
Approche bayésienne pour la décomposition conjointe d'une séquence de spectres de photo-électrons

Vincent Mazet¹, Sylvain Faisan¹, Antoine Masson^{2,3},
Marc-André Gaveau^{2,3}, Lionel Poisson^{2,3},
Jean-Michel Mestdagh^{2,3}

1. ICube, Université de Strasbourg, CNRS, UMR 7357
300 boulevard Sébastien Brant, BP 10413, F-67412 Illkirch cedex
{vincent.mazet,faisan}@unistra.fr
2. CNRS, IRAMIS, SPAM, Laboratoire Francis Perrin, URA 2453
F-91191 Gif-sur-Yvette
antoine.masson@epfl.ch,
{marc-andre.gaveau,lionel.poisson,jean-michel.mestdagh}@cea.fr
3. CEA, IRAMIS, SPAM, Laboratoire Francis Perrin, URA 2453
F-91191 Gif-sur-Yvette
(Antoine Masson est actuellement à l'EPFL, 1015 Lausanne, Suisse)

RÉSUMÉ. Ce travail traite de la décomposition d'une séquence temporelle de spectres de photo-électrons en une somme de raies dont on estime les positions, amplitudes et largeurs. Comme les raies évoluent (lentement) dans le temps, la décomposition est effectuée conjointement sur toute la séquence afin de prendre en compte cette information temporelle. À cette fin, nous avons développé un modèle bayésien où un champ de Markov favorise une évolution douce des raies. L'approche est non supervisée et un échantillonneur de Gibbs couplé à un schéma de recuit simulé permet d'estimer le maximum a posteriori. Nous montrons la pertinence de cette approche par rapport à une méthode dans laquelle les spectres sont décomposés séparément et présentons une application sur données réelles de photo-électrons.

ABSTRACT. This work deals with the decomposition of a temporal sequence of photoelectron spectra into a sum of peaks whose positions, amplitudes and widths are estimated. Since the peaks exhibit a (slow) evolution with time, the decomposition is performed jointly on the whole sequence to take this temporal information into account. To this end, we have developed a Bayesian model where a Markov random field favors a smooth evolution of peaks. The approach is unsupervised and a Gibbs sampler within a simulated annealing scheme enables to estimate the maximum a posteriori. We show the relevance of this approach compared with a method in which the spectra are decomposed separately and present an application on real photoelectron data.

MOTS-CLÉS : décomposition d'une séquence de signaux spectroscopiques, inférence bayésienne, méthodes de Monte Carlo par chaînes de Markov (MCMC), recuit simulé, spectroscopie de photo-électrons.

KEYWORDS: spectroscopic signal sequence decomposition, Bayesian inference, Markov chain Monte Carlo (MCMC) methods, simulated annealing, photoelectron spectroscopy.

DOI:10.3166/TS.30.9-34 © 2013 Lavoisier

Extended Abstract

Introduction

Time-resolved photoelectron spectroscopy (Stolow *et al.*, 2004) studies the energy relaxation occurring after absorption of a photon by a molecule, an atom or a blend of both. The absorption results in the ionization of the studied system, thereby ejecting a so-called photoelectron. The distribution of the photoelectrons according to their energy is measured at different times to get a temporal sequence of photoelectron spectra (Gloaguen *et al.*, 2005). The goal is to determine how the energy, intensity and shape (*i.e.* centers, amplitudes and widths) of the peaks evolve through the sequence, indicating the changes undergone by the studied system. More information on the technique as well as the experimental setup can be found in (Masson *et al.*, 2010).

A photoelectron spectrum can be modeled as the sum of peaks superimposed on a continuum. The peak parameters (centers, amplitudes and widths) are to be estimated and the peaks have to be associated so as to characterize their evolution.

As far as we know, nobody has addressed the problem of the decomposition of a sequence of spectra, but many works provide solutions to similar problems. Approaches related to source separation (convolutive or not) (Nguyen Thi, Jutten, 1995) or spectral unmixing (Keshava, Mustard, 2002) cannot be directly used because the peaks exhibit various locations and shapes through the sequence. However, the decomposition of a unique spectrum has been intensively studied in the past (Donoho, Tsaig, 2006 ; Sahnoun *et al.*, 2011 ; Fischer, Dose, 2001 ; Gulam Razul *et al.*, 2003 ; Haan, Godsill, 2002 ; Mazet, 2005).

We recently show (Mazet, 2011) that a sequential approach, in which the spectra has decomposed separately, is inappropriate. On the contrary, a joint decomposition approach, where spectra are decomposed simultaneously, favors a smooth change of the peak parameters by regularizing their evolution, thus providing coherent and consistent results. In addition, a joint approach is also able to classify the peaks, thus giving the possibility to follow them through the sequence. However, the method proposed in (Mazet, 2011) needs to set the hyperparameters and was not applied on real data.

Model and Algorithm

The proposed approach is set in a Bayesian framework. A spectrum \mathbf{y}_s ($s \in \{1, \dots, S\}$) is modeled as the noisy sum of K Gaussian peaks (with centers $c_{s,k}$, amplitude $a_{s,k}$, and width $w_{s,k}$) and an exponential continuum:

$$(\mathbf{y}_s)_n = \sum_{k=1}^K a_{s,k} \exp\left(-\frac{(n - c_{s,k})^2}{2w_{s,k}^2}\right) + \alpha_s \exp\left(-\frac{n}{\beta_s}\right) + (\mathbf{b}_s)_n,$$

The peak number K is supposed to be known and identical in each spectrum. Besides, we need to introduce the notion of *tracks*: a track is the association of S peaks, each one in a different spectrum, and we suppose that the peak evolution in one track is smooth.

A Markovian prior models, for each track, the evolution of the centers, the amplitudes and the width. In the same way, a Markovian prior is also used for the continuum parameters. The noise is supposed to be zero-mean, white, and Gaussian; the hyper-parameters (noise variance, prior strengths) need to be small so they are supposed to be distributed according to an inverse gamma distribution.

A Gibbs sampler is used to sample the posterior distribution (Robert, Casella, 2004). But, the Gibbs sampler is known to be able to be trapped in local minima, that is why it is here included into a simulated annealing framework (Geman, Geman, 1984; Kirkpatrick *et al.*, 1983).

Results

The proposed approach is compared on synthetic signals with the one proposed in (Mazet, 2011): its computation time is longer (which can be explained by the fact that more variables are needed to be simulated) but it yields better estimations. We show that the proposed joint approach, which is able to take into account the estimation of the contiguous spectra, favors certain zones in the solution space: in other words, the decomposition of a spectrum helps the decomposition of the contiguous spectra. In addition, the proposed approach is able to manage crossings and peak overlaps.

The proposed approach was also performed on real photoelectron data. The experiments presented in this paper study the relaxation of an atom of barium (Ba) on an aggregate (droplet isolated from several hundred atoms) of argon (Ar): it is question of understanding the mechanism that leads the excited barium atom to lose its energy through an interaction with an inert and cold medium (34 K). The results are shown figures 7 and 8: we see that the peak evolution is smooth and that the peaks are well classified.

1. Introduction

La spectroscopie de photo-électrons résolue en temps (Stolow *et al.*, 2004) est une méthode permettant d'étudier le processus de relaxation de l'énergie qui apparaît après absorption d'un photon par une molécule, un atome ou un assemblage des deux. La détection de cette dynamique repose sur l'ionisation du système étudié, éjectant ainsi un électron, appelé alors photo-électron. L'énergie cinétique de ce dernier porte une information sur l'état du système au moment de son ionisation : à énergie de photon incident constante, l'électron est émis avec d'autant moins d'énergie cinétique que le système a relaxé son énergie électronique en énergie thermique. Un spectre de photo-électrons est la distribution des électrons en fonction de leur énergie (un exemple est représenté figure 6). L'acquisition d'un spectre de photo-électrons en fonction du temps permet donc de suivre l'évolution du phénomène de relaxation (Gloaguen *et al.*, 2005) : on aboutit alors à une séquence temporelle de spectres de photo-électrons (la figure 3 représente une séquence simulée).

Le spectre de photo-électrons peut être modélisé comme une somme de raies (qui informe donc sur la distribution d'énergie des électrons du système physique considéré) superposée à un continuum. L'objectif de cet article est de déterminer comment l'énergie, la forme et l'intensité des raies évoluent dans le temps – donc à travers la séquence – indiquant ainsi les modifications électroniques que subit le système étudié. Pour cela, il faut décomposer chaque spectre en raies dont on estime les centres, amplitudes et largeurs (c'est un problème de décomposition spectroscopique) et il faut également associer les raies entre elles au fil du temps de manière à caractériser l'évolution temporelle des différentes raies. Ces deux étapes sont ici effectuées de manière jointe.

Un exemple particulièrement simple et informatif apparaît dans (Gloaguen *et al.*, 2005) : plusieurs états électroniques interviennent après que la molécule de tétrakis (diméthylamino-éthylène) ait été excitée électroniquement par une impulsion laser, celle-ci peuple un état de valence V ; ce dernier relaxe en totalité vers un état de Rydberg R qui lui-même relaxe vers un état zwitterionique Z. L'expérience rapportée dans (Gloaguen *et al.*, 2005) est très simple car le laser interroge uniquement ces trois états et ceux-ci conduisent à trois raies parfaitement individualisées dans le spectre de photo-électrons : on observe alors que la raie associée à R est un transitoire, peuplé au cours du temps par la décroissance de la raie V, se dépeuplant lui-même au profit de Z. Les travaux en cours en dynamique réactionnelle portent sur des situations plus complexes. Par exemple, (Poisson *et al.*, 2008 ; 2010) considèrent des situations maintenant très courantes où, d'une part, le nombre et la position des raies pour chaque spectre ne sont pas connus et, d'autre part, ces paramètres ne sont pas constants au cours de la séquence (ils peuvent varier d'un spectre à l'autre). Par ailleurs, les raies sont larges et empiètent les unes sur les autres dans un contexte où le rapport signal-sur-bruit est faible. Dans cet article, nous considérons ce type de situation complexe à la différence que le nombre de raies de chaque spectre ne varie pas au cours de la séquence et qu'il est supposé connu.

La décomposition spectroscopique ne peut pas se poser comme un problème de séparation de source (convolutif ou non) (Nguyen Thi, Jutten, 1995) ou de démixage spectral (*spectral unmixing*) (Keshava, Mustard, 2002), car les spectres sont modélisés dans ces approches comme des fonctions de signaux élémentaires invariants (appelés *sources*). Dans notre cas, les spectres sont modélisés par des raies dont les positions et formes varient au cours de la séquence.

Notons toutefois que (Gobinet *et al.*, 2009) proposent différents types de prétraitements pour que les méthodes traditionnelles de séparation de source puissent être utilisées pour effectuer la décomposition d'une séquence de signaux spectroscopiques. Parmi ces prétraitements, citons par exemple : un calibrage qui permet de supprimer plusieurs bruits liés à l'acquisition, la suppression de la dérive de ligne de base, l'alignement des pics et l'homogénéisation de la largeur des pics. Or, l'alignement des pics et l'homogénéisation de leur largeur nécessitent une connaissance a priori sur la localisation des pics. Il existe des alternatives (Tomasi *et al.*, 2004) qui ne nécessitent pas d'a priori mais elles sont supervisées et n'ont pas donné de résultats satisfaisants.

À notre connaissance, aucun autre travail n'a déjà traité de la décomposition conjointe d'une séquence de spectres, à savoir du traitement de tous les spectres de la séquence dans leur ensemble et non séparément. Toutefois, certaines applications pour lesquelles des solutions ont été proposées y ressemblent. En chromatographie en phase liquide couplée à la spectrométrie de masse (LC-MS : *liquid chromatography-mass spectrometry*), (Katajamaa, Orešič, 2005) ont développé un programme dans lequel les spectres sont traités séparément et les centres des raies sont estimés comme étant les maxima locaux. Il en résulte une méthode peu fiable si le rapport signal-à-bruit est faible ou si des raies se superposent. De plus, nous avons récemment montré qu'une approche dans laquelle les spectres sont décomposés conjointement plutôt que séquentiellement est bien plus performante (Mazet, 2011) ; nous détaillons ce point dans la suite. En temps-fréquence, plusieurs travaux ont introduit des algorithmes d'approximation parcimonieuse pour décomposer des représentations de signaux sonores en atomes temps-fréquence ; citons par exemple les algorithmes *Matching Pursuit* de (Mallat, Zhang, 1993) et *Basis Pursuit* de (Chen *et al.*, 1995). Des méthodes stochastiques ont également été proposées : (Davy *et al.*, 2006) par exemple, utilisent l'algorithme RJMCMC (*reversible jump Monte Carlo Markov chain*) sur des représentations temps-fréquence de signaux musicaux. Dans cette application, les objets à estimer sont des fonctions bidimensionnelles (les atomes temps-fréquence) ; cela reviendrait dans notre cas à considérer les objets de nos données comme étant les trajectoires d'une raie à travers la séquence, ce qui impliquerait un modèle trop complexe. Notre idée est donc de se ramener à des objets unidimensionnels (les raies) et de les associer pour former les trajectoires dont on peut quantifier l'évolution. En sismique-réflexion multicanal, (Roueff *et al.*, 2004b ; 2004a) ont développé une approche pour séparer les ondes sismiques à l'aide de l'algorithme de ligne de partage des eaux sur la transformée en ondelettes des données, mais cette approche n'est pas paramétrique et est donc mal adaptée à notre problème où l'on peut modéliser les raies. (Idier, Gousard, 1993) ont développé un modèle bayésien où un a priori markovien permet de probabiliser la présence d'une raie à certaines positions. Dans notre cas, nous propo-

sons un modèle dans lequel un a priori markovien permet de probabiliser directement le vecteur des inconnus. Cela permet d'avoir une méthode plus précise sur le positionnement des raies (rien n'impose qu'elles soient centrées sur les échantillons de la grille), de régulariser sur d'autres paramètres que les centres (dans notre cas : les largeurs et les amplitudes) et enfin de disposer d'un modèle plus riche pour modéliser les évolutions des raies (on peut, par exemple, favoriser les évolutions constantes, linéaires, etc.).

Les travaux précédemment cités n'étant pas directement adaptables à notre problème, il nous faut proposer une méthode originale. Or, il existe une littérature abondante traitant de la décomposition d'un *unique* spectre. Les premiers travaux proposent des approches très simples, comme une estimation des paramètres au sens des moindres carrés (Caruana *et al.*, 1986) ou des méthodes de gradient (Allen, McMeeking, 1978 ; Fraser, Suzuki, 1966). Seulement, dans le cas de données réelles, bruitées, et sur lesquelles apparaissent de nombreux recouvrements de raies, ces méthodes sont complètement inefficaces. Des approches plus récentes se placent dans un cadre bayésien pour estimer les paramètres des raies en utilisant des méthodes de Monte Carlo par chaînes de Markov (MCMC) (Fischer, Dose, 2001 ; Gulam Razul *et al.*, 2003 ; Haan, Godsill, 2002 ; Mazet, 2005). Les résultats obtenus sont très satisfaisants et c'est la raison pour laquelle nous nous inspirons de ces travaux pour développer notre algorithme. Plus récemment, des approches de décomposition ont été traitées comme un problème d'approximation parcimonieuse (Donoho, Tsaig, 2006 ; Sahnoun *et al.*, 2011).

L'approche que nous proposons dans cet article se place dans un cadre bayésien puisque ce sont des méthodes performantes et relativement simples à mettre en œuvre. Nous avons récemment proposé (Mazet, 2011) une telle méthode et montré qu'une approche séquentielle, dans laquelle les spectres sont décomposés indépendamment les uns des autres, n'est pas appropriée¹. En effet, la décomposition de deux spectres contigus (c'est-à-dire enregistrés à deux temps adjacents) peut aboutir à deux décompositions très différentes alors que les spectres sont très similaires. Nous avons également montré l'intérêt d'une décomposition conjointe, qui prend en compte le fait que les raies évoluent lentement dans le temps (au fil de la séquence) :

- la décomposition d'un spectre est guidée par les décompositions voisines grâce à l'a priori d'évolution douce des paramètres ;
- les décompositions sont cohérentes car elles évoluent doucement ;
- une décomposition conjointe fournit une classification des raies, donnant ainsi la possibilité de suivre les raies dans la séquence.

Cependant, l'approche proposée dans (Mazet, 2011) n'a été validée que sur données simulées et plusieurs limitations sont apparues lors de l'application sur données réelles.

1. (Idier, Goussard, 1993) sont également arrivés à la même conclusion.

En effet, les données réelles contiennent plus de spectres que les données synthétiques utilisées dans (Mazet, 2011) : cela conduit à une augmentation significative de la taille de l'espace de recherche et l'algorithme proposé dans (Mazet, 2011) échoue ; une solution est présentée dans la section 3. En outre, le continuum n'ayant pas été modélisé dans (Mazet, 2011), nous proposons ici de le faire par une fonction exponentielle (équation (1)). Enfin, la méthode proposée dans (Mazet, 2011) requiert le réglage d'hyperparamètres. La variabilité des données réelles rend ce réglage difficile et c'est pourquoi la méthode proposée dans cet article est non supervisée.

Le modèle et les lois a priori sont détaillés dans la section 2. L'estimateur du maximum a posteriori (MAP) est obtenu à l'aide d'un algorithme MCMC couplé à une approche de recuit simulé (section 3). Finalement, les performances de la méthode sont illustrées sur des données synthétiques (section 4) et réelles (section 5).

2. Modèle bayésien

Dans la suite de l'article, les variables notées en gras sont des vecteurs ou des matrices, sinon elles correspondent à des scalaires.

Les données à traiter correspondent à S spectres composés chacun de N échantillons. Un spectre \mathbf{y}_s ($s \in \{1, \dots, S\}$) est modélisé comme la somme bruitée de K raies gaussiennes et d'une fonction exponentielle (Masson *et al.*, 2010) :

$$(\mathbf{y}_s)_n = \sum_{k=1}^K a_{s,k} \exp\left(-\frac{(n - c_{s,k})^2}{2w_{s,k}^2}\right) + \alpha_s \exp\left(-\frac{n}{\beta_s}\right) + (\mathbf{b}_s)_n, \quad (1)$$

où $(\mathbf{y}_s)_n$ est le n -ème élément ($n \in \{1, \dots, N\}$) du vecteur \mathbf{y}_s . Les paramètres de la k -ème raie du spectre s sont son centre $c_{s,k}$, son amplitude $a_{s,k}$ et sa largeur $w_{s,k}$, les paramètres du continuum sont α_s et β_s et \mathbf{b}_s est un bruit additif modélisant les erreurs de modèle et de mesure. Le nombre K de raies est supposé connu et identique pour chaque spectre. Des séquences réelles de photo-électrons regroupent typiquement une centaine de spectres ($S \approx 100$), chacun d'une taille $N \approx 100$ et on peut s'attendre à trouver $K \approx 5$ raies. Dans la suite, nous utiliserons les notations ϕ_s et γ_s qui représentent respectivement la somme des raies et le continuum pour le spectre s , l'équation (1) devient alors :

$$\mathbf{y}_s = \phi_s + \gamma_s + \mathbf{b}_s \quad (2)$$

avec

$$(\phi_s)_n = \sum_{k=1}^K a_{s,k} \exp\left(-\frac{(n - c_{s,k})^2}{2w_{s,k}^2}\right), \quad (3)$$

$$(\gamma_s)_n = \alpha_s \exp\left(-\frac{n}{\beta_s}\right). \quad (4)$$

En outre, dans le but de pouvoir suivre l'évolution des raies au cours de la séquence et de régulariser leur évolution, il est nécessaire d'introduire la notion de trajectoire : une trajectoire est l'association de S raies, chacune dans un spectre différent, et on suppose que l'évolution des raies d'une même trajectoire est douce. Aussi, dans les données, il y a K trajectoires. On suppose dans l'équation (1) que les raies ne sont pas encore identifiées aux trajectoires : par exemple, les positions des raies de la trajectoire k ne sont pas forcément $\{c_{1,k}, c_{2,k}, \dots, c_{S,k}\}$. En fait, l'étiquette de chaque raie est définie grâce à la variable $z_{s,k} \in \{1, \dots, K\}$, si bien que les positions des raies de la trajectoire k correspondent au vecteur

$$\mathbf{c}^k = \{c_{s,k_0} | z_{s,k_0} = k, \forall s, \forall k_0\}. \quad (5)$$

De la même manière, on notera respectivement \mathbf{a}^k et \mathbf{w}^k les amplitudes et les largeurs de la trajectoire k . Au final, les variables d'intérêt à estimer sont : $\mathbf{c} \in \mathbb{R}^{S \times K}$, $\mathbf{a} \in \mathbb{R}^{S \times K}$, $\mathbf{w} \in \mathbb{R}^{S \times K}$, $\mathbf{z} \in \mathbb{R}^{S \times K}$ auxquelles il faut associer les paramètres du continuum $\boldsymbol{\alpha} \in \mathbb{R}^S$ et $\boldsymbol{\beta} \in \mathbb{R}^S$ et les six hyperparamètres qui seront introduits dans la section 2.1.

2.1. Lois a priori

Le bruit est supposé blanc gaussien de moyenne nulle et de variance r_b (constante pour tous les spectres) :

$$\forall s, n, \quad (\mathbf{b}_s)_n | r_b \sim \mathcal{N}(0, r_b). \quad (6)$$

Les positions des raies de la trajectoire \mathbf{c}^k ($\forall k \in \{1, \dots, K\}$) sont supposées évoluer lentement. On souhaite donc régulariser la solution à l'aide d'un a priori de douceur. Aussi, \mathbf{c}^k est modélisé à l'aide d'un champ de Markov unidimensionnel gaussien (Geman, Yang, 1995 ; Rue, Held, 2005). Plus exactement, la loi a priori est tronquée afin d'interdire certaines valeurs de \mathbf{c}^k :

$$p(\mathbf{c}^k | \mathbf{z}) \propto \frac{1}{r_c^{S/2}} \exp\left(-\frac{1}{2r_c} \|D\mathbf{c}^k\|^2\right) \mathbb{I}_{[1,N]}(\mathbf{c}^k), \quad (7)$$

où r_c est un hyperparamètre qui caractérise l'évolution du paramètre \mathbf{c}^k , $\|\cdot\|$ représente la norme L^2 et $\mathbb{I}_{[1,N]}(\mathbf{c}^k)$ est la fonction indicatrice qui permet de définir le support du paramètre \mathbf{c}^k : $\mathbb{I}_{[1,N]}(\mathbf{c}^k) = 1$ si $\mathbf{c}^k \in [1, N]^S$, 0 sinon. La matrice D définit une dérivée discrète (une dérivée d'ordre un pénalise toute évolution, une dérivée d'ordre deux favorise une évolution linéaire) et code la structure des voisinages : par exemple, si D correspond à la matrice de la dérivée première, alors les cliques correspondent à deux éléments consécutifs du vecteur \mathbf{c}^k . Par ailleurs, \mathbf{z} apparaît implicitement dans l'équation (7) car \mathbf{c}^k dépend de \mathbf{z} (cf. équation (5)).

Comme l'espace $[1, N]$ est suffisamment grand par rapport aux faibles valeurs de r_c , la troncature est négligeable et donc la densité de probabilité réelle $p(\mathbf{c}^k | \mathbf{z})$ peut

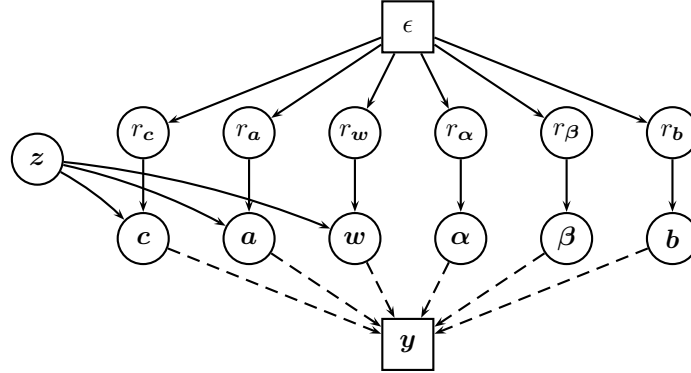


Figure 1. Graphe du modèle bayésien. Les rectangles et les cercles représentent respectivement les variables connues (ou fixées) et les variables aléatoires inconnues. Les flèches en traits continus indiquent des dépendances stochastiques entre variables alors que les flèches en pointillés indiquent des dépendances déterministes

être approximée par l'équation (7). Enfin, les vecteurs des positions \mathbf{c}^k de chaque trajectoire sont supposés mutuellement indépendants.

De la même manière, on définit les lois a priori pour les amplitudes \mathbf{a}^k et les largeurs \mathbf{w}^k des trajectoires ($\forall k \in \{1, \dots, K\}$), et les paramètres du continuum, à savoir α et β . Le support de ces variables est alors \mathbb{R}^+ .

Les étiquettes sont supposées être distribuées uniformément sur l'ensemble \mathcal{S}_K des permutations de $\{1, \dots, K\}$:

$$\forall s, \quad \mathbf{z}_s \sim \mathcal{U}_{\mathcal{S}_K}. \quad (8)$$

La variance du bruit r_b suit une loi inverse gamma :

$$r_b \sim \mathcal{IG}(\epsilon, \epsilon) \quad (9)$$

où ϵ est très petit afin que l'a priori ressemble à la distribution de Jeffreys traditionnellement utilisée mais malheureusement impropre.

Enfin, les hyperparamètres r_c, r_a, r_w (relatifs à l'évolution de la position, de l'amplitude et de la largeur des raies au cours des séquences), r_α ou r_β (relatifs à l'évolution du continuum) doivent être faibles pour permettre une évolution en douceur. Nous choisissons donc un a priori similaire à celui utilisé pour la variance du bruit :

$$\begin{aligned} r_c &\sim \mathcal{IG}(\epsilon, \epsilon), & r_a &\sim \mathcal{IG}(\epsilon, \epsilon), & r_w &\sim \mathcal{IG}(\epsilon, \epsilon), \\ r_\alpha &\sim \mathcal{IG}(\epsilon, \epsilon), & r_\beta &\sim \mathcal{IG}(\epsilon, \epsilon), \end{aligned} \quad (10)$$

avec ϵ très petit (cf. équation (9)).

Le graphe de dépendance entre les différents paramètres est représenté figure 1.

Avant de détailler la loi a posteriori résultant des lois a priori présentées ci-avant, il est intéressant de remarquer que le terme d'attache aux données obtenu par l'a priori sur le bruit permet de lier les paramètres d'un même spectre tandis que le terme de régularisation (équation (7)) permet de lier les paramètres d'une même trajectoire. Ainsi, d'une certaine manière, chacun de ces deux a priori joue un rôle sur l'un des deux axes des données (l'axe vertical des temps ou l'axe horizontal des énergies).

2.2. Lois a posteriori

La loi a posteriori globale s'exprime à l'aide du théorème de Bayes :

$$\begin{aligned}
p(\mathbf{c}, \mathbf{a}, \mathbf{w}, \boldsymbol{\alpha}, \boldsymbol{\beta}, \mathbf{z}, r_{\mathbf{b}}, r_{\mathbf{c}}, r_{\mathbf{a}}, r_{\mathbf{w}}, r_{\boldsymbol{\alpha}}, r_{\boldsymbol{\beta}} | \mathbf{y}) &\propto \\
\frac{1}{r_{\mathbf{b}}^{NS/2}} \exp\left(-\frac{1}{2r_{\mathbf{b}}} \sum_{s=1}^S \|\mathbf{y}_s - \boldsymbol{\phi}_s - \boldsymbol{\gamma}_s\|^2\right) &\times \frac{e^{-\epsilon/r_{\mathbf{b}}}}{r_{\mathbf{b}}^{\epsilon+1}} \mathbb{I}_{\mathbb{R}^+}(r_{\mathbf{b}}) \\
\times \frac{1}{r_{\mathbf{c}}^{KS/2}} \exp\left(-\frac{1}{2r_{\mathbf{c}}} \sum_{l=1}^K \|D\mathbf{c}^l\|^2\right) &\mathbb{I}_{[1,N]}(\mathbf{c}) \times \frac{e^{-\epsilon/r_{\mathbf{c}}}}{r_{\mathbf{c}}^{\epsilon+1}} \mathbb{I}_{\mathbb{R}^+}(r_{\mathbf{c}}) \\
\times \frac{1}{r_{\mathbf{a}}^{KS/2}} \exp\left(-\frac{1}{2r_{\mathbf{a}}} \sum_{l=1}^K \|D\mathbf{a}^l\|^2\right) &\mathbb{I}_{\mathbb{R}^+}(\mathbf{a}) \times \frac{e^{-\epsilon/r_{\mathbf{a}}}}{r_{\mathbf{a}}^{\epsilon+1}} \mathbb{I}_{\mathbb{R}^+}(r_{\mathbf{a}}) \\
\times \frac{1}{r_{\mathbf{w}}^{KS/2}} \exp\left(-\frac{1}{2r_{\mathbf{w}}} \sum_{l=1}^K \|D\mathbf{w}^l\|^2\right) &\mathbb{I}_{\mathbb{R}^+}(\mathbf{w}) \times \frac{e^{-\epsilon/r_{\mathbf{w}}}}{r_{\mathbf{w}}^{\epsilon+1}} \mathbb{I}_{\mathbb{R}^+}(r_{\mathbf{w}}) \\
\times \frac{1}{r_{\boldsymbol{\alpha}}^{S/2}} \exp\left(-\frac{1}{2r_{\boldsymbol{\alpha}}} \|D\boldsymbol{\alpha}\|^2\right) &\mathbb{I}_{\mathbb{R}^+}(\boldsymbol{\alpha}) \times \frac{e^{-\epsilon/r_{\boldsymbol{\alpha}}}}{r_{\boldsymbol{\alpha}}^{\epsilon+1}} \mathbb{I}_{\mathbb{R}^+}(r_{\boldsymbol{\alpha}}) \\
\times \frac{1}{r_{\boldsymbol{\beta}}^{S/2}} \exp\left(-\frac{1}{2r_{\boldsymbol{\beta}}} \|D\boldsymbol{\beta}\|^2\right) &\mathbb{I}_{\mathbb{R}^+}(\boldsymbol{\beta}) \times \frac{e^{-\epsilon/r_{\boldsymbol{\beta}}}}{r_{\boldsymbol{\beta}}^{\epsilon+1}} \mathbb{I}_{\mathbb{R}^+}(r_{\boldsymbol{\beta}}) \\
\times \prod_s \mathbb{I}_{\mathcal{S}_K}(\mathbf{z}_s) &
\end{aligned} \tag{11}$$

où, on le rappelle, $\boldsymbol{\phi}_s$ et $\boldsymbol{\gamma}_s$ représentent respectivement la somme des raies et le continuum du spectre s . Nous utilisons un échantillonneur de Gibbs pour simuler cette loi a posteriori (section 3). Il consiste à simuler chaque variable selon sa loi a posteriori conditionnelle (Robert, Casella, 2004). Ces dernières sont obtenues directement à partir de la loi a posteriori globale et elles sont présentées ci-après.

Les centres, positions et β_s sont donc distribués selon

$$\forall s, k \quad p(c_{s,k} | \dots) \propto \exp\left(-\frac{\|\mathbf{y}_s - \phi_s - \gamma_s\|^2}{2r_b} - \frac{\|D\mathbf{c}^l\|^2}{2r_c}\right) \mathbb{I}_{[1,N]}(c_{s,k}), \quad (12)$$

$$p(w_{s,k} | \dots) \propto \exp\left(-\frac{\|\mathbf{y}_s - \phi_s - \gamma_s\|^2}{2r_b} - \frac{\|D\mathbf{w}^l\|^2}{2r_w}\right) \mathbb{I}_{\mathbb{R}^+}(w_{s,k}), \quad (13)$$

$$p(\beta_s | \dots) \propto \exp\left(-\frac{\|\mathbf{y}_s - \phi_s - \gamma_s\|^2}{2r_b} - \frac{\|D\beta\|^2}{2r_\beta}\right) \mathbb{I}_{\mathbb{R}^+}(\beta_s) \quad (14)$$

avec $l = z_{s,k}$ et « ... » représente l'ensemble des variables à l'exception de celle dont on considère la probabilité.

Les amplitudes et α_s sont distribués suivant une loi normale à support positif (cf. annexe Annexe A) :

$$\forall s, k \quad p(a_{s,k} | \dots) \propto \exp\left(-\frac{(a_{s,k} - \mu_{s,k})^2}{2\rho_{s,k}}\right) \mathbb{I}_{\mathbb{R}^+}(a_{s,k}), \quad (15)$$

$$p(\alpha_s | \dots) \propto \exp\left(-\frac{(\alpha_s - \lambda_s)^2}{2\nu_s}\right) \mathbb{I}_{\mathbb{R}^+}(\alpha_s) \quad (16)$$

où (voir l'annexe Annexe A pour les expressions des autres variables) :

$$\rho_{s,k} = \frac{r_a r_b}{r_a \boldsymbol{\varphi}_{s,k}^T \boldsymbol{\varphi}_{s,k} + r_b A_{s,k}} \quad \mu_{s,k} = \rho_{s,k} \left(\frac{\mathbf{x}_s^T \boldsymbol{\varphi}_{s,k}}{r_b} - \frac{B_{s,k}}{2r_a} \right) \quad (17)$$

$$\nu_s = \frac{r_a r_b}{r_a \boldsymbol{\psi}_s^T \boldsymbol{\psi}_s + r_b E_s} \quad \lambda_s = \nu_s \left(\frac{(\mathbf{y}_s - \phi_s)^T \boldsymbol{\psi}_s}{r_b} - \frac{F_s}{2r_a} \right) \quad (18)$$

Les étiquettes sont distribuées selon

$$p(\mathbf{z} | \dots) \propto \exp\left[-\sum_l \left(\frac{\|D\mathbf{c}^l\|^2}{2r_c} + \frac{\|D\mathbf{a}^l\|^2}{2r_a} + \frac{\|D\mathbf{w}^l\|^2}{2r_w} \right)\right] \prod_s \mathbb{I}_{S_K}(\mathbf{z}_s). \quad (19)$$

Enfin, les lois a posteriori conditionnelles des hyperparamètres sont des lois inverse gamma :

$$r_b | \dots \sim \mathcal{IG}(NS/2 + \epsilon, \sum_s \|\mathbf{y}_s - \phi_s - \gamma_s\|^2/2 + \epsilon), \quad (20)$$

$$r_c | \dots \sim \mathcal{IG}(KS/2 + \epsilon, \sum_l \|D\mathbf{c}^l\|^2/2 + \epsilon), \quad (21)$$

$$r_a | \dots \sim \mathcal{IG}(KS/2 + \epsilon, \sum_l \|D\mathbf{a}^l\|^2/2 + \epsilon), \quad (22)$$

$$r_w | \dots \sim \mathcal{IG}(KS/2 + \epsilon, \sum_l \|D\mathbf{w}^l\|^2/2 + \epsilon), \quad (23)$$

$$r_\alpha | \dots \sim \mathcal{IG}(S/2 + \epsilon, \|D\alpha\|^2/2 + \epsilon), \quad (24)$$

$$r_\beta | \dots \sim \mathcal{IG}(S/2 + \epsilon, \|D\beta\|^2/2 + \epsilon). \quad (25)$$

3. Algorithme MCMC et recuit simulé

Un échantillonneur de Gibbs permet d'échantillonner la loi a posteriori : c'est en effet un algorithme souvent utilisé pour simuler des champs de Markov (Wang *et al.*, 2000), en particulier dans le cadre de la décomposition spectroscopique (Fischer, Dose, 2001 ; Gulam Razul *et al.*, 2003 ; Haan, Godsill, 2002). L'échantillonneur de Gibbs consiste à simuler les inconnues par bloc suivant leur loi a posteriori conditionnellement aux autres blocs du modèle ; on peut en trouver une description dans (Robert, Casella, 2004). Pour notre implémentation, chaque bloc de variable correspond en fait à chaque scalaire inconnu, excepté pour les étiquettes z qui sont échantillonnées toutes ensemble. Nous avons choisi d'échantillonner les blocs dans un ordre aléatoire à chaque itération de l'échantillonneur (*random sweep*). Les techniques de simulation de chaque bloc dépendent alors de la complexité des lois conditionnelles associées à ces blocs.

Ainsi, les paramètres $c_{s,k}$, $w_{s,k}$ et β_s ($\forall s, k$) sont échantillonnés suivant un algorithme de Metropolis-Hastings à marche aléatoire (Robert, Casella, 2004) : pour chacun de ces paramètres, un candidat est généré suivant une loi normale centrée sur la valeur courante du paramètre et de variance fixée a priori. La loi normale est restreinte au support que peuvent prendre les paramètres.

Les lois conditionnelles des paramètres $a_{s,k}$ et α_s ($\forall s, k$) étant des gaussiennes, elles peuvent être simulées directement. De même, les hyperparamètres r_c , r_a , r_w , r_α , r_β et r_b suivent des lois inverse gamma et peuvent donc aussi être échantillonnés directement, par exemple à l'aide d'une méthode d'acceptation-rejet (Devroye, 1986).

Enfin, les étiquettes sont échantillonnées à l'aide de l'algorithme symétrique de Metropolis-Hastings décrit ci-après. La simulation de la loi $p(z|\dots)$ n'est en effet pas évidente car c'est une loi discrète sur un ensemble de $K!S$ valeurs. Pour améliorer la simulation de la loi, 100 candidats sont proposés à chaque balayage de l'échantillonneur de Gibbs.

En pratique, l'échantillonneur de Gibbs peut rester bloqué dans un minimum local (Robert, Casella, 2004), c'est pourquoi il est inclus ici dans un schéma de recuit simulé qui est un algorithme populaire pour l'optimisation de champs de Markov (Geman, Geman, 1984 ; Kirkpatrick *et al.*, 1983). Ce schéma consiste à simuler non pas la loi a posteriori conditionnelle $p(\cdot)$ mais la loi $p^{1/T}(\cdot)$ où T est la « température » qui décroît continûment au cours des itérations (une descente de température géométrique est utilisée). De cette manière, l'estimation du MAP est simplement l'échantillon de la chaîne de Markov qui maximise la loi a posteriori. Toutefois, le recuit simulé n'est pas implémenté sur les variables $a_{s,k}$ et α_s ainsi que sur les hyperparamètres r_c , r_a , r_w , r_α , r_β et r_b car il peut introduire des problèmes numériques et de divergence : cette approche peut être vue comme une méthode de recuit simulé avec différentes températures (Perret *et al.*, 2011).

Algorithme 1. Échantillonneur des étiquettes

1. Simuler :

$$k_1, k_2 \sim \mathcal{U}_{\{1, \dots, K\}} \text{ avec } k_1 \neq k_2,$$

$$s_1, s_2 \sim \mathcal{U}_{\{1, \dots, S\}} \text{ (avec potentiellement } s_1 = s_2);$$

2. Poser comme candidat : $\tilde{z} = z$, sauf que les étiquettes k_1 et k_2 entre $s = s_1$ et $s = s_2$ sont permutées ;

3. Accepter le candidat avec la probabilité $\min(\zeta, 1)$ avec

$$\zeta = \frac{p(\tilde{z} | \dots)}{p(z | \dots)} = \exp \left[\sum_l \left(\frac{\|Dc^l\|^2 - \|D\tilde{c}^l\|^2}{2r_c} + \frac{\|Da^l\|^2 - \|D\tilde{a}^l\|^2}{2r_a} + \frac{\|Dw^l\|^2 - \|D\tilde{w}^l\|^2}{2r_w} \right) \right].$$

4. Résultats sur données simulées

Pour quantifier les apports de la méthode de décomposition conjointe non supervisée proposée, nous la comparons aux méthodes de décomposition séquentielle et de décomposition conjointe supervisée présentées dans (Mazet, 2011). La seule modification apportée à ces deux dernières méthodes est la prise en compte du continuum. Ainsi, les hyperparamètres sont fixés comme indiqués dans (Mazet, 2011) ; la décomposition séquentielle ne permet évidemment pas d'estimer les étiquettes des raies ni d'introduire une douceur sur leur évolution ; et la décomposition conjointe supervisée simule les étiquettes avec l'algorithme proposé dans (Mazet, 2011). Chaque méthode effectue l'estimation grâce à un échantillonneur de Gibbs couplé à du recuit simulé sur 5 000 itérations afin d'estimer le MAP (utiliser plus d'itérations n'a pas conduit à une amélioration significative des résultats).

Les trois méthodes ont été testées sur 30 séquences simulées avec $S = 10$, $K = 4$, $I = 5000$ itérations, D est la dérivée seconde et les températures initiales et finales du recuit simulé sont respectivement fixées à 50 et 1. Les simulations ont été effectuées avec Matlab 7.10.0 sur un Pentium E5200 à 2,5 GHz sous Ubuntu 10.04. Le code Matlab et les données sont disponibles librement à l'adresse lsiit-miv.u-strasbg.fr/mazet/jointdec. Les performances ont été comparées en calculant l'erreur quadratique moyenne entre les signaux observés \mathbf{y} et reconstruits $\hat{\mathbf{y}}$:

$$\text{EQM} = \frac{1}{SN} \sum_{s=1}^S \sum_{n=1}^N ((\mathbf{y}_s)_n - (\hat{\mathbf{y}}_s)_n)^2. \quad (26)$$

Tableau 1. Performance des décompositions séquentielle (DS), conjointe supervisée (DCS) et conjointe non supervisée (DCNS, méthode proposée) pour 5 000 itérations. L'EQM correspond à l'erreur quadratique moyenne calculée entre les signaux mesurés et reconstruits

	DS	DCS	DCNS
EQM	$11,54 \cdot 10^{-2}$	$7,40 \cdot 10^{-2}$	$3,78 \cdot 10^{-2}$
temps de calcul	96.95 s	118.56 s	179.83 s

La méthode proposée est la plus longue en termes de temps de calcul puisqu'elle possède plus de paramètres à estimer que les deux autres méthodes et que la simulation des étiquettes est plus longue que dans la version non supervisée. Cependant, elle permet d'obtenir de meilleures estimations.

Ainsi, les méthodes de décomposition séquentielle et conjointe non supervisée obtiennent des EQM respectives de $11,54 \cdot 10^{-2}$ et $3,78 \cdot 10^{-2}$: notre approche est donc plus performante. Or, la décomposition séquentielle devrait en théorie être meilleure en termes d'EQM puisque son critère est, justement, l'EQM alors que le critère de notre méthode contient, en plus de l'EQM, un terme de régularisation. Aussi, les résultats obtenus s'expliquent essentiellement par le fait que la décomposition séquentielle reste bloquée dans des minima locaux. En revanche, l'approche conjointe, qui tient compte des estimations obtenues sur les spectres voisins grâce à l'a priori de douceur, favorise la recherche de raies dans certaines zones de l'espace : en d'autres termes, l'estimation obtenue pour un spectre aide l'estimation des spectres voisins. Cela met également en évidence le bon fonctionnement de la simulation des étiquettes.

Un exemple d'une de ces simulations est représenté figures 2 et 3. Dans l'ordre lexicographique : centres c , amplitudes a , largeurs w (matrices de taille $S \times K$), paramètres du continuum α et β (vecteurs de taille S) en fonction de l'instant d'acquisition s . Les lignes en pointillés représentent les paramètres réels, les croix et les cercles représentent respectivement les estimations obtenues avec les décompositions séquentielle et conjointe (les étiquettes des raies sont représentées par les couleurs). Certaines croix sont en dehors des axes. Sur cet exemple, les raies évoluent de manière volontairement importante et il y a beaucoup de chevauchements de raies. Il est clair qu'une décomposition séquentielle n'est pas satisfaisante (figure 2) ; en effet, les paramètres estimés (croix noires) sont parfois éloignés des trajectoires réelles (certaines estimations ne sont d'ailleurs pas visibles car en dehors des axes). Au contraire, la décomposition conjointe estime relativement bien les raies (les cercles sont souvent proches des trajectoires) et permet en outre de suivre leur évolution. L'estimation obtenue est meilleure que celle de la décomposition séquentielle, mais on peut cependant noter quelques écarts avec les trajectoires réelles :

- la raie verte sur le spectre $s = 1$ (figure 3, zone 1) est mal estimée en amplitude et en largeur car la raie réelle correspondante est située à l'une des extrémités du spectre. L'estimation est donc délicate car seule une partie de la raie existe dans les données ;

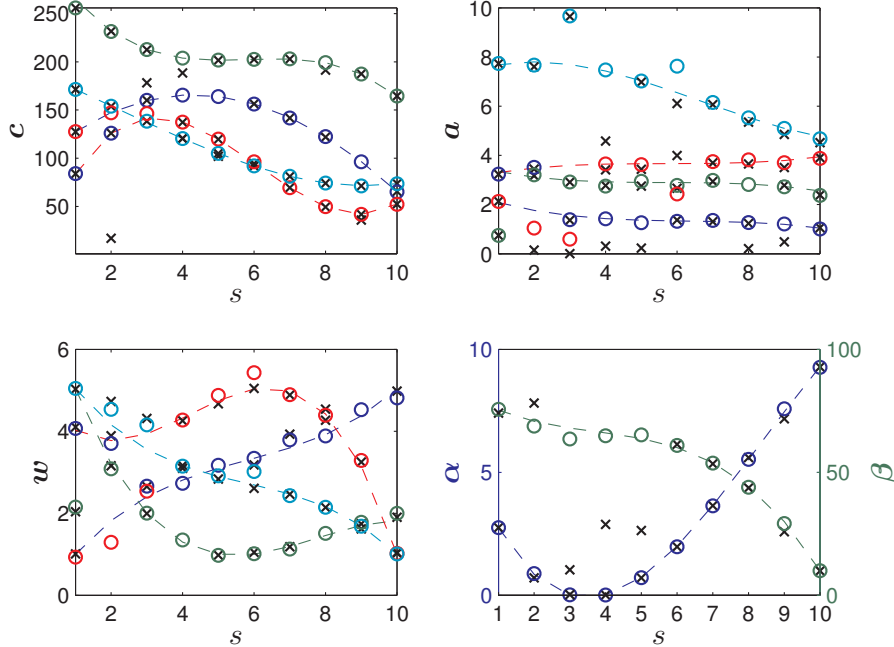


Figure 2. Exemple d'estimation obtenue avec les décompositions séquentielle et conjointe non supervisée. Les couleurs des figures de l'article sont visibles sur la version en ligne

– un problème d'estimation apparaît également sur le spectre $s = 6$ entre les raies rouge et cyan (figure 3, zone 2). Pour pouvoir augmenter l'amplitude de la raie rouge, il faudrait simultanément diminuer celle de la raie cyan : c'est typiquement un problème dû à l'échantillonneur de Gibbs qui met à jour chaque variable séparément (des solutions à ce problème ont été proposées par (Ge *et al.*, 2011) par exemple);

– enfin, il existe une permutation sur les deux premiers spectres entre les trajectoires rouge et bleu foncé (figure 2, centres; figure 3, zone 3). Les trajectoires réelles ont des évolutions brusques (cf. positions de la trajectoire rouge) alors que l'algorithme a proposé une estimation correspondant à des trajectoires globalement plus linéaires (au détriment d'une variation importante de l'amplitude de la trajectoire rouge). L'analyse de tels résultats où les trajectoires sont entremêlées est très délicate car il existe plusieurs solutions envisageables.

Enfin, la méthode proposée surpasse également la décomposition conjointe supervisée (dont l'EQM est $7,40 \cdot 10^{-2}$). Cela est dû au fait que le réglage des hyperparamètres n'est pas simple. En effet, ils contrôlent l'importance relative des termes de données et de régularisation et cela implique qu'ils doivent être réglés en fonction du rapport signal-sur-bruit et de l'évolution attendue des raies. Cela met clairement en évidence les apports de la méthode non supervisée.

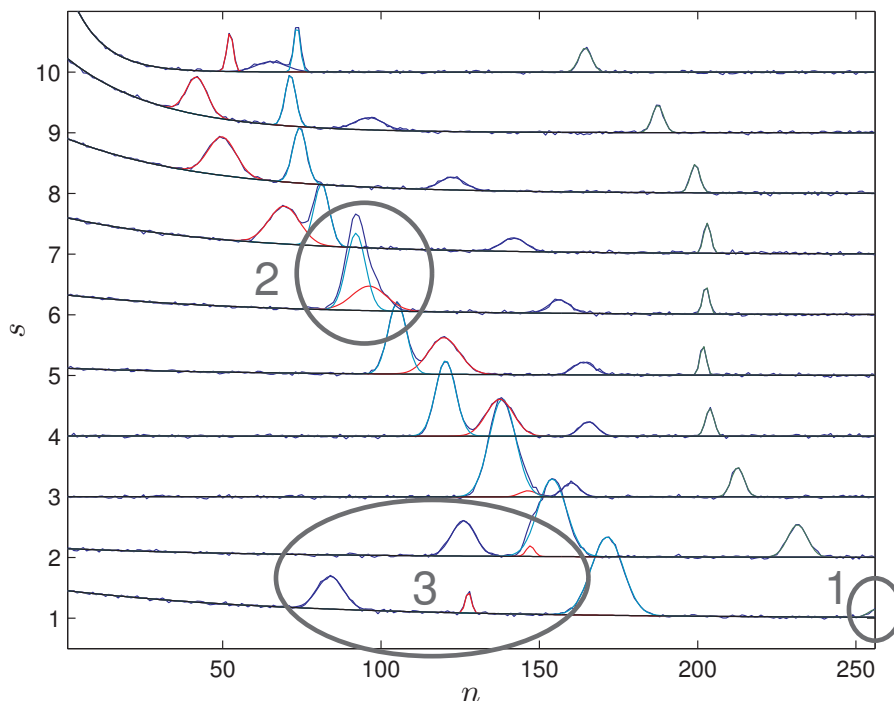


Figure 3. Exemple d'estimation obtenue avec la décomposition conjointe non supervisée. Les données sont représentées en bleu, les raies estimées en couleur (identiques à celles de la figure 2) et le continuum en noir

5. Résultats sur spectres de photo-électrons

L'expérience porte sur l'étude de la relaxation d'un atome de baryum (Ba) sur un agrégat (gouttelette isolée de plusieurs centaines d'atomes) d'argon (Ar) : il s'agit, pour ce système Ba-Ar_n, de comprendre le mécanisme qui va conduire l'atome de baryum excité à perdre son énergie à travers une interaction avec un milieu inerte et froid (34 K).

5.1. Dispositif expérimental

Le dispositif expérimental est schématisé figure 4 ; une présentation détaillée peut être trouvée dans (Masson *et al.*, 2010). L'expérience est effectuée au SLIC (*Saclay Laser-Matter Interaction Center*) sur le laser femtoseconde LUCA.

Un jet pulsé d'agrégats d'argon (environ 800 atomes en moyenne) est créé par détente supersonique (figure 4, (1)). Ce jet passe à travers un four chauffé (figure 4, (2)) qui dépose des atomes de baryum par collision sur les agrégats (la probabilité d'avoir plusieurs atomes de baryum déposés sur les agrégats d'argon est faible). Le jet chargé

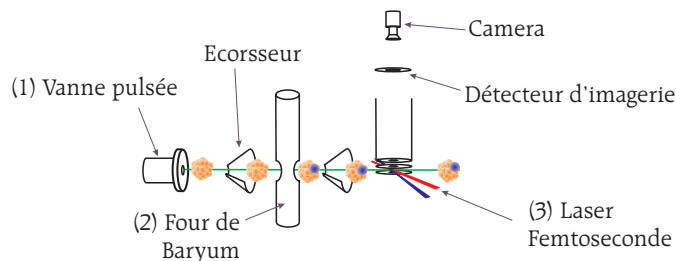


Figure 4. Dispositif expérimental. L'agrégat d'argon est représenté par le complexe orange et l'atome de baryum par le point bleu

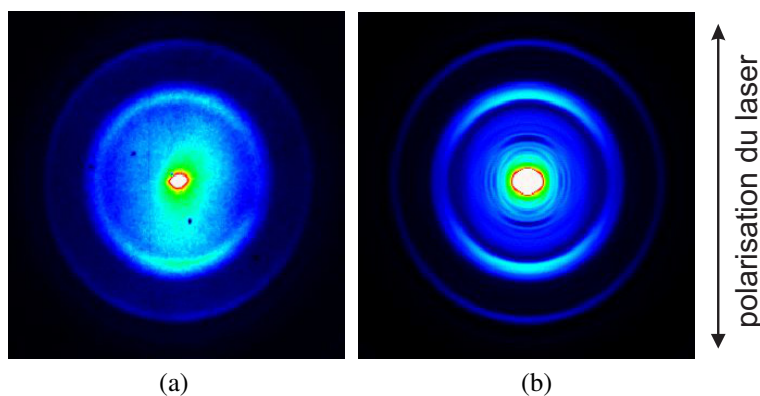


Figure 5. Images brute (a) et inversée (b) pour un délai de 70 ps. La distance radiale de l'image inversée est proportionnelle à la vitesse des photo-électrons et l'angle correspond à leur direction d'émission par rapport à la polarisation du laser de sonde qui est verticale sur la figure

continue ensuite vers la zone d'interaction de deux lasers. Le premier laser femtoseconde de 266 nm (appelé « pompe ») excite les atomes de baryum déposés (figure 4, (3)) et le second laser de 400 nm ou 800 nm (appelé « sonde »), de même longueur temporelle, direction et polarisation horizontale, vient alors sonder l'état du système à un instant variable après l'excitation. Les photo-électrons émis sont distribués sur une sphère centrée sur les contre-ions résiduels, cette distribution est projetée verticalement par un champ électrique sur un détecteur sensible en position (technique de cartographie d'imagerie de vitesse (Eppink, Parker, 1997)). L'acquisition d'une multitude d'événements, obtenus pour un même délai entre la pompe et la sonde, conduit à la représentation de la projection des distributions d'électrons : on obtient des images brutes comme celle représentée figure 5(a). La direction de polarisation du laser de sonde induit une symétrie cylindrique d'émission des électrons selon son axe. La distribution radiale représente la distribution de vitesse des électrons et la distribution angulaire représente l'anisotropie de la distribution des électrons (dans le cadre de ce travail, seule la distribution radiale est considérée). La transformée d'Abel

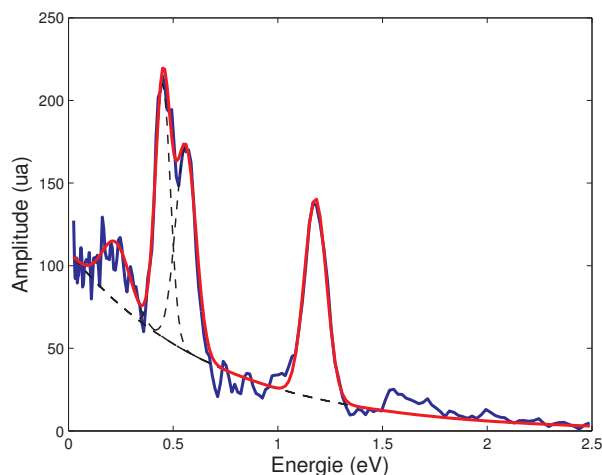


Figure 6. Exemple de spectre de photo-électrons (en bleu) obtenu à partir de l'image inversée 5(b); ua : unité arbitraire. Cet exemple représente également la décomposition obtenue avec notre méthode dans le cas de $K = 4$ raies (en noir) et le spectre reconstruit (en rouge)

inverse permet alors de remonter de la projection d'une distribution ayant une telle symétrie vers la coupe de cette distribution (figure 5(b)), donc vers une représentation du nombre de photo-électrons émis en fonction de leur vitesse : le spectre de photo-électrons. La procédure pBASEX (Garcia *et al.*, 2004) permet de réaliser cette inversion. Le spectre de photo-électrons obtenu à partir de l'image inversée de la figure 5(b) est représenté figure 6. Le suivi de ces spectres en fonction du délai entre la pompe et la sonde donne alors une information sur le processus de relaxation de l'état électronique peuplé par la pompe. Afin d'optimiser le temps de l'acquisition, l'échantillonnage temporel du délai entre la pompe et la sonde n'est pas constant mais varie dans les différentes régions explorées. Une expérience complète consiste à enregistrer un spectre pour plusieurs délais pompe-sonde pour obtenir finalement les données que nous traitons (figure 8).

Dans le cas d'un système complexe, possédant plusieurs voies de déséxcitation (perte d'énergie) ou plusieurs états initialement peuplés, le spectre de photo-électrons apparaît comme la somme de tous les processus prenant place en parallèle. C'est ce qu'il se produit lors de l'étude de la relaxation de l'atome de baryum du système Ba-Ar_n. Les spectres obtenus peuvent être modélisés comme la superposition de raies dont les propriétés (position, amplitude, largeur) peuvent varier en fonction du délai entre la pompe et la sonde.

5.2. Résultats obtenus

La séquence traitée est constituée de 44 spectres (couvrant une durée de 3,47 ps), chacun étant constitué de $N = 182$ mesures (de 0,02 eV à 2,52 eV) avec des échantillonnages en temps (s) et en énergie (n) irréguliers (l'échantillonnage temporel irrégulier n'a toutefois pas été pris en compte afin de simplifier l'algorithme et parce que cela ne semble pas porter préjudice à la qualité des résultats). Sur la figure 8, l'axe horizontal correspond à s et l'axe vertical à n et on cherche à quantifier l'évolution des raies en fonction de s .

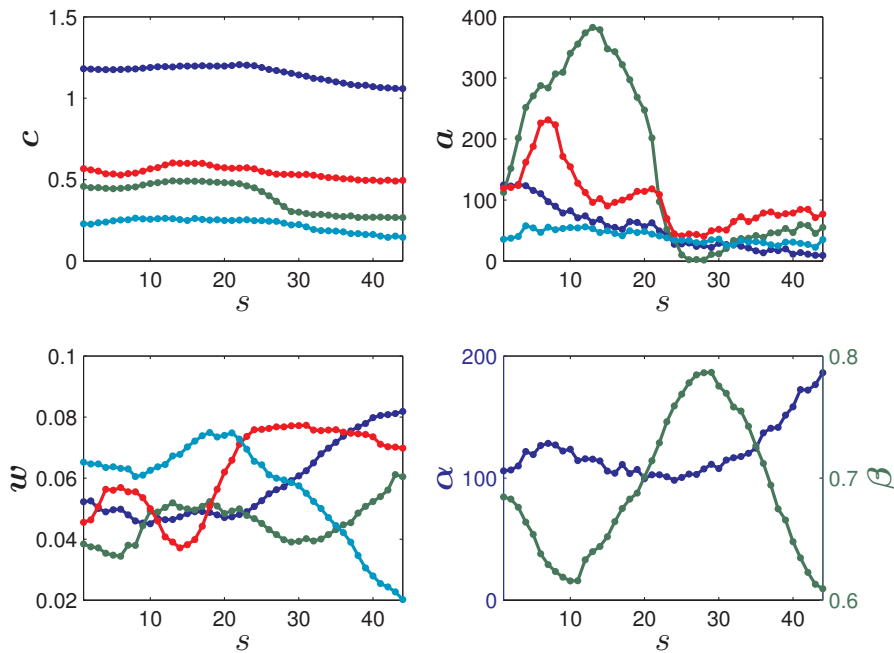


Figure 7. Paramètres estimés (cf. légende de la figure 2) : centres c , amplitudes a et largeurs w des raies ; paramètres α et β du continuum. Chaque couleur correspond à une étiquette

L'approche proposée a été exécutée avec $K = 4$ raies, $I = 10^4$ itérations et en utilisant une dérivée d'ordre un (priviliégiant donc aucune évolution, équation (7)). La température du recuit simulé décroît de 10 à 0,1. Les résultats sont représentés figures 7 et 8 : on y voit que l'évolution des paramètres est assez lisse et que les raies sont bien classées (il n'y a pas de permutation entre les raies). Cela valide l'algorithme utilisé pour l'échantillonnage des étiquettes alors que la méthode proposée dans (Mazet, 2011) ne fournit pas de bons résultats lorsque S est grand. Il apparaît également que les raies évoluent plus doucement que si l'estimation avait été séquentielle, telle qu'elle a été implémentée dans (Masson *et al.*, 2010) (résultat non présenté). On s'aperçoit par ailleurs que l'énergie d'une raie varie dans le temps : son énergie aug-

mente légèrement de 1,18 à 1,20 eV entre 0 et 0,4 ps puis elle diminue et atteint 1,05 eV à 2,5 ps. Après 2,5 ps, la raie a une amplitude très faible et une largeur très importante, ce qui est typique d'un pic qui disparaît. C'est une limitation de l'approche dans la mesure où nous avons fait l'hypothèse, qui n'est pas toujours vérifiée, que le nombre de raies est constant dans le temps. Malgré tout, les résultats permettent de confirmer quantitativement les observations qualitatives exprimées dans (Masson *et al.*, 2010), à savoir que l'énergie d'un niveau électronique du système Ba-Ar_n varie dans le temps. Cela suggère que le niveau énergétique correspondant est très sensible à l'environnement des atomes d'argon, et que le baryum est solvaté dans le groupe d'atomes d'argon.

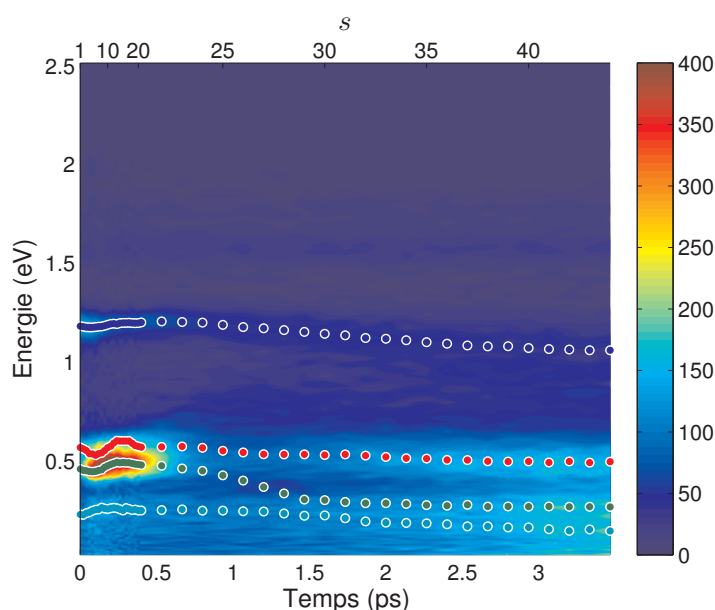


Figure 8. Spectres de photo-électrons : les données sont tracées en fonction de l'énergie des électrons (n , axe vertical) et du temps (s , axe horizontal) et les valeurs des spectres sont indiquées par l'échelle de couleur. Les points correspondent aux centres estimés des quatre raies et les couleurs de ces points représentent leur étiquette

6. Conclusion

Nous avons présenté une approche originale pour estimer les paramètres des raies d'une séquence de signaux spectroscopiques. L'idée principale est d'effectuer une décomposition en traitant la séquence entière plutôt que chaque spectre séparément. Pour cela, des champs de Markov ont été introduits dans le modèle bayésien pour adoucir l'évolution des paramètres estimés. L'approche proposée est complètement non supervisée et la solution est obtenue à l'aide d'un algorithme MCMC combiné à un schéma

de recuit simulé. Des simulations ont montré la pertinence de cette approche et les résultats obtenus sur des spectres de photo-électrons ont confirmé le point de vue des experts.

Les futurs travaux se concentreront principalement sur l'estimation du nombre de raies à l'aide de l'algorithme RJMCMC (Green, 1995), ce point étant à l'heure actuelle la principale limitation de l'algorithme qui peut le rendre inefficace sur certaines données.

Remerciements

Ce travail est soutenu par l'Institut des sciences de l'information et de leurs interactions du CNRS dans le cadre du PEPS Rupture « SpectroDec ».

Bibliographie

- Allen G., McMeeking R. (1978). Deconvolution of spectra by least-squares fitting. *Analytica Chimica Acta*, vol. 103, n° 1, p. 73-108.
- Caruana R., Searle R., Heller T., Shupack S. (1986). Fast algorithm for the resolution of spectra. *Analytical Chemistry*, vol. 58, p. 1162-1167.
- Chen S., Donoho D., Saunders M. (1995). *Atomic decomposition by basis pursuit*. Rapport technique. Stanford University.
- Davy M., Godsill S., Idier J. (2006, avril). Bayesian analysis of polyphonic western tonal music. *Journal of the Acoustical Society of America*, vol. 119, n° 4, p. 2498-2517.
- Devroye L. (1986). *Non-uniform random variate generation*. Springer-Verlag. <http://luc.devroye.org/rnbookindex.html>
- Donoho D., Tsai Y. (2006). *Fast solution of l_1 -norm minimization problems when the solution may be sparse*. Rapport technique. Stanford University.
- Eppink A., Parker D. (1997). Velocity map imaging of ions and electrons using electrostatic lenses: Application in photoelectron and photofragment ion imaging of molecular oxygen. *Review of Scientific Instruments*, vol. 68, n° 9, p. 3477-3484.
- Fischer R., Dose V. (2001). Analysis of mixtures in physical spectra. *Bayesian methods*, p. 145-154.
- Fraser R., Suzuki E. (1966). Resolution of overlapping absorption bands by least squares procedures. *Analytical Chemistry*, vol. 38, n° 12, p. 1770-1773.
- Garcia G., Nahon L., Powis I. (2004). Two-dimensional charged particle image inversion using a polar basis function expansion. *Review of Scientific Instruments*, vol. 75, n° 11, p. 4989-4996.
- Ge D., Idier J., Le Carpentier E. (2011). Enhanced sampling schemes for MCMC based blind Bernoulli-Gaussian deconvolution. *Signal Processing*, vol. 91, n° 4, p. 759-772.
- Geman D., Yang C. (1995). Nonlinear image recovery with half-quadratic regularization. *IEEE Transactions on Image Processing*, vol. 4, n° 7, p. 932-946.

- Geman S., Geman D. (1984). Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 6, p. 721-741.
- Gloaguen E., Mestdagh J.-M., Poisson L., Lepetit F., Visticot J.-P., *et al.* (2005). Experimental evidence for ultrafast electronic relaxation in molecules, mediated by diffuse states. *Journal of the American Chemical Society*, vol. 127, n° 47, p. 16529-16534.
- Gobinet C., Vrabie V., Manfait M., Piot O. (2009, mai). Preprocessing methods of Raman spectra for source extraction on biomedical samples: Application on paraffin-embedded skin biopsies. *IEEE Transactions on Biomedical Engineering*, vol. 56, n° 5, p. 1371-1382.
- Green P. (1995). Reversible jump Markov chain Monte Carlo computation and Bayesian model determination. *Biometrika*.
- Gulam Razul S., Fitzgerald W., Andrieu C. (2003). Bayesian model selection and parameter estimation of nuclear emission spectra using RJMCMC. *Nuclear Instruments and Methods in Physics Research A*, vol. 497, p. 492-510.
- Haan N., Godsill S. (2002). Bayesian models for DNA sequencing. In *Ieee international conference on acoustics, speech, and signal processing*, p. IV-4020-IV-4023.
- Idier J., Goussard Y. (1993, juillet). Markov modeling for Bayesian restoration of two-dimensional layered structures. *IEEE Transactions on Information Theory*, vol. 39, n° 4, p. 1356-1373.
- Katajamaa M., Orešič M. (2005). Processing methods for differential analysis of LC/MS profile data. *BMC Bioinformatics*, vol. 6, n° 1, p. 179-191. <http://www.biomedcentral.com/1471-2105/6/179>
- Keshava N., Mustard J. (2002, janvier). Spectral unmixing. *IEEE Signal Processing Magazine*, vol. 19, n° 1, p. 44-57.
- Kirkpatrick S., Gelatt Jr. C., Vecchi M. (1983). Optimization by simulated annealing. *Science*, vol. 220, p. 671-680.
- Mallat S., Zhang Z. (1993, décembre). Matching pursuits with time-frequency dictionaries. *IEEE Transactions on Signal Processing*, vol. 41, n° 12, p. 3397 -3415.
- Masson A., Poisson L., Gaveau M.-A., Soep B., Mestdagh J.-M., *et al.* (2010). Dynamics of highly excited barium atoms deposited on large argon clusters. I. General trends. *The Journal of Chemical Physics*, vol. 133, n° 5 (054307).
- Mazet V. (2005). *Développement de méthodes de traitement de signaux spectroscopiques : estimation de la ligne de base et du spectre de raies*. Thèse de doctorat non publiée, Université Henri Poincaré, Nancy 1.
- Mazet V. (2011). Joint Bayesian decomposition of a spectroscopic signal sequence. *IEEE Signal Processing Letters*, vol. 18, n° 3, p. 191-184.
- Nguyen Thi H.-L., Jutten C. (1995). Blind source separation for convolutive mixtures. *Signal Processing*, vol. 45, n° 2, p. 209-229.
- Perret B., Mazet V., Collet C., Slezak E. (2011). Hierarchical multispectral galaxy decomposition using a MCMC algorithm with multiple temperature simulated annealing. *Pattern Recognition*, vol. 44, n° 6, p. 1328-1342.

- Poisson L., Maksimenska R., Soep B., Mestdagh J.-M., Parker D., *et al.* (2010). Unusual quantum interference in the S1 state of DABCO and observation of intramolecular vibrational redistribution. *Journal of Physical Chemistry A*, vol. 114, p. 3313-3319.
- Poisson L., Mestdagh E. G. J.-M., Soep B., Gonzalez A., Chergui M. (2008). Direct observation of microscopic solvation at the surface of clusters by ultrafast photoelectron imaging. *Journal of Physical Chemistry A*, vol. 112, n° 39, p. 9200-9210.
- Robert C., Casella G. (2004). *Monte Carlo statistical methods* (2^e éd.). Springer.
- Roueff A., Chanussot J., Mars J., Nguyen M.-Q. (2004a). Partition d'une séquence d'images temps-échelle pour la séparation d'ondes dans un profil sismique. *Traitement du signal*, vol. 21, p. 17-36.
- Roueff A., Chanussot J., Mars J., Nguyen M.-Q. (2004b). Unsupervised separation of seismic waves using the watershed algorithm on time-scale images. *Geophysical Prospecting*, vol. 52, p. 287-300.
- Rue H., Held L. (2005). *Gaussian Markov random fields: Theory and applications*. Chapman & Hall/CRC.
- Sahnoun S., Djermoune E.-H., Soussen C., Brie D. (2011). Sparse multiresolution modal estimation. In *Ieee statistical signal processing workshop*. Nice, France.
- Stolow A., Bragg A., Neumark D. (2004). Femtosecond time-resolved photoelectron spectroscopy. *Chemical Reviews*, vol. 104, n° 4, p. 1719-1757.
- Tomasi G., van den Berg F., Andersson C. (2004). Correlation optimized warping and dynamic time warping as preprocessing methods for chromatographic data. *Journal of Chemometrics*, vol. 18, p. 231-241.
- Wang L., Liu J., Li S. (2000). MRF parameter estimation by MCMC method. *Pattern Recognition*, vol. 33, n° 11, p. 1919-1925.

Annexe A. Calcul des lois a posteriori conditionnelles de $\mathbf{a}_{s,k}$ et α_s

Dans cette annexe, nous montrons que la loi a posteriori de l'amplitude $\mathbf{a}_{s,k}$ ($\forall s, k$) est une loi normale à support positif. Le calcul est similaire pour les paramètres α_s du continuum.

L'équation (11) nous permet d'isoler la loi a posteriori de $\mathbf{a}_{s,k}$ conditionnelle :

$$p(\mathbf{a}_{s,k} | \dots) \propto \exp\left(-\frac{1}{2r_b} \|\mathbf{y}_s - \phi_s - \gamma_s\|^2 - \frac{1}{2r_a} \|D\mathbf{a}^l\|^2\right) \mathbb{1}_{\mathbb{R}^+}(\mathbf{a}_{s,k}) \quad (27)$$

avec $l = z_{s,k}$.

En notant $(\varphi_{s,k})_n = \exp\left(-\frac{(n - c_{s,k})^2}{2w_{s,k}^2}\right)$ la forme normalisée de la raie d'amplitude $a_{s,k}$ et

$$\mathbf{x}_{s,k} = \mathbf{y}_s - \phi_s - \gamma_s + a_{s,k}\varphi_{s,k} \quad (28)$$

$$\Leftrightarrow (\mathbf{x}_{s,k})_n = (\mathbf{y}_s)_n - \sum_{\substack{m=1 \\ m \neq k}}^K a_{s,m} \exp\left(-\frac{(n - c_{s,m})^2}{2w_{s,m}^2}\right) - (\gamma_s)_n, \quad (29)$$

il vient :

$$\|\mathbf{y}_s - \phi_s - \gamma_s\|^2 = \|\mathbf{x}_{s,k} - a_{s,k}\varphi_{s,k}\|^2 = \mathbf{x}_s^T \mathbf{x}_{s,k} - 2a_{s,k}\mathbf{x}_s^T \varphi_{s,k} + a_{s,k}^2 \varphi_{s,k}^T \varphi_{s,k}. \quad (30)$$

Par ailleurs,

$$\|D\mathbf{a}^l\|^2 = A_{s,k}a_{s,k}^2 + B_{s,k}a_{s,k} + C_{s,k} \quad (31)$$

où les valeurs de $A_{s,k}$, $B_{s,k}$ et $C_{s,k}$ dépendent de l'ordre de la dérivée. Lorsque D est une dérivée première ou seconde, les valeurs de $A_{s,k}$ et $B_{s,k}$ sont celles résumées dans les tableaux 2 et 3 (on verra dans la suite que la valeur de $C_{s,k}$ n'est pas utilisée).

Tableau 2. Coefficients $A_{s,k}$ et $B_{s,k}$ lorsque D est une dérivée première ($l = z_{s,k}$)

s	$A_{s,k}$	$B_{s,k}$
1	1	$-2a_2^l$
$2, \dots, S-1$	2	$-2(a_{s-1}^l + a_{s+1}^l)$
S	1	$-2a_{S-1}^l$

Tableau 3. Coefficients $A_{s,k}$ et $B_{s,k}$ lorsque D est une dérivée seconde ($l = z_{s,k}$)

s	$A_{s,k}$	$B_{s,k}$
1	1	$-4a_2^l + 2a_3^l$
2	5	$-4a_1^l - 8a_3^l + 2a_4^l$
$3, \dots, S-2$	6	$2a_{s-2}^l - 8a_{s-1}^l - 8a_{s+1}^l + 2a_{s+2}^l$
$S-1$	5	$-4a_S^l - 8a_{S-2}^l + 2a_{S-3}^l$
S	1	$-4a_{S-1}^l + 2a_{S-2}^l$

Ainsi :

$$\begin{aligned}
 & \frac{1}{2r_{\mathbf{b}}}\|\mathbf{y}_s - \phi_s - \gamma_s\|^2 + \frac{1}{2r_{\mathbf{a}}}\|D\mathbf{a}^l\|^2 \\
 &= \frac{1}{2r_{\mathbf{b}}}\left(\mathbf{x}_s^T \mathbf{x}_{s,k} - 2a_{s,k} \mathbf{x}_s^T \boldsymbol{\varphi}_{s,k} + a_{s,k}^2 \boldsymbol{\varphi}_{s,k}^T \boldsymbol{\varphi}_{s,k}\right) \\
 & \quad + \frac{1}{2r_{\mathbf{a}}}\left(A_{s,k} a_{s,k}^2 + B_{s,k} a_{s,k} + C_{s,k}\right) \\
 &= a_{s,k}^2 \left(\frac{\boldsymbol{\varphi}_{s,k}^T \boldsymbol{\varphi}_{s,k}}{2r_{\mathbf{b}}} + \frac{A_{s,k}}{2r_{\mathbf{a}}}\right) + a_{s,k} \left(-\frac{2\mathbf{x}_s^T \boldsymbol{\varphi}_{s,k}}{2r_{\mathbf{b}}} + \frac{B_{s,k}}{2r_{\mathbf{a}}}\right) \\
 & \quad + \left(\frac{\mathbf{x}_s^T \mathbf{x}_{s,k}}{2r_{\mathbf{b}}} + \frac{C_{s,k}}{2r_{\mathbf{a}}}\right), \tag{32}
 \end{aligned}$$

les termes entre parenthèses étant des scalaires. Or, comme

$$ax^2 + bx + c = a \left(x + \frac{b}{2a}\right)^2 + \left(c - \frac{b^2}{4a}\right) \tag{33}$$

l'équation (32) s'écrit :

$$\begin{aligned}
 & \frac{1}{2r_{\mathbf{b}}}\|\mathbf{y}_s - \phi_s - \gamma_s\|^2 + \frac{1}{2r_{\mathbf{a}}}\|D\mathbf{a}^l\|^2 \\
 &= \left(\frac{\boldsymbol{\varphi}_{s,k}^T \boldsymbol{\varphi}_{s,k}}{2r_{\mathbf{b}}} + \frac{A_{s,k}}{2r_{\mathbf{a}}}\right) \left(a_{s,k} + \frac{\left(-\frac{2\mathbf{x}_s^T \boldsymbol{\varphi}_{s,k}}{2r_{\mathbf{b}}} + \frac{B_{s,k}}{2r_{\mathbf{a}}}\right)}{2\left(\frac{\boldsymbol{\varphi}_{s,k}^T \boldsymbol{\varphi}_{s,k}}{2r_{\mathbf{b}}} + \frac{A_{s,k}}{2r_{\mathbf{a}}}\right)}\right)^2 \\
 & \quad + \left(\left(\frac{\mathbf{x}_s^T \mathbf{x}_{s,k}}{2r_{\mathbf{b}}} + \frac{C_{s,k}}{2r_{\mathbf{a}}}\right) - \frac{\left(-\frac{2\mathbf{x}_s^T \boldsymbol{\varphi}_{s,k}}{2r_{\mathbf{b}}} + \frac{B_{s,k}}{2r_{\mathbf{a}}}\right)^2}{4\left(\frac{\boldsymbol{\varphi}_{s,k}^T \boldsymbol{\varphi}_{s,k}}{2r_{\mathbf{b}}} + \frac{A_{s,k}}{2r_{\mathbf{a}}}\right)}\right) \tag{34}
 \end{aligned}$$

La loi a posteriori conditionnelle de $\mathbf{a}_{s,k}$ (équation (27)) devient (le second terme de l'équation (34) étant constant par rapport à $a_{s,k}$, il peut être ignoré) :

$$p(\mathbf{a}_{s,k} | \dots) \propto \exp\left(-\frac{1}{2\rho_{s,k}}(a_{s,k} - \mu_{s,k})^2\right) \mathbb{I}_{\mathbb{R}^+}(\mathbf{a}_{s,k}) \tag{35}$$

où

$$\frac{1}{2\rho_{s,k}} = \frac{\boldsymbol{\varphi}_{s,k}^T \boldsymbol{\varphi}_{s,k}}{2r_{\mathbf{b}}} + \frac{A_{s,k}}{2r_{\mathbf{a}}} \quad \Leftrightarrow \quad \rho_{s,k} = \frac{r_{\mathbf{a}} r_{\mathbf{b}}}{r_{\mathbf{a}} \boldsymbol{\varphi}_{s,k}^T \boldsymbol{\varphi}_{s,k} + r_{\mathbf{b}} A_{s,k}} \tag{36}$$

$$\mu_{s,k} = \frac{\frac{2\mathbf{x}_s^T \boldsymbol{\varphi}_{s,k}}{2r_{\mathbf{b}}} - \frac{B_{s,k}}{2r_{\mathbf{a}}}}{2\left(\frac{\boldsymbol{\varphi}_{s,k}^T \boldsymbol{\varphi}_{s,k}}{2r_{\mathbf{b}}} + \frac{A_{s,k}}{2r_{\mathbf{a}}}\right)} \quad \Leftrightarrow \quad \mu_{s,k} = \rho_{s,k} \left(\frac{\mathbf{x}_s^T \boldsymbol{\varphi}_{s,k}}{r_{\mathbf{b}}} - \frac{B_{s,k}}{2r_{\mathbf{a}}}\right) \tag{37}$$

C'est donc une loi normale de moyenne $\mu_{s,k}$ et de variance $\rho_{s,k}$ tronquée sur \mathbb{R}^+ .

Avec des calculs similaires, on vérifie que α_s suit également une loi normale de moyenne λ_s et de variance ν_s tronquée sur \mathbb{R}^+ :

$$p(\alpha_s | \dots) \propto \exp\left(-\frac{1}{2\nu_s}(\alpha_s - \lambda_s)^2\right) \mathbb{I}_{\mathbb{R}^+}(\alpha_s) \quad (38)$$

où

$$\nu_s = \frac{r_{\alpha} r_{\mathbf{b}}}{r_{\alpha} \psi_s^T \psi_s + r_{\mathbf{b}} E_s} \quad (39)$$

$$\lambda_s = \nu_s \left(\frac{(\mathbf{y}_s - \phi_s)^T \psi_s}{r_{\mathbf{b}}} - \frac{F_s}{2r_{\alpha}} \right) \quad (40)$$

avec $E_s = A_{s,k}$ ($A_{s,k}$ ne dépend en réalité pas de k), $(\psi_s)_n = \exp\left(-\frac{n}{\beta_s}\right)$ et les valeurs de F_s sont similaires aux valeurs de $B_{s,k}$ ($\forall k$) en remplaçant $a_{s,k}$ ($\forall k$) par α_s dans les expressions.

Vincent Mazet a obtenu sa thèse en 2005 à l'Université Henri Poincaré, Nancy 1 sous la direction de David Brie. Il est depuis septembre 2006 maître de conférences au Laboratoire ICube (anciennement LSIIIT) du CNRS et de l'Université de Strasbourg. Vincent Mazet s'intéresse désormais au traitement de signaux spectroscopiques et au traitement d'images hyperspectrales.

Sylvain Faisan est ingénieur de l'École nationale supérieure de physique de Strasbourg (2001) et a obtenu en décembre 2004 le grade de docteur de l'Université de Strasbourg I en traitement d'images. Depuis 2005, il est maître de conférences à Télécom Physique Strasbourg (Université de Strasbourg). Ses thèmes de recherche incluent la modélisation statistique des images et du signal.

Antoine Masson a obtenu sa thèse en 2012 au CEA Saclay dans le groupe Dynamique réactionnelle. Il est depuis en post-doc au LCPM de l'EPFL. Il s'intéresse désormais à la spectroscopie de molécule ionique d'intérêt biologique en phase gaz.

Marc-André Gaveau est diplômé de l'École Centrale de Paris, docteur en 1984 et ingénieur-chercheur au CEA depuis 1977. Il étudie la dynamique réactionnelle et la spectroscopie d'espèces isolées sur agrégats classiques de van der Waals de grande taille et a récemment développé la première source française d'agrégats d'hélium pour étendre ces travaux à des milieux quantiques isolés.

Lionel Poisson est diplômé de l'École Normale Supérieure (Paris), chercheur au CNRS depuis 2003 et habilité à diriger des recherches depuis 2008. Après un post-doctorat dans le groupe de D.M. Neumark à Berkeley, Lionel Poisson est devenu spécialiste de la dynamique ultrarapide de systèmes isolés. Il a obtenu le prix jeune chercheur de la Division de Chimie Physique (SCF-SFP) en 2011.

Jean-Michel Mestdagh est ancien élève de l'École Normale Supérieure (Paris) et chercheur au CNRS depuis 1976. Il a développé les premières activités de recherche en France sur la femtochimie en phase gazeuse et étudie désormais la dynamique réactionnelle de systèmes à nombreux degrés de liberté électroniques et géométriques. Il dirige le groupe de « Dynamique réactionnelle » du Laboratoire Francis Perrin.