire entrer une molécule en résonance si sa fréquence est égale à la fréquence de vibration de la molécule. En fonction de la configuration des atomes dans la molécule, cette vibration peut impliquer une déformation du nuage électronique : la molécule se comporte donc comme un dipôle oscillant (à l'instar d'une antenne) qui a pour propriété d'émettre des photons dans toutes les directions et dont la fréquence peut être différente de celle du champ électrique incident.

Dans le cas où les photons incident et diffusé ont la même énergie (diffusion élastique, le cas le plus probable), on parle de diffusion Rayleigh. Si au contraire les photons incident et diffusé n'ont pas la même énergie (diffusion inélastique), on parle de diffusion Raman [23]. En particulier,

- la diffusion Stokes correspond au cas où le photon diffusé a une énergie plus faible que le photon incident. Le photon incident a donc cédé à la molécule au repos une quantité d'énergie correspondant à l'énergie de vibration nécessaire à la transition entre un état de basse énergie et un état excité. Par conséquent, le photon diffusé a une longueur d'onde plus grande que celle du photon incident (d'après l'équation (1.1)) ;
- la diffusion anti-Stokes, au contraire, concerne le cas où le photon diffusé a une énergie plus grande que le photon incident : la molécule dans un état excité a cédé au photon incident une quantité d'énergie correspondant à la différence d'énergie de vibration entre l'état excité et l'état fondamental. La longueur d'onde du photon diffusé est alors plus petite que celle du photon incident.

La figure 1.9 illustre ces trois types de diffusions. Les diffusions Rayleigh (élastique), Stokes et anti-Stokes (inélastiques) sont représentées respectivement à gauche, au centre et à droite.

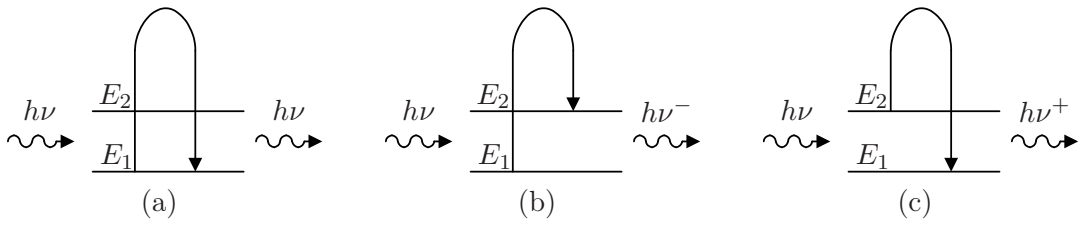


FIG. 1.9 – Diffusions Rayleigh (a), Stokes (b) et anti-Stokes (c).

Sur un spectre Raman, les raies Stokes et anti-Stokes sont à égales distances du pic de Rayleigh. En effet, pour un mode de vibration normal particulier, l'énergie perdue par le photon incident lors de la diffusion Stokes est la même (en valeur absolue) que l'énergie gagnée lors de la diffusion anti-Stokes. En revanche, l'amplitude des raies Stokes et anti-Stokes est différente car le phénomène de diffusion anti-Stokes nécessite que les molécules soient dans un état excité. Or, l'état d'excitation des molécules est lié à la température : plus la température est élevée, plus les molécules sont excitées. À température ambiante, le nombre de molécules à l'état excité est faible, c'est la raison pour laquelle les chimistes préfèrent le plus souvent travailler avec le spectre de diffusion Stokes, qui ne dépend pas de la température. Notons également que l'on peut utiliser la diffusion Stokes et anti-Stokes pour déterminer la température en mesurant le rapport des intensités de ces deux diffusions.

Par ailleurs, la diffusion Stokes ne concerne qu'un photon sur plusieurs millions. C'est pourquoi il est nécessaire d'avoir une source d'énergie puissante, comme le laser (Raman, pour sa part, utilisait la lumière du soleil à travers un télescope). De plus, il est nécessaire d'avoir une source monochromatique car le spectre consiste en une différence d'énergie entre la longueur d'onde du photon incident et les raies de diffusion Stokes et anti-Stokes. Encore une fois, le laser permet de remplir cette condition.

1.5.3 Instrumentation

Le schéma de principe de la spectroscopie Raman est représenté figure 1.10 (pour des raisons de clarté, l'instrumentation a été simplifiée à son strict minimum). Ici, une source monochromatique (un

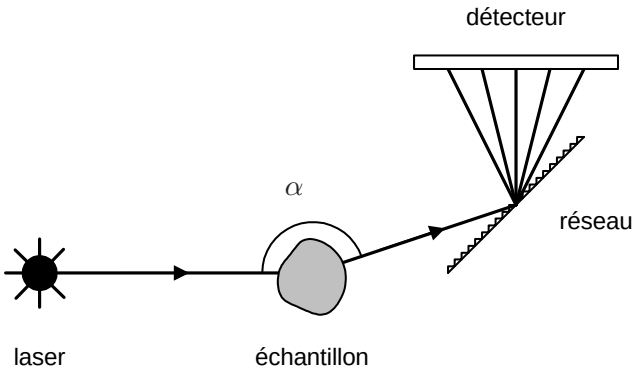


Fig. 1.10 – Schéma de principe de la spectroscopie Raman.

laser) émet un rayonnement sur un échantillon. Le rayonnement diffusé est polychromatique et est décomposé grâce au réseau en un spectre, qui est mesuré à l'aide d'une barrette CCD (*charged-coupled device*). L'un des inconvénient de la spectroscopie Raman est la faible intensité du phénomène : les temps d'acquisition sont alors très longs et peuvent atteindre plus d'une heure ! L'avantage de la rétrodiffusion, qui consiste à mesurer le rayon diffusé à $\alpha = 0^{\circ}$ du rayon incident, est de récupérer un grand nombre de photons, et donc de diminuer le temps d'acquisition.

1.6 Ligne de base

Souvent, des déviations d'intensité apparaissent sur le spectre. Elles sont causées par des phénomènes intrinsèques ou étrangers à l'expérience et sont à l'origine d'une *ligne de base*.

En spectroscopie infrarouge, la ligne de base peut être la conséquence de la diffusion d'un rayonnement infrarouge causée par les hétérogénéités ou les impuretés de l'échantillon ou tout simplement de la lumière extérieure.

En spectroscopie Raman, elle est souvent due soit à la diffusion Rayleigh, soit à la fluorescence de certaines molécules organiques provenant de l'échantillon lui-même ou de contamination extérieure (la fluorescence correspond à une émission de lumière à une certaine longueur d'onde par une substance irradiée à une autre longueur d'onde).

La ligne de base est principalement caractérisée par le fait qu'elle varie plus lentement que les raies. Cependant, la distinction entre ligne de base et raies n'est, en général, pas facile à établir. En outre, la ligne de base peut être d'intensité très grande par rapport à l'intensité des raies et, par conséquent, gêner considérablement l'analyse du spectre. Il faut alors l'éliminer par des méthodes expérimentales ou numériques pour permettre une bonne estimation des paramètres du spectre. Le chapitre 2 traite de ce sujet et propose une méthode basée sur la minimisation d'une fonction asymétrique pour estimer puis supprimer la ligne de base du spectre.

1.7 Modélisation du problème

Dans cette section, nous proposons une modélisation du problème qui permettra d'une part de poser les formulations des méthodes proposées dans cette thèse et d'autre part de créer un simulateur qui sera ensuite utilisé pour tester les méthodes d'estimation de la ligne de base et des paramètres du spectre présentées dans cette thèse. Ces spectres simulés, pour lesquels les différents signaux et paramètres sont parfaitement connus, permettront d'évaluer et de comparer les performances des méthodes.

Notons $\mathbf{y} \in \mathbb{R}^N$ le signal spectroscopique mesuré. Nous considérons la modélisation généralement admise (et justifiée ci-après) que le *spectre* $\mathbf{y}$ est la somme de trois signaux :

- un *spectre pur* $\mathbf{v} \in \mathbb{R}^N$;
- une *ligne de base* $\mathbf{z} \in \mathbb{R}^N$;
- un *bruit* $\mathbf{b} \in \mathbb{R}^N$.

Les modèles de chacun de ces trois signaux sont détaillés dans les sections suivantes. Ces signaux sont de taille N et échantillonnés sur une grille $\{1, \dots, N\}$. On note $n \in \{1, \dots, N\}$ la variable discrète correspondant au nombre d'onde. Dans certains cas, nous aurons besoin d'une variable continue du nombre d'onde, nous la noterons $t \in [1, N]$.

Dans le chapitre 2 (estimation de la ligne de base), nous regroupons dans le vecteur $\mathbf{w} \in \mathbb{R}^N$ le *spectre bruité*, c'est à dire la somme du spectre pur et du bruit :

$$\mathbf{w} = \mathbf{v} + \mathbf{b}.$$

Le spectre bruité est alors considéré comme un « bruit », puisqu'il correspond au signal ne contenant pas l'information intéressante. En revanche, dans les chapitres 3 et 4 (estimation des raies), on suppose que le spectre est sans ligne de base (ou qu'elle a été supprimée) et donc que les données correspondent au spectre bruité $\mathbf{w}$.

La figure 1.11 représente le modèle adopté dans cette thèse. Pour plus de précisions sur les notations, nous avons fait la distinction entre la somme du spectre pur et du bruit et la somme du spectre bruité et de la ligne de base. Cette distinction n'a bien sûr aucun fondement physique et ne représente que la structure du modèle qui sera exploitée ici.

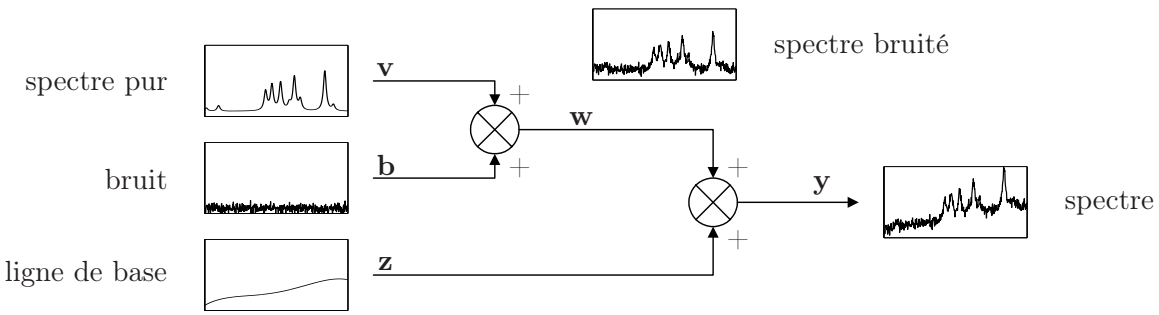


Fig. 1.11 – Modélisation du problème.

1.7.1 Modélisation du spectre pur

Un spectre pur $\mathbf{v} \in \mathbb{R}^{+N}$, sans bruit ni ligne de base, est composé d'un ensemble de « raies » (appelées aussi bandes). Ces raies ne sont jamais infiniment étroites. Plusieurs facteurs importants contribuent à la forme et à la largeur des raies dans le spectre [54] :

- l'élargissement naturel est dû au fait que les niveaux d'énergie des molécules sont plus ou moins étalés (ce phénomène s'explique à l'aide de la physique quantique [54, section 2.3.1]). C'est un élargissement homogène puisque chaque atome ou chaque molécule se comporte de la même manière, conduisant à un profil lorentzien. L'élargissement naturel est souvent très petit par rapport aux autres causes d'élargissement ;
- l'élargissement Doppler est la conséquence de l'effet Doppler dû au mouvement des particules. En effet, la fréquence à laquelle se produisent les transitions d'un état d'énergie à un autre dépend de la vitesse de la particule par rapport au détecteur. L'élargissement est hétérogène puisque les particules ne se déplacent pas toutes de la même façon. Il conduit à un profil gaussien ;
- l'élargissement dû à la pression se produit lorsque des collisions interviennent entre les molécules en phase gazeuse : il y a alors un échange d'énergie qui conduit à étaler les niveaux d'énergie des molécules. Cet élargissement est homogène et produit habituellement un profil lorentzien ;
- enfin, la résolution de l'appareil et, en spectroscopie infrarouge, le procédé d'apodisation (voir section 1.4.3) contribuent à élargir les raies du spectre.

En fonction de l'importance de ces phénomènes, le profil de la raie peut donc être considéré gaussien, lorentzien, ou une combinaison de ces deux formes. Une combinaison classique est la fonction de Voigt (convolution d'une gaussienne et d'une lorentzienne) ou, plus couramment, son approximation : la fonction de pseudo-Voigt (somme pondérée d'une gaussienne et d'une lorentzienne) [58]. Parfois, la fonction de Pearson VII est également utilisée [82].

Dans le cas de la spectroscopie Raman, on peut considérer les raies lorentziennes ; dans le cas de la spectroscopie infrarouge une forme gaussienne convient. Dans la suite de la thèse, nous choisissons arbitrairement de considérer les raies lorentziennes qui ne possèdent qu'un seul paramètre de forme (c'est-à-dire un seul degré de liberté) ; il est facile d'adapter les méthodes à une forme gaussienne. Des formes plus évoluées, telle qu'une fonction de pseudo-Voigt, seraient traitées de la même manière, à la différence qu'il y aurait plus d'un paramètre de forme. La forme simple de la lorentzienne est choisie afin de ne pas compliquer l'exposé des méthodes.

Par ailleurs, notons que la lorentzienne (ou fonction de Lorentz) et la fonction de Cauchy correspondent à la même fonction, dont la transformée de Fourier inverse est la fonction de Laplace. Le terme « fonction de Cauchy » est habituellement utilisé en statistique (on parle alors de distribution de Cauchy) et le terme « fonction de Lorentz » (ou lorentzienne) en physique.

Parfois, il est possible que les raies n'aient pas toutes la même largeur : cela est dû aux hétérogénéités de l'échantillon. Il convient donc, pour certains spectres, de considérer un modèle plus évolué qui prenne en compte ce phénomène. Dans cette thèse, nous considérons en première approximation que les raies sont toutes identiques (chapitre 3) ; un modèle plus réaliste est ensuite utilisé dans le chapitre 4.

Il importe également d'ajouter à ce modèle la fonction d'appareil qui convolue le spectre pur par une fonction que l'on peut prendre gaussienne. Cet effet résulte des déformations optiques et du détecteur du spectromètre. Toutefois cet effet est très faible et peut être négligé. De plus, dans le cas de la spectroscopie infrarouge où les raies du spectre pur sont gaussiennes, la convolution avec la fonction d'appareil gaussienne ne fera qu'élargir la largeur des raies, qui restent gaussiennes (car la convolution de deux gaussiennes est une gaussienne). Pour ces raisons, nous ne considérerons pas ici l'effet de la fonction d'appareil.

En résumé, le spectre pur est modélisé par une somme de K raies lorentziennes :

$$\forall n \in \{1, \ldots, N\}, \qquad \mathbf{v}_n = \sum_{k=1}^{K} \mathbf{a}_k \frac{\mathbf{s}_k^2}{\mathbf{s}_k^2 + (n - \mathbf{c}_k)^2}, \tag{1.2}$$

où les vecteurs $\mathbf{c} = \{\mathbf{c}_1, \ldots, \mathbf{c}_K\}$, $\mathbf{a} = \{\mathbf{a}_1, \ldots, \mathbf{a}_K\}$ et $\mathbf{s} = \{\mathbf{s}_1, \ldots, \mathbf{s}_K\}$ représentent respectivement les nombres d'onde, intensités et largeurs des raies lorentziennes. Dans le cas particulier où les raies sont identiques, on a alors, pour tout $k \in \{1, \ldots, K\}$, $\mathbf{s}_k = s$.

En spectroscopie optique, les raies sont toutes de même signe car elles correspondent à un nombre de photons en plus (spectre d'émission dans lequel les raies sont positives) ou en moins (spectre d'absorption dans lequel les raies sont négatives). C'est pourquoi, sauf exception, nous considérons dans cette thèse des spectres dont les raies sont toutes positives. Dès lors, on a $\mathbf{a} \in \mathbb{R}^{+K}$. Par ailleurs, les nombres d'onde des raies ne sont pas obligatoirement situés sur la grille d'échantillonnage. On a donc : $\mathbf{c} \in [1, N]^K$.

Le spectre pur est simulé de la manière suivante. Le nombre de raies K est tout d'abord déterminé en calculant la partie entière d'une variable aléatoire normale centrée sur λN et d'écart-type 1. λ correspond au rapport du nombre de raies sur le nombre total de points, il est traditionnellement compris entre 0 et 0,1 pour un signal impulsionnel et est fixé par l'utilisateur. K raies sont ensuite générées aléatoirement suivant la loi $\mathcal{U}_{[1,N]}$. Les amplitudes de ces raies sont simulées suivant une loi normale à support positif $\mathcal{N}^+(0, r_{\mathbf{x}})$, où $r_{\mathbf{x}}$ correspond à la variance de l'amplitude des raies (fixée par l'utilisateur). Enfin, le spectre pur est calculé comme la somme de K lorentziennes d'amplitudes $\mathbf{a}_k$ et centrées sur $\mathbf{c}_k$. Les largeurs des lorentziennes sont fixées par l'utilisateur.

1.7.2 Modélisation de la ligne de base

Comme on l'a vu dans la section 1.6, la ligne de base varie lentement. C'est la raison pour laquelle elle est, en général, modélisée par un polynôme d'ordre faible [45, 69, 77, 110]. Par conséquent, l'effet de filtrage par la fonction d'appareil peut être négligé puisque la réponse en fréquence de la ligne de base est beaucoup plus concentrée dans les basses fréquences que celle de la fonction d'appareil.

En outre, comme la ligne de base provient de lumière intrinsèque ou extérieure à l'expérience, elle constitue un signal qui vient s'ajouter au spectre.

Dès lors, la ligne de base est modélisée par un signal additif $\mathbf{z} \in \mathbb{R}^N$ correspondant à un polynôme d'ordre 4 ou 5. L'ordre du polynôme est choisi aléatoirement de manière équiprobable, et dont les coefficients sont distribués suivant une gaussienne de moyenne nulle et de variance 1. La ligne de base est enfin normalisée en amplitude (pour n'être ni trop faible, ni trop grande), notamment en ajoutant une constante afin que les échantillons du spectre soient d'amplitude positive.

1.7.3 Modélisation du bruit

Bien que le spectre infrarouge soit beaucoup moins bruité que le spectre Raman, il s'avère nécessaire de prendre en compte un signal modélisant le bruit dans le simulateur. Le bruit a une origine physique puisque la spectroscopie revient à un comptage de photons : on peut donc modéliser le bruit par un processus de Poisson. De plus, la chaîne de mesure introduit une erreur sur le signal, notamment, en spectroscopie Raman, à cause de la barrette CCD qui par ailleurs peut avoir un comportement non-linéaire lorsqu'elle arrive à saturation. À cela vient s'ajouter la nécessité de considérer notre modèle imparfait : le bruit doit également modéliser les incertitudes du modèle.

L'hypothèse classique utilisée en traitement du signal est de considérer le bruit comme étant additif, blanc, i.i.d., gaussien et centré [59]. Ce choix est motivé par trois faits.

Premièrement, la spectroscopie consistant en un comptage de photons, un modèle poissonnien serait plus adapté. Cependant, lorsque le nombre de photons est relativement important, le modèle gaussien

est une bonne approximation. En effet, l'hypothèse gaussienne est justifiée par référence au théorème de la limite centrale qui dit, sous des conditions assez larges, que si le bruit résulte d'un grand nombre d'effets élémentaires cumulés, « aléatoires » et indépendants, la distribution gaussienne est une bonne approximation de sa véritable distribution de fréquence.

Deuxièmement, la modélisation gaussienne du bruit ne vient pas d'une hypothèse sur le caractère aléatoire du bruit mais résulte d'un mode de représentation d'une information a priori incomplète propre à l'inférence bayésienne [59]. Cette modélisation n'est sûrement pas le modèle le plus proche de la réalité, et nous ne prétendons pas qu'elle l'est. En fait, nous supposons plusieurs choses sur le bruit :

- il peut prendre toute valeur réelle ;
- il est de moyenne nulle ;
- les grandes valeurs du bruit sont moins probables que les petites.

Autrement dit, cela revient à considérer que le bruit est de moyenne nulle et d'écart-type fini. Par contre, nous n'avons aucune idée de l'existence des cumulants d'ordre supérieur. Dans ces conditions, le choix le moins compromettant est celui d'une distribution gaussienne.

Troisièmement, comme la distribution du bruit permet d'obtenir la fonction de vraisemblance du modèle, l'hypothèse d'une distribution gaussienne permet de simplifier les calculs.

Pour toutes ces raisons, le bruit $\mathbf{b} \in \mathbb{R}^N$ est supposé additif, blanc, i.i.d. et gaussien de moyenne nulle et de variance $r_{\mathbf{b}}\mathbf{I}$:

$$\mathbf{b} \sim \mathcal{N}(\mathbf{0}, r_{\mathbf{b}}\mathbf{I}).$$

Néanmoins, une amélioration du modèle serait de prendre en compte le fait que la variance du bruit est fonction de l'amplitude du signal spectroscopique : c'est d'ailleurs une bonne approximation d'un bruit poissonnien [101].

Chapitre 2

Estimation de la ligne de base

2.1 Introduction

Ce chapitre a pour objet l'estimation de la ligne de base (encore appelé fond ou arrière-plan) du signal spectroscopique afin de pouvoir la supprimer et faciliter ainsi l'analyse du spectre de raies. La figure 2.1 représente le spectre Raman d'un échantillon de gibbsite $Al(OH)_3$ et une estimation possible de la ligne de base par la méthode proposée dans ce chapitre[1]. Il est alors possible, en soustrayant cette

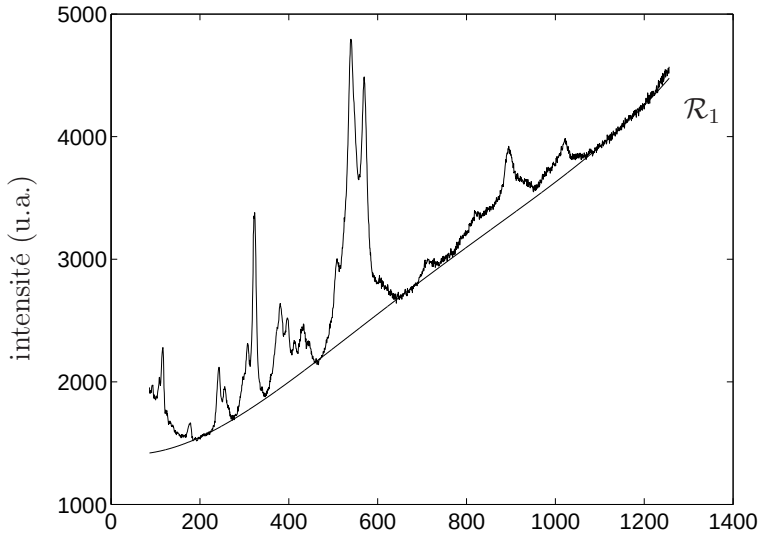


FIG. 2.1 – Spectre Raman de gibbsite et une estimation de la ligne de base.

estimation du signal spectroscopique, de déterminer les paramètres des raies de manière plus précise (c'est l'objet des chapitres 3 et 4).

Ce travail a fait l'objet de publications en congrès [74, 79] et en revue [77]. Le programme Matlab est librement téléchargeable à l'adresse `www.iris.cran.uhp-nancy.fr/francais/idsys/Personnes/Perso_Mazet/backest-fr.htm`.

Une première hypothèse de travail est de considérer que le spectre mesuré est la somme de la ligne de base $\mathbf{z}$ et du signal utile contenant les raies $\mathbf{w}$. Une deuxième hypothèse est la variation lente de la ligne de base par rapport au reste du signal (raies et bruit).

L'idée développée dans ce chapitre est d'estimer la ligne de base par un polynôme d'ordre assez

[1] Plus précisément, il s'agit de la minimisation d'une parabole tronquée asymétrique dont le seuil est fixé à 20 u.a. et l'ordre du polynôme à 4 (« u.a. » signifie « unité arbitraire »).

faible (relativement au nombre de données N), dont les coefficients sont estimés en minimisant une fonction-coût φ adaptée au problème.

Ce chapitre est organisé comme suit.

Tout d'abord, une synthèse bibliographique des méthodes d'estimation de la ligne de base est proposée en section 2.2.

Ensuite, dans la section 2.3, nous proposons une modélisation du problème puis nous présentons quatre fonctions-coût qui empêcheront les raies d'être trop influentes sur l'estimation. Ainsi, la fonction de Huber et la parabole tronquée sont considérées, puis étendues au cas asymétrique afin d'améliorer l'estimation dans le cas de spectres où les raies sont de même signe. Ces fonctions n'étant pas quadratiques (voire pas convexes), la minimisation du critère n'est pas directe. C'est pourquoi nous utilisons une méthode de minimisation semi-quadratique [42, 60]. L'approche proposée est ensuite comparée à d'autres méthodes de minimisation de fonctions-coût asymétriques puis aux moindres carrés tamisés (*least trimmed squares*) [100] et enfin à la méthode proposée par Lieber *et al.* [69] pour laquelle une interprétation mathématique est donnée.

Dans la section 2.4, nous appliquons l'approche proposée à des spectres simulés afin d'illustrer les performances de la méthode. Nous discutons également les avantages et inconvénients respectifs des différentes formes des fonctions-coût. L'influence des paramètres du signal (nombre de points N et taux de contamination Q) sont étudiés et quelques indications sont données pour le choix des paramètres de la méthode (le seuil s et l'ordre du polynôme p).

L'algorithme est enfin appliqué sur des spectres infrarouges et Raman expérimentaux dans la section 2.5.

2.2 Méthodes de correction de la ligne de base

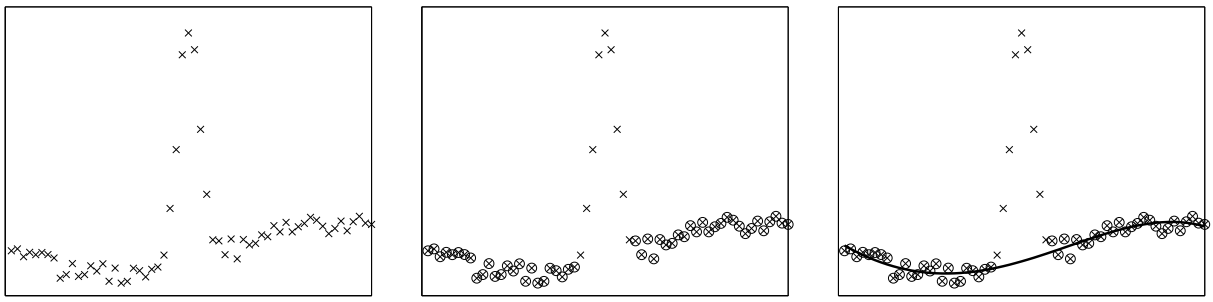
Il n'est pas toujours possible d'éliminer complètement la ligne de base d'un spectre de manière expérimentale. Bien que plusieurs méthodes aient été développées, il faut parfois faire appel à des méthodes numériques pour la supprimer complètement. La plupart de ces méthodes font l'hypothèse que le spectre observé est la somme de la ligne de base et du signal utile contenant les raies du spectre. La correction consiste donc à estimer la ligne de base puis à la soustraire au spectre mesuré. Cette section présente un rapide état de l'art des techniques numériques actuelles.

La méthode la plus simple est l'application de la dérivée première ou seconde qui corrige la composante continue ou linéaire de la ligne de base [4]. Cependant, la plupart des spectres (et en particulier les spectres infrarouge ou Raman) n'ont pas une ligne de base linéaire et, de plus, peuvent être trop bruités pour permettre l'utilisation de cette méthode.

La plupart des logiciels commercialisés (tels qu'Origin[2] ou PeakFit[3]), ainsi que certaines méthodes proposées dans la littérature [45, 110], estime la ligne de base par le polynôme qui minimise le critère des moindres-carrés sur un sous-ensemble de points du spectre défini par l'utilisateur, et considéré comme appartenant à la ligne de base.

Par exemple, la figure 2.2 représente un cas d'école où l'on considère une partie d'un spectre simulé (2.2(a)), la sélection des points que l'utilisateur considère appartenir à la ligne de base seulement et pas aux raies (2.2(b)) et enfin l'estimation de la ligne de base par la méthode des moindres carrés (2.2(c)). Ainsi, si les points sont correctement sélectionnés, cette méthode donne des résultats très satisfaisants. Cela peut être interprété par la possibilité du modèle polynomial de modéliser correctement la plupart des lignes de base [45, 69, 110]. Mais la sélection de ce sous-ensemble de points n'est pas toujours évidente et est une étape lourde et longue (en effet, cette opération ne peut s'effectuer que sur un seul spectre à la fois, ce qui s'avère fastidieux lorsque l'on travaille avec une séquence de plusieurs spectres).

[2] OriginLab Corporation, www.originlab.com.
[3] Systat Software Inc., www.systat.com.



(a) Spectre simulé. (b) Sélection des points. (c) Estimation de la ligne de base.

FIG. 2.2 – Exemple d'estimation de la ligne de base par la méthode des moindres carrés sur un sous-ensemble de points.

C'est pourquoi il est nécessaire de développer des méthodes permettant une sélection automatique des points ou qui soient insensibles aux raies.

Ainsi, l'algorithme itératif proposé par Lieber *et al.* [69] calcule l'estimation des moindres-carrés sur un signal qui est modifié à chaque itération jusqu'à devenir une estimation de la ligne de base. Cet algorithme est détaillé dans la section 2.3.4.3 où nous montrons qu'il est un cas particulier de notre approche.

Eilers [33] utilise un algorithme de moindres-carrés asymétriques (ALS, *asymmetric least squares*) en minimisant une fonction non symétrique φ associée à un paramètre de régularisation :

$$\mathcal{J}(\mathbf{t}) = \sum_{n=1}^{N} \varphi\left(\mathbf{y}_n - (\mathbf{N}\mathbf{t})_n\right) + \lambda \sum_{n=1}^{N} \left(\Delta^2 (\mathbf{N}\mathbf{t})_n\right)^2 .$$

Nous proposons dans la section 2.3.4.1 un lien entre cette méthode et notre approche.

Enfin, nous avons proposé récemment d'utiliser une méthode de moindres-carrés tamisés [79]. C'est une méthode de régression robuste pour calculer l'estimation des moindres-carrés sur un sous-ensemble de points sélectionnés automatiquement. Elle a été appliquée à l'origine dans le domaine de l'estimation robuste aux valeurs aberrantes par Rousseeuw *et al.* [98, 100]. Dans la section 2.3.4.2, nous montrons que cette méthode est équivalente à la méthode proposée.

La transformée en ondelettes est devenue un outil très utilisé pour la correction de la ligne de base [9, 28, 70, 108]. La méthode consiste dans un premier temps à appliquer une transformée en ondelettes (traditionnellement, ce sont les ondelettes Daubechies ou Symlet qui sont utilisées) au spectre, à partir de laquelle les coefficients de l'ondelette sont calculés ; puis, dans un deuxième temps, à séparer la ligne de base supposée être dans les basses fréquences (coefficients d'approximation), des raies et du bruit supposés être dans les hautes fréquences (coefficients de détail). Une telle approche suppose donc implicitement que la ligne de base est complètement séparée du reste du spectre dans le domaine de la transformée. De surcroît, il n'est pas souvent évident de déterminer automatiquement à quel niveau de décomposition les coefficients d'ondelette correspondent à la ligne de base.

Les techniques MSC (*multiplicative scatter correction*) [52, 112] sont utilisées sur des séquences de spectres. Ces séquences correspondent à un ensemble de spectres du même échantillon enregistrés en faisant varier un paramètre (la température, la pression, la quantité d'eau, le pH, etc.). L'hypothèse de travail est que, pour un nombre d'onde donné, l'évolution de la ligne de base selon la séquence est linéaire. La méthode MSC permet d'estimer les coefficients de cette évolution linéaire par des méthodes de minimisation de moindres-carrés.

Toujours dans le cas de séquences de spectres, les méthodes DOSC (*direct orthogonal signal correction*) [72, 102, 111, 112] utilisent deux matrices : la matrice de la séquence de spectre et la matrice de

concentration qui contient les valeurs du paramètre variant. Ici, on considère que la ligne de base est la même pour tous les spectres. De fait, les méthodes DOSC cherchent à supprimer, à partir de cette séquence, les variations qui sont autant que possible orthogonales à la matrice de concentration. Aussi, la matrice de la séquence des spectres est projetée sur la matrice de concentration, la décomposant ainsi en deux parties orthogonales, l'une dans l'espace de concentration (les spectres estimés), l'autre dans l'espace orthogonal (la ligne de base estimée).

Cependant, ces deux méthodes de pré-traitement nécessitent d'avoir enregistré une séquence de plusieurs spectres et, dans le cas des méthodes DOSC, d'en connaître la matrice de concentration. Or, nous souhaitons développer ici une méthode d'estimation de la ligne de base d'un seul spectre.

Enfin, certains auteurs ont posé le problème de l'estimation de la ligne de base dans un cadre bayésien [35, 47, 48]. La ligne de base y est modélisée par des splines cubiques dont la position des nœuds et les coefficients des splines sont à estimer. Pour cela, un a priori de douceur est posé sur la ligne de base. Or, la loi a posteriori résultante est difficilement optimisable, c'est pourquoi les méthodes MCMC sont utilisées pour obtenir une estimation, mais elles sont alors la cause d'un coût de calcul important. Notons que dans [47, 48], l'estimation de la ligne de base est couplée à la déconvolution des raies.

En modélisant le problème de manière différente, nous évitons d'avoir à utiliser les méthodes MCMC, ce qui permet d'obtenir un algorithme plus rapide.

2.3 Méthode proposée

2.3.1 Modélisation du problème

Comme nous l'avons vu dans le chapitre 1, le signal spectroscopique $\mathbf{y} \in \mathbb{R}^N$ peut être modélisé comme la somme d'une ligne de base $\mathbf{z} \in \mathbb{R}^N$ et du spectre bruité $\mathbf{w} \in \mathbb{R}^N$:

$$\mathbf{y} = \mathbf{z} + \mathbf{w}. \tag{2.1}$$

Le signal $\mathbf{z}$ est modélisé par un polynôme d'ordre γ, puisqu'une fonction polynômiale permet de modéliser la plupart des lignes de base [45, 69, 110]. Ainsi, il s'écrit sous la forme $\mathbf{z} = \mathbf{N}\mathbf{t}$ où

$$\mathbf{N} = \begin{pmatrix} \mathbf{n}_1^0 & \cdots & \mathbf{n}_1^\gamma \\ \vdots & & \vdots \\ \mathbf{n}_N^0 & \cdots & \mathbf{n}_N^\gamma \end{pmatrix} \quad \text{et} \quad \mathbf{t} = \begin{pmatrix} \mathbf{t}_0 \\ \vdots \\ \mathbf{t}_\gamma \end{pmatrix}$$

représentent respectivement la matrice de Vandermonde des nombres d'onde et les coefficients du polynôme.

Dans cette section, le signal $\mathbf{w}$ est considéré comme un bruit, puisque pour l'instant il ne contient pas l'information recherchée ; son traitement sera considéré dans les chapitres 3 et 4.

Aucun modèle particulier n'est nécessaire sur ce signal ; cependant nous supposerons qu'il est constitué de raies de positions, d'amplitudes et de formes différentes, plus d'un signal gaussien centré de variance $r_{\mathbf{b}}$ de moyenne nulle représentant les erreurs de mesure et les incertitudes du modèle. Comme le signal est constitué de raies, il peut être considéré comme parcimonieux, c'est-à-dire qu'il possède un grand nombre de valeurs proches de zéro, et peu de valeurs de grande amplitude. Par la suite, nous considérerons les cas où les raies peuvent être de signes différents, puis le cas où elles sont toutes positives.

Le modèle ainsi obtenu va nous permettre de proposer les fonctions-coût appropriées.

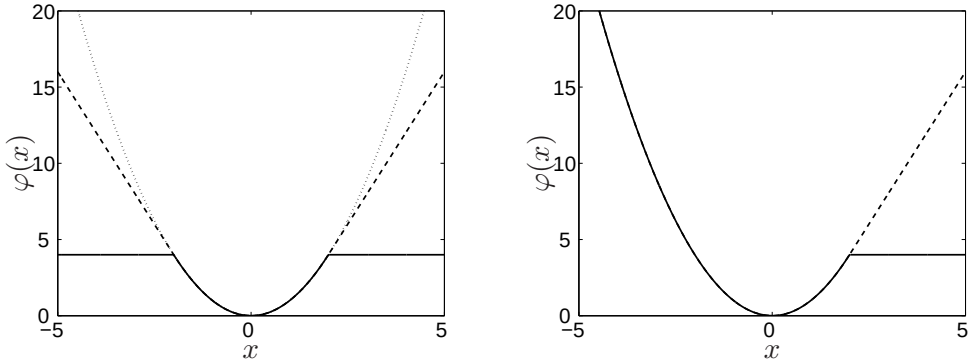
2.3.2 Détermination des fonctions-coût

La méthode proposée dans ce chapitre consiste à trouver les coefficients polynômiaux $\mathbf{t}$ qui minimisent un critère de la forme :

$$\mathcal{J}(\mathbf{t}) = \sum_{n=1}^{N} \varphi\big(\mathbf{y}_n - (\mathbf{N}\mathbf{t})_n\big), \tag{2.2}$$

où φ est l'une des fonctions-coût présentées ci-dessous.

Tout d'abord, considérons l'estimation de la ligne de base par l'approche classique des moindres carrés. Cette approche consiste à trouver les coefficients $\mathbf{t}$ qui minimisent l'erreur des moindres carrés entre la ligne de base estimée et le signal, c'est-à-dire que la fonction-coût est la fonction parabole : $\varphi(x) = x^2$ (figure 2.3(a)).



(a) Fonctions-coût symétriques. (b) Fonctions-coût asymétriques.

FIG. 2.3 – Fonctions-coût proposées pour l'estimation de la ligne de base avec $s = 2$ ($\cdots$: parabole, - - : fonction de Huber, — : parabole tronquée).

Dans le cas du problème considéré, cette approche donne des résultats inexploitables puisqu'une parabole donne un coût quadratique à chaque valeur $\mathbf{y}_n - (\mathbf{N}\widehat{\mathbf{t}})_n$. Dès lors, les raies, caractérisées par une valeur $\mathbf{y}_n - (\mathbf{N}\mathbf{t})_n$ très grande, sont affectés d'un coût très important, déformant alors le polynôme vers ces grandes valeurs pour tenter de minimiser le critère (2.2). La figure 2.4 considère le spectre de la gibbsite (déjà représenté figure 2.1) pour lequel la ligne de base est estimée grâce à une fonction-coût quadratique[4] : cet exemple met en évidence le fait que l'estimation tend vers les grandes valeurs pour réduire le coût du critère $\mathcal{J}$, impliquant alors une mauvaise estimation de la ligne de base.

Une autre manière de voir la limitation de l'approche des moindres carrés est de considérer l'approche probabiliste. En effet, utiliser une fonction-coût quadratique est équivalent à supposer les échantillons du signal $\mathbf{w}$ i.i.d. et distribués suivant une gaussienne de moyenne nulle. Cela n'est évidemment pas le cas ici puisque $\mathbf{w}$ est la somme de raies et d'un bruit gaussien : la distribution des éléments de $\mathbf{w}$ peut donc être vue comme la convolution entre une loi normale et une loi Bernoulli-gaussienne (le chapitre 3 développe plus en détail cette modélisation).

Pour surmonter ce problème, nous proposons d'utiliser des fonctions-coût dont le coût est moindre pour les grandes valeurs. Ces fonctions doivent être quadratiques au voisinage de zéro, c'est-à-dire lorsque la ligne de base et le signal sont proches (cela satisfait l'interprétation bayésienne d'un bruit gaussien centré sur zéro), mais doivent croître plus lentement qu'une parabole à partir d'un seuil s afin que les raies soient moins influentes sur l'estimation de la ligne de base. Le choix de ce seuil est discuté section 2.4.4.

Dans [77], les deux fonctions-coût suivantes, initialement proposées dans le contexte de l'estimation

[4] l'ordre du polynôme est égal à 4.

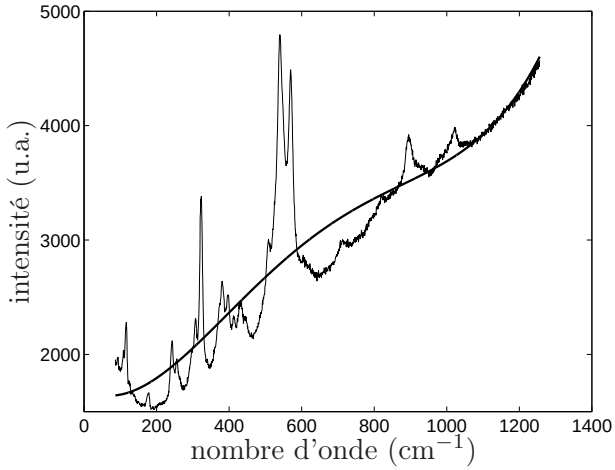


FIG. 2.4 – Estimation des moindres carrés de la ligne de base d'un spectre Raman de gibbsite.

robuste aux valeurs aberrantes [57, 99], sont considérées (cf. figure 2.3(a)) :

– la fonction de Huber :

$$\forall x \in \mathbb{R}, \quad \varphi_{\mathrm{H}}(x) = \begin{cases} x^2 & \text{si } |x| < s, \\ 2s|x| - s^2 & \text{sinon}\,; \end{cases} \tag{2.3}$$

– la parabole tronquée :

$$\forall x \in \mathbb{R}, \quad \varphi_{\mathrm{PT}}(x) = \begin{cases} x^2 & \text{si } |x| < s, \\ s^2 & \text{sinon.} \end{cases} \tag{2.4}$$

Alors que la fonction-coût des moindres carrés est quadratique partout, la fonction de Huber et la parabole tronquée sont respectivement linéaires et constantes au delà de s. De plus, dans le cas de la parabole tronquée, les points dont la distance à l'estimation est supérieure à s ont tous le même coût. En d'autres termes, cela signifie qu'une grande raie affecte l'estimation de la même façon qu'une petite.

Les fonctions-coût précédentes sont symétriques, ce qui signifie qu'elles diminuent le coût des grandes valeurs positives, mais également des grandes valeurs négatives, ce qui est approprié aux spectres où les raies peuvent être positives ou négatives.

Dans le cas particulier de la spectroscopie optique où les raies sont toutes positives, les valeurs négatives n'ont pas à avoir un coût faible. Pour cela, nous proposons d'utiliser la variante asymétrique des fonctions-coût précédentes (cf. figure 2.3(b)) :

– la fonction de Huber asymétrique :

$$\forall x \in \mathbb{R}, \quad \varphi_{\mathrm{HA}}(x) = \begin{cases} x^2 & \text{si } x < s, \\ 2sx - s^2 & \text{sinon}\,; \end{cases} \tag{2.5}$$

– la parabole tronquée asymétrique :

$$\forall x \in \mathbb{R}, \quad \varphi_{\mathrm{PTA}}(x) = \begin{cases} x^2 & \text{si } x < s, \\ s^2 & \text{sinon.} \end{cases} \tag{2.6}$$

Ces fonctions-coût donnent toujours un coût faible aux raies positives, mais sont quadratiques dans toute la partie négative, prenant en compte le fait que, dans la partie négative, le signal $\mathbf{w}$ n'est constitué que de bruit gaussien.

Bien entendu, ces deux fonctions-coût ne sont pas les seules qui peuvent être utilisées pour l'estimation de la ligne de base. Par exemple, dans [74, 79], nous avons également proposé les fonctions-coût

hyperboliques ou de Cauchy. Cependant, il n'est pas possible d'utiliser la variante asymétrique de la fonction de Cauchy avec la méthode proposée dans ce chapitre car cette fonction-coût ne vérifie pas les conditions nécessaires pour utiliser les algorithmes de minimisation semi-quadratique (voir section suivante).

2.3.3 Minimisation semi-quadratique

Dans le cas d'une fonction-coût quadratique, le critère (2.2) correspond au critère des moindres carrés dont la minimisation est directe et donne l'expression explicite suivante :

$$\widehat{\mathbf{t}} = \left(\mathbf{N}^T\mathbf{N}\right)^{-1}\mathbf{N}^T\mathbf{y}.$$

En revanche, la minimisation de fonctions-coût non-quadratiques n'est pas directe. Aussi, pour minimiser le critère (2.2), nous proposons d'utiliser un algorithme de minimisation semi-quadratique, qui consiste en une méthode déterministe et itérative simplifiant l'optimisation d'un critère non-quadratique [19, 42, 60] ; c'est une méthode très simple à mettre en œuvre.

Notons que la méthode des moindres carrés repondérés [114] est un des premiers exemple d'approche semi-quadratique développé dans le cas unidimensionnel pour la déconvolution LP. Une comparaison entre cette méthode et la minimisation quadratique peut être trouvée dans [60].

2.3.3.1 Principe

L'idée de la minimisation semi-quadratique est d'introduire une variable auxiliaire $\mathbf{d} = \{\mathbf{d}_1, \cdots, \mathbf{d}_N\}^T$ de façon à obtenir un critère augmenté $\mathcal{K}$ admettant le même minimum que $\mathcal{J}$, soit, pour tout $\mathbf{t}$,

$$\min_{\mathbf{d}} \mathcal{K}(\mathbf{t}, \mathbf{d}) = \mathcal{J}(\mathbf{t}). \tag{2.7}$$

Pour des raisons que nous expliquerons plus loin, nous choisissons d'utiliser la forme de Geman et Yang pour laquelle la variable auxiliaire est additive [42] :

$$\mathcal{K}(\mathbf{t}, \mathbf{d}) = \sum_{n=1}^{N} \frac{1}{2}\left((\mathbf{y}_n - (\mathbf{N}\mathbf{t})_n - \mathbf{d}_n)^2 + \psi(\mathbf{d}_n)\right).$$

Notons que le nouveau critère $\mathcal{K}$ est quadratique en $\mathbf{t}$ et convexe en $\mathbf{d}$, justifiant le nom de « critère semi-quadratique ». La fonction ψ est choisie de manière à vérifier l'équation (2.7). Pour cela, elle doit satisfaire certaines conditions détaillées dans [42, section III.A].

On appelle *paire de Legendre* le couple de fonction (f, g) telles que f et g soient convexes et vérifient :

$$f(u) = \sup_v(uv - g(v)) \qquad \text{et} \qquad g(v) = \sup_u(uv - f(u)).$$

Or, d'après [42, théorème 2], si

$$f(u) = u^2/2 - \varphi(u) \qquad \text{et} \qquad g(v) = v^2/2 + \psi(v)$$

est une paire de Legendre, alors :

$$\mathcal{J}(\mathbf{t}) = \inf_{\mathbf{d}} \mathcal{K}(\mathbf{t}, \mathbf{d}).$$

On obtient par ailleurs :

$$\varphi(u) = \sup_v\left(\frac{(u-v)^2}{2} + \psi(v)\right) \qquad \text{et} \qquad \psi(v) = \sup_u\left(-\frac{(u-v)^2}{2} + \varphi(u)\right).$$

En fait, nous optons pour le critère de Geman et Yang dans lequel la fonction-coût φ est altérée par un changement d'échelle (voir [60, section III.B]) : on considère dorénavant la fonction $\varphi_\alpha = \alpha\varphi$. Ainsi, en supposant que φ satisfasse la condition suivante :

$$\exists\, \alpha_{\max} \,/\, \forall \alpha \in [0;\alpha_{\max}[, \qquad f_\alpha(u) = u^2/2 - \alpha\varphi(u) \text{ est strictement convexe,} \tag{2.8}$$

le critère augmenté $\mathcal{K}$ prend alors la forme suivante :

$$\mathcal{K}(\mathbf{t},\mathbf{d}) = \frac{1}{\alpha}\sum_{n=1}^{N}\frac{1}{2}\left((\mathbf{y}_n - (\mathbf{N}\mathbf{t})_n - \mathbf{d}_n)^2 + \psi_\alpha(\mathbf{d}_n)\right),$$

où

$$\psi_\alpha(v) = \sup_u\left(-\frac{(u-v)^2}{2} + \alpha\varphi(u)\right).$$

La condition (2.8) est vérifiée si $f_\alpha(x)$ est strictement convexe, ce qui est équivalent à $f''_\alpha(x) \geq 0$, soit :

$$1 - \alpha\varphi''(x) \geq 0 \qquad \Leftrightarrow \qquad \alpha \leq 1/\varphi''(x).$$

Comme :

$$\varphi''_{\mathrm{H}}(x) = \begin{cases} 2 & \text{si } |x| < s, \\ 0 & \text{sinon,} \end{cases} \qquad \varphi''_{\mathrm{PT}}(x) = \begin{cases} 2 & \text{si } |x| < s, \\ 0 & \text{sinon,} \end{cases}$$

$$\varphi''_{\mathrm{HA}}(x) = \begin{cases} 2 & \text{si } x < s, \\ 0 & \text{sinon,} \end{cases} \qquad \varphi''_{\mathrm{PTA}}(x) = \begin{cases} 2 & \text{si } x < s, \\ 0 & \text{sinon,} \end{cases}$$

on en déduit que les fonctions-coût proposées vérifient la condition (2.8) avec $\alpha_{\max} = 1/2$.

2.3.3.2 Algorithme d'optimisation

Dans [18, 19, 60], l'algorithme déterministe LEGEND est appliqué pour la minimisation de $\mathcal{K}$ dans le cadre de l'estimation de contours en traitement d'image. Notons que dans le cas du critère de Geman et Reynolds [41], l'algorithme ARTUR est utilisé pour assurer l'optimisation du critère [18, 19, 60]. Or, φ ne vérifie pas les conditions permettant d'utiliser ARTUR (en particulier, la symétrie de la fonction-coût n'est forcément pas vérifiée pour les fonctions asymétriques) ; c'est pourquoi nous utilisons la forme de Geman et Yang et l'algorithme LEGEND.

L'algorithme LEGEND est un algorithme de descente par bloc (ou de relaxation) qui estime alternativement $\mathbf{t}$ et $\mathbf{d}$ en utilisant la convexité de $\mathcal{K}$ en $\mathbf{t}$ quand $\mathbf{d}$ est fixé et vice-versa :

- Lorsque $\mathbf{d}$ est fixé, $\mathcal{K}$ est quadratique et sa minimisation donne une expression explicite de $\widehat{\mathbf{t}}$:

$$\widehat{\mathbf{t}} = \left(\mathbf{N}^T\mathbf{N}\right)^{-1}\mathbf{N}^T(\mathbf{y}+\mathbf{d}), \tag{2.9}$$

  lorsque la matrice $\mathbf{N}^T\mathbf{N}$ est inversible. Cette expression est intéressante pour deux raisons. Tout d'abord, elle peut s'interpréter comme étant l'estimateur des moindres carrés sur le signal $\mathbf{y}+\mathbf{d}$. Ensuite, d'un point de vue calculatoire, la matrice $\left(\mathbf{N}^T\mathbf{N}\right)^{-1}$ peut être calculée une fois pour toutes au début de l'algorithme.
- lorsque $\mathbf{t}$ est fixé, le minimum de $\mathcal{K}$ est atteint pour [18, 19, 60] :

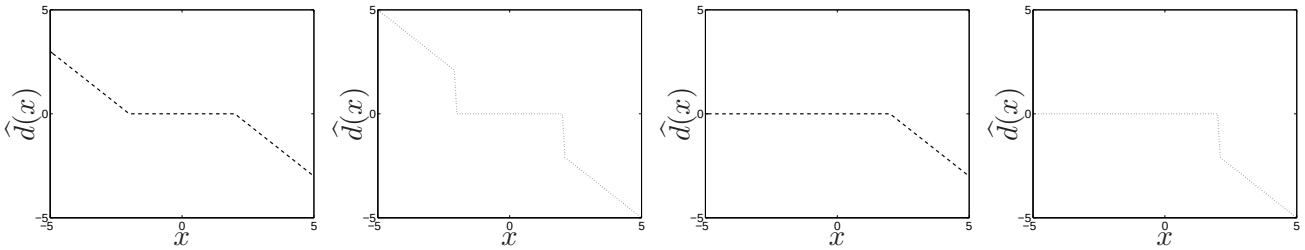
$$\forall n \in \{1,\ldots,N\}, \quad \widehat{\mathbf{d}}_n = -\varepsilon_n + \alpha\varphi'(\varepsilon_n) \tag{2.10}$$

  où $\varepsilon_n = \mathbf{y}_n - (\mathbf{N}\mathbf{t})_n$. Ainsi, on a pour chacune des fonction-coût dans le cas où $\alpha = \alpha_{\max} = 1/2$:

$$\widehat{d}_{\mathrm{H}}(x) = \begin{cases} -x - s & \text{si } x \leq -s, \\ 0 & \text{si } |x| < s, \\ -x + s & \text{si } x \geq s, \end{cases} \qquad \widehat{d}_{\mathrm{PT}}(x) = \begin{cases} 0 & \text{si } |x| < s, \\ -x & \text{sinon,} \end{cases}$$

$$\widehat{d}_{\mathrm{HA}}(x) = \begin{cases} 0 & \text{si } x < s, \\ -x + s & \text{sinon,} \end{cases} \qquad \widehat{d}_{\mathrm{PTA}}(x) = \begin{cases} 0 & \text{si } x < s, \\ -x & \text{sinon.} \end{cases}$$

La figure 2.5 représente les quatre fonctions $\widehat{d}(x)$. Il est à noter que ces fonctions effectuent un seuillage sur $\varepsilon_n = \mathbf{y}_n - (\mathbf{Nt})_n$ pour les fonctions-coût symétriques et sur $|\varepsilon|$ pour les fonctions-coût asymétriques. Ce seuillage est « dur » dans le cas de la parabole tronquée (les coefficients inférieur au seuil sont mis à zéro) et « doux » dans le cas de la fonction de Huber (identique au seuillage dur, mais sans discontinuité dans la courbe).



(a) Fonction de Huber symétrique. (b) Parabole tronquée symétrique. (c) Fonction de Huber asymétrique. (d) Parabole tronquée asymétrique.

FIG. 2.5 – Fonctions $\widehat{d}(x)$ ($s = 2$).

En résumé, l'algorithme LEGEND consiste à :

1. initialiser $\widehat{\mathbf{t}}^0$ (ici, on choisit l'estimation des moindres carrés) ;
2. répéter jusqu'à convergence :
   (a) $\widehat{\mathbf{d}}^i = \arg\min_{\mathbf{d}} \mathcal{K}\left(\widehat{\mathbf{t}}^{i-1}, \mathbf{d}\right)$ (équation (2.10)),
   (b) $\widehat{\mathbf{t}}^i = \arg\min_{\mathbf{t}} \mathcal{K}\left(\mathbf{t}, \widehat{\mathbf{d}}^i\right)$ (équation (2.9)).

Dans [18], Charbonnier préconise de ne pas effectuer le test d'arrêt sur l'évolution du critère $\mathcal{K}$, car le calcul peut être très coûteux. Nous suivons donc sa démarche et utilisons comme critère d'arrêt l'évolution relative entre deux estimées successives :

$$\frac{\left|\left|\widehat{\mathbf{z}}^{(i)} - \widehat{\mathbf{z}}^{(i-1)}\right|\right|^2}{\left|\left|\widehat{\mathbf{z}}^{(i-1)}\right|\right|^2} < \epsilon,$$

où ϵ est une valeur prédéfinie par l'utilisateur. Afin d'accélérer la convergence de l'algorithme, la valeur de α doit être très proche de $\alpha_{\max}$ [60]. Pour cette raison, nous choisissons en pratique $\alpha = 0,99 \times \alpha_{\max}$.

Notons que puisque la fonction de Huber est convexe, la convergence de l'algorithme vers le minimum global est assuré [18, 19, 60]. Au contraire, la parabole tronquée étant non convexe, le critère $\mathcal{K}$ peut avoir un minimum local, même s'il est convexe en $\mathbf{t}$ pour $\mathbf{d}$ fixé et vice-versa : nous ne pouvons donc pas garantir la convergence vers le minimum global dans le cas de la parabole tronquée. Cependant, dans toutes les simulations effectuées, les lignes de base estimées restent très proches des lignes de base simulées ce qui nous laisse penser que le minimum global (ou un minimum local très proche du minimum global) est atteint à chaque fois. Cet comportement a également été observé expérimentalement dans [19].

2.3.4 Comparaison avec d'autres méthodes

2.3.4.1 Méthodes basées sur des fonctions-coût asymétriques

L'idée de minimiser une fonction-coût asymétrique a déjà été proposée par Eilers [33] (dont la méthode est également appliquée au problème de l'estimation de la ligne de base) et Koenker *et al.* [65].

Ces deux articles proposent de minimiser le critère suivant :

$$\mathcal{J}(\mathbf{t}) = \sum_{n=1}^{N} \varphi\left(\mathbf{y}_n - (\mathbf{N}\mathbf{t})_n\right) + \lambda \sum_{n=1}^{N} \xi\left(\Delta^2(\mathbf{N}\mathbf{t})_n\right)$$

où φ correspond à la fonction-coût asymétrique et le second terme (terme de régularisation) contrôle la douceur de la solution à travers la minimisation de la dérivée seconde de l'estimation. Le paramètre positif λ fixe le poids du terme de régularisation : plus λ est grand, plus l'estimation est douce. Dans [33], on a :

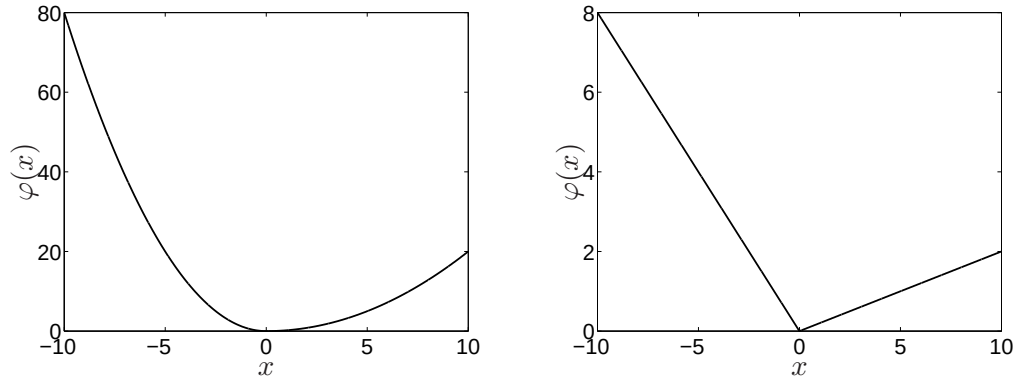
$$\varphi(x) = \begin{cases} \tau x^2 & \text{si } x > 0, \\ (1-\tau)x^2 & \text{sinon} \end{cases} \qquad \text{et} \qquad \xi(x) = x^2$$

avec $\tau \in [0, 1]$. Dans [65],

$$\varphi(x) = x(\tau - \mathbb{1}_{x<0}) \qquad \text{et} \qquad \xi(x) = |x|,$$

où $\tau \in [0, 1]$ et l'estimation est restreinte à la classe des splines cubiques.

La figure 2.6 représente les fonctions-coût utilisées par ces deux auteurs, dans le cas où $\tau = 0, 2$.



(a) Fonction-coût de Eilers. (b) Fonction-coût de Koenker *et al.*

FIG. 2.6 – Fonctions-coût asymétriques proposées par Eilers et Koenker *et al.* ($\tau = 0, 2$).

L'asymétrie de ces fonctions-coût leur permet donc d'être utilisée pour calculer le polynôme estimant la ligne de base. En effet, on retrouve l'idée qui consiste à affecter un coût plus grand aux valeurs négatives qu'aux valeurs positives. De plus, elle sont comparables à la fonction de Huber asymétrique car, comme elles n'ont pas de partie constante, les raies de grande amplitude restent plus influentes que les petites sur l'estimation.

Dans l'approche proposée dans ce chapitre, la douceur de la ligne de base estimée est contrôlée en fixant l'ordre du polynôme à une valeur suffisamment faible. En revanche, Eilers et Koenker *et al.* assurent la douceur de l'estimation grâce au terme de régularisation de la fonction-coût : la douceur est alors contrôlée par le paramètre λ. Ainsi, l'estimation n'est pas restreinte à une fonction polynômiale, ce qui peut s'avérer être intéressant dans le cas de lignes de base très irrégulières. Avec la méthode proposée, il faudrait augmenter l'ordre du polynôme, ce qui peut introduire des problèmes numériques.

2.3.4.2 Méthode des moindres carrés tamisés

Minimiser une parabole tronquée asymétrique est équivalent à l'approche des moindres carrés tamisés (LTS, *least trimmed squares*) proposée par Rousseeuw [98, 100].

En effet, la méthode des moindres carrés tamisés consiste à minimiser le critère des moindres carrés sur un sous-ensemble de données. Dans l'exemple de notre application, cela consiste à sélectionner les points du spectre qui n'appartiennent qu'à la ligne de base, en ignorant les raies. En cela, la méthode des moindres carrés tamisés est équivalente aux méthodes d'estimation de la ligne de base

présentes dans la plupart des logiciels de chimiométrie, à cela près que la sélection du sous-ensemble de points est automatique. Par ailleurs, la méthode des moindres carrés tamisés revient à assigner un coût constant aux raies, puisqu'ils ne sont pas influents dans le calcul de l'estimation. Le nombre de points du sous-ensemble étant fixé au début, l'algorithme cherche alors le sous-ensemble qui minimise l'erreur quadratique moyenne. La convergence est atteinte lorsque le sous-ensemble reste identique d'une itération à l'autre. Pour éviter une recherche exhaustive de ce sous-ensemble, une méthode rapide (FAST-LTS) est proposée dans [100].

Comme l'approche proposée dans ce chapitre donne un coût constant aux raies et un coût quadratique aux autres points, elle définie également deux sous-ensembles de points. En supposant que les points sont affectés de la même manière dans les deux approches, le résultat sera exactement le même : on peut donc dire que ces méthodes sont équivalentes. En fait, la seule différence entre elles vient de la méthode de recherche du sous-ensemble.

Ainsi, il apparaît que la méthode proposée est équivalente à celle utilisée dans la plupart des logiciels, la différence étant que la recherche de ce sous-ensemble est automatique.

Afin de permettre une étude quantitative de ces deux approches, cinquante spectres de 256 points ont été simulés avec un RSB (rapport signal-à-bruit) de 15 dB. Dans le cadre de l'estimation de la ligne de base, le RSB est calculé comme le rapport des énergies de la ligne de base $\mathbf{z}$ sur le spectre bruité $\mathbf{w}$:

$$\mathrm{RSB} = 10 \log \left(\frac{\sum_{n=1}^{N} \mathbf{z}_n^2}{\sum_{n=1}^{N} \mathbf{w}_n^2} \right).$$

Pour toutes les simulations, la ligne de base a été simulée avec un polynôme d'ordre 6. Les temps de calcul moyens ainsi que l'EQM (erreur quadratique moyenne) sont présentés dans le tableau 2.1. L'EQM est calculée grâce à la formule suivante :

$$\mathrm{EQM} = \frac{1}{N} \sum_{n=1}^{N} (\mathbf{z}_n - \widehat{\mathbf{z}}_n)^2 .$$

	temps de calcul moyen	EQM moyenne
méthode proposée	63 ms	0,17
FAST-LTS	2278 ms	0,23

Tab. 2.1 – Comparaison des performances entre la méthode proposée et les moindres carrés tamisés.

Il apparaît alors que les deux méthodes donnent des résultats équivalents (les EQM sont du même ordre de grandeur), mais que notre approche est plus rapide [74].

2.3.4.3 Méthode de Lieber *et al.*

Lieber *et al.* [69] ont proposé un algorithme itératif basé sur la méthode des moindres carrés pour estimer la ligne de base d'un spectre par un polynôme, en éliminant les raies. Leur étude est restreinte aux spectres Raman pour lesquels les raies sont positives.

L'estimation est calculée à partir du signal $\mathbf{y}$, redéfini à chaque itération. Ainsi, à l'estimation i, l'algorithme consiste à :

1. calculer une estimation de la ligne de base ($\mathbf{N}\widehat{\mathbf{t}}$) par le polynôme minimisant le critère des moindres carrés :

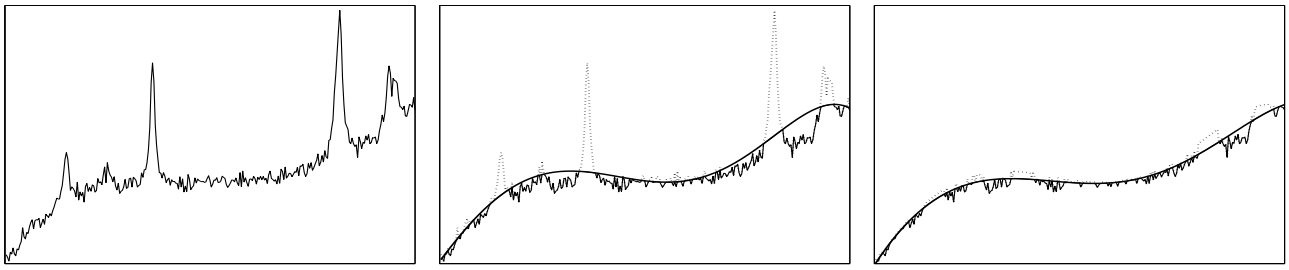
$$\mathcal{J}(\mathbf{t}) = \sum_{n=1}^{N} (\mathbf{y}_n - (\mathbf{N}\mathbf{t})_n)^2 \ ;$$

2. redéfinir $\mathbf{y}$ tel que, pour tout point $n \in \{1, \ldots, N\}$;

$$\begin{cases} \mathbf{y}_n = \mathbf{y}_n & \text{si } \mathbf{y}_n < (\mathbf{N}\widehat{\mathbf{t}})_n, \\ \mathbf{y}_n = (\mathbf{N}\widehat{\mathbf{t}})_n & \text{sinon ;} \end{cases}$$

3. retour en 1.

En d'autres termes, pour chaque point du spectre, $\mathbf{y}$ est fixé égal à l'amplitude de la ligne de base estimée si l'intensité du spectre en ce point est plus grande que l'estimation ; sinon il est inchangé. La figure 2.7 représente le spectre original suivi des deux premières itérations de l'algorithme.



(a) Spectre original. (b) Première estimation de la ligne de base et $\mathbf{y}$ recalculé. (c) Deuxième estimation de la ligne de base et $\mathbf{y}$ recalculé.

FIG. 2.7 – Les deux premières itérations de l'algorithme de Lieber *et al.* sur un spectre simulé.

Cette méthode est équivalente à la minimisation d'une parabole tronquée asymétrique où $s = 0$ car LEGEND estime la ligne de base en minimisant le critère des moindres carrés sur le signal $\mathbf{y} + \mathbf{d}$ (équation (2.9)) :

$$\widehat{\mathbf{t}} = \left(\mathbf{N}^T\mathbf{N}\right)^{-1}\mathbf{N}^T(\mathbf{y} + \mathbf{d})$$

en supposant que la matrice $\mathbf{N}^T\mathbf{N}$ est inversible. À partir de l'équation (2.10), et en considérant un échantillon particulier $n \in \{1, \ldots, N\}$,

$$\begin{aligned} \mathbf{y}_n + \widehat{\mathbf{d}}_n &= \mathbf{y}_n - \varepsilon_n + \alpha\varphi'(\varepsilon_n), \\ &= (\mathbf{N}\widehat{\mathbf{t}})_n + \alpha\varphi'(\mathbf{y}_n - (\mathbf{N}\widehat{\mathbf{t}})_n). \end{aligned} \tag{2.11}$$

Or, en choisissant la valeur limite $\alpha = \alpha_{\max} = 1/2$ et la parabole tronquée asymétrique, nous obtenons :

$$\alpha\varphi'(\mathbf{y}_n - (\mathbf{N}\widehat{\mathbf{t}})_n) = \begin{cases} \mathbf{y}_n - (\mathbf{N}\widehat{\mathbf{t}})_n & \text{si } \mathbf{y}_n - (\mathbf{N}\widehat{\mathbf{t}})_n < s, \text{ c'est-à-dire } \mathbf{y}_n < (\mathbf{N}\widehat{\mathbf{t}})_n + s, \\ 0 & \text{sinon.} \end{cases}$$

Ainsi, l'équation (2.11) devient, pour $s = 0$:

$$\mathbf{y}_n + \widehat{\mathbf{d}}_n = \begin{cases} \mathbf{y}_n & \text{si } \mathbf{y}_n < (\mathbf{N}\widehat{\mathbf{t}}), \\ (\mathbf{N}\widehat{\mathbf{t}})_n & \text{sinon.} \end{cases}$$

En résumé, l'estimation des moindres carrés est calculée sur le signal égal, pour chaque échantillon n, à $\mathbf{y}_n$ si $\mathbf{y}_n$ est plus petit que la ligne de base estimée, sinon à $(\mathbf{N}\widehat{\mathbf{t}})_n$.

Dès lors, la méthode présentée dans [69] est un cas particulier de celle présentée dans ce chapitre où le seuil de la parabole tronquée asymétrique est égal à zéro. Or, une valeur nulle du seuil fait tendre l'estimation vers le bas du spectre (voir section 2.4.4). Pour limiter ce problème, Lieber *et al.* ont proposé de supposer la convergence atteinte lorsqu'un nombre maximal de points modifiés est atteint.

2.4 Influence et choix des paramètres

En pratique, la fonction Matlab `polyfit` a été utilisée pour calculer les coefficients du polynôme. De plus, le vecteur $\mathbf{t}$ a été centré et redimensionné à $[-1;1]$ afin d'éviter les problèmes numériques.

Afin de quantifier le résultat des simulations, nous rappelons l'expression de l'erreur quadratique moyenne (EQM) :

$$\text{EQM} = \frac{1}{N}\sum_{n=1}^{N}(\mathbf{z}_n - \widehat{\mathbf{z}}_n)^2 .$$

2.4.1 Choix de la fonction-coût

Afin de comparer les estimations obtenues grâce aux quatre fonctions-coûts, nous avons procédé à une étude quantitative en créant pour les deux types de spectre (raies de même signe ou non) deux cents simulations. Pour chaque simulation, la ligne de base a été estimée en utilisant les quatre fonctions-coût dont les seuils ont été choisis comme étant ceux donnant la meilleure estimation en terme d'EQM (on parle de *seuil optimal*). Pour chaque fonction-coût, la moyenne des EQM entre les lignes de base réelles et estimées a été calculée et est reportée dans le tableau 2.2.

	raies positives ou négatives	raies positives seulement
fonction de Huber symétrique	$49,1 \times 10^{-5}$	$241,8 \times 10^{-5}$
parabole tronquée symétrique	$2,7 \times 10^{-5}$	$21,4 \times 10^{-5}$
fonction de Huber asymétrique	$1129,1 \times 10^{-5}$	$14,8 \times 10^{-5}$
parabole tronquée asymétrique	$6387,1 \times 10^{-5}$	$2,0 \times 10^{-5}$

TAB. 2.2 – Comparaison de l'EQM entre les lignes de base réelles et estimées pour les quatre fonctions-coût.

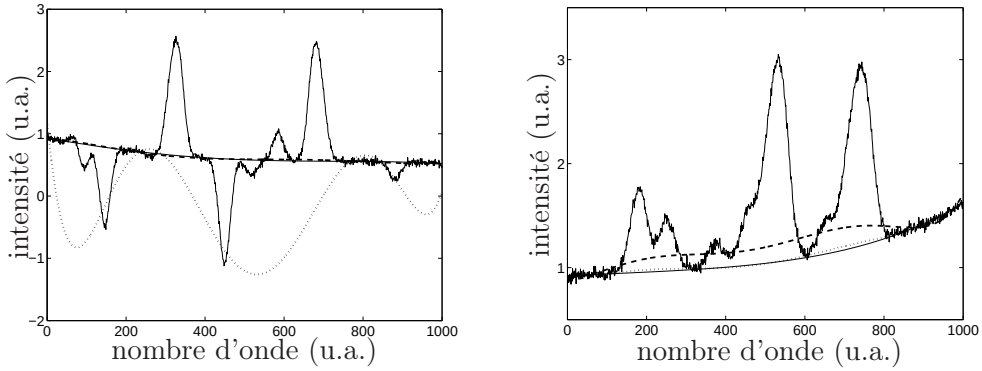
Nous avons vu dans la section 2.3.2 que les fonctions-coût asymétriques ont été introduites dans le cas où les raies du spectre sont de même signe, alors que les fonctions-coût symétriques sont plus adaptées aux spectres possédant des raies positives et négatives. Les statistiques obtenues confirment cette remarque. De plus, lorsque la symétrie de la fonction-coût est correctement choisie, la parabole tronquée donne une meilleure estimation que la fonction de Huber. Cela s'explique par le fait que, dans le cas de la parabole tronquée, toutes les raies dont l'amplitude est supérieure au seuil s ont un coût constant. Ainsi, quelle que soit leur amplitude, les raies n'influent pas du tout sur l'estimation. En revanche, la fonction de Huber ne possède pas de partie constante et, par conséquent, plus les raies sont grandes, plus elles perturbent l'estimation.

Afin d'illustrer l'influence de la symétrie, nous avons comparé les estimations obtenues avec une parabole tronquée symétrique et sa variante asymétrique dans le cas de spectres contenant des raies positives et négatives (figure 2.8(a)) ou seulement des raies positives (figure 2.8(b)). Pour chaque fonction-coût, le seuil correspond au seuil optimal. Dans le cas d'un spectre contenant des raies positives et négatives, la meilleure estimation est donnée par la forme symétrique de la fonction-coût, alors que pour un spectre contenant des raies de même signes, c'est la forme asymétrique qui donne le meilleur résultat.

Les figures 2.9(a) et 2.9(b) présentent les deux types de spectres pour lesquels les lignes de base sont estimées grâce aux deux fonctions-coût étudiées (la symétrie étant correctement choisie et le seuil correspondant au seuil optimal). Ces deux simulations illustrent le fait que la parabole tronquée donne la meilleure estimation pourvu que la symétrie de la fonction-coût soit correctement choisie.

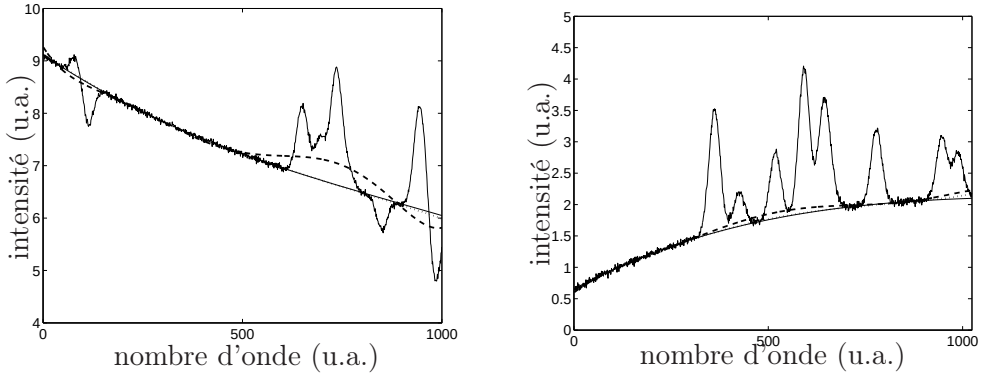
Les seuils optimaux des quatre fonctions-coût sont les suivants :

– $s = 0,02$ pour la fonction de Huber symétrique ;



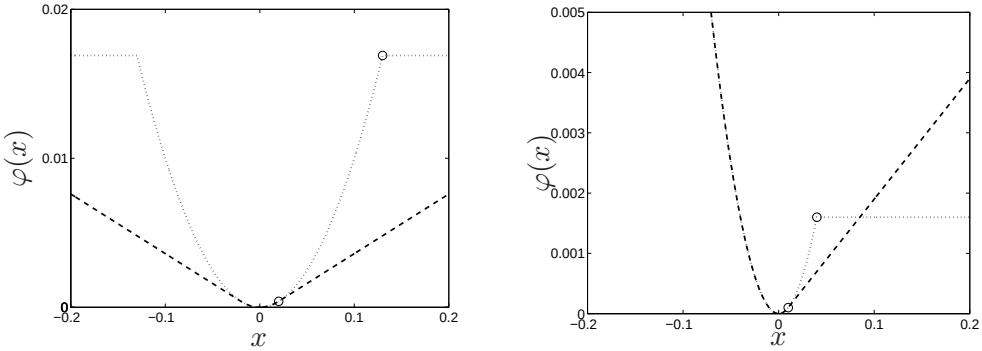
(a) Raies positives et négatives. (b) Raies positives.

FIG. 2.8 – Comparaison de l'estimation de la ligne de base avec les formes symétrique (- -) et asymétrique (. . .) de la parabole tronquée.



(a) Raies positives et négatives. (b) Raies seulement positives.

FIG. 2.9 – Comparaison des lignes de base estimées par la fonction de Huber (- -) et la parabole tronquée asymétrique (· · ·). La ligne de base réelle est représentée par une ligne pleine.



(a) Fonctions-coût de la figure 2.9(a). (b) Fonctions-coût de la figure 2.9(b).

FIG. 2.10 – Fonctions-coût correspondant aux estimations de la figure 2.9. Les cercles correspondent aux valeurs des seuils.

– $s = 0,13$ pour la parabole tronquée symétrique ;
– $s = 0,01$ pour la fonction de Huber asymétrique ;
– $s = 0,04$ pour la parabole tronquée asymétrique.

Les fonctions-coût correspondantes sont représentées figures 2.10(a) et 2.10(b). Pour les deux types de spectres, le seuil de la fonction de Huber est toujours inférieur au seuil de la parabole tronquée. Cela est dû au fait que, comme les raies sont influentes sur l'estimation, le seuil doit être plus petit afin de réduire leur coût.

2.4.2 Influence du taux de contamination

Nous définissons le *taux de contamination* Q comme le rapport entre le nombre de points appartenant à une raie sur le nombre total de points du spectre. Sur un spectre simulé, on considère qu'un point appartient à une raie si l'amplitude de ce point sur le spectre pur est supérieure à 1 % de l'amplitude maximale du spectre pur, soit :

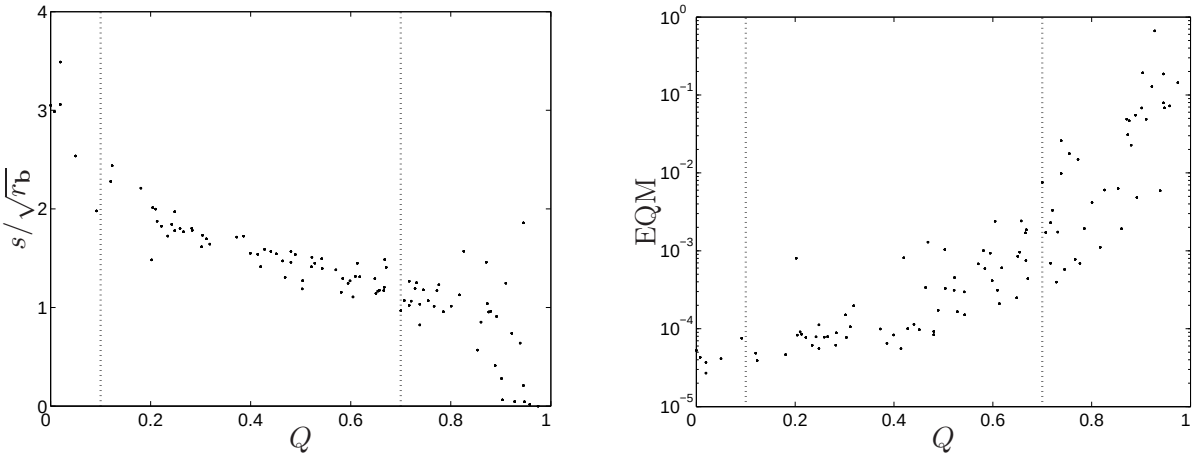
$$Q = \frac{1}{N} \sum_{n=1}^{N} \mathbb{1}_{\mathbf{v}_n > \max(\mathbf{v})/100}.$$

En pratique, on ne peut calculer qu'une valeur approchée du taux de contamination.

Ainsi, plus le nombre de raies ou plus la largeur de celles-ci est grande, plus le taux de contamination est grand et, inévitablement, plus il est difficile d'estimer correctement la ligne de base puisqu'il y a moins de points appartenant à la ligne de base seule.

Afin d'étudier l'influence du taux de contamination, la simulation suivante a été faite. Cent spectres purs (sans ligne de base ni bruit) ont été simulés. Pour chacun d'eux, dix lignes de base et dix bruits gaussiens centrés de variance $r_{\mathbf{b}}$ ont été générés puis ajoutés, résultant alors en dix spectres différents mais ayant le même taux de contamination. Puis la ligne de base a été estimée en minimisant la parabole tronquée asymétrique et la moyenne des dix seuils optimaux et des dix EQM ont été enregistrés.

La figure 2.11(a) représente le seuil normalisé $s/\sqrt{r_{\mathbf{b}}}$, et la figure 2.11(b) les performances en termes d'EQM en fonction du taux de contamination Q.



(a) Seuil optimal. (b) Erreur quadratique moyenne.

FIG. 2.11 – Influence du taux de contamination.

De ces simulations, nous pouvons définir trois parties :

– la première (pour $Q < 10$ % environ) correspond à un taux de contamination très bas, c'est-à-dire qu'il y a peu de raies et qu'elles sont de faible largeur. La valeur du seuil optimal peut être supérieure à deux fois l'écart-type du bruit sans que l'estimation en soit altérée, puisque l'EQM est relativement faible (inférieure à 10^{-4}) ;

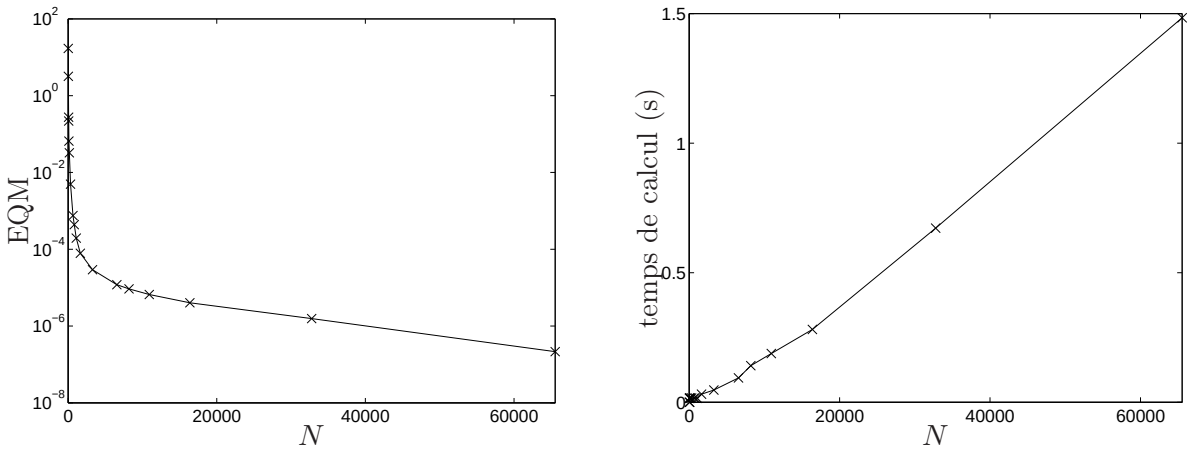
– la seconde (jusqu'à 70 % environ) correspond à la majorité des spectres. Elle fait apparaître une décroissante linéaire du rapport $s/\sqrt{r_\mathrm{b}}$ de 2 à 1, alors que l'EQM augmente doucement. La figure 2.11(a) peut donc aider dans le choix de la valeur du seuil ;
– la troisième partie (pour un taux de contamination supérieur à 70 % environ) correspond aux spectres ne contenant quasiment que des raies. Pour ceux-ci, il est très difficile d'estimer la ligne de base. La meilleure solution reste à diminuer le seuil a une valeur très petite afin d'obtenir une estimation très proche du bas du spectre. Cependant, cette estimation ne sera pas satisfaisante car l'EQM sera trop importante.

En conclusion, la méthode proposée estime correctement la ligne de base si le taux de contamination n'est pas trop grand (typiquement inférieur à 70 %). Dès lors, pour avoir un taux de contamination faible, il ne faut pas forcément travailler que sur la partie spectrale utile mais inclure également des zones sans raies.

2.4.3 Influence de la longueur du spectre

Afin d'étudier l'influence du nombre de points N sur la qualité de l'estimation, nous avons créé un spectre de $N = 65\,536$ points, qui a ensuite été ré-échantillonné plusieurs fois à des valeurs différentes afin d'obtenir des signaux ayant des longueurs différentes mais possédant la même ligne de base, le même niveau de bruit et le même taux de contamination.

Pour chaque signal, la ligne de base a été estimée en minimisant la parabole tronquée asymétrique et l'EQM correspondante a été calculée (voir figure 2.12(a)).



(a) Erreur quadratique moyenne. (b) Temps de calcul.

FIG. 2.12 – Influence de la longueur du spectre.

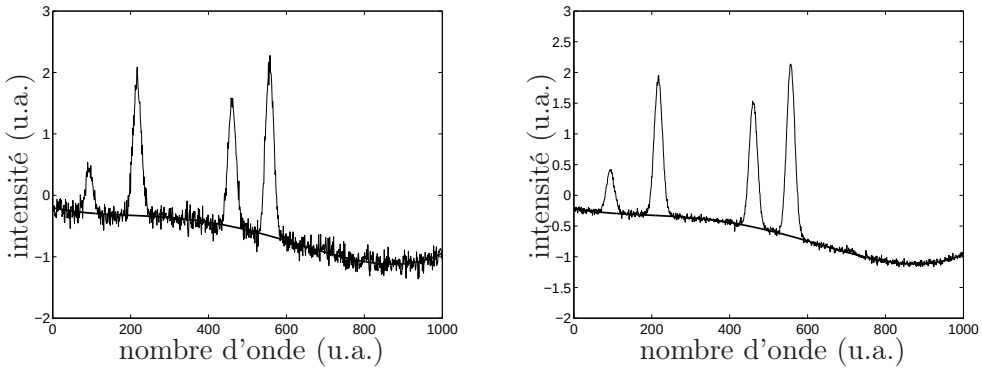
Il apparaît que la valeur de l'EQM diminue lorsque le nombre de points augmente, mais, à partir de quelques milliers de points, son évolution devient très faible (rappelons que l'axe des ordonnées de la figure 2.12(a) est en échelle logarithmique). On en conclue que la longueur du spectre (tant qu'elle est suffisamment grande) n'influe pas de manière significative sur la qualité de l'estimation.

La figure 2.12(b) représente le temps de calcul de l'algorithme en fonction du nombre de points N : le temps de calcul est approximativement linéaire par rapport au nombre de points et reste raisonnable même pour de grands spectres (moins de deux secondes pour un signal de 65 536 points).

2.4.4 Choix du seuil

Comme nous l'avons vu dans la section 2.4.2, le seuil s doit être choisi en fonction du taux de contamination Q et du niveau de bruit $\sqrt{r_\mathrm{b}}$ afin d'obtenir une estimation satisfaisante. Ce point

est illustré figure 2.13 qui présente deux signaux correspondants au même spectre (dont le taux de contamination est d'environ 35 %) mais avec deux niveaux de bruit différents. L'estimation de la ligne de base par la parabole tronquée asymétrique est représenté par une ligne pleine.



(a) Niveau de bruit important ($\sqrt{r_{\mathbf{b}}} \simeq 0,10, s = 0,24$).

(b) Niveau de bruit faible ($\sqrt{r_{\mathbf{b}}} \simeq 0,03, s = 0,05$).

FIG. 2.13 – Influence du bruit sur le choix du seuil.

Sur la figure 2.13(a), $\sqrt{r_{\mathbf{b}}}$ est fixé à 0,1, et le seuil optimal est égal à 0,24, soit environ 2,5 fois l'écart-type du bruit. Sur la figure 2.13(b), $\sqrt{r_{\mathbf{b}}}$ est fixé à 0,03, et le seuil optimal est égal à 0,05, c'est-à-dire environ 1,5 fois l'écart-type du bruit.

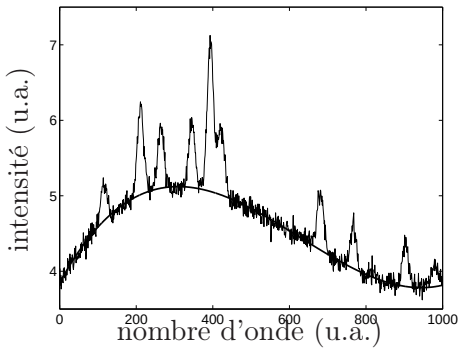
Ainsi, les valeurs optimales du seuil sont égales à environ deux fois l'écart-type du bruit, ce qui est du même ordre de grandeur que le seuil optimal de la figure 2.11(a) qui est légèrement plus petit que deux fois l'écart-type du bruit pour un taux de contamination correspondant à celui du spectre.

Afin d'étudier l'influence du seuil sur l'estimation, nous considérons la figure 2.14 qui représente un spectre avec un taux de contamination de 40 % et l'estimation obtenue avec la parabole tronquée pour trois seuils différents :

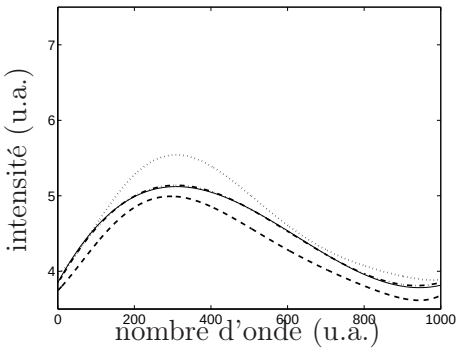
- le seuil qui donne la meilleure estimation est deux fois plus grand que l'écart-type du bruit (c'est à dire 0,14). Cela est en accord avec la figure 2.11(a) ;
- si le seuil fixé trop grand (ici, 10), alors l'estimation tend vers l'estimation des moindres carrés puisque la fonction-coût tend vers une parabole lorsque le seuil tend vers l'infini ;
- Au contraire, si le seuil est trop petit (ici, 0,1), alors l'estimation se déplace vers le bas du spectre. Cela est dû au fait qu'en choisissant une valeur du seuil faible, les coûts négatifs sont quadratiques alors que les coûts positifs tendent vers zéro. Ainsi, tous les points positifs ont un coût très proche de zéro et sont privilégiés par rapport aux points négatifs. Or, pour une fonction-coût symétrique, le résultat sera différent. En effet, lorsque le seuil tend vers zéro, la parabole tronquée symétrique tend vers une fonction-coût constante qui ne privilégie aucune estimation par rapport à la suivante : l'estimation obtenue est donc celle des moindres carrés, qui correspond à l'initialisation. Dans le cas de la fonction de Huber symétrique, l'estimation tend à minimiser l'erreur en valeur absolue, et pour laquelle l'estimation sera également non satisfaisante.

La figure 2.15 montre l'évolution de l'EQM en fonction du seuil normalisé $s/\sqrt{r_{\mathbf{b}}}$ pour trois taux de contamination différents (12 %, 36 % et 90 %). Cette simulation montre certains aspects de la robustesse par rapport au choix du seuil, complétant ainsi l'analyse faite section 2.4.2 :

- nous avons montré que pour un faible taux de contamination ($Q \leq 10$ %), le seuil peut être fixé plus grand que deux fois l'écart-type du bruit. Notons que, d'après la figure 2.15, un seuil plus grand ne dégrade pas l'estimation de manière significative puisque l'EQM reste à peu près constante au delà de deux fois l'écart-type du bruit. En fait, lorsque le seuil augmente, la fonction-coût tend vers une parabole et l'estimation tend alors vers l'estimation des moindres carrés. Or l'estimateur des moindres carrés est le meilleur estimateur (au sens de l'EQM) lorsqu'il n'y a pas



(a) Spectre simulé et ligne de base réelle.



(b) Ligne de base réelle et ses trois estimations ($-\cdot$: $s = 0,14$, $\cdots$: $s = 10$, - - : $s = 0,1$.

FIG. 2.14 – Estimation de la ligne de base pour trois seuils différents. La ligne de base réelle est représenté par une ligne pleine et $\sqrt{r_{\mathbf{b}}} \simeq 0,07$.

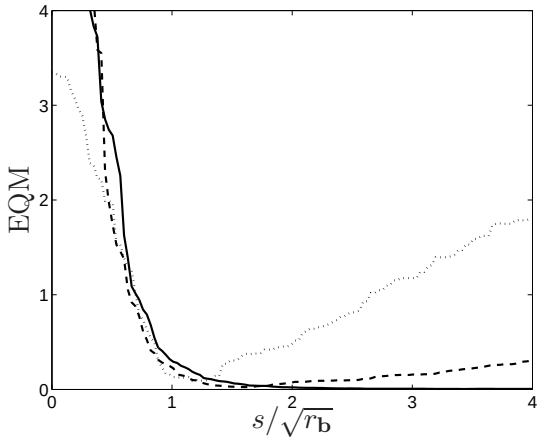


FIG. 2.15 – Évolution de l'EQM par rapport au seuil normalisé pour trois taux de contamination différents (— : 12 %, - - : 36 %, $\cdots$: 90 %).

de raies car alors tous les points du signal sont considérés comme des points de la ligne de base ;
- pour un taux de contamination moyen (10 % $\leq Q \leq$ 70 %), la qualité de l'estimation reste acceptable même lorsque le seuil s'éloigne de la valeur optimale. Par exemple, pour un taux de contamination de 36 %, le seuil peut être choisi dans l'intervalle $[\sqrt{r_{\mathbf{b}}}; 4\sqrt{r_{\mathbf{b}}}]$ sans affecter significativement l'EQM. Notons que l'intervalle des valeurs acceptables de s diminue quand le taux de contamination augmente ;
- enfin, pour un taux de contamination élevé ($Q \geq$ 70 %), la figure 2.15 montre que le seuil doit être choisi précisément, puisque l'EQM augmente significativement pour une faible variation de s. De toute façon, dans ce cas, la qualité de l'estimation est mauvaise, faisant apparaître les limites de la méthode proposée.

2.4.5 Choix de l'ordre du polynôme

Quelques méthodes ont envisagé d'estimer l'ordre du polynôme automatiquement, tel le critère AIC [2]. Cependant, nous croyons qu'il peut être choisi comme un paramètre défini par l'utilisateur, permettant ainsi de donner un degré de liberté à l'algorithme, car l'ordre du polynôme permet d'ajuster la douceur de la ligne de base estimée. Il est donc clair que γ doit être choisi en fonction de la ligne de base à estimer, mais également en fonction du besoin de l'utilisateur.

En effet, certains spectres peuvent avoir une zone qui peut être considérée comme faisant partie ou

non de la ligne de base. Par exemple, la figure 2.16 représente un spectre simulé où la « bosse » autour de 700 u.a. peut être interprétée comme étant une raie ou appartenant à la ligne de base. En ajustant l'ordre du polynôme, il est possible de tenir compte ou non de cette zone du spectre.

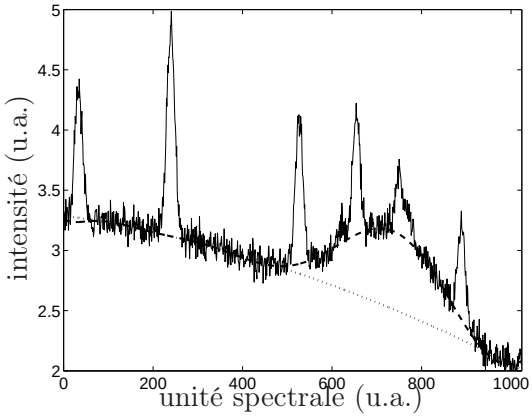


FIG. 2.16 – Estimation de la ligne de base pour deux ordre de polynôme différents ($\cdots$: $\gamma = 2$, - - : $\gamma = 10$).

2.4.6 Conclusion

Dans cette section, nous avons donné des indications sur le choix de la fonction-coût et des paramètres de la méthode ; nous avons également étudié l'influence du taux de contamination et de la longueur du spectre. Ainsi, la parabole tronquée asymétrique est à préférer pour la spectroscopie vibrationnelle, le seuil et l'ordre du polynôme permettant à l'utilisateur d'estimer la ligne de base souhaitée. Par ailleurs, pour obtenir une estimation correcte, il est conseillé de travailler sur des signaux suffisamment grands et possédant suffisamment de zones sans raies.

2.5 Application sur des spectres réels

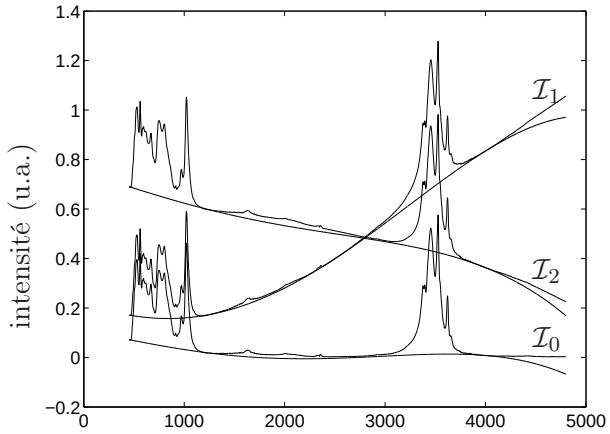
Nous avons montré dans les sections précédentes que la parabole tronquée asymétrique est la fonction-coût qui donne les meilleurs résultats pour estimer la ligne de base d'un spectre simulé contenant seulement des raies positives. Elle est maintenant appliquée sur les spectres infrarouge et Raman de gibbsite présentés section 1.2.

2.5.1 Spectres infrarouge

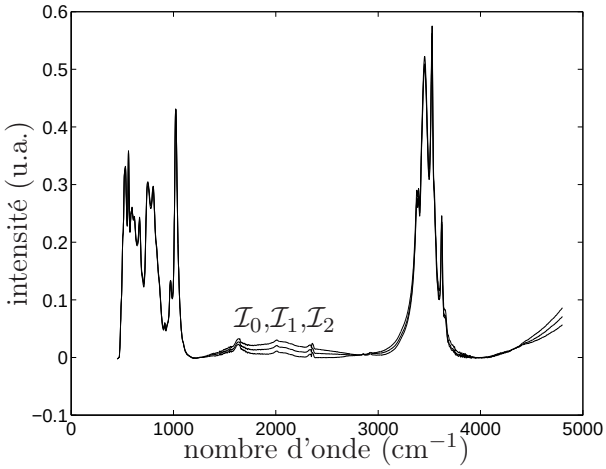
Comme les trois spectres infrarouges représentés figure 2.17(a) sont composés de la même quantité de rayonnement infrarouge mais de lignes de base différentes, les performances de la méthode peuvent être évaluées par son habilité à retourner seulement les raies du spectre infrarouge. Les lignes de base sont estimées par un polynôme d'ordre 4, et un seuil $s = 0{,}001$. Après avoir supprimé les lignes de base estimées (figure 2.17(b)), les trois spectres sont quasiment identiques, conformément à ce qui était attendu. Cependant, l'estimation obtenue ne permet pas de corriger parfaitement la ligne de base dans la zone située entre 1200 cm^{-1} et 3000 cm^{-1}, ainsi qu'au delà de 4500 cm^{-1}.

Pour que les trois spectres corrigés soient pratiquement identiques, il faudrait augmenter l'ordre du polynôme, mais alors les spectres corrigés auraient des valeurs négatives et la ligne de base risque d'être plus « chahutée » et donc moins conforme à l'estimation attendue.

Pour ces raisons, nous choisissons d'estimer la ligne de base de manière à corriger correctement les zones ayant un réel intérêt chimique. Ainsi, du point de vue du chimiste, le mauvais résultat de la correction de la ligne de base au delà de 4500 cm^{-1} n'est pas très important puisque cette zone ne



(a) Spectres expérimentaux et estimation des lignes de base.



(b) Spectres corrigés.

FIG. 2.17 – Correction de la ligne de base avec la parabole tronquée asymétrique sur des spectres infrarouge de gibbsite.

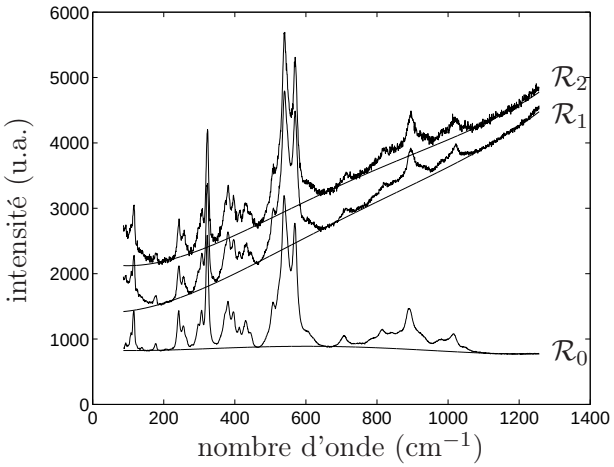
contient pas de raies. De même, la zone centrale du spectre correspond à des raies de vapeur d'eau et de dioxyde de carbone présents dans l'air : elle est donc sans information sur l'échantillon de gibbsite.

2.5.2 Spectres Raman

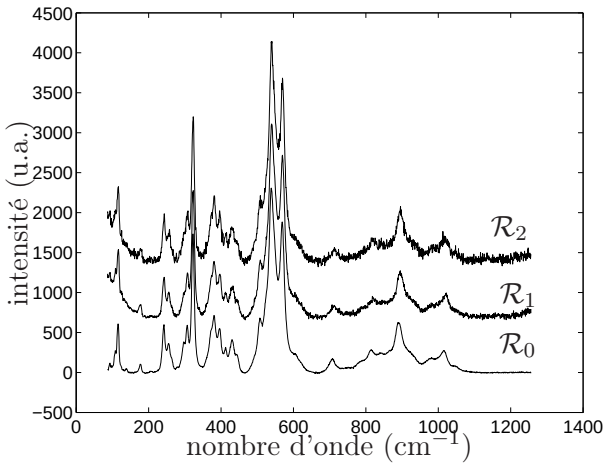
Pour les spectres Raman $\mathcal{R}_0$, $\mathcal{R}_1$ et $\mathcal{R}_2$, les lignes de base de fluorescence ont été estimées par des polynômes d'ordre 4 et des seuils respectifs de 10, 20 et 30 (figure 2.18). Bien que le spectre $\mathcal{R}_0$ corresponde au spectre de gibbsite pure, il existe une ligne de base qu'il a fallu supprimer pour que les trois spectres corrigés soient similaires.

En effet, les spectres Raman obtenus après correction de la ligne de base sont très proches les uns des autres (figure 2.18(b)) et correspondent aux spectres attendus.

Toutefois, la ligne de base due à la diffusion Rayleigh et qui est présente aux faibles nombres d'onde



(a) Spectres expérimentaux et estimation des lignes de base.



(b) Spectres corrigés (et décalés).

FIG. 2.18 – Correction de la ligne de base avec la parabole tronquée asymétrique sur des spectres Raman de gibbsite.

des spectres $\mathcal{R}_1$ et $\mathcal{R}_2$ n'est pas correctement estimée. Or, si l'ordre du polynôme est ajusté à une valeur plus grande afin d'estimer cette partie de la ligne de base, la nouvelle ligne de base estimée ne correspondra pas à celle attendue sur le reste du spectre, puisque, comme dans le cas des spectres infrarouges, elle sera plus irrégulière. Pour corriger ce problème, il existe deux possibilités. La première consiste à estimer une deuxième ligne de base sur la partie du spectre qui pose problème. La deuxième possibilité est, après estimation des raies, de considérer que les raies estimées les plus larges — donc celles évoluant le plus lentement — constituent en fait une ligne de base résiduelle. Cette dernière méthode est évidemment assez approximative, néanmoins elle peut être utilisée si rien d'autre ne peut être fait. Un exemple de cette possibilité est présenté sur ces spectres Raman, dans la section 4.6.2.

2.6 Conclusion

Dans ce chapitre, nous avons présenté une méthode itérative pour estimer la ligne de base de signaux spectroscopiques par le polynôme minimisant une fonction-coût appropriée au problème. La méthode est alors bien adaptée à une large gamme de spectroscopies (infrarouge, Raman, mais aussi UV-Visible, RMN, ...) puisqu'elle ne requiert aucun modèle sur les raies.

Plusieurs fonctions-coût ont été utilisées pour éviter aux raies d'être trop influentes sur l'estimation, comme c'est le cas de la parabole, qui est la fonction-coût de la méthode des moindres carrés. Ainsi, les fonctions de Huber et la parabole tronquée ont été proposées : elles sont quadratiques au voisinage de zéro (modélisant un bruit gaussien) et respectivement linéaires et constantes au delà d'un seuil (modélisant la distribution des raies). Mais comme elles ne sont pas quadratiques, le critère est minimisé par un algorithme de minimisation semi-quadratique.

De manière générale, la parabole tronquée donne les meilleurs estimations. De plus, les fonctions-coût asymétriques donnent les meilleurs résultats en spectroscopie optique où les raies sont toutes positives, alors que les fonctions-coût symétriques sont plus adaptées aux signaux où les raies peuvent être à la fois positives et négatives.

Deux paramètres doivent être fixés pour implémenter cette méthode : le seuil de la fonction-coût et l'ordre du polynôme.

La valeur optimale du seuil dépend de l'écart-type du bruit et du taux de contamination. Dans la majorité des cas pratiques, le seuil peut être fixé entre une et deux fois l'écart-type du bruit pour donner une estimation satisfaisante, mais il n'y a pas d'argument formel pour justifier cette règle empirique. Cependant, la robustesse de la méthode à cet intervalle de valeur a été confirmée expérimentalement, tant que le taux de contamination reste suffisamment faible.

Concernant l'ordre du polynôme, il doit être choisi en fonction de la douceur de la ligne de base à estimer. Même si certaines méthodes peuvent être développées pour l'estimer, nous pensons qu'il peut être laissé au choix de l'utilisateur, permettant ainsi de considérer les « bosses » du spectres comme faisant partie de la ligne de base ou non.

Des tests ont été faits pour la correction de la ligne de base sur des spectres Raman et infrarouges : dans tous les cas traités, la méthode a donné des résultats satisfaisants et exploitables.

De plus, l'algorithme reste très rapide, même pour des signaux très grands.

Enfin, même si la ligne de base de la plupart des spectres est correctement estimée, l'utilisation de splines plutôt qu'un polynôme d'ordre fixé pourrait permettre d'estimer des lignes de base plus irrégulières, mais alors l'estimation de l'emplacement des abscisses risquerait de compliquer sérieusement la méthode.

Une autre perspective de travail consiste à combiner les fonctions-coût proposées dans ce chapitre avec le paramètre de régularisation des critères de Eilers et Koenker *et al.* (voir section 2.3.4.1). En effet, cela permettrait d'estimer la ligne de base par une courbe quelconque : il n'y aurait alors pas d'ordre de polynôme à fixer. L'intérêt est de pouvoir faire varier continûment la ligne de base estimée, alors

qu'avec une ligne de base polynômiale, l'évolution n'est pas toujours continue. Cete procédure serait alors très pratique pour éviter les problèmes rencontrés lors de l'estimation des lignes de base réelles (section 2.5). Par ailleurs, la minimisation semi-quadratique donne ici une approche très attractive pour la minimisation de ce nouveau critère.

Chapitre 3

Estimation des raies par déconvolution impulsionnelle positive myope

3.1 Formulation du problème

3.1.1 Modélisation du spectre

Dans le chapitre précédent, nous avons proposé une méthode originale pour corriger la ligne de base d'un spectre. Nous pouvons maintenant aborder le cœur du problème : l'estimation des raies du spectre, c'est-à-dire l'estimation de leurs nombre d'onde, intensité et paramètres de forme. L'approche présentée dans ce chapitre a fait l'objet d'une publication en congrès [78].

Désormais, les données sont le spectre bruité $\mathbf{w}$ qui correspond au spectre $\mathbf{y}$ dont on a corrigé la ligne base (s'il y en a une). Par ailleurs, comme nous l'avons précisé dans la section 1.7.1, nous considérons uniquement le cas de raies de forme lorentzienne : le choix d'une forme différente, telle une gaussienne ou une pseudo-Voigt, sera traité de manière équivalente. Par conséquent, l'estimation des paramètres de forme des raies revient seulement à estimer la largeur s de la lorentzienne.

Dans ce chapitre, nous considérons l'estimation des paramètres du spectre comme un problème de déconvolution impulsionnelle positive myope[1] car nous pouvons montrer, sous certaines hypothèses simplificatrices, que le spectre pur $\mathbf{v}$ correspond à la convolution entre un signal impulsionnel positif $\mathbf{x}$ et une réponse impulsionnelle lorentzienne $\mathbf{h}$. En effet, rappelons l'expression du spectre pur (équation (1.2)) :

$$\forall n \in \{1, \dots, N\}, \qquad \mathbf{v}_n = \sum_{k=1}^{K} \mathbf{a}_k \frac{\mathbf{s}_k^2}{\mathbf{s}_k^2 + (n - \mathbf{c}_k)^2}.$$

Une première simplification est de supposer toutes les raies identiques. Dès lors, on a pour tout $k \in \{1, \dots, K\}$, $\mathbf{s}_k = s$. Une deuxième simplification consiste à discrétiser les positions des pics, ce qui revient à poser pour tout k : $\mathbf{c}_k = m \in \{1, \dots, N\}$. L'équation précédente se réécrit alors :

$$\mathbf{v}_n = \sum_{m=1}^{N} \mathbf{x}_m \frac{s^2}{s^2 + (n - m)^2},$$

où $\mathbf{x}_m$ code l'absence ou la présence — et dans ce cas l'amplitude (positive) — d'une raie lorentzienne centrée en m. Dès lors, en notant, pour tout $n \in \{1, \dots, N\}$,

$$\mathbf{h}_n = \frac{s^2}{s^2 + n^2},$$

[1] On préférera le terme « myope » à « aveugle » car on considère que la déconvolution impulsionnelle aveugle correspond au cas où absolument aucune connaissance a priori sur la réponse impulsionnelle n'est disponible. Si tel était le cas, le problème serais insoluble (voir en particulier le problème des indéterminations, section 3.2).

on obtient :

$$\mathbf{v}_n = \sum_{m=1}^{N} \mathbf{x}_m \mathbf{h}_{n-m}.$$

Cette expression correspond à la convolution du signal $\mathbf{x}$ et de la réponse impulsionnelle $\mathbf{h}$ lorentzienne :

$$\mathbf{v} = \mathbf{x} \star \mathbf{h}$$

que l'on peut écrire sous forme matricielle :

$$\mathbf{v} = \mathbf{H}\mathbf{x},$$

où $\mathbf{H} \in \mathbb{R}^{N \times N}$ est la matrice de Toeplitz de la réponse impulsionnelle $\mathbf{h}$:

$$\mathbf{H} = \begin{pmatrix} \mathbf{h}_0 & \dots & \mathbf{h}_{-N+1} \\ \vdots & \ddots & \vdots \\ \mathbf{h}_{N-1} & \dots & \mathbf{h}_0 \end{pmatrix}$$

La figure 3.1 illustre cette modélisation en considérant des signaux continus.

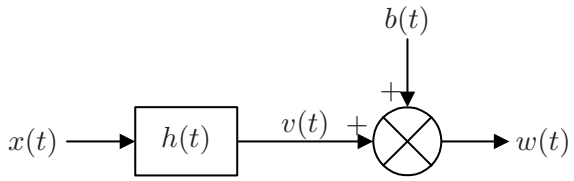


FIG. 3.1 – Déconvolution impulsionnelle positive myope.

Le problème de l'estimation des raies du spectre devient alors un problème de déconvolution impulsionnelle positive myope puisqu'il faut estimer le signal impulsionnel $\mathbf{x}$ et la réponse impulsionnelle $\mathbf{h}$; c'est donc un problème de détection-estimation dans lequel il faut détecter les pics et estimer leur amplitude [59]. En outre, nous considérons le cas non supervisé dans lequel les différents hyperparamètres du problème sont inconnus et doivent également être estimés.

En résumé, l'approche par déconvolution impulsionnelle positive myope suppose deux hypothèses simplificatrices importantes : d'une part que les raies sont toutes de même forme et de mêmes paramètres de forme, d'autre part qu'elles sont situées aux « instants » d'échantillonnage. L'approche considérée dans ce chapitre n'est donc qu'une modélisation grossière du spectre, mais elle consiste en une approximation simple de la réalité. Par ailleurs, la déconvolution impulsionnelle myope fournit un cadre d'étude largement étudié car cette modélisation est très répandue.

Cependant, on peut s'attendre à deux problèmes majeurs. Le premier est que les positions des raies estimées ne seront pas précises puisqu'elles se situeront en des nombres d'onde forcément discrets. Le deuxième problème est que, si le spectre présente des raies de forme ou de paramètres de forme différents, alors les raies larges seront estimées par plusieurs pics. Pour ces raisons, il convient donc d'analyser l'estimation obtenue avec prudence (cf. section 3.6).

3.1.2 Méthodes de déconvolution impulsionnelle myope

Le problème de la déconvolution impulsionnelle myope a tout d'abord été résolu à l'aide d'algorithmes déterministes. Les premières applications ont été développées en géophysique où l'on cherchait à estimer la réflectivité $\mathbf{x}$ du sol et l'ondelette incidente $\mathbf{h}$.

Ainsi, un algorithme de descente par bloc est proposé par Chi *et al.* [21]. Cet algorithme estime la réflectivité, l'ondelette et les paramètres du problème (excepté la variance des amplitudes des pics) au

sens du maximum de vraisemblance. La réflectivité est modélisée par un processus Bernoulli-gaussien (voir section 3.3.1.1) et le modèle général est traité à l'aide d'une représentation d'état.

Gautier *et al.* [37] modélisent la réflectivité par un processus blanc de loi gaussienne généralisée et l'ondelette est supposée douce et centrée ; le bruit est supposé gaussien. Le critère composite s'écrit donc, à partir des quatre hypothèses précédentes :

$$\mathcal{J}(\mathbf{x},\mathbf{h}) = ||\mathbf{w}-\mathbf{H}\mathbf{x}||^2 + \alpha\sum_n |\mathbf{x}_n|^p + \beta\mathbf{h}^T\mathbf{Q}\mathbf{h} + \gamma\mathbf{h}^T\mathbf{h},$$

où α, β et γ sont des paramètres de régularisation et $\mathbf{Q}$ est une matrice qui permet d'introduire un a priori de douceur (on peut prendre par exemple une matrice de Toeplitz d'un filtre dérivé). En ce qui concerne la réflectivité, l'approche choisie est la déconvolution L2LP (pour plus de détails, consulter [59, section 5.2.3]).

Kaaresen *et al.* [63] proposent un algorithme déterministe de descente par bloc, alternant une étape d'estimation de la réflectivité et une étape d'estimation MAP de l'ondelette. L'estimation de la réflectivité est calculée grâce à l'algorithme IWM (*Iterated Window Maximization*), qui est un algorithme de maximisation de vraisemblance généralisée, c'est-à-dire qu'il consiste à maximiser la probabilité de toutes les quantités probabilistes connaissant les données et les paramètres déterministes. L'originalité de l'approche est de décomposer le signal sur plusieurs fenêtres, ce choix est justifié par le fait que lorsqu'un pic apparaît, la réflectivité ne sera modifiée qu'au voisinage de ce pic. Cette approche fournit de bons résultats sur des longs signaux, mais le nombre de pics est supposé connu.

Des travaux plus récents ont abordé le problème de la déconvolution impulsionnelle myope en faisant appel à des algorithmes stochastiques.

Ainsi, Rosec *et al.* [96, 97] proposent d'utiliser la déconvolution impulsionnelle myope pour l'amélioration de la résolution d'images sismiques marines. Pour cela, la réflectivité est modélisée par un mélange de deux gaussiennes (généralisant le modèle Bernoulli-gaussien) et la réponse impulsionnelle par un processus MA.

Dans un premier temps, les auteurs proposent d'estimer les paramètres du problème (c'est-à-dire les hyperparamètres et les valeurs de la réponse impulsionnelle). Deux versions de l'algorithme sont proposées. Dans la première, les paramètres sont estimés par maximum de vraisemblance, à l'aide d'une version stochastique de l'algorithme EM (SEM [12] ou SAEM [13]). Le choix de cet algorithme est motivé par la complexité numérique de l'optimisation. La deuxième version correspond à une approche bayésienne en posant des lois a priori sur les paramètres. L'utilisation de l'échantillonneur de Gibbs permet alors de simuler les lois des paramètres, ces paramètres sont ensuite estimés au sens de l'EAP (espérance a posteriori).

Une fois les paramètres estimés, il reste le problème de la déconvolution proprement dit. Dans un contexte déterministe, l'estimation de $\mathbf{x}$ fait face à un problème de combinatoire car il faut tester les 2^N combinaisons que peut prendre le vecteur codant l'occurrence des pics. Pour éviter cela, les auteurs proposent plusieurs algorithmes qui maximisent la loi a posteriori, comme l'algorithme ICM (*Iterative Conditionnal Mode*) [5], MPM (*Maximum Posterior Mode*) [15] ou du recuit simulé [93].

En particulier, l'algorithme MPM consiste à simuler le signal $\mathbf{x}$ grâce à l'échantillonneur de Gibbs, puis à prendre une décision pour obtenir un estimateur impulsionnel du signal. Nous proposons dans ce chapitre de simuler le signal de la façon similaire, mais le calcul de l'estimateur est différent. Une discussion à propos de la méthode de décision de Rosec *et al.* est proposée dans la section 3.5.

On peut aussi résoudre le problème de la déconvolution impulsionnelle myope dans un cadre totalement bayésien en estimant conjointement le signal et les paramètres du problème.

Ainsi, en sismique-réflexion, Cheng *et al.* [20] ont modélisé la séquence de réflexion $\mathbf{x}$ par un processus Bernoulli-gaussien et le vecteur $\mathbf{h}$ des amplitudes de l'ondelette est distribué suivant une gaussienne dont la moyenne et la matrice de covariance sont fixées. Notons que les auteurs ne considèrent pas explicitement l'existence de la variable cachée $\mathbf{q}$ représentant l'occurrence des pics du processus

Bernoulli-gaussien. Les auteurs utilisent ensuite l'échantillonneur de Gibbs pour éviter la minimisation directe trop compliquée de la loi a posteriori. Enfin, Cheng *et al.* proposent d'estimer $\mathbf{x}_n$ au sens de l'EAP. Nous verrons dans la section 3.5 que cet estimateur peut être inadéquat.

3.1.3 Approche retenue

Inférence bayésienne

Un problème est bien posé au sens de Hadamard si la solution existe, est unique et continue par rapport au données (une petite perturbation sur les données n'engendre pas une grande perturbation sur la solution) [26, 51, 59]. Or, l'estimation des raies d'un spectre est typiquement un problème inverse mal posé. La résolution de ce type de problème impose donc d'inclure des connaissances a priori fortes pour limiter l'espace des solutions, comme par exemple imposer de restaurer un signal impulsionnel ou considérer la réponse impulsionnelle de forme connue. Cela consiste en la régularisation du problème.

Dans ce contexte, l'approche bayésienne fournit un cadre particulièrement attractif pour résoudre ce type de problème [59, chapitre 3]. En effet, elle permet d'inclure dans le modèle des connaissances (voire des méconnaissances) afin de limiter le domaine des solutions ; ces connaissances peuvent d'ailleurs être qualitatives, comme par exemple le caractère impulsionnel de $\mathbf{x}$. En outre, et contrairement aux méthodes de régularisation plus « classiques », elle offre des réponses au choix des hyperparamètres du problème et de l'optimisation du critère [59]. Le problème de l'estimation des raies étant relativement complexe, l'inférence bayésienne apporte donc une méthodologie simple et efficace.

En conséquence, le choix de traiter le problème dans un cadre totalement bayésien où toutes les inconnues sont estimées en même temps nous a séduit par la facilité de sa mise en œuvre.

Modèle hiérarchique

L'approche bayésienne nécessite de fixer des lois a priori sur les variables d'intérêt (qui sont ici $\mathbf{x}$ et $\mathbf{h}$). Or, il est rare que l'information connue a priori soit suffisamment riche pour définir exactement une loi a priori. Par exemple, nous choisissons le modèle traditionnel d'un processus Bernoulli-gaussien pour modéliser $\mathbf{x}$ (voir section 3.3.1.1), mais nous ne possédons pas de connaissance suffisante pour fixer exactement les paramètres de cette loi. Il est alors intéressant d'inclure cette méconnaissance dans le modèle bayésien et d'estimer les paramètres (le problème est donc non supervisé).

Cela est possible en fixant une nouvelle loi a priori (une *hyper-loi a priori*) sur ces *hyperparamètres* : le modèle possède alors une couche supplémentaire contenant les paramètres de cette hyper-loi. Ce modèle en couches est appelé dans la littérature *modèle hiérarchique* (voir [44, chapitres 2 et 19] et [94, chapitre 10]) et la loi $p(\mathbf{w})$ se décompose en :

$$p(\mathbf{w}) = \int p(\mathbf{w}|\mathbf{x}, s, r_{\mathbf{b}}) p(\mathbf{x}|\lambda, r_{\mathbf{a}}) p(s) p(\lambda) p(r_{\mathbf{a}}) p(r_{\mathbf{b}}) d\mathbf{x}\, ds\, d\lambda\, dr_{\mathbf{a}}\, dr_{\mathbf{b}}$$

où λ et $r_{\mathbf{a}}$ sont les paramètres du processus Bernoulli-gaussien et $r_{\mathbf{b}}$ est la variance du bruit.

Il est possible d'utiliser autant de couches que l'on souhaite, mais un modèle hiérarchique de deux couches suffit dans la majorité des cas. On introduit généralement des lois peu informatives au dernier niveau de la hiérarchie.

L'approche bayésienne hiérarchique s'oppose à l'approche bayésienne empirique pour laquelle les paramètres des lois a priori sont fixés à partir des observations (ce qui revient à leur affecter une impulsion de Dirac comme loi a priori), tandis que l'approche bayésienne hiérarchique modélise le manque d'information par des lois a priori moins informatives qu'une impulsion de Dirac, rendant ainsi l'analyse bayésienne plus robuste par rapport aux choix arbitraires des hyperparamètres. En d'autres termes, la solution sera moins sensible aux variations de valeur des hyperparamètres.

La loi des hyperparamètres est généralement une loi conjuguée ou une loi non-informative [94, p. 464].

- Une loi a priori conjuguée permet d'obtenir une loi a posteriori conditionnelle de même forme, mais dont les paramètres sont fonction de la vraisemblance et de l'a priori. Par exemple, si la fonction de vraisemblance est une loi normale $x|\sigma \sim \mathcal{N}(0,\sigma)$, la loi a priori conjuguée naturelle sur σ est une loi inverse gamma $\sigma \sim \mathcal{IG}(\alpha,\beta)$; ainsi, la loi a posteriori est la loi inverse gamma $\sigma|x \sim \mathcal{IG}(\alpha+1/2, x^2/2+\beta)$. Dans [94, p. 121], on peut trouver un tableau regroupant les lois conjuguées pour différentes fonctions de vraisemblance.

  L'intérêt des lois conjuguées est, entre autre, de limiter l'influence des données x (puisque seuls les paramètres de la loi a posteriori varient, la forme restant identique à l'a priori), mais surtout de simplifier la charge numérique [94, section 3.3.2]. Cependant, les lois conjuguées ne sont pas forcément les lois a priori les plus robustes et, de plus, dans le cas d'un manque total d'information, l'intérêt des lois a priori conjuguées n'est que purement analytique : il est impossible de justifier leur choix et les hyperparamètres ne peuvent être déterminés.

- Aussi, sans connaissance a priori, la loi a priori ne peut donc être déduite que de la fonction de vraisemblance, puisque c'est la seule information valable : on l'appelle alors loi non-informative [94, section 3.5]. Cependant, malgré leur dénomination, les lois non informatives intègrent une certaine connaissance dans le modèle : c'est pourquoi elles doivent être considérées plutôt comme des lois a priori par défaut si aucune connaissance n'est disponible. Ainsi, certaines lois non-informatives peuvent être plus ou moins performantes que d'autres, mais elle n'en sont pas pour autant moins informatives [94, p. 127].

  La loi non-informative la plus simple est bien évidemment l'a priori de Laplace qui est une loi uniforme. Elle a cependant l'inconvénient d'être impropre et non-invariante en cas de changement de variable. En effet, si l'a priori sur une variable quelconque θ est uniforme, alors l'a priori sur $\eta = g(\theta)$ devrait l'être également, ce qui n'est pas le cas puisque $\pi(\eta) = \left|\frac{d}{d\eta} g^{-1}(\eta)\right|$. C'est pour cette raison que, en cas d'absence d'information sur les hyperparamètres du problème, l'approche bayésienne traditionnelle est celle de Jeffreys [64, 94]. Celle-ci consiste à poser comme loi a priori sur les paramètres θ la mesure induite par l'information de Fisher :

$$p(\theta) \propto I(\theta)^{1/2} = \mathbb{E}_{x|\theta}\left[\left(\frac{\partial \ln p(x|\theta)}{\partial \theta}\right)^2\right]^{1/2}.$$

  Toutefois, Jeffreys précise que cette « règle générale » peut conduire à des lois a priori qui contredisent d'autres principes d'invariance (en particulier lorsque la dimension du paramètre θ augmente) [64, 61, p. 182]. Aussi, Jeffreys recommande de considérer les paramètres de position (comme la moyenne) indépendants des paramètres d'échelle (écart-type, variance, ...) et donc de les traiter séparément. Plus généralement, lorsque le problème fait apparaître d'autres types de paramètres, Jeffreys propose de considérer les paramètres de position séparément des autres paramètres.

  Dans [113], on peut trouver un catalogue de lois non-informatives.

Le choix entre une loi conjuguée ou non-informative doit donc prendre en compte le niveau de connaissance sur l'hyperparamètre et le degré de complexité de l'algorithme que l'on est prêt à assumer. Quoi qu'il en soit, dans la majorité des cas, la loi a priori est suffisamment non-informative pour ne pas avoir une influence très importante sur l'estimation.

Dans ces conditions, nous proposons dans la suite de poser des lois non-informatives de Jeffreys sur les hyperparamètres qui est le choix traditionnel. Cependant, le choix d'une autre loi a priori sera envisagé dans le cas où l'a priori de Jeffreys ne convient pas (c'est le cas pour s, voir plus loin), ou bien lorsqu'une dégénérescence a pu être observée. Dans ce dernier cas, nous proposons d'utiliser une loi a priori conjuguée afin d'éviter cette dégénérescence et de faciliter les calculs.

Méthode d'optimisation

Le modèle probabiliste étant obtenu, il reste à optimiser la loi a posteriori jointe afin de calculer les estimations des inconnues $\boldsymbol{\theta}$. Du fait de sa complexité, la loi a posteriori ne permet pas d'obtenir des expressions explicites des variables qui l'optimisent. Par ailleurs, une optimisation globale est également difficilement réalisable.

Aussi, l'utilisation de l'algorithme EM [27] pourrait être envisagée. Cet algorithme itératif permet d'estimer le vecteur $\boldsymbol{\theta}$ en introduisant des variables cachées $\mathbf{x}$. Dans une première étape, l'espérance conditionnelle de la log-vraisemblance de $\boldsymbol{\theta}$ est calculée (étape E : *expectation*) :

$$\begin{aligned} Q(\boldsymbol{\theta}, \boldsymbol{\theta}^{(i-1)}) &= \mathbb{E}_{\mathbf{x}|\mathbf{y},\boldsymbol{\theta}^{(i-1)}}\big[\log p(\mathbf{y}, \mathbf{x}|\boldsymbol{\theta})\big], \\ &= \int \log p(\mathbf{y}, \mathbf{x}|\boldsymbol{\theta}) p(\mathbf{x}|\mathbf{y}, \boldsymbol{\theta}^{(i-1)}) d\mathbf{x}. \end{aligned}$$

Puis, une estimation $\boldsymbol{\theta}^{(i)}$ est obtenue par maximisation de cette espérance en $\boldsymbol{\theta}$ (étape M : *maximization*). Cet algorithme est donc déterministe et tend à augmenter la vraisemblance à chaque itération ; il ne garantit donc pas la convergence vers l'optimum global. Enfin, une fois que $\boldsymbol{\theta}$ est estimé par la dernière valeur de $\boldsymbol{\theta}^{(i)}$, il faut estimer $\mathbf{x}$. Puisque les hyperparamètres sont maintenant connus, le problème est simplifié : on peut par exemple utiliser l'algorithme SMLR (*Single Most Likely Replacement*) de Kormylo et Mendel [66, 17] ou la méthode HT (*Hunt Threshold*) de Mazet *et al.* [73][2]. Toutefois, l'algorithme EM nécessite le calcul d'une espérance et donc d'intégrales qui peuvent être de dimension très grande.

Il existe des versions stochastiques de cet algorithme qui permettent de remplacer le calcul de l'espérance par une simulation de la loi $p(\mathbf{x}|\mathbf{y}, \boldsymbol{\theta}^{(i-1)})$, puis en approchant l'estimateur $\widehat{\boldsymbol{\theta}}$ par la moyenne des variables $\boldsymbol{\theta}^{(i)}$. En particulier, l'algorithme SEM [12] permet de ne pas être attiré par les maxima locaux. En revanche, peu de résultats de convergence sont disponibles.

Cependant, lorsque ni le vecteur $\boldsymbol{\theta}$, ni les variables cachées $\mathbf{x}$ ne peuvent être estimées à partir d'une expression explicite, les méthodes du type EM n'ont plus d'intérêt. Les techniques de simulations stochastiques par chaînes de Markov (algorithmes MCMC) s'avèrent être alors un choix pertinent et élégant (voir annexe A). En effet, compte-tenu de la factorisation de $p(\mathbf{w})$, l'utilisation de l'échantillonneur de Gibbs est facile et naturelle ; il a l'avantage de pouvoir estimer ensemble le signal impulsionnel $\mathbf{x}$, la réponse impulsionnelle $\mathbf{h}$ et les hyperparamètres du problème (en fait, aucune distinction n'est faite entre variables de la première couche du modèle hiérarchique et hyperparamètres). Cependant, il convient de faire attention à l'intégrabilité des lois a posteriori (voir annexe A, page 130). Par ailleurs, il a été montré que, dans le cas de l'estimation des paramètres de mélange de gaussiennes — qui est un problème semblable à celui traité dans les chapitres 3 et 4 (cf. section 4.1.3.1), la convergence des méthodes MCMC est meilleure que celle des algorithmes EM et SEM dans le sens où, en moyenne, les chaînes de Markov construites à l'aide des méthodes MCMC convergent plus souvent et plus rapidement vers l'optimum global que les deux méthodes précédentes [30].

3.1.4 Organisation du chapitre

Le chapitre est organisé comme suit. Tout d'abord, nous abordons l'un des problèmes majeurs en déconvolution myope : le problème d'indétermination. Par exemple, comme $\mathbf{x} \star \mathbf{h} = \mathbf{h} \star \mathbf{x}$, il est impossible de différencier lequel de ces deux signaux correspond à l'entrée du système et lequel correspond au filtre. Cette indétermination est facilement levée puisqu'on suppose que $\mathbf{x}$ est impulsionnel et que $\mathbf{h}$ est une lorentzienne. Cependant, il convient de fixer d'autres conditions pour lever toutes les indéterminations en déconvolution myope : c'est l'objectif de la section 3.2.

Nous pouvons alors traiter le problème de l'estimation des raies du spectre en considérant le problème de la déconvolution impulsionnelle positive myope non supervisée. Nous proposons donc dans

[2] Cet article fait suite aux recherches effectuées en DEA et est reproduit dans l'annexe E.

la section 3.3 un modèle probabiliste pour obtenir les lois a posteriori conditionnelles des différentes variables à estimer.

L'échantillonneur de Gibbs permet alors de simuler chaque variable suivant sa loi a posteriori. Il n'est pas toujours évident de simuler de telles lois. Aussi, l'implémentation de l'échantillonneur de Gibbs et les choix adoptés sont discutés dans la section 3.4.

L'échantillonneur de Gibbs génère une chaîne de Markov pour chaque variable, il faut alors trouver un estimateur qui permet d'obtenir un résultat exploitable. Dans le cas des paramètres du filtre et des hyperparamètres, une estimation au sens de l'espérance a posteriori convient : elle est approchée en calculant simplement la moyenne de la chaîne de Markov. Par contre, l'estimation du signal impulsionnel est moins évidente. Aussi, après avoir présenté quelques estimateurs existants, nous proposons dans la section 3.5 un estimateur original qui nous semble plus pertinent.

Nous pouvons alors tester l'approche proposée dans ce chapitre sur un signal simulé (section 3.6). Ce signal est construit de telle manière à faire apparaître les principales difficultés auxquelles peut être confrontée une méthode d'estimation des raies d'un spectre. Nous montrons en particulier que l'estimateur proposé dans la section 3.5 s'avère plus pertinent que ceux proposés dans la littérature. La méthode proposée est également testée sur un signal simulé où les raies ne sont pas toutes de même largeur : cela permet d'étudier le comportement de la méthode face à ce type de spectre.

Enfin, la section 3.7 conclue ce chapitre.

3.2 Indéterminations en déconvolution myope

En l'absence de toute connaissance a priori sur l'entrée x et le filtre h d'un système, la déconvolution myope ne permet pas de discriminer les couples

$$(x, h) \qquad \text{et} \qquad (x \star h_1, h_2)$$

avec $h = h_1 \star h_2$. Une solution extrême pourrait même être $(x \star h, \delta)$ où la réponse impulsionnelle est le filtre identité. Ce problème, appelé *indétermination sur le contenu spectral*, est illustré par la figure 3.2.

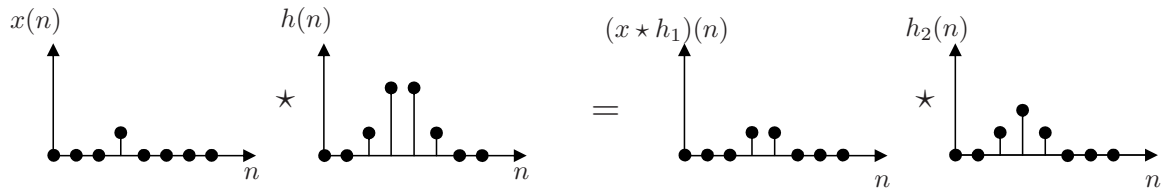


FIG. 3.2 – Exemple d'indétermination sur le contenu spectral.

Dans notre cas, l'hypothèse de blancheur de l'entrée (due à la modélisation Bernoulli-gaussienne, voir section suivante) permet de lever cette indétermination, sauf si, bien sûr, le filtre h_1 est un filtre passe-tout.

On peut distinguer trois cas particuliers d'indétermination sur le contenu spectral, détaillés ci-dessous [59]. En particulier, le problème des indéterminations temporelle et en amplitude a déjà été signalé dans de nombreux travaux [20, 96], sans pour autant être étudié plus en détail.

Indétermination sur la phase

Ici, l'entrée et la réponse impulsionnelle peuvent être déphasés d'une valeur opposée, fréquence par fréquence. Dans le domaine fréquentiel, on ne peut pas différencier les couples :

$$\big(X(f), H(f)\big) \qquad \text{et} \qquad \big(X(f)\mathrm{e}^{i\varphi(f)}, H(f)e^{-i\varphi(f)}\big),$$

où $\varphi(f)$ est une fonction quelconque. L'indétermination sur la phase correspond à un cas particulier d'indétermination sur le contenu spectral avec, dans le domaine fréquentiel :

$$H_1(f) = e^{+i\varphi(f)} \qquad \text{et} \qquad H_2(f) = H(f)e^{-i\varphi(f)}.$$

La figure 3.3 présente un cas d'indétermination sur la phase.

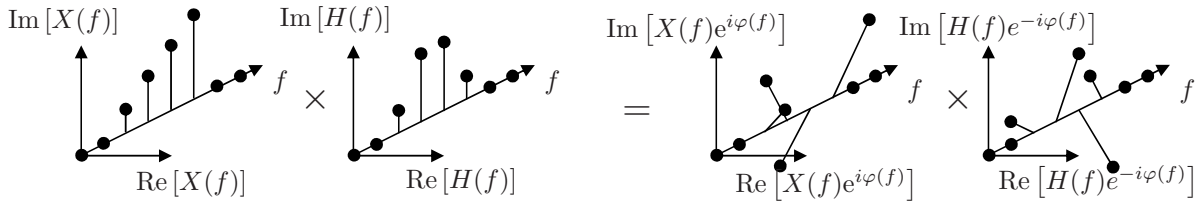


FIG. 3.3 – Exemple d'indétermination sur la phase.

Encore une fois, l'hypothèse d'un signal Bernoulli-gaussien permet de lever cette indétermination. En effet, un signal i.i.d. Bernoulli-gaussien convolué par un filtre passe-tout $\mathrm{e}^{i\varphi(f)}$ n'est plus Bernoulli-gaussien.

Indétermination sur l'amplitude

Dans ce cas, on ne peut pas différencier les cas

$$(x, h) \qquad \text{et} \qquad (\alpha x, h/\alpha).$$

On peut interpréter l'indétermination sur l'amplitude comme un cas particulier d'indétermination sur le contenu spectral, où $h_1 = \alpha$ et $h_2 = h/\alpha$. Dans le cas où $\alpha = -1$, on peut également l'interpréter comme une indétermination sur la phase avec $\varphi(f) = \pi$. La figure 3.4 présente un cas d'indétermination sur l'amplitude.

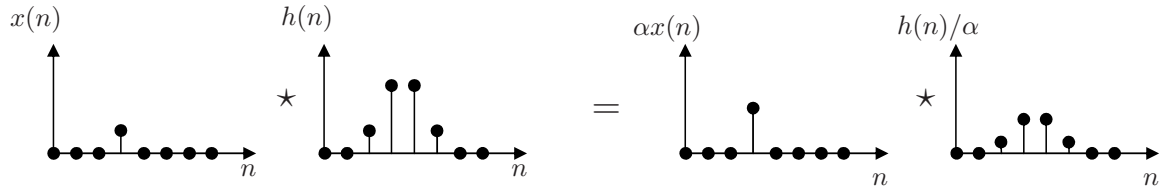


FIG. 3.4 – Exemple d'indétermination sur l'amplitude ($\alpha = 2$).

On peut même envisager que, lorsque α évolue autour de zéro, les signaux s'inversent au cours des itérations, la moyenne empirique de x et h tendant alors vers zéro, donnant une estimation au sens de l'EAP nulle ! En pratique heureusement, ce problème apparaît rarement. De plus, nous ne considérons ici que des pics positifs, impliquant $\alpha > 0$. Cela ne résout pas le problème de l'indétermination sur l'amplitude mais empêche les signaux de s'inverser et, ainsi, de donner une estimation nulle.

À notre connaissance, la totalité des études citant le problème d'indétermination sur l'amplitude lève le problème en normalisant l'énergie ou l'amplitude de la réponse impulsionnelle [20, 96]. Dès lors, nous choisissons de fixer le maximum de la réponse impulsionnelle à 1, et donc de laisser la variance de l'amplitude des pics $r_{\mathbf{a}}$ libre (l'hyperparamètre $r_{\mathbf{a}}$ doit alors être estimé).

Indétermination temporelle

L'indétermination temporelle correspond au fait que l'entrée et la réponse impulsionnelle peuvent être décalées dans le temps d'une valeur opposée. On ne peut donc pas différencier les cas :

$$\big(x(n), h(n)\big) \qquad \text{et} \qquad \big(x(n+\tau), h(n-\tau)\big).$$

L'indétermination temporelle peut être vue comme un cas particulier de l'indétermination sur le contenu spectral avec, dans le domaine fréquentiel :

$$H_2(f) = H(f)e^{-i2\pi f\tau} \qquad \text{et} \qquad H_1(f) = e^{+i2\pi f\tau}.$$

L'indétermination temporelle est également un cas particulier d'indétermination sur la phase où $\varphi(f) = 2\pi f\tau$. La figure 3.5 présente un cas d'indétermination temporelle.

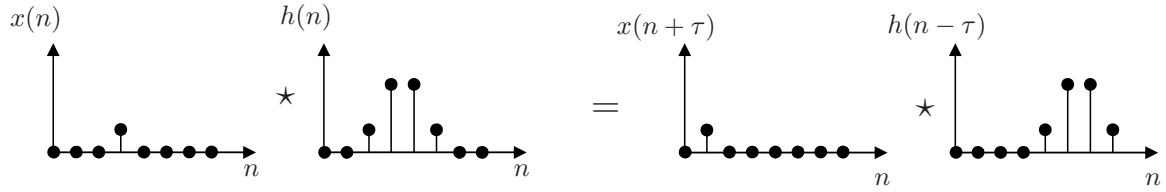


FIG. 3.5 – Exemple d'indétermination temporelle ($\tau = 2$).

Nous levons cette indétermination en fixant le maximum de la réponse impulsionnelle en zéro ; dans notre application, cela revient à dire que la réponse impulsionnelle $\mathbf{h}$ est centrée. Comme par ailleurs nous avons fixé le maximum de la réponse impulsionnelle à 1, nous avons donc : $\mathbf{h}_0 = 1$.

3.3 Modélisation probabiliste

3.3.1 Lois a priori

3.3.1.1 Loi a priori du signal impulsionnel x

Un signal impulsionnel est couramment modélisé par un processus Bernoulli-gaussien [16, 17, 20, 66] qui consiste à représenter $\mathbf{x}$ par le couple $(\mathbf{q}, \mathbf{a})$ où $\mathbf{q}$ fixe l'occurrence des pics et $\mathbf{a}$ leur amplitude. $\mathbf{q}_n$ suit une loi de Bernoulli de paramètre λ :

$$\mathbf{q}_n \sim \mathcal{B}er(\lambda), \tag{3.1}$$

et $\mathbf{a}_n$ suit une gaussienne de moyenne nulle et de variance $r_{\mathbf{a}}\mathbf{q}_n$. De plus, nous nous restreignons ici au cas où les impulsions sont positives : nous posons donc comme loi a priori sur $\mathbf{a}_n$ une loi normale à support positif (cf. annexe D) :

$$\mathbf{a}_n \sim \mathcal{N}^+(0, r_{\mathbf{a}}\mathbf{q}_n) \tag{3.2}$$

avec la convention que $\mathcal{N}^+(0,0)$ est une impulsion de Dirac. En d'autres termes, $\mathbf{a}_n$ est distribué suivant une gaussienne à support positif de variance $r_{\mathbf{a}}$ si $\mathbf{q}_n = 1$ et suivant une impulsion de Dirac centrée en zéro si $\mathbf{q}_n = 0$:

$$\mathbf{a}_n \sim \begin{cases} \mathcal{N}^+(0, r_{\mathbf{a}}) & \text{si } \mathbf{q}_n = 1, \\ \delta_0(\mathbf{a}_n) & \text{si } \mathbf{q}_n = 0. \end{cases}$$

On peut donc déterminer la loi a priori de $\mathbf{x}_n$ (loi a priori jointe de $(\mathbf{q}_n, \mathbf{a}_n)$) :

$$\begin{aligned} p(\mathbf{x}_n|\lambda, r_{\mathbf{a}}) &= p(\mathbf{a}_n, \mathbf{q}_n|\lambda, r_{\mathbf{a}}) \\ &= p(\mathbf{a}_n|\mathbf{q}_n, r_{\mathbf{a}})p(\mathbf{q}_n|\lambda), \end{aligned}$$

donc :

$$\mathbf{x}_n \sim \lambda\mathcal{N}^+(0, r_{\mathbf{a}}) + (1-\lambda)\delta_0(\mathbf{x}_n). \tag{3.3}$$

Remarquons que nous aurions pu poser une loi gamma sur les amplitudes $\mathbf{a}_n$, comme dans [47, 48]. Cependant, l'intérêt d'utiliser une loi normale à support positif plutôt qu'une loi gamma est multiple. Tout d'abord, la loi normale est un bon modèle pour coder le fait qu'il y a plus de pics de petite amplitude que de pics de grande amplitude ; la loi gamma suppose en outre que les pics ne peuvent

pas avoir une amplitude proche de zéro, ce qui ne reflète pas la réalité. Par ailleurs, une loi gamma résulterait en une loi a posteriori plus complexe, d'autant plus qu'il faudrait fixer deux paramètres au lieu d'un seul pour la loi normale à support positif. Enfin, le modèle considéré permet facilement de revenir à des signaux dont les pics sont positifs et négatifs, puisqu'il suffit de ne pas imposer la contrainte de positivité.

3.3.1.2 Loi a priori de λ

En pratique, lorsque le nombre réel de raies devient très faible, il apparaît que la solution qui consiste à avoir un grand nombre de pics de petite amplitude devient très probable : λ est alors très proche de 1. C'est typiquement un problème de dégénérescence. Il est donc nécessaire de fixer un a priori sur λ qui permette de limiter le nombre de pics. Nous choisissons d'utiliser une loi a priori conjuguée qui indique que le nombre de pics le plus probable est nul. Cela est bien évidemment faux, mais aura simplement pour conséquence de faire tendre l'estimation vers une valeur faible de λ. Dans ce cas, l'a priori est une loi beta :

$$\lambda \sim \mathcal{B}e(1, N+1). \tag{3.4}$$

La figure 3.6 illustre le fait que l'on souhaite faire tendre λ vers 0 (avec ici une faible valeur de la longueur du signal : $N = 8$). Lorsque N devient grand, cet effet s'accentue.

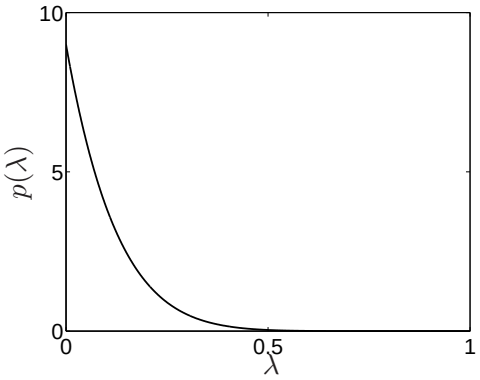


FIG. 3.6 – Loi a priori sur λ ($N = 8$).

3.3.1.3 Loi a priori de $r_{\mathbf{a}}$

Sans aucun a priori sur le paramètre $r_{\mathbf{a}}$, sa loi a posteriori s'écrit :

$$r_{\mathbf{a}} \sim \mathcal{IG}\left(\frac{N_{\mathbf{q}}}{2} - 1, \frac{\mathbf{x}_n^T \mathbf{x}_n}{2}\right),$$

où $N_{\mathbf{q}}$ correspond au nombre de pics non nuls :

$$N_{\mathbf{q}} = \sum_{n=1}^{N} \mathbf{q}_n. \tag{3.5}$$

Il apparaît alors que lorsque $N_{\mathbf{q}} \leq 2$ le premier paramètre de la loi inverse gamma n'est plus strictement positif. Pour éviter ce problème, nous posons une loi a priori conjuguée sur $r_{\mathbf{a}}$, donc une inverse gamma (cf. figure 3.7) :

$$r_{\mathbf{a}} \sim \mathcal{IG}(\alpha_{\mathbf{a}}, \beta_{\mathbf{a}}), \tag{3.6}$$

où $\alpha_{\mathbf{a}} > 0$ et $\beta_{\mathbf{a}} > 0$. L'espérance et la variance d'une telle loi sont (cf. annexe D) :

$$\mathbb{E}_{r_{\mathbf{a}}}[r_{\mathbf{a}}] = \frac{\beta_{\mathbf{a}}}{\alpha_{\mathbf{a}} - 1}, \qquad (\text{si } \alpha > 1),$$

$$\mathrm{Var}_{r_{\mathbf{a}}}[r_{\mathbf{a}}] = \frac{\beta_{\mathbf{a}}^2}{(\alpha_{\mathbf{a}} - 1)^2(\alpha_{\mathbf{a}} - 2)} \qquad (\text{si } \alpha > 2).$$

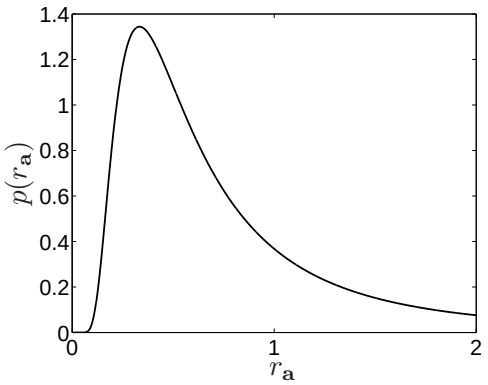


FIG. 3.7 – Loi a priori sur $r_{\mathbf{a}}$ ($\varepsilon = 10^{-6}$, $\mathbb{E}_{r_{\mathbf{a}}}[r_{\mathbf{a}}] = 1$)

L'a priori a été choisi pour éviter les erreurs numériques et non pour influer le résultat de l'estimation : il doit donc être le moins informatif possible. À ce titre, la variance doit être choisie la plus grande possible et l'espérance de manière à ne pas affecter l'estimation. On choisit donc de prendre l'espérance égale à une estimation grossière de la variance de l'amplitude des pics. Celle-ci n'étant pas connue a priori, il y a trois façon de la déterminer :

- elle peut être mesurée approximativement à partir du spectre ;
- on peut calculer par exemple $\mathbb{E}_{r_{\mathbf{a}}}[r_{\mathbf{a}}] = \mathbf{w}^T\mathbf{w}/N$ [6], mais ce choix consiste à calculer la variance des amplitudes de chaque point et pas seulement des pics ;
- le spectre est multiplié par un coefficient de façon à avoir un signal dont la variance approximative des pics est de 1.

C'est cette dernière méthode qui est utilisée ici. Il n'est pas nécessaire d'avoir une valeur exacte puisque l'a priori est peu informatif, donc peu influent sur l'estimation. En effet, des simulations ont montré que même si cela n'est pas fait, l'estimation des amplitudes des raies ne sera pas trop influencée par la loi a priori.

Dès lors, puisque $\mathrm{Var}_{r_{\mathbf{a}}}[r_{\mathbf{a}}] = \mathbb{E}_{r_{\mathbf{a}}}[r_{\mathbf{a}}]^2/(\alpha_{\mathbf{a}} - 2)$, il faut que $\alpha_{\mathbf{a}}$ soit le plus petit possible pour avoir la variance la plus grande possible. Nous choisissons donc $\alpha_{\mathbf{a}} = 2 + \varepsilon$, où $\varepsilon \in \mathbb{R}^+$ et $\varepsilon \ll 1$. D'après l'expression de l'espérance on en déduit : $\beta_{\mathbf{a}} = \mathbb{E}_{r_{\mathbf{a}}}[r_{\mathbf{a}}](\alpha_{\mathbf{a}} - 1) = 1 + \varepsilon$.

3.3.1.4 Loi a priori de la réponse impulsionnelle h

Cheng *et al.* [20] choisissent un a priori gaussien sur l'amplitude des composantes $\mathbf{h}_n$. Le vecteur $\mathbf{h}$ est ensuite estimé en bloc. Cependant, comme le problème de la déconvolution myope est délicat à cause du grand nombre de variables à estimer relativement au nombre de données, nous souhaitons coder le maximum d'information sur la réponse impulsionnelle.

Ainsi, une première amélioration de la méthode de Cheng *et al.* consiste par exemple à introduire un a priori de douceur sur la réponse impulsionnelle pour éviter les réponses impulsionnelles trop chahutées, ou encore d'estimer seulement la moitié de la réponse impulsionnelle si on sait qu'elle est symétrique (ou antisymétrique).

Or, dans certaines applications (comme en spectroscopie), on peut supposer que la structure de $\mathbf{h}$ est connue : ainsi, seuls les paramètres de la structure sont à estimer, limitant ainsi le nombre d'inconnues

du problème. Dès lors, en considérant une réponse impulsionnelle lorentzienne :

$$\forall\, n \in \mathbb{R}, \quad h(n) = \frac{s^2}{s^2 + n^2},$$

il suffit simplement d'estimer le paramètre s. L'avantage de cette modélisation est qu'il n'y a qu'un seul paramètre à estimer. De plus, il n'est pas nécessaire de fixer la longueur de la réponse impulsionnelle, contrairement à Cheng *et al.* qui doivent estimer autant de variables que le nombre de points de la réponse impulsionnelle. Par ailleurs, puisqu'ils utilisent une réponse impulsionnelle finie, il doivent la considérer suffisamment grande pour réduire les erreurs numériques. Aussi, imposer la structure de la réponse impulsionnelle connue est un a priori très fort, limitant ainsi le nombre de solutions possibles et donc améliorant l'estimation.

Aucune connaissance particulière n'est disponible sur s, si ce n'est qu'il doit être positif. En effet, la positivité est imposée pour éviter d'éventuelles aberrations lors du calcul de son estimation au sens de l'EAP car si s était alternativement positif et négatif, la réponse impulsionnelle n'en serait pas affectée, mais on aurait alors $\widehat{s} \to 0$!

Par ailleurs, nous souhaitons avoir des lois a posteriori conditionnelles intégrables, à défaut d'avoir une loi a posteriori conjointe intégrable (voir annexe A, page 130). Il s'avère qu'en choisissant une loi non-informative de Laplace cette intégrabilité est vérifiée dans la majorité des cas (voir plus loin), alors qu'avec une loi non-informative de Jeffreys, l'intégrabilité serait plus difficile à prouver.

Pour ces raisons, on pose comme a priori sur s :

$$s \sim \mathcal{U}_{\mathbb{R}^+}. \tag{3.7}$$

3.3.1.5 Loi a priori du bruit b

Le bruit $\mathbf{b}$ est supposé i.i.d., gaussien de moyenne nulle et de variance $r_{\mathbf{b}}\mathbf{I}$:

$$\mathbf{b} \sim \mathcal{N}(\mathbf{0}, r_{\mathbf{b}}\mathbf{I}). \tag{3.8}$$

Comme nous n'avons aucune connaissance de $r_{\mathbf{b}}$, nous choisissons de lui affecter une loi a priori de Jeffreys qui est calculée en tenant compte des propositions de Jeffreys [61, 64], c'est-à-dire en considérant les autres paramètres du problème fixes. On obtient alors :

$$r_{\mathbf{b}} \sim 1/r_{\mathbf{b}}.$$

Une telle loi permet de conserver l'intégrabilité de la loi a posteriori conditionnelle de $r_{\mathbf{b}}$.

3.3.2 Graphe acyclique orienté

Il est commode d'utiliser une représentation graphique des modèles hiérarchiques en les représentant par un *graphe acyclique orienté* (*directed acyclic graph*) [44, 68]. La figure 3.8 représente le graphe acyclique orienté du modèle proposé dans ce chapitre.

Chaque variable du modèle correspond à un nœud du graphe. Les flèches correspondent aux liens entre les variables : les flèches continues représentent les dépendances probabilistes alors que les flèches en pointillés représentent les relations déterministes. Les carrés représentent les quantités fixées ou observées, et les cercles représentent les quantités inconnues. Enfin, les structures répétitives, telle que $\mathbf{x}_n$, sont représentées sur des rectangles « empilés ». Pour rappel, on a également reporté les lois a priori de chacune des quantités inconnues (« $\mathcal{BG}$ » signifie distribution Bernoulli-gaussienne et « $\mathcal{J}$ » représente la loi non-informative de Jeffreys).

Ainsi, un graphe acyclique orienté permet, tout en représentant le modèle de manière graphique, d'écrire la loi a posteriori du modèle très facilement, même pour des modèles complexes [44, chapitre 2].

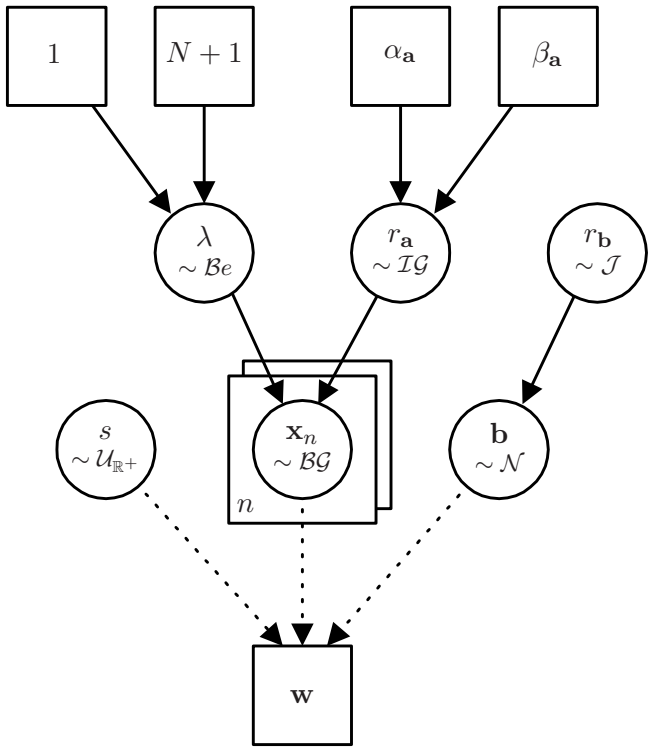


FIG. 3.8 – Graphe acyclique orienté du modèle de déconvolution impulsionnelle positive myope.

3.3.3 Lois a posteriori

Il y a donc $2N+4$ paramètres à estimer :

$$\boldsymbol{\theta} = \{\mathbf{x}, s, \lambda, r_{\mathbf{a}}, r_{\mathbf{b}}\} = \{\mathbf{q}, \mathbf{a}, s, \lambda, r_{\mathbf{a}}, r_{\mathbf{b}}\}.$$

La loi a priori sur $\mathbf{b}$ permet d'écrire la fonction de vraisemblance :

$$p(\mathbf{w}|\mathbf{x}, s, \lambda, r_{\mathbf{a}}, r_{\mathbf{b}}) = p(\mathbf{w}|\mathbf{x}, s, r_{\mathbf{b}}) = \frac{1}{(2\pi r_{\mathbf{b}})^{N/2}} \exp\left(-\frac{1}{2r_{\mathbf{b}}}||\mathbf{w} - \mathbf{H}\mathbf{x}||^2\right),$$

où $\mathbf{H}$ est la matrice de Toeplitz de la réponse impulsionnelle, fonction de s. La dépendance de $\mathbf{H}$ ou de $\mathbf{h}$ par rapport à s est implicite pour éviter de surcharger les notations.

La règle de Bayes permet d'obtenir l'expression de la loi a posteriori globale :

$$p(\mathbf{x}, s, \lambda, r_{\mathbf{a}}, r_{\mathbf{b}}|\mathbf{w}) \propto p(\mathbf{w}|\mathbf{x}, s, r_{\mathbf{b}}) p(\mathbf{a}|\mathbf{q}, r_{\mathbf{a}}) p(\mathbf{q}|\lambda) p(s) p(\lambda) p(r_{\mathbf{a}}) p(r_{\mathbf{b}}),$$

d'où :

$$\begin{aligned} p(\mathbf{x}, s, \lambda, r_{\mathbf{a}}, r_{\mathbf{b}}|\mathbf{w}) \propto & \frac{1}{(2\pi r_{\mathbf{b}})^{N/2}} \exp\left(-\frac{1}{2r_{\mathbf{b}}}||\mathbf{w} - \mathbf{H}\mathbf{x}||^2\right) \\ & \times \prod_{n=1}^{N} \left[\lambda\sqrt{\frac{2}{\pi r_{\mathbf{a}}}} \exp\left(-\frac{\mathbf{x}_n^2}{2r_{\mathbf{a}}}\right) \mathbb{1}_{\mathbb{R}^+}(\mathbf{x}_n) + (1-\lambda)\delta_0(\mathbf{x}_n)\right] \\ & \times \frac{\lambda^0(1-\lambda)^N}{B(1, N+1)} \mathbb{1}_{[0,1]}(\lambda) \times \frac{\beta_{\mathbf{a}}^{\alpha_{\mathbf{a}}}}{\Gamma(\alpha_{\mathbf{a}})} \frac{\mathrm{e}^{-\beta_{\mathbf{a}}/r_{\mathbf{a}}}}{r_{\mathbf{a}}^{\alpha_{\mathbf{a}}+1}} \mathbb{1}_{\mathbb{R}^+}(r_{\mathbf{a}}) \times \mathbb{1}_{\mathbb{R}^+}(s) \times \frac{1}{r_{\mathbf{b}}}. \end{aligned}$$

L'intégrabilité de la loi a posteriori n'a pas été vérifiée car son expression mathématique est trop complexe. Comme il est précisé dans l'annexe A (page 130), pour être sûr que l'intégrale de cette

fonction soit finie, il faudrait choisir des lois a priori propres puisque la fonction de vraisemblance est bornée. Or, nous avons choisi de poser des lois de Laplace et de Jeffreys sur s et $r_{\mathbf{b}}$ puisque poser des lois a priori non-informatives constitue le choix traditionnel dans le cas d'un modèle bayésien hiérarchique. Malheureusement, ces deux lois a priori sont impropres. Nous ne nous sommes pas attardé plus en détail sur ce point car, dans la configuration actuelle, les chaînes de Markov de chaque variable ont un comportement normal qui n'indique aucune dégénérescence : en effet, nous avons pris soin d'obtenir des lois a posteriori conditionnelles intégrables.

3.3.3.1 Loi a posteriori du signal impulsionnel x

Nous travaillons sur chaque élément de $\mathbf{x}$ séparément car les calculs et l'implémentation sont plus faciles ; en effet, cela évite d'avoir à travailler avec des matrices de dimension $N \times N$.

La loi a posteriori de $\mathbf{x}_n$ s'écrit :

$$\begin{aligned} p(\mathbf{x}_n|\mathbf{w}, \mathbf{x}_{-n}, s, \lambda, r_{\mathbf{a}}, r_{\mathbf{b}}) &\propto p(\mathbf{w}|\mathbf{x}_n, \mathbf{x}_{-n}, s, \lambda, r_{\mathbf{a}}, r_{\mathbf{b}}) p(\mathbf{x}_n|\mathbf{x}_{-n}, s, \lambda, r_{\mathbf{a}}, r_{\mathbf{b}}), \\ &\propto p(\mathbf{w}|\mathbf{x}, s, r_{\mathbf{b}}) p(\mathbf{x}_n|\lambda, r_{\mathbf{a}}), \\ &\propto \frac{1}{(2\pi r_{\mathbf{b}})^{N/2}} \exp\left(-\frac{1}{2r_{\mathbf{b}}}||\mathbf{w} - \mathbf{H}\mathbf{x}||^2\right) \\ &\quad \times \left[\lambda\sqrt{\frac{2}{\pi r_{\mathbf{a}}}} \exp\left(-\frac{\mathbf{x}_n^2}{2r_{\mathbf{a}}}\right) \mathbb{1}_{\mathbb{R}^+}(\mathbf{x}_n) + (1-\lambda)\delta_0(\mathbf{x}_n)\right]. \end{aligned}$$

On montre (cf. annexe B) que :

$$\mathbf{x}_n \sim \lambda_{0,n}\mathcal{N}^+(\mu_{0,n}, \rho_{0,n}) + \lambda_{1,n}\mathcal{N}^+(\mu_{1,n}, \rho_{1,n}), \tag{3.9}$$

ce qui revient à :

$$\mathbf{q}_n \sim \mathcal{B}er(\lambda_{1,n}), \tag{3.10}$$

$$\mathbf{a}_n \sim \begin{cases} \mathcal{N}^+(\mu_{0,n}, \rho_{0,n}) & \text{si } \mathbf{q}_n = 0, \\ \mathcal{N}^+(\mu_{1,n}, \rho_{1,n}) & \text{si } \mathbf{q}_n = 1. \end{cases} \tag{3.11}$$

avec

$$\lambda_{1,n} = \left[1 + \frac{1-\lambda}{\lambda}\sqrt{\frac{r_{\mathbf{a}}}{\rho_{1,n}}} \exp\left(-\frac{\mu_{1,n}^2}{2\rho_{1,n}}\right)\right]^{-1},$$

$$\lambda_{0,n} = \left[1 + \frac{\lambda}{1-\lambda}\sqrt{\frac{\rho_{1,n}}{r_{\mathbf{a}}}} \exp\left(\frac{\mu_{1,n}^2}{2\rho_{1,n}}\right)\right]^{-1},$$

$$\mu_{1,n} = \frac{\rho_{1,n}}{r_{\mathbf{b}}}\mathbf{e}_n^T\mathbf{H}_n, \qquad \mu_{0,n} = 0, \qquad \rho_{1,n} = \frac{r_{\mathbf{a}} r_{\mathbf{b}}}{r_{\mathbf{b}} + r_{\mathbf{a}}\mathbf{H}_n^T\mathbf{H}_n}, \qquad \rho_{0,n} = 0,$$

$$\mathbf{e}_n = \mathbf{w} - (\mathbf{H}\mathbf{x} - \mathbf{H}_n\mathbf{x}_n).$$

Remarquons que lorsqu'une partie du signal est négative et d'amplitude relativement importante, on a $\mu_{1,n} < 0$ et par conséquent $\lambda_{1,n}$ proche de 1. Dès lors, la probabilité pour que $\mathbf{q}_n = 1$ est grande : un pic est donc estimé, même s'il est d'amplitude quasi-nulle. Pour éviter cela, nous forçons $\mu_{1,n}$ à zéro s'il est négatif.

3.3.3.2 Loi a posteriori du paramètre de Bernoulli λ

La loi a posteriori de λ s'écrit :

$$\begin{aligned} p(\lambda|\mathbf{w}, \mathbf{x}, s, r_{\mathbf{a}}, r_{\mathbf{b}}) &\propto p(\mathbf{w}|\mathbf{x}, s, \lambda, r_{\mathbf{a}}, r_{\mathbf{b}}) p(\mathbf{x}, s, r_{\mathbf{a}}, r_{\mathbf{b}}|\lambda) p(\lambda), \\ &\propto p(\mathbf{w}|\mathbf{x}, s, r_{\mathbf{b}}) p(\mathbf{x}|\lambda) p(\lambda), \\ &\propto \prod_{n=1}^{N} p(\mathbf{x}_n|\lambda) p(\lambda). \end{aligned}$$

D'où :

$$
\begin{aligned}
p(\lambda|\mathbf{w},\mathbf{x},s,r_{\mathbf{a}},r_{\mathbf{b}}) &= \prod_{n=1}^{N}\left[\lambda\sqrt{\frac{2}{\pi r_{\mathbf{a}}}}\exp\left(-\frac{\mathbf{x}_n^2}{2r_{\mathbf{a}}}\right)\mathbb{1}_{\mathbb{R}^+}(\mathbf{x}_n)+(1-\lambda)\delta_0(\mathbf{x}_n)\right]\times\frac{\lambda^0(1-\lambda)^N}{B(1,N+1)}\mathbb{1}_{[0,1]}(\lambda),\\
&\propto \prod_{\substack{n=1\\ \mathbf{q}_n=1}}^{N}\left[\lambda\sqrt{\frac{2}{\pi r_{\mathbf{a}}}}\exp\left(-\frac{\mathbf{x}_n^2}{2r_{\mathbf{a}}}\right)\mathbb{1}_{\mathbb{R}^+}(\mathbf{x}_n)\right]\prod_{\substack{n=1\\ \mathbf{q}_n=0}}^{N}\left[(1-\lambda)\delta_0(\mathbf{x}_n)\right]\times\frac{\lambda^0(1-\lambda)^N}{B(1,N+1)}\mathbb{1}_{[0,1]}(\lambda),\\
&\propto \lambda^{N_{\mathbf{q}}}(1-\lambda)^{N-N_{\mathbf{q}}}\frac{\lambda^0(1-\lambda)^N}{B(1,N+1)}\mathbb{1}_{[0,1]}(\lambda).
\end{aligned}
$$

Par conséquent, la loi a posteriori de λ est une loi beta :

$$
\lambda\sim\mathcal{B}e\big(N_{\mathbf{q}}+1,2N-N_{\mathbf{q}}+1\big). \tag{3.12}
$$

3.3.3.3 Loi a posteriori de la variance des amplitudes $r_{\mathbf{a}}$

La loi a posteriori de $r_{\mathbf{a}}$ s'écrit :

$$
\begin{aligned}
p(r_{\mathbf{a}}|\mathbf{w},\mathbf{x},s,\lambda,r_{\mathbf{b}}) &\propto p(\mathbf{w}|\mathbf{x},s,\lambda,r_{\mathbf{b}},r_{\mathbf{a}})p(\mathbf{x},s,\lambda,r_{\mathbf{b}}|r_{\mathbf{a}})p(r_{\mathbf{a}}),\\
&\propto p(\mathbf{w}|\mathbf{x},s,r_{\mathbf{b}})p(\mathbf{x}|r_{\mathbf{a}})p(r_{\mathbf{a}}),\\
&\propto \prod_{n=1}^{N}p(\mathbf{x}_n|r_{\mathbf{a}})p(r_{\mathbf{a}}).
\end{aligned}
$$

D'où :

$$
\begin{aligned}
p(r_{\mathbf{a}}|\mathbf{w},\mathbf{x},s,\lambda,r_{\mathbf{b}}) &\propto \prod_{n=1}^{N}\left[\lambda\sqrt{\frac{2}{\pi r_{\mathbf{a}}}}\exp\left(-\frac{\mathbf{x}_n^2}{2r_{\mathbf{a}}}\right)\mathbb{1}_{\mathbb{R}^+}(\mathbf{x}_n)+(1-\lambda)\delta_0(\mathbf{x}_n)\right]\times\frac{\beta_{\mathbf{a}}^{\alpha_{\mathbf{a}}}}{\Gamma(\alpha_{\mathbf{a}})}\frac{\mathrm{e}^{-\beta_{\mathbf{a}}/r_{\mathbf{a}}}}{r_{\mathbf{a}}^{\alpha_{\mathbf{a}}+1}}\mathbb{1}_{\mathbb{R}^+}(r_{\mathbf{a}}),\\
&\propto \prod_{\substack{n=1\\ \mathbf{q}_n=1}}^{N}\left[\lambda\sqrt{\frac{2}{\pi r_{\mathbf{a}}}}\exp\left(-\frac{\mathbf{x}_n^2}{2r_{\mathbf{a}}}\right)\mathbb{1}_{\mathbb{R}^+}(\mathbf{x}_n)\right]\times\frac{\beta_{\mathbf{a}}^{\alpha_{\mathbf{a}}}}{\Gamma(\alpha_{\mathbf{a}})}\frac{\mathrm{e}^{-\beta_{\mathbf{a}}/r_{\mathbf{a}}}}{r_{\mathbf{a}}^{\alpha_{\mathbf{a}}+1}}\mathbb{1}_{\mathbb{R}^+}(r_{\mathbf{a}}),\\
&\propto \left(\lambda\sqrt{\frac{2}{\pi r_{\mathbf{a}}}}\right)^{N_{\mathbf{q}}}\exp\left(-\frac{1}{2r_{\mathbf{a}}}\sum_{\substack{n=1\\ \mathbf{q}_n=1}}^{N}\mathbf{x}_n^2\right)\times\frac{\beta_{\mathbf{a}}^{\alpha_{\mathbf{a}}}}{\Gamma(\alpha_{\mathbf{a}})}\frac{\mathrm{e}^{-\beta_{\mathbf{a}}/r_{\mathbf{a}}}}{r_{\mathbf{a}}^{\alpha_{\mathbf{a}}+1}}\mathbb{1}_{\mathbb{R}^+}(r_{\mathbf{a}}),\\
&\propto \frac{1}{r_{\mathbf{a}}^{N_{\mathbf{q}}/2+\alpha_{\mathbf{a}}+1}}\exp\left(-\frac{1}{2r_{\mathbf{a}}}\sum_{\substack{n=1\\ \mathbf{q}_n=1}}^{N}\mathbf{x}_n^2-\frac{\beta_{\mathbf{a}}}{r_{\mathbf{a}}}\right)\mathbb{1}_{\mathbb{R}^+}(r_{\mathbf{a}}).
\end{aligned}
$$

En remarquant que $\mathbf{x}_n=0$ si $\mathbf{q}_n=0$, on peut simplifier la somme :

$$
\sum_{\substack{n=1\\ \mathbf{q}_n=1}}^{N}\mathbf{x}_n^2=\sum_{n=1}^{N}\mathbf{x}_n^2=\mathbf{x}_n^T\mathbf{x}_n
$$

Donc $r_{\mathbf{a}}$ est distribué suivant une inverse gamma :

$$
r_{\mathbf{a}}\sim\mathcal{IG}\left(\frac{N_{\mathbf{q}}}{2}+\alpha_{\mathbf{a}},\frac{\mathbf{x}_n^T\mathbf{x}_n}{2}+\beta_{\mathbf{a}}\right), \tag{3.13}
$$

avec, on le rappelle (section 3.3.1.3) :

$$
\alpha_{\mathbf{a}}=2+\varepsilon,\qquad \beta_{\mathbf{a}}=1+\varepsilon,\qquad \varepsilon\ll 1.
$$

3.3.3.4 Loi a posteriori de la largeur s

La loi a posteriori de s s'écrit :

$$\begin{aligned} p(s|\mathbf{w},\mathbf{x},\lambda,r_{\mathbf{a}},r_{\mathbf{b}}) &\propto p(\mathbf{w}|\mathbf{x},s,\lambda,r_{\mathbf{a}},r_{\mathbf{b}})p(s), \\ &\propto p(\mathbf{w}|\mathbf{x},s,r_{\mathbf{b}})p(s), \end{aligned}$$

d'où :

$$s \sim \exp\left(-\frac{1}{2r_{\mathbf{b}}}||\mathbf{w}-\mathbf{H}\mathbf{x}||^2\right)\mathbb{1}_{\mathbb{R}^+}(s). \tag{3.14}$$

où s intervient dans la matrice de Toeplitz $\mathbf{H}$.

On peut montrer que, en pratique, la loi a posteriori conditionnelle de s (et, par extension, des paramètres de forme de la raie) est bien intégrable quelle que soit la forme de la raie. En effet, $p(s|\mathbf{w},\mathbf{x},\lambda,r_{\mathbf{a}},r_{\mathbf{b}})$ peut être majorée par la fonction $\exp(-\alpha s)$ ($\alpha \in \mathbb{R}^+$) qui est intégrable car :

$$\begin{aligned} \exists\alpha\in\mathbb{R}^+, \forall s\in\mathbb{R}^+, \qquad & \exp\left(-\frac{1}{2r_{\mathbf{b}}}||\mathbf{w}-\mathbf{H}\mathbf{x}||^2\right)\mathbb{1}_{\mathbb{R}^+}(s) < \exp(-\alpha s), \\ \Leftrightarrow \quad & -\frac{1}{2r_{\mathbf{b}}}||\mathbf{w}-\mathbf{H}\mathbf{x}||^2 < -\alpha s, \\ \Leftrightarrow \quad & ||\mathbf{w}-\mathbf{H}\mathbf{x}||^2 > 2r_{\mathbf{b}}\alpha s \end{aligned}$$

Or, $||\mathbf{w}-\mathbf{H}\mathbf{x}||^2$ est toujours strictement positif quelle que soit la valeur du paramètre de forme. En effet, ce terme s'annule seulement lorsque $\mathbf{w}=\mathbf{H}\mathbf{x}$, ce qui est n'est pas possible en pratique lorsque le signal $\mathbf{w}$ est bruité. En particulier, l'inégalité est évidente puisque le bruit crée des échantillons négatifs et que le spectre pur est forcément positif. Par conséquent, il est possible de trouver une valeur de α pour laquelle $p(s|\mathbf{w},\mathbf{x},\lambda,r_{\mathbf{a}},r_{\mathbf{b}})$ est minimisée par $\exp(-\alpha s)$, prouvant alors l'intégrabilité de la loi a posteriori conditionnelle.

3.3.3.5 Loi a posteriori de la variance du bruit $r_{\mathbf{b}}$

La loi a posteriori de $r_{\mathbf{b}}$ s'écrit :

$$\begin{aligned} p(r_{\mathbf{b}}|\mathbf{w},\mathbf{x},s,\lambda,r_{\mathbf{a}}) &\propto p(\mathbf{w}|\mathbf{x},s,\lambda,r_{\mathbf{a}},r_{\mathbf{b}})p(r_{\mathbf{b}}), \\ &\propto p(\mathbf{w}|\mathbf{x},s,r_{\mathbf{b}})p(r_{\mathbf{b}}), \end{aligned}$$

d'où :

$$p(r_{\mathbf{b}}|\mathbf{w},\mathbf{x},s,\lambda,r_{\mathbf{a}}) \propto \frac{1}{(2\pi r_{\mathbf{b}})^{N/2}}\exp\left(-\frac{1}{2r_{\mathbf{b}}}||\mathbf{w}-\mathbf{H}\mathbf{x}||^2\right)\frac{1}{r_{\mathbf{b}}}.$$

Ainsi, la loi a priori de $r_{\mathbf{b}}$ est une inverse gamma :

$$r_{\mathbf{b}} \sim \mathcal{IG}\left(\frac{N}{2}, \frac{||\mathbf{w}-\mathbf{H}\mathbf{x}||^2}{2}\right). \tag{3.15}$$

3.4 Implémentation de l'échantillonneur de Gibbs

Le modèle hiérarchique permet d'utiliser naturellement l'échantillonneur de Gibbs. Nous utilisons la version à balayage aléatoire (cf. section A.3.2), c'est-à-dire que les variables $\mathbf{x}_n$, s, λ, $r_{\mathbf{a}}$ et $r_{\mathbf{b}}$ sont échantillonnées dans un ordre aléatoire à chaque itération.

Il existe peu d'études sur le choix des valeurs initiales d'une méthode MCMC [44, p. 13 et p. 431]. Elles peuvent être fixées par l'utilisateur ou générées à partir de leur loi a posteriori. Cependant, les paramètres des lois a posteriori ne sont pas connus puisqu'ils sont eux aussi des variables du modèle.

C'est pourquoi nous laissons à l'utilisateur le choix des valeurs initiales. De toute manière, pour une chaîne de Markov irréductible, l'échantillonneur de Gibbs est assuré de converger vers la distribution d'équilibre quelles que soient les valeurs initiales (mais le temps de convergence peut être réduit en initialisant la chaîne à une valeur raisonnable). Heureusement, il est en général assez facile de déterminer de manière grossière la variance du bruit et des pics, le nombre de pics, et enfin les paramètres de la réponse impulsionnelle : en effet, des mesures approximatives sur le signal de mesure permettent de donner un ordre de grandeur aux valeurs initiales assez proche des valeurs réelles des paramètres, accélérant ainsi la convergence de l'algorithme.

Échantillonnage de $\mathbf{x}_n$

L'échantillonnage de $\mathbf{x}_n$ est assez simple puisqu'il suffit d'échantillonner les variables $\mathbf{q}_n$ puis $\mathbf{a}_n$ dont les lois a posteriori sont facilement échantillonnables :

1. calcul des paramètres $\lambda_{1,n}$, $\mu_{1,n}$ et $\rho_{1,n}$;
2. échantillonner $\mathbf{q}_n \sim \mathcal{B}er(\lambda_{1,n})$;
3. si $\mathbf{q}_n = 1$, échantillonner $\mathbf{a}_n \sim \mathcal{N}^+(\mu_{1,n}, \rho_{1,n})$, sinon $\mathbf{a}_n = 0$.

La génération de variables aléatoires distribuées suivant une gaussienne à support positif a fait l'objet de plusieurs travaux présentés dans l'annexe C. On y propose également un algorithme d'acceptation-rejet mixte utilisant plusieurs lois candidates et dont le taux d'acceptation moyen est très élevé, donnant ainsi une méthode d'échantillonnage rapide.

La génération de $\mathbf{q}_n$ se fait au moyen de méthodes classiques [93].

Échantillonnage de s

Puisque la loi a posteriori de s n'est pas directement échantillonnable, nous proposons d'utiliser un algorithme de Metropolis-Hastings pour en générer des échantillons. Plus précisément, nous utilisons un algorithme de Metropolis-Hastings à marche aléatoire car la forme de la loi n'est pas connue, et donc nous ne pouvons pas proposer une loi candidate satisfaisante (cf. annexe A).

En général, la loi candidate d'un algorithme de Metropolis-Hastings à marche aléatoire est uniforme, gaussienne ou de Student [93, 103]. Nous choisissons une loi candidate gaussienne car elle est à support infini et simulable facilement. Elle est centrée sur $s^{(i-1)}$, valeur du paramètre à l'itération précédente, et de variance r_s choisie au préalable. Comme la largeur s d'une lorentzienne est positive, nous choisissons une loi candidate normale définie sur $\mathbb{R}^+$: $\mathcal{N}^+(s^{(i-1)}, r_s)$. Finalement, comme la loi candidate est symétrique[3], l'algorithme obtenu est un algorithme de Metropolis (cf. annexe A).

L'efficacité de l'algorithme de Metropolis-Hastings à marche aléatoire dépend beaucoup du paramètre d'échelle r_s de la loi candidate. Si la variance est trop faible, la chaîne de Markov risque de converger trop lentement ; au contraire, si la variance est trop grande, l'algorithme va rejeter un grand nombre de propositions. Gelman *et al.* [39] recommandent par exemple de chercher à atteindre un taux d'acceptation de 0,5 pour les modèles de dimension 1. En général, le réglage de la variance est heuristique. Dès lors, nous proposons simplement de fixer dès le départ la valeur de r_s pour que le taux d'acceptation de l'algorithme de Metropolis-Hastings à marche aléatoire soit d'environ 0,5.

Le calcul du taux d'acceptation de l'algorithme de Metropolis peut conduire à des erreurs numériques. Pour éviter cela, nous en proposons maintenant une implémentation particulière. Le taux d'acceptation de l'algorithme de Metropolis s'écrit :

$$\alpha = \min\left\{1, \frac{p(\widetilde{s})}{p(s^{(i-1)})}\right\}$$

[3] Pour s'en convaincre, il faut remarquer que $\mathbb{1}_{\mathbb{R}^+}(\widetilde{s}) = \mathbb{1}_{\mathbb{R}^+}(s^{(i-1)})$ (où $\widetilde{s}$ est le candidat) puisque, par construction, $\widetilde{s}$ et $s^{(i-1)}$ sont positifs.

où $\widetilde{s}$ est le candidat. D'un point de vue algorithmique, il est inutile de calculer le minimum entre 1 et le rapport des deux probabilités puisqu'accepter $\widetilde{s}$ avec la probabilité α est implémenté en échantillonnant une variable aléatoire uniforme $u \sim \mathcal{U}_{[0,1]}$ et en acceptant $\widetilde{s}$ si et seulement si $u < \alpha$, ce qui revient simplement à calculer le rapport :

$$u < \frac{p(\widetilde{s})}{p(s^{(i-1)})}.$$

Ainsi, au niveau de l'algorithme, $\widetilde{s}$ est accepté si et seulement si

$$\begin{aligned}
&\Leftrightarrow && u < \frac{\exp\left(-||\mathbf{w} - \widetilde{\mathbf{H}}\mathbf{x}||^2/2r_{\mathbf{b}}\right)}{\exp\left(-||\mathbf{w} - \mathbf{H}^{(i-1)}\mathbf{x}||^2/2r_{\mathbf{b}}\right)},\\
&\Leftrightarrow && u < \exp\left(-\frac{1}{2r_{\mathbf{b}}}\left(||\mathbf{w} - \widetilde{\mathbf{H}}\mathbf{x}||^2 - ||\mathbf{w} - \mathbf{H}^{(i-1)}\mathbf{x}||^2\right)\right),\\
&\Leftrightarrow && -2r_{\mathbf{b}}\ln(u) > ||\mathbf{w} - \widetilde{\mathbf{H}}\mathbf{x}||^2 - ||\mathbf{w} - \mathbf{H}^{(i-1)}\mathbf{x}||^2,
\end{aligned}$$

où $\widetilde{\mathbf{H}}$ correspond à la matrice $\mathbf{H}$ calculée avec $\widetilde{s}$ et $\mathbf{H}^{(i-1)}$ est la matrice $\mathbf{H}$ à l'itération précédente. L'intérêt d'une telle implémentation consiste simplement à éviter les erreurs de débordement de précision lors du calcul du rapport entre les deux exponentielles.

Échantillonnage des hyperparamètres λ, $r_{\mathbf{a}}$ et $r_{\mathbf{b}}$

Enfin, pour l'échantillonnage des hyperparamètres λ, $r_{\mathbf{a}}$ et $r_{\mathbf{b}}$, il existe des méthodes classiques telles que celles données dans [93] et présentées dans l'annexe D. Pour l'implémentation, nous utilisons la boîte à outils Matlab Stixbox [56].

3.5 Choix des estimateurs

L'estimateur au sens de l'espérance a posteriori (EAP) est un estimateur naturel à implémenter à partir d'une chaîne de Markov. En effet, il suffit d'approcher cet estimateur par une moyenne de la chaîne de Markov obtenue :

$$\widehat{\boldsymbol{\theta}}^{\text{EAP}} = \mathbb{E}_{\boldsymbol{\theta}}[\boldsymbol{\theta}] \approx \frac{1}{I - I_0 + 1}\sum_{i=I_0}^{I}\boldsymbol{\theta}^{(i)}$$

(I_0 étant la longueur de la période de transition). C'est pourquoi l'estimateur EAP est utilisé afin d'obtenir une estimation de la réponse impulsionnelle : il suffit de calculer la moyenne de $s^{(i)}$. L'estimateur EAP est également utilisé si l'on souhaite obtenir une estimation des hyperparamètres λ, $r_{\mathbf{a}}$ et $r_{\mathbf{b}}$.

Par contre, l'estimation du signal impulsionnel n'est pas aussi évidente. Dans cette section, nous nous concentrons sur le choix de cet estimateur. Dans un premier temps, nous présentons les deux estimateurs proposés par Cheng *et al.* [20] et Rosec *et al.* [97] qui utilisent une modélisation et une résolution du problème proche de celle utilisée dans ce chapitre. Nous appliquons ces estimateurs à la chaîne de Markov générée et montrons qu'ils peuvent, dans certains cas, conduire à une interprétation erronée. Nous proposons dans un deuxième temps un estimateur original permettant d'obtenir une meilleure interprétation des résultats.

Dans la suite, on suppose donc que l'échantillonneur de Gibbs a généré les chaînes de Markov de $\mathbf{x}$ et $\mathbf{q}$. Pour simplifier, nous supposons que la période de transition de la chaîne de Markov est négligeable ou a été supprimée : la chaîne est considérée stationnaire de $i = 1$ à $i = I$.

Considérons l'exemple académique d'une chaîne de Markov telle que $I = 4$ (figure 3.9) et où des échantillons placés n'importe où sur $\{1, \ldots, N\}$ sont répertoriés par les indices $n_1, \ldots, n_6$. En n_1 et n_4, les pics (d'amplitudes respectives 1 et 1/2) sont toujours générés ; en n_2, un pic d'amplitude 1 n'est généré qu'à la première itération ; en n_3 aucun pic n'est généré ; et enfin un pic d'amplitude 1 est généré

alternativement en n_5 et n_6 (qui sont deux échantillons très proches : ils sont consécutifs ou séparés par un pixel), signifiant qu'il y a probablement un pic d'amplitude 1 situé entre n_5 et n_6, mais qu'il est difficile de le positionner avec précision. Les résultats fournis par les estimateurs de Cheng *et al.* et Rosec *et al.*, ainsi que le nôtre, sont également représentés sur la figure et discutés ci-après.

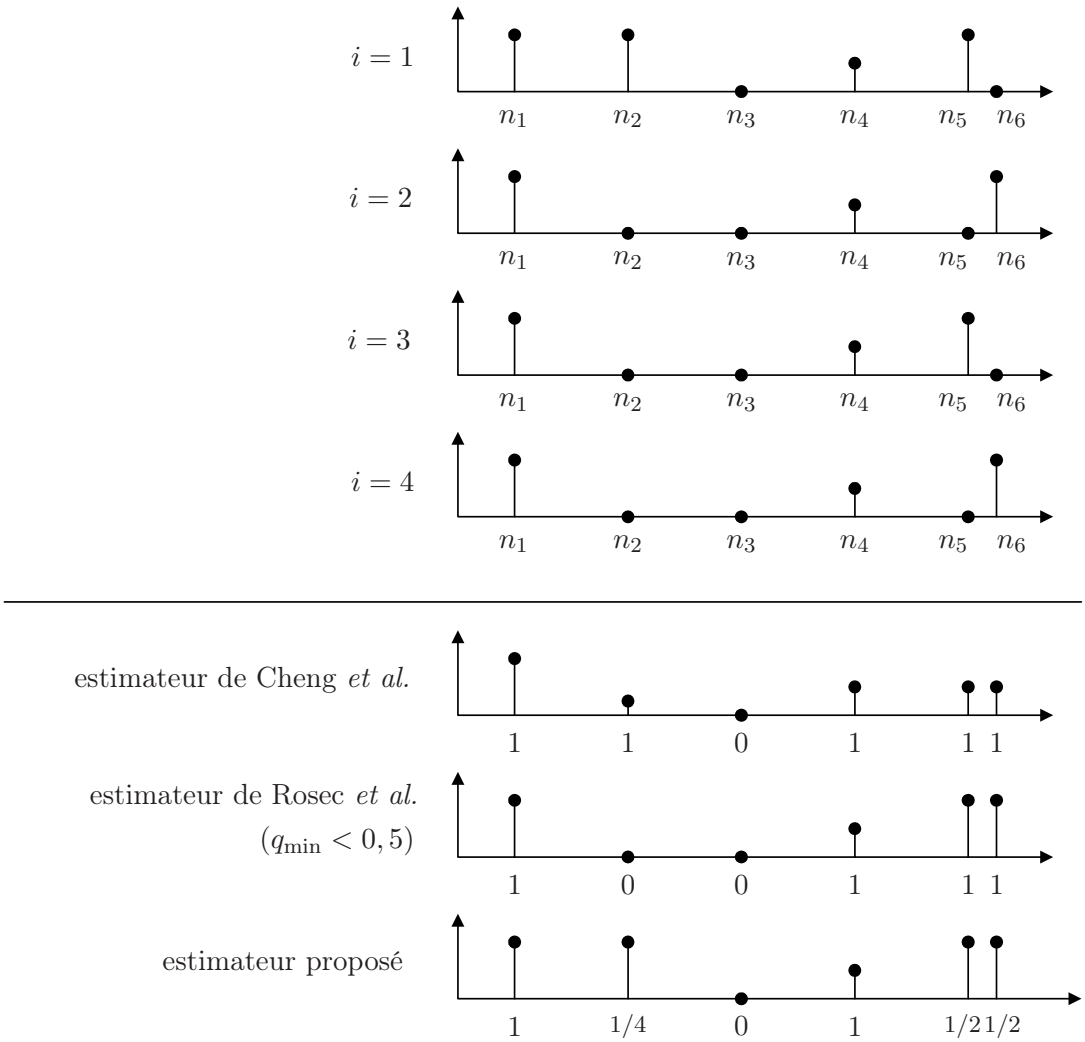


FIG. 3.9 – Comparaison des différents estimateurs d'un processus Bernoulli-gaussien.

3.5.1 Estimateur de Cheng *et al.*

Dans [20], les auteurs ne considèrent pas le vecteur $\mathbf{q}$ des occurrences des pics et proposent d'estimer $\mathbf{x}_n$ au sens de l'EAP. Cette quantité n'étant pas calculable explicitement, l'estimateur est approché par :

$$\widehat{\mathbf{x}}_n \triangleq \frac{1}{I} \sum_{i=1}^{I} \mathbf{x}_n^{(i)}.$$

Le résultat fourni par cet estimateur sur l'exemple de la figure 3.9 donne un résultat cohérent en n_1, n_3 et n_4 seulement. En revanche, en n_2, cet estimateur considère qu'il y a un pic d'amplitude $1/4$, ce qui n'est pas le cas : il n'y a sans doute aucun pic à cet endroit, le pic généré à la première itération de la chaîne de Markov pouvant être considéré comme un artefact. De plus, l'analyse de l'estimation indique qu'il y a deux pics d'amplitude $1/2$ en n_5 et n_6, alors qu'il n'y en a qu'un seul. Cependant, ces deux pics étant très proches, on peut supposer qu'il n'y en a en fait qu'un d'amplitude double, c'est-à-dire d'amplitude 1.

3.5.2 Estimateur de Rosec *et al.*

Rosec *et al.* [97], pour leur part, prennent en compte le vecteur $\mathbf{q}$ et l'estiment en testant la valeur de la somme des échantillons de la chaîne de Markov :

$$\widehat{\mathbf{q}}_n \triangleq \begin{cases} 1 & \text{si } \frac{1}{I}\sum_{i=1}^{I} \mathbf{q}_n^{(i)} > q_{\min}, \\ 0 & \text{sinon,} \end{cases}$$

où $q_{\min}$ est un seuil fixé a priori. En particulier, le choix $q_{\min} = 0{,}5$ n'estime à 1 que les pics qui apparaissent pendant plus de la moitié des itérations de la chaîne de Markov. Cela revient donc à estimer $\mathbf{q}$ au sens du MAP (maximum a posteriori). Si par contre $q_{\min} = 0$, alors l'estimateur de Rosec correspond à l'estimateur de Cheng.

L'estimateur de l'amplitude est ensuite défini par :

$$\widehat{\mathbf{x}}_n \triangleq \begin{cases} \dfrac{1}{\sum_{i=1}^{I} \mathbf{q}_n^{(i)}} \displaystyle\sum_{\substack{i=1 \\ \mathbf{q}_n^{(i)}=1}}^{I} \mathbf{x}_n^{(i)} & \text{si } \widehat{\mathbf{q}}_n = 1, \\ \dfrac{1}{\sum_{i=1}^{I} (1-\mathbf{q}_n^{(i)})} \displaystyle\sum_{\substack{i=1 \\ \mathbf{q}_n^{(i)}=0}}^{I} \mathbf{x}_n^{(i)} & \text{si } \widehat{\mathbf{q}}_n = 0, \end{cases}$$

En d'autre termes, si le nombre de pics générés au point n est trop faible ($\widehat{\mathbf{q}}_n = 0$), cet estimateur considère qu'il n'y a aucun pic. Si au contraire il y a suffisamment de pics générés ($\widehat{\mathbf{q}}_n = 1$), alors l'estimation de l'amplitude correspond à la moyenne des amplitudes des pics générés. La difficulté vient alors du choix de la valeur de $q_{\min}$.

Cet estimateur, qui constitue l'étape de décision de l'algorithme MPM, est également utilisé par Bourguignon *et al.* [6].

L'estimateur appliqué à la chaîne de Markov de la figure 3.9 donne un résultat correct pour les pics en n_1, n_3 et n_4, ainsi qu'en n_2 où l'artefact est correctement ignoré. Cependant, avec $q_{\min} < 0,5$, il estime deux pics d'amplitude 1 en n_5 et n_6 alors qu'il n'y en a qu'un seul. Si au contraire $q_{\min} > 0,5$, alors aucun pic ne sera estimé en n_5 et n_6, ce qui n'est pas non plus une solution acceptable.

3.5.3 Estimateur proposé

La principale limitation des estimateurs précédents provient de la prise de décision sous-jacente à l'estimateur. Pour le premier estimateur, tous les pics sont considérés comme étant significatifs, alors que pour le deuxième, seuls les plus probables sont pris en compte. Afin d'améliorer l'analyse de la chaîne de Markov, nous proposons un troisième estimateur constitué de deux indices représentant respectivement la probabilité d'occurrence d'un pic et son amplitude :

$$\widetilde{\mathbf{q}}_n \triangleq \frac{1}{I}\sum_{i=1}^{I} \mathbf{q}_n^{(i)}, \qquad \widetilde{\mathbf{x}}_n \triangleq \frac{1}{\sum_{i=1}^{I} \mathbf{q}_n^{(i)}} \sum_{\substack{i=1 \\ \mathbf{q}_n^{(i)}=1}}^{I} \mathbf{x}_n^{(i)}.$$

Nous ne considérons pas ces deux quantités comme des estimateurs à proprement parler de $\mathbf{q}$ et de $\mathbf{x}$. En effet, nous pensons qu'un estimateur doit conserver les contraintes les plus fortes imposées à la quantité estimée. Ainsi, un estimateur de $\mathbf{q}$ devrait être à valeurs dans $\{0,1\}$ et fournir une estimation de l'occurrence des pics. Ce n'est manifestement pas le cas de $\widetilde{\mathbf{q}}$ puisqu'il permet d'obtenir une estimation de la *probabilité* d'occurrence des pics. C'est pour cette raison que nous parlons d'indice et non d'estimateur.

Notons que pour $I \to \infty$, ces quantités reviennent à :

$$\widetilde{\mathbf{q}}_n \to \mathbb{E}_{\mathbf{q}_n}[\mathbf{q}_n], \qquad \widetilde{\mathbf{x}}_n \to \mathbb{E}_{\mathbf{x}_n}[\mathbf{x}_n|\mathbf{q}_n = 1].$$

Ainsi, dans le cas de la figure 3.9, les pics en n_1, n_3 et n_4 sont correctement estimés, et la méthode estime un pic d'amplitude 1 et de probabilité d'existence $1/4$ en n_2. Le caractère éphémère de l'artefact est donc représenté par une faible occurrence. Enfin, en n_5 et n_6 deux pics d'amplitude 1 sont estimés, mais de probabilité d'existence $1/2$, ce qui reflète mieux le comportement de la chaîne de Markov. Comme les deux pics sont très proches, on peut suivre l'idée de Rosec *et al.* [97] qui conseillent de remplacer ces pics par leur centre de gravité, c'est-à-dire que la position du vrai pic (respectivement son amplitude) correspond au barycentre des positions (respectivement des amplitudes) des pics, où les poids sont les probabilité d'occurrence, soit, pour notre exemple :

$$\text{position} = \frac{n_5\widetilde{\mathbf{q}}_{n_5} + n_6\widetilde{\mathbf{q}}_{n_6}}{\widetilde{\mathbf{q}}_{n_5} + \widetilde{\mathbf{q}}_{n_6}},$$
$$\text{amplitude} = \frac{\widetilde{\mathbf{x}}_{n_5}\widetilde{\mathbf{q}}_{n_5} + \widetilde{\mathbf{x}}_{n_6}\widetilde{\mathbf{q}}_{n_6}}{\widetilde{\mathbf{q}}_{n_5} + \widetilde{\mathbf{q}}_{n_6}}.$$

On peut donc penser qu'il n'y a qu'un seul pic situé en $(n_5 + n_6)/2$ et d'amplitude $(1 + 1)/2 = 1$. Si le vrai pic était situé plus près de n_5 que de n_6, alors la chaîne de Markov aurait logiquement généré plus de pics en n_5 qu'en n_6. La probabilité d'occurrence du pic en n_5 serait donc supérieure à celle en n_6 : c'est pour cette raison que le calcul de l'estimateur doit prendre en compte la probabilité d'occurrence.

Ces deux indices permettent ainsi un analyse plus fine de la chaîne de Markov que les estimateurs proposés par Cheng *et al.* [20] et Rosec *et al.* [97]. Contrairement à ceux-ci, nous préférons laisser la prise de décision à l'utilisateur. C'est une étape plus longue et plus fastidieuse puisqu'elle n'est pas automatique ; en revanche, elle permet d'éviter certaines mauvaises interprétations du résultat.

En fait, tout se passe comme si l'estimateur proposé regroupait en un seul signal presque toute l'information de la chaîne de Markov. En effet, la seule information perdue est l'ordre des itérations, mais cette information n'est pas nécessaire pour le calcul de l'estimateur.

3.6 Application sur des spectres simulés

3.6.1 Cas de raies identiques

Afin de valider notre méthode et d'illustrer la pertinence de l'estimateur, nous avons simulé un spectre sur lequel a été appliquée la méthode proposée dans ce chapitre[4]. Le spectre est sans ligne de base, de longueur $N = 500$ et sur lequel 10 raies ont été placées : cela correspond donc à un paramètre de Bernoulli $\lambda = 0,02$ et la variance de l'amplitude est $r_{\mathbf{a}} \approx 0,84$. Les raies sont toutes des lorentziennes de largeur $s = 6$, et le bruit a pour variance $r_{\mathbf{b}} = 5,25 \times 10^{-3}$, correspondant à un RSB de 20,5 dB environ. Le spectre et les raies sont représentés sur la figure 3.10 et les positions et amplitudes sont répertoriées dans le tableau 3.1.

Ce spectre permet de faire apparaître les principales difficultés auxquelles l'algorithme peut être confronté : raie d'intensité très faible, raies très proches et d'intensité très différentes ou similaires, etc.

La méthode proposée a été appliquée sur ce spectre simulé, avec les valeurs initiales suivantes :

$$\lambda = 0,5, \qquad r_{\mathbf{a}} = 2, \qquad s = 10, \qquad r_{\mathbf{b}} = 0,1,$$

et un spectre initial nul.

Nous avons effectué $I = 4000$ itérations de l'échantillonneur de Gibbs ; la figure 3.11 représente l'évolution des chaînes de $\mathbf{q}$, $r_{\mathbf{a}}$, λ, s et $r_{\mathbf{b}}$. La longueur de la période de transition a été choisie à $I_0 = 1000$ itérations. Le temps de calcul nécessaire a été d'environ 870 secondes.

[4] Les simulations et les calculs ont été réalisés avec le logiciel Matlab 6.5, sur un Pentium 4 à 1,8 GHz et 256 Mo de mémoire RAM.

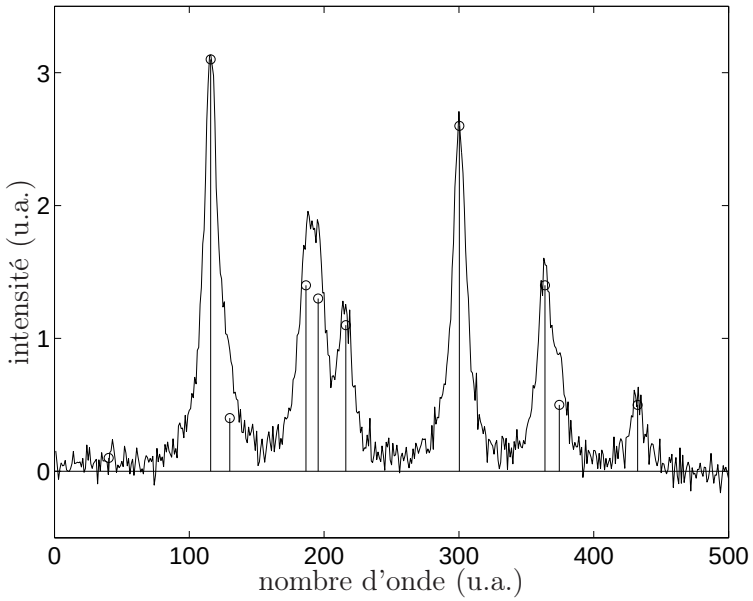


FIG. 3.10 – Spectre bruité simulé.

k	1	2	3	4	5	6	7	8	9	10
$\mathbf{c}^*$	40,2	115,9	130	186,5	195,6	216,1	300,3	363,8	374,4	432,5
$\mathbf{a}^*$	0,1	3,1	0,4	1,4	1,3	1,1	2,6	1,4	0,5	0,5

TAB. 3.1 – Positions et amplitudes des raies du spectre pur.

Estimation des hyperparamètres

Les hyperparamètres de la méthode ont été estimés au sens de l'EAP. Le tableau 3.2 donne les valeurs réelles et estimées des hyperparamètres, ainsi que l'écart-type et le biais des estimations.

	valeur réelle	**estimation**	**écart-type**	**biais**
$N_{\mathbf{q}}$	10	15,1270	1,2299	5,1270
λ	0,02	0,0107	0,0027	−0,0093
$r_{\mathbf{a}}$	0,8484	1,0064	0,3691	0,1580
s	6	5,8293	0,0754	−0,1707
$r_{\mathbf{b}}$	$5{,}25\times10^{-3}$	$5{,}44\times10^{-3}$	$0{,}36\times10^{-3}$	$0{,}19\times10^{-3}$

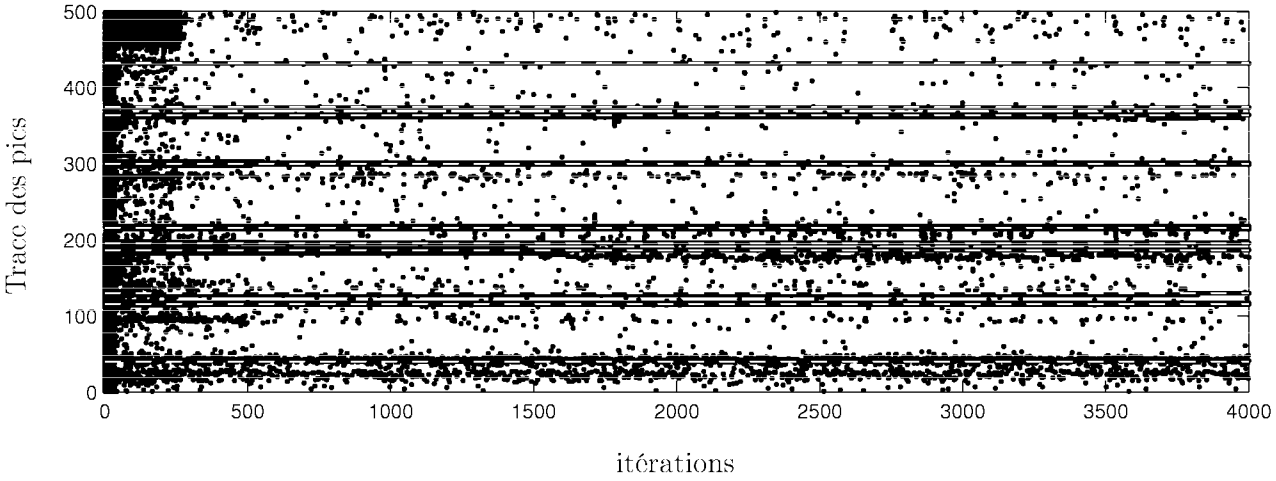
TAB. 3.2 – Estimation des hyperparamètres.

Nous pouvons constater que la largeur de la lorentzienne et la variance du bruit sont correctement estimés. Par contre, cette simulation a montré que $N_{\mathbf{q}}$ et $r_{\mathbf{a}}$ sont estimés avec un biais positif et λ avec un biais négatif. Ce comportement a également été observé dans d'autres simulations. En particulier, nous pensons que le biais sur $N_{\mathbf{q}}$ peut s'expliquer par le fait que certaines grandes raies sont parfois estimées par deux pics assez proches (voir plus loin) et que le biais négatif sur λ est peut être dû à la loi a priori qui serait trop informative.

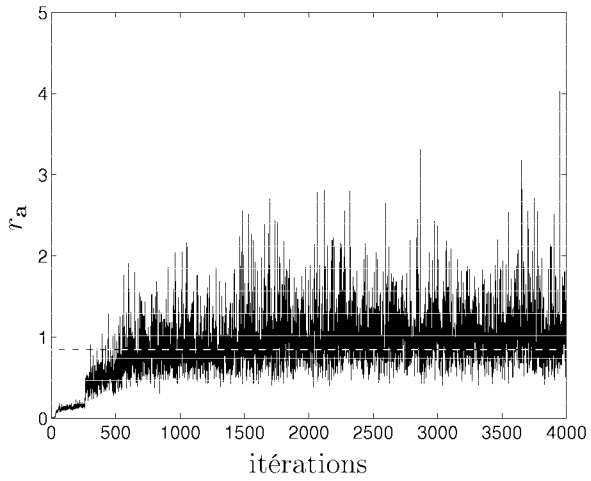
Estimation du spectre pur

L'estimation du spectre pur ($\widetilde{\mathbf{q}}$ et $\widetilde{\mathbf{x}}$) est présentée sur la figure 3.12.

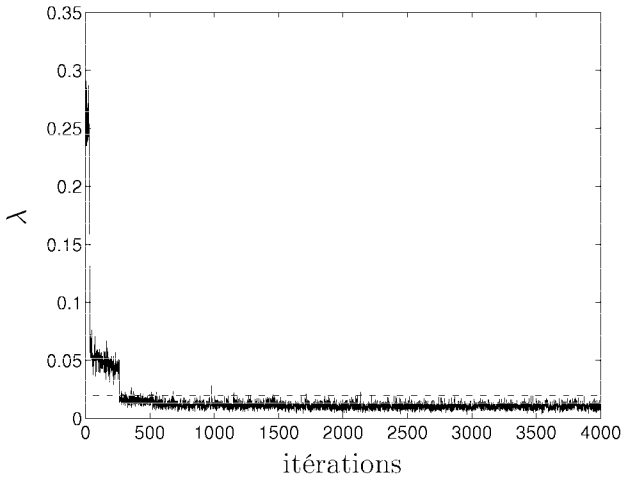
La raie 1, de petite amplitude, est estimée par un ensemble de pics de petites amplitudes et de probabilité d'occurrence de 10 % environ. Cet étalement est sans doute dû au fait que la méthode a du mal à positionner précisément la raie dans le bruit ambiant. Nous sommes donc dans le cas étudié



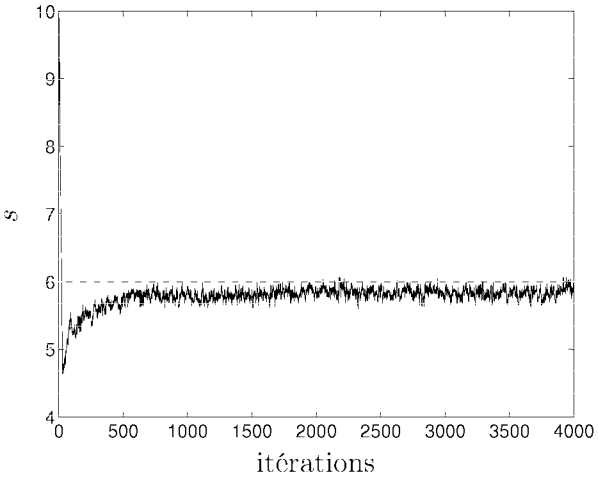
(a) Trace des pics.



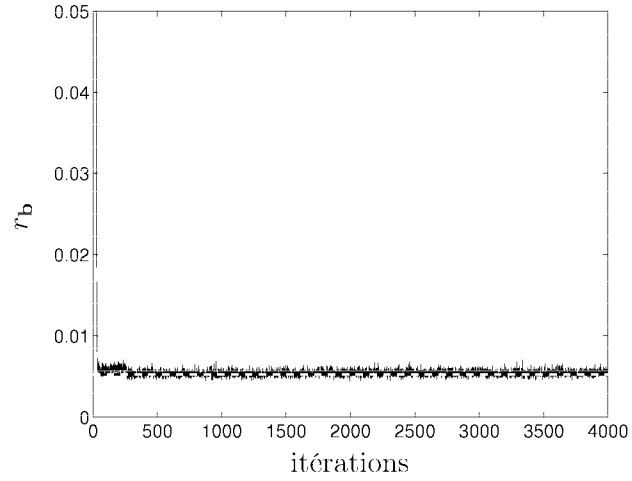
(b) Chaîne de Markov de $r_{\mathbf{a}}$.



(c) Chaîne de Markov de λ.



(d) Chaîne de Markov de s.



(e) Chaîne de Markov de $r_{\mathbf{b}}$.

FIG. 3.11 Évolution des chaînes de Markov (tirets : valeurs réelles).

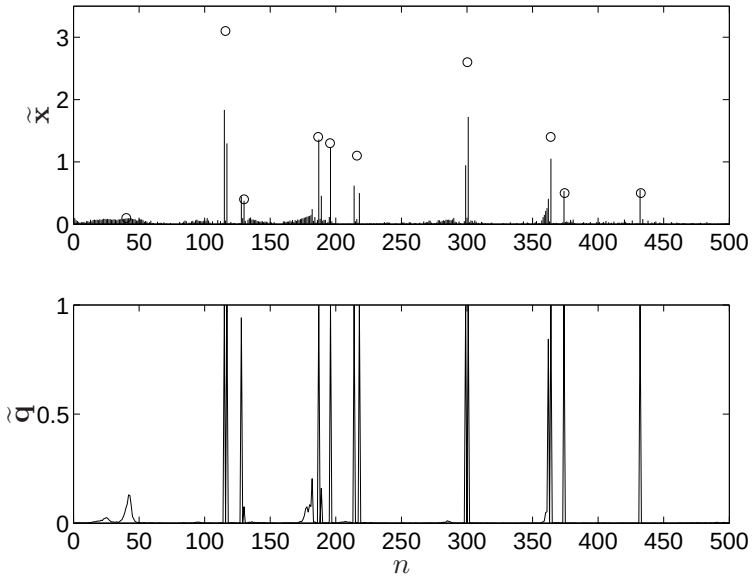


FIG. 3.12 – Estimation du signal impulsionnel. Les ∘ indiquent les raies réelles.

dans la section 3.5 pour les pics n_5 et n_6. Par conséquent, nous utilisons la procédure proposée pour recombiner cet ensemble de pic en un seul.

Par ailleurs, on s'aperçoit que les raies 2, 6, 7 et 8 ont été estimées par deux pics successifs ou séparés par un échantillon. Le fait d'avoir une raie estimée par deux pics est malheureusement assez fréquent en déconvolution impulsionnelle myope [6, 97]. Il s'explique par la lenteur de convergence de l'échantillonneur de Gibbs qui est attiré par des modes locaux et est incapable de passer par une solution de plus faible probabilité [14, 93]. En effet, comme l'échantillonneur de Gibbs ne simule qu'une variable à la fois, la solution qui consisterait à supprimer l'un des deux pics puis à mettre à jour les paramètres du pic restant conduirait à une probabilité plus faible que la probabilité de l'itération considérée : il est donc difficile pour la chaîne de s'échapper de ce mode local. Ce type de situation est une limitation de la méthode.

Bourguignon *et al.* [6] proposent une modification de l'échantillonneur de Gibbs en ajoutant une étape de simulation à l'algorithme. Cette étape consiste à inverser les valeurs de $\mathbf{q}$ de deux échantillons successifs choisis au hasard, puis à recalculer l'amplitude des échantillons par une méthode déterministe. Ce mouvement est accepté s'il accroît la loi conditionnelle $p(\mathbf{x}|\mathbf{w}, \lambda, r_{\mathbf{a}}, s, r_{\mathbf{b}})$. En pratique, cette procédure est proposée avec une probabilité décroissante exponentiellement, assurant la convergence de la chaîne vers la loi a posteriori. Cependant, aucune étude théorique n'a encore été réalisée par rapport à cette procédure.

Nous préférons appliquer une méthode manuelle qui consiste à remplacer les deux pics par un pic dont la position correspond à la moyenne des positions des pics pondérées par leur amplitude, et l'amplitude à la somme des amplitudes pondérées par la probabilité d'occurrence.

Le tableau 3.3 présente l'estimation finale, après avoir regroupé les pics.

k	1	2	3	4	5	6	7	8	9	10
$\mathbf{c}^*$	40,2	115,9	130	186,5	195,6	216,1	300,3	363,8	374,4	432,5
$\mathbf{a}^*$	0,1	3,1	0,4	1,4	1,3	1,1	2,6	1,4	0,5	0,5
$\widehat{\mathbf{c}}$	41,77	115,82	128	187	196	215,78	300,29	363,48	374	432
$\widehat{\mathbf{a}}$	0,09	3,14	0,44	1,36	1,23	1,12	2,67	1,46	0,53	0,55

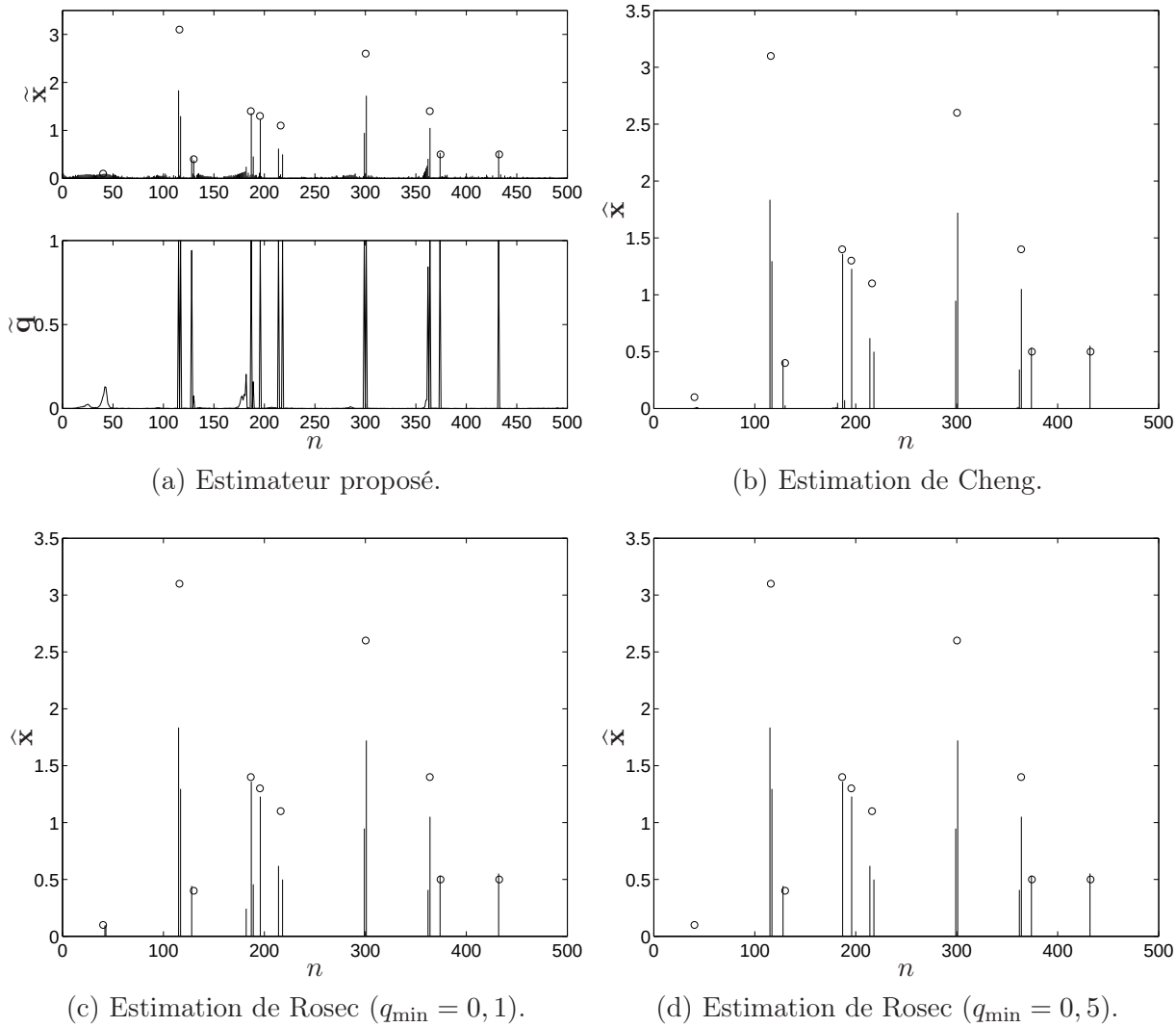
TAB. 3.3 – Estimation du signal impulsionnel après interprétation.

Le résultat de cette interprétation de l'estimation est très satisfaisant. En particulier, l'analyse visuelle de l'estimateur qui consiste à regrouper deux pics très proches donne un résultat très bon. Toutefois, on pourrait penser qu'il y a également un pic d'amplitude 0,16 en 179,78, car, comme la raie 1, il y a une zone de probabilité d'occurrence assez importante aux environ de $n = 180$.

En conclusion, la méthode d'estimation a permis, une fois que l'estimation ait été interprétée, de retrouver toutes les raies du signal, y compris la raie 1 noyée dans le bruit. Cependant, il est clair que la qualité de l'analyse est très fortement liée à la sélection de la zone relative à un pic, étape qui à l'heure actuelle est laissée à l'appréciation de l'utilisateur. Le développement d'une technique plus objective reste un problème ouvert et le sujet de recherches futures.

Comparaison avec les estimateurs de Cheng *et al.* et Rosec *et al.*

À titre de comparaison, les estimateurs de Cheng et Rosec ont été appliqués sur la chaîne de Markov issue de notre algorithme : les résultats sont présentés figure 3.13.



(a) Estimateur proposé.

(b) Estimation de Cheng.

(c) Estimation de Rosec ($q_{\min} = 0,1$).

(d) Estimation de Rosec ($q_{\min} = 0,5$).

FIG. 3.13 – Estimation du signal impulsionnel avec les estimateurs de Cheng *et al.*, Rosec *et al.* et le nôtre.

De ces trois résultats, on peut déjà remarquer que l'estimateur de Cheng *et al.* (figure 3.13(b)) ne fournit pas un résultat impulsionnel puisqu'il y a un pic à presque tous les échantillons. Pour cette raison, l'estimateur de Rosec *et al.* est préférable : en fonction du seuil $q_{\min}$, on peut obtenir plus ou moins de pics dans le signal estimé. Ainsi, avec $q_{\min} = 0,1$, on retrouve bien la première raie d'intensité

très faible (bien qu'elle soit estimée par plusieurs pics), mais de faux pics apparaissent également autour de la quatrième raie. Ces pics disparaissent si le seuil est augmenté, mais cela a pour conséquence de faire également disparaître l'estimation du premier pic.

Ces applications montrent, comme il a été dit de manière théorique dans la section précédente, qu'il est parfois préférable de ne pas utiliser de prise de décision automatique.

3.6.2 Cas de raies différentes

Nous avons voulu étudier le comportement de la méthode par déconvolution impulsionnelle myope dans le cas où les raies du spectre sont de largeurs différentes. Pour cela, nous avons simulé un spectre identique au spectre précédent mais dont les raies sont de largeurs différentes. La figure 3.14 représente le spectre et le tableau 3.4 reporte les paramètres des dix motifs.

La méthode proposée a été appliquée sur ce spectre simulé, avec les mêmes valeurs initiales que précédemment :

$$\lambda = 0,5, \qquad r_{\mathbf{a}} = 2, \qquad s = 10, \qquad r_{\mathbf{b}} = 0,1,$$

et un spectre initial nul.

Nous avons effectué $I = 4000$ itérations de l'échantillonneur de Gibbs et la longueur de la période de transition a été choisie à $I_0 = 1000$ itérations. Le temps de calcul nécessaire a été d'environ 877 secondes, ce qui est équivalent au temps de calcul qu'il a fallu dans le cas d'un spectre où les raies sont identiques.

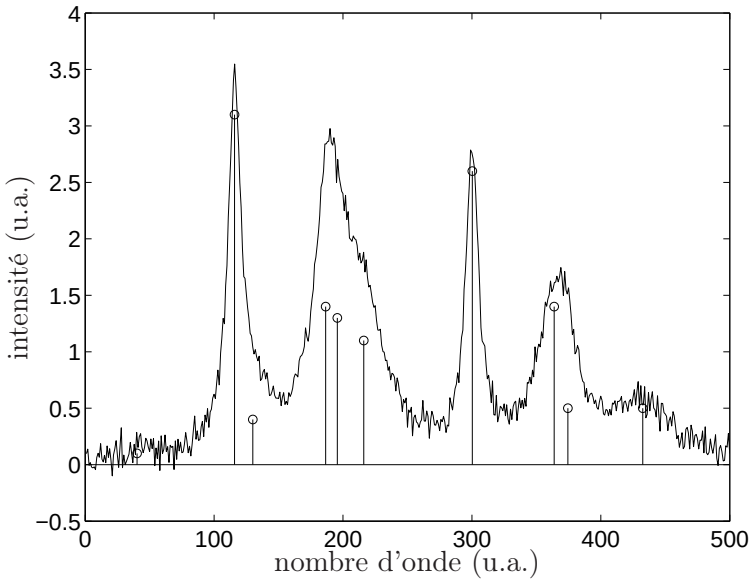


FIG. 3.14 – Spectre bruité simulé.

k	1	2	3	4	5	6	7	8	9	10
$\mathbf{c}^*$	40,2	115,9	130	186,5	195,6	216,1	300,3	363,8	374,4	432,5
$\mathbf{a}^*$	0,1	3,1	0,4	1,4	1,3	1,1	2,6	1,4	0,5	0,5
$\mathbf{s}^*$	10	6	15	10	15	20	6	15	6	30

TAB. 3.4 – Positions et amplitudes des raies du spectre pur.

L'estimation obtenue est représentée figure 3.15. Comme ce résultat est difficilement analysable, nous avons également représenté son interprétation figure 3.16 en regroupant les pics de la même manière que dans la section 3.6.1.

Tout d'abord, on s'aperçoit que les raies 2 et 7, qui sont les deux raies de largeur la plus faible, sont correctement estimé par un pic dont la position et l'amplitude sont bien estimés. Cela s'explique

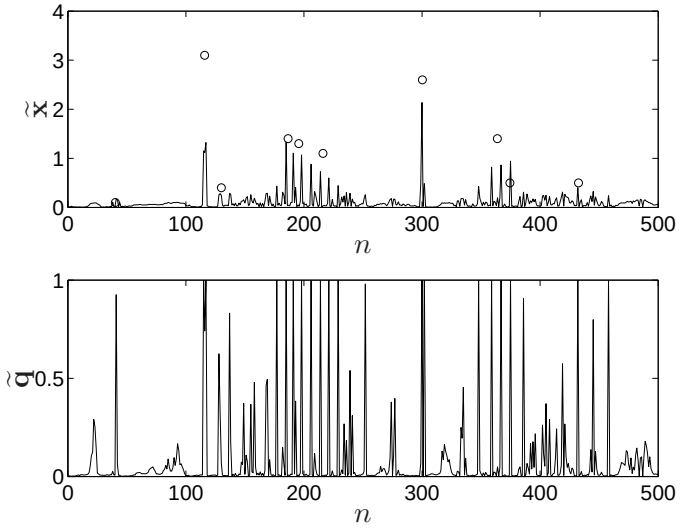


FIG. 3.15 – Estimation du signal impulsionnel. Les ∘ indiquent les raies réelles.

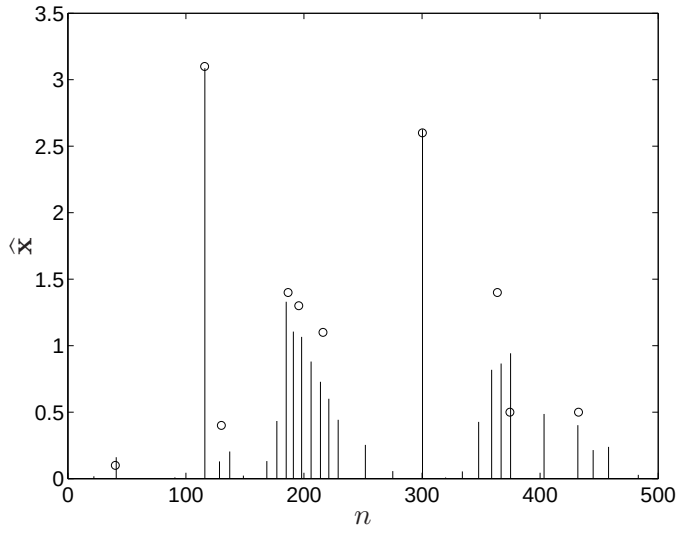


FIG. 3.16 – Interprétation de l'estimation du signal impulsionnel. Les ∘ indiquent les raies réelles.

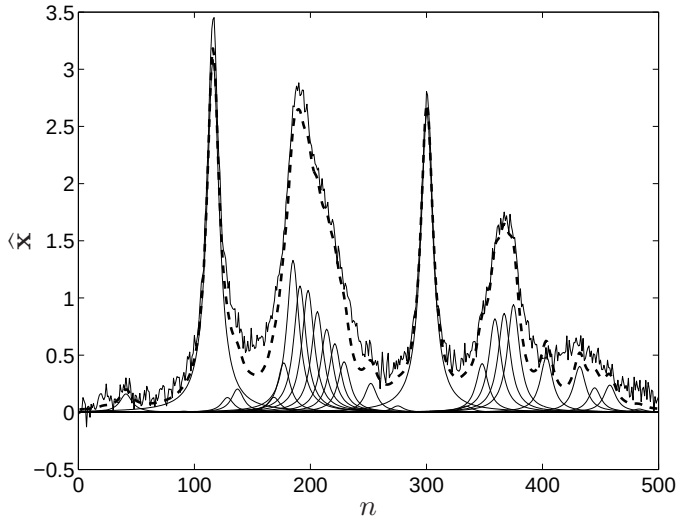


FIG. 3.17 – Raies estimées et spectre reconstruit (tirets).

par le fait que la seule manière possible pour la méthode de reconstruire correctement le spectre est d'estimer au mieux les raies les plus fines, puis d'estimer les autres à l'aide d'une combinaison de pics fins. En effet, la largeur estimée est égale à 6,34 et les raies plus larges du spectre sont estimées par plusieurs pics.

Notons que la raie 9 est également de largeur 6. Si elle n'est pas estimée correctement, c'est probablement parce que la raie 8 en est très proche, et qu'il est difficile pour l'algorithme de séparer correctement ces deux raies.

Le spectre reconstruit est représenté figure 3.17. La qualité de la reconstruction dépend de l'interprétation des pics qui est une étape subjective.

3.7 Conclusion

Dans ce chapitre, le problème de l'estimation des raies du spectre est vu comme un problème de déconvolution impulsionnelle positive myope non supervisée. Cette formulation implique deux simplifications importantes, à savoir que les raies estimées sont de même forme et de même paramètres de forme et qu'elles sont situées sur la grille d'échantillonnage. Ces simplifications permettent de résoudre le problème à l'aide d'une méthodologie très répandue et d'obtenir de bons résultats, pourvu que les vraies raies soient de formes identiques.

La méthode été développée dans un cadre totalement bayésien. En particulier, le spectre pur a été modélisé par un processus Bernoulli-gaussien restreint au support positif afin de ne générer que des pics d'amplitude positive. À cet égard, nous avons développé une méthode originale de simulation d'une loi normale tronquée grâce à un algorithme d'acceptation-rejet mixte (voir annexe C). Par ailleurs, nous avons également introduit dans ce chapitre le problème d'indéterminations en déconvolution myope. Dans notre cas, les indéterminations sont levées en posant comme hypothèses le caractère blanc — et plus particulièrement Bernoulli-gaussien — de $\mathbf{x}$ et le fait que $\mathbf{h}$ soit centré et d'amplitude maximale un.

Les raies étant de forme connue, seuls ses paramètres de forme sont à simuler ; cela permet non seulement de réduire le nombre de variables du problème mais également d'introduire dans le modèle le maximum d'information a priori. Dès lors, les performances de l'algorithme en termes de temps de calcul et d'estimation sont améliorées par rapport à une approche où aucune connaissance a priori n'est introduite sur $\mathbf{h}$.

Les hyperparamètres ont également été modélisés par des lois a priori, constituant en cela une approche bayésienne hiérarchique et permettant d'introduire des connaissances qualitatives. Nous avons utilisé deux types de lois a priori. D'une part, des lois a priori conjuguées ont été utilisées pour éviter d'avoir à faire à des problèmes de dégénérescence et permettre d'obtenir des lois a posteriori classiques et donc facilement simulables. D'autre part, des lois non-informatives de Laplace ou de Jeffreys ont été proposées afin de garantir l'intégrabilité des lois a posteriori conditionnelles, à défaut de l'intégrabilité de la loi a posteriori conjointe.

Les inconnues ont été simulées grâce à l'échantillonneur de Gibbs qui est une méthode d'optimisation naturelle dans le cas d'un modèle bayésien hiérarchique. Un estimateur original a alors été proposé en calculant deux indices qui permettent de condenser toute l'information intéressante de la chaîne de Markov dans un seul signal. Une analyse de ce signal permet, en particulier, de retrouver des pics qui ont été difficilement générés par la chaîne de Markov.

Enfin, la méthode a été appliquée sur un spectre simulé afin d'illustrer ses performances ainsi que la pertinence de l'estimateur. Il a cependant été noté que, parfois, une raie peut être estimée par deux pics. Ce problème provient de la lenteur de convergence de l'échantillonneur de Gibbs et constitue donc un inconvénient de la méthode. Néanmoins, une analyse visuelle permet de regrouper ces deux pics, résultant alors en une estimation correcte de la raie.

Malgré tout, cette approche possède trois limitations particulièrement gênantes.

La première limitation réside dans le fait que, à cause de la modélisation Bernoulli-gaussienne, les raies estimées ne sont situées qu'aux « instants » d'échantillonnage. Il serait intéressant de mettre en place une méthode qui puisse estimer les positions des raies en n'importe quel nombre d'onde. Un processus de Poisson, qui peut être considéré comme un cas limite du processus de Bernoulli avec un pas de discrétisation nul, pourrait résoudre ce problème [67].

La deuxième limitation réside dans l'opération de convolution qui suppose implicitement que toutes les raies sont identiques. Bien que l'approche développée dans ce chapitre donne une estimation peu satisfaisante mais que l'on sait toutefois interpréter, il est tout de même préférable de développer une méthode qui puisse prendre en compte les différences de forme des raies.

Enfin, la troisième limitation concerne le nombre de variables à estimer. En effet, dans un spectre de N échantillons, seuls λN échantillons correspondent à des pics d'amplitude non nulle. La méthode proposée ici nécessite tout de même d'estimer les $(1-\lambda)N$ échantillons nuls. Ce problème est bien sûr intrinsèque à la modélisation Bernoulli-gaussienne. Or, cette modélisation considère implicitement que $\lambda \approx 0$, sans quoi le signal ne possède plus son caractère impulsionnel et le processus Bernoulli-gaussien devient un modèle inadapté. Il est donc dommage de simuler tous les échantillons du signal sachant qu'un grand nombre est nul.

Le chapitre suivant propose une modélisation différente qui permet d'éliminer ces trois limitations : le problème est alors considéré comme une décomposition en motifs élémentaires.

Chapitre 4

Estimation des raies par décomposition en motifs élémentaires

4.1 Formulation du problème

4.1.1 Modélisation du spectre

Nous proposons dans ce chapitre une alternative à la méthode d'estimation des raies proposée dans le chapitre 3. Cette nouvelle approche consiste à exploiter directement l'équation (1.2) sans faire de simplification ni sur les positions des raies, ni sur leur paramètres de forme. Nous considérons donc le spectre pur comme une somme de *motifs* (des lorentziennes) dont la position, l'amplitude et les paramètres sont à estimer :

$$\forall\, n \in \{1, \dots, N\}, \qquad \mathbf{v}_n = \sum_{k=1}^{K} \mathbf{a}_k \frac{\mathbf{s}_k^2}{\mathbf{s}_k^2 + (n - \mathbf{c}_k)^2},$$

soit :

$$\mathbf{v} = \sum_{k=1}^{K} \mathbf{a}_k \mathbf{h}(\mathbf{c}_k, \mathbf{s}_k) \tag{4.1}$$

où :

$$\mathbf{h}(\mathbf{c}_k, \mathbf{s}_k) = \begin{pmatrix} \mathbf{h}_1(\mathbf{c}_k, \mathbf{s}_k) \\ \vdots \\ \mathbf{h}_N(\mathbf{c}_k, \mathbf{s}_k) \end{pmatrix} \qquad \text{et} \qquad \mathbf{h}_n(\mathbf{c}_k, \mathbf{s}_k) = \frac{\mathbf{s}_k^2}{\mathbf{s}_k^2 + (n - \mathbf{c}_k)^2}.$$

En temps continu, l'approche utilisée ici peut être modélisée par le système représenté figure 4.1, où $h_i(t)$ est une lorentzienne de largeur $\mathbf{s}_i$.

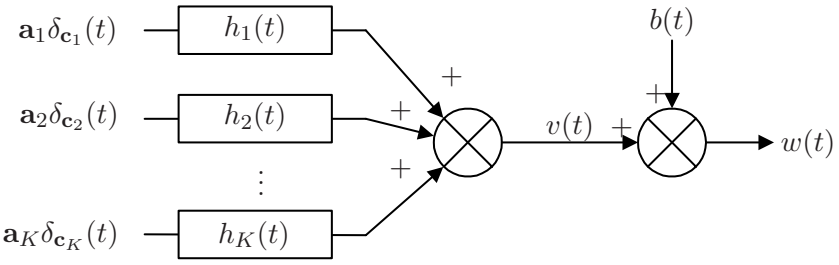


FIG. 4.1 – Décomposition en motifs élémentaires dans le cas de motifs différents.

Le problème de l'estimation des raies du spectre consiste donc désormais en l'estimation du nombre de motifs K, de leur position, amplitude et paramètres : c'est donc clairement un problème de décomposition en motifs élémentaires.

Cette approche peut bien sûr être appliquée aux spectres dont les motifs sont tous de même largeur. Pour cela, il faut considérer que tous les éléments de $\mathbf{s}$ sont identiques, donc que $\mathbf{s}_k = s$ pour tout $k \in \{1, \dots, K\}$. Si cela n'est pas fait, l'estimation risque d'être moins bonne puisque le fait de ne pas considérer les motifs identiques constitue une connaissance a priori en moins. De plus, comme il y a plus de variables à estimer, l'estimation sera plus lente.

Grâce à cette modélisation, les limitations du modèle précédent ont été supprimées : les motifs ne sont pas forcés en des positions discrètes, les paramètres de forme peuvent être différents et le nombre de variables est réduit. Désormais, la difficulté provient du choix de la forme du motif : en spectroscopie infrarouge et Raman, un motif lorentzien ou gaussien convient la plupart du temps, mais dans certains cas les raies ne sont pas symétriques ou elles ont une forme plus compliquée qu'une simple lorentzienne ou gaussienne (par exemple une fonction de Voigt). En outre, ce modèle considère que le spectre est échantillonné de manière régulière, ce qui n'est malheureusement pas toujours le cas pour des raisons techniques.

4.1.2 Méthodes de décomposition en motifs élémentaires

Dans [83, 84], Mohammad-Djafari pose le problème de la décomposition en motifs élémentaires dans un cadre bayésien, en supposant entre autres que la forme des réponses impulsionnelles est connue. L'optimisation du critère a posteriori est effectué grâce à une version modifiée de l'algorithme de Newton-Raphson ou grâce à un algorithme de type acceptation-rejet, au risque de tomber dans un minimum local. L'estimation du nombre de motifs K n'est cependant pas optimale puisqu'elle consiste à répéter K fois l'algorithme en testant toutes les valeurs de K, et en sélectionnant celle qui minimise le critère. De fait, l'algorithme est très long à converger.

La complexité du problème fait maintenant préférer l'usage de méthodes MCMC. Citons parmi les travaux publiés ceux de Fischer *et al.* [34], de Gulam Razul *et al.* [47, 48] en spectroscopie nucléaire et de Haan *et al.* [49, 50] en génétique (séquençage d'ADN). Ces travaux sont similaires puisqu'ils sont posés dans un cadre bayésien et utilisent le même algorithme d'optimisation, à savoir l'algorithme RJMCMC (*Reversible Jump MCMC*) de Green [46, 91]. En effet, une méthode MCMC classique ne peut pas être utilisée car la dimension du modèle peut varier d'une itération à l'autre. L'algorithme RJMCMC effectue des « sauts » entre des espaces de dimensions différentes, permettant ainsi au nombre de motifs de varier.

Les différences principales entre ces méthodes viennent du choix des a priori des paramètres du problème et de l'application.

Malheureusement, à part l'estimation du nombre de motifs qui est effectuée en calculant le maximum a posteriori marginal (MMAP), ces travaux ne détaillent pas l'estimateur utilisé pour le signal impulsionnel. Pourtant, c'est une étape indispensable et qui, dans le cas de la décomposition en motifs élémentaires, fait apparaître des problèmes de permutation d'indices très gênants pour l'estimation du signal. Dans la section 4.4, nous discutons des méthodes de ré-indexage traditionnellement utilisées en décomposition de mélange de lois et proposons une méthode originale et efficace dans le cadre de notre estimation.

4.1.3 Approche retenue

Pour les mêmes raisons que dans le chapitre précédent, l'inférence bayésienne couplée à une méthode MCMC est utilisée pour apporter le maximum de connaissances a priori et résoudre le problème d'optimisation.

Toutefois, une méthode MCMC classique ne peut s'appliquer ici, en particulier parce que la loi jointe des paramètres n'est pas stationnaire car le nombre de motifs (et donc de variables à estimer) peut varier d'une itération à l'autre. En d'autres termes, le nombre d'inconnues est également une inconnue. Nous proposons dans ce chapitre une approche originale pour laquelle le modèle est de dimension

fixe, permettant alors d'utiliser un algorithme MCMC classique (en l'occurrence, l'échantillonneur de Gibbs).

L'idée est de considérer le nombre de motifs fixe et égal à un nombre maximal $K_{\max}$. Ce nombre est fixé a priori par l'utilisateur et choisi plus grand que le nombre réel de motifs K, mais inférieur au nombre de points N. Pour éviter d'obtenir un signal estimé de $K_{\max}$ motifs, nous nous inspirons de la modélisation Bernoulli-gaussienne (section 3.3.1.1) en introduisant le vecteur $\mathbf{q} \in \{0,1\}^{K_{\max}}$ codant l'occurrence des motifs. Ainsi :

$$\forall k \in \{1,\dots,K_{\max}\}, \qquad \begin{cases} \mathbf{q}_k = 1 \;\Rightarrow\; \mathbf{a}_k \neq 0 \;\Rightarrow\; \text{le motif } k \text{ est présent dans le spectre,} \\ \mathbf{q}_k = 0 \;\Rightarrow\; \mathbf{a}_k = 0 \;\Rightarrow\; \text{le motif } k \text{ est absent du spectre.} \end{cases}$$

Si le motif k est présent, il a pour position $\mathbf{c}_k$, pour amplitude $\mathbf{a}_k$ et pour largeur $\mathbf{s}_k$. En revanche, s'il est absent, cela revient à le négliger car il n'apparaît pas dans le spectre. L'amplitude de ce motif est nulle mais sa position et sa largeur ne le sont pas : ce choix est motivé par la méthode de simulation utilisée (voir section 4.3).

On note K le nombre de motifs présents et $\mathbf{x} = (\mathbf{q}, \mathbf{a})$ (contrairement au chapitre précédent, les vecteurs sont maintenant de dimension $K_{\max}$). Toujours en s'inspirant de la modélisation Bernoulli-gaussienne, les occurrences des motifs $\mathbf{q}_k$ sont distribuées suivant une loi de Bernoulli de paramètre λ et les amplitudes $\mathbf{a}_k$ sont distribuées suivant une loi normale à support positif. L'objectif est donc d'estimer le triplet $(\mathbf{c}, \mathbf{x}, \mathbf{s})$ et les hyperparamètres du problème.

L'équation 4.1 se réécrit alors :

$$\mathbf{v} = \sum_{k=1}^{K_{\max}} \mathbf{x}_k \mathbf{h}(\mathbf{c}_k, \mathbf{s}_k),$$

soit sous forme matricielle :

$$\mathbf{v} = \mathbf{G}\mathbf{x}, \tag{4.2}$$

où $\mathbf{G}$ est la matrice :

$$\mathbf{G} = \big(\mathbf{h}(\mathbf{c}_1, \mathbf{s}_1) \dots \mathbf{h}(\mathbf{c}_K, \mathbf{s}_K)\big).$$

Grâce à cette modélisation, les motifs peuvent maintenant se situer en dehors de la grille d'échantillonnage ($\mathbf{c}_k$ peut prendre n'importe quelle valeur entre 1 et N) et être différents les uns des autres. En outre, le nombre de variables à estimer est inférieur à l'approche précédente puisqu'il est ici de $4K_{\max} + 3$ ($\mathbf{q}$, $\mathbf{c}$, $\mathbf{a}$, $\mathbf{s}$, λ, $r_{\mathbf{a}}$ et $r_{\mathbf{b}}$), alors qu'avec un a priori Bernoulli-gaussien, il est de $2N + 4$. En effet, comme on suppose qu'il y a peu de motifs, on peut raisonnablement considérer que $K_{\max} < \frac{1}{2}N$. Par conséquent, la méthode sera plus rapide (à condition que les simulations de l'échantillonneur de Gibbs soient de durées similaires pour les deux approches) et l'estimation de meilleure qualité (car il y a moins d'inconnues pour le même nombre de données).

Plutôt qu'une décomposition en motifs élémentaires, une déconvolution impulsionnelle positive myope dans laquelle le signal impulsionnel est modélisé par un processus Poisson-gaussien aurait pu être envisagée [67]. Cette modélisation permet non seulement de garantir que les raies sont situées en dehors de la grille discrète mais également de réduire le nombre de variables. Cependant, l'approche par déconvolution ne permet pas d'avoir des raies de forme différentes. En outre, la dimension du modèle est variable.

Il est possible d'établir un lien entre la déconvolution impulsionnelle (myope ou non) et la décomposition en motifs élémentaires. En effet, toute déconvolution impulsionnelle peut être vue et traitée comme une décomposition en motifs élémentaires dans laquelle les motifs sont identiques et situés aux « instants » d'échantillonnage. Ce lien a déjà été signalé par Mohammad-Djafari [83, 84]. La méthode proposée dans ce chapitre nous paraît alors être une alternative intéressante à la déconvolution impulsionnelle puisque le nombre de variables à estimer est plus faible.

On peut alors effectuer une comparaison des notations entre la déconvolution Bernoulli-gaussienne et la décomposition en motifs élémentaires (tableau 4.1).

	déconvolution Bernoulli-gaussienne	décomposition en motifs élémentaires
Nombre de pics/motifs	$N_{\mathbf{q}}$	K
Nombre maximal de pics/motifs	N	$K_{\max}$
Positions des pics/centres de motifs	$\{1,\dots,N\}$	$\mathbf{c}$
Amplitude des pics/des motifs	$\mathbf{x}=(\mathbf{q},\mathbf{a})$	$\mathbf{x}=(\mathbf{q},\mathbf{a})$

TAB. 4.1 – Correspondance des notations entre les approches des chapitres 3 et 4.

4.1.3.1 Lien avec la décomposition de mélange de lois

D'une certaine manière, le problème de la décomposition en motifs élémentaires ressemble au problème de décomposition en mélange de lois (*analysis of mixtures models*) qui est un problème largement traité dans la littérature (voir par exemple [14, 46, 91, 104]). Le parallèle entre ces deux approches a d'ailleurs été noté dans [34].

Le modèle d'un mélange de loi pour des observations indépendantes $\boldsymbol{\chi}=\{\boldsymbol{\chi}_1,\dots,\boldsymbol{\chi}_N\}$ considère le cas où les données sont distribuées suivant la loi :

$$\forall\, n\in\{1,\dots,N\},\qquad \boldsymbol{\chi}_n\sim p(\boldsymbol{\chi}_n|\boldsymbol{\pi},\boldsymbol{\xi})=\sum_{k=1}^{K}\boldsymbol{\pi}_k f_k(\boldsymbol{\chi}_n|\boldsymbol{\xi}_k),$$

où les $\boldsymbol{\pi}_k$ sont les proportions du mélange (elles suivent une loi de Dirichlet, c'est-à-dire qu'elles sont comprises entre 0 et 1 et de somme unité) et $\{f_k\}$ est une famille de densités de paramètres $\boldsymbol{\xi}_k$. L'objectif de ce problème consiste en l'estimation des paramètres $\boldsymbol{\theta}_k=\{\boldsymbol{\pi}_k,\boldsymbol{\xi}_k\}$ à partir de N réalisations, ce qui est très semblable au problème que l'on traite dans ce chapitre. Notons qu'une contrainte forte du problème consiste à considérer la somme des proportions égale à l'unité.

La différence principale entre ces deux modélisations réside essentiellement dans l'expression de la fonction de vraisemblance et, par conséquent, dans l'expression des lois a posteriori. Malgré cette différence, les méthodes de résolution ainsi que certains problèmes sont communs [34].

Ainsi, un grand nombre d'articles traitant de l'une ou l'autre de ces deux approches utilisent une approche bayésienne et un algorithme MCMC pour l'optimisation de la loi a posteriori. En général, comme l'ordre du modèle K est inconnu, l'algorithme RJMCMC est proposé.

De surcroît, ces deux approches font face au problème de « permutation d'indices » (*label-switching problem*, parfois appelé en français *renversement d'étiquetage*) qui définit le fait que, d'une itération à l'autre, deux paramètres ($\boldsymbol{\theta}_k$ et $\boldsymbol{\theta}_{k'}$ par exemple) peuvent permuter. Dès lors, le calcul de l'estimateur est très délicat. Beaucoup de travaux se contentent d'introduire une contrainte d'identifiabilité (par exemple imposer un ordre inaltérable sur les variables), comme cela a été proposé dans [31]. La section 4.4 présente les méthodes de ré-indexage proposées dans la littérature ainsi qu'une alternative à ces méthodes qui a l'avantage notamment de considérer un nombre de motifs variables en exploitant le modèle original traité dans ce chapitre.

4.1.4 Organisation du chapitre

Le chapitre est organisé comme suit. Dans un premier temps, le choix des lois a priori est considéré dans la section 4.2. Ces lois sont, pour la plupart, similaires aux lois proposées dans le chapitre précédent.

L'implémentation de l'échantillonneur de Gibbs est ensuite présentée dans la section 4.3. Elle reste également très semblable à l'implémentation du chapitre 3 pour la plupart des variables.

L'estimation du signal impulsionnel est ensuite discutée dans la section 4.4. En effet, peu de travaux traitant de la décomposition en motifs élémentaires décrivent précisément la procédure à appliquer pour obtenir une estimation exploitable à partir de la chaîne de Markov. Or, cette étape primordiale pour l'analyse du spectre n'est pas évidente. En effet, une simple estimation au sens de l'espérance a posteriori est parfois inadéquate, du fait du problème de permutation d'indices. Aussi, nous présentons plus en détail ce problème dans la section 4.4, puis, après avoir donné une liste des méthodes permettant de résoudre ce problème, nous proposons une procédure originale qui permet d'obtenir une estimation satisfaisante du signal impulsionnel.

La méthode proposée est ensuite appliquée sur des signaux simulés et réels (sections 4.5 et 4.6). Afin de comparer avec l'approche par déconvolution impulsionnelle myope, nous travaillons avec les mêmes signaux simulés que dans le chapitre 3. Les signaux réels correspondent aux signaux de gibbsite dont la ligne de base a été corrigée à l'aide de l'approche présentée dans le chapitre 2.

Enfin, la section 4.7 conclue le chapitre.

4.2 Modélisation probabiliste

4.2.1 Lois a priori

4.2.1.1 Loi a priori de la position des motifs c

Nous supposons que les motifs sont situés aléatoirement sur $[1, N]$. Pour cela, une loi a priori uniforme est utilisée pour $\mathbf{c}_k$:

$$\mathbf{c}_k \sim \mathcal{U}_{[1,N]}. \tag{4.3}$$

Cette loi ne correspond pas à une loi non-informative due à un manque de connaissance, mais reflète bien la volonté de modéliser la position des motifs sur l'intervalle $[1, N]$ sans privilégier un emplacement par rapport à un autre.

4.2.1.2 Loi a priori de l'amplitude des motifs x

De même que dans le chapitre 3, un a priori Bernoulli-gaussien à support positif est utilisé pour $\mathbf{x}_k$. Ainsi :

$$\mathbf{q}_k \sim \mathcal{B}er(\lambda), \qquad \text{et} \qquad \mathbf{a}_k \sim \mathcal{N}^+(0, r_\mathbf{a}\mathbf{q}_k),$$

d'où :

$$\mathbf{x}_k \sim \lambda\mathcal{N}^+(0, r_\mathbf{a}) + (1-\lambda)\delta_0(\mathbf{x}_k). \tag{4.4}$$

4.2.1.3 Lois a priori de la largeur s

Comme dans le cas d'un motif unique, aucun a priori n'est posé pour $\mathbf{s}_k$, si ce n'est toujours un a priori de positivité. La loi a priori s'écrit donc :

$$\mathbf{s}_k \sim \mathcal{U}_{\mathbb{R}^+}. \tag{4.5}$$

4.2.1.4 Autres lois a priori

Les lois a priori sur λ, $r_\mathbf{a}$ et $r_\mathbf{b}$ sont similaires à celles du chapitre précédent :

$$\lambda \sim \mathcal{B}e(1, K+1), \tag{4.6}$$
$$r_\mathbf{a} \sim \mathcal{IG}(\alpha_\mathbf{a}, \beta_\mathbf{a}), \tag{4.7}$$
$$r_\mathbf{b} \sim 1/r_\mathbf{b}. \tag{4.8}$$

où K est le nombre de motifs existants, c'est-à-dire tels que $\mathbf{q}_k = 1$:

$$K = \sum_{k=1}^{K_{\max}} \mathbf{q}_k$$

et

$$\alpha_{\mathbf{a}} = 2 + \varepsilon, \qquad \beta_{\mathbf{a}} = 1 + \varepsilon, \qquad \varepsilon \ll 1.$$

4.2.2 Graphe acyclique orienté

Le graphe acyclique orienté du modèle est représenté figure 4.2.

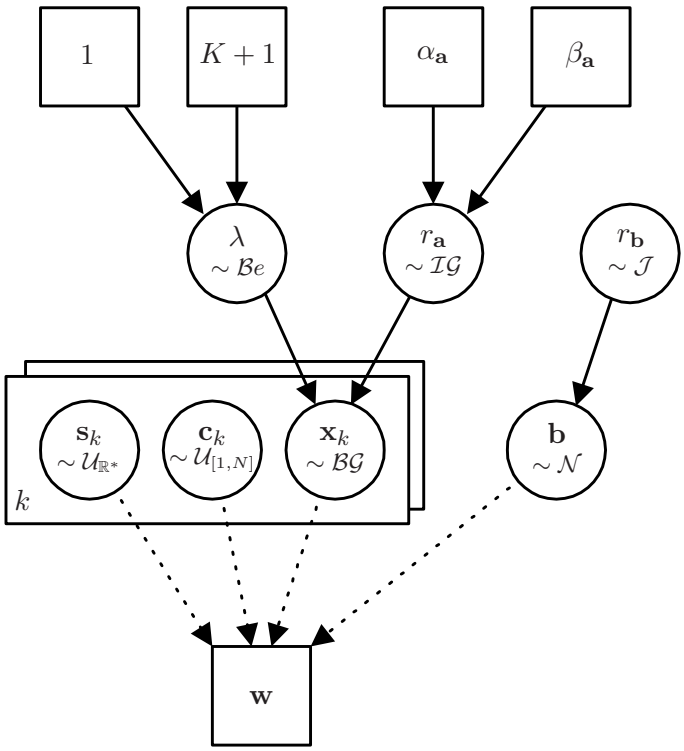


FIG. 4.2 – Graphe acyclique orienté du modèle de décomposition en motifs élémentaires.

4.2.3 Lois a posteriori

Il y a $4K_{\max} + 3$ paramètres à estimer :

$$\boldsymbol{\theta} = \{\mathbf{q}, \mathbf{c}, \mathbf{a}, \mathbf{s}, \lambda, r_{\mathbf{a}}, r_{\mathbf{b}}\}.$$

Dans le cas (le plus courant) où $K_{\max} < \frac{1}{2}N$, la modélisation proposée ici permet donc d'avoir moins de variables à estimer que dans le cas d'une modélisation Bernoulli-gaussienne pour le même nombre de données.

D'après la règle de Bayes la loi a posteriori globale s'écrit :

$$\begin{aligned} p(\mathbf{c}, \mathbf{x}, s, \lambda, r_{\mathbf{a}}, r_{\mathbf{b}} | \mathbf{w}) \propto & \frac{1}{(2\pi r_{\mathbf{b}})^{N/2}} \exp\left(-\frac{1}{2r_{\mathbf{b}}} ||\mathbf{w} - \mathbf{G}\mathbf{x}||^2\right) \\ & \times \prod_{k=1}^{K_{\max}} \left[\lambda \sqrt{\frac{2}{\pi r_{\mathbf{a}}}} \exp\left(-\frac{\mathbf{x}_k^2}{2r_{\mathbf{a}}}\right) \mathbb{1}_{\mathbb{R}^+}(\mathbf{x}_k) + (1-\lambda)\delta_0(\mathbf{x}_k)\right] \\ & \times \prod_{k=1}^{K_{\max}} \mathbb{1}_{[1,N]}(\mathbf{c}_k) \times \prod_{k=1}^{K_{\max}} \mathbb{1}_{\mathbb{R}^+}(\mathbf{s}_k) \\ & \times \frac{\lambda^0 (1-\lambda)^{K_{\max}}}{B(1, K_{\max}+1)} \mathbb{1}_{[0,1]}(\lambda) \times \frac{\beta_{\mathbf{a}}^{\alpha_{\mathbf{a}}}}{\Gamma(\alpha_{\mathbf{a}})} \frac{\mathrm{e}^{-\beta_{\mathbf{a}}/r_{\mathbf{a}}}}{r_{\mathbf{a}}^{\alpha_{\mathbf{a}}+1}} \mathbb{1}_{\mathbb{R}^+}(r_{\mathbf{a}}) \times \frac{1}{r_{\mathbf{b}}}, \end{aligned}$$

où :

$$\mathbf{G} = \big(\mathbf{h}(\mathbf{c}_1, \mathbf{s}_1) \dots \mathbf{h}(\mathbf{c}_{K_{\max}}, \mathbf{s}_{K_{\max}})\big).$$

Comme dans le chapitre précédent, et pour les mêmes raisons, nous n'avons pas vérifié l'intégrabilité de la loi a posteriori (voir également page 130).

4.2.3.1 Loi a posteriori de la position des motifs $\mathbf{c}_k$

La loi a posteriori de $\mathbf{c}_k$ s'écrit :

$$\begin{aligned} p(\mathbf{c}_k|\mathbf{w}, \mathbf{c}_{-k}, \mathbf{x}, s, \lambda, r_\mathbf{a}, r_\mathbf{b}) &= p(\mathbf{w}|\mathbf{c}_k, \mathbf{c}_{-k}, \mathbf{x}, s, \lambda, r_\mathbf{a}, r_\mathbf{b}) p(\mathbf{c}_k|\mathbf{c}_{-k}, \mathbf{x}, s, \lambda, r_\mathbf{a}, r_\mathbf{b}), \\ &= p(\mathbf{w}|\mathbf{c}, \mathbf{x}, s, r_\mathbf{b}) p(\mathbf{c}_k), \end{aligned}$$

d'où :

$$\mathbf{c}_k \sim \exp\left(-\frac{1}{2r_\mathbf{b}}||\mathbf{w} - \mathbf{G}\mathbf{x}||^2\right) \mathbb{1}_{[1,N]}(\mathbf{c}_k)$$

où $\mathbf{c}_k$ intervient dans la matrice $\mathbf{G}$.

4.2.3.2 Loi a posteriori de l'amplitude de motifs $\mathbf{x}_k$

La loi a posteriori de $\mathbf{x}_k$ s'écrit :

$$\begin{aligned} p(\mathbf{x}_k|\mathbf{w}, \mathbf{c}, \mathbf{x}_{-k}, s, \lambda, r_\mathbf{a}, r_\mathbf{b}) &= p(\mathbf{w}|\mathbf{x}_k, \mathbf{c}, \mathbf{x}_{-k}, s, \lambda, r_\mathbf{a}, r_\mathbf{b}) p(\mathbf{x}_k|\mathbf{c}, \mathbf{x}_{-k}, s, \lambda, r_\mathbf{a}, r_\mathbf{b}), \\ &= p(\mathbf{w}|\mathbf{c}, \mathbf{x}, s, r_\mathbf{b}) p(\mathbf{x}_k|\lambda, r_\mathbf{a}), \\ &= \frac{1}{(2\pi r_\mathbf{b})^{N/2}} \exp\left(-\frac{1}{2r_\mathbf{b}}||\mathbf{w} - \mathbf{G}\mathbf{x}||^2\right) \\ &\quad \times \left[\lambda\sqrt{\frac{2}{\pi r_\mathbf{a}}} \exp\left(-\frac{\mathbf{x}_k^2}{2r_\mathbf{a}}\right) \mathbb{1}_{\mathbb{R}^+}(\mathbf{x}_k) + (1-\lambda)\delta_0(\mathbf{x}_k)\right]. \end{aligned}$$

On montre (en s'inspirant de l'annexe B) que :

$$\mathbf{x}_k \sim \lambda_{0,k}\mathcal{N}^+(\mu_{0,k}, \rho_{0,k}) + \lambda_{1,k}\mathcal{N}^+(\mu_{1,k}, \rho_{1,k}), \tag{4.9}$$

ce qui revient à :

$$\mathbf{q}_k \sim \mathcal{B}er(\lambda_{1,k}), \tag{4.10}$$

$$\mathbf{a}_k \sim \begin{cases} \mathcal{N}^+(\mu_{0,k}, \rho_{0,k}) & \text{si } \mathbf{q}_k = 0, \\ \mathcal{N}^+(\mu_{1,k}, \rho_{1,k}) & \text{si } \mathbf{q}_k = 1. \end{cases} \tag{4.11}$$

avec

$$\lambda_{1,k} = \left[1 + \frac{1-\lambda}{\lambda}\sqrt{\frac{r_\mathbf{a}}{\rho_{1,k}}} \exp\left(-\frac{\mu_{1,k}^2}{2\rho_{1,k}}\right)\right]^{-1},$$

$$\lambda_{0,k} = \left[1 + \frac{\lambda}{1-\lambda}\sqrt{\frac{\rho_{1,k}}{r_\mathbf{a}}} \exp\left(\frac{\mu_{1,k}^2}{2\rho_{1,k}}\right)\right]^{-1},$$

$$\mu_{1,k} = \frac{\rho_{1,k}}{r_\mathbf{b}}\mathbf{e}_k^T\mathbf{G}_k, \qquad \mu_{0,k} = 0, \qquad \rho_{1,k} = \frac{r_\mathbf{a} r_\mathbf{b}}{r_\mathbf{b} + r_\mathbf{a}\mathbf{G}_k^T\mathbf{G}_k}, \qquad \rho_{0,k} = 0,$$

$$\mathbf{e}_k = \mathbf{w} - (\mathbf{G}\mathbf{x} - \mathbf{G}_k\mathbf{x}_k).$$

Comme dans le chapitre précédent (section 3.3.3.1), nous forçons $\mu_{1,k}$ à zéro s'il est négatif. Si cela n'est pas fait, l'algorithme risque d'estimer un très grand nombre de motifs à l'endroit où le spectre est négatif, ce qui aura pour conséquence de faire tendre K vers $K_{\max}$. Cela n'est pas acceptable car si le nombre maximal de motifs est atteint, cela signifie qu'il est trop faible. Or, ne pas forcer $\mu_{1,k}$ à zéro fera toujours tendre K vers $K_{\max}$.

4.2.3.3 Loi a posteriori de la largeur $\mathbf{s}_k$

La loi a posteriori de $\mathbf{s}_k$ s'écrit :

$$p(\mathbf{s}_k|\mathbf{w},\mathbf{c},\mathbf{x},\mathbf{s}_{-k},\lambda,r_{\mathbf{a}},r_{\mathbf{b}}) \propto \exp\left(-\frac{1}{2r_{\mathbf{b}}}||\mathbf{w}-\mathbf{G}\mathbf{x}||^2\right)\mathbb{1}_{\mathbb{R}^+}(\mathbf{s}_k).$$

où $\mathbf{s}_k$ intervient implicitement dans $\mathbf{G}$. On peut prouver l'intégrabilité de cette loi a posteriori conditionnelle de la même façon que dans le chapitre précédent.

4.2.3.4 Loi a posteriori des autres variables

Elles sont similaires au lois établies dans le chapitre précédent :

$$\lambda \sim \mathcal{B}e\big(K+1, 2K_{\max}-K+1\big), \tag{4.12}$$

$$r_{\mathbf{a}} \sim \mathcal{IG}\left(\frac{K}{2}+\alpha_{\mathbf{a}}, \frac{\mathbf{x}_n^T\mathbf{x}_n}{2}+\beta_{\mathbf{a}}\right), \tag{4.13}$$

$$r_{\mathbf{b}} \sim \mathcal{IG}\left(\frac{N}{2}, \frac{||\mathbf{w}-\mathbf{G}\mathbf{x}||^2}{2}\right). \tag{4.14}$$

4.3 Implémentation de l'échantillonneur de Gibbs

Pour les mêmes raisons que dans le chapitre précédent, nous utilisons un échantillonneur de Gibbs à balayage aléatoire et l'initialisation est laissée à l'utilisateur (voir section 3.4).

L'échantillonnage de $\mathbf{x}_k$ et des autres hyperparamètres est similaire à celui du chapitre précédent. De même, nous utilisons un algorithme d'acceptation-rejet à marche aléatoire pour simuler chaque composante du vecteur $\mathbf{s}$.

En revanche, la loi a posteriori de $\mathbf{c}_k$ n'est pas directement échantillonnable.

Dans le cas où le motif est présent ($\mathbf{q}_k = 1$), nous utilisons un algorithme de Metropolis-Hastings à marche aléatoire dont la loi candidate est une gaussienne centrée sur $\mathbf{c}_k^{(i-1)}$, position du motif à l'itération précédente, et de variance $r_{\mathbf{c}}$ choisie au préalable. De plus, la loi est bornée puisque $\mathbf{c}_k \in [1, N]$. De même qu'avec $\mathbf{s}_k$, le taux d'acceptation

$$\alpha = \min\left\{1, \frac{p(\widetilde{\mathbf{c}}_k)}{p(\mathbf{c}_k^{(i-1)})}\right\}$$

peut s'écrire, afin d'éviter les erreurs de débordement de précision :

$$-2r_{\mathbf{b}}\ln(u) > ||\mathbf{w}-\widetilde{\mathbf{G}}\mathbf{x}||^2 - ||\mathbf{w}-\mathbf{G}^{(i-1)}\mathbf{x}||^2$$

où $\widetilde{\mathbf{G}}$ correspond à la matrice $\mathbf{G}$ calculée avec $\widetilde{\mathbf{c}} = \{\mathbf{c}_1,\ldots,\widetilde{\mathbf{c}}_k,\ldots,\mathbf{c}_{K_{\max}}\}$ et $\mathbf{G}^{(i-1)}$ est la matrice $\mathbf{G}$ à l'itération précédente. L'utilisation d'un algorithme de Metropolis-Hastings à marche aléatoire est motivé par le fait que, puisque le motif est présent dans le spectre, alors il estime un motif du spectre. L'objectif est donc de définir au mieux sa position : de petites perturbations autour de la valeur courante permettent d'affiner l'estimation.

Si au contraire le motif est absent du spectre ($\mathbf{q}_k = 0$), nous ne disposons d'aucune connaissance pour déterminer l'emplacement d'un motif à estimer. C'est pourquoi nous utilisons un algorithme de Metropolis-Hastings dont la loi candidate est uniforme : $\mathcal{U}_{[1,N]}$. L'intérêt de mettre à jour la position d'un motif absent est d'explorer l'espace entier, permettant ainsi d'estimer un motif du signal qui n'aurait pas encore été retrouvé.

4.4 Méthode de ré-indexage

Comme dans le chapitre précédent, les hyperparamètres du problème sont estimés au sens de l'EAP. L'estimation des raies du spectre peut être calculée au sens de l'EAP seulement si la chaîne de Markov le permet. Or, cela n'est possible que si le signal $\mathbf{w}$ est peu bruité et que les raies sont bien séparées les unes des autres. Dans le cas contraire, le problème de permutation d'indices apparaît inévitablement et l'estimation du signal impulsionnel devient délicate (voir par exemple les études de Celeux [10], Holmes *et al.* [55], Richardson et Green [91] ou Stephens [104]). Ce type de problème est très courant mais très difficile à résoudre (surtout si le modèle est à dimension variable), c'est d'ailleurs encore un problème ouvert [10].

Considérons l'exemple académique d'un signal de $N = 500$ points, très bruité (RSB ≈ 6 dB) et possédant $K = 3$ raies de même largeur ($\mathbf{s} = \{6, 6, 6\}$) dont les positions et amplitudes sont reportées ci-dessous (cf. figure 4.3) :

$$\mathbf{c}_1 = 100, \qquad \mathbf{c}_2 = 150, \qquad \mathbf{c}_3 = 300,$$
$$\mathbf{a}_1 = 0,1, \qquad \mathbf{a}_2 = 0,1, \qquad \mathbf{a}_3 = 1.$$

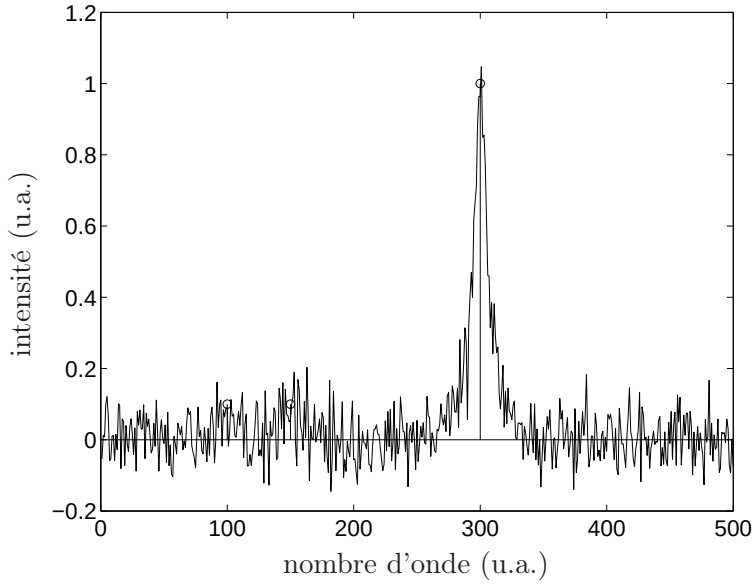


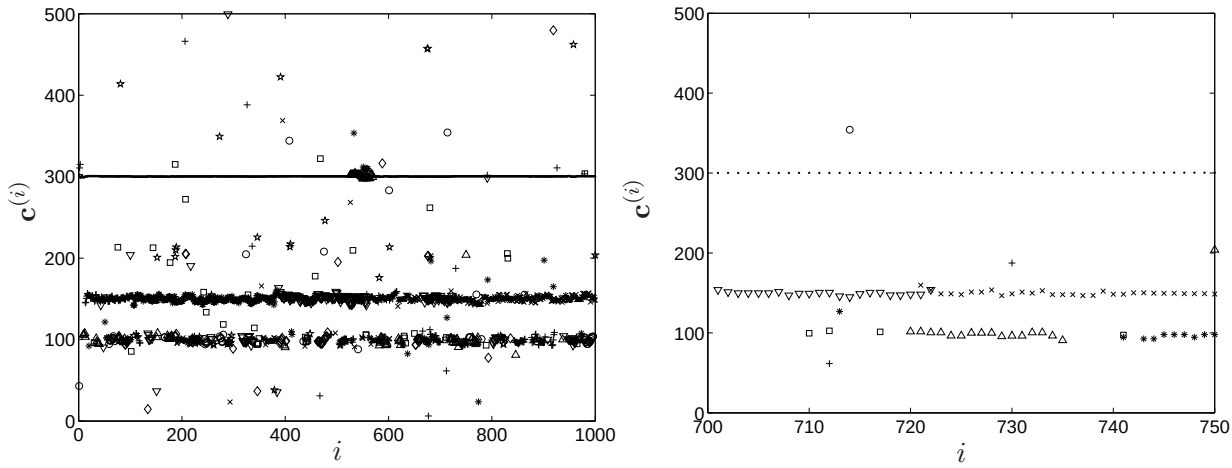
FIG. 4.3 – Exemple académique de spectre posant des problèmes de permutation d'indices.

Nous avons voulu estimer les raies de ce spectre avec la méthode proposée en considérant un nombre maximal de $K_{\max} = 10$ motifs. La figure 4.4(a) représente la chaîne de Markov de longueur $I = 1000$ des positions $\mathbf{c}$ des motifs. Chaque position est représentée par un symbole particulier.

Les trois raies du spectre sont bien retrouvées, mais, comme on le voit sur la figure 4.4(b), les raies situées en 100 et 150 ne sont pas estimées par le même motif à chaque itération. Il y a donc, au cours des itérations, des permutations entre les motifs. En fait, la permutation entre deux motifs revient à une permutation des indices k de ces motifs. De plus, il y a également, au cours des itérations, apparition et disparition de motifs (cf. figure 4.4(b)).

Le problème de permutation d'indices est dû à deux phénomènes.

D'une part, s'il n'y a pas d'information a priori qui permette de distinguer les variables des motifs entre elles, alors la valeur de la loi a posteriori est la même pour toutes les permutations sur les indices k de ces variables : on dit que la loi a posteriori est symétrique par rapport aux variables. Dans



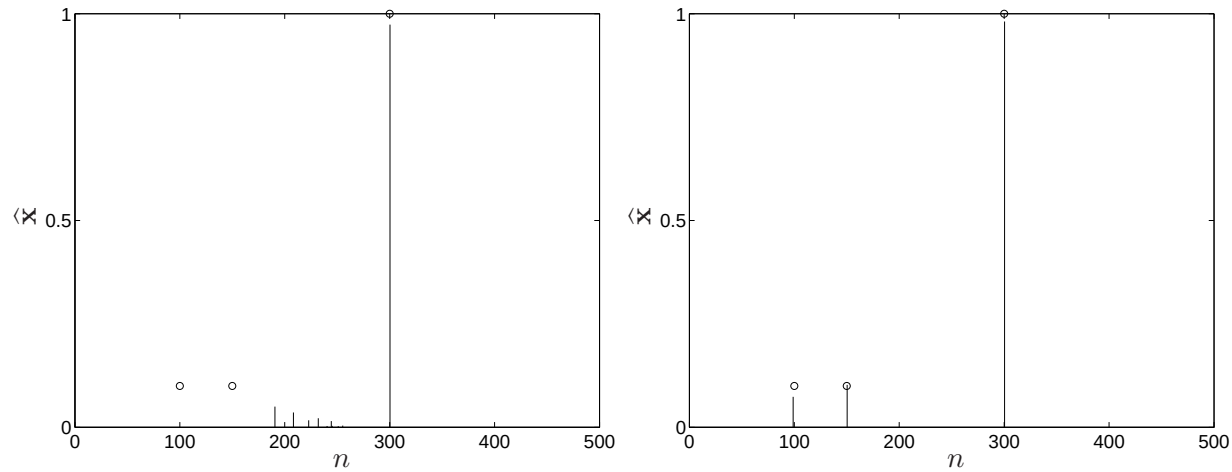
(a) Chaîne de Markov de $\mathbf{c}$ (chaîne entière). (b) Chaîne de Markov de $\mathbf{c}$ (zoom).

FIG. 4.4 – Chaîne de Markov de $\mathbf{c}$.

notre application, cela s'illustre par le fait que deux motifs peuvent être intervertis, c'est-à-dire qu'une permutation entre $\boldsymbol{\theta}_k$ et $\boldsymbol{\theta}_{k'}$ $(k \neq k')$ ne change pas la valeur de la loi a posteriori $p(\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_K|\mathbf{y})$. Il est donc impossible de distinguer les deux motifs, et donc de les empêcher de permuter.

D'autre part, la méthode de simulation n'empêche pas non plus les permutations puisqu'elle peut explorer les $k!$ possibilités de permutation à chaque itération : la permutation des indices est d'ailleurs une preuve de mélange correct de la chaîne de Markov [104, p. 30].

En cas de permutation d'indices, il est clair que l'estimateur au sens de l'EAP est totalement inefficace. En effet, l'estimation au sens de l'EAP à partir de la chaîne de Markov considérée est représentée figure 4.5(a) : elle est inexploitable sauf, évidemment, pour le grand pic. C'est pourquoi

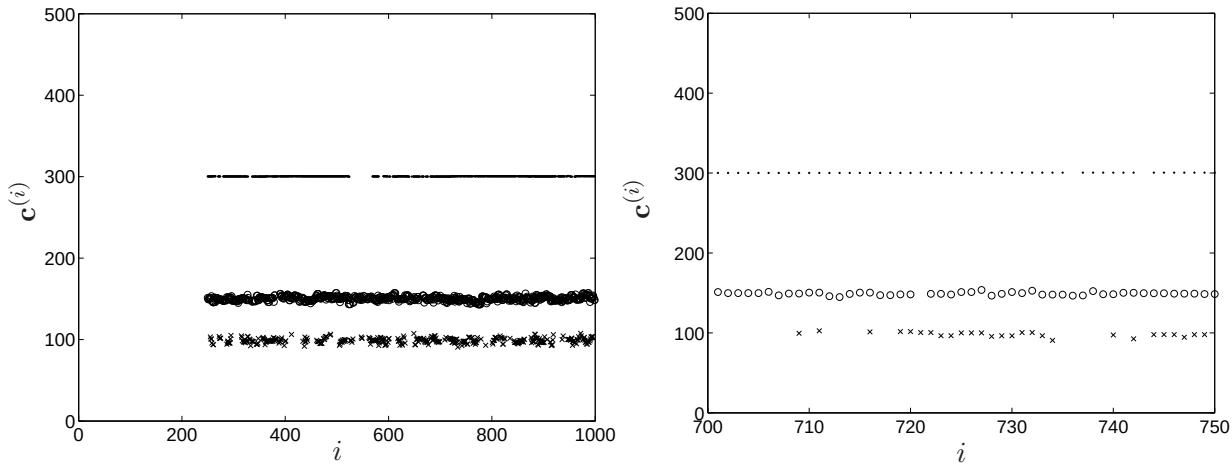


(a) Sans ré-indexage. (b) Avec ré-indexage.

FIG. 4.5 – Estimation du signal impulsionnel. Les ∘ indiquent les raies réelles.

il est indispensable de ré-indexer les motifs avant de pouvoir calculer l'estimateur au sens de l'EAP. L'estimation obtenue grâce à la méthode de ré-indexage proposée à la fin de cette section est représentée figure 4.5(b) : le résultat est maintenant satisfaisant. Les figures 4.6(a) et 4.6(b) représentent la chaîne de Markov après ré-indexage.

Cette section a pour but de présenter et discuter de plusieurs méthodes de ré-indexage. Par souci de



(a) Chaîne de Markov de $\mathbf{c}$ (chaîne entière). (b) Chaîne de Markov de $\mathbf{c}$ (zoom).

FIG. 4.6 – Chaîne de Markov de $\mathbf{c}$ après ré-indexage.

clarté, on suppose dans la suite que la période de transition est négligeable ou qu'elle a été supprimée, et donc que la chaîne de Markov considérée évolue de $i = 1$ à $i = I$.

Dans un premier temps, nous présentons deux méthodes plus ou moins intuitives et basées sur un raisonnement simple. Puis nous détaillons les méthodes qui ont été proposées dans le cadre de la décomposition de lois, en reprenant la classification de Dias et Wedel [30] :

- contrainte d'identifiabilité [31, 91] ;
- méthode de Celeux [11] ;
- méthode de Stephens [104, 105, 106] ;
- méthode de Celeux, Hurn et Robert [14].

Les trois dernières méthodes sont construites sur le même schéma : elles cherchent à ré-indexer chaque motif de manière à diminuer un coût. Les différences principales entre ces méthodes résident dans l'expression du coût, dans la recherche des indices optimaux et dans l'initialisation. En conservant ce schéma, nous proposons une méthode originale pour le calcul de l'estimateur $(\widehat{\mathbf{c}}, \widehat{\mathbf{x}}, \widehat{\mathbf{s}})$ qui permet d'obtenir un très bon résultat pour notre application.

4.4.1 Méthode « en deux temps »

Une méthode très simple consiste à lancer une première fois l'algorithme afin de calculer une estimation du nombre de motifs $\widehat{K}$. Un estimateur traditionnel est celui du MAP marginal (MMAP) [48] qui revient à choisir la valeur de K qui revient le plus souvent dans la chaîne de Markov : dans l'exemple considéré plus haut, on aurait $\widehat{K}^{\mathrm{MMAP}} = 3$. Puis, l'algorithme est lancé une seconde fois en fixant le nombre de motifs connus, c'est-à-dire en supposant que $K_{\max} = \widehat{K}^{\mathrm{MMAP}}$ et que les amplitudes des motifs sont distribuées suivant une simple gaussienne à support positif (et pas un processus Bernoulli-gaussien à support positif), puisqu'alors les $K_{\max}$ motifs sont sensés exister.

Cependant, il n'est pas garantit que cette méthode d'estimation empêche la permutation de deux motifs. De plus, il serait dommage d'utiliser cette méthode d'estimation car l'information contenue dans la première chaîne de Markov contient déjà toute l'information utile pour calculer une estimation du signal impulsionnel. Générer une seconde chaîne de Markov s'avère alors redondant.

4.4.2 Méthode « à histogramme »

Une autre solution[1] consiste tout d'abord à estimer le nombre de motifs $\widehat{K}$ (par exemple au sens du MMAP). On peut alors tracer l'histogramme de chaque quantité (position, amplitude et paramètres

[1] Suite à une discussion avec Manuel Davy.

de forme) en considérant seulement les itérations pour lesquelles le nombre de motifs est égal à $\widehat{K}$. Les $\widehat{K}$ plus grands pics de chaque histogramme correspondent à la quantité estimée.

Cependant, la largeur des barres de l'histogramme est importante et peut influencer considérablement l'estimation. En effet, des barres trop larges diminuent la précision et des barres trop fines ne permettent pas de faire ressortir les pics de l'histogramme. De plus, cette méthode s'avère inefficace pour l'estimation des amplitudes car il est très délicat de déterminer les maxima de l'histogramme généré à partir de la chaîne de Markov des amplitudes. Quand bien même il serait possible d'estimer les positions, les amplitudes et les largeurs, comment recombiner les triplets position/amplitude/largeur ?

4.4.3 Contrainte d'identifiabilité

L'une des premières véritables méthode de ré-indexage a été d'imposer une *contrainte d'identifiabilité* sur les paramètres qui ne soit vérifiée que pour une seule permutation des indices [31, 91].

Une telle contrainte impose une dissymétrie sur la loi a priori et donc sur la loi a posteriori : celle-ci n'est alors plus invariante par permutation d'indices. Une contrainte d'identifiabilité serait par exemple d'imposer un ordre sur les positions des motifs :

$$\mathbf{c}_1 < \mathbf{c}_2 < \cdots < \mathbf{c}_K.$$

Cet ordre est établi de manière arbitraire et n'est pas le reflet d'une réalité physique : c'est pour cela qu'on parle parfois de contrainte d'identifiabilité *artificielle* [55].

Notons qu'en déconvolution impulsionnelle, on peut considérer qu'une contrainte d'identifiabilité est imposée implicitement sur les positions puisque les pics ne peuvent pas permuter les uns les autres.

Afin de respecter la contrainte d'identifiabilité, les indices des paramètres sont permutés à chaque itération de l'échantillonneur de Gibbs. Cette opération peut être effectuée une fois que la chaîne de Markov a convergé [104, proposition 3.1] ou au sein de la méthode de simulation, par exemple en générant les positions $\mathbf{c}_k$ entre $\mathbf{c}_{k-1}$ et $\mathbf{c}_{k+1}$ (avec la convention que $\mathbf{c}_{-1} = 1$ et $\mathbf{c}_{K+1} = N$). En fait, cela revient à poser $\mathbf{c}_k \sim \mathcal{U}_{[\mathbf{c}_{k-1}, \mathbf{c}_{k+1}]}$ comme a priori. La position est alors estimée au sens de l'EAP en l'approximant par sa moyenne [104].

Cependant, dans notre cas, il est difficile d'imposer une contrainte d'identifiabilité à la fin des itérations de l'échantillonneur de Gibbs, car le nombre de motifs présents dans le spectre peut varier.

Par ailleurs, certains choix de contraintes d'identifiabilité peuvent être inefficaces pour supprimer la symétrie de la loi a posteriori, et le problème persiste. En effet, il est maintenant reconnu qu'imposer une contrainte d'identifiabilité peut tout de même conduire à des estimations non satisfaisantes : des études empiriques ont montré que cette méthode de ré-indexage peut résulter en un biais sur l'estimation [14, 30, 105]. En particulier, imposer une contrainte d'ordre sur les positions ne garantit pas que le loi a posteriori marginale de chaque position ne sera pas multimodale [55, 106].

Enfin, si une contrainte d'identifiabilité est imposée pendant la simulation, elle risque d'augmenter dramatiquement le temps de convergence de l'algorithme. Considérons par exemple le cas illustré figure 4.7 sur lequel est représenté l'estimation du signal impulsionnel à une itération particulière de la chaîne de Markov. Ici, le nombre de motifs est $K = 3$ et on a choisi également $K_{\max} = 3$. Ce choix n'est certainement pas judicieux, mais il permet de mettre en évidence la lenteur de convergence lorsqu'une contrainte d'identifiabilité est imposée.

Seules les deux premières raies (a et b) sont correctement estimées par les motifs $k = 1$ et $k = 3$. Comme les positions des motifs sont contraintes de respecter l'ordre $\mathbf{c}_1 < \mathbf{c}_2 < \mathbf{c}_3$, la seule solution pour obtenir une estimation correcte du signal est que le motif $k = 3$ se déplace pour estimer la raie c, et qu'il soit remplacé par le motif $k = 2$.

Mais cette solution risque de prendre énormément de temps car l'échantillonneur de Gibbs doit passer par une solution de faible probabilité pour arriver dans une solution de plus forte probabilité,

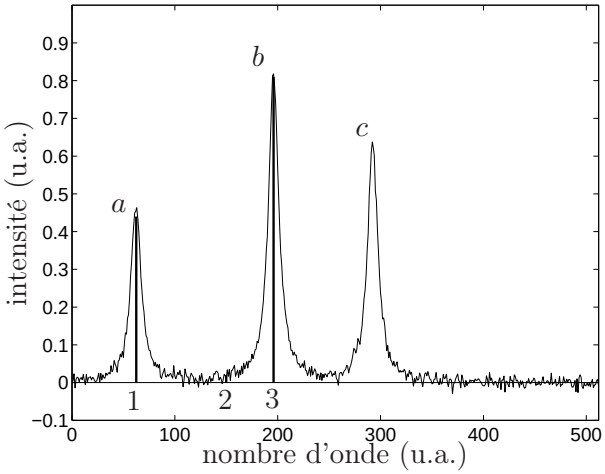


FIG. 4.7 – Exemple de convergence lente en imposant une contrainte d'identifiabilité.

et, comme on l'a déjà vu, l'échantillonneur de Gibbs à du mal à franchir des « vallées » de faible probabilité [14]. Au contraire, si aucun ordre n'est imposé, la raie c sera sans problème estimée par le motif $k = 2$ (pourvu bien sûr qu'il puisse se déplacer suffisamment vite).

4.4.4 Méthode de Celeux

Celeux [11, 14] propose un algorithme de ré-indexage qui consiste, à chaque itération, à permuter les indices de manière à minimiser pour chaque paramètre la distance entre la valeur courante et la moyenne sur les itérations précédentes. Cette approche a l'avantage de ne pas nécessiter le stockage complet de la chaîne de Markov car il peut être effectué en ligne.

On considère donc la chaîne de Markov $\{\boldsymbol{\theta}^{(i)}\}$ qui représente la séquence des quantités à estimer $\boldsymbol{\theta} = \{\mathbf{c}, \mathbf{a}, \mathbf{s}\}$. Le vecteur $\boldsymbol{\theta}$ est donc de taille $3K_{\max}$. Dans la suite, on note $\boldsymbol{\theta}_l$ ($l \in \{1, \ldots, 3K_{\max}\}$) un élément de $\boldsymbol{\theta}$.

L'algorithme est tout d'abord initialisé à partir des I_1 premiers vecteurs $\boldsymbol{\theta}^{(1)}, \ldots, \boldsymbol{\theta}^{(I_1)}$ simulés (on rappelle qu'il n'y a pas de période de transition). La valeur de I_1 est typiquement d'une centaine d'échantillons. Bien que cette valeur influe peu sur l'estimation, elle doit être suffisamment grande pour s'assurer que l'estimation initiale est une approximation correcte des moyennes a posteriori, et suffisamment petite pour éviter les permutations d'indices.

Nous pensons que ce point constitue l'inconvénient principal de l'approche. En effet, il n'est pas toujours évident de sélectionner dans la chaîne de Markov une centaine d'échantillons consécutifs pour lesquels il n'y a pas de permutation d'indices, en particulier lorsque les échantillons sont de petites amplitudes (voir l'exemple illustré par la figure 4.4).

Les valeurs initiales des moyennes $\bar{\boldsymbol{\theta}}_l^{[0]}$ et des variances $\check{\boldsymbol{\theta}}_l^{[0]}$ de chaque paramètre sont alors calculées grâce aux équations suivantes :

$$\bar{\boldsymbol{\theta}}_l^{[0]} = \frac{1}{I_1} \sum_{i=1}^{I_1} \boldsymbol{\theta}_l^{(i)},$$

$$\check{\boldsymbol{\theta}}_l^{[0]} = \frac{1}{I_1} \sum_{i=1}^{I_1} \left(\boldsymbol{\theta}_l^{(i)} - \bar{\boldsymbol{\theta}}_l^{[0]} \right)^2,$$

pour tout $l \in \{1, \ldots, 3K_{\max}\}$. On peut supposer que lorsque le nombre de motifs est variable, alors les calculs ne considèrent que les motifs présents.

La procédure de ré-indexage débute alors à partir de l'itération $I_1 + 1$. Ainsi, pour toute itération $i = I_1 + j$ $(j \geq 1)$, l'algorithme de ré-indexage suit le schéma suivant :

1. détermination de la permutation σ_j qui minimise

$$\sigma_j = \arg\min_{\sigma_l} \frac{\left(\sigma_l\left(\boldsymbol{\theta}_l^{(I_1+j)}\right) - \bar{\boldsymbol{\theta}}_l^{[j-1]}\right)^2}{\check{\boldsymbol{\theta}}_l^{[j-1]}},$$

   Le vecteur $\boldsymbol{\theta}^{(I_1+j)}$ est alors redéfini en prenant en compte la permutation des indices ;

2. mise à jour des moyennes et variance de chaque paramètre :

$$\bar{\boldsymbol{\theta}}_l^{[j]} = \frac{1}{I_1 + j} \sum_{i=1}^{I_1+j} \boldsymbol{\theta}_l^{(i)}, \tag{4.15}$$

$$\check{\boldsymbol{\theta}}_l^{[j]} = \frac{1}{I_1 + j} \sum_{i=1}^{I_1+j} \left(\boldsymbol{\theta}_l^{(i)} - \bar{\boldsymbol{\theta}}_l^{[j]}\right)^2. \tag{4.16}$$

La première étape de l'algorithme consiste à trouver la permutation qui minimise la distance normalisée

$$\left(\sigma_l(\boldsymbol{\theta}_l^{(I_1+j)}) - \bar{\boldsymbol{\theta}}_l^{[j-1]}\right)^2 / \check{\boldsymbol{\theta}}_j^{[j-1]}.$$

Or, cette recherche n'est pas évidente et une recherche exhaustive parmi les $K_{\max}!$ combinaisons est à écarter à cause du coût de calcul. Une solution consiste à tester une permutation au hasard à chaque itération de l'algorithme et à l'accepter si elle réduit la distance normalisée. Malheureusement, cela diminue la vitesse de convergence.

4.4.5 Méthode de Stephens

Stephens [104, 105, 106] propose, dans le cadre de l'estimation des paramètres d'un mélange de loi, un algorithme de ré-indexage dont l'idée est de minimiser une fonction-coût $\mathcal{L}(a, \boldsymbol{\theta})$ correspondant au coût d'une action a sur les paramètres $\boldsymbol{\theta}$. Il faut alors choisir une action $\widehat{a}$ qui minimise le risque suivant, défini comme l'espérance de la fonction-coût :

$$\mathcal{R}(a) = \mathbb{E}_{\boldsymbol{\theta}}[\mathcal{L}(a, \boldsymbol{\theta})].$$

En outre, la fonction coût doit être invariante aux permutations d'indices. Stephens choisit d'imposer cette invariance en restreignant les fonctions-coût à la forme :

$$\mathcal{L}(a, \boldsymbol{\theta}) = \min_{\nu} \mathcal{L}_0(a, \nu(\boldsymbol{\theta}))$$

où ν est une fonction de permutation. Des exemples de fonction-coût peuvent être trouvés dans [104, 105, 106].

La valeur optimale de a est :

$$a^* = \arg\min_a \mathcal{R}(a) = \arg\min_a \int \mathcal{L}(a, \boldsymbol{\theta}) p(\boldsymbol{\theta}) d\boldsymbol{\theta}.$$

Comme cette intégrale ne peut pas être calculée exactement, le risque $\mathcal{R}(a)$ est approximé par le risque de Monte Carlo

$$\widetilde{\mathcal{R}}(a) = \frac{1}{I} \sum_{i=1}^{I} \mathcal{L}(a, \boldsymbol{\theta}^{(i)}) = \frac{1}{I} \sum_{i=1}^{I} \min_{\nu_i} \mathcal{L}_0(a, \nu_i(\boldsymbol{\theta}^{(i)})) = \min_{\nu_1,\ldots,\nu_I} \sum_{i=1}^{I} \mathcal{L}_0(a, \nu_i(\boldsymbol{\theta}^{(i)})),$$

qui est minimisé par l'algorithme suivant ([104, algorithme 3.4] [106, algorithme 4.1]) :

1. choisir les valeurs initiales de $\nu_1, \ldots, \nu_I$ (on peut poser par exemple la permutation identité) ;
2. effectuer jusqu'à la convergence :
   (a) choisir $\widehat{a}$ qui minimise
   $$\sum_{i=1}^{I} \mathcal{L}_0(a, \nu_i(\boldsymbol{\theta}^{(i)})),$$
   (b) pour $i = 1, \ldots, I$, choisir la permutation ν_i qui minimise
   $$\mathcal{L}_0(\widehat{a}, \nu_i(\boldsymbol{\theta}^{(i)})).$$

Le coût de calcul de l'algorithme dépend de la complexité de $\mathcal{L}_0$. L'algorithme converge vers un point fixe puisque chaque itération diminue le coût $\widetilde{\mathcal{R}}(a)$ et qu'il y a un nombre fini de permutation $\nu_1, \ldots, \nu_I$. Cependant, la convergence vers le minimum global n'est pas garantie. En fait, l'estimation obtenue dépend de la valeur initiale.

Comme pour la méthode de Celeux, une recherche exhaustive de la permutation qui minimise $\mathcal{L}_0(\widehat{a}, \nu_i(\boldsymbol{\theta}^{(i)}))$ n'est pas directe Toutefois, la recherche de la permutation optimale peut être traitée comme un problème de transport (*transportation problem*) [107], connu également comme un problème d'affectation linéaire (*assignment problem*), qui peut être résolu rapidement.

4.4.6 Méthode de Celeux, Hurn et Robert

Enfin, Celeux *et al.* [14] proposent d'utiliser une fonction-coût $\mathcal{L}(\boldsymbol{\theta}, \widehat{\boldsymbol{\theta}})$ invariante par permutation d'indices et de trouver l'estimateur de Bayes correspondant :

$$\widehat{\boldsymbol{\theta}}^* = \arg\min_{\widehat{\boldsymbol{\theta}}} \mathbb{E}_{\boldsymbol{\theta}}[\mathcal{L}(\boldsymbol{\theta}, \widehat{\boldsymbol{\theta}})].$$

Des exemples de fonctions-coût peuvent être trouvés dans [14]. La minimisation de cette équation n'est pas forcément explicite : on peut alors utiliser une approximation de l'espérance par la somme

$$\mathbb{E}_{\boldsymbol{\theta}}[\mathcal{L}(\boldsymbol{\theta}, \widehat{\boldsymbol{\theta}})] \approx \frac{1}{I} \sum_{i=1}^{I} \mathcal{L}(\boldsymbol{\theta}^{(i)}, \widehat{\boldsymbol{\theta}}).$$

L'utilisation de méthodes d'optimisation stochastiques (tel l'algorithme du recuit simulé [93]) permet de minimiser l'espérance, mais cela engendre un coût de calcul supplémentaire alors que la méthode de décomposition en motifs élémentaires est déjà assez lourde du fait de l'échantillonneur de Gibbs.

4.4.7 Méthode proposée

Parmi les méthodes de ré-indexage précédentes, nous avons vu que les méthodes intuitives sont inadéquates. En outre, imposer une contrainte d'identifiabilité souffre de plusieurs inconvénients, dont celui de ne pas supprimer la symétrie de la loi a posteriori. La méthode de Celeux est plus performante mais les étapes d'initialisation et de recherche de la permutation optimale ne contribuent malheureusement pas à diminuer la vitesse de convergence. Le choix de la fonction-coût et l'étape d'initialisation sont les deux inconvénients de la méthode de Stephens. Enfin, l'inconvénient principal de la méthode de Celeux, Hurn et Robert est son coût de calcul.

Par ailleurs, comme il a été noté dans [55, 106], la méthode de Celeux correspond en fait à une version en ligne de la méthode de Stephens dans laquelle la fonction-coût est :

$$\mathcal{L}_0(\boldsymbol{\theta}, \bar{\boldsymbol{\theta}}, \check{\boldsymbol{\theta}}) = -\ln \left[\prod_{l=1}^{3K_{\max}} \mathcal{N}(\boldsymbol{\theta}_l | \bar{\boldsymbol{\theta}}_l, \check{\boldsymbol{\theta}}_l) \right].$$

La minimisation de cette expression en $\bar{\boldsymbol{\theta}}$ et $\check{\boldsymbol{\theta}}$ donne les expressions explicites (4.15) et (4.16). Celeux propose alors une méthode en ligne qui à l'avantage de ne pas nécessiter de sauvegarder la totalité de la chaîne de Markov, alors que la méthode de Stephens est itérative et est utilisée une fois que la convergence de l'algorithme MCMC est atteinte. Néanmoins, l'initialisation de ces deux méthodes est semblable puisqu'elle consiste à choisir l'identité comme permutation initiale. Stephens préconise toutefois d'utiliser plusieurs valeurs initiales pour se soustraire de l'influence de l'initialisation [104].

Quoi qu'il en soit, l'inconvénient majeur de ces deux méthodes de ré-indexage réside dans le fait qu'elles sont difficilement adaptables au cas où le nombre de motifs est variable.

La méthode de ré-indexage que nous proposons[2] s'inspire de la méthode de Stephens, mais plusieurs contributions permettent d'en améliorer certains aspects ou d'en éviter les inconvénients. Nous cherchons donc à minimiser la fonction-coût suivante, qui est également celle utilisée par Celeux [11] :

$$\mathcal{L}_0(\mathbf{c},\mathbf{x},\mathbf{s},\bar{\mathbf{c}},\check{\mathbf{c}},\bar{\mathbf{x}},\check{\mathbf{x}},\bar{\mathbf{s}},\check{\mathbf{s}}) = -\ln\left[\prod_{k=1}^{K_{\max}} \mathcal{N}(\mathbf{c}_k|\bar{\mathbf{c}}_k,\check{\mathbf{c}}_k)\mathcal{N}(\mathbf{x}_k|\bar{\mathbf{x}}_k,\check{\mathbf{x}}_k)\mathcal{N}(\mathbf{s}_k|\bar{\mathbf{s}}_k,\check{\mathbf{s}}_k)\right].$$

Le choix d'utiliser des gaussiennes permet de simplifier les calculs : en effet, la minimisation de la fonction-coût conduit à des expressions explicites simples (voir plus loin). La méthode proposée se distingue de celle de Stephens par les trois points suivants :

- tout d'abord, une estimation approximative de la solution est obtenue à partir des maxima des histogrammes des quantités d'intérêt. Cette première estimation permet d'initialiser la méthode à une valeur plus proche de l'optimum global qu'une simple permutation identité ;
- par ailleurs, le ré-indexage n'est pas obtenu en permutant les motifs mais en tentant de les ré-indexer un par un de manière à minimiser la fonction-coût $\mathcal{L}_0$;
- enfin, la méthode proposée prend en compte le fait que le nombre de motifs présents dans le spectre peut varier.

Description de l'algorithme

Le nombre de motifs est tout d'abord estimé au sens du MMAP. La suite de la méthode est une procédure itérative sur $k = 1,\ldots,\widehat{K}^{\text{MMAP}}$.

La *première étape* consiste à créer l'histogramme des quantités simulées. L'histogramme est donc en trois dimensions (positions, amplitudes et largeurs) ; dans le cas de motifs plus complexes, il faut ajouter autant de dimensions qu'il y a de paramètres de forme. C'est d'ailleurs l'inconvénient majeur de ce type d'approche, puisque tous les logiciels ne permettent pas d'utiliser des tableaux de plus de deux dimensions, et que le stockage d'un tel tableau est très gourmand en mémoire. Malgré tout, l'histogramme permet d'obtenir une approximation proche de l'optimum et n'est utilisé que pour l'initialisation. C'est pourquoi la taille des cellules n'est pas très importante : chaque dimension est composée de cent cellules, sauf celle des positions pour laquelle on utilise N cellules.

Le maximum de l'histogramme est ensuite déterminé et ses coordonnées $\bar{\boldsymbol{\theta}}_k = (\bar{\mathbf{c}}_k,\bar{\mathbf{x}}_k,\bar{\mathbf{s}}_k)$ constituent une première approximation du résultat. On initialise également les variables $\check{\mathbf{c}}_k$, $\check{\mathbf{x}}_k$ et $\check{\mathbf{s}}_k$ à une valeur très grande ; ces variables correspondent aux variances des échantillons simulés. La suite de la méthode consiste à affiner la valeur des moyennes qui constitueront les estimées.

La *deuxième étape* de l'algorithme correspond à l'étape 2(a) de la méthode de Stephens. Elle consiste donc, pour $i = 1,\ldots,I$, à minimiser la fonction-coût $\mathcal{L}_0$ en $(\mathbf{c}^{(i)},\mathbf{x}^{(i)},\mathbf{s}^{(i)})$, ce qui revient à minimiser :

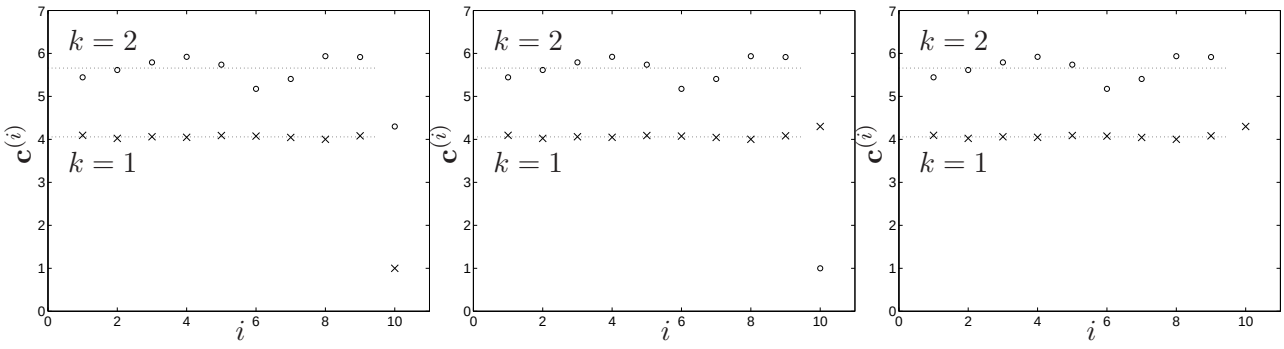
$$\mathcal{L}_0(\mathbf{c}^{(i)},\mathbf{x}^{(i)},\mathbf{s}^{(i)},\bar{\mathbf{c}},\check{\mathbf{c}},\bar{\mathbf{x}},\check{\mathbf{x}},\bar{\mathbf{s}},\check{\mathbf{s}}) = \sum_{k=1}^{K}\left[\frac{(\mathbf{c}_k^{(i)}-\bar{\mathbf{c}}_k)^2}{\check{\mathbf{c}}_k} + \frac{(\mathbf{x}_k^{(i)}-\bar{\mathbf{x}}_k)^2}{\check{\mathbf{x}}_k} + \frac{(\mathbf{s}_k^{(i)}-\bar{\mathbf{s}}_k)^2}{\check{\mathbf{s}}_k}\right].$$

[2] Cette méthode a été établie suite à une discussion avec El-Hadi Djermoune.

La méthode de Stephens consiste à chercher la permutation qui minimise ce terme, c'est-à-dire à arranger les indices des motifs pour toute itération. Or, la méthode proposée ré-indexe les motifs les uns après les autres (rappelons que l'algorithme est itératif pour $k = 1, \ldots, \widehat{K}$) et pas tous ensemble à l'aide d'une permutation. Aussi, nous proposons une alternative à la méthode de Stephens en cherchant le motif l qui minimise la distance

$$D(\boldsymbol{\theta}_l^{(i)}, \bar{\boldsymbol{\theta}}_k) = \frac{(\mathbf{c}_l^{(i)} - \bar{\mathbf{c}}_k)^2}{\check{\mathbf{c}}_k} + \frac{(\mathbf{x}_l^{(i)} - \bar{\mathbf{x}}_k)^2}{\check{\mathbf{x}}_k} + \frac{(\mathbf{s}_l^{(i)} - \bar{\mathbf{s}}_k)^2}{\check{\mathbf{s}}_k}.$$

Cela revient à chercher, pour chaque itération de la chaîne de Markov, le motif l qui est le plus proche de la moyenne $\bar{\boldsymbol{\theta}}_k = (\bar{\mathbf{c}}_k, \bar{\mathbf{x}}_k, \bar{\mathbf{s}}_k)$. La recherche doit toutefois s'effectuer sur les motifs présents dans le spectre, c'est-à-dire tels que $\mathbf{q}_l^{(i)} = 1$. Par ailleurs, si le motif le plus près de la moyenne en est tout de même trop éloigné, il convient de ne pas le sélectionner. En effet, considérons l'exemple illustré figure 4.8. La figure 4.8(a) représente la chaîne de Markov des positions de deux motifs, que



(a) Indices initiaux.

(b) Ré-indexage sans prendre en compte Δ.

(c) Ré-indexage en prenant en compte Δ.

FIG. 4.8 – Exemple de ré-indexage avec la méthode proposée.

nous dénoterons $\times$ et $\circ$ pour plus de commodité. Les estimations des deux motifs à l'itération 9 sont représentées par des pointillés ($k = 1$ et $k = 2$). Pour une raison quelconque, le motif $\mathbf{c}_\times^{(10)}$ est très éloigné de la valeur de l'estimation calculée jusqu'à cette itération.

Comme nous l'avons dit, la méthode proposée effectue une boucle sur le nombre de motifs estimés $\widehat{K}$. L'ordre des motifs est sélectionné dans l'ordre décroissant des amplitudes des barres de l'histogramme. C'est la raison pour laquelle les motifs dont la chaîne de Markov a la plus faible variance sont susceptibles d'être estimés avant ceux dont la chaîne de Markov a la plus grande variance. Dans notre exemple, le premier motif estimé sera donc le motif $\times$. Dès lors, à l'itération $i = 10$, la position la plus proche de la moyenne $k = 1$ est en fait $\mathbf{c}_\circ^{(10)}$: cette quantité est donc ré-indexée, comme c'est illustré sur la figure 4.8(b). Par conséquent, la position $\mathbf{c}_\times^{(10)}$ sera forcément attribuée à la moyenne $k = 2$ (figure 4.8(b)). Ce ré-indexage n'est évidemment pas satisfaisant. En outre, il ne correspond pas à la meilleure permutation.

Aussi, nous pensons que le motif le plus proche de la moyenne ne doit être sélectionné que s'il n'en est pas trop éloigné : il faut que la condition $D(\boldsymbol{\theta}_l^{(i)}, \bar{\boldsymbol{\theta}}_k) < \Delta^2$ soit vérifiée, où Δ est une constante que nous choisissons égale à cinq. Ce choix garantit que les motifs sélectionnés seront distants de la moyenne d'au maximum cinq fois l'écart-type des échantillons sélectionnés. Il est donc possible qu'aucun motif ne soit sélectionné à certaines itérations si le motif le plus proche de $\bar{\boldsymbol{\theta}}_k$ en est trop éloigné. C'est le cas dans l'exemple considéré où, à l'itération 10, aucun motif n'est affecté à la moyenne $k = 2$ puisque la position $\mathbf{c}_\times^{(10)}$ est trop éloignée de la moyenne (figure 4.8(c)). Ce comportement peut se rencontrer dans le cas de raies particulièrement difficiles à retrouver. En outre, cela permet de gérer le cas où le nombre de motifs est variable.

En résumé, on sélectionne le motif l tel que, pour chaque itération de la chaîne de Markov :

– ce motif soit le plus proche de la moyenne $\bar{\boldsymbol{\theta}}_k$;
– la distance à cette moyenne soit inférieure à une valeur Δ^2 ;
– le motif est présent dans le spectre ($\mathbf{q}_l^{(i)} = 1$).

En d'autres termes, on cherche $\boldsymbol{\theta}_l^{(i)} = \{\mathbf{c}_l^{(i)}, \mathbf{x}_l^{(i)}, \mathbf{s}_l^{(i)}\}$ qui vérifie :

$$\begin{cases} D(\boldsymbol{\theta}_l^{(i)}, \bar{\boldsymbol{\theta}}_k) \leq D(\boldsymbol{\theta}_{l'}^{(i)}, \bar{\boldsymbol{\theta}}_k) & \text{pour tout } l' \in \{1, \dots, K_{\max}\}, \\ D(\boldsymbol{\theta}_l^{(i)}, \bar{\boldsymbol{\theta}}_k) < \Delta^2, \\ \mathbf{q}_l^{(i)} = 1. \end{cases} \tag{4.17}$$

On suppose dans la suite que L motifs ont été sélectionnés et nous regroupons sous le vecteur $\mathbf{l}_k = \{\mathbf{l}_k^{(1)}, \dots, \mathbf{l}_k^{(L)}\}$ l'ensemble des indices de ces motifs.

La *troisième étape* de l'algorithme correspond à l'étape 2.(b) de la méthode de Stephens. Elle consiste à mettre à jour les paramètres $\bar{\boldsymbol{\theta}}_k$ et $\check{\boldsymbol{\theta}}_k$. Là encore, le cas gaussien permet d'obtenir des expressions explicites des minimiseurs. Ainsi, dans le cas des positions, on a :

$$\bar{\mathbf{c}}_k = \frac{1}{L} \sum_{j=1}^{L} \mathbf{c}_{\mathbf{l}_k^{(j)}}^{(j)}, \qquad \check{\mathbf{c}}_k = \frac{1}{L} \sum_{j=1}^{L} \left(\mathbf{c}_{\mathbf{l}_k^{(j)}}^{(j)} - \bar{\mathbf{c}}_k \right)^2,$$

On procède de même pour toutes les autres quantités.

Enfin, à partir des moyennes et variances mises à jour, on sélectionne à nouveau les motifs les plus proches, et on réitère ainsi de suite les deuxième et troisième étapes jusqu'à ce que la sélection $\mathbf{l}_k$ ne varie plus. Le motif estimé a alors pour paramètre $\bar{\boldsymbol{\theta}}_k$. L'algorithme reboucle jusqu'à $k = \widehat{K}$, après avoir mis à zéro les probabilités d'occurrence $\mathbf{q}_l^{(i)}$ des motifs sélectionnés, pour éviter qu'ils soient sélectionnés une nouvelle fois.

L'algorithme de la méthode de ré-indexage proposée est représenté figure 4.9.

En conclusion, la méthode de ré-indexage proposée minimise une fonction-coût comparable à celle utilisée par Stephens, mais la procédure de minimisation est différente. En effet, l'initialisation permet de démarrer l'algorithme à une valeur plus proche de l'optimum qu'une simple permutation identité ; malheureusement cette étape d'initialisation souffre d'un coût en mémoire très important. Par ailleurs, plutôt que de chercher pour chaque itération quelle est la permutation qui minimise la fonction-coût, nous préférons effectuer une boucle sur le nombre de motifs estimé $\widehat{K}$ en cherchant, à chaque itération, quel est le motif le plus proche de l'estimation actuelle. Enfin, en prenant en compte l'occurrence des motifs ainsi qu'une distance maximale à l'estimation, l'algorithme permet de prendre en compte le fait que le nombre de motifs varie.

Cette méthode ne bénéficie d'aucune preuve mathématique permettant de garantir son bon fonctionnement, mais toutes les expériences pratiques ont montré que l'estimation obtenue était satisfaisante. Cette méthode est également un peu longue, surtout sur les spectres réels pour lesquels le nombre de motifs et surtout le nombre d'itérations peuvent être important. Elle reste toutefois plus rapide qu'une recherche exhaustive de permutation.

4.5 Application sur des spectres simulés

4.5.1 Cas de raies identiques

Nous avons tout d'abord testé la méthode proposée dans ce chapitre sur le même spectre que celui présenté dans la section 3.6.1. Nous rappelons les caractéristiques du spectre dans la figure 4.10 et le tableau 4.2. L'objectif est de vérifier que la méthode proposée ici peut s'adapter au cas où les motifs sont identiques. Pour cela, nous utilisons la variante de l'approche qui considère que, pour tout $k \in \{1, \dots, K\}$, $\mathbf{s}_k = s$.

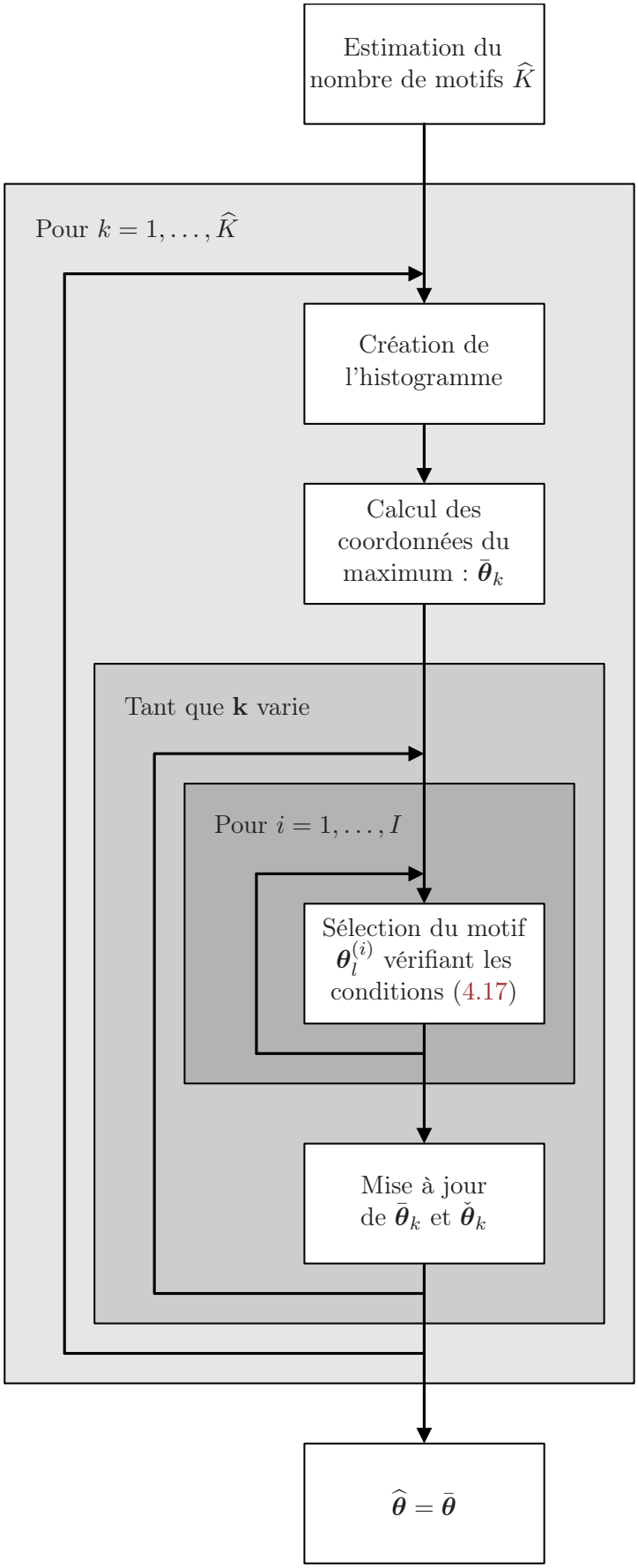


FIG. 4.9 – Algorithme proposé pour le ré-indexage.

La méthode a été initialisée avec les valeurs initiales suivantes :

$$K_{\max} = 30, \qquad r_{\mathbf{a}} = 2, \qquad s = 10, \qquad r_{\mathbf{b}} = 0{,}1,$$

et un spectre initial contenant 15 motifs répartis uniformément et d'amplitudes nulles.

Nous avons effectué $I = 4000$ itérations de l'échantillonneur de Gibbs ; les figures 4.11 et 4.12 représentent l'évolution des chaînes de $\mathbf{q}$, $\mathbf{a}$, K, $r_{\mathbf{a}}$, s et $r_{\mathbf{b}}$. La longueur de la période de transition a été choisie à $I_0 = 1000$ itérations. Le temps de calcul a été d'environ 238 secondes soit presque quatre fois moins que la méthode par déconvolution impulsionnelle myope.

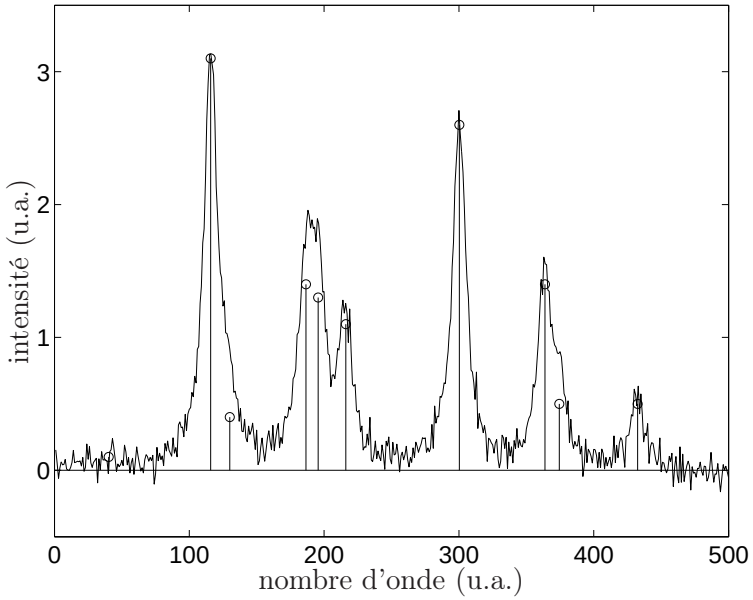


FIG. 4.10 – Spectre bruité simulé.

k	1	2	3	4	5	6	7	8	9	10
$\mathbf{c}^*$	40,2	115,9	130	186,5	195,6	216,1	300,3	363,8	374,4	432,5
$\mathbf{a}^*$	0,1	3,1	0,4	1,4	1,3	1,1	2,6	1,4	0,5	0,5

TAB. 4.2 – Positions et amplitudes des raies du spectre pur.

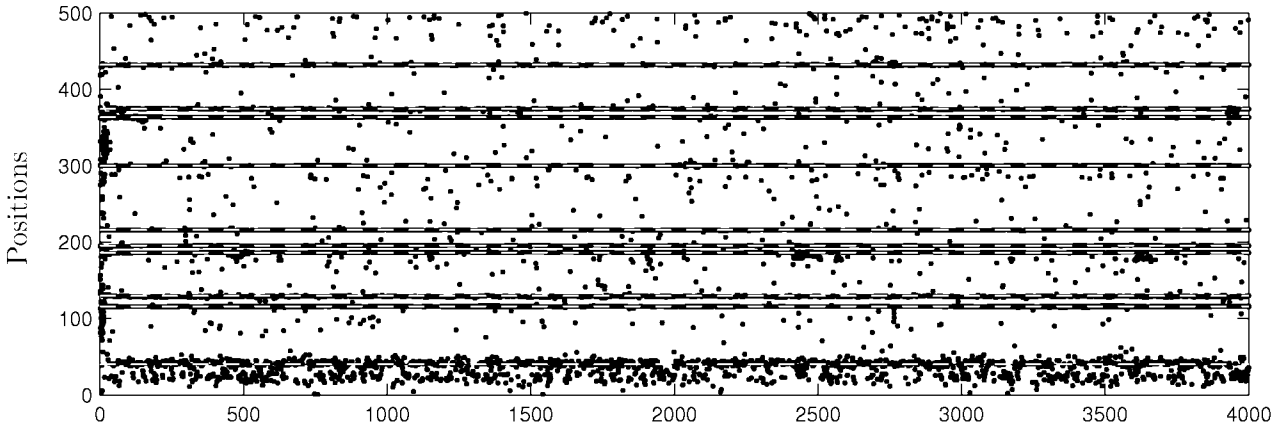
Estimation des hyperparamètres

Les hyperparamètres de la méthode ont été estimés au sens de l'EAP. Le tableau 4.3 donne les valeurs réelles et estimées des hyperparamètres, ainsi que l'écart-type et le biais de l'estimation.

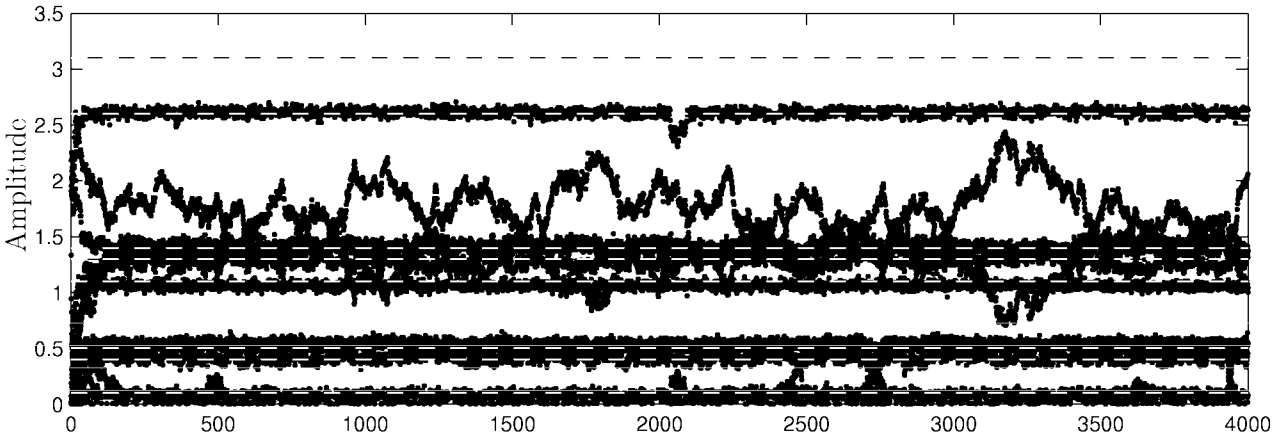
	valeur réelle	**estimation**	**écart-type**	**biais**
K	10	11,1593	0,7741	1,1593
λ	10/30	0,1982	0,0524	−0,1352
$r_{\mathbf{a}}$	0,8484	1,6157	0,6979	0,7673
s	6	6,0732	0,0999	0,0732
$r_{\mathbf{b}}$	$5{,}25\times10^{-3}$	$5{,}3424\times10^{-3}$	$0{,}3554\times10^{-3}$	$0{,}0924\times10^{-3}$

TAB. 4.3 – Estimation des hyperparamètres.

Comme pour la méthode par déconvolution impulsionnelle myope, les estimations de s et de $r_{\mathbf{b}}$ sont correctes. En revanche, l'estimation de λ est inférieure à la valeur réelle et l'estimation de $r_{\mathbf{a}}$ est

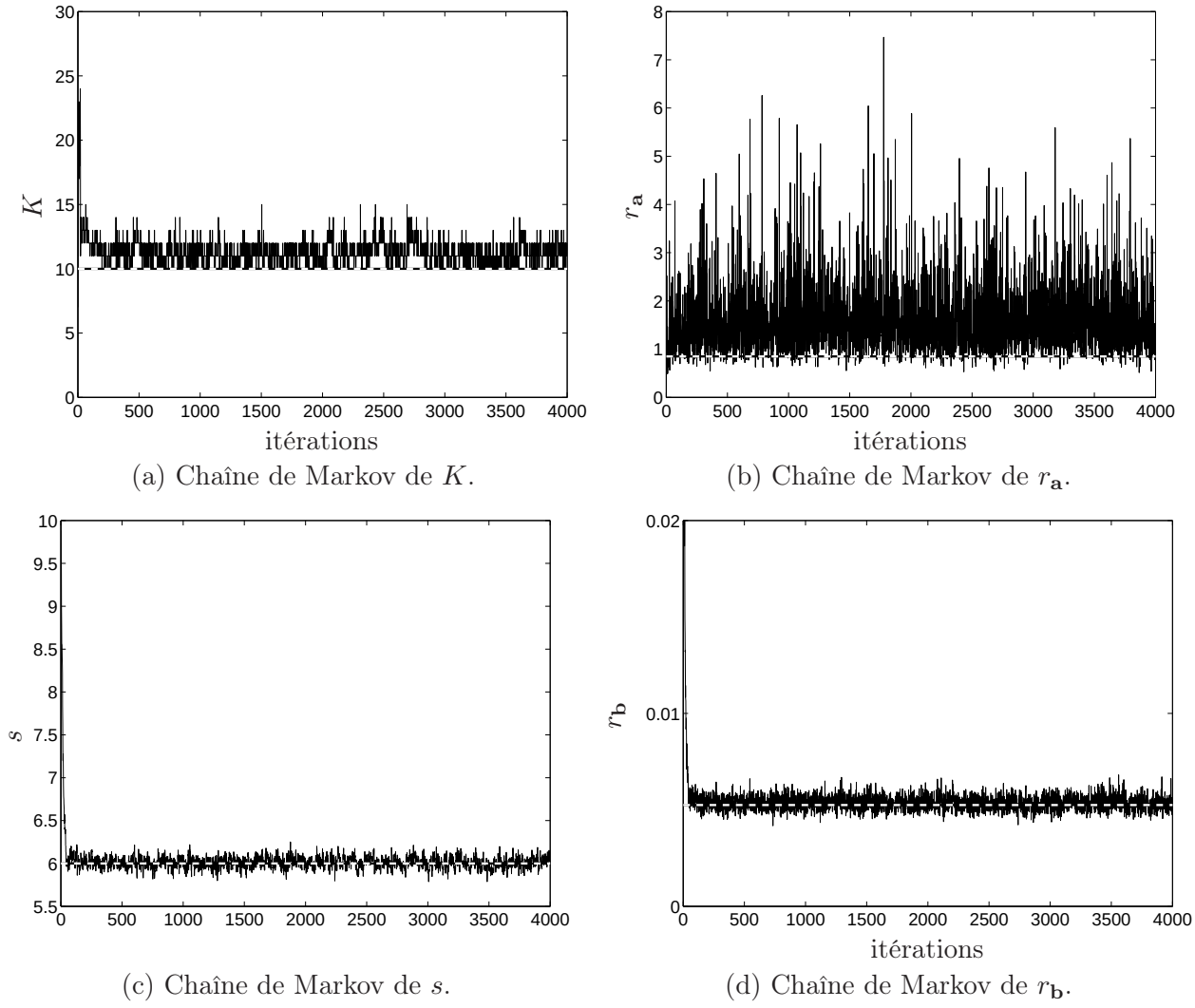


(a) Positions des motifs.



(b) Amplitude des motifs.

FIG. 4.11 – Évolution des chaînes de Markov des paramètres du spectre (tirets : valeurs réelles).



(a) Chaîne de Markov de K.

(b) Chaîne de Markov de $r_{\mathbf{a}}$.

(c) Chaîne de Markov de s.

(d) Chaîne de Markov de $r_{\mathbf{b}}$.

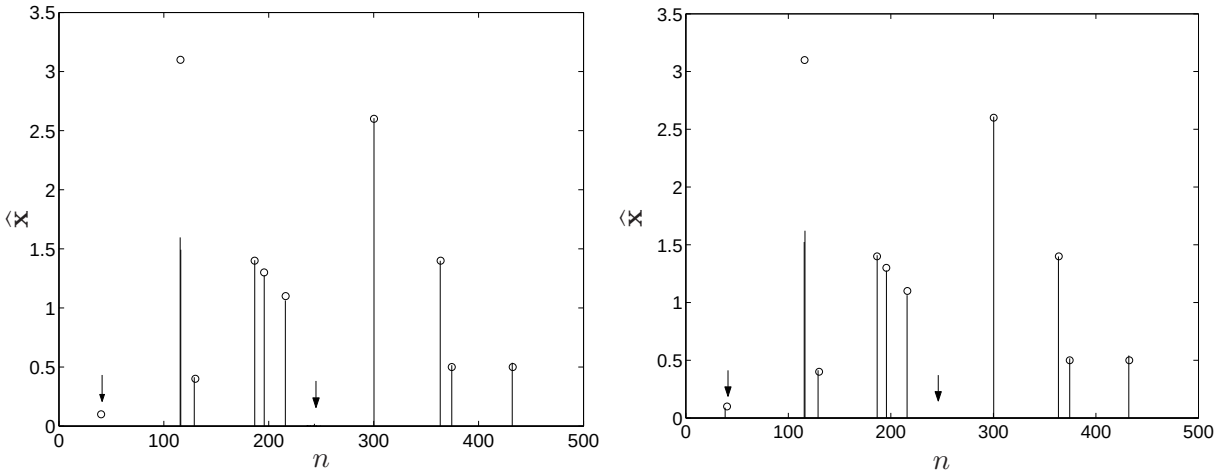
FIG. 4.12 – Évolution des chaînes de Markov de s et des hyperparamètres (tirets : valeurs réelles).

supérieure (avec une variance très importante). En outre, la valeur de $\widehat{K}$ est supérieure à la valeur réelle, mais elle reste toute de même du même ordre de grandeur.

Par conséquent, d'un point de vue de la qualité de l'estimation des hyperparamètres, la décomposition en motifs élémentaires est une méthode équivalente à la déconvolution impulsionnelle myope.

Estimation du spectre pur

La figure 4.13(a) représente l'estimation du spectre pur obtenue avant avoir effectué l'étape de ré-indexage. L'estimation du spectre pur après ré-indexage est présentée figure 4.13(b). La méthode de ré-indexage a duré 132 secondes qu'il faut rajouter aux 238 secondes de la décomposition proprement dit. Cela fait un total de 370 secondes, à comparer aux 870 secondes de l'approche par déconvolution impulsionnelle myope.



(a) Avant ré-indexage. (b) Après ré-indexage.

FIG. 4.13 – Estimation du signal impulsionnel. Les ∘ indiquent les raies réelles.

Globalement, l'estimation initiale du spectre pur est satisfaisante. Quelques défauts sont cependant corrigés grâce à la méthode de ré-indexage (cf. flèches) : certaines amplitudes sont mieux estimées, les faux motifs ont disparus (il sont très petits et situés au niveau de $n = 245$) et, surtout, la première raie est retrouvée. Cela confirme que le ré-indexage est nécessaire et efficace. Toutefois, la raie 2 est toujours estimée par deux motifs, localisés respectivement en 115,52 et 116,19 et d'amplitudes 1,52 et 1,62.

Nous pensons qu'une modification de la méthode de ré-indexage proposée pourrait combiner deux raies proches en une seule. En effet, on peut ajouter une quatrième condition aux équations (4.17) qui consiste à sélectionner tous les motifs dont la distance $D(\boldsymbol{\theta}_l^{(i)}, \bar{\boldsymbol{\theta}}_k)$ est inférieure à une certaine valeur. Si alors plusieurs motifs sont sélectionnés à une même itération, ils sont remplacés par un motif unique, comme cela est fait pour l'instant manuellement.

Le tableau 4.4 rapporte les positions et amplitudes des motifs estimés, après ré-indexage, et en regroupant les deux motifs estimant la raie 2. En comparant le résultat de cette estimation avec celui

k	1	2	3	4	5	6	7	8	9	10
$\mathbf{c}^*$	40,2	115,9	130	186,5	195,6	216,1	300,3	363,8	374,4	432,5
$\mathbf{a}^*$	0,1	3,1	0,4	1,4	1,3	1,1	2,6	1,4	0,5	0,5
$\widehat{\mathbf{c}}$	38,35	115,87	129,05	186,60	195,66	215,84	300,33	363,45	374,46	432,07
$\widehat{\mathbf{a}}$	0,09	3,14	0,41	1,41	1,28	1,04	2,61	1,40	0,52	0,54

TAB. 4.4 – Estimation du signal impulsionnel après ré-indexage.

obtenu par déconvolution impulsionnelle myope, on constate que l'approche traitée dans ce chapitre donne des estimations légèrement meilleures en termes de biais.

4.5.2 Cas de raies différentes

La méthode est maintenant testée sur un spectre similaire au spectre simulé précédent, mais dont les motifs sont de largeurs différentes. La figure 4.14 représente le spectre et le tableau 4.5 reporte les paramètres des dix motifs.

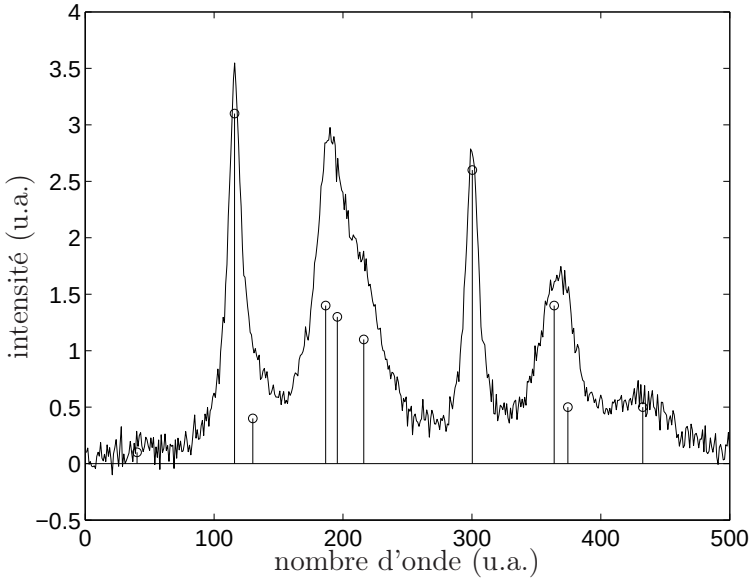


FIG. 4.14 – Spectre bruité simulé.

k	1	2	3	4	5	6	7	8	9	10
$\mathbf{c}^*$	40,2	115,9	130	186,5	195,6	216,1	300,3	363,8	374,4	432,5
$\mathbf{a}^*$	0,1	3,1	0,4	1,4	1,3	1,1	2,6	1,4	0,5	0,5
$\mathbf{s}^*$	10	6	15	10	15	20	6	15	6	30

TAB. 4.5 – Positions et amplitudes des raies du spectre pur.

Nous avons appliqué la méthode par décomposition en motifs élémentaires (motifs différents) sur ce spectre simulé, avec les valeurs initiales suivantes :

$$K_{\max} = 30, \qquad r_{\mathbf{a}} = 2, \qquad r_{\mathbf{b}} = 0,1,$$

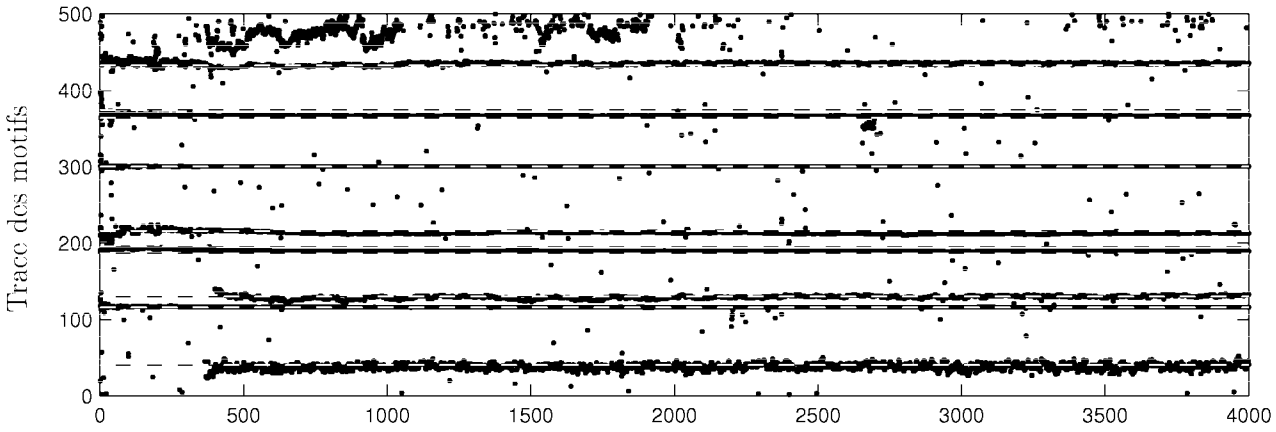
et un spectre initial contenant 15 motifs répartis uniformément, d'amplitudes nulles et de largeurs égales à 20.

Nous avons effectué $I = 4000$ itérations de l'échantillonneur de Gibbs ; les figures 4.15 et 4.16 représentent l'évolution des chaînes de $\mathbf{q}$, $\mathbf{a}$, $\mathbf{s}$, K, $r_{\mathbf{a}}$ et $r_{\mathbf{b}}$. La longueur de la période de transition a été choisie à $I_0 = 1000$ itérations.

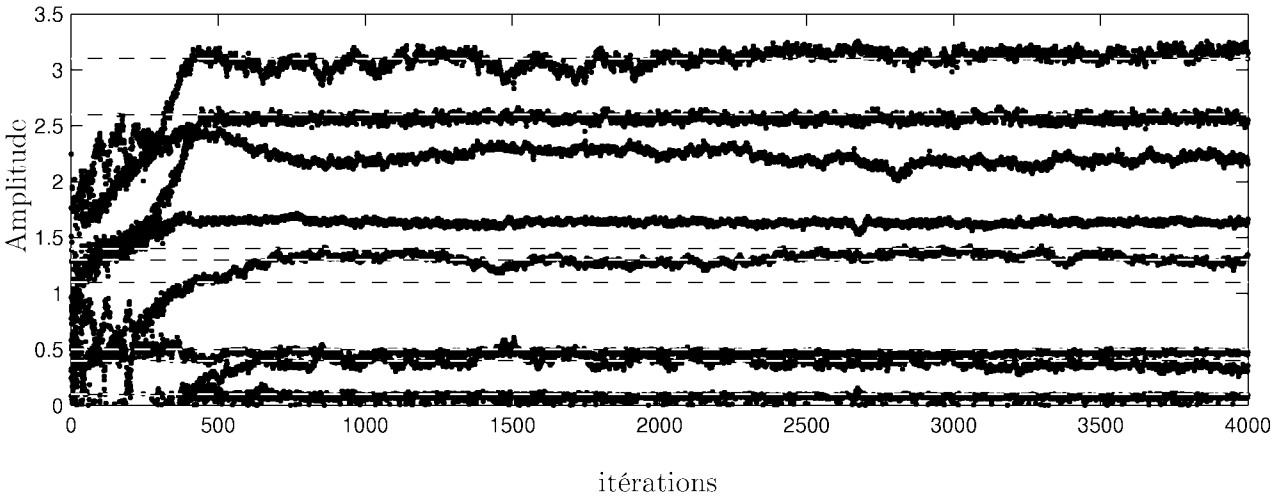
Le temps de calcul nécessaire a été d'environ 376 secondes, soit plus que la méthode où les motifs sont identiques puisqu'il y a plus de variables à estimer.

Estimation des hyperparamètres

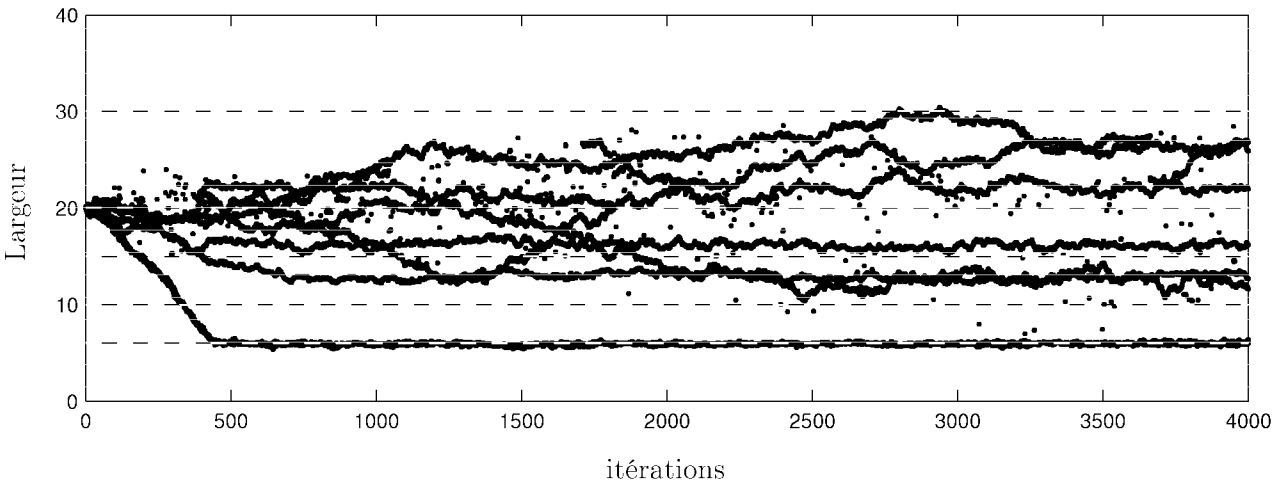
Les hyperparamètres de la méthode ont été estimés au sens de l'EAP. Le tableau 4.6 donne les valeurs réelles et estimées des hyperparamètres, ainsi que l'écart-type et le biais de l'estimation.



(a) Trace des motifs.

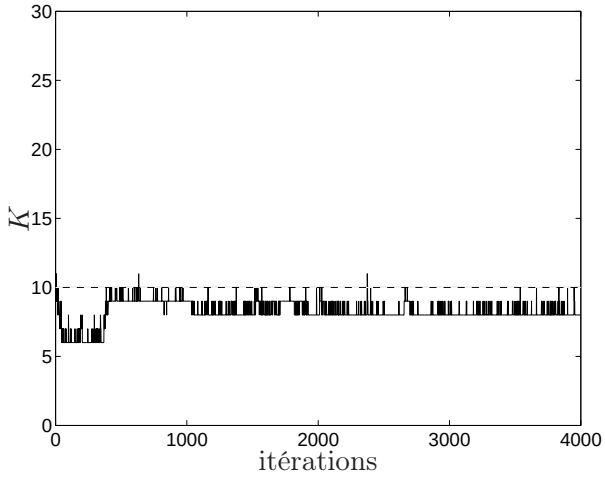


(b) Amplitude des motifs.

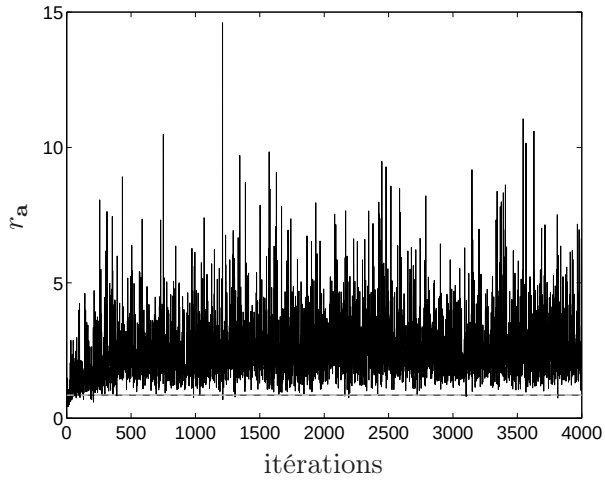


(c) Largeur des motifs.

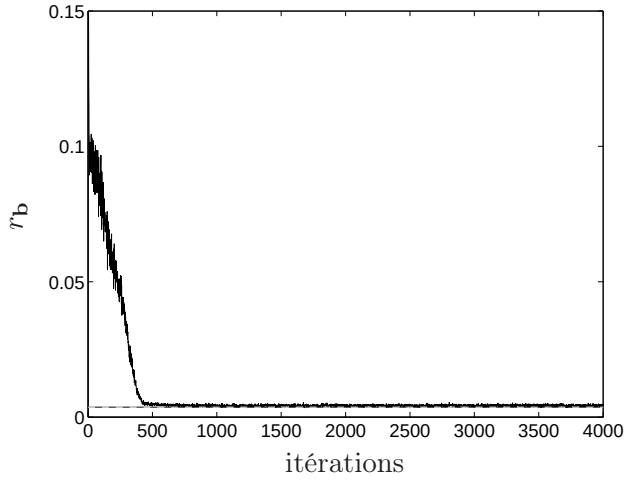
FIG. 4.15 – Évolution des chaînes de Markov des paramètres du spectre (tirets : valeurs réelles).



(a) Chaîne de Markov de K.



(b) Chaîne de Markov de $r_{\mathbf{a}}$.



(c) Chaîne de Markov de $r_{\mathbf{b}}$.

FIG. 4.16 – Évolution des chaînes de Markov des hyperparamètres (tirets : valeurs réelles).

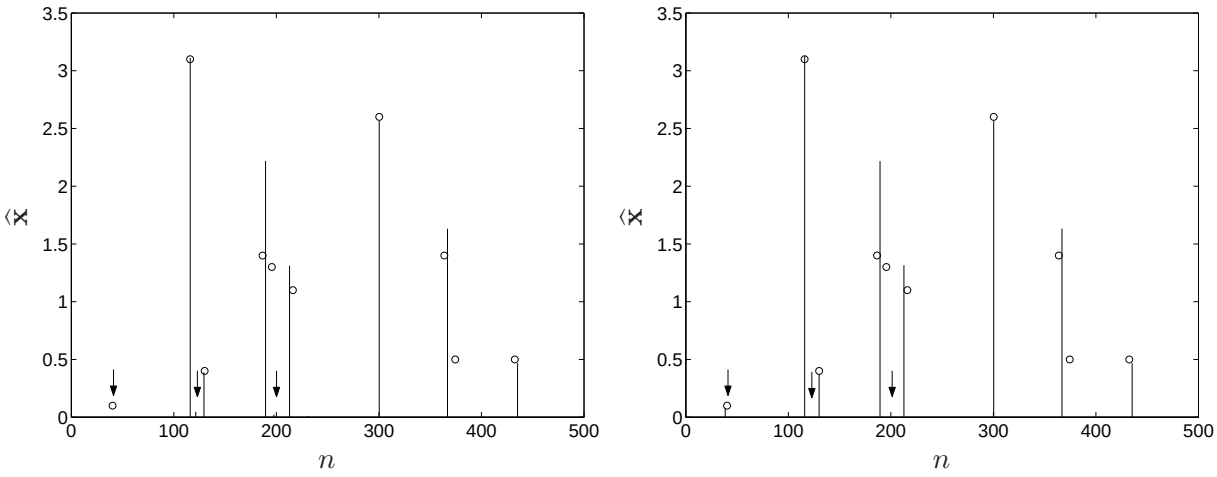
	valeur réelle	**estimation**	**écart-type**	**biais**
K	10	8,2286	0,4447	$-1{,}7714$
λ	$10/30$	0,1486	0,0451	$-0{,}1847$
$r_{\mathbf{a}}$	0,8484	2,7119	1,2942	1,8635
$r_{\mathbf{b}}$	3.6793×10^{-3}	$4{,}2591\times10^{-3}$	$0{,}2757\times10^{-3}$	$0{,}5799\times10^{-3}$

TAB. 4.6 – Estimation des hyperparamètres.

Comme dans les simulations précédentes, la valeur de $r_{\mathbf{b}}$ est bien estimée et les estimations de λ et $r_{\mathbf{a}}$ sont respectivement inférieures et supérieures aux vraies valeurs. En revanche, le nombre de motifs estimé $\widehat{K}$ est inférieur à la valeur réelle. Ce dernier point s'explique par le fait que la méthode a plus de mal à distinguer deux motifs trop proches et qu'ils sont en fait estimés par un seul motif (voir plus loin).

Estimation du spectre pur

La figure 4.17(a) représente l'estimation du spectre pur obtenue avant d'avoir effectué l'étape de ré-indexage. L'estimation du spectre pur après ré-indexage est présentée figure 4.17(b). Le ré-indexage a duré 88 secondes à rajouter au temps de calcul de la décomposition, soit 464 secondes au total. Notons que le ré-indexage est plus rapide que dans la simulation précédente ; cela s'explique en partie parce que le nombre de motifs estimés est plus faible.



(a) Avant ré-indexage. (b) Après ré-indexage.

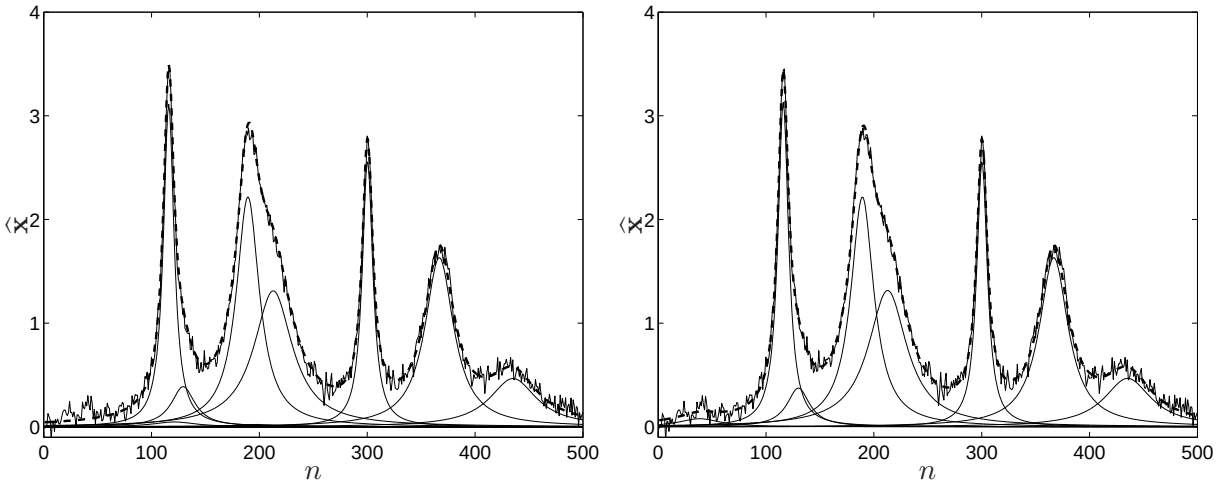
FIG. 4.17 – Estimation du signal impulsionnel. Les ∘ indiquent les raies réelles.

Le tableau 4.7 rapporte les positions et amplitudes des raies estimées.

k	1	2	3	4	5	6	7	8	9	10
$\mathbf{c}^*$	40,2	115,9	130	186,5	195,6	216,1	300,3	363,8	374,4	432,5
$\mathbf{a}^*$	0,1	3,1	0,4	1,4	1,3	1,1	2,6	1,4	0,5	0,5
$\mathbf{s}^*$	10	6	15	10	15	20	6	15	6	30
$\widehat{\mathbf{c}}$	38,44	115,98	130,03	189,42		212,78	300,34	366,87		435,34
$\widehat{\mathbf{a}}$	0,08	3,13	0,37	2,22		1,32	2,56	1,63		0,47
$\widehat{\mathbf{s}}$	23,41	5,99	12,70	12,91		21,67	5,92	16,19		26,57

TAB. 4.7 – Estimation du signal impulsionnel après ré-indexage.

Globalement, les raies sont correctement estimées. En outre, la méthode de ré-indexage permet de retrouver la raie 1. Toutefois, sa largeur estimée est beaucoup plus large qu'en réalité. Ce biais est sans



(a) Avant ré-indexage. (b) Après ré-indexage.

FIG. 4.18 – Spectre reconstruit (tirets).

doute dû à la faible amplitude de la raie qui ne permet pas de l'estimer avec précision. Par ailleurs, les raies 4, 5 et 6 d'une part et 8 et 9 d'autre part sont respectivement estimées par deux et un seul motif. Néanmoins, cela n'empêche pas le spectre d'être correctement reconstruit (voir figure 4.18(b)).

En conclusion, même si cette méthode ne donne pas une estimation parfaite, elle reste néanmoins préférable à la méthode par déconvolution impulsionnelle myope dont l'inconvénient majeur est qu'elle ne prend pas en compte le fait que les raies sont différentes.

4.6 Application sur des spectres réels

Dans cette section, nous présentons les estimations des raies des spectres expérimentaux de gibbsite (voir section 1.2). Rappelons que pour chaque type de spectroscopie (infrarouge ou Raman), les spectres correspondent au même échantillon de gibbsite, mais avec une ligne de base différente. La ligne de base de ces spectres a été estimée puis soustraite grâce à la méthode du chapitre 2. Une fois corrigés, les spectres sont à peu près similaires, seules quelques différences peuvent malgré tout induire des différences sur les estimations.

Pour plus de facilité, les spectres ont été découpés en plusieurs zones, et chacune a été traitée séparément : cela permet d'une part d'obtenir un résultat plus rapidement car le nombre d'itérations peut être plus faible, d'autre part de limiter le phénomène de permutation d'indices.

L'estimation a été effectuée grâce à la méthode présentée dans ce chapitre car c'est l'approche la plus rapide et dont la modélisation est la plus proche de la réalité. Pour chaque spectre, nous avons effectué $I = 10\,000$ itérations et la période de transition a été choisie à $I_0 = 5000$ itérations. Les valeurs initiales correspondent à un spectre nul avec $\lambda = 0,5$, $r_{\mathbf{a}} = 2$ et $r_{\mathbf{b}} = 0,1$. Le nombre maximal de pics $K_{\max}$ dépend du spectre.

Par ailleurs, il est conseillé d'initialiser les motifs de manière à ce qu'ils soient très larges (par l'intermédiaire de la variance de la gaussienne dans le cas d'un spectre infrarouge ou de la largeur de la lorentzienne en spectroscopie Raman). En effet, en partant avec une valeur faible, le spectre risque d'être estimé par un grand nombre de motifs très fins, ou alors la convergence sera très lente.

4.6.1 Spectres infrarouge

Les spectres infrarouge ont été découpés en trois zones dont les intervalles sont (en cm^{-1}) :

$$[450\,;\,1206], \qquad [1206\,;\,2801], \qquad [2801\,;\,3981].$$

Sur chaque zone, la méthode par décomposition en motifs élémentaires a été appliquée en considérant un nombre maximal de $K_{\max} = 40$ motifs, de forme gaussienne. Les résultats des estimations sont présentés sur les figures 4.19, 4.20 et 4.21 (pages 110 à 112), et les tableaux 4.8, 4.9 et 4.10 récapitulent les caractéristiques des raies estimées, à savoir les nombres d'onde, intensités et aires. L'aire d'une lorentzienne est calculée par la formule suivante :

$$\mathcal{A}_L = a \int_{-\infty}^{+\infty} \frac{s^2}{s^2 + x^2} dx = as \int_{-\infty}^{+\infty} \frac{1}{1 + y^2} dy = as \left(\arctan(+\infty) + \frac{\pi}{2} \right) = as\pi.$$

Tous les spectres sont bien reconstruits, sauf les quelques raies très fines présentent sur les zones 2 des spectres. De plus, comme la correction de la ligne de base dans cette zone n'est pas optimale (cf. section 2.5.1), seules quelques raies estimées sont retrouvées dans les trois spectres. Toutefois, cette zone centrale du spectre n'intéresse pas les chimistes car il ne s'y trouve aucune information intéressante sur l'échantillon (la plupart des raies correspondent d'ailleurs à l'eau et au dioxyde de carbone présents dans l'air au moment de l'expérience : cette partie du spectre n'est donc pas parfaitement reproductible d'une mesure à l'autre).

En ce qui concerne les deux autres zones, l'estimation est très satisfaisante, et plusieurs raies sont retrouvées dans les trois spectres. Par exemple, la troisième zone correspond à l'élongation O–H. Le tableau 4.10 (page 112) présente dans la dernière colonne les raies qui ont une signification physique. Ainsi, les raies « G » correspondent à des modes de vibration de la gibbsite, les raies « B » à ceux de la bayérite (dont la molécule est également $Al(OH)_3$ mais dont la configuration géométrique est différente de la gibbsite) et enfin « S » est la raie de l'eau en surface de l'échantillon. Toutes les autres raies (très larges ou peu intenses) qui ne sont pas marquées n'ont pas de réelle signification physique. Elles correspondent soit à un résidu de ligne de base qui n'aurait pas été corrigé (les plus larges), soit à des raies de vapeur d'eau, soit enfin à une raie annexe qui permet de compenser le caractère asymétrique de raies plus importantes (c'est le cas pour la raie marquée « A » qui peut être associée à la raie située en 3525 cm^{-1}).

4.6.2 Spectres Raman

Les spectres Raman ont été découpés en quatre zones dont les intervalles sont (en cm^{-1}) :

$$[86\,;\,192], \qquad [193\,;\,463], \qquad [464\,;\,665], \qquad [665\,;\,1175].$$

Une décomposition en motifs élémentaires a été appliquée sur chaque zone de chaque spectre en considérant un maximum de $K_{\max} = 20$ motifs lorentziens. Les résultats sont présentés figures 4.23, 4.24, 4.25 et 4.26 (pages 114 à 117). Les tableaux 4.12, 4.13, 4.14 et 4.15 récapitulent les nombres d'onde, intensités, largeurs et aire des raies estimées. L'aire d'une gaussienne est égale à :

$$\mathcal{A}_G = a \int_{-\infty}^{+\infty} \exp\left(-\frac{x^2}{2s^2} \right) dx = a\sqrt{2\pi s^2}.$$

Les estimations obtenues sont très satisfaisantes, d'une part parce que les motifs retrouvés sont, à quelques exceptions près, reproduits dans chaque spectre (et leurs paramètres sont assez proches d'un spectre à l'autre), d'autre part parce que les raies retrouvées ont une réelle signification chimique.

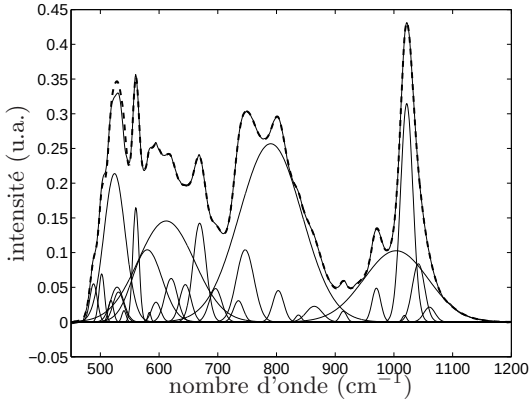
Les raies situées entre 350 et 650 cm^{-1} correspondent à une déformation d'angle O–Al–O, et les raies situées au delà de 700 cm^{-1} correspondent à la déformation d'angle Al–O–H. En particulier, dans la troisième zone, on estime bien les deux raies situées aux alentours de 539 cm^{-1} et 570 cm^{-1}. En revanche, pour les spectres $\mathcal{R}_0$ et $\mathcal{R}_2$, la petite raie située en 556 cm^{-1} n'a aucune signification chimique, alors qu'au contraire il semble manquer une raie en 490 cm^{-1}.

Toutefois, certaines raies sont parfois estimées par plusieurs motifs très proches, mais dans la majorité des cas cela est dû à l'asymétrie de la raie estimée. Par exemple, dans la zone 3, les raies situées en 524 cm^{-1} et 539 cm^{-1} semblent reproduire un même schéma d'asymétrie.

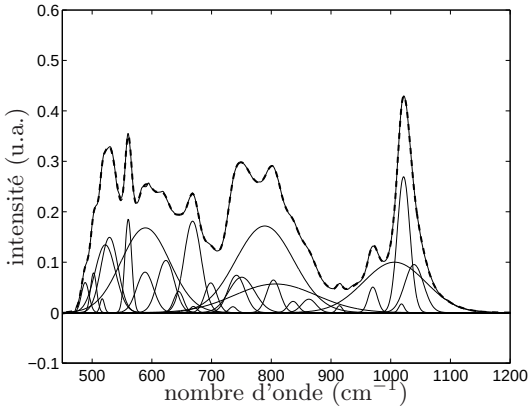
Spectres infrarouge — zone 1

$\mathcal{I}_0$				$\mathcal{I}_1$				$\mathcal{I}_2$			
$\hat{\mathbf{c}}$	$\hat{\mathbf{a}} \times 1000$	$\hat{\mathbf{s}}$	$\mathcal{A}_L$	$\hat{\mathbf{c}}$	$\hat{\mathbf{a}} \times 1000$	$\hat{\mathbf{s}}$	$\mathcal{A}_L$	$\hat{\mathbf{c}}$	$\hat{\mathbf{a}} \times 1000$	$\hat{\mathbf{s}}$	$\mathcal{A}_L$
488,30	55,21	6,34	19,93	488,30	59,76	6,38	20,05	488,40	50,77	6,33	19,88
502,20	69,50	4,76	14,95	502,30	78,96	5,10	16,02	501,70	54,87	4,00	12,56
517,90	31,26	4,38	13,76	516,80	27,61	3,62	11,36				
524,10	213,85	18,01	56,58	522,00	134,54	17,08	53,66	521,60	173,06	18,25	57,34
				529,00	149,38	14,33	45,03	527,50	121,91	14,75	46,34
528,40	50,28	11,18	35,12								
530,80	43,19	9,14	28,70								
540,20	16,51	3,56	11,17								
560,30	164,99	5,26	16,51	560,40	185,51	5,64	17,71	560,40	174,06	5,40	16,97
579,70	104,10	25,30	79,48	588,50	80,32	14,16	44,47	584,00	119,87	33,52	105,31
583,30	13,82	2,72	8,55	589,10	168,04	40,56	127,44	587,50	30,05	10,94	34,37
594,70	28,63	7,76	24,38	623,50	103,53	14,08	44,23	619,70	19,78	6,60	20,73
612,10	145,56	47,96	150,66	644,80	41,89	8,25	25,92	636,90	162,35	53,62	168,44
620,50	62,80	10,68	33,55								
644,70	53,93	9,71	30,50								
668,90	142,48	11,04	34,68	668,10	181,41	14,35	45,09	649,50	13,92	7,83	24,59
696,00	48,29	10,79	33,90	669,60	12,31	5,73	18,01	669,20	93,76	9,27	29,14
735,10	30,77	8,38	26,33	698,60	58,96	10,35	32,51	697,60	37,68	13,04	40,97
746,60	104,05	14,89	46,79	735,70	11,72	6,11	19,21	736,10	44,59	10,35	32,53
790,10	256,85	53,14	166,93	741,60	74,12	14,23	44,70	743,10	111,48	14,72	46,25
				750,40	70,30	21,46	67,43	760,00	34,11	11,01	34,60
				789,00	171,80	48,49	152,35	789,80	211,11	38,06	119,57
802,90	45,37	9,38	29,48	803,50	64,73	11,34	35,62	803,70	39,53	9,30	29,22
837,40	10,11	5,21	16,36	806,90	56,87	68,70	215,84	826,80	60,71	55,90	175,62
864,20	22,70	13,79	43,32	836,50	22,35	8,73	27,41	837,50	18,00	7,56	23,76
913,70	15,55	5,72	17,98	862,90	26,75	12,11	38,03	862,00	26,38	12,30	38,63
								868,70	18,57	26,63	83,65
								913,90	14,19	4,82	15,14
				914,20	13,79	4,65	14,61	955,40	48,07	32,82	103,11
970,10	48,88	7,45	23,39	970,00	50,73	7,74	24,31	970,30	57,08	8,35	26,23
1003,70	102,69	54,66	171,71	1006,40	100,06	52,18	163,93	1018,20	30,55	5,78	18,15
1017,90	9,52	3,81	11,98	1018,10	17,26	4,66	14,65				
1021,80	314,79	10,45	32,84	1021,60	269,64	10,51	33,03	1022,30	263,50	10,68	33,55
1041,70	83,92	10,40	32,67	1039,10	95,30	15,63	49,09	1025,50	117,38	33,93	106,58
1060,40	21,40	9,66	30,34					1041,90	59,28	13,19	41,45
								1096,30	10,12	31,31	98,36

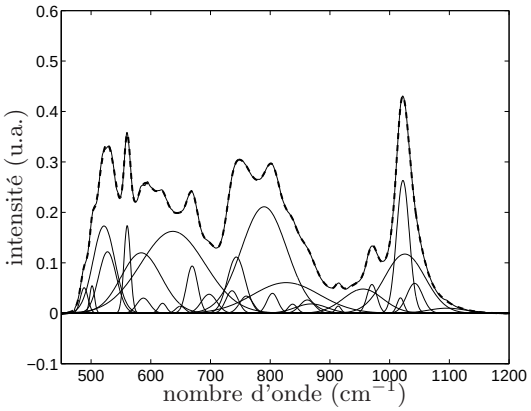
TAB. 4.8 – Raies estimés pour les spectres infrarouge (zone 1).



(a) $\mathcal{I}_0$.



(b) $\mathcal{I}_1$.



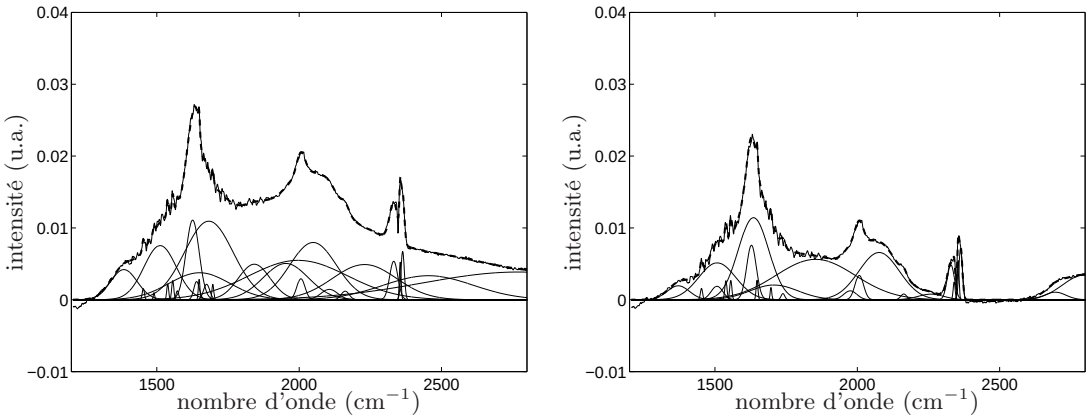
(c) $\mathcal{I}_2$.

FIG. 4.19 – Zone 1 des spectres infrarouge : spectre corrigé, raies estimées et spectre reconstruit (tirets).

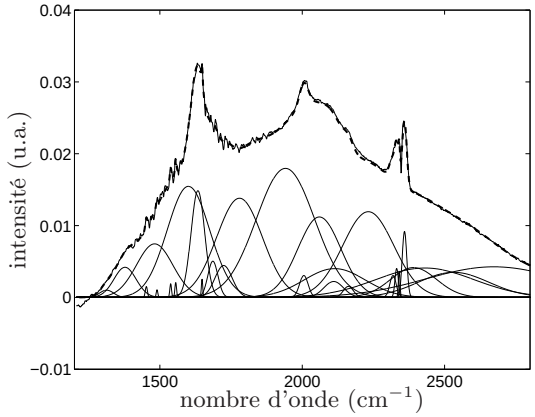
Spectres infrarouge — zone 2

$\mathcal{I}_0$				$\mathcal{I}_1$				$\mathcal{I}_2$			
$\widehat{\mathbf{c}}$	$\widehat{\mathbf{a}} \times 1000$	$\widehat{\mathbf{s}}$	$\mathcal{A}_L$	$\widehat{\mathbf{c}}$	$\widehat{\mathbf{a}} \times 1000$	$\widehat{\mathbf{s}}$	$\mathcal{A}_L$	$\widehat{\mathbf{c}}$	$\widehat{\mathbf{a}} \times 1000$	$\widehat{\mathbf{s}}$	$\mathcal{A}_L$
1384,30	4,20	46,57	0,61	1371,10	1,95	38,89	122,19	1312,90	1,02	24,23	76,12
1452,50	1,56	3,88	0,02	1452,70	1,58	3,94	12,37	1377,50	4,20	38,95	122,35
1490,20	0,99	2,52	0,01	1507,10	1,88	20,60	64,71	1452,40	1,52	3,75	11,78
1511,70	7,54	56,77	1,35	1508,20	5,14	64,94	204,03	1481,60	7,46	63,78	200,36
1538,10	2,21	3,39	0,02	1538,30	2,65	3,88	12,18	1490,00	1,06	2,81	8,82
1555,90	2,67	3,72	0,03	1555,60	2,69	3,79	11,91	1537,70	1,91	2,97	9,33
1572,10	1,24	4,75	0,02	1628,10	7,59	17,00	53,39	1555,50	2,09	2,88	9,05
1625,60	11,09	22,87	0,80	1635,30	11,43	53,58	168,32	1600,20	15,46	74,76	234,86
1639,50	2,54	9,74	0,08	1647,70	2,73	3,18	10,00	1633,40	14,86	24,11	75,75
1648,50	2,87	2,81	0,03	1696,80	1,71	2,76	8,68	1647,90	2,53	2,39	7,51
1650,20	3,76	106,77	1,26	1707,90	2,05	83,55	262,48	1686,10	5,05	17,59	55,26
1674,90	2,15	11,05	0,07	1738,60	0,86	6,44	20,22	1723,70	4,45	30,25	95,03
1682,70	10,94	85,20	2,93	1853,60	5,63	138,47	435,02	1779,70	13,81	80,46	252,76
1696,80	2,14	3,20	0,02	1975,30	1,24	23,01	72,29	1941,00	17,97	102,56	322,20
1842,20	4,95	55,77	0,87								
1952,90	5,06	79,42	1,26								
1997,00	5,49	173,51	2,99								
2005,90	2,92	14,41	0,13	2006,70	3,38	15,10	47,45	2005,40	3,05	14,95	46,97
2048,90	7,98	81,71	2,05	2076,60	6,58	71,13	223,47	2060,10	11,20	75,40	236,86
2106,00	1,45	29,29	0,13	2163,00	0,80	10,64	33,43	2110,40	2,23	34,20	107,43
2161,40	1,20	14,17	0,05	2258,80	0,78	45,19	141,97	2114,20	4,02	107,18	336,72
2229,70	4,91	112,40	1,73					2161,90	1,54	17,82	55,99
2331,90	5,36	14,44	0,24	2328,30	4,87	12,79	40,18	2231,90	11,94	89,52	281,22
2354,80	5,23	3,02	0,05	2340,00	2,48	3,46	10,86	2318,50	3,05	11,28	35,44
2363,40	6,75	5,88	0,12	2354,30	5,47	3,11	9,76	2331,50	4,04	6,04	18,99
				2362,70	7,19	6,10	19,16	2340,30	3,67	2,83	8,89
2454,00	3,35	157,04	1,65	2695,70	1,06	37,09	116,53	2359,10	9,19	8,06	25,31
2769,70	3,86	335,20	4,06	2796,50	3,44	74,56	234,25	2391,20	4,16	80,34	252,38
								2431,20	4,10	171,97	540,26
								2531,90	3,50	123,53	388,07
								2673,60	4,26	238,34	748,75

TAB. 4.9 – Raies estimés pour les spectres infrarouge (zone 2).



(a) $\mathcal{I}_0$. (b) $\mathcal{I}_1$.



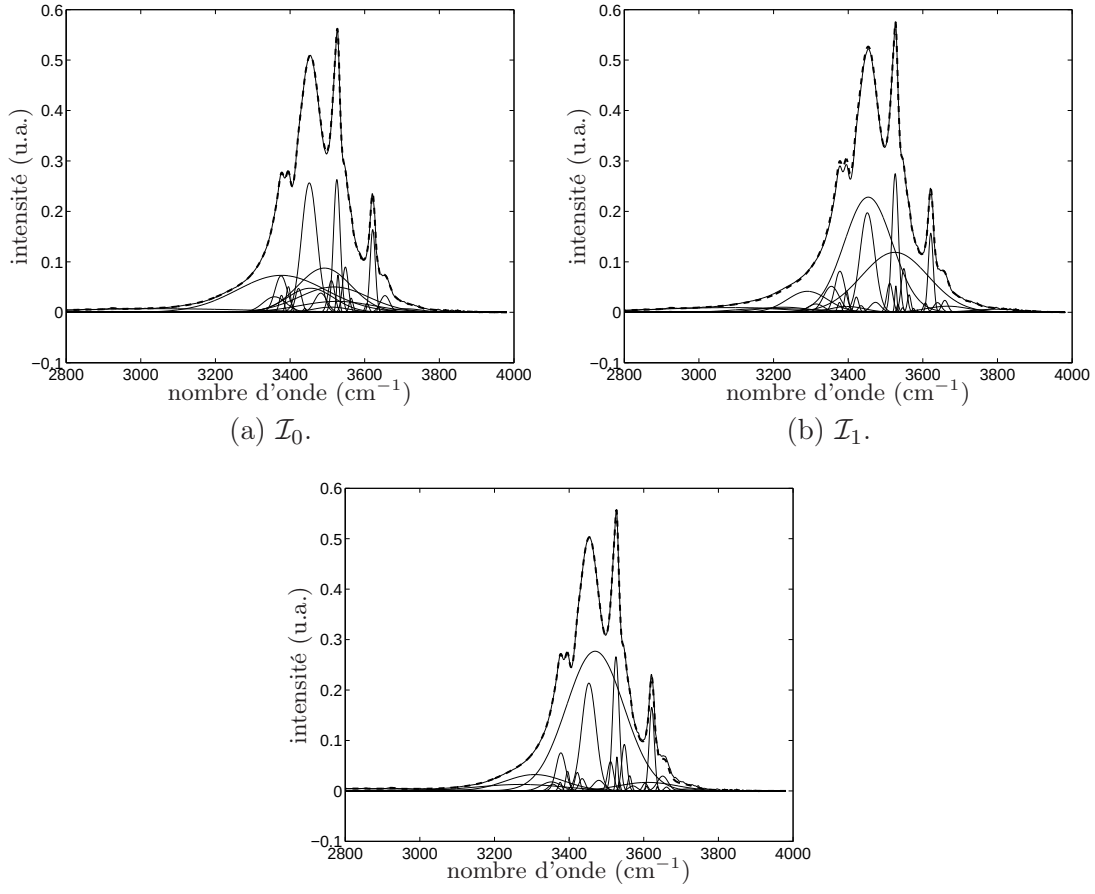
(c) $\mathcal{I}_2$.

FIG. 4.20 – Zone 2 des spectres infrarouge : spectre corrigé, raies estimées et spectre reconstruit (tirets).

Spectres infrarouge — zone 3

$\mathcal{I}_0$				$\mathcal{I}_1$				$\mathcal{I}_2$				
$\hat{\mathbf{c}}$	$\hat{\mathbf{a}} \times 1000$	$\hat{\mathbf{s}}$	$\mathcal{A}_L$	$\hat{\mathbf{c}}$	$\hat{\mathbf{a}} \times 1000$	$\hat{\mathbf{s}}$	$\mathcal{A}_L$	$\hat{\mathbf{c}}$	$\hat{\mathbf{a}} \times 1000$	$\hat{\mathbf{s}}$	$\mathcal{A}_L$	
2923,40	1,76	10,32	0,06	2923,7	1,74	10,31	32,39	2829,7	4,51	207,37	651,46	
3046,50	7,06	249,72	5,54	3082,4	9,22	209,75	658,95	2923,8	1,81	10,72	33,67	
3358,30	30,68	30,88	2,98	3171,4	7,21	79,87	250,93	3270,6	12,79	117,72	369,84	
3375,90	72,23	16,52	3,75	3226,5	10,09	68,54	215,34	3308,3	32,50	77,91	244,76	
				3291	41,56	54,25	170,44	3352,1	17,85	26,83	84,29	
				3313,6	16,49	21,71	68,20	3356,6	13,29	11,55	36,29	
				3354,9	51,45	21,23	66,71					
3377,00	33,59	6,31	0,67	3376,6	19,48	5,02	15,78	3376	16,97	4,81	15,12	G
3377,10	73,22	118,25	27,20	3377,7	81,39	13,84	43,49	3377,9	75,49	12,89	40,49	
3395,60	50,98	5,91	0,95	3395,6	42,46	5,40	16,96	3395,8	38,95	5,05	15,86	G
				3398,3	11,52	45,44	142,76					
3422,80	47,16	8,87	1,31	3421,9	30,26	6,30	19,79	3421,9	36,63	6,58	20,69	
3439,80	7,47	4,54	0,11	3434,3	13,13	7,51	23,60	3434,9	24,39	8,39	26,37	
3451,60	256,55	20,51	16,53	3451,3	197,24	19,51	61,30	3453	213,78	19,01	59,71	S
				3453,8	228,30	64,17	201,61	3469,7	276,93	77,02	241,96	
3454,50	47,71	62,55	9,38	3473,1	19,69	14,39	45,19	3479,5	21,00	14,60	45,86	
3482,30	37,96	16,53	1,97									
3492,40	87,61	70,85	19,50									
3510,70	10,27	46,96	1,52									
3510,80	62,96	9,03	1,79	3511,3	57,64	8,00	25,12	3511	58,17	8,14	25,56	A
3517,40	50,28	92,37	14,59	3525	118,88	89,10	279,92					
3525,40	263,29	8,85	7,32	3525,9	274,88	8,51	26,73	3525,4	266,17	8,43	26,48	G
3525,60	21,63	99,19	6,74	3527,7	52,39	2,86	8,99	3528,1	67,31	3,25	10,22	
3528,40	73,95	3,96	0,92	3533,4	15,41	3,33	10,47	3534,6	18,92	2,67	8,39	
				3545,8	9,09	3,27	10,28					
3548,20	89,72	7,89	2,22	3549,1	86,71	7,29	22,89	3547,9	92,22	7,26	22,79	B
3563,60	28,00	5,16	0,45					3562,4	30,15	4,82	15,15	
3606,50	17,40	5,24	0,29	3563,3	34,76	4,91	15,43	3571,1	9,93	11,17	35,08	
				3574,9	7,74	3,27	10,28	3603,6	15,51	6,77	21,28	
				3585,6	4,73	1,30	4,09	3608	9,54	2,90	9,12	
				3606,8	18,32	5,67	17,80	3611,3	16,93	77,93	244,83	
3621,50	164,31	6,92	3,57	3621,2	156,96	6,65	20,88	3621,3	165,74	6,91	21,70	G
3636,60	11,35	2,47	0,09	3639,3	19,17	10,45	32,83	3636,3	8,97	2,06	6,46	
3654,30	33,14	12,79	1,33	3659	23,66	8,45	26,54	3650,7	29,40	14,39	45,19	B
3710,80	6,90	42,31	0,92	3667,1	12,01	57,07	179,31	3660,9	7,68	6,00	18,84	
3825,70	2,73	41,65	0,36	3738,3	2,82	11,53	36,21	3718,2	3,93	24,42	76,70	
				3798,7	6,04	64,92	203,95	3823,4	1,31	29,33	92,14	

Tab. 4.10 – Raies estimés pour les spectres infrarouge (zone 3).



(a) $\mathcal{I}_0$. (b) $\mathcal{I}_1$.

(c) $\mathcal{I}_2$.

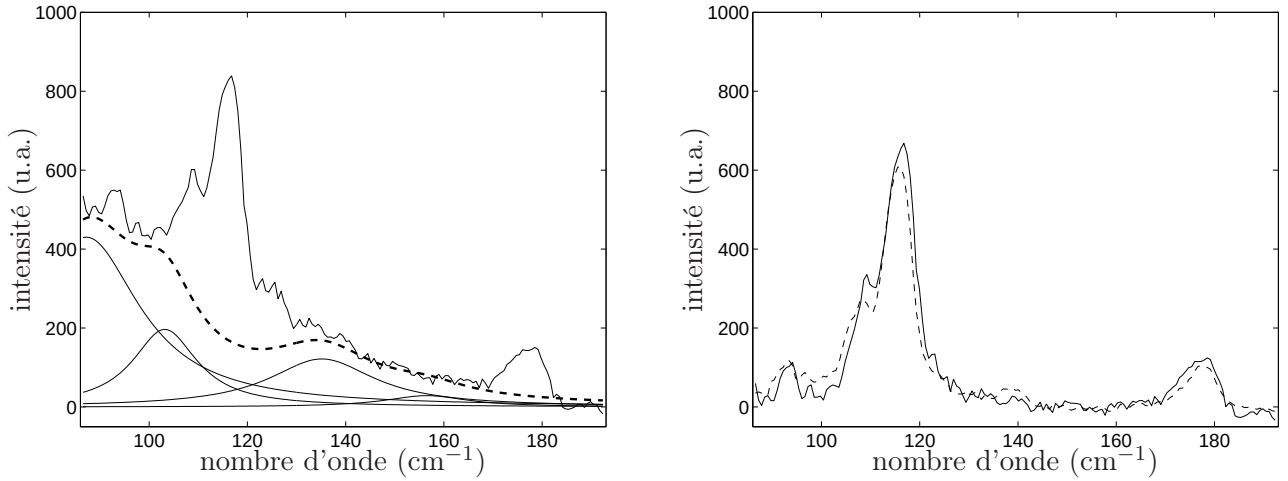
Fig. 4.21 – Zone 3 des spectres infrarouge : spectre corrigé, raies estimées et spectre reconstruit (tirets).

Comme nous l'avons évoqué dans la section 2.5.2, la ligne de base due à la diffusion Rayleigh des spectres $\mathcal{R}_1$ et $\mathcal{R}_2$ n'est pas correctement estimée. Une des possibilité de correction de cette ligne de base est de considérer que les motifs les plus larges qui composent le signal constituent en fait une ligne de base résiduelle.

Nous montrons que cette méthode peut être appliquée en dernier recours. Considérons l'exemple du spectre $\mathcal{R}_1$. Il semble que quatre raies correspondent à la ligne de base résiduelle. En effet, elles n'apparaissent pas sur le spectre $\mathcal{R}_0$ pour lequel il n'existe pas de ligne de base résiduelle et elles sont très larges, ce qui signifie qu'elles évoluent moins vite que les vrais motifs (c'était, rappelons-le, la principale caractéristique de la ligne de base). Les paramètres de ces quatres raies sont répertoriés dans le tableau 4.11.

$\widehat{\mathbf{c}}$	$\widehat{\mathbf{a}}$	$\widehat{\mathbf{s}}$
87,24	430,18	13,13
103,18	196,73	8,00
135,15	121,65	13,11
156,49	28,68	10,31

TAB. 4.11 – Raies pouvant constituer la ligne de base résiduelle du spectre $\mathcal{R}_1$.



(a) Ligne de base estimée à partir de quatre motifs.

(b) Comparaison entre les spectres corrigés $\mathcal{R}_1$ (traits pleins) et $\mathcal{R}_0$ (tirets).

FIG. 4.22 – Estimation de la ligne de base sur un spectre Raman de gibbsite.

La figure 4.22(a) représente ces quatres raies ainsi que leur somme, qui est donc une estimation de la ligne de base résiduelle, et la figure 4.22(b) présente le spectre $\mathcal{R}_1$ corrigé par cette nouvelle estimation ainsi que le spectre $\mathcal{R}_0$. Les deux spectres sont assez semblables : cette procédure peut donc être utilisée comme estimation de la ligne de base, mais une estimation par la méthode présentée dans le chapitre 2 reste préférable.

4.6.3 Améliorations possibles

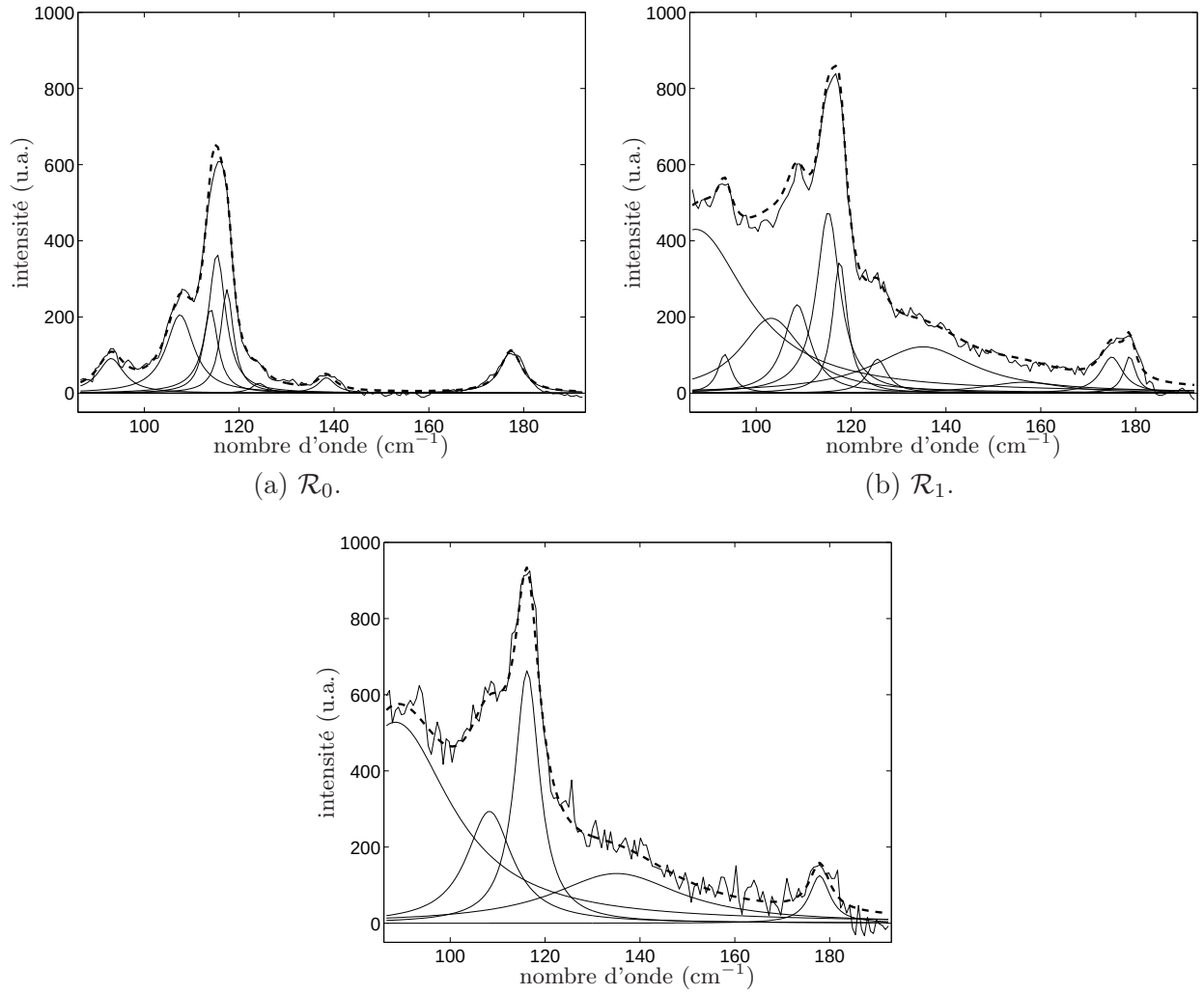
Outre le fait de prendre en compte des motifs asymétriques ou de forme encore plus évoluée, plusieurs modifications sont possibles pour améliorer les estimations. Elles consistent à rajouter des informations très précises sur le spectre qui seront intégrées grâce à l'approche bayésienne. Ces informations sont spécifiques à chaque spectre et proviennent des connaissances a priori qu'ont les chimistes sur l'échantillon étudié.

Ainsi, il est possible de préciser l'ordre de grandeur des nombres d'onde et des largeurs des raies les plus grandes et les mieux résolues. Les informations sont codées en imposant sur les paramètres

Spectres Raman — zone 1

$\mathcal{R}_0$				$\mathcal{R}_1$				$\mathcal{R}_2$			
$\widehat{\mathbf{c}}$	$\widehat{\mathbf{a}}$	$\widehat{\mathbf{s}}$	$\mathcal{A}_G/10^3$	$\widehat{\mathbf{c}}$	$\widehat{\mathbf{a}}$	$\widehat{\mathbf{s}}$	$\mathcal{A}_G/10^3$	$\widehat{\mathbf{c}}$	$\widehat{\mathbf{a}}$	$\widehat{\mathbf{s}}$	$\mathcal{A}_G/10^3$
				87,24	430,18	13,13	14,16	88,44	527,91	14,24	18,84
93,00	91,20	3,30	0,75	93,36	101,97	1,76	0,45				
				103,18	196,73	8,00	3,95				
107,53	205,96	3,53	1,82	108,65	232,36	3,35	1,95	108,26	293,82	5,85	4,31
114,00	222,00	1,82	1,02	115,19	476,14	2,96	3,53	116,21	663,11	3,15	5,24
115,33	365,43	2,20	2,01								
117,43	271,49	1,65	1,12	117,65	346,93	1,85	1,61				
124,21	26,16	1,85	0,12	125,58	89,63	2,46	0,55				
138,40	40,73	2,15	0,22	135,15	121,65	13,11	4,00	135,10	130,55	16,54	5,41
				156,49	28,68	10,31	0,74				
				174,97	95,01	2,85	0,68				
177,37	111,27	2,76	0,77	178,74	96,70	1,50	0,36	177,90	124,94	2,62	0,82

TAB. 4.12 – Raies estimées pour les spectres Raman (zone 1).



(a) $\mathcal{R}_0$. (b) $\mathcal{R}_1$.

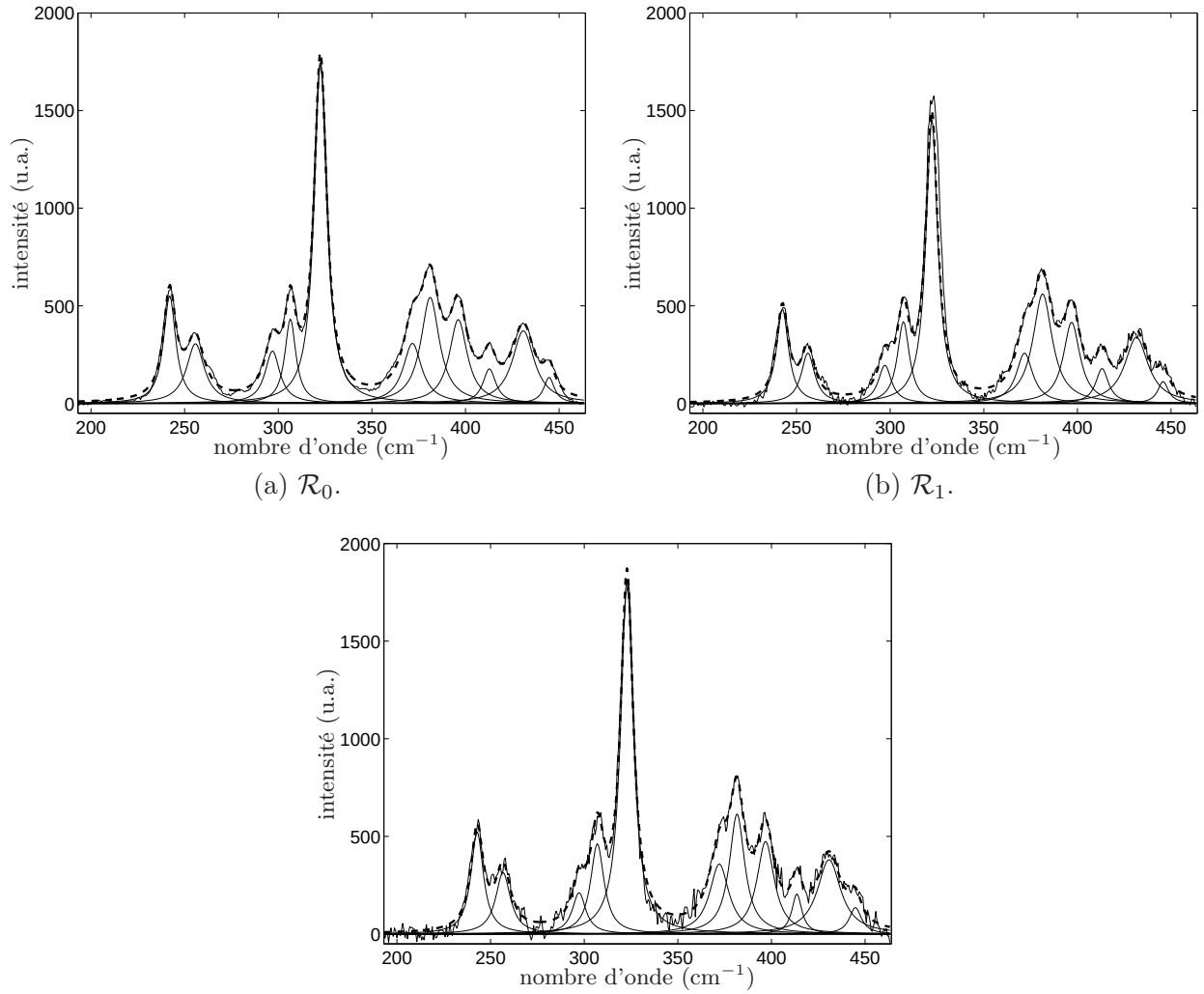
(c) $\mathcal{R}_2$.

FIG. 4.23 – Zone 1 des spectres Raman : spectre corrigé, raies estimées et spectre reconstruit (tirets).

Spectres Raman — zone 2

$\mathcal{R}_0$				$\mathcal{R}_1$				$\mathcal{R}_2$			
$\widehat{\mathbf{c}}$	$\widehat{\mathbf{a}}$	$\widehat{\mathbf{s}}$	$\mathcal{A}_G/10^3$	$\widehat{\mathbf{c}}$	$\widehat{\mathbf{a}}$	$\widehat{\mathbf{s}}$	$\mathcal{A}_G/10^3$	$\widehat{\mathbf{c}}$	$\widehat{\mathbf{a}}$	$\widehat{\mathbf{s}}$	$\mathcal{A}_G/10^3$
242,03	550,76	3,96	5,47	242,51	483,50	3,88	4,70	242,85	522,84	4,03	5,28
255,79	306,22	5,81	4,46	256,12	258,30	4,34	2,81	256,67	315,41	4,76	3,76
296,93	269,16	4,80	3,24	297,20	196,30	4,26	2,10	297,25	212,24	4,34	2,31
306,49	432,78	3,56	3,86	307,10	419,10	3,89	4,09	307,02	462,43	4,36	5,06
322,55	1754,40	4,04	17,75	322,27	1455,20	3,80	13,85	322,85	1816,47	4,04	18,40
371,66	308,15	6,06	4,68	371,77	258,30	5,35	3,47	372,16	358,99	6,26	5,64
381,14	544,08	6,01	8,20	381,43	561,90	6,26	8,82	381,63	615,68	4,99	7,70
396,16	430,26	5,79	6,25	396,95	416,40	5,38	5,61	396,89	474,96	5,58	6,65
412,75	177,49	4,21	1,87	413,16	178,80	4,17	1,87	413,57	205,28	3,25	1,67
430,87	372,54	7,13	6,66	431,52	339,10	7,58	6,44	430,76	380,88	7,59	7,24
444,66	133,88	3,47	1,16	445,78	112,70	3,47	0,98	444,84	135,93	3,95	1,35

TAB. 4.13 – Raies estimées pour les spectres Raman (zone 2).



(a) $\mathcal{R}_0$. (b) $\mathcal{R}_1$.

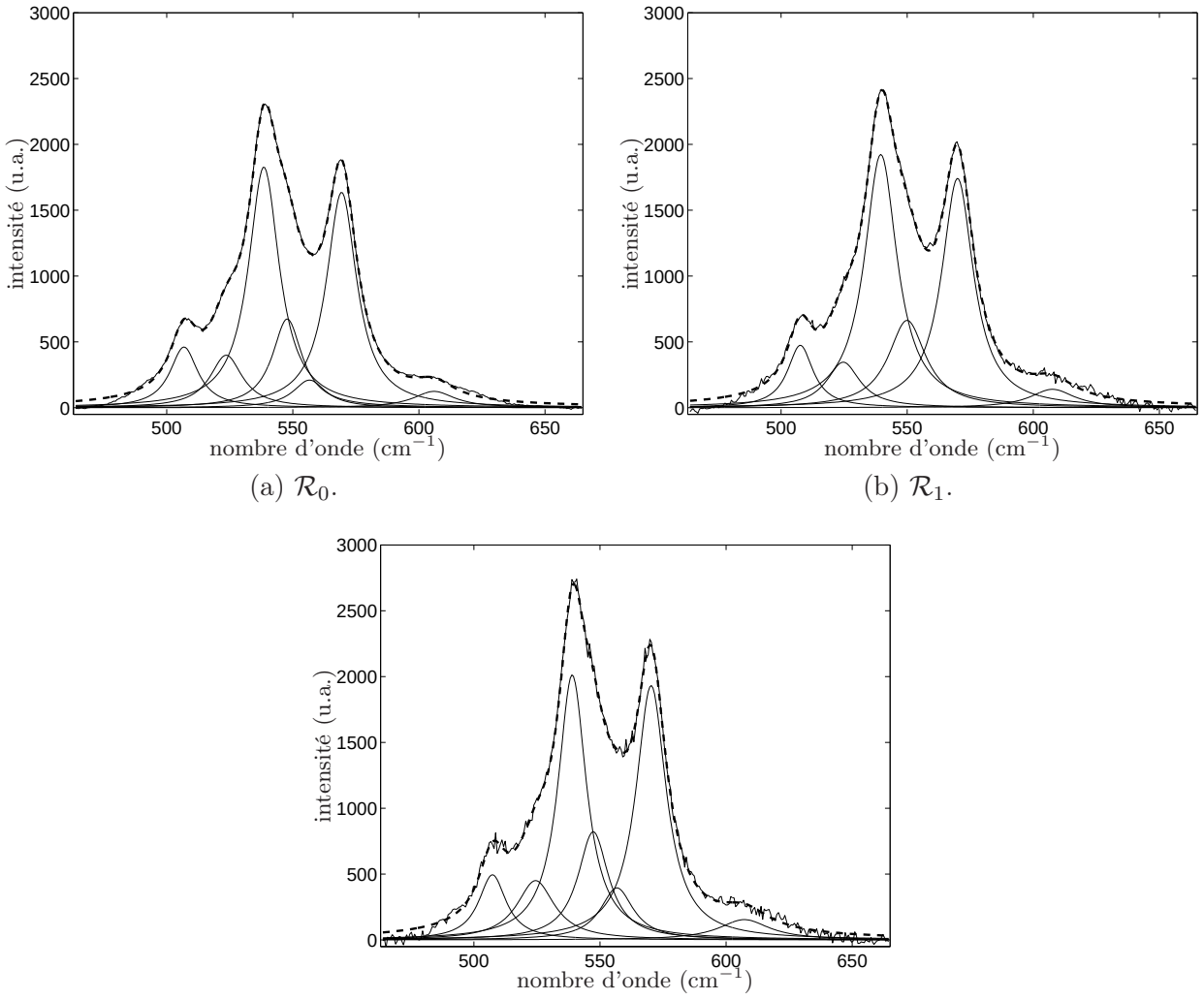
(c) $\mathcal{R}_2$.

FIG. 4.24 – Zone 2 des spectres Raman : spectre corrigé, raies estimées et spectre reconstruit (tirets).

Spectres Raman — zone 3

$\mathcal{R}_0$				$\mathcal{R}_1$				$\mathcal{R}_2$			
$\widehat{\mathbf{c}}$	$\widehat{\mathbf{a}}$	$\widehat{\mathbf{s}}$	$\mathcal{A}_G/10^3$	$\widehat{\mathbf{c}}$	$\widehat{\mathbf{a}}$	$\widehat{\mathbf{s}}$	$\mathcal{A}_G/10^3$	$\widehat{\mathbf{c}}$	$\widehat{\mathbf{a}}$	$\widehat{\mathbf{s}}$	$\mathcal{A}_G/10^3$
506,87	461,08	6,74	7,79	507,65	474,78	6,16	7,33	507,38	495,10	6,63	8,23
523,62	399,37	7,78	7,79	524,69	346,28	8,03	6,97	524,50	451,10	9,07	10,26
538,55	1828,99	7,39	33,86	539,56	1924,32	7,84	37,80	539,01	2013,90	6,79	34,29
547,79	674,23	7,11	12,02	549,97	663,75	8,99	14,96	547,32	822,20	7,34	15,13
556,62	210,00	7,88	4,15					556,71	395,80	7,95	7,89
569,36	1636,76	7,30	29,95	570,00	1742,95	7,62	33,27	570,28	1932,90	7,28	35,27
605,99	124,22	10,02	3,12	607,60	140,10	10,66	3,74	607,21	154,40	12,20	4,72

TAB. 4.14 – Raies estimées pour les spectres Raman (zone 3).



(a) $\mathcal{R}_0$. (b) $\mathcal{R}_1$.

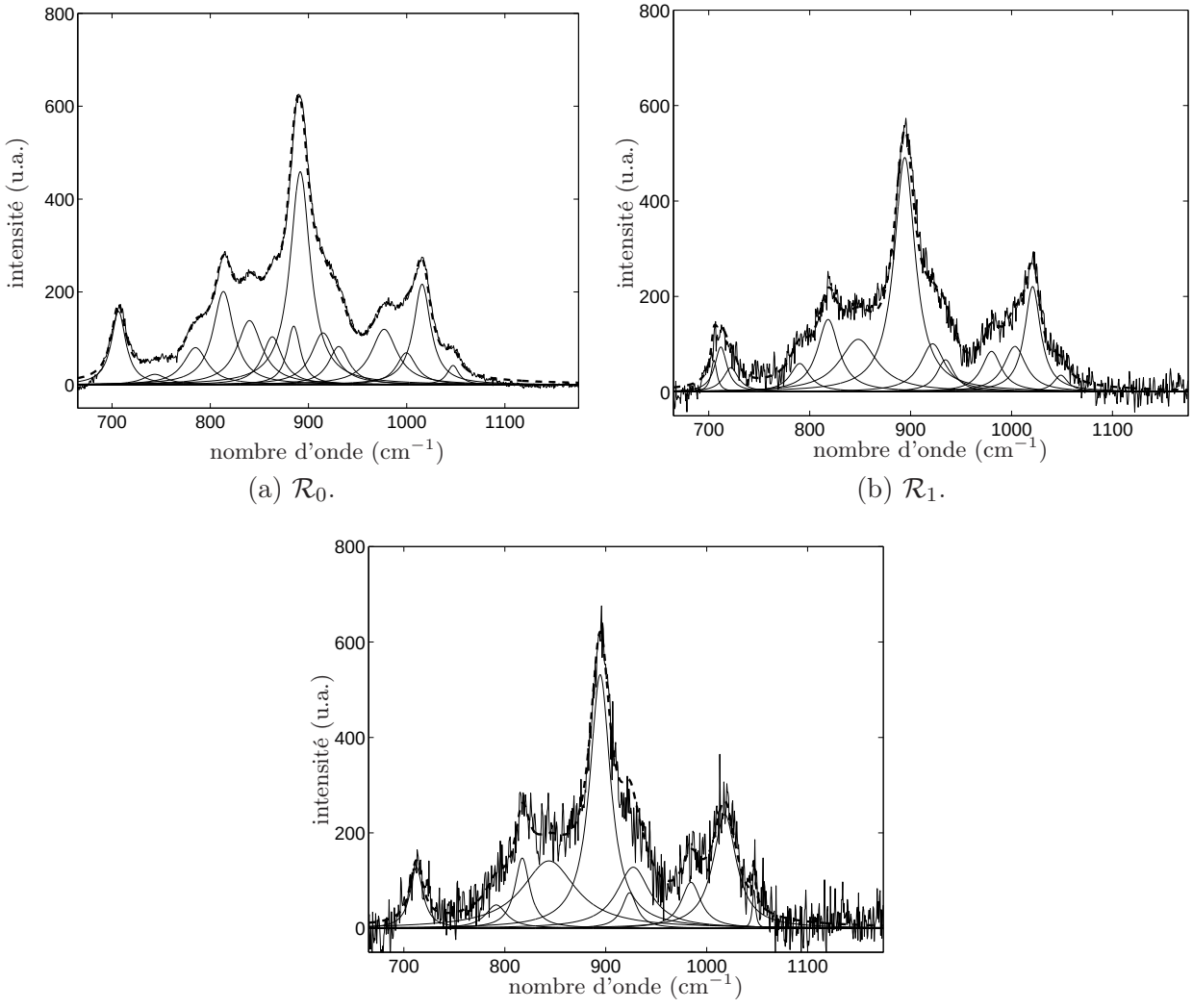
(c) $\mathcal{R}_2$.

FIG. 4.25 – Zone 3 des spectres Raman : spectre corrigé, raies estimées et spectre reconstruit (tirets).

Spectres Raman — zone 4

$\mathcal{R}_0$				$\mathcal{R}_1$				$\mathcal{R}_2$			
$\widehat{\mathbf{c}}$	$\widehat{\mathbf{a}}$	$\widehat{\mathbf{s}}$	$\mathcal{A}_G/10^3$	$\widehat{\mathbf{c}}$	$\widehat{\mathbf{a}}$	$\widehat{\mathbf{s}}$	$\mathcal{A}_G/10^3$	$\widehat{\mathbf{c}}$	$\widehat{\mathbf{a}}$	$\widehat{\mathbf{s}}$	$\mathcal{A}_G/10^3$
706,67	158,38	8,88	3,52	705,25	66,36	3,52	0,59	712,51	127,89	7,63	2,44
				712,01	94,34	7,26	1,72				
				721,93	50,83	9,04	1,15				
743,65	22,78	14,66	0,84								
784,96	80,17	14,79	2,97	790,20	59,49	11,62	1,73	791,49	49,15	12,83	1,58
813,46	201,12	12,74	6,42	818,34	151,88	12,84	4,89	817,35	147,38	8,93	3,30
839,96	138,36	15,45	5,36	848,12	110,21	22,43	6,20	843,76	141,33	32,48	11,51
863,16	103,50	12,25	3,18								
885,12	126,64	8,19	2,60								
891,64	459,12	12,34	14,20	894,19	491,34	12,53	15,43	894,53	531,67	12,42	16,55
914,88	111,84	14,61	4,10	921,86	100,89	14,80	3,74	923,64	74,65	9,00	1,68
930,92	82,60	12,62	2,61	934,64	67,42	10,33	1,75	927,37	128,18	20,05	6,44
977,15	119,64	15,95	4,78	980,24	85,12	11,90	2,54	984,54	96,57	11,91	2,88
999,73	69,15	12,11	2,10	1003,39	95,25	13,22	3,16				
1015,75	216,82	10,49	5,70	1020,93	220,82	9,57	5,30	1017,96	238,93	14,00	8,39
1047,28	41,89	7,31	0,77	1048,92	35,10	8,62	0,76	1046,90	81,66	1,83	0,38

TAB. 4.15 – Raies estimées pour les spectres Raman (zone 4).



(a) $\mathcal{R}_0$. (b) $\mathcal{R}_1$.

(c) $\mathcal{R}_2$.

FIG. 4.26 – Zone 4 des spectres Raman : spectre corrigé, raies estimées et spectre reconstruit (tirets).

de certaines raies une loi a priori plus précise que celle utilisée jusqu'à présent. Ainsi, plutôt que d'affecter une loi uniforme sur les positions, on peut poser un a priori gaussien pour imposer une position particulière sur les grandes raies. De même, en modifiant les paramètres de la loi a priori sur s (qui est une inverse gamma), on peut imposer un « intervalle » très précis de variation de la largeur des raies. Par exemple, en spectroscopie Raman, les lorentziennes ont une largeur à mi-hauteur moyenne de 12 ± 5 cm^{-1}. On peut donc poser comme loi a priori sur s (égale à la moitié de la largeur à mi-hauteur) une inverse gamma de moyenne 6 cm^{-1} et d'écart-type 2,5 cm^{-1}. De plus, la fonction d'appareil a une largeur de 2,7 cm^{-1} : on ne peut donc pas trouver de raies plus fines que cette largeur. Une solution consiste donc à introduire une contrainte supplémentaire pour interdire certaines largeurs. Cette contrainte peut également être introduite en jouant sur les paramètres de la loi a priori de s.

4.7 Conclusion

Ce chapitre propose une alternative à l'approche par déconvolution impulsionnelle myope pour l'estimation des raies du spectre. Elle consiste en une méthode originale de décomposition d'un signal en motifs élémentaires. Cette approche s'inspire du processus Bernoulli-gaussien dans le sens où elle considère un nombre de motifs maximal $K_{\max}$ affectés d'une « variable d'occurrence » $\mathbf{q}$ qui indique pour chaque motif s'il est présent ou non dans le spectre. Dès lors, la dimension du problème est fixe et il est donc possible d'utiliser des méthodes MCMC classiques, comme l'échantillonneur de Gibbs.

Ainsi, cette méthode n'a pas les trois limitations principales d'une déconvolution impulsionnelle myope en permettant d'estimer des raies de paramètres de forme différents et situées en dehors de la grille discrète d'échantillonnage, tout en étant plus rapide puisque le nombre de variables à estimer est plus faible.

Un problème important a cependant été rencontré lors de l'estimation des paramètres des raies. En effet, il peut arriver que deux motifs permutent au cours de la simulation, c'est-à-dire qu'il y a une permutation des indices de ces motifs. Cette permutation d'indices est rendue possible par le fait que la loi a posteriori est symétrique par rapport aux variables. Nous avons donc proposé une méthode originale qui permet de ré-indexer les motifs, permettant par la suite d'estimer les paramètres de raies au sens de l'EAP. Cette méthode de ré-indexage considère en outre que des motifs peuvent apparaître ou disparaître d'une itération à l'autre.

La méthode proposée a ensuite été testée sur les spectres simulés du chapitre 3 afin de comparer les performances des deux approches. Dans le cas où les raies du spectres sont identiques, nous conseillons de considérer les motifs identiques : cela permet d'inclure dans le modèle une connaissance a priori supplémentaire qui restreint l'espace des solutions et donc améliore l'estimation. Par ailleurs, cela permet de diminuer le nombre de variables à simuler.

Enfin, la méthode a été appliquée sur les signaux infrarouge et Raman de gibbsite dont la ligne de base a été corrigée grâce à l'algorithme du chapitre 2. Il est montré que, d'une manière générale, la méthode réussit à obtenir des estimations identiques pour chaque spectre, tant que la ligne de base est correctement estimée. Or, ce n'est pas le cas pour la zone 2 des spectres infrarouge et la zone 1 des spectres Raman, mais nous avons montré qu'il était possible, en dernier recours, de considérer la somme des raies les plus larges comme une estimation de la ligne de base.

Plusieurs extensions sont envisageables suites à ce travail.

En effet, le problème rencontré dans le chapitre précédent où une raie pouvait être estimée par deux pics est toujours présent ici. Pour l'éliminer, une solution serait de modifier la loi a priori sur les positions des motifs, en imposant une distance minimale entre deux motifs, par exemple à l'aide d'un processus de Poisson. La méthode pourrait également être étendue au cas où le spectre est échantillonné irrégulièrement (pour des raisons techniques par exemple).

Par ailleurs, avec une approche telle que celle présentée ici, on considère que les motifs sont tous de la même forme, en l'occurrence une lorentzienne. Or, il serait intéressant de considérer des systèmes où les motifs peuvent être de formes différentes. Par exemple, on peut considérer le cas d'un mélange

de motifs gaussiens et lorentziens. Une solution consiste alors à regrouper les différents motifs en une seule expression. Dans le cas d'un mélange de gaussiennes et de lorentziennes, on peut modéliser chaque motif par une fonction de pseudo-Voigt (voir [58] et la section 1.7.1) :

$$h(n) = \kappa \exp\left(-\frac{n^2}{2\sigma^2}\right) + (1-\kappa)\frac{s^2}{s^2+n^2},$$

avec $\kappa \in [0,1]$. Cette expression possède trois paramètres à estimer, à savoir κ, σ et s. Par conséquent, le nombre de variables à estimer sera alors de $6K_{\max}+3$:

$$\boldsymbol{\theta} = \{\mathbf{q}, \mathbf{c}, \mathbf{a}, \boldsymbol{\kappa}, \boldsymbol{\sigma}, \mathbf{s}, \lambda, r_{\mathbf{a}}, r_{\mathbf{b}}\}.$$

Une dernière perspective de ce travail est la suivante. Dans certains types de spectres, il est possible que certains motifs soient de même forme et de même paramètres. En fait, le spectre est constitué de plusieurs familles de motifs. Il y a alors moins de motifs différents que de raies. L'approche développée dans ce chapitre peut être appliquée, mais elle sera incapable de regrouper les raies de même paramètres de forme. Aussi, les recherches futures pourront s'intéresser à développer une méthode pour estimer un signal impulsionnel de K raies filtrées séparément par l'un des L motifs ($L \leq K$).

Une première méthode serait d'utiliser l'approche précédente couplée à une méthode de classification qui regrouperait les paramètres des motifs en L groupes. Une deuxième méthode, qui nous semble plus intéressante, est de coder la variable $\mathbf{q}$ sur $\{0,1,2,\ldots,L\}$. Ainsi, une absence de motif serait toujours codée par $\mathbf{q}_k = 0$ et l'existence d'un motif par $\mathbf{q}_k \neq 0$. La valeur de $\mathbf{q}_k$ indique à quel motif correspond la raie k. Cette approche pourra être modélisée par le système représenté figure 4.27 où h_l représente une lorentzienne de largeur $\mathbf{s}_l$ et tel que $K_L = K$. Les entrées $\sum_k \mathbf{a}_{l,k}\delta_{\mathbf{c}_{l,k}}(t)$ ($l \in \{0,\ldots,L\}$) sont ici

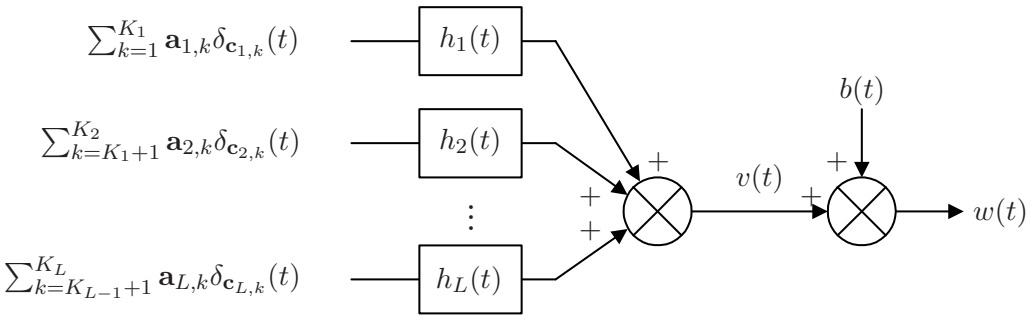


FIG. 4.27 – Modélisation de l'approche par décomposition en motifs élémentaires dans le cas de plusieurs groupes de motifs.

des signaux impulsionnels. Ainsi, cette modélisation peut être considérée comme une généralisation des approches précédentes, puisque le cas $L = 1$ correspond à la déconvolution impulsionnelle myope (en supposant les $\mathbf{c}_{l,k}$ discrets) et le cas $L = K$ correspond à la décomposition en motifs élémentaires présentée dans ce chapitre (puisqu'alors le système contient K réponses impulsionnelles h_l et que les entrées ne contiennent qu'une impulsion).

Conclusion

Ce travail de thèse a consisté à développer des méthodes numériques pour l'analyse de spectres, plus particulièrement en spectroscopie infrarouge et Raman.

Dans un premier temps, nous avons considéré le problème de la correction de la ligne de base. Ce problème a été résolu à l'aide d'une méthode déterministe consistant à estimer la ligne de base par le polynôme minimisant une fonction-coût. Cette fonction-coût doit croître moins vite qu'une parabole au delà d'un certain seuil, ceci afin d'éviter d'affecter un coût trop important aux raies. Nous avons donc considéré le cas de la fonction de Huber et de la parabole tronquée. De plus, nous avons montré que les versions asymétriques de ces fonctions-coût sont particulièrement bien adaptées aux spectres dont les raies sont positives (ce qui est le cas des spectres infrarouge et Raman) puisque l'asymétrie suppose que les raies sont toutes de mêmes signes. Il s'est alors avéré que la parabole tronquée asymétrique est la fonction-coût qui donne la meilleure estimation pour les spectres considérés. L'estimation a été obtenue grâce à l'algorithme de minimisation semi-quadratique LEGEND.

En outre, la méthode obtenue permet à l'opérateur de choisir l'ordre du polynôme et la valeur du seuil, laissant ainsi des degrés de liberté pour l'estimation de la ligne de base. L'influence de ces deux paramètres a été étudiée à travers des simulations et des indications sur le choix de leur valeur ont été données. Nous avons également testé le comportement de la méthode en fonction des paramètres du spectre. Enfin, des applications sur des spectres expérimentaux infrarouge et Raman de gibbsite ont montré l'efficacité de la méthode face à différentes lignes de base. Cette méthode a l'intérêt d'être très rapide et de permettre à l'utilisateur de fixer ses propres paramètres afin d'estimer la ligne de base attendue.

Deux perspectives s'ouvrent à ce travail. Tout d'abord, plutôt que de considérer un modèle polynomial où la douceur de la ligne de base est contrôlée par l'ordre du polynôme, nous pourrions introduire un terme de régularisation dans le critère qui permettrait un contrôle continu de la douceur de la ligne de base. Une seconde perspective serait de modéliser la ligne de base par des splines afin d'estimer des lignes de base plus irrégulières.

Dans un deuxième temps, nous avons considéré l'estimation des raies du spectre, c'est-à-dire l'estimation des nombres d'onde, intensités et paramètres de forme des raies. Pour cela, deux approches ont été présentées. Elles se placent toutes les deux dans un cadre probabiliste permettant d'introduire le maximum d'information a priori.

La première méthode considère le problème comme un problème de déconvolution impulsionnelle positive myope. Pour cela, un modèle bayésien hiérarchique a été développé, notamment en considérant le spectre pur comme un processus Bernoulli-gaussien convolué avec une réponse impulsionnelle de forme connue (dans ce document, nous avons considéré l'exemple d'une lorentzienne). L'échantillonneur de Gibbs est une approche directe pour simuler la loi a posteriori à partir des lois conditionnelles du modèle hiérarchique. À cet égard, nous avons été amené à proposer un algorithme d'acceptation-rejet mixte pour simuler une loi normale tronquée unilatéralement. Nous avons également proposé un estimateur original du signal impulsionnel en introduisant deux indices ; cet estimateur a été comparé à d'autres sur un exemple académique et un spectre simulé, montrant que l'estimateur proposé est plus efficace et qu'il permet notamment de retrouver des raies de très petite amplitude. En contrepartie, une interprétation de ces indices est nécessaire pour restaurer effectivement un signal impulsionnel, étape

qui actuellement est laissée à l'appréciation de l'utilisateur. En tout état de cause, dans le cadre de l'analyse de signaux de spectroscopie, cette première méthode souffre de deux inconvénients majeurs. D'une part les raies estimées sont forcément situées sur la grille d'échantillonnage et d'autre part elles ont toutes une largeur identique. Par ailleurs, la modélisation Bernoulli-gaussienne suppose un signal parcimonieux, c'est-à-dire ayant un grand nombre de valeurs nulles, mais impose de simuler tous ces échantillons, même ceux inutiles. Ceci a pour conséquence un coût de calcul élevé et une perte d'efficacité de la procédure d'estimation.

Une deuxième approche est alors présentée afin de résoudre ces deux problèmes. Celle-ci aborde le problème sous l'angle de la décomposition en motifs élémentaires, c'est-à-dire que le spectre est modélisé par une somme de motifs de forme connue. Dès lors, un parallèle entre déconvolution impulsionnelle myope et décomposition en motifs élémentaire est possible. Ce type de modélisation permet, d'une part, de placer les centres des raies estimées hors des « instants » d'échantillonnage et, d'autre part, de pouvoir les considérer toutes de largeurs différentes. De surcroît, le nombre de variables à estimer est plus petit qu'en déconvolution impulsionnelle myope. Cependant, une difficulté est de gérer un modèle dont la dimension peut varier d'une itération à l'autre. Pour éviter cela, nous nous sommes inspirés du processus Bernoulli-gaussien en considérant un nombre maximal de motifs qui peuvent, grâce à une variable d'occurrence, être présents ou absents du signal. Cette approche permet de maintenir la dimension du modèle constante et donc d'utiliser l'échantillonneur de Gibbs. Cependant, une difficulté supplémentaire résulte des permutations d'indices qui peuvent apparaître dans les chaînes de Markov des quantités simulées, interdisant alors une estimation naïve, au sens de l'EAP par exemple. Nous avons donc proposé une méthode de ré-indexage qui a l'avantage de pouvoir considérer que des motifs puissent apparaître ou disparaître. Une analyse de performance sur des signaux simulés confirme que cette méthode est plus rapide et plus efficace que l'approche par déconvolution impulsionnelle. Une application a des signaux expérimentaux a donné des résultats satisfaisants et exploitables par les physico-chimistes.

Toutefois, une limitation de ces deux approches d'estimation du spectre de raies réside dans le fait qu'une raie est parfois estimée par deux pics très proches : une perspective intéressante serait d'utiliser une autre loi a priori sur les positions. Les travaux futurs chercheront également à rendre le modèle actuel plus adapté aux spectres expérimentaux. Tout d'abord, on pourra s'intéresser à des motifs plus complexes, tels que des fonctions de Voigt ou des motifs asymétriques. Par ailleurs, on pourra également développer une méthode qui prend en compte le fait que l'échantillonnage du spectre n'est pas toujours régulier. Enfin, on pourra considérer une modélisation dans laquelle les motifs peuvent être regroupés en famille. Comme nous l'avons proposé dans la section 4.7, une modélisation intéressante serait de considérer que la variable d'occurrence peut prendre plusieurs valeurs correspondants soit à une absence du motif, soit à la famille à laquelle appartient ce motif.

Annexe A

Méthodes de Monte Carlo par chaînes de Markov

Cette annexe a pour but de présenter les algorithmes de Monte Carlo par chaînes de Markov utilisés dans cette thèse, en ne présentant que les notions utilisées. Pour de plus amples informations, le lecteur pourra se référer aux publications ayant inspiré cette annexe [3, 36, 49, 96, 103, 109], et en particulier les ouvrages de Gilks *et al.* [44], de Meyn et Tweedie [81] et de Robert [93].

A.1 Introduction

Les algorithmes présentés dans les chapitres 3 et 4 font face à trois types de problèmes :
- le calcul d'intégrales de grandes dimensions ;
- l'optimisation de critères complexes ;
- la simulation de variables aléatoires.

Le problème du calcul d'intégrales multiples intervient en particulier lors des calculs d'estimateurs du type EAP ou d'estimation de marginales ; de même, le problème de l'optimisation de critères (ou de lois a posteriori) est lié au calcul d'estimateurs du type MAP.

En pratique, ces problèmes sont très souvent impossibles à résoudre du fait de leur complexité calculatoire. La plupart des méthodes classiques sont généralement peu satisfaisantes car elles sont difficiles à implémenter puisque les calculs doivent se faire dans un espace de dimension élevée.

Les méthodes MCMC (*Markov Chain Monte Carlo*) se présentent comme une alternative aux méthodes classiques dans le cas où ces dernières s'avèrent inefficaces. Le principe des méthodes MCMC est de créer une chaîne de Markov $\{\boldsymbol{\theta}^{(i)}\}$ $(i \in \{1, \ldots, I\})$ dont les éléments sont distribués asymptotiquement selon une distribution d'intérêt π, puis d'utiliser une estimation de Monte Carlo pour approcher la solution recherchée. Ainsi, sous certaines conditions, la distribution des éléments de la chaîne de Markov converge vers la distribution d'intérêt lorsque le nombre d'itérations tend vers l'infini.

Nous présentons ci-dessous l'intérêt des méthodes MCMC dans le cadre des trois problèmes présentés.

Calcul d'intégrales de grandes dimensions

Exceptions faites de cas académiques, il est souvent très difficile d'effectuer analytiquement le calcul d'intégrales dans le cadre de l'inférence bayésienne. En effet, dès que la dimension de $\boldsymbol{\theta}$ devient importante, les méthodes classiques d'intégration nécessite des temps de calculs rédhibitoires (généralement exponentiels en la dimension de l'espace) et la prise en compte de contraintes sur le domaine d'optimisation peut être très complexe.

Ainsi, le calcul d'intégrales du type :

$$\mathbb{E}_{\boldsymbol{\theta}}[f(\boldsymbol{\theta})] = \int_E f(\boldsymbol{\theta})\pi(\boldsymbol{\theta})d\boldsymbol{\theta}$$

peut être approché par l'estimateur suivant :

$$\widehat{\mathbb{E}}_{\boldsymbol{\theta}}[f(\boldsymbol{\theta})] = \frac{1}{I}\sum_{i=1}^{I} f\left(\boldsymbol{\theta}^{(i)}\right),$$

où $\{\boldsymbol{\theta}^{(i)}\}_{i\in\{1,\ldots,I\}}$ est une séquence i.i.d. de I échantillons distribués selon la densité $\pi(\boldsymbol{\theta})$. Si I est suffisamment grand, la loi forte des grands nombres implique la convergence presque sûre de $\widehat{\mathbb{E}}_{\boldsymbol{\theta}}[f(\boldsymbol{\theta})]$ vers $\mathbb{E}_{\boldsymbol{\theta}}[f(\boldsymbol{\theta})]$. De plus, si la variance $\mathrm{Var}_{\boldsymbol{\theta}}[f(\boldsymbol{\theta})]$ est finie, le théorème de la limite centrale s'applique, indiquant que la convergence de l'estimateur de Monte Carlo ne dépend ni de la forme de $f(\boldsymbol{\theta})$, ni de la dimension du domaine d'intégration, mais seulement du nombre d'itérations I [96].

Optimisation de critères complexes

De même que pour le calcul d'intégrales de grandes dimensions, il est souvent très difficile de calculer l'optimum d'une fonction de grande dimension dans le cadre de l'inférence bayésienne. En outre, les méthodes d'optimisation classiques (du type gradient) à base de procédures déterministes sont sensibles à l'initialisation et risquent de conduire à des optima locaux.

Ainsi, par exemple, les méthodes MCMC proposent plusieurs algorithmes pour calculer le MAP :

- si la maximisation se fait composante par composante, on utilise généralement une méthode graphique en visualisant l'approximation de la fonction à maximiser, c'est-à-dire l'histogramme des valeurs simulées ;
- des méthodes automatisées (méthodes de crible) permettent de procéder de la même façon : il s'agit d'algorithmes itératifs qui parcourent l'espace et conserve la plus grande valeur simulée de toutes les itérations ;
- le recuit simulé est une méthode itérative qui permet de converger vers le minimum global de $\pi(\boldsymbol{\theta})$. Pour échapper à l'attraction des minima locaux, la distribution π est lentement modifiée au cours des itérations par l'introduction d'un paramètre appelé *température* qui dépend de i. Il a été montré qu'une décroissance logarithmique de la température permet à la chaîne de Markov de converger presque sûrement vers le minimum global [93].

Simulation de variables aléatoires

Pour mettre en œuvre les méthodes MCMC dans les deux cas précédents, il faut être capable de simuler les échantillons des chaînes de Markov. Dans le cas où la distribution π est standard, de nombreux algorithmes ont été proposés et doivent être utilisés de préférence aux méthodes MCMC, beaucoup plus longues à converger. Parmi ces méthodes, citons la méthode d'inversion de la fonction de répartition, l'algorithme de Box et Muller, les méthodes de mélanges, d'acceptation-rejet, les méthodes générales pour les densités log-concaves, sans oublier les générateurs de variables uniformes. Une présentation plus détaillée de ces méthodes peut être trouvée dans [29, 93].

Dans le cas où ces méthodes ne peuvent être appliquées, les méthodes MCMC proposent plusieurs algorithmes dont la section A.3 présentent celles utilisées dans cette thèse.

A.2 Chaînes de Markov

On note π la distribution d'intérêt qui correspond à la loi de distribution des éléments de la chaîne de Markov et définie sur un espace E appelé espace d'état. La famille des parties mesurables de E est notée $\mathcal{E}$.

Une *chaîne de Markov* est une suite de variables aléatoires $\{\boldsymbol{\theta}^{(i)}\}$ ($i \in \{1,\ldots,I\}$) dont la distribution de probabilité de $\boldsymbol{\theta}^{(i)}$ sachant toutes les valeurs précédentes $\boldsymbol{\theta}_1,\ldots,\boldsymbol{\theta}_{k-1}$ est la même que la distribution de probabilité de $\boldsymbol{\theta}_k$ sachant seulement $\boldsymbol{\theta}_{k-1}$. Ainsi, pour tout $A \in \mathcal{E}$:

$$P(\boldsymbol{\theta}_k \in A|\boldsymbol{\theta}_1,\ldots,\boldsymbol{\theta}_{k-1}) = P(\boldsymbol{\theta}_k \in A|\boldsymbol{\theta}_{k-1}).$$

Une chaîne de Markov est définie par deux composantes :
– sa distribution initiale $p(\boldsymbol{\theta}^{(0)})$ (ou la valeur initiale $\boldsymbol{\theta}^{(0)}$ de la chaîne) ;
– et son noyau de transition :

$$T(\boldsymbol{\theta}, A) = P\left(\boldsymbol{\theta}^{(i+1)} \in A | \boldsymbol{\theta}^{(i)} = \boldsymbol{\theta}\right).$$

Afin de pouvoir définir le noyau de transition, il faut que la chaîne de Markov vérifie le condition d'homogénéité. Une chaîne de Markov est *homogène* si ses caractéristiques statistiques n'évoluent pas au cours des itérations, c'est-à-dire que la distribution π ne dépend pas de i. Un exemple classique de chaîne de Markov non homogène est l'algorithme du recuit simulé où, d'une itération à l'autre, les statistiques des transitions évoluent en fonction de la température.

Ainsi, $T(\boldsymbol{\theta}, \cdot)$ est la loi de la chaîne de Markov après une itération, sachant l'état précédent égal à $\boldsymbol{\theta}$. En posant $T^1(\boldsymbol{\theta}, A) = T(\boldsymbol{\theta}, A)$, on définit le noyau de i transitions par la relation de récurrence suivante :

$$T^i(\boldsymbol{\theta}, A) = \int_E T^{i-1}(\boldsymbol{\tau}, A) P(\boldsymbol{\theta}, d\boldsymbol{\tau}).$$

A.2.1 Propriétés des chaînes de Markov

Nous présentons maintenant les propriétés fondamentales que doivent vérifier les chaînes de Markov produites par les algorithmes MCMC présentés pour assurer leur convergence. Seules les notions nécessaires à la convergence des algorithmes de la section A.3 sont présentées ici ; une présentation plus détaillée peut être trouvée dans [93, 103, 109].

Invariance

Une loi π est *invariante* (ou *stationnaire*) pour le noyau de transition T si :

$$\forall A \in \mathcal{E}, \qquad \pi(A) = \int_E T(\boldsymbol{\theta}, A) \pi(d\boldsymbol{\theta}).$$

Cela signifie que si $\boldsymbol{\theta}^{(i)}$ est distribué suivant π, alors $\boldsymbol{\theta}^{(i+1)}$ et les suivants seront distribués marginalement selon π.

Lors de la construction d'algorithmes MCMC, il est souvent délicat de trouver des noyaux vérifiant cette propriété d'invariance. Mais une condition suffisante et souvent plus facile à démontrer est la condition de réversibilité.

On dit qu'un noyau de transition T est π-*réversible* si :

$$\forall (A, B) \in \mathcal{E} \times \mathcal{E}, \qquad \int_A T(\boldsymbol{\theta}, B) \pi(d\boldsymbol{\theta}) = \int_B T(\boldsymbol{\theta}, A) \pi(d\boldsymbol{\theta}),$$

ce qui signifie que la probabilité d'aller en B en partant de A est égale à la probabilité d'aller en A partant de B.

Irréductibilité

Soit φ une mesure de probabilité sur $(E, \mathcal{E})$. Une chaîne de Markov est dite φ-*irréductible* si :

$$\forall \boldsymbol{\theta} \in E, \forall A \in \mathcal{E}, \varphi(A) > 0 \qquad \Rightarrow \qquad \exists i \in \mathbb{N}^* / P(\boldsymbol{\theta}^{(i)} \in A | \boldsymbol{\theta}^{(0)} = \boldsymbol{\theta}) > 0.$$

Cette propriété signifie que tous les ensembles de probabilité non nulle peuvent être atteints à partir de tout point de départ dans l'espace en un nombre fini d'itérations.

Apériodicité

Une chaîne de Markov est dite *apériodique* si :

$$\forall\, \boldsymbol{\theta} \in E, \qquad \text{p.g.c.d.}\{m \geq 1 | P(\boldsymbol{\theta}^{(m)} = \boldsymbol{\theta} | \boldsymbol{\theta}^{(0)} = \boldsymbol{\theta})\} = 1.$$

L'hypothèse d'apériodicité élimine donc les noyaux qui induisent un comportement périodique des trajectoires et donc évite une comportement « déterministe » de la chaîne de Markov.

Récurrence au sens de Harris

Soit φ une mesure de probabilité sur $(E, \mathcal{E})$. Une chaîne ϕ-irréductible est dite *récurrente au sens de Harris* si :

$$\forall A \in \mathcal{E}, \forall\, \boldsymbol{\theta} \in A, \varphi(A) > 0 \qquad \Rightarrow \qquad P\left(\boldsymbol{\theta}_i \in A \text{ infiniment souvent} | \boldsymbol{\theta}_0 = \boldsymbol{\theta}\right) = 1.$$

La récurrence au sens de Harris signifie que presque sûrement, les trajectoires de la chaîne de Markov repassent une infinité de fois dans A.

Ergodicité

Les quatre propriétés précédentes permettent de définir la propriété d'ergodicité : une chaîne de Markov admettant une distribution invariante et qui de plus est irréductible, apériodique et récurrente au sens de Harris est dite *ergodique*.

A.2.2 Théorèmes de convergence

Avant de présenter les deux théorèmes de convergence que doivent vérifier les algorithmes MCMC utilisés dans cette thèse, il est nécessaire d'introduire la définition de la norme de la variation totale permettant de mesurer la distance entre deux mesures de probabilité.

La *norme de la variation totale* est définie par :

$$||\varphi_1 - \varphi_2||_{\text{VT}} = \sup_{A \in \mathcal{E}} |\varphi_1(A) - \varphi_2(A)|.$$

où φ_1 et φ_2 sont deux mesures de probabilité de $(E, \mathcal{E})$.

Théorème A.1
Soit une chaîne de Markov de noyau de transition T, π-invariante, irréductible et apériodique. Alors, pour π-presque tout point de départ $\boldsymbol{\theta}^{(0)}$,

$$\lim_{i \to +\infty} ||T^i(\cdot | \boldsymbol{\theta}^{(0)}) - \pi(\cdot)||_{\text{VT}} = 0.$$

La propriété de récurrence au sens de Harris permet d'obtenir une convergence plus forte car le théorème précédent peut alors s'appliquer pour tout point de départ de E.

Théorème A.2
Soit une chaîne de Markov de noyau de transition T, π-invariante, irréductible et apériodique récurrence au sens de Harris (elle est donc ergodique). Alors, pour tout point de départ $\boldsymbol{\theta}^{(0)} \in E$,

$$\lim_{i \to +\infty} ||T^i(\cdot | \boldsymbol{\theta}^{(0)}) - \pi(\cdot)||_{\text{VT}} = 0.$$

L'objectif est maintenant de construire des algorithmes MCMC qui produisent des chaînes de Markov vérifiant l'un ou l'autre de ces deux théorèmes. C'est le cas des algorithmes présentés dans la section suivante qui, au pire, vérifient la convergence pour π-presque tout point de départ.

A.3 Méthodes de Monte Carlo par chaînes de Markov

A.3.1 Algorithme de Metropolis-Hastings

L'algorithme de Metropolis-Hastings consiste à accepter avec un certain taux d'acceptation des échantillons générés à l'aide d'une *distribution instrumentale* (ou *candidate*) $q(\widetilde{\boldsymbol{\theta}}|\boldsymbol{\theta})$. Il peut être vu comme le dénominateur commun des méthodes MCMC classiques, qui n'en sont que des cas particuliers [3].

L'algorithme de Metropolis-Hastings suit le schéma suivant :

1. initialiser $\boldsymbol{\theta}^{(0)}$ (à une valeur éventuellement aléatoire) ;
2. pour chaque itération $i \in \{1, \ldots, I\}$:
   (a) simuler $\widetilde{\boldsymbol{\theta}} \sim q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(i-1)})$,
   (b) calculer le taux d'acceptation

$$\alpha = \min\left\{1, \frac{\pi(\widetilde{\boldsymbol{\theta}})}{\pi(\boldsymbol{\theta}^{(i-1)})} \frac{q(\boldsymbol{\theta}^{(i-1)}|\widetilde{\boldsymbol{\theta}})}{q(\widetilde{\boldsymbol{\theta}}|\boldsymbol{\theta}^{(i-1)})}\right\},$$

   (c) accepter $\widetilde{\boldsymbol{\theta}}$ avec la probabilité α :

$$\boldsymbol{\theta}^{(i)} = \begin{cases} \widetilde{\boldsymbol{\theta}} & \text{si } u < \alpha, \text{ avec } u \sim \mathcal{U}_{[0,1]} \text{ (acceptation)}, \\ \boldsymbol{\theta}^{(i-1)} & \text{sinon (rejet) ;} \end{cases}$$

3. $i \leftarrow i + 1$ et aller en 2.

La distribution instrumentale est soit disponible analytiquement (à une constante près), soit symétrique (c'est-à-dire telle que $q(\widetilde{\boldsymbol{\theta}}|\boldsymbol{\theta}) = q(\boldsymbol{\theta}|\widetilde{\boldsymbol{\theta}})$).

De plus, pour améliorer le taux d'acceptation α de l'algorithme, elle doit être simulable rapidement et doit être choisie en fonction de la distribution d'intérêt. En particulier, q doit être une bonne approximation de π et doit couvrir tout le support de π. En effet, si le support de la distribution instrumentale est trop petit, certaines zones du support de π ne seront pas explorées et donc aucun échantillon n'y sera simulé ; cependant, une distribution instrumentale à support trop large créerait trop de rejets, et ralentirait la convergence de l'algorithme. En général, les règles de choix des distributions instrumentales sont heuristiques.

Nous ne détaillerons pas ici les propriétés de convergence de l'algorithme de Metropolis-Hastings, qui peuvent être trouvées dans [3, 93].

L'algorithme de Metropolis-Hastings ne génère pas d'échantillons i.i.d., en particulier parce que la probabilité d'acceptation de $\widetilde{\boldsymbol{\theta}}$ dépend de $\boldsymbol{\theta}^{(i-1)}$, et peut être à l'origine d'apparitions multiples d'une même valeur puisque le rejet de $\widetilde{\boldsymbol{\theta}}$ conduit à reproduire $\boldsymbol{\theta}^{(i-1)}$ à l'itération i.

Lorsque la dimension de l'espace est trop importante, il peut être délicat de définir une distribution instrumentale q convenable, et celle-ci peut conduire à un taux d'acceptation faible. Dans ce cas, il est possible de subdiviser $\boldsymbol{\theta}$ en plusieurs vecteurs de tailles plus raisonnables et de les simuler au moyen de M densités instrumentales. Cet algorithme est appelé *algorithme de Metropolis-Hastings « un à la fois »* [3, 96].

En pratique, différents choix de q permettent de définir plusieurs variantes de l'algorithme de Metropolis-Hastings que nous présentons maintenant.

Algorithme de Metropolis-Hastings indépendant

Dans ce cas, la distribution instrumentale est indépendante de l'état $\boldsymbol{\theta}^{(i-1)}$, c'est-à dire que l'on a $q(\widetilde{\boldsymbol{\theta}}|\boldsymbol{\theta}^{(i-1)}) = q(\widetilde{\boldsymbol{\theta}})$, conduisant à une probabilité d'acceptation :

$$\alpha = \min\left\{1, \frac{\pi(\widetilde{\boldsymbol{\theta}})}{\pi(\boldsymbol{\theta}^{(i-1)})}\frac{q(\boldsymbol{\theta}^{(i-1)})}{q(\widetilde{\boldsymbol{\theta}})}\right\}.$$

Algorithme de Metropolis-Hastings à marche aléatoire

Contrairement à l'approche précédente, une seconde possibilité est de générer un candidat à partir de $\boldsymbol{\theta}^{(i-1)}$, plus précisément en y ajoutant une perturbation aléatoire. Autrement dit, la distribution instrumentale est la forme :

$$q(\widetilde{\boldsymbol{\theta}}|\boldsymbol{\theta}^{(i-1)}) = q(\widetilde{\boldsymbol{\theta}} - \boldsymbol{\theta}^{(i-1)}),$$

et la probabilité d'acceptation devient :

$$\alpha = \min\left\{1, \frac{\pi(\widetilde{\boldsymbol{\theta}})}{\pi(\boldsymbol{\theta}^{(i-1)})}\frac{q(\boldsymbol{\theta}^{(i-1)} - \widetilde{\boldsymbol{\theta}})}{q(\widetilde{\boldsymbol{\theta}} - \boldsymbol{\theta}^{(i-1)})}\right\}.$$

Dans le cas particulier où q est symétrique ($q(\widetilde{\boldsymbol{\theta}} - \boldsymbol{\theta}^{(i-1)}) = q(\boldsymbol{\theta}^{(i-1)} - \widetilde{\boldsymbol{\theta}})$), la probabilité d'acceptation se simplifie en :

$$\alpha = \min\left\{1, \frac{\pi(\widetilde{\boldsymbol{\theta}})}{\pi(\boldsymbol{\theta}^{(i-1)})}\right\}.$$

Cette forme de l'algorithme est appelée algorithme de Metropolis en référence à la méthode originelle de Metropolis [80].

La distribution instrumentale est généralement une loi uniforme, gaussienne ou de Student. Le problème est alors de choisir les paramètres de la distribution afin d'explorer le support de manière efficace, et tout particulièrement le paramètre d'échelle. En effet, s'il est trop faible, le taux d'acceptation sera élevé, mais la chaîne de Markov convergera trop lentement puisque ses pas seront trop petits. Au contraire, s'il est trop grand, l'algorithme va rejeter un grand nombre de propositions impliquant un faible taux d'acceptation.

Par exemple, considérons l'exemple académique suivant, où l'on cherche à simuler la loi normale $\mathcal{N}(0,1)$ par un algorithme de Metropolis-Hastings à marche aléatoire. La valeur initiale est fixée à $\theta^{(0)} = 2$, et on choisit une loi uniforme comme loi candidate. Les figures suivantes illustrent les 1000 premiers échantillons dans le cas de la loi $\mathcal{U}_{[-0,1;0,1]}$ (figure A.1(a)) et de la loi $\mathcal{U}_{[-100;100]}$ (figure A.1(b)), de plus grand paramètre d'échelle. Il est alors évident que ni l'une ni l'autre des lois candidates n'est adaptée : cette exemple met bien en lumière l'importance du paramètre d'échelle dans le cas d'un algorithme de Metropolis-Hastings à marche aléatoire.

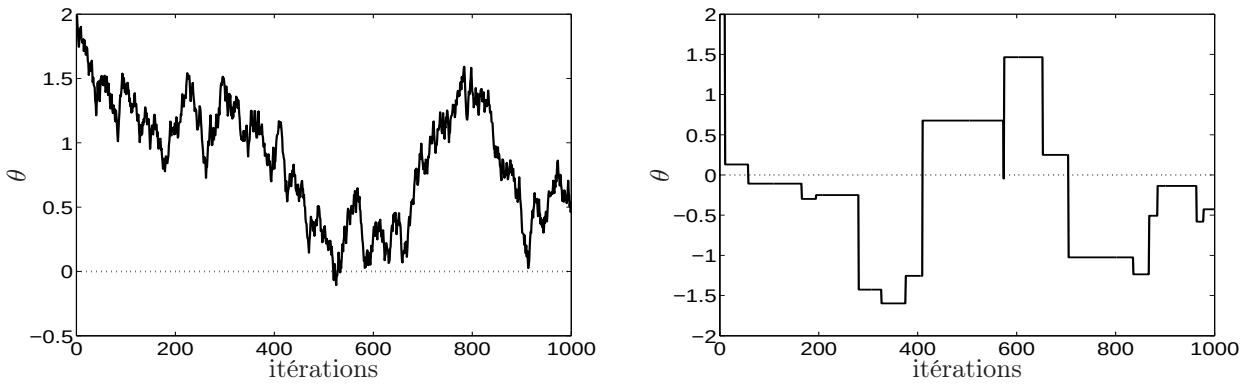
En général, ce réglage est heuristique. Dans [39] et [95], les auteurs recommandent par exemple de chercher à atteindre un taux d'acceptation de 0,5 pour les modèles de dimension 1 ou 2, et de 0,25 pour les modèles de dimension plus élevée. Une mise en œuvre de « calibration algorithmique » dans le cas multidimensionnel a été proposée par Müller [86], c'est-à-dire que des modifications successives du facteur d'échelle sont effectuées jusqu'à atteindre un taux d'acceptation voisin de 0,25.

Dans cette thèse, nous nous contentons de choisir une valeur du paramètre d'échelle afin que le taux d'acceptation suive les recommandations précédentes.

A.3.2 Échantillonneur de Gibbs

En considérant le vecteur $\boldsymbol{\theta} = \{\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_N\}$, l'échantillonneur de Gibbs est une approche intéressante de simulation dans le cas où les densités conditionnelles $\pi_n(\boldsymbol{\theta}_n|\boldsymbol{\theta}_{-n})$ sont connues et simulables.

Le schéma de simulation de l'échantillonneur de Gibbs est le suivant :



(a) Loi candidate : $\mathcal{U}_{[-0,1;0,1]}$. (b) Loi candidate : $\mathcal{U}_{[-100;100]}$.

FIG. A.1 – Exemple de simulation de la loi normale $\mathcal{N}(0,1)$ par un algorithme de Metropolis-Hastings à marche aléatoire.

1. initialiser $\boldsymbol{\theta}^{(0)} = (\boldsymbol{\theta}_1^{(0)}, ..., \boldsymbol{\theta}_N^{(0)})$ (à des valeurs éventuellement aléatoires) ;
2. pour chaque itération $i \in \{1, \ldots, I\}$: simuler

$$\begin{aligned}
\boldsymbol{\theta}_1^{(i)} &\sim \pi_1(\boldsymbol{\theta}_1 | \boldsymbol{\theta}_2^{(i-1)}, ..., \boldsymbol{\theta}_N^{(i-1)}), \\
\boldsymbol{\theta}_2^{(i)} &\sim \pi_2(\boldsymbol{\theta}_2 | \boldsymbol{\theta}_1^{(i)}, \boldsymbol{\theta}_3^{(i-1)}, ..., \boldsymbol{\theta}_N^{(i-1)}), \\
&\vdots \\
\boldsymbol{\theta}_N^{(i)} &\sim \pi_N(\boldsymbol{\theta}_N | \boldsymbol{\theta}_2^{(i)}, ..., \boldsymbol{\theta}_{N-1}^{(i)}) \ ;
\end{aligned}$$

3. $i \leftarrow i + 1$ et aller en 2.

L'échantillonneur de Gibbs est nécessairement multidimensionnel, avec un nombre de variables fixe. Notons cependant que les quantités $\boldsymbol{\theta}_n$ ne sont pas forcément des scalaires. De plus, comme elles sont simulées à partir de leur loi conditionnelle, le taux d'acceptation de l'algorithme est de un car tous les échantillons simulés sont acceptés.

Dans l'algorithme présenté, le balayage des éléments du vecteur $\boldsymbol{\theta}$ est déterministe (ici, les simulations sont effectuées dans l'ordre, du premier au N^{e} élément de $\boldsymbol{\theta}$) : c'est un exemple de balayage systématique. Mais il est possible d'effectuer un balayage aléatoire où l'ordre de simulation est choisi aléatoirement à chaque itération. L'étape 2 de l'échantillonneur de Gibbs devient alors :

générer une permutation σ sur $\{1, \ldots, N\}$,

simuler :

$$\begin{aligned}
\boldsymbol{\theta}_{\sigma_1}^{(i)} &\sim \pi_{\sigma_1}(\boldsymbol{\theta}_{\sigma_1} | \boldsymbol{\theta}_j^{(i-1)}, j \neq \sigma_1), \\
&\vdots \\
\boldsymbol{\theta}_{\sigma_N}^{(i)} &\sim \pi_{\sigma_N}(\boldsymbol{\theta}_{\sigma_N} | \boldsymbol{\theta}_j^{(i)}, j \neq \sigma_N).
\end{aligned}$$

L'intérêt de ce type de balayage est que la chaîne de Markov produite est réversible et de distribution stationnaire la loi a posteriori [93, p. 188] : c'est cette version qui est utilisée dans cette thèse.

Par ailleurs, l'échantillonneur de Gibbs peut être vu comme un cas particulier de l'algorithme de Metropolis-Hastings « un à la fois » pour lequel les distributions instrumentales de chaque composante $\boldsymbol{\theta}_n$ sont égales à leur distribution conditionnelle. Plus exactement, il correspond à la composition de N algorithmes de Metropolis-Hastings de probabilités d'acceptation uniformément égales à 1 [93, théorème 5.1].

Les propriétés de convergence de l'échantillonneur de Gibbs ne sont pas traitées ici, mais sont détaillées dans [93].

Dans cette thèse, nous utilisons des modèles bayésiens hiérarchiques (voir section 3.1.3), c'est-à-dire que la loi a posteriori peut s'écrire sous la forme :

$$\pi(\boldsymbol{\theta}) = \pi_1(\boldsymbol{\theta}_1)\pi_2(\boldsymbol{\theta}_2|\boldsymbol{\theta}_1)\dots\pi_N(\boldsymbol{\theta}_N|\boldsymbol{\theta}_1,\dots,\boldsymbol{\theta}_{N-1}) \ ;$$

faisant ainsi de l'échantillonneur de Gibbs une approche naturelle pour la simulation de $\boldsymbol{\theta} = \{\boldsymbol{\theta}_1,\dots,\boldsymbol{\theta}_N\}$.

Toutefois, cette facilité peut conduire dans certains cas à utiliser l'échantillonneur de Gibbs lorsque la loi a posteriori est impropre, c'est-à-dire d'intégrale divergente [3, 53, 93]. En effet, il peut arriver que les lois conditionnelles extraites de la loi a posteriori soient clairement simulables, mais que la loi a posteriori jointe ne soit pas intégrable. Cette question se pose dans les sections 3.3.3 et 4.2.3 où l'on choisit des lois a priori sur les variables du système pour construire ensuite les lois a posteriori conditionnelles qui seront échantillonnées par l'échantillonneur de Gibbs. Cependant, la complexité de la loi a posteriori jointe empêche de vérifier analytiquement son intégrabilité. L'intégrabilité peut toutefois être vérifiée en prenant garde de travailler avec des lois a priori propres. Or, comme nous avons fait le choix traditionnel d'utiliser des a priori non-informatifs (et impropres) sur certains hyperparamètres, nous avons plutôt cherché à obtenir des lois a posteriori conditionnelles intégrables en choisissant correctement ces lois a priori, sans quoi nous avons pu constaté, parfois, des divergences des chaînes de Markov.

A.4 Contrôle de convergence des algorithmes MCMC

L'implémentation d'une méthode MCMC nécessite de fixer la longueur I de la chaîne de Markov (c'est-à-dire le nombre d'itérations de la méthode). Par ailleurs, il convient d'écarter les I_0 premières itérations du calcul de l'estimateur car leur distribution est très différente de π. La période située aux itérations $i < I_0$ est appelée période de transition. Elle dépend principalement de la valeur initiale $\boldsymbol{\theta}^{(0)}$ et du noyau de transition de l'algorithme.

La méthode la plus simple pour déterminer ces deux variables consiste à analyser visuellement la trace de la chaîne de Markov et à déterminer soi-même les valeurs I et I_0. I_0 doit être fixé de telle sorte que tous les échantillons $\{\boldsymbol{\theta}^{(i)}\}$ tels que $i > I_0$ peuvent être considérés comme étant distribués suivant π. I est ensuite choisi de manière à avoir un nombre d'échantillons suffisant pour permettre une estimation de bonne qualité.

Dans cette thèse, nous utilisons cette procédure par souci de simplicité. Cependant, un grand nombre d'outils plus formels pour la détermination de I existent : on les regroupe sous le terme de « tests de convergence » (*convergence diagnostics*). Une partie de ces approches proposent également une valeur de I_0. Nous présentons maintenant quelques techniques courantes ; un panorama plus détaillé peut être trouvé dans [8, 25, 93, 109].

Les méthodes fondées sur l'analyse graphique de la chaîne de Markov sont parmi les plus simples car elles consistent en des méthodes empiriques effectuées sur les échantillons simulés. Ainsi, une condition nécessaire de convergence est la stationnarité d'une suite construite à partir des échantillons $\boldsymbol{\theta}^{(i)}$. On peut prendre par exemple la moyenne cumulée des échantillons ou le rapport entre la distribution d'intérêt et la distribution estimée.

L'approche de Raftery et Lewis [90] est très populaire (voir également [44, chapitre 7]). Elle consiste à calculer, à partir de la suite à deux états $z_i = \mathbb{1}_{\boldsymbol{\theta}^{(i)} \leq \underline{\boldsymbol{\theta}}}$ (où $\underline{\boldsymbol{\theta}}$ est une valeur arbitraire choisie dans le support de π), le nombre d'itérations I, la longueur de la période de transition I_0 et le pas minimum d'échantillonnage i_0 de la chaîne de Markov (la chaîne de Markov n'est échantillonnée que toutes les i_0 itérations).

La méthode de Mykland *et al.* [87] est une méthode de regénération, c'est-à-dire qu'il existe des instants $T_0 \leq T_1 \leq \ldots$ tels que, à chaque T_j, le futur de la chaîne de Markov est indépendant du passé et identiquement distribué. La convergence est observée graphiquement après avoir simulé I échantillons, puis en traçant T_j/T_I en fonction de j/I. L'un des principe de base est de n'enregistrer qu'une itération sur i_0 afin de construire une série d'échantillons approximativement i.i.d.

Enfin, d'autres méthodes font appel à la simulation de chaînes de Markov en parallèle. En effet, il n'est généralement pas possible de diagnostiquer la convergence d'une chaîne de Markov à partir d'une seule séquence, car celle-ci peut rester dans une région très influencée par l'initialisation, même après un grand nombre d'itérations. En simulant plusieurs chaînes, la variabilité et la dépendance aux conditions initiales sont réduites et la convergence peut être établie en comparant les estimations des quantités simulées de chaque chaîne. Une approche assez populaire est la méthode de Gelman et Rubin [40]. Cependant, le choix d'une analyse sur plusieurs chaînes ou d'une chaîne unique très longue reste un problème ouvert, et les études dans la littérature sont contradictoires quant à ce choix [44, 103, 109]. En effet, il n'est généralement pas possible d'effectuer des inférences valides si le nombre d'échantillons est trop faible : il faut donc que les chaînes de Markov soient suffisamment longues, ce qui augmente le coût de calcul. Par ailleurs, simuler une chaîne de Markov très longue permet de mieux explorer la distribution d'intérêt π : il y a plus de chances de trouver de nouveaux modes et de détecter des zones de faible probabilité.

Concluons cette section en précisant que, finalement, il n'est pas possible de dire avec certitude lorsqu'une chaîne de Markov (de longueur finie) a convergé, c'est-à-dire lorsqu'elle est représentative d'une distribution stationnaire [25, 93]. En outre, la prédominance d'une méthode sur une autre dépend fortement du modèle et du problème considérés. Ces méthodes restent encore exploratoires et ont parfois des fondements théoriques limités : l'étude des tests de convergence reste un problème ouvert [93].

Annexe B

Calcul de la loi a posteriori de $\mathbf{x}_n$

On rappelle l'expression de la loi a posteriori de $\mathbf{x}_n$ (section 3.3.3.1) :

$$p(\mathbf{x}_n|\mathbf{w}, \mathbf{x}_{-n}, \lambda, r_\mathbf{a}, s, r_\mathbf{b}) \propto p(\mathbf{w}|\mathbf{x}, s, r_\mathbf{b}) p(\mathbf{x}_n|\lambda, r_\mathbf{a}).$$

En considérant l'impulsion de Dirac comme une gaussienne de variance nulle r_0, la vraisemblance et la densité a priori s'écrivent respectivement :

$$\begin{aligned} p(\mathbf{w}|\mathbf{x}, s, r_\mathbf{b}) &= \frac{1}{(2\pi r_\mathbf{b})^{N/2}} \exp\left(-\frac{1}{2r_\mathbf{b}}||\mathbf{w} - \mathbf{H}\mathbf{x}||^2\right), \\ p(\mathbf{x}_n|\lambda, r_\mathbf{a}) &= (1-\lambda)\sqrt{\frac{2}{\pi r_0}} \exp\left(-\frac{\mathbf{x}_n^2}{2r_0}\right) \mathbb{1}_{\mathbb{R}^+}(\mathbf{x}_n) + \lambda\sqrt{\frac{2}{\pi r_\mathbf{a}}} \exp\left(-\frac{\mathbf{x}_n^2}{2r_\mathbf{a}}\right) \mathbb{1}_{\mathbb{R}^+}(\mathbf{x}_n). \end{aligned}$$

Ainsi, en regroupant ces deux expressions, on obtient :

$$p(\mathbf{x}_n|\mathbf{w}, \mathbf{x}_{-n}, \lambda, r_\mathbf{a}, s, r_\mathbf{b}) \propto \sum_{l=0,1} \lambda_l \sqrt{\frac{2}{\pi r_l}} \exp\left(-Q_{l,n}\right) \mathbb{1}_{\mathbb{R}^+}(\mathbf{x}_n),$$

avec :

$$\lambda_0 = 1 - \lambda, \qquad \lambda_1 = \lambda, \qquad r_1 = r_\mathbf{a} \qquad \text{et} \qquad Q_{l,n} = \frac{1}{2r_\mathbf{b}}||\mathbf{w} - \mathbf{H}\mathbf{x}||^2 + \frac{\mathbf{x}_n^2}{2r_l}.$$

Le terme $1/(2\pi r_\mathbf{b})^{N/2}$ a été supprimé puisqu'il est indépendant de l et n'est donc qu'une constante de l'expression.

En posant $\mathbf{e}_n = \mathbf{w} - (\mathbf{H}\mathbf{x} - \mathbf{H}_n\mathbf{x}_n)$ afin d'extraire le terme $\mathbf{x}_n$, $Q_{l,n}$ devient :

$$Q_{l,n} = \frac{1}{2r_\mathbf{b}}||\mathbf{e}_n - \mathbf{H}_n\mathbf{x}_n||^2 + \frac{\mathbf{x}_n^2}{2r_l},$$

qui, en développant, s'écrit :

$$Q_{l,n} = a\mathbf{x}_n^2 + b\mathbf{x}_n + c = a(\mathbf{x}_n + b/2a)^2 + (c - b^2/4a),$$

avec :

$$a = \frac{r_\mathbf{b} + r_l\mathbf{H}_n^T\mathbf{H}_n}{2r_l r_\mathbf{b}}, \qquad b = -\frac{\mathbf{e}_n^T\mathbf{H}_n}{r_\mathbf{b}}, \qquad \text{et} \qquad c = \frac{\mathbf{e}_n^T\mathbf{e}_n}{2r_\mathbf{b}}.$$

On obtient donc l'expression suivante :

$$Q_{l,n} = \frac{r_\mathbf{b} + r_l\mathbf{H}_n^T\mathbf{H}_n}{2r_l r_\mathbf{b}} \left(\mathbf{x}_n - \frac{r_l r_\mathbf{b}\mathbf{e}_n^T\mathbf{H}_n}{r_\mathbf{b}(r_\mathbf{b} + r_l\mathbf{H}_n^T\mathbf{H}_n)}\right)^2 + \left(\frac{\mathbf{e}_n^T\mathbf{e}_n}{2r_\mathbf{b}} - \left(\frac{\mathbf{e}_n^T\mathbf{H}_n}{r_\mathbf{b}}\right)^2 \frac{r_l r_\mathbf{b}}{2(r_\mathbf{b} + r_l\mathbf{H}_n^T\mathbf{H}_n)}\right),$$

et ainsi,

$$p(\mathbf{x}_n|\mathbf{w},\mathbf{x}_{-n},\lambda,r_{\mathbf{a}},s,r_{\mathbf{b}}) \propto \sum_{l=0,1} \frac{\lambda_{l,n}}{\sqrt{2\pi\rho_{l,n}}} \exp\left(-\frac{(\mathbf{x}_n-\mu_{l,n})^2}{2\rho_{l,n}}\right) \mathbb{1}_{\mathbb{R}^+}(\mathbf{x}_n),$$

où :

$$\begin{aligned}
\frac{\lambda_{l,n}}{\sqrt{2\pi\rho_{l,n}}} &= \lambda_l\sqrt{\frac{2}{\pi r_l}} \exp\left(-\frac{\mathbf{e}_n^T\mathbf{e}_n}{2r_{\mathbf{b}}} + \left(\frac{\mathbf{e}_n^T\mathbf{H}_n}{r_{\mathbf{b}}}\right)^2 \frac{r_l r_{\mathbf{b}}}{2(r_{\mathbf{b}}+r_l\mathbf{H}_n^T\mathbf{H}_n)}\right),\\
\frac{1}{2\rho_{l,n}} &= \frac{r_{\mathbf{b}}+r_l\mathbf{H}_n^T\mathbf{H}_n}{2r_l r_{\mathbf{b}}},\\
\mu_{l,n} &= \frac{r_l r_{\mathbf{b}}\mathbf{e}_n^T\mathbf{H}_n}{r_{\mathbf{b}}(r_{\mathbf{b}}+r_l\mathbf{H}_n^T\mathbf{H}_n)}.
\end{aligned}$$

Or, dans l'expression de $\lambda_{l,n}$, le terme $\exp\left(-\mathbf{e}_n^T\mathbf{e}_n/2r_{\mathbf{b}}\right)$ est constant et peut donc être supprimé. Les expressions précédentes se simplifient en :

$$\lambda_{l,n} = 2\lambda_l\sqrt{\frac{\rho_{l,n}}{r_l}} \exp\left(\frac{\mu_{l,n}^2}{2\rho_{l,n}}\right), \qquad \rho_{l,n} = \frac{r_l r_{\mathbf{b}}}{r_{\mathbf{b}}+r_l\mathbf{H}_n^T\mathbf{H}_n}, \qquad \mu_{l,n} = \frac{\rho_{l,n}}{r_{\mathbf{b}}}\mathbf{e}_n^T\mathbf{H}_n.$$

De plus, afin d'obtenir une densité de probabilité d'intégrale 1, il est nécessaire que $0 \leq \lambda_{l,n} \leq 1$. On redéfinit donc $\lambda_{l,n}$ en posant :

$$\lambda_{l,n} = \frac{\widetilde{\lambda}_{l,n}}{\widetilde{\lambda}_{l,n}+\widetilde{\lambda}_{1-l,n}} = \frac{1}{1+\widetilde{\lambda}_{1-l,n}/\widetilde{\lambda}_{l,n}}, \qquad \text{avec} \qquad \widetilde{\lambda}_{l,n} = 2\lambda_l\sqrt{\frac{\rho_{l,n}}{r_l}} \exp\left(\frac{\mu_{l,n}^2}{2\rho_{l,n}}\right).$$

Enfin, la loi a posteriori de $\mathbf{x}_n$ a pour expression :

$$p(\mathbf{x}_n|\mathbf{w},\mathbf{x}_{-n},\lambda,r_{\mathbf{a}},s,r_{\mathbf{b}}) \sim \lambda_{0,n}\mathcal{N}^+\big(\mu_{0,n},\rho_{0,n}\big) + \lambda_{1,n}\mathcal{N}^+\big(\mu_{1,n},\rho_{1,n}\big) \tag{B.1}$$

où les expressions des paramètres se simplifient en (rappelons que $r_0 = 0$) :

$$\begin{aligned}
\lambda_{0,n} &= \left[1+\frac{\lambda}{1-\lambda}\sqrt{\frac{\rho_{1,n}}{r_{\mathbf{a}}}} \exp\left(\frac{\mu_{1,n}^2}{2\rho_{1,n}}\right)\right]^{-1}, & \rho_{0,n} &= 0, & \mu_{0,n} &= 0,\\
\lambda_{1,n} &= \left[1+\frac{1-\lambda}{\lambda}\sqrt{\frac{r_{\mathbf{a}}}{\rho_{1,n}}} \exp\left(-\frac{\mu_{1,n}^2}{2\rho_{1,n}}\right)\right]^{-1}, & \rho_{1,n} &= \frac{r_{\mathbf{a}}r_{\mathbf{b}}}{r_{\mathbf{b}}+r_{\mathbf{a}}\mathbf{H}_n^T\mathbf{H}_n}, & \mu_{1,n} &= \frac{\rho_{1,n}}{r_{\mathbf{b}}}\mathbf{e}_n^T\mathbf{H}_n.
\end{aligned}$$

Annexe C

Simulation d'une distribution normale à support positif

On considère dans cette annexe le problème de la génération d'une variable aléatoire suivant une loi normale à support positif de paramètres μ et σ^2, que l'on notera :

$$x \sim \mathcal{N}^+(\mu, \sigma^2).$$

Nous présentons dans cette annexe une approche originale, qui a notamment fait l'objet de publications en congrès [75, 76]. Le programme Matlab est librement téléchargeable à l'adresse `www.iris.cran.uhp-nancy.fr/francais/idsys/Personnes/Perso_Mazet/rpnorm-fr.htm`.

La génération de variables aléatoires distribuées suivant une loi normale à support positif est un problème rencontré dans les chapitres 3 et 4, lors de la simulation des intensités des raies. Cette densité de probabilité correspond à une densité gaussienne de moyenne μ et de variance σ^2 tronquée en zéro et normalisée afin que l'intégrale soit égale à 1 (cf. figure C.1). Elle a pour expression :

$$f(x) = \frac{1}{C} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \mathbb{1}_{\mathbb{R}^+}(x) \qquad \text{où} \qquad C = \sqrt{\frac{\pi\sigma^2}{2}} \left[1 + \operatorname{erf}\left(\frac{\mu}{\sqrt{2\sigma^2}}\right)\right]$$

et erf est la fonction erreur définie telle que :

$$\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x \exp(-t^2) dt.$$

Notons que la valeur de C n'est pas nécessaire pour la mise en œuvre de l'algorithme.

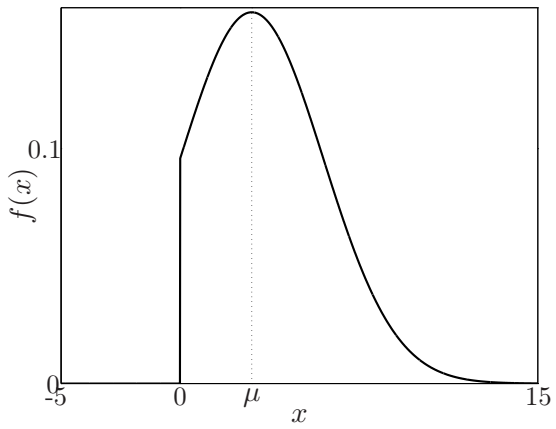


FIG. C.1 – Loi normale à support positif ($\mu = 3$, $\sigma^2 = 9$).

La loi normale tronquée $\mathcal{N}^+(\mu, \sigma^2)$ a pour moyenne et variance :

$$\mathbb{E}_x[x] = \mu + \sqrt{\frac{2\rho}{\pi}} \frac{\exp\left(-\mu^2/2\rho\right)}{1 + \operatorname{erf}\left(\mu/\sqrt{2\rho}\right)}, \qquad \operatorname{Var}_x[x] = \rho + \frac{\mu^2}{4} - \left(\frac{\mu}{2} + \sqrt{\frac{2\rho}{\pi}} \frac{\exp\left(-\mu^2/2\rho\right)}{1 + \operatorname{erf}\left(\mu/\sqrt{2\rho}\right)}\right)^2.$$

Remarque

Si f est tronquée à $t \neq 0$, la méthode peut être adaptée par simple translation de la variable aléatoire générée X : $Y = X + t$. De même, sans perdre en généralité, on peut se restreindre au cas $\sigma^2 = 1$, puisque tous les autres cas se déduisent par changement d'échelle $Y = X\sigma$: nous supposerons donc dans la suite que la loi est tronquée en zéro et que $\sigma^2 = 1$.

C.1 Méthodes existantes

La technique d'inversion de la fonction de répartition proposée par Devroye [29] ou Gelfand *et al.* [38] consiste à générer une variable uniforme $u \sim \mathcal{U}_{[0,1]}$, puis à calculer :

$$x = \mu + \sqrt{2\sigma^2}\operatorname{erf}^{-1}\left[u + \operatorname{erf}\left(\mu/\sqrt{2\sigma^2}\right)(u-1)\right].$$

Cette méthode a l'avantage de fournir une solution explicite mais suppose que les fonctions erf et erf^{-1} sont parfaitement calculables. En pratique, l'utilisation de ces fonctions peut se révéler inefficace car elles ne peuvent être qu'approchées numériquement. Dès lors, l'erreur d'approximation devient importante lorsque $|\mu|$ est trop grand (voir la section C.4 et les références [43, 92]).

Une autre approche consiste à utiliser un algorithme d'acceptation-rejet. La loi candidate la plus simple est évidemment la loi normale, mais celle-ci n'est adaptée que dans le cas où μ est suffisamment grand (cf. figure C.3). Au contraire, la loi exponentielle proposée par Robert [92] n'est bien adaptée que lorsque $-\mu$ est suffisamment grand (cf. figure C.3). En effet, lorsque $\mu \to -\infty$, la loi normale tronquée tend à ressembler à une loi exponentielle [43].

Finalement, comme la forme de la loi varie en fonction de μ et σ^2, Geweke propose d'utiliser un algorithme d'*acceptation-rejet mixte*, qui est une méthode basée sur l'algorithme d'acceptation-rejet mais utilisant plusieurs lois candidates, chacune étant adaptée aux formes particulières de la loi cible [43]. Ainsi, il faut déterminer quelle est la loi candidate générant le moins de rejets en fonction des paramètres de la loi cible (ici, μ et σ^2). Geweke propose deux lois candidates (les lois normale et exponentielle) pour simuler une loi normale tronquée unilatéralement et détermine les intervalles d'utilisation de ces lois de manière empirique.

En s'inspirant de l'algorithme d'acceptation-rejet mixte de Geweke, nous proposons une méthode où les intervalles d'utilisation des lois candidates sont déterminés en fonction du taux d'acceptation théorique, car une détermination empirique dépend du code utilisé et du logiciel de calcul. De plus, afin d'améliorer les performances de l'algorithme (en particulier au niveau de $\mu = 0$), nous proposons d'utiliser quatre lois candidates.

Notons que cette idée d'algorithme d'acceptation-rejet mixte peut s'adapter sur toute loi cible qui n'est pas directement échantillonnable par les techniques classiques, et dont la forme varie en fonction de ses paramètres.

C.2 Algorithme d'acceptation-rejet mixte

Dans cette section, nous présentons l'algorithme d'acceptation-rejet mixte dans le cas général, c'est-à-dire sans se restreindre au cas particulier de la simulation d'une loi normale tronquée. Pour cela, nous considérons tout d'abord l'algorithme d'acceptation-rejet classique, puis la détermination de la loi candidate et de la constante M et enfin l'algorithme d'acceptation-rejet mixte.

C.2.1 Algorithme d'acceptation-rejet classique

Rappelons que l'algorithme d'acceptation-rejet requiert la détermination d'une loi candidate g et d'une constante M telles que

$$\forall x \in S, \quad M \geq f(x)/g(x) \tag{C.1}$$

où S est le support de f. La loi candidate $g(x)$ est nécessairement non nulle sur S afin que M soit fini et donc que le taux d'acceptation moyen soit non nul (cf. équation (C.3)). L'algorithme d'acceptation-rejet suivant permet de générer une variable aléatoire x distribuée suivant f [29, 92, 93] :

1. générer $z \sim g(z)$ et $u \sim \mathcal{U}_{[0,1]}$;
2. calculer $\rho(z) = f(z)/Mg(z)$;
3. si $u \leq \rho(z)$: $x = z$ (acceptation),
   sinon : retour en 1 (rejet).

C.2.2 Détermination des lois candidates

Le choix des lois candidates est déterminant pour les performances de la méthode.

Tout d'abord, elles doivent être facilement échantillonnables, sinon l'algorithme d'acceptation-rejet perd son intérêt. En particulier, elles doivent être simulables avec un taux d'acceptation de 1, sinon le taux d'acceptation moyen correspondant doit être pris en compte dans le calcul du taux d'acceptation moyen global.

Précisons également que le choix des lois candidates doit prendre en compte la complexité de l'algorithme : certaines lois intéressantes a priori s'avéreront inadéquates car la détermination de M ou le calcul de ses paramètres peut être difficile, gourmand en temps de calcul, voire impossible ! Ainsi, on peut choisir des lois courantes (les lois normale et exponentielle) ou construire des lois particulières afin d'obtenir des lois candidates convenables (une loi normale couplée à la loi uniforme pour notre exemple).

C.2.3 Détermination de M

Pour le choix de M, n'importe quelle constante vérifiant l'équation (C.1) convient ; cependant, il est nécessaire que M soit la plus petite possible pour avoir un taux d'acceptation moyen élevé (équation (C.3)). La valeur optimale de M est donc :

$$M = \max_{x \in S} f(x)/g(x).$$

On peut alors déterminer le taux d'acceptation ρ :

$$\rho(x) = f(x)/Mg(x). \tag{C.2}$$

Enfin, le taux d'acceptation moyen $\overline{\rho} = \mathbb{E}_x[\rho(x)]$ définit une mesure d'efficacité de l'algorithme et permet de déterminer les intervalles d'utilisation des lois candidates :

$$\overline{\rho} \triangleq \int \rho(x)g(x)dx = \frac{1}{M}\int f(x)dx = \frac{1}{M}. \tag{C.3}$$

Remarquons que si $g(x)$ est proche de zéro, alors M devient très grand et le taux d'acceptation moyen diminue. L'efficacité de l'algorithme dépend donc de l'adéquation entre f et g ; en particulier, pour que M reste borné, il faut que g ait des queues plus lourdes que celle de f, sans toutefois que leur rapport soit trop grand, au risque d'augmenter M [93].

L'expression du taux d'acceptation dépend parfois des paramètres de la loi candidate, comme par exemple le paramètre α de la loi exponentielle (voir section suivante). Ces valeurs de paramètres sont à calculer afin de maximiser le taux d'acceptation.

C.2.4 Algorithme d'acceptation-rejet mixte

Parmi l'ensemble des lois candidates utilisées, une seule fournit le meilleur taux d'acceptation moyen pour des paramètres μ et σ^2 donnés. Il faut donc déterminer les intervalles sur ces paramètres pour déterminer quelle est la meilleure loi candidate à utiliser sur l'intervalle considéré. L'algorithme proposé est donc le même que l'algorithme d'acceptation-rejet avec une étape préliminaire qui consiste à choisir la meilleure loi candidate g_i en fonction des paramètres de f, puis à calculer la constante M correspondante :

1. sélectionner la loi candidate g_i donnant le plus grand taux d'acceptation moyen en fonction des paramètres de f puis calculer la constante M_i correspondante ;
2. générer $z \sim g_i(z)$ et $u \sim \mathcal{U}_{[0,1]}$;
3. calculer $\rho(z) = f(z)/M_i g_i(z)$;
4. si $u \leq \rho(z)$: $x = z$ (acceptation),
   sinon : retour en 3 (rejet).

C.3 Application à la simulation d'une loi normale tronquée

Quatre lois candidates (représentées figure C.2) sont proposées pour simuler la loi cible f :

① La loi normale :

$$g_1(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \ ; \tag{C.4}$$

② La loi normale couplée à la loi uniforme (définie pour $\mu \geq 0$) :

$$g_2(x) = \frac{\mathbb{1}_{\mathbb{R}^+}(x)}{\mu + \sqrt{\frac{\pi\sigma^2}{2}}} \begin{cases} 1 & \text{si } 0 \leq x < \mu, \\ \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) & \text{si } x \geq \mu \ ; \end{cases} \tag{C.5}$$

C'est une distribution définie sur $\mathbb{R}^+$, uniforme sur $[0, \mu[$ et qui suit la loi normale $\mathcal{N}(\mu, \sigma^2)$ sur $[\mu, +\infty[$,

③ La loi normale tronquée en la moyenne (définie pour $\mu \leq 0$) :

$$g_3(x) = \frac{2}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \mathbb{1}_{[\mu,+\infty[(x)} \ ; \tag{C.6}$$

④ La loi exponentielle [92] :

$$g_4(x) = \alpha \exp(-\alpha x) \mathbb{1}_{\mathbb{R}^+}(x) \tag{C.7}$$

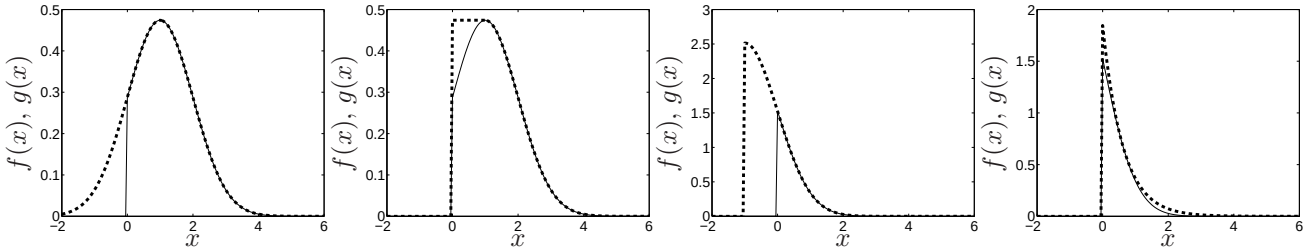
où la valeur de α correspond à la valeur qui maximise le taux d'acceptation moyen (cf. section C.5) :

$$\alpha = \left(\sqrt{\mu^2 + 4\sigma^2} - \mu\right) / 2\sigma^2. \tag{C.8}$$

La section C.6 propose des méthodes pour générer ces lois candidates.

Le choix de ces lois est motivé par le fait que les lois g_1 et g_4 permettent d'obtenir un taux d'acceptation moyen très important lorsque $|\mu| \gg 0$, et les lois g_2 et g_3 permettent d'améliorer le taux d'acceptation moyen autour de zéro (cf. figure C.3). Pour chacune de ces lois, nous pouvons calculer l'expression des constantes M (et ainsi les taux d'acceptation moyens) :

$$\begin{aligned} M_1 &= \sqrt{2\pi\sigma^2}/C, \\ M_2 &= \left(\mu + \sqrt{\pi\sigma^2/2}\right)/C, \\ M_3 &= \sqrt{2\pi\sigma^2}/2C, \\ M_4 &= \exp\left(\frac{\alpha}{2}(2\mu + \alpha\sigma^2)\right)/\alpha C. \end{aligned}$$



(a) Loi normale g_1 ($\mu = 1$). (b) Loi normale couplée à la loi uniforme g_2 ($\mu = 1$). (c) Loi normale tronquée en la moyenne g_3 ($\mu = -1$). (d) Loi exponentielle g_4 ($\mu = -1$).

FIG. C.2 – Lois candidates proposées pour la simulation d'une loi normale à support positif (— : $f(x)$, $\cdots$: $Mg(x)$).

Le calcul des taux d'acceptation $\rho(x)$ est alors direct :

$$\rho_1(x) = \begin{cases} 1 & \text{si } x \geq 0, \\ 0 & \text{sinon,} \end{cases}$$

$$\rho_2(x) = \begin{cases} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) & \text{si } 0 \leq x < \mu, \\ 1 & \text{si } x \geq \mu, \end{cases}$$

$$\rho_3(x) = \begin{cases} 1 & \text{si } x \geq 0, \\ 0 & \text{sinon,} \end{cases}$$

$$\rho_4(x) = \exp\left(-\frac{(x-\mu)^2}{2\sigma^2} - \frac{\alpha}{2}(2\mu - 2x + \alpha\sigma^2)\right).$$

Les taux d'acceptation moyens pour chacune des quatre lois candidates ont été représentés en fonction de μ figure C.3 (on rappelle que l'on considère $\sigma^2 = 1$).

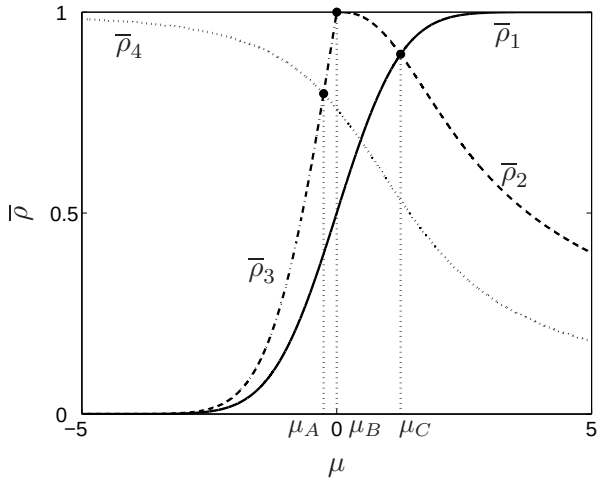


FIG. C.3 – Taux d'acceptation moyens pour les quatre lois candidates en fonction de μ ($\sigma^2 = 1$).

Nous pouvons alors trouver les expressions des trois points d'intersections qui délimitent les zones. Les calculs sont directs pour μ_B et μ_C, mais on ne peut que déterminer une valeur approchée de μ_A. En effet, d'après l'équation (C.8), on a :

$$\mu = (1 - \alpha^2\sigma^2)/\alpha,$$

et, en égalant les taux d'acceptation moyen des lois g_3 et g_4, on obtient :

$$\alpha\sigma \exp(\alpha^2\sigma^2/2) = e\sqrt{2/\pi}.$$

Cette dernière équation définit $\alpha\sigma$ sous une forme implicite. On peut en déduire que $\alpha\sigma$ s'exprime sous la forme

$$\alpha\sigma = \sqrt{W(2e^2/\pi)} \approx 1,1367,$$

où W est la fonction de Lambert, définie comme l'inverse multivaluée de la fonction $f(w) = we^w$ [24].

Ainsi :

$$\mu_A \approx -0,257\sigma, \qquad \mu_B = 0, \qquad \mu_C = \sigma\sqrt{\pi/2}.$$

Le taux d'acceptation moyen le plus faible correspond à $\mu = \mu_A$; il est égal à 0,797 environ, ce qui est un très bon taux d'acceptation : l'algorithme obtenu est donc très rapide car il génère très peu de rejet.

C.4 Simulations numériques

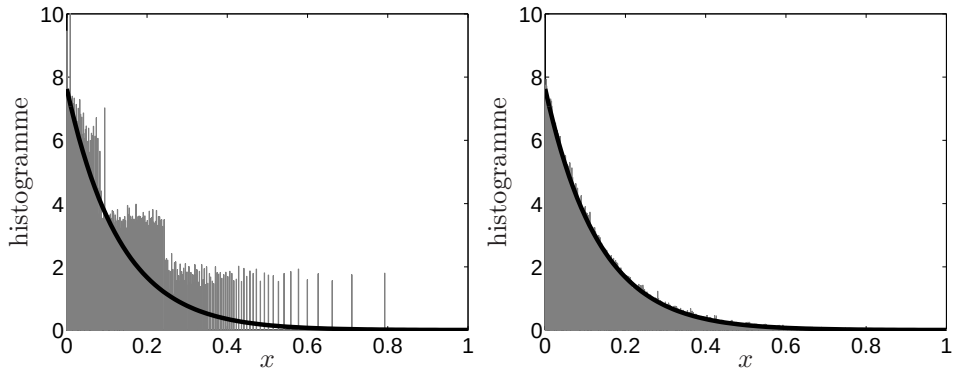
La fonction erf n'est pas explicite ; cependant, il en existe des approximations, comme par exemple celles de Cody [22] (qui correspond à la fonction de Matlab que nous utilisons ici) ou Abramowitz et Stegun [1] qui proposent une approximation exacte à 10^{-8} près.

Les variables uniformes et normales sont générées avec `rand` et `randn` respectivement.

C.4.1 Inversion de la fonction de répartition

La méthode d'inversion [29, 38], bien qu'elle soit explicite et que le taux d'acceptation soit de 1, s'avère inefficace dans certains cas pratiques ou $-\mu$ est trop grand (voir [43] et [93, p. 24]). Cela est dû à la fonction erf dont le calcul engendre des approximations créant alors des problèmes numériques.

Pour illustrer ce propos, considérons la génération de 100 000 variables normales positives avec le paramètre $\mu = -7,5$ grâce à la méthode d'inversion et l'approche proposée. Les histogrammes correspondants sont représentés figure C.4 (les éventuelles valeurs infinies générées par la méthode d'inversion n'ont pas été représentées sur l'histogramme). Il est clair que la méthode d'inversion s'avère inefficace.



(a) Méthode d'inversion. (b) Approche proposée.

FIG. C.4 – Histogrammes de 100 000 variables ($\mu = -7,5$, $\sigma^2 = 1$), et la distribution normale tronquée correspondante.

Cependant, les performances de la méthode d'inversion pourraient être améliorées en utilisant une meilleure approximation de la fonction erf, mais cela entraînerait un coup calculatoire conséquent. De plus, pour certaines valeurs des paramètres, la méthode d'inversion peut générer des valeurs négatives ou infinies. Si elles restent relativement peu nombreuses, le problème peut être résolu en générant de

nouvelles variables, mais, dans certains cas, la méthode peut ne donner que des valeurs inexploitables (par exemple avec $\mu = -8,5$) : la méthode d'inversion s'avère alors inappropriée.

C.4.2 Comparaison avec des algorithmes d'acceptation-rejet classiques

Comparons maintenant l'approche proposée avec trois algorithmes d'acceptation-rejet classiques utilisant les lois candidates g_1, g_2 et g_4. Nous comparons le temps de calcul pour différents μ, ce qui, en général, est relativement proportionnel au taux d'acceptation moyen. Les temps de simulation de 10 000 variables sont présentés dans le tableau C.1.

	approche proposée	acceptation-rejet		
		g_1	g_2	g_4
$\mu = -2$	0,36 s	11,094 s	($\mu < 0$!)	0,39 s
$\mu = 0$	0,406 s	0,485 s	0,422 s	0,454 s
μ =0,5	0,406 s	0,375 s	0,422 s	0,515 s
$\mu = 2$	0,281 s	0,297 s	0,531 s	0,813 s

TAB. C.1 – Comparaison des temps de calcul pour l'approche proposée et les algorithmes d'acceptation-rejet utilisant les lois candidates g_1, g_2 et g_4.

Dans tous les cas, les temps de calcul sont cohérents avec les taux d'acceptation moyens de la figure C.3, sauf pour $\mu = 0,5$ où la loi normale g_1 est plus rapide que la loi normale couplée à la loi uniforme g_2. Cela est dû au fait que générer une variable normale est beaucoup plus rapide que générer une variable suivant la loi g_2, et que le taux d'acceptation moyen de la méthode proposée n'est pas assez grand par rapport à celui de la loi normale. Mais comme cette différence dépend de la qualité de la programmation, du langage et du programme utilisés, il est difficile de déterminer les différents intervalles à partir des temps de calcul réels. C'est pourquoi nous préférons le faire à partir des taux d'acceptation moyens théoriques, ce qui permet tout de même d'obtenir de bonnes valeurs pour les intervalles. De toute manière, la différence est faible et particulière au cas $\mu = 0,5$.

C.5 Détails des calculs de M et ρ dans le cas de la loi exponentielle

Les calculs des constantes M et des taux d'acceptation $\rho(x)$ pour les lois g_1, g_2, et g_3 étant directs, ils ne seront pas traités ici.

Pour la loi exponentielle g_4, on peut calculer le rapport $f(x)/g(x)$ pour $x \geq 0$:

$$\frac{f(x)}{g(x)} = \frac{1}{C\alpha} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2} + \alpha x\right),$$

pour lequel le maximum est atteint pour $x = \alpha\sigma^2 + \mu$. En remplaçant cette valeur de x dans l'expression du rapport, on obtient l'expression de la meilleure constante M :

$$M = \frac{1}{C\alpha} \exp\left(\frac{\alpha}{2}\left(2\mu + \alpha\sigma^2\right)\right).$$

Le taux d'acceptation est alors :

$$\rho(x) = \exp\left(-\frac{(x-\mu)^2}{2\sigma^2} - \frac{\alpha}{2}\left(2\mu - 2x + \alpha\sigma^2\right)\right).$$

Il reste maintenant à calculer la valeur de α qui maximise le taux d'acceptation moyen $\overline{\rho} = 1/M$. Pour cela, on calcule la dérivée de $\overline{\rho}$:

$$\frac{\partial\overline{\rho}}{\partial\alpha} = C \exp\left(-\frac{\alpha}{2}\left(2\mu + \alpha\sigma^2\right)\right)(1 - \alpha\mu - \alpha^2\sigma^2).$$

Comme $\alpha > 0$, la racine valable de la dérivée est donc :

$$\alpha = \left(\sqrt{\mu^2 + 4\sigma^2} - \mu\right) / 2\sigma^2.$$

C.6 Simulation des lois candidates

Plusieurs méthodes générant un échantillon suivant la loi normale g_1 sont présentées dans [29, 89, 93]. En particulier, sous Matlab, nous utilisons la fonction `randn`.

Un échantillon x distribué suivant la loi g_2 est situé soit dans la partie uniforme (d'aire $\mathcal{A}_u$), soit dans la partie gaussienne (d'aire $\mathcal{A}_g$). Comme ces aires sont égales à :

$$\mathcal{A}_u = \frac{\mu}{\mu + \sqrt{\pi\sigma^2/2}} \qquad \text{et} \qquad \mathcal{A}_g = \frac{\sqrt{\pi\sigma^2/2}}{\mu + \sqrt{\pi\sigma^2/2}},$$

elles sont de somme unité. L'algorithme suivant permet de générer un échantillon x suivant la loi g_2 :

1. générer $u \sim \mathcal{U}_{[0,1]}$,
2. si $u < \mathcal{A}_u$, générer $v \sim \mathcal{U}_{[0,1]}$ puis calculer $x = \mu v$,
   sinon, générer $v \sim \mathcal{N}(0, \sigma^2)$ puis calculer $x = |v| + \mu$.

Pour générer un échantillon x suivant la loi g_4, il suffit de décaler de μ la valeur absolue d'une variable gaussienne centrée :

$$x = |y| + \mu \qquad \text{où} \qquad y \sim \mathcal{N}(0, \sigma^2).$$

Enfin, la méthode d'inversion de la fonction de répartition permet de simuler la loi exponentielle g_4 en générant $u \sim \mathcal{U}_{[0,1]}$, puis en calculant

$$x = -\ln(u)/\alpha.$$

Annexe D

Distributions de probabilité usuelles

Cette annexe donne les expressions des densités et de leurs deux premiers moments utilisées dans cette thèse. Pour plus d'information, le lecteur pourra se reporter à [93, annexe 1], principale référence de cette annexe.

Loi normale $\mathcal{N}(\mu, \sigma^2)$

$$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right),$$

$$\mathbb{E}_x[x] = \mu, \qquad \mathrm{Var}_x[x] = \sigma^2.$$

La fonction `randn` de Matlab permet de générer des variables aléatoires gaussiennes.

Loi normale à support positif $\mathcal{N}^+(\mu, \sigma^2)$

$$p(x) = \frac{1}{C} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \mathbb{1}_{\mathbb{R}^+}(x)$$

$$\text{où} \quad C = \sqrt{\frac{\pi\sigma^2}{2}} \left[1 + \mathrm{erf}\left(\frac{\mu}{\sqrt{2\sigma^2}}\right)\right],$$

$$\mathbb{E}_x[x] = \mu + \sqrt{\frac{2\sigma^2}{\pi}} \frac{\exp\left(-\mu^2/2\sigma^2\right)}{1 + \mathrm{erf}\left(\mu/\sqrt{2\sigma^2}\right)}, \qquad \mathrm{Var}_x[x] = \sigma^2 + \frac{\mu^2}{4} - \left(\frac{\mu}{2} + \sqrt{\frac{2\sigma^2}{\pi}} \frac{\exp\left(-\mu^2/2\sigma^2\right)}{1 + \mathrm{erf}\left(\mu/\sqrt{2\sigma^2}\right)}\right)^2.$$

L'annexe C propose une approche originale pour échantillonner une loi normale à support positif.

Loi gamma $\mathcal{G}(\alpha, \beta)$

$(\alpha, \beta > 0)$

$$p(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} \exp\left(-\beta x\right) \mathbb{1}_{\mathbb{R}^+}(x),$$

$$\mathbb{E}_x[x] = \frac{\alpha}{\beta}, \qquad \mathrm{Var}_x[x] = \frac{\alpha}{\beta^2}.$$

La fonction `rgamma` de la boîte à outils Stixbox [56] permet de générer des variables aléatoires distribuées suivant une loi gamma grâce à une méthode d'acceptation-rejet.

Loi inverse gamma $\mathcal{IG}(\alpha, \beta)$

$(\alpha, \beta > 0)$

$$p(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} \frac{\exp(-\beta/x)}{x^{\alpha+1}} \mathbb{1}_{\mathbb{R}^+}(x),$$

$$\mathbb{E}_x[x] = \frac{\beta}{\alpha - 1}, \qquad \mathrm{Var}_x[x] = \frac{\beta^2}{(\alpha-1)^2(\alpha-2)} \quad (\text{avec } \alpha > 2).$$

Pour générer des variables aléatoires distribuées suivant une loi inverse gamma, la fonction `rgamma` de la boîte à outils Stixbox [56] est utilisée, puisque la loi inverse gamma est la distribution de x^{-1} quand $x \sim \mathcal{G}(\alpha, \beta)$. La fonction `rgamma` utilise une méthode d'acceptation-rejet.

Loi exponentielle $\mathcal{E}(\alpha)$

$(\alpha > 0)$

$$p(x) = \alpha e^{-\alpha x} \mathbb{1}_{\mathbb{R}^+}(x),$$

$$\mathbb{E}_x[x] = \frac{1}{\alpha}, \qquad \mathrm{Var}_x[x] = \frac{1}{\alpha^2}.$$

La loi exponentielle est un cas particulier de la loi gamma : $\mathcal{G}(1, \alpha)$. En pratique, un échantillon suivant une loi exponentielle est généré grâce à la méthode d'inversion de la fonction de répartition. Ainsi, il suffit de générer $u \sim \mathcal{U}_{[0,1]}$, puis de calculer :

$$x = -\ln(u)/\alpha.$$

Loi beta $\mathcal{B}e(\alpha, \beta)$

$(\alpha, \beta > 0)$

$$p(x) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1} \mathbb{1}_{[0,1]}(x),$$

$$\mathbb{E}_x[x] = \frac{\alpha}{\alpha+\beta}, \qquad \mathrm{Var}_x[x] = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}.$$

La fonction `rbeta` de la boîte à outils Stixbox [56] permet de générer des variables aléatoires distribuées suivant une loi beta en utilisant une méthode d'inversion de la fonction de répartition. La fonction de répartition est calculée grâce à une fonction Matlab (`betainc`) définie dans [1, section 26.5].

Il est également possible de calculer $x_1/(x_1+x_2)$ où $x_1 \sim \mathcal{G}(\alpha, 1)$ et $x_2 \sim \mathcal{G}(\beta, 1)$ comme alternative à l'approche précédente pour éviter les problèmes numériques.

Loi de Bernoulli $\mathcal{B}er(\lambda)$

$$p(x) = (1-\lambda)\delta_0(x) + \lambda\delta_1(x),$$

$$\mathbb{E}_x[x] = \lambda, \qquad \mathrm{Var}_x[x] = \lambda - \lambda^2.$$

Pour générer une variable aléatoire suivant une loi de Bernoulli, il suffit de générer une variable aléatoire uniforme $u \sim \mathcal{U}_{[0,1]}$, puis de tester si elle est inférieure à λ :

$$x = \begin{cases} 1 & \text{si } u < \lambda, \\ 0 & \text{sinon.} \end{cases}$$

Annexe E

Déconvolution impulsionnelle par filtrage de Hunt et une méthode de seuillage

Cette annexe reproduit l'article « Sparse spike train deconvolution using the Hunt filter and a thresholding method » paru dans *IEEE Signal Processing Letters* (volume 11, numéro 5, mai 2004, p. 486–489). Cet article a été publié suite aux recherches effectuées pendant mon DEA et qui portaient sur la déconvolution impulsionnelle de signaux de décharges partielles.

486 IEEE SIGNAL PROCESSING LETTERS, VOL. 11, NO. 5, MAY 2004

Sparse Spike Train Deconvolution Using the Hunt Filter and a Thresholding Method

Vincent Mazet, David Brie, and Cyrille Caironi

***Abstract*—A new deconvolution method of sparse spike trains is presented. It is based on the coupling of the Hunt filter with a thresholding. We show that a good model for the probability density function of the Hunt filter output is a Gaussian mixture, from which we derive the threshold that minimizes the probability of errors. Based on an interpretation of the method as a maximum *a posteriori* (MAP) estimator, the hyperparameters are estimated using a joint MAP approach. Simulations show that this method performs well at a very low computation time.**

***Index Terms*—Bernoulli–Gaussian, Hunt filter, sparse spike train deconvolution, thresholding.**

I. Introduction

We consider the problem of sparse spike train deconvolution which is classically stated as follows (in the sequel, $\boldsymbol{v}_i$ will represent the ith element of vector $\boldsymbol{v}$ and $\boldsymbol{M}_{i,j}$ the element (i,j) of the matrix $\boldsymbol{M}$): given some observation sequence $\boldsymbol{y} \in \mathbb{R}^N$, find the sparse spike train sequence $\boldsymbol{x} \in \mathbb{R}^N$ such as

$$\boldsymbol{y} = \boldsymbol{H}\boldsymbol{x} + \boldsymbol{n}$$

where $\boldsymbol{H} \in \mathbb{R}^{N\times N}$ is a known impulse response matrix, and $\boldsymbol{n} \in \mathbb{R}^N$ a noise term assumed to be Gaussian and independent, identically distributed (i.i.d.) whose probability density function (pdf) is $p(\boldsymbol{n}) = \mathcal{N}(\boldsymbol{0}, r_n \boldsymbol{I}_N)$, r_n being the noise variance. To handle the sparse nature of $\boldsymbol{x}$, it is modeled as an i.i.d. Bernoulli–Gaussian (BG) sequence [1], [2], i.e.,

$$\forall k, \quad p(\boldsymbol{x}_k) = \lambda \mathcal{N}(0, r_x) + (1-\lambda)\delta(\boldsymbol{x}_k)$$

where r_x is the pulse variance, $\lambda \ll 1$ is the Bernoulli parameter (ratio between pulse and sample number), and δ is the Dirac mass centered on zero. In the sequel, $\boldsymbol{\theta} = \{\lambda, r_x, r_n\}$ gathers the hyperparameters of the problem.

Among the many applications of sparse spike deconvolution, we would like to mention partial discharge analysis, which has motivated this work [3]. The final goal was to develop algorithms that can be embedded to monitor the insulation of high-voltage AC motors. In addition, in this particular application, one has to process very large signals (typically $N = 2^{16}$ points). These are the reasons that have led us to develop an approach fast and simple to implement needing only low memory space.

On the one hand, combinatory algorithms like single mst likely replacement (SMLR) [1], [2], iterated window maximization (IWM) [4], or others given in [5] are well-suited to the problem. Although they give very satisfactory estimations, they are very slow and difficult to implement on embedded systems.

On the other hand, the Hunt filter [6] is very fast and simple to implement, but it does not yield satisfactory results, mainly because of the underlying prior, which assumes that the restored signal is Gaussian. To overcome this drawback, we propose to associate the Hunt filter with a thresholding procedure to actually restore a sparse spike train signal resulting in the so-called Hunt–threshold (HT) method.

The letter is organized as follows. Section II presents the HT method: first, the Hunt filter is briefly recalled; then, we derive the approximate pdf of the Hunt filter output from which we determine the threshold that minimizes the probability of errors. In Section III, based on an interpretation of the HT method as a maximum *a posteriori* (MAP) estimator, we propose to estimate the hyperparameters using a joint MAP (JMAP) approach [7]. In Section IV, the performances of the HT method are evaluated and compared to those of the SMLR of [1]. Finally, Section V gives the conclusions and perspectives of this letter.

II. HT Method

A. Hunt Filter

The Hunt filter [6] results from the Phillips and Twomey criterion minimization, which corresponds to the discrete time formulation of the Tikhonov criterion

$$\mathcal{J}_{\mathrm{PT}} = (\boldsymbol{y} - \boldsymbol{H}\boldsymbol{x})^T(\boldsymbol{y} - \boldsymbol{H}\boldsymbol{x}) + \alpha \boldsymbol{x}^T \boldsymbol{D}^T \boldsymbol{D}\boldsymbol{x}$$

where $\boldsymbol{D} = \boldsymbol{I}_N$ so as to favor the restoration of zero-value signal.

Minimizing $\mathcal{J}_{\mathrm{PT}}$ yields a first estimation $\tilde{\boldsymbol{x}}$ of $\boldsymbol{x}$

$$\tilde{\boldsymbol{x}} = \boldsymbol{G}\boldsymbol{y} \quad \text{where} \quad \boldsymbol{G} = (\boldsymbol{H}^T\boldsymbol{H} + \alpha \boldsymbol{I}_N)^{-1}\boldsymbol{H}^T. \tag{1}$$

The matrix $\boldsymbol{H}$ being supposed circulant, its (discrete) Fourier transform is a diagonal matrix. Thus, the algorithm may be efficiently implemented using the fast Fourier transform, $\tilde{\boldsymbol{x}}$ being then obtained by an inverse (discrete) Fourier transform. This results in the so-called Hunt filter.

A Bayesian interpretation [5] shows that the Hunt filter is equivalent to a MAP approach with a Gaussian prior: it yields $\alpha = r_n/r_x$, where r_x is the variance of the Gaussian signal to restore. In the sequel, we will also take $\alpha = r_n/r_x$ with r_x

Manuscript received April 17, 2003; revised July 3, 2003. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Dimitris A. Pados.

V. Mazet and D. Brie are with the Centre de Recherche en Automatique de Nancy, Centre National de la Recherche Scientifique, UMR 7039, Université Henri Poincaré, 54506 Vandoeuvre-lès-Nancy Cedex, France (e-mail: vincent.mazet@cran.uhp-nancy.fr; david.brie@cran.uhp-nancy.fr).

C. Caironi is with Alstom Moteurs, 54250 Champigneulles, France (e-mail: cyrille.caironi@tde.alstom.com).

Digital Object Identifier 10.1109/LSP.2004.826655

the variance of the pulses of the BG sequence to favor the pulse amplitude estimation.

B. Approximate PDF of $\tilde{\boldsymbol{x}}$

The pdf $p(\tilde{\boldsymbol{x}})$ of $\tilde{\boldsymbol{x}} = \boldsymbol{G}\boldsymbol{y} = \boldsymbol{G}\boldsymbol{H}\boldsymbol{x} + \boldsymbol{G}\boldsymbol{n}$ is needed to calculate the threshold t (Section II-C).

As $(\boldsymbol{G}\boldsymbol{n})_k = \sum_{i=1}^{N} \boldsymbol{G}_{k,i}\boldsymbol{n}_i$, we have

$$p\left((\boldsymbol{G}\boldsymbol{n})_k\right) = p\left(\sum_{i=1}^{N} \boldsymbol{G}_{k,i}\boldsymbol{n}_i\right) = \overset{\star}{\prod_{i=\{1,\ldots,N\}}} p(\boldsymbol{G}_{k,i}\boldsymbol{n}_i).$$

$\overset{\star}{\prod}$ is introduced to improve the readibility of the letter

$$\overset{\star}{\prod_{i=\{1,\ldots,N\}}} f_i(x) = [f_1 \star \ldots \star f_N](x).$$

Since $p(\boldsymbol{n}_k) = \mathcal{N}(0, r_n)$ and $\mathcal{N}(0,\sigma_1^2) \star \mathcal{N}(0,\sigma_2^2) = \mathcal{N}(0, \sigma_1^2 + \sigma_2^2)$, we get

$$p\left((\boldsymbol{G}\boldsymbol{n})_k\right) = \mathcal{N}\left(0, r_n \sum_{i=1}^{N} \boldsymbol{G}_{k,i}^2\right). \tag{2}$$

Similarly, since $p(\boldsymbol{x}_k) = \lambda\mathcal{N}(0, r_x) + (1-\lambda)\delta(\boldsymbol{x}_k)$ and $(\boldsymbol{G}\boldsymbol{H}\boldsymbol{x})_k = \sum_{i=1}^{N} (\boldsymbol{G}\boldsymbol{H})_{k,i}\boldsymbol{x}_i$, we have

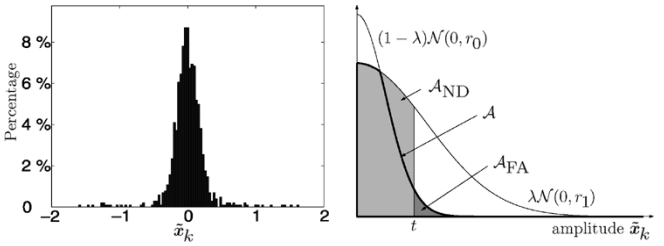
$$\begin{aligned} p\left((\boldsymbol{G}\boldsymbol{H}\boldsymbol{x})_k\right) &= \overset{\star}{\prod_{i=\{1,\ldots,N\}}} p\left((\boldsymbol{G}\boldsymbol{H})_{k,i}\boldsymbol{x}_i\right) \\ &= \overset{\star}{\prod_{i=\{1,\ldots,N\}}} \left[\lambda\mathcal{N}\left(0, r_x(\boldsymbol{G}\boldsymbol{H})_{k,i}^2\right) + (1-\lambda)\delta(\boldsymbol{x}_i)\right]. \end{aligned}$$

To simplify the problem, we now need two assumptions.

1) We suppose that the off-diagonal terms of $\boldsymbol{G}\boldsymbol{H}$ are negligible as compared to the diagonal terms. This is nothing but considering the values $(\boldsymbol{G}\boldsymbol{H}\boldsymbol{x})_k$ decorrelated. We note that the higher the SNR is (i.e., α decreases), the more valid this approximation is. Indeed, when α decreases, $\boldsymbol{G}\boldsymbol{H}$ tends to $\boldsymbol{I}_N$. With this assumption, we have $r_x(\boldsymbol{G}\boldsymbol{H})_{k,i} = 0 \; \forall \, k \neq i$. The validity of this approximation depends both on the SNR and the impulse response. Typically, for the impulse response used in this letter (see Section IV), the approximation is very realistic for a SNR greater than 15 dB. But if the frequency contents of the impulse response becomes more concentrated in low frequencies, the SNR should be greater;
2) We suppose that the diagonal terms of $\boldsymbol{G}\boldsymbol{G}^T$ and $\boldsymbol{G}\boldsymbol{H}\boldsymbol{H}^T\boldsymbol{G}^T$ are constant, i.e., we neglect the boundary effects.

Thus, considering that a Gaussian with zero variance is a Dirac mass, we get

$$p\left((\boldsymbol{G}\boldsymbol{H}\boldsymbol{x})_k\right) = \lambda\mathcal{N}\left(0, r_x(\boldsymbol{G}\boldsymbol{H})_{k,k}^2\right) + (1-\lambda)\delta(\boldsymbol{x}_k). \tag{3}$$



(a) (b)

Fig. 1. PDF of $\tilde{\boldsymbol{x}}$. (a) Histogram of $\tilde{\boldsymbol{x}}$ for 1024 samples. (b) Areas of false alarms and nondetections.

From (2) and (3), we obtain

$$\begin{aligned} p(\tilde{\boldsymbol{x}}_k) &= p\left((\boldsymbol{G}\boldsymbol{H}\boldsymbol{x})_k\right) \star p\left((\boldsymbol{G}\boldsymbol{n})_k\right) \\ &= \lambda\mathcal{N}\left(0, r_x(\boldsymbol{G}\boldsymbol{H})_{k,k}^2 + r_n\sum_{i=1}^{N}\boldsymbol{G}_{k,i}^2\right) \\ &\quad + (1-\lambda)\mathcal{N}\left(0, r_n\sum_{i=1}^{N}\boldsymbol{G}_{k,i}^2\right). \end{aligned}$$

As $r_n\sum_{i=1}^{N}\boldsymbol{G}_{k,i}^2 = r_n(\boldsymbol{G}\boldsymbol{G}^T)_{k,k}$ and $r_x(\boldsymbol{G}\boldsymbol{H})_{k,k}^2 = r_x\sum_{i=1}^{N}(\boldsymbol{G}\boldsymbol{H})_{k,i}^2 = r_x(\boldsymbol{G}\boldsymbol{H}\boldsymbol{H}^T\boldsymbol{G}^T)_{k,k}$, it follows that the pdf is a Gaussian mixture centered on zero

$$p(\tilde{\boldsymbol{x}}_k) = \lambda\mathcal{N}(0, r_1) + (1-\lambda)\mathcal{N}(0, r_0) \tag{4}$$

with

$$r_1 = r_x(\boldsymbol{G}\boldsymbol{H}\boldsymbol{H}^T\boldsymbol{G}^T)_{k,k} + r_n(\boldsymbol{G}\boldsymbol{G}^T)_{k,k} \tag{5}$$

$$r_0 = r_n(\boldsymbol{G}\boldsymbol{G}^T)_{k,k}, \qquad r_0 < r_1. \tag{6}$$

The histogram of $\tilde{\boldsymbol{x}}$ obtained on a 1024-sample signal [Fig. 1(a)] clearly shows that the Gaussian mixture is a good model for $p(\tilde{\boldsymbol{x}}_k)$.

C. Thresholding

The thresholding separates pulses [term $\lambda\mathcal{N}(0, r_1)$] from noise [term $(1-\lambda)\mathcal{N}(0, r_0)$]. The estimated signal may be expressed as $\widehat{\boldsymbol{x}}_k = \tilde{\boldsymbol{x}}_k$ if $|\tilde{\boldsymbol{x}}_k| > t$, $\widehat{\boldsymbol{x}}_k = 0$ otherwise. Performing such a (hard) thresholding, one can make two kinds of errors. *False alarms* correspond to wrongly detected pulses, while *nondetections* correspond to missed pulses. The threshold t is chosen to minimize the probability of error $\{p_{\mathrm{FA}} + p_{\mathrm{ND}}\}$ where p_{FA} and p_{ND} are the probabilities of false alarm and nondetection. Fig. 1(b) shows the positive part of the two Gaussians. The area $\mathcal{A}_{\mathrm{FA}}$ represents half the probability p_{FA} and $\mathcal{A}_{\mathrm{ND}}$ half the probability p_{ND}. It appears that $\mathcal{A}_{\mathrm{ND}} + \mathcal{A}_{\mathrm{FA}}$ is minimal and equal to $\mathcal{A}$ when t corresponds to the intersection point of the two Gaussians, which gives

$$t = \sqrt{\ln\left(\frac{\lambda}{1-\lambda}\sqrt{\frac{r_0}{r_1}}\right)\frac{2r_1r_0}{r_0 - r_1}}. \tag{7}$$

III. Hyperparameter Estimation

To address the hyperparameter estimation problem, we first need to interpret the HT method in a Bayesian framework. In [8] and [9], it is shown that the MAP estimator remains unchanged

for data varying in a neighborhood if and only if the prior pdf is nonsmooth (i.e., its derivative is discontinuous) in this neighborhood. This results in a thresholding effect of the MAP estimator under nonsmooth priors. Our prior being BG, the MAP estimation implies a thresholding, that allows us to use the HT method as a MAP estimator. The BG estimation $\widehat{\boldsymbol{x}}$ may be interpreted as the pointwise multiplication: $\forall k$, $\widehat{\boldsymbol{x}}_k = \tilde{\boldsymbol{x}}_k \widehat{\boldsymbol{q}}_k$, $\tilde{\boldsymbol{x}}$ corresponding to the Hunt filter output and $\widehat{\boldsymbol{q}}$ being an estimate of the i.i.d. Bernoulli sequence with parameter λ, controlling the occurrence of pulses in $\boldsymbol{x}$. So, we consider the MAP estimation of $(\boldsymbol{x}, \boldsymbol{q})$, which maximizes the following joint posterior pdf:

$$p(\boldsymbol{x}, \boldsymbol{q}|\boldsymbol{y}, \boldsymbol{\theta}) \propto p(\boldsymbol{y}|\boldsymbol{x}, \boldsymbol{q}, \boldsymbol{\theta}) p(\boldsymbol{x}|\boldsymbol{q}, \boldsymbol{\theta}) p(\boldsymbol{q}|\boldsymbol{\theta}) \tag{8}$$

where

- $p(\boldsymbol{y}|\boldsymbol{x}, \boldsymbol{q}, \boldsymbol{\theta}) = p(\boldsymbol{y}|\boldsymbol{x}, \boldsymbol{\theta}) = \mathcal{N}(\boldsymbol{H}\boldsymbol{x}, r_n \boldsymbol{I}_N)$ because $p(\boldsymbol{n}|\boldsymbol{\theta}) = \mathcal{N}(\boldsymbol{0}, r_n \boldsymbol{I}_N)$ and $\boldsymbol{y} = \boldsymbol{H}\boldsymbol{x} + \boldsymbol{n}$;
- $p(\boldsymbol{x}|\boldsymbol{q}, \boldsymbol{\theta}) = \mathcal{N}(\boldsymbol{0}, r_x \boldsymbol{Q})$ where $\boldsymbol{Q} = \mathrm{diag}\{\boldsymbol{q}\}$;
- $p(\boldsymbol{q}|\boldsymbol{\theta}) = \lambda^{N_q}(1-\lambda)^{N-N_q}$, N_q being the number of nonzero samples of $\boldsymbol{q}$.

To estimate the three hyperparameters gathered into $\boldsymbol{\theta} = \{\lambda, r_x, r_n\}$, we consider the following JMAP [7]:

$$\begin{aligned}(\tilde{\boldsymbol{x}}, \widehat{\boldsymbol{q}}, \widehat{\boldsymbol{\theta}}) &= \arg\max_{(\boldsymbol{x}, \boldsymbol{q}, \boldsymbol{\theta})} p(\boldsymbol{x}, \boldsymbol{q}, \boldsymbol{\theta}|\boldsymbol{y}) \\ &= \arg\max_{(\boldsymbol{x}, \boldsymbol{q}, \boldsymbol{\theta})} p(\boldsymbol{y}|\boldsymbol{x}, \boldsymbol{q}, \boldsymbol{\theta}) p(\boldsymbol{x}|\boldsymbol{q}, \boldsymbol{\theta}) p(\boldsymbol{q}|\boldsymbol{\theta})\end{aligned}$$

where $p(\boldsymbol{\theta})$ does not appear because it is considered as uniform (no *a priori* on the hyperparameters: it is in fact an ML estimator). This optimization problem is solved using an iterative procedure

$$\begin{cases} \left(\tilde{\boldsymbol{x}}^{(i)}, \widehat{\boldsymbol{q}}^{(i)}\right) = \arg\max_{(\boldsymbol{x}, \boldsymbol{q})} p\left(\boldsymbol{x}, \boldsymbol{q}, \lambda^{(i-1)}, r_x^{(i-1)}, r_n^{(i-1)}|\boldsymbol{y}\right) \\ \left(\widehat{\lambda}^{(i)}, \widehat{r_x}^{(i)}, \widehat{r_n}^{(i)}\right) = \arg\max_{(\lambda, r_x, r_n)} p\left(\boldsymbol{x}^{(i)}, \boldsymbol{q}^{(i)}, \lambda, r_x, r_n|\boldsymbol{y}\right). \end{cases}$$

The first optimization problem is approximatively solved using the HT method, and the hyperparameters are estimated using (8), with the assumption that $\boldsymbol{x}$ is a BG signal

$$\begin{aligned}\widehat{\lambda} &= \arg\max_{\lambda} p(\boldsymbol{q}|\lambda) = \frac{N_q}{N} \\ \widehat{r_x} &= \arg\max_{r_x} p(\boldsymbol{x}|\boldsymbol{q}, r_x) = \frac{1}{N_q}\|\widehat{\boldsymbol{x}}_k\|^2 \\ \widehat{r_n} &= \arg\max_{r_n} p(\boldsymbol{y}|\boldsymbol{x}, \boldsymbol{q}, r_n) = \frac{1}{N}\|\boldsymbol{y} - \boldsymbol{H}\widehat{\boldsymbol{x}}\|^2.\end{aligned}$$

Rather than using (5) and (6) to estimate r_0 and r_1, we propose to estimate them directly from $\tilde{\boldsymbol{x}}$ as

$$(\widehat{r_0}, \widehat{r_1}) = \arg\max_{(r_0, r_1)} p(\tilde{\boldsymbol{x}}|\boldsymbol{q}, r_0, r_1).$$

From (4), we get for all k

$$\begin{aligned}p(\tilde{\boldsymbol{x}}_k|\boldsymbol{q}, r_0, r_1) &= \boldsymbol{q}_k \mathcal{N}(0, r_1) + (1 - \boldsymbol{q}_k)\mathcal{N}(0, r_0) \\ &= \mathcal{N}\left(0, r_1 \boldsymbol{q}_k + r_0(1 - \boldsymbol{q}_k)\right).\end{aligned}$$

Then

$$p(\tilde{\boldsymbol{x}}|\boldsymbol{q}, r_0, r_1) = \mathcal{N}\left(\boldsymbol{0}, r_1 \boldsymbol{Q} + r_0(\boldsymbol{I}_N - \boldsymbol{Q})\right).$$

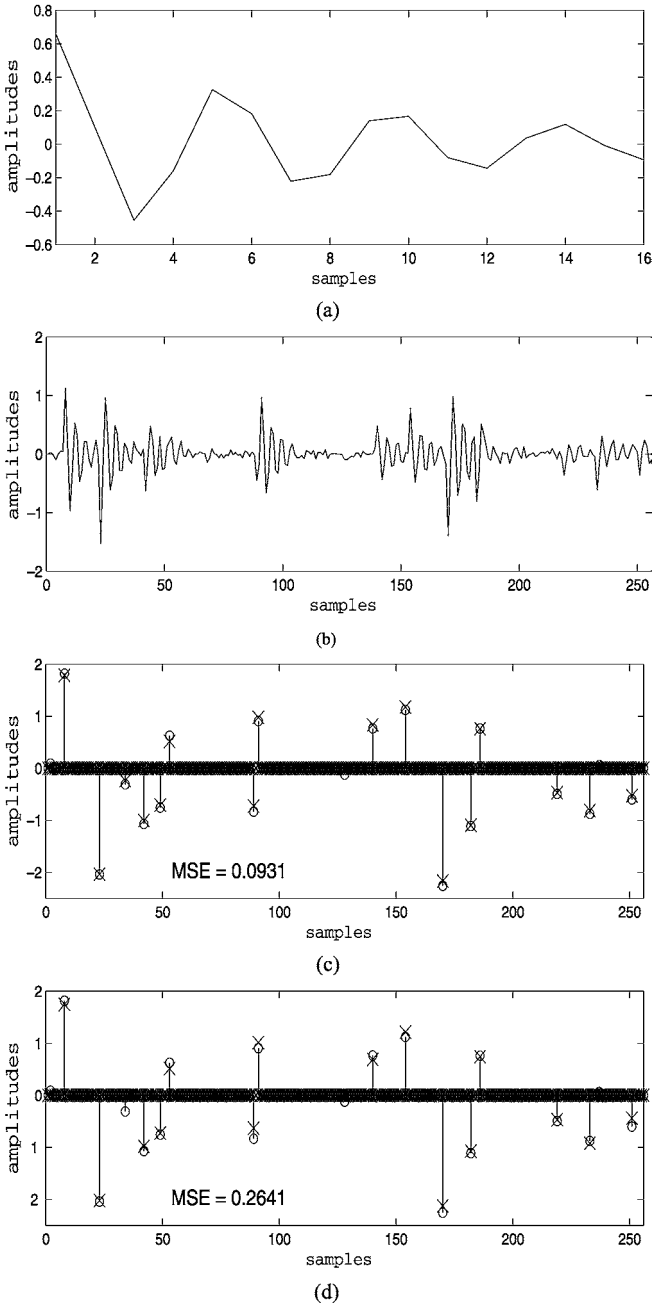


Fig. 2. SMLR and HT estimates. ∘ represents real pulses, and × estimated ones. MSE: mean square error on the detected pulses. (a) Impulse response ($\boldsymbol{h}$). (b) Data ($\boldsymbol{y}$). (c) SMLR estimate. (d) HT estimate.

Straightforward calculations yield the following explicit expressions:

$$\widehat{r_1} = \frac{1}{N_q}\sum_{k=1}^{N} \tilde{\boldsymbol{x}}_k^2 \widehat{\boldsymbol{q}}_k, \quad \widehat{r_0} = \frac{1}{N - N_q}\sum_{k=1}^{N} \tilde{\boldsymbol{x}}_k^2 (1 - \widehat{\boldsymbol{q}}_k).$$

This approach has been chosen because it leads to a faster algorithm and gives better results than those obtained by using (5) and (6).

λ, r_x, and r_n are initialized to arbitrary values, while r_0 and r_1 are initialized using (5) and (6). The convergence test is made by comparing the signal and hyperparameter values from one iteration to the next. The procedure stops when they differ no

TABLE I
COMPUTATION TIME, PERCENTAGE OF DETECTED PULSES (DP), AND FALSE ALARMS (FA) FOR SEVERAL SNR

SNR = 20 dB	SMLR	23.39 s	92.0 % DP	2.0 % FA
	HT	0.16 s	87.0 % DP	5.0 % FA
SNR = 15 dB	SMLR	26.94 s	88.17 % DP	2.15 % FA
	HT	0.19 s	81.72 % DP	4.3 % FA
SNR = 10 dB	SMLR	37.59 s	69.07 % DP	3.09 % FA
	HT	0.23 s	61.85 % DP	8.24 % FA
SNR = 5 dB	SMLR	105.64 s	42.9 % DP	3.7 % FA
	HT	0.31 s	40.2 % DP	1.9 % FA

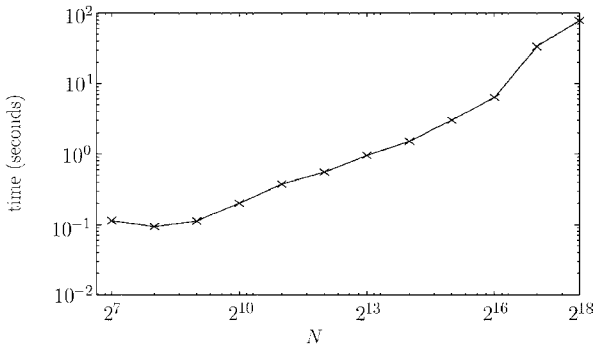


Fig. 3. Mean computation time.

more than 1%. Extensive simulations have shown the very satisfactory behavior of the procedure.

IV. SIMULATIONS RESULTS

Some simulations were carried out to assess the performances of the HT method and to compare them to those of the SMLR of [1]. The two methods were implemented in Matlab 6.5 on a Pentium III at 800 MHz. Fig. 2 shows the results achieved by the two methods on a 256-sample signal with a 16-sample impulse response and the following real hyperparameters: $\lambda = 0.08$, $r_x = 1.5$, SNR $= 15$ dB. Comparing the results, it appears that the SMLR performs slightly better than the HT method. Table I gives the statistics obtained on ten simulations. They confirm the superiority of the SMLR at the price of a very high computation time as compared to the HT method. We also note that the performances of the two methods become comparable as the SNR increases. Fig. 3 shows the mean computation time evolution of the HT method as a function of N, confirming that the computational burden remains reasonable, even for very large signals.

V. CONCLUSION

The HT method is a sparse spike train deconvolution consisting in coupling the Hunt filter with a thresholding. We show that when the signal $\boldsymbol{x}$ is BG, a Gaussian mixture is a good model for the pdf of the Hunt filter output. From this result, we derive the threshold that minimizes the probability of errors. We then propose to use the HT method as a MAP estimator and to jointly estimate the hyperparameters of the problem. Simulations show that the method performs well at a very low computation time, making this approach very well-suited to process very large signals and to be implemented on embedded systems. Future work will be directed to find which global criterion is minimized.

REFERENCES

[1] F. Champagnat, Y. Goussard, and J. Idier, "Unsupervised deconvolution of sparse spike trains using stochastic approximation," *IEEE Trans. Signal Processing*, vol. 44, pp. 2988–2998, Dec. 1996.

[2] J. J. Kormylo and J. M. Mendel, "Maximum likelihood detection and estimation of Bernoulli–Gaussian processes," *IEEE Trans. Inform. Theory*, vol. IT-28, pp. 482–488, May 1982.

[3] V. Mazet, D. Brie, and C. Caironi, "Déconvolution impulsionnelle par filtre de Hunt et seuillage," in *Proc. GRETSI*, Sept. 2003.

[4] K. F. Kaaresen, "Deconvolution of sparse spike trains by iterated window maximization," *IEEE Trans. Signal Processing*, vol. 45, pp. 1173–1183, May 1997.

[5] J. Idier, *Approche bayésienne pour les problèmes inverses*. Paris, France: Hermès Science, 2001.

[6] B. R. Hunt, "The inverse problem of radiography," *Math. Biosci.*, vol. 8, pp. 161–179, 1970.

[7] A. Mohammad-Djafari, "Joint estimation of parameters and hyperparameters in a Bayesian approach of solving inverse problems," in *Proc. ICASSP*, Munich, Germany, Apr. 1997, pp. 2837–2840.

[8] P. Moulin and J. Liu, "Analysis of multiresolution image denoising schemes using generalized Gaussian and complexity priors," *IEEE Trans. Inform. Theory*, vol. 45, pp. 909–919, Apr. 1999.

[9] M. Nikolova, "Local strong homogeneity of a regularized estimator," *SIAM J. Appl. Math.*, vol. 61, no. 2, pp. 633–658, 2000.

Bibliographie

[1] M. Abramowitz et I.A. Stegun, éditeurs. *Handbook of Mathematical Functions*. Dover Publications, 1974.

[2] H. Akaike. A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19 p. 716–723, 1974.

[3] C. Andrieu, A. Doucet et P. Duvaut. Méthodes de Monte Carlo par chaînes de Markov appliquées au traitement du signal. rapport interne ETIS-URA 2235 97-n°03, ENSEA, 1997.

[4] D. Bertrand et É. Dufour, éditeurs. *La spectroscopie infrarouge et ses applications analytiques*. Éditions Tec & Doc, 2000.

[5] J. Besag. Statistical analysis of dirty pictures. *Journal of the Royal Statistical Society, Series B*, 48(3) p. 259–279, 1986.

[6] S. Bourguignon et H. Carfantan. Bernoulli-Gaussian spectral analysis of unevenly spaced astrophysical data. Dans *Actes de IEEE Workshop Statistical Signal Processing*, papier #238, Bordeaux, France, 17–20 juillet 2005.

[7] E.G. Brame et J.G. Grasselli, éditeurs. *Infrared and Raman Spectroscopy*, volume A de *Practical Spectroscopy*. Marcel Dekker Inc., 1976.

[8] S.P. Brooks et G.O. Roberts. Convergence assesment techniques for Markov chain Monte Carlo. *Statistics and Computing*, 8 p. 319–335, 1998.

[9] T.T. Cai, D. Zhang et D. Ben-Amotz. Enhanced chemical classification of Raman images using multiresolution wavelet transformation. *Applied Spectroscopy*, 55(9) p. 1124–1130, 2001.

[10] G. Celeux. Contribution to the discussion of paper by Richardson and Green (1997). *Journal of the Royal Statistical Society, Series B*, 59 p. 775–776, 1997.

[11] G. Celeux. Bayesian inference for mixtures : the label switching problem. Dans R. Payne et P. Green, éditeurs, *COMPSTAT 98*, p. 227–232. Physica-Verlag, 24–28 août 1998.

[12] G. Celeux et J. Diebolt. The SEM algorithm : a probabilistic teacher algorithm derived from the EM algorithm for the mixture problem. *Computationnal Statistics Quaterly*, 2 p. 73–82, 1985.

[13] G. Celeux et J. Diebolt. A stochastic approximation type EM algorithm for the mixture problem. *Stochastics and Stochastics Reports*, 41 p. 119–134, 1992.

[14] G. Celeux, M. Hurn et C.P. Robert. Computational and inferential difficulties with mixture posterior distributions. *Journal of the American Statistical Association*, 95(451) p. 957–970, 2000.

[15] B. Chalmond. An iterative Gibbsian technique for reconstruction of m-ary images. *Pattern recognition*, 22(6) p. 747–761, 1989.

[16] F. Champagnat. *Déconvolution impulsionnelle et extensions pour la caractérisation des milieux inhomogènes en échographie*. Thèse de doctorat, Université de Paris Sud, décembre 1993.

[17] F. Champagnat, Y. Goussard et J. Idier. Unsupervised deconvolution of sparse spike trains using stochastic approximation. *IEEE Transactions on Signal Processing*, 44(12) p. 2988–2998, décembre 1996.

[18] P. Charbonnier. *Reconstruction d'image : régularisation avec prise en compte des discontinuités*. Thèse de doctorat, Université de Nice-Sophia Antipolis, septembre 1994.

[19] P. Charbonnier, L. Blanc-Féraud, G. Aubert et M. Barlaud. Two deterministic half-quadratic regularization algorithms for computed imaging. Dans *Actes de IEEE International Conference on Image Processing (ICIP)*, volume 2, p. 168–172, Austin, Texas, États-Unis, 13-16 novembre 1994.

[20] Q. Cheng, R. Chen et T.-H. Li. Simultaneous wavelet estimation and deconvolution of reflection seismic signals. *IEEE Transactions on Geoscience and Remote Sensing*, 34(2) p. 377–384, mars 1996.

[21] C.-Y. Chi, J.M. Mendel et D. Hampson. A computationally fast approach to maximum-likelihood deconvolution. *Geophysics*, 49(5) p. 550–565, 1984.

[22] W.J. Cody. Rational Chebyshev approximations for the error function. *Mathematics of Computation*, 22 p. 631–638, 1969.

[23] N.B. Colthup, L.H. Daly et S.E. Wiberley. *Introduction to Raman and infrared spectroscopy.* Academic Press, 1964.

[24] R.M. Corless, G.H. Gonnet, D.E.G. Hare, D.J. Jeffrey et D.E. Knuth. On the Lambert W function. *Advances in Computational Mathematics*, 5 p. 329–359, 1996.

[25] M.K. Cowles et B.P. Carlin. Markov chain Monte Carlo convergence diagnostics : a comparative review. *Journal of the American Statistical Association*, 91 p. 883–904, 1996.

[26] G. Demoment. Déconvolution des signaux. Cours de l'École Supérieure d'Électricité (Supélec), 1986.

[27] A.P. Dempster, N.M. Laird et D.B. Rubin. Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society, Series B*, 39 p. 1–38, 1977.

[28] U. Depczynski, K. Jetter, K. Molt et A. Niemöller. The fast wavelet tranform on compact intervals as a tool in chemometrics. I. Mathematical background. *Chemometrics and Intelligent Laboratory Systems*, 39 p. 19–27, 1997.

[29] L. Devroye. *Non-uniform random variate generation.* Springer-Verlag, New York, 1986[1].

[30] J.G. Dias et M. Wedel. An empirical comparison of EM, SEM and MCMC performance for problematic Gaussian mixture likelihoods. *Statistics and Computing*, 14 p. 323–332, 2004.

[31] J. Diebolt et C.P. Robert. Estimation of finite mixture distributions through Bayesian sampling. *Journal of the Royal Statistical Society, Series B*, 2(56) p. 363–375, 1994.

[32] B. Dippel. *Grundlagen, Techniken und Anwendungen zur Raman-Spektroskopie*, 2001. www.raman.de.

[33] P.H.C. Eilers. Parametric time warping. *Analytical Chemistry*, 76(2) p. 404–411, 2004.

[34] R. Fischer et V. Dose. Analysis of mixtures in physical spectra. Dans *Actes de ISBA 2000 (6th world meeting of the International Society for Bayesian Analysis)*, Heraklion, Crête, Grèce, 28 mai–1er juin 2000.

[35] R. Fischer, K.M. Hanson, V. Dose et W. von der Linden. Background estimation in experimental spectra. *Physical Review E*, 61 p. 1152–1161, 2000.

[36] W.J. Fitzgerald. Markov chain Monte Carlo methods with applications to signal processing. *Signal Processing*, 81 p. 3–18, 2001.

[37] S. Gautier, J. Idier, F. Champagnat, A. Mohammad-Djafari et B. Lavayssière. Traitement d'échogrammes ultrasonores par déconvolution aveugle. Dans *Actes de GRETSI 1997*, p. 1431–1434, 1997.

[38] A.E. Gelfand, A.F.M. Smith et T.-M. Lee. Bayesian analysis of constrained parameter and truncated problems using Gibbs sampling. *Journal of the American Statistical Association*, 87 p. 523–532, 1992.

[1] téléchargeable à l'adresse : http ://jeff.cs.mcgill.ca/~luc/rnbookindex.html

[39] A. Gelman, G.O. Roberts et W.R. Gilks. Efficient metropolis jumping rules. Dans J.M. Bernardo, J.O. Berger, A.P. Dawid et A.F.M. Smith, éditeurs, *Bayesian Statistics 5*, p. 599–608. Oxford University Press, 1996.

[40] A. Gelman et D.B. Rubin. Inference from iterative simulation using multiple sequences. *Statistical Science*, 7 p. 457–511, 1992. Avec discussion.

[41] D. Geman et G. Reynolds. Constrained restoration and the recovery of discontinuities. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(3) p. 367–383, 1992.

[42] D. Geman et C. Yang. Nonlinear image recovery with half-quadratic regularization. *IEEE Transactions on Image Processing*, 4 p. 932–946, 1995.

[43] J. Geweke. Efficient simulation from the multivariate normal and Student-t distributions subject to linear constraints. Dans *Computing Science and Statistics, proceedings of the 23rd Symposium on the Interface*, p. 571–578, 1991.

[44] W.R. Gilks, S. Richardson et D.J. Spiegelhalter, éditeurs. *Markov chain Monte Carlo in practice*. Chapman & Hall, 1996.

[45] R.P. Goehner. Background subtract subroutine for spectral data. *Analytical Chemistry*, 50(8) p. 1223–1225, 1978.

[46] P.J. Green. Reversible jump Markov chain Monte Carlo computation and Bayesian model determination. *Biometrika*, 82(4) p. 711–732, 1995.

[47] S. Gulam Razul, W.J. Fitzgerald et C. Andrieu. Bayesian deconvolution in nuclear spectroscopy using RJMCMC. Dans *Actes de European Signal Processing Conference (EUSIPCO)*, volume 2, p. 1309–1312, 2002.

[48] S. Gulam Razul, W.J. Fitzgerald et C. Andrieu. Bayesian model selection and parameter estimation of nuclear emission spectra using RJMCMC. *Nuclear Instruments and Methods in Physics Research A*, 497 p. 492–510, 2003.

[49] N.M. Haan. *Statistical Models and Algorithms for DNA Sequencing*. Thèse de doctorat, Université de Cambridge, juin 2001.

[50] N.M. Haan et S.J. Godsill. Bayesian model for DNA sequencing. Dans *Actes de IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, volume 4, p. 4020–4023, Orlando, Floride, États-Unis, 2002.

[51] J. Hadamard. Sur les problèmes aux dérivées partielles et leur signification physique. *Princeton University Bulletin*, 13, 1902.

[52] I.S. Helland, T.Næs et T. Isaksson. Related versions of the multiplicative scatter correction method for preprocessing spectroscopic data. *Chemometrics and Intelligent Laboratory Systems*, 29 p. 233–241, 1995.

[53] J.P. Hobert et G. Casella. The effect of improper priors on Gibbs sampling in hierarchical linear mixed models. *Journal of the American Statistical Association*, 91(436) p. 1461–1473, 1996.

[54] J.M. Hollas. *Spectroscopie*. Dunod, 1998. Édition française (traduction de S. Delacoudre).

[55] C.C. Holmes, A. Jasra et D.A. Stephens. Markov chain Monte Carlo methods and the label switching problem in Bayesian mixture modeling. *Statistical Science*, 20(1) p. 50–67, 2005.

[56] A. Holtsberg. Boîte à outils Matlab « Stixbox ». www.maths.lth.se/matstat/stixbox/. version 1.29.

[57] P.J. Huber. *Robust Statistics*. Wiley, New York, 1981.

[58] T. Ida, M. Ando et H. Toraya. Extended pseudo-Voigt function for approximating the Voigt profile. *Journal of Applied Spectroscopy*, 33 p. 1311–1316, 2000.

[59] J. Idier, éditeur. *Approche bayésienne pour les problèmes inverses*. Traité IC2, Série traitement du signal et de l'image. Hermès, Paris, novembre 2001.

[60] J. Idier. Convex half-quadratic criteria and interacting auxiliary variables for image restoration. *IEEE Transactions on Image Processing*, 10 p. 1001–1009, 2001.

[61] H. Jeffreys. *Theory of Probability*. Oxford University Press, 3e édition, 1961.

[62] M.C. Jodin, F. Gaboriaud et B. Humbert. Repercussions of size heterogeneity on the measurement of specific surface areas of colloidal minerals : Combination of macroscopic and microscopic analyses. *American Mineralogist*, 89 p. 1456–1462, 2004.

[63] K.F. Kaaresen et T. Taxt. Multichannel blind deconvolution of seismic signals. *Geophysics*, 63(6) p. 2093–2107, 1998.

[64] E. Kass et L. Wasserman. The selection of prior distributions by formal rules. *Journal of the American Statistical Association*, 91(435) p. 1343–1370, 1996.

[65] R. Koenker, P. Ng et S. Portnoy. Quantile smoothing splines. *Biometrika*, 81 p. 673–680, 1994.

[66] J.J. Kormylo et J.M. Mendel. Maximum likelihood detection and estimation of Bernoulli-Gaussian processes. *IEEE Transactions on Information Theory*, 28(3) p. 482–488, mai 1982.

[67] H. Kwakernaak. Estimation of pulse height and arrival times. *Automatica*, 16(4) p. 367–377, 1980.

[68] S.L. Lauritzen et D.J. Spiegelhalter. Local computations with probabilities on graphical structures and application to expert systems. *Journal of the Royal Statistical Society, Series B*, 50(2) p. 157–224, 1988.

[69] C.A. Lieber et A. Mahadevan-Jansen. Automated method for subtraction of fluorescence from biological Raman spectra. *Applied Spectroscopy*, 57 p. 1363–1367, 2003.

[70] B.F. Liu, Y. Sera, N. Matsubara, K. Otsuka et S. Terabe. Signal denoising and baseline correction by discrete wavelet transform for microchip capillary electrophoresis. *Electrophoresis*, 24 p. 3260–3265, 2003.

[71] D.A. Long. *Raman Spectroscopy*. McGraw-Hill, 1977.

[72] J. Luypaert, S. Heuerding, S. de Jong et D.L. Massart. An evaluation of direct orthogonal signal correction and other preprocessing methods for the classification of clinical study lots of a dermatological cream. *Journal of Pharmaceutical Biomedical Analysis*, 30 p. 453–466, 2002.

[73] V. Mazet, D. Brie et C. Caironi. Sparse spike train deconvolution using the Hunt filter and a thresholding method. *IEEE Signal Processing Letters*, 11(5) p. 486–489, mai 2004.

[74] V. Mazet, D. Brie et J. Idier. Baseline spectrum estimation using half-quadratic minimization. Dans *Actes de European Signal Processing Conference (EUSIPCO)*, p. 305–308, Vienne, Autriche, septembre 2004.

[75] V. Mazet, D. Brie et J. Idier. Simulation of positive normal variables using several proposal distributions. Dans *Actes de IEEE Workshop Statistical Signal Processing*, papier #190, Bordeaux, France, 17–20 juillet 2005.

[76] V. Mazet, D. Brie et J. Idier. Simuler une distribution normale à support positif à partir de plusieurs lois candidates. Dans *Actes de GRETSI 2005*, Louvain-la-Neuve, Belgique, 6–9 septembre 2005.

[77] V. Mazet, C. Carteret, D. Brie, J. Idier et B. Humbert. Background removal from spectra by designing and minimising a non-quadratic cost function. *Chemometrics and Intelligent Laboratory Systems*, 76(2) p. 121–133, avril 2005.

[78] V. Mazet, J. Idier et D. Brie. Déconvolution impulsionnelle positive myope. Dans *Actes de GRETSI 2005*, Louvain-la-Neuve, Belgique, 6–9 septembre 2005.

[79] V. Mazet, J. Idier, D. Brie, B. Humbert et C. Carteret. Estimation de l'arrière-plan de spectres par différentes méthodes dérivées des moindres carrés. Dans *Actes de Chimiométrie 2003*, p. 173–176, Paris, France, décembre 2003.

[80] N. Metropolis, A.W. Rosenbluth, M.N. Rosenbluth, A.H. Teller et E. Teller. Equations of state calculations by fast computing machines. *Journal of Chemical Physics*, 21(6) p. 1087–1092, 1953.

[81] S.P. Meyn et R.L. Tweedie. *Markov chains and stochastic stability*. Springer-Verlag, 1993.

[82] A.G. Michette et S.J. Pfauntsch. Laser plasma X-ray line spectra fitted using the Pearson VII function. *Journal of Physics D, Applied Physics*, 33 p. 1186–1190, 2000.

[83] A. Mohammad-Djafari. A Bayesian estimation method for detection, localisation and estimation of superposed sources in remote sensing. Dans *Actes de SPIE 97*, volume 3163, San Diego, Californie, États Unis, 24 juillet–1er août 1997.

[84] A. Mohammad-Djafari. Une méthode bayésienne pour la localisation et la séparation de sources de formes connues. Dans *Actes de GRETSI 97*, p. 135–139, Grenoble, France, 1997.

[85] S. Moussaoui. *Séparation de sources non négatives — application au traitement des signaux de spectroscopie*. Thèse de doctorat, Université Henri Poincaré, Nancy 1, décembre 2005.

[86] P. Müller. A generic approach to posterior integration and Gibbs sampling. Rapport technique # 91-09, Université de Purdue, West Lafayette, Indiana, États-Unis, 1991.

[87] P. Mykland, L. Tierney et B. Yu. Regeneration in Markov chain samplers. *Journal of the American Statistical Association*, 90 p. 233–241, 1995.

[88] N. Phambu, B. Humbert et A. Burneau. Relation between the infrared spectra and the lateral specific surface areas of gibbsite samples. *Langmuir*, 16 p. 6200–6207, 2000.

[89] W.H. Press, S.A. Teukolsky, W.T. Vetterling et B.P. Flannery. *Numerical Recipes in C*. Cambrige University Press, 2e édition, 1992.

[90] A.E. Raftery et S. Lewis. How many iterations in the Gibbs sampler ? Dans J.O. Berger, J.M. Bernardo, A.P. Dawid et A.F.M. Smith, éditeurs, *Bayesian Statistics 4*, p. 763–773. Oxford University Press, 1992.

[91] S. Richardson et P.J. Green. On Bayesian analysis of mixtures with an unknown number of components. *Journal of the Royal Statistical Society, Series B*, 59 p. 731–792, 1997.

[92] C.P. Robert. Simulation of truncated normal variables. *Statistics and Computing*, 5 p. 121–125, 1995.

[93] C.P. Robert. *Méthodes de Monte Carlo par chaînes de Markov*. Economica, 1996.

[94] C.P. Robert. *The Bayesian Choice*. Springer, 2e édition, 2001.

[95] G.O. Roberts, A. Gelman et W.R. Gilks. Weak convergence and optimal scaling of random walk Metropolis algorithms. *The Annals of Applied Probability*, 7(1) p. 110–120, 1997.

[96] O. Rosec. *Déconvolution aveugle multicapteur en sismique réflexion marine Très Haute Résolution*. Thèse de doctorat, Université de Bretagne Occidentale, avril 2000.

[97] O. Rosec, J.-M. Boucher, B. Nsiri et T. Chonavel. Blind marine seismic deconvolution using statistical MCMC methods. *IEEE Journal of Oceanic Engineering*, 28 p. 502–512, 2003.

[98] P.J. Rousseeuw. Least median of squares regression. *Journal of the American Statistical Association*, 79 p. 871–880, 1984.

[99] P.J. Rousseeuw et A.M. Leroy. *Robust Regression and Outlier Detection*. Series in Applied Probability and Statistics. Wiley-Interscience, New York, 1987.

[100] P.J. Rousseeuw et K. Van Driessen. Computing LTS regression for large data sets. Rapport technique, Université d'Anvers, 1999.

[101] K. Sauer et C. Bouman. A local update strategy for iterative reconstruction from projections. *IEEE Transactions on Signal Processing*, 41(2) p. 534–548, février 1993.

[102] J. Sjöblom, O. Svensson, M. Josefson, H. Kullberg et S. Wold. An evaluation of orthogonal signal correction applied to calibration transfer of near infrared spectra. *Chemometrics and Intelligent Laboratory Systems*, 44 p. 229–244, 1998.

[103] S. Sénécal. *Méthodes de simulation Monte-Carlo par chaînes de Markov pour l'estimation de modèles. Application en séparation de sources et en égalisation.* Thèse de doctorat, Institut National Polytechnique de Grenoble, novembre 2002.

[104] M. Stephens. *Bayesian methods for mixtures of normal distribution.* Thèse de doctorat, Magdalen College, Université d'Oxford, 1997.

[105] M. Stephens. Discussion of the paper by Richardson and Green "On bayesian analysis of mixtures with an unknown number of components". *Journal of the Royal Statistical Society, Series B*, 59(4) p. 768–769, 1997.

[106] M. Stephens. Dealing with label-switching in mixture models. *Journal of the Royal Statistical Society, Series B*, 62 p. 795–809, 2000.

[107] H.A. Taha. *Operations research : an introduction.* Macmillan Publishing Company, 4e édition, 1989.

[108] H.-W. Tan et S.D. Brown. Wavelet analysis applied to removing non-constant, varying spectroscopic background in multivariate calibration. *Journal of Chemometrics*, 16 p. 228–240, 2002.

[109] L. Tierney. Markov chain for exploring posterior distributions. *The Annals of Statistics*, 22(4) p. 1701–1762, 1994.

[110] T.J. Vickers, R.E. Wambles et C.K. Mann. Curve fitting and linearity : data processing in Raman spectroscopy. *Applied Spectroscopy*, 55 p. 389–393, 2001.

[111] J.A. Westerhuis, S. de Jong et A.K. Smilde. Direct orthogonal signal correction. *Chemometrics and Intelligent Laboratory Systems*, 56 p. 13–25, 2001.

[112] S. Wold, H. Antti, F. Lindgren et J. Ohman. Orthogonal signal correction of near-infrared spectra. *Chemometrics and Intelligent Laboratory Systems*, 44 p. 175–185, 1998.

[113] R. Yang et J.O. Berger. A catalog of noninformative priors. ISDS, Duke University, Durham, Caroline du Nord, États-Unis, 1997. Discussion Paper 97-42.

[114] R. Yarlagadda, J. Bee Bednar et T.L. Watt. Fast algorithms for l_p deconvolution. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 33(1) p. 174–182, février 1985.

Université HENRI POINCARÉ
NANCY 1

Monsieur **MAZET Vincent**

DOCTORAT DE L'UNIVERSITE HENRI POINCARE, NANCY 1

en AUTOMATIQUE, TRAITEMENT DU SIGNAL & GENIE INFORMATIQUE

VU, APPROUVÉ ET PERMIS D'IMPRIMER N°1140

Nancy, le 22 décembre 2005

Le Président de l'Université

UNIVERSITÉ HENRI POINCARÉ, NANCY I
LE PRESIDENT

J.P. FINANCE

Université Henri Poincaré, Nancy I
24-30 rue Lionnois - B.P. 3069 - 54013 NANCY Cédex
Tél. : 03 83 68 20 00 - Fax : 03 83 68 21 00

Résumé

Cette thèse s'inscrit dans le cadre d'une collaboration entre le CRAN (UMR 7039) et le LCPME (UMR 7564) dont l'objectif est de développer des méthodes d'analyse de signaux spectroscopiques.

Dans un premier temps est proposée une méthode déterministe qui permet d'estimer la ligne de base des spectres par le polynôme qui minimise une fonction-coût non quadratique (fonction de Huber ou parabole tronquée). En particulier, les versions asymétriques sont particulièrement bien adaptées pour les spectres dont les raies sont positives. Pour la minimisation, on utilise l'algorithme de minimisation semi-quadratique LEGEND.

Dans un deuxième temps, on souhaite estimer le spectre de raies : l'approche bayésienne couplée aux techniques MCMC fournit un cadre d'étude très efficace. Une première approche formalise le problème en tant que déconvolution impulsionnelle myope non supervisée. En particulier, le signal impulsionnel est modélisé par un processus Bernoulli-gaussien à support positif ; un algorithme d'acceptation-rejet mixte permet la simulation de lois normales tronquées. Une alternative intéressante à cette approche est de considérer le problème comme une décomposition en motifs élémentaires. Un modèle original est alors introduit ; il a l'intérêt de conserver l'ordre du système fixe. Le problème de permutation d'indices est également étudié et un algorithme de ré-indexage est proposé.

Les algorithmes sont validés sur des spectres simulés puis sur des spectres infrarouge et Raman réels.

Mots clés : spectroscopie infrarouge, spectroscopie Raman, correction de la ligne de base, minimisation semi-quadratique, approche bayésienne, déconvolution impulsionnelle positive myope, décomposition en motifs élémentaires, simulation de loi normale tronquée.

Abstract

This thesis is part of a collaboration between the CRAN (UMR 7039) and the LCPME (UMR 7564) aiming at developing analysis methods for spectroscopic signals.

Firstly, a deterministic method is proposed. It allows to estimate the baseline as the polynomial minimizing a non-quadratic cost function (Huber function, truncated quadratic). In particular, asymmetrical cost functions are well adapted for spectra whose peaks are positive. The minimization is achieved by the half-quadratic minimisation algorithm LEGEND.

Secondly, we consider the problem of the spectral peaks estimation : the Bayesian approach coupled with MCMC methods provides a very powerful theoretical framework. In a first approach, the problem is considered as an unsupervised blind positive sparse spike deconvolution. The sparse spike train is modelled as a Bernoulli-Gaussian process with positive support ; an original mixed accept-reject algorithm allows the simulation of positive normal variables. An interesting alternative method consists in considering the problem as a decomposition into elementary signals. An original model is introduced, allowing to keep the model dimension fixed. The label switching problem is also studied and a relabelling algorithm is proposed.

The methods are applied to both simulated and experimental spectra (infrared and Raman).

Keywords : infrared spectroscopy, Raman spectroscopy, baseline estimation, half-quadratic minimisation, Bayesian approach, positive sparse spike train deconvolution, decomposition into elementary signals, simulation of positive normal variables.

Erratum

Développement de méthodes de traitement de signaux spectroscopiques : estimation de la ligne de base et du spectre de raies

Vincent Mazet

15 mai 2007

- Les adresses Internet renvoyant sur mon site web ont changé. Ces adresses sont désormais : `http ://lsiit-miv.u-strasbg.fr/lsiit/perso/mazet/...`
- **p. 26** La véritable expression d'une lorentzienne est en fait :

$$f(x) = \frac{1}{\pi}\frac{s}{s^2 + (x - x_0)^2}.$$

  Toutefois, tout au long de la thèse, nous utilisons une lorentzienne multipliée par un scalaire :

$$f(x) = \frac{s^2}{s^2 + (x - x_0)^2}.$$

  Cette expression permet d'avoir une amplitude maximale de 1 (en x_0), le paramètre $\mathbf{a}$ donnant alors la véritable amplitude du pic lorentzien.
- **p. 60** Dans la première équation de la section 3.3.1.3, il faut remplacer $\mathbf{x}_n^T\mathbf{x}_n$ par $\mathbf{x}^T\mathbf{x}$.
- **p. 64** Il n'est pas nécessaire de forcer $\mu_{1,n}$ à zéro s'il est négatif (voir correction pages 133–134).
- **p. 65** Dans l'équation (3.13) et l'équation précédente, il faut remplacer $\mathbf{x}_n^T\mathbf{x}_n$ par $\mathbf{x}^T\mathbf{x}$.
- **p. 83** Dernière équation de la page, il faut lire :

$$\mathbf{G} = (\mathbf{h}(\mathbf{c}_1, \mathbf{s}_1) \dots \mathbf{h}(\mathbf{c}_{K_{\max}}, \mathbf{s}_{K_{\max}}))$$

- **p. 85** L'a priori sur λ (équation (4.6)), est inexact : il faut lire $\lambda \sim \mathcal{B}e(1, K_{\max} + 1)$.
- **p. 86 (figure 4.2)** Dans le graphe acyclique orienté, $\mathbf{s}_k$ est distribué suivant $\mathcal{U}_{\mathbb{R}^+}$ et non $\mathcal{U}_{\mathbb{R}^*}$.
- **p. 87** Il n'est pas nécessaire de forcer $\mu_{1,n}$ à zéro s'il est négatif (voir correction pages 133–134).
- **p. 88** Dans l'équation (4.13), il faut remplacer $\mathbf{x}_n^T\mathbf{x}_n$ par $\mathbf{x}^T\mathbf{x}$.
- **p. 90** Deuxième ligne, la loi a posteriori est en fait $p(\boldsymbol{\theta}_1, \dots, \boldsymbol{\theta}_{K_{\max}}|\mathbf{y})$.
- **p. 95** Dernier paragraphe avant la section 4.4.6 : il manque un point entre « directe » et « Toutefois ».
- **p. 96** Dans la dernière équation, la somme va de $k = 1$ à $K_{\max}$ (et non K).
- **p. 110–112 (tableaux 4.8–4.10)** Les troisièmes colonnes de chaque spectre ne correspondent pas à $\widehat{s}$ mais à $\widehat{\sigma}$, l'écart-type des raies gaussiennes.
- **p. 133 (et tout au long de la thèse)** La notation $\mathbf{e}_n = \mathbf{w} - (\mathbf{H}\mathbf{x} - \mathbf{H}_n\mathbf{x}_n)$ est incohérente. Il faudrait plutôt utiliser la notation $\mathbf{e}^n$ puisque cette quantité n'est pas un scalaire mais bien un vecteur.
- **p. 133–134** La première équation de la page 134 est fausse : le coefficient devant l'exponentielle est égal à la variable C définie page 135. Dès lors, les calculs suivants sont inexacts. Les valeurs de $\mu_{0,n}$, $\mu_{1,n}$, $\rho_{0,n}$ et $\rho_{1,n}$ restent valables, mais on a :

$$\lambda_{1,n} = \left[1 + \frac{1-\lambda}{\lambda}\sqrt{\frac{r_{\mathbf{a}}}{\rho_{1,n}}}\exp\left(-\frac{\mu_{1,n}^2}{2\rho_{1,n}}\right)\frac{1}{1 + \operatorname{erf}\left(\mu_{1,n}/\sqrt{2\rho_{1,n}}\right)}\right]^{-1}.$$

Ainsi, il n'est plus nécessaire de forcer $\mu_{1,n}$ à zéro s'il est négatif comme cela est indiqué pages 64 et 87. Les détails des calculs exacts sont téléchargeables sur mon site.

- **p. 136** Dans l'expression de la moyenne et de la variance d'une loi normale tronquée, il faut remplacer ρ par σ^2.

Pour terminer, il existe un grand nombre de fautes d'orthographe dissimulées. Le premier lecteur qui m'enverra la liste complète se verra remettre un exemplaire de ma thèse dédicacée.

Annexe B

Calcul de la loi a posteriori de $\mathbf{x}_n$ (correction)

On rappelle l'expression de la loi a posteriori de $\mathbf{x}_n$ (section 3.3.3.1) :

$$p(\mathbf{x}_n|\mathbf{w}, \mathbf{x}_{-n}, \lambda, r_\mathbf{a}, s, r_\mathbf{b}) \propto p(\mathbf{w}|\mathbf{x}, s, r_\mathbf{b}) p(\mathbf{x}_n|\lambda, r_\mathbf{a}).$$

En considérant l'impulsion de Dirac comme une gaussienne de variance nulle r_0, la vraisemblance et la densité a priori s'écrivent respectivement :

$$\begin{aligned} p(\mathbf{w}|\mathbf{x}, s, r_\mathbf{b}) &= \frac{1}{(2\pi r_\mathbf{b})^{N/2}} \exp\left(-\frac{1}{2r_\mathbf{b}}||\mathbf{w} - \mathbf{H}\mathbf{x}||^2\right), \\ p(\mathbf{x}_n|\lambda, r_\mathbf{a}) &= (1-\lambda)\sqrt{\frac{2}{\pi r_0}} \exp\left(-\frac{\mathbf{x}_n^2}{2r_0}\right) \mathbb{1}_{\mathbb{R}^+}(\mathbf{x}_n) + \lambda\sqrt{\frac{2}{\pi r_\mathbf{a}}} \exp\left(-\frac{\mathbf{x}_n^2}{2r_\mathbf{a}}\right) \mathbb{1}_{\mathbb{R}^+}(\mathbf{x}_n). \end{aligned}$$

Ainsi, en regroupant ces deux expressions, on obtient :

$$p(\mathbf{x}_n|\mathbf{w}, \mathbf{x}_{-n}, \lambda, r_\mathbf{a}, s, r_\mathbf{b}) \propto \sum_{l=0,1} \lambda_l \sqrt{\frac{2}{\pi r_l}} \exp\left(-Q_{l,n}\right) \mathbb{1}_{\mathbb{R}^+}(\mathbf{x}_n),$$

avec :

$$\lambda_0 = 1 - \lambda, \qquad \lambda_1 = \lambda, \qquad r_1 = r_\mathbf{a} \qquad \text{et} \qquad Q_{l,n} = \frac{1}{2r_\mathbf{b}}||\mathbf{w} - \mathbf{H}\mathbf{x}||^2 + \frac{\mathbf{x}_n^2}{2r_l}.$$

Le terme $1/(2\pi r_\mathbf{b})^{N/2}$ a été supprimé puisqu'il est indépendant de l et n'est donc qu'une constante de l'expression.

En posant $\mathbf{e}_n = \mathbf{w} - (\mathbf{H}\mathbf{x} - \mathbf{H}_n\mathbf{x}_n)$ afin d'extraire le terme $\mathbf{x}_n$, $Q_{l,n}$ devient :

$$Q_{l,n} = \frac{1}{2r_\mathbf{b}}||\mathbf{e}_n - \mathbf{H}_n\mathbf{x}_n||^2 + \frac{\mathbf{x}_n^2}{2r_l},$$

qui, en développant, s'écrit :

$$Q_{l,n} = a\mathbf{x}_n^2 + b\mathbf{x}_n + c = a(\mathbf{x}_n + b/2a)^2 + (c - b^2/4a),$$

avec :

$$a = \frac{r_\mathbf{b} + r_l\mathbf{H}_n^T\mathbf{H}_n}{2r_l r_\mathbf{b}}, \qquad b = -\frac{\mathbf{e}_n^T\mathbf{H}_n}{r_\mathbf{b}}, \qquad \text{et} \qquad c = \frac{\mathbf{e}_n^T\mathbf{e}_n}{2r_\mathbf{b}}.$$

On obtient donc l'expression suivante :

$$Q_{l,n} = \frac{r_\mathbf{b} + r_l\mathbf{H}_n^T\mathbf{H}_n}{2r_l r_\mathbf{b}} \left(\mathbf{x}_n - \frac{r_l r_\mathbf{b}\mathbf{e}_n^T\mathbf{H}_n}{r_\mathbf{b}(r_\mathbf{b} + r_l\mathbf{H}_n^T\mathbf{H}_n)}\right)^2 + \left(\frac{\mathbf{e}_n^T\mathbf{e}_n}{2r_\mathbf{b}} - \left(\frac{\mathbf{e}_n^T\mathbf{H}_n}{r_\mathbf{b}}\right)^2 \frac{r_l r_\mathbf{b}}{2(r_\mathbf{b} + r_l\mathbf{H}_n^T\mathbf{H}_n)}\right),$$

et ainsi,

$$p(\mathbf{x}_n|\mathbf{w},\mathbf{x}_{-n},\lambda,r_{\mathbf{a}},s,r_{\mathbf{b}}) \propto \sum_{l=0,1} \frac{\gamma_{l,n}}{\sqrt{\pi\rho_{l,n}/2}\left[1+\operatorname{erf}\left(\mu_{l,n}/\sqrt{2\rho_{l,n}}\right)\right]} \exp\left(-\frac{(\mathbf{x}_n-\mu_{l,n})^2}{2\rho_{l,n}}\right) \mathbb{1}_{\mathbb{R}^+}(\mathbf{x}_n),$$

où :

$$\frac{\gamma_{l,n}}{\sqrt{\pi\rho_{l,n}/2}\left[1+\operatorname{erf}\left(\mu_{l,n}/\sqrt{2\rho_{l,n}}\right)\right]} = \lambda_l\sqrt{\frac{2}{\pi r_l}}\exp\left(-\frac{\mathbf{e}_n^T\mathbf{e}_n}{2r_{\mathbf{b}}}+\left(\frac{\mathbf{e}_n^T\mathbf{H}_n}{r_{\mathbf{b}}}\right)^2\frac{r_l r_{\mathbf{b}}}{2(r_{\mathbf{b}}+r_l\mathbf{H}_n^T\mathbf{H}_n)}\right),$$

$$\frac{1}{2\rho_{l,n}} = \frac{r_{\mathbf{b}}+r_l\mathbf{H}_n^T\mathbf{H}_n}{2r_l r_{\mathbf{b}}},$$

$$\mu_{l,n} = \frac{r_l r_{\mathbf{b}}\mathbf{e}_n^T\mathbf{H}_n}{r_{\mathbf{b}}(r_{\mathbf{b}}+r_l\mathbf{H}_n^T\mathbf{H}_n)}.$$

Or, dans l'expression de $\gamma_{l,n}$, le terme $\exp\left(-\mathbf{e}_n^T\mathbf{e}_n/2r_{\mathbf{b}}\right)$ est constant et peut donc être supprimé. Les expressions précédentes se simplifient en :

$$\gamma_{l,n} = \lambda_l\sqrt{\frac{\rho_{l,n}}{r_l}}\exp\left(\frac{\mu_{l,n}^2}{2\rho_{l,n}}\right)\left[1+\operatorname{erf}\left(\frac{\mu_{l,n}}{\sqrt{2\rho_{l,n}}}\right)\right], \qquad \rho_{l,n} = \frac{r_l r_{\mathbf{b}}}{r_{\mathbf{b}}+r_l\mathbf{H}_n^T\mathbf{H}_n}, \qquad \mu_{l,n} = \frac{\rho_{l,n}}{r_{\mathbf{b}}}\mathbf{e}_n^T\mathbf{H}_n.$$

De plus, afin d'obtenir une densité de probabilité d'intégrale 1, il est nécessaire que $\gamma_{0,n}+\gamma_{1,n}=1$. On définit alors $\lambda_{l,n}$ en posant :

$$\lambda_{l,n} = \frac{\gamma_{l,n}}{\gamma_{l,n}+\gamma_{1-l,n}} = \frac{1}{1+\gamma_{1-l,n}/\gamma_{l,n}}.$$

Enfin, la loi a posteriori de $\mathbf{x}_n$ a pour expression :

$$p(\mathbf{x}_n|\mathbf{w},\mathbf{x}_{-n},\lambda,r_{\mathbf{a}},s,r_{\mathbf{b}}) \sim \lambda_{0,n}\mathcal{N}^+\left(\mu_{0,n},\rho_{0,n}\right)+\lambda_{1,n}\mathcal{N}^+\left(\mu_{1,n},\rho_{1,n}\right) \tag{B.1}$$

où les expressions des paramètres se simplifient en (rappelons que $r_0=0$) :

$$\lambda_{0,n} = 1-\lambda_{1,n}, \qquad \lambda_{1,n} = \left[1+\frac{1-\lambda}{\lambda}\sqrt{\frac{r_{\mathbf{a}}}{\rho_{1,n}}}\exp\left(-\frac{\mu_{1,n}^2}{2\rho_{1,n}}\right)\frac{1}{1+\operatorname{erf}\left(\mu_{1,n}/\sqrt{2\rho_{1,n}}\right)}\right]^{-1},$$

$$\rho_{0,n} = 0, \qquad \mu_{0,n} = 0, \qquad \rho_{1,n} = \frac{r_{\mathbf{a}}r_{\mathbf{b}}}{r_{\mathbf{b}}+r_{\mathbf{a}}\mathbf{H}_n^T\mathbf{H}_n}, \qquad \mu_{1,n} = \frac{\rho_{1,n}}{r_{\mathbf{b}}}\mathbf{e}_n^T\mathbf{H}_n.$$