

# Apprentissage supervisé robuste de caractéristiques de classes. Application en protéomique

Pascal SZACHERSKI<sup>1,2</sup>, Jean-François GIOVANNELLI<sup>1</sup>, Laurent GERFAULT<sup>2</sup>, Pierre GRANGEAT<sup>2</sup>

<sup>1</sup>IMS (Université Bordeaux 1 – IPB – CNRS), 351 rue de la Liberté, 33405 Talence Cedex, France

<sup>2</sup>CEA Leti, MINATEC Campus, DTBS, 17 rue des Martyrs, 38054 Grenoble Cedex, France

<sup>1</sup>prénom.nom@ims-bordeaux.fr, <sup>2</sup>prénom.nom@cea.fr

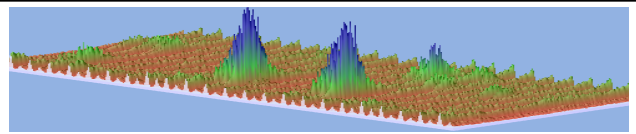
**Résumé** – La protéomique est un domaine en pleine expansion, et son utilisation est notamment envisagée dans le diagnostic de maladies comme le cancer. Un diagnostic basé sur une classification nécessite la connaissance des distributions des biomarqueurs dans chaque classe. Nous développons dans ce document l'apprentissage de ces distributions à partir de données indirectes. L'acquisition hiérarchique des données fait intervenir des variabilités biologique et technologique. Le cadre des problèmes inverses intègre les variabilités technologique (instrumentale) et biologique qui impactent les données et tire profit de la nature hiérarchique du modèle. La robustesse est démontrée grâce à la comparaison de deux méthodes: apprentissage avec inversion globale et apprentissage avec inversion réduite où les variables instrumentales sont connues. Pour cela, nous mesurons la proximité des deux distributions apprises avec la distribution vraie par la divergence de Kullback-Leibler.

**Abstract** – Proteomics is an expanding domain, and its application is highly considered in diagnosis of diseases like cancer. For a diagnosis based on classification, the knowledge of the biomarker distribution of each state is necessary. In this paper, we develop a learning method using indirect data. Within the hierarchical acquisition of the data, biological and technical variabilities occur. By placing our work into an inverse-problem-framework, we take into account the variabilities which are impacting the data, and use the hierarchical structure of the model. Furthermore, we underline the robustness by comparing this learning using global inversion with an ideal learning using reduced inversion where all instrumental variables are known in advance. This will be quantified by the distance of the learnt distributions with the true one, given by the Kullback-Leibler-divergence.

## 1 Introduction

Les protéines d'un organisme peuvent être sur- ou sous-exprimées suivant son état biologique. Le profil protéique, fourni à partir des expressions des protéines, permet de procéder à un diagnostic, une détection de maladie précoce, un suivi thérapeutique, de décrire l'efficacité d'un médicament, etc. Cependant, les biomarqueurs, i.e. les protéines différenciellement exprimées spécifiques à une maladie, sont souvent difficilement détectables du fait de leur faible concentration et du grand rapport sur la dynamique des concentrations entre les protéines prédominantes et les protéines cibles qui peut atteindre jusqu'à  $10^8$ . L'intérêt de l'utilisation d'un cadre bayésien à partir de mesures d'un couplage de chromatographie liquide (LC) et de spectrométrie de masse (MS) pour la quantification robuste a été montré [1]. Le modèle physique de l'instrument LC-MS a été réalisé dans [2] et nous sert de base de travail. Les étapes de la chaîne d'analyse introduisent la structure hiérarchique du modèle utilisé.

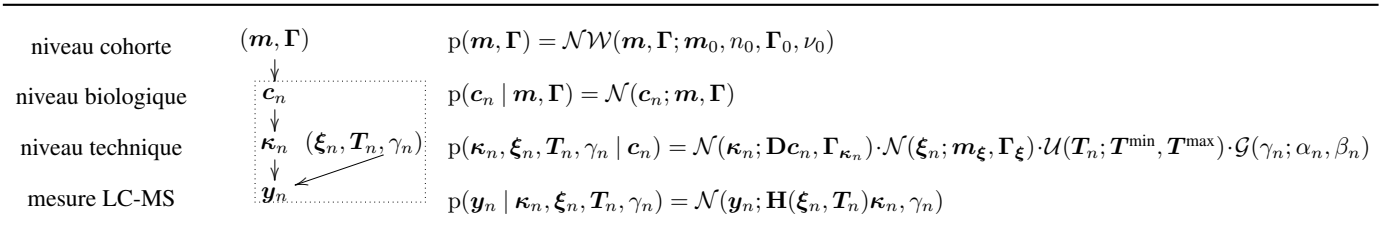
Les tâches de classification, telles que la détection ou le diagnostic [3], nécessitent de connaître les caractéristiques des états biologiques d'une manière robuste par rapport aux variabilités mentionnées. Elles doivent être apprises sur des mesures étiquetées provenant de cohortes identifiées. Un apprentissage



**Fig. 1** Exemple d'un spectre LC-MS [2].

supervisé sur les données brutes LC-MS en forme d'images sans étapes d'estimation est une méthode incertaine car elle n'intègre pas les variabilités technologique et biologique sous-jacentes. Un autre moyen est d'estimer les concentrations moléculaires des biomarqueurs, et d'apprendre les caractéristiques à partir de ces grandeurs. Cela intègre certes la variabilité biologique, et la variabilité technologique peut être traitée dans l'estimation des concentrations. Mais on ne peut pas associer dans une même estimation les deux variabilités. Elles sont traitées de manière indépendante.

Dans l'approche présentée ci-après, nous extrayons d'un jeu de données d'apprentissage les paramètres de la distribution des biomarqueurs tout en intégrant les variabilités technologique et biologique. Nous disposons de mesures indirectes de la grandeur d'intérêt, d'où le recours au cadre méthodologique des problèmes inverses. L'acquisition se fait dans une cascade de processus séquentiels : un modèle direct hiérarchique explique alors les données observées. Ces processus étant incer-



**Fig. 2** – Schéma du modèle hiérarchique direct et lois *a priori* conditionnelles associées à l’intérieur d’une classe donnée.

tains, nous intégrons les variabilités à l’aide d’une modélisation probabiliste. L’estimation globale dans un cadre bayésien est réputée plus robuste que l’estimation séparée « paramètre par paramètre » [4]. De plus, cela nous permet d’utiliser la nature hiérarchique du modèle [5].

## 2 Apprentissage des caractéristiques

L’apprentissage de caractéristiques de classes peut être vu comme la résolution d’un problème d’estimation paramétrique à partir d’un modèle direct hiérarchique. Nous supposons que la distribution d’une classe  $B$  est normale multivariée de moyenne  $\mathbf{m}_B \in \mathbb{R}^P$  et de précision  $\mathbf{\Gamma}_B \in \mathbb{R}^{P \times P}$ ,  $P \geq 1$  étant le nombre de biomarqueurs considérés et  $B$  l’étiquette de la classe considérée. Ainsi, nous pouvons en effet étendre le modèle hiérarchique que nous avons proposé dans [3] pour introduire un niveau cohorte. Ceci complète la hiérarchie et nous permet d’estimer pour chaque classe sa moyenne  $\mathbf{m}_B$  et sa matrice de précision  $\mathbf{\Gamma}_B$  à partir d’une cohorte de  $N$  mesures.

La mesure LC-MS  $\mathbf{y}_n$  est étiquetée par l’état biologique  $\mathcal{B}_n$  associé. Chaque mesure est acquise indépendamment des autres ce qui permet de regrouper les expériences par classe :

$$\mathcal{Y}_B = \{\mathbf{y}_n | \mathcal{B}_n = B, n = 1, \dots, N\}$$

où  $B$  représente l’étiquette de la classe donnée. Par exemple, il peut s’agir d’une cohorte contrôle d’échantillons biologiques « sain » et d’une cohorte cas d’échantillons biologiques « pathologiques ». Les classes étant connues, l’apprentissage global se sépare naturellement en autant de processus d’apprentissages qu’il y a de classes grâce à l’indépendance des données :

$$p(\mathbf{m}_{B_1}, \mathbf{\Gamma}_{B_1}, \dots, \mathbf{m}_{B_b}, \mathbf{\Gamma}_{B_b} | \mathcal{Y}_{B_1}, \dots, \mathcal{Y}_{B_b}) = p(\mathbf{m}_{B_1}, \mathbf{\Gamma}_{B_1} | \mathcal{Y}_{B_1}) \cdots p(\mathbf{m}_{B_b}, \mathbf{\Gamma}_{B_b} | \mathcal{Y}_{B_b}). \quad (1)$$

C’est la raison pour laquelle nous pouvons nous restreindre dans ce qui suit à un apprentissage des caractéristiques d’une classe.

### 2.1 Problème direct

Suivant la décomposition des molécules analysées, une chaîne d’analyse est décrite par un modèle hiérarchique schématisé sur la Fig. 2.

Les caractéristiques des classes  $\mathbf{m}$  et  $\mathbf{\Gamma}$  déterminent aléatoirement la concentration des protéines  $\mathbf{c}_n$  du  $n^{\text{ème}}$  individu,  $n = 1, \dots, N$ . Les protéines sont découpées par un enzyme

en peptides de concentration  $\mathbf{\kappa}_n = \mathbf{D}\mathbf{c}_n + \boldsymbol{\varepsilon}_{\mathbf{\kappa}}$ , perturbé par un bruit de digestion. Le gain de digestion est décrit par la matrice  $\mathbf{D}$ . Les peptides sont ensuite séparés par chromatographie, caractérisé par leurs temps de rétention  $\mathbf{T}_n$ . Dans la suite, ils sont ionisés par *electrospray* en vue de l’introduction dans le spectromètre de masse. Lors de l’ionisation, les peptides subissent un gain noté  $\boldsymbol{\xi}_n$ . Puis, sont acquises les données  $\mathbf{y}_n = \mathbf{H}(\boldsymbol{\xi}_n, \mathbf{T}_n)\mathbf{\kappa}_n + \mathbf{b}_n$  où  $\mathbf{H}$  est linéaire en  $\boldsymbol{\xi}_n$ , perturbées par un bruit blanc gaussien centré  $\mathbf{b}_n$  d’inverse-variance  $\gamma_n$ . La matrice de système  $\mathbf{H}$  de toutes les expériences traduit la relation linéaire entre les peptides et la sortie.

Suite à la relation bilinéaire de  $\boldsymbol{\xi}_n$  et  $\mathbf{\kappa}_n$  par rapport la sortie  $\mathbf{y}_n$ , il existe une indétermination entre les deux paramètres. Nous utilisons un étalon interne, PSAQ [6], qui nous permet de monitorer le gain et donc de lever l’indétermination.

### 2.2 Loi jointe

Du fait des incertitudes et variabilités lors de la préparation et à l’intérieur de l’instrument, les valeurs des paramètres à chaque étape ne sont connues qu’à travers des probabilités. Elles sont explicitées dans la Fig. 2 et ont été choisies pour décrire au mieux les processus tout en restant simples à utiliser. Le temps de rétention prenant sa valeur dans  $[\mathbf{T}^{\min}, \mathbf{T}^{\max}]$  de manière équiprobable, sa loi *a priori* est donnée par une loi uniforme  $\mathcal{U}(\mathbf{T}^{\min}, \mathbf{T}^{\max})$ . La loi pour le couple  $(\mathbf{m}, \mathbf{\Gamma})$  est la loi Normale-Wishart qui est le produit d’une loi normale de moyenne  $\mathbf{m}_0$  et de précision  $n_0\mathbf{\Gamma}$  et d’une loi Wishart de matrice d’échelle  $\mathbf{\Gamma}_0$  avec  $\nu_0$  degrés de liberté :

$$p(\mathbf{m}, \mathbf{\Gamma}) = \frac{|\mathbf{\Gamma}_0|^{-\nu_0/2}}{2^{\nu_0 P/2} (2\pi)^P \cdot g_P(\nu_0/2)} |\mathbf{\Gamma}|^{(\nu_0 - P + 1)/2} \exp\left(-\frac{1}{2} [\text{tr}(\mathbf{\Gamma}_0^{-1}\mathbf{\Gamma}) + n_0(\mathbf{m}_0 - \mathbf{m})^T \mathbf{\Gamma}(\mathbf{m}_0 - \mathbf{m})]\right) \quad (2)$$

où  $g_P(\cdot)$  est la fonction Gamma  $P$ -dimensionnelle et  $\text{tr}(\cdot)$  la fonction trace de matrice.

La loi jointe s’exprime alors, compte tenu de l’indépendance des expériences et des indépendances conditionnelles liées à la hiérarchie [7, sect. 10.2], par

$$p(\mathbf{m}, \mathbf{\Gamma}, \mathbf{c}_{1:N}, \mathbf{\kappa}_{1:N}, \gamma_{1:N}, \boldsymbol{\xi}_{1:N}, \mathbf{T}_{1:N}, \mathbf{y}_{1:N}) = p(\mathbf{m}, \mathbf{\Gamma}) \cdot \prod_{n=1}^N \left[ p(\mathbf{y}_n | \mathbf{\kappa}_n, \gamma_n) \cdot p(\gamma_n) p(\boldsymbol{\xi}_n) p(\mathbf{T}_n) p(\mathbf{\kappa}_n | \mathbf{c}_n) p(\mathbf{c}_n | \mathbf{m}, \mathbf{\Gamma}) \right] \quad (3)$$

où l’indice  $1 : N$  regroupe tous les indices  $n = 1, \dots, N$ .

**Alg. 1** – Résumé de l’algorithme d’estimation des caractéristiques.

<p>1. Initialisation de l’algorithme.</p> <p>2. Boucle de Gibbs :</p> <p style="padding-left: 20px;">pour <math>k = 1</math> à <math>K_0 + K</math></p> <p style="padding-left: 40px;">(a) pour <math>n = 1</math> à <math>N</math>, échantillonner :</p> <p style="padding-left: 80px;">i. <math>\gamma_n^{(k+1)} \sim p(\gamma_n   \mathbf{y}_n, \boldsymbol{\kappa}_n^{(k)})</math></p> <p style="padding-left: 80px;">ii. <math>\boldsymbol{\xi}_n^{(k+1)} \sim p(\boldsymbol{\xi}_n   \mathbf{y}_n, \boldsymbol{\kappa}_n^{(k)}, \mathbf{T}_n^{(k)}, \gamma_n^{(k+1)})</math></p>	<p style="padding-left: 40px;">iii. <math>\mathbf{T}_n^{(k+1)} \sim p(\mathbf{T}_n   \mathbf{y}_n, \boldsymbol{\kappa}_n^{(k)}, \boldsymbol{\xi}_n^{(k+1)}, \gamma_n^{(k+1)})</math></p> <p style="padding-left: 40px;">iv. <math>\boldsymbol{\kappa}_n^{(k+1)} \sim p(\boldsymbol{\kappa}_n   \mathbf{y}_n, \gamma_n^{k+1}, \mathbf{T}_n^{(k+1)}, \boldsymbol{\xi}_n^{(k+1)}, \mathbf{c}_n^{(k)})</math></p> <p style="padding-left: 40px;">v. <math>\mathbf{c}_n^{(k+1)} \sim p(\mathbf{c}_n   \boldsymbol{\kappa}_n^{(k+1)}, \mathbf{m}^{(k)}, \boldsymbol{\Gamma}^{(k)})</math></p> <p style="padding-left: 40px;">(b) échantillonner <math>(\mathbf{m}^{(k+1)}, \boldsymbol{\Gamma}^{(k+1)}) \sim p(\mathbf{m}, \boldsymbol{\Gamma}   \mathbf{c}_{1:N}^{(k+1)})</math></p> <p>3. Approximation du EAP :</p> $[\hat{\mathbf{m}}, \hat{\boldsymbol{\Gamma}}] = \frac{1}{K} \sum_{k=K_0+1}^{K_0+K} [\mathbf{m}^{(k)}, \boldsymbol{\Gamma}^{(k)}]$
--	---

### 2.3 Estimateur des caractéristiques

L’objectif est d’estimer les caractéristiques de l’état biologique. Un des critères d’évaluation des estimations les plus utilisés est le coût quadratique qui évalue la norme au carré de la différence entre la vraie valeur et son estimation. L’estimateur bayésien associé à ce coût est l’estimateur moyenne *a posteriori* (EAP) [7, sect. 2.5]. Dans le cadre présent, il s’écrit

$$[\bar{\mathbf{m}}, \bar{\boldsymbol{\Gamma}}] = \int [\mathbf{m}, \boldsymbol{\Gamma}] p(\mathbf{m}, \boldsymbol{\Gamma} | \mathbf{y}_{1:N}) d(\mathbf{m}, \boldsymbol{\Gamma}). \quad (4)$$

Ce calcul nécessite de marginaliser les paramètres intermédiaires  $\boldsymbol{\theta}_{1:N} = [\mathbf{c}_{1:N}, \boldsymbol{\kappa}_{1:N}, \boldsymbol{\xi}_{1:N}, \mathbf{T}_{1:N}, \gamma_{1:N}]$  du processus :

$$[\bar{\mathbf{m}}, \bar{\boldsymbol{\Gamma}}] = \int [\mathbf{m}, \boldsymbol{\Gamma}] p(\mathbf{m}, \boldsymbol{\Gamma}, \boldsymbol{\theta}_{1:N} | \mathbf{y}_{1:N}) d(\mathbf{m}, \boldsymbol{\Gamma}) d\boldsymbol{\theta}_{1:N}. \quad (5)$$

Le calcul de cette intégrale est analytiquement impossible. C’est la raison pour laquelle nous allons approcher la loi *a posteriori* jointe  $p(\mathbf{m}, \boldsymbol{\Gamma}, \boldsymbol{\theta}_{1:N} | \mathbf{y}_{1:N})$  en utilisant une méthode MCMC [8], [5, ch. 11]. Ceci permettra de marginaliser les paramètres dans l’équation (5) et ensuite d’approcher l’estimateur de (4).

## 3 Mise en œuvre algorithmique

À l’intérieur de l’approche MCMC, nous adoptons une structure de Gibbs qui permet de transformer un problème global en une succession de sous-problèmes plus simples [5, sect. 11.4], [8, ch. 10]. Ainsi, nous ne faisons pas de tirage global « direct » de  $[\mathbf{m}, \boldsymbol{\Gamma}, \boldsymbol{\theta}_{1:N}]$  sous la loi *a posteriori* jointe ; nous tirons plutôt d’abord l’échantillon pour le couple  $(\mathbf{m}, \boldsymbol{\Gamma})$ , puis l’échantillon pour  $\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \dots, \boldsymbol{\theta}_N$ . Comme  $\boldsymbol{\theta}_n = [\mathbf{c}_n, \boldsymbol{\kappa}_n, \boldsymbol{\xi}_n, \mathbf{T}_n, \gamma_n]$  est un vecteur, la structure de Gibbs s’applique également à celui-ci pour  $n = 1, \dots, N$ .

L’échantillonneur de Gibbs nécessite les lois *a posteriori* conditionnelles de chaque paramètre que nous déduisons de la loi jointe définie dans Éq. (3). Pour les concentrations (protéines et peptides), le gain et la précision du bruit, les lois *a posteriori* conditionnelles sont de la même famille que leurs lois *a priori* grâce à leur conjugaison par la fonction de vraisemblance associée [5, sect. 2.4]. Ceci permet un échantillonnage direct de ces paramètres sous leurs lois *a posteriori* conditionnelles respectives puisqu’elles sont classiques. Autrement dit, l’algorithme ne met pas à jour la loi, mais uniquement les hyperparamètres qui caractérisent la loi. Quant au temps de rétention où il n’y a pas de conjugaison des lois, une étape de

*Metropolis-Hasting* (MH) est intégrée dans la boucle de Gibbs [5, sect. 11.5]. Le principe est le suivant : à l’itération courante, la valeur est  $\mathbf{T}_n^c$ . L’algorithme propose *aléatoirement* une nouvelle valeur  $\mathbf{T}_n^p$  sous une loi instrumentale  $\psi(\mathbf{T}_n^p, \mathbf{T}_n^c)$ . La proposition est acceptée avec la probabilité

$$\delta = \min \left\{ 1; \frac{\varphi(\mathbf{T}_n^p) \psi(\mathbf{T}_n^c, \mathbf{T}_n^p)}{\varphi(\mathbf{T}_n^c) \psi(\mathbf{T}_n^p, \mathbf{T}_n^c)} \right\} \quad (6)$$

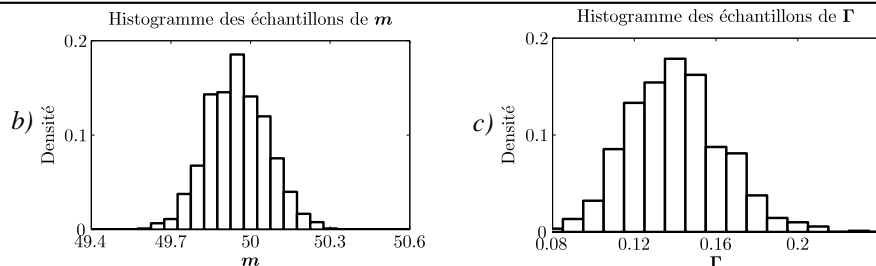
qui dépend bien évidemment des valeurs courante et proposée ainsi que de la loi cible  $\varphi$  (loi *a posteriori* conditionnelle pour  $\mathbf{T}_n$ ) et de la loi instrumentale. Il existe de nombreuses variantes en fonction du choix de la loi instrumentale [5, 9, 7]. Un bon compromis entre rapidité des propositions et rapidité de la convergence est un MH à Marche Aléatoire. Dans ce cas, la loi instrumentale  $\psi(\mathbf{T}_n^p, \mathbf{T}_n^c)$  est un noyau gaussien de moyenne  $\mathbf{T}_n^c$  et de précision  $\boldsymbol{\Gamma}_{RW}$ . Le paramètre d’algorithme  $\boldsymbol{\Gamma}_{RW}$  doit être choisi de manière à proposer des valeurs suffisamment éloignées de la valeur courante pour explorer, mais en même temps à ne pas rejeter trop souvent les propositions. En effet, plus le voisinage est petit, plus la probabilité d’acceptation est grande ; plus le voisinage est grand, plus l’algorithme explore l’espace des paramètres, mais plus il les rejette. La symétrie de  $\psi$  et le choix de la loi *a priori* uniforme pour  $\mathbf{T}_n$  fait que la probabilité d’acceptation s’exprime simplement comme rapport de vraisemblances.

On approche ainsi un tirage sous la loi *a posteriori* jointe. La marginalisation des paramètres intermédiaires est obtenue en ne gardant que les tirages des paramètres d’intérêt. Enfin, le calcul du EAP du couple  $(\mathbf{m}, \boldsymbol{\Gamma})$  revient à moyenniser les tirages marginalisés. Les premiers tirages ne se font pas sous la loi *a posteriori* jointe [5, sect. 11.6]. Pour ne pas fausser le calcul de l’estimateur, nous ne prenons en compte les tirages qu’après un temps de chauffe de  $K_0$  itérations. L’algorithme est résumé dans Alg. 1.

## 4 Résultats

Nous présentons les résultats obtenus sur une protéine composée d’un seul peptide. La concentration vraie utilisée dans la cohorte d’apprentissage a été modélisée par une loi gaussienne  $p^*(c) = \mathcal{N}(c; m^*, \gamma^*)$  de moyenne  $m^* = 50$  (unité arbitraire) de précision  $\gamma^* = 0.2$ . Les autres paramètres ont une variabilité de 2% pour le temps de rétention et de 10% pour le gain du système. L’inverse-variance du bruit de mesure de simulation

$N$	$d(p^*, \tilde{p})$	$d(p^*, \hat{p})$
10	0.1149	0.1151
20	0.0570	0.0566
50	0.0387	0.0392
100	0.0329	0.0334
200	0.0303	0.0309



**Fig. 3** – a) Résultats moyens sur 250 expériences du test de performance de l’apprentissage des distributions avec les variables technologiques connues ( $\tilde{p}$ ) et estimées  $\hat{p}$ .  $d(p^*, \cdot)$  désigne la divergence de Kullback-Leibler avec la distribution vraie  $p^*$ . b) Histogramme des tirages aléatoires pour  $m \equiv m \in \mathbb{R}$ . c) Histogramme des tirages aléatoires pour  $\Gamma \equiv \gamma \in \mathbb{R}$ .

est choisi uniformément entre 0.01 et 1 ; i.e. la puissance du bruit varie entre 1 et 100.

La divergence de Kullback-Leibler entre les distributions vraie et estimée est utilisée pour évaluer les performances de l’approche présentée [5, Annexe B]. Ceci est fait en comparant deux stratégies. Premièrement, lors de l’inversion, les paramètres instruments sont fixés aux valeurs vraies pour connaître les performances de l’approche idéale de la méthode. Deuxièmement, tous les paramètres instruments sont estimés lors de l’inversion. Les distributions estimées sont notées respectivement  $\tilde{p}$  et  $\hat{p}$ . Les résultats de ces évaluations pour différentes tailles de cohorte sont présentés dans la table 3a) qui donne une moyenne sur 250 réalisations. On constate deux choses :

- i) Plus la taille  $N$  de la cohorte est grande, plus la distribution apprise est proche de la distribution initiale et plus la divergence de Kullback-Leibler décroît.
- ii) Les performances de l’inversion complète, grâce à l’apprentissage robuste, sont proches du cas idéal pour lequel uniquement les paramètres biologiques sont incertains.

Les figures 3b) et c) présentent les histogrammes marginaux des tirages pour les caractéristiques des classes pour 1000 itérations.

## 5 Conclusion

Ce document présente une méthode d’apprentissage à partir de données LC-MS issues de la protéomique. Notre méthode d’apprentissage est robuste par rapport aux variabilités qui surviennent lors de l’acquisition des données, grâce à l’utilisation du contexte des problèmes inverses. Nous proposons l’utilisation d’une approche bayésienne dont le problème direct est résolu par l’échantillonnage stochastique via un algorithme MCMC. Nous comparons la stratégie d’inversion globale, i.e. d’autres paramètres intervenants, à celle d’inversion réduite, i.e. les paramètres instruments sont connus. Nous montrons la robustesse de la méthode présentée par quasi-égalité des distances des distributions à la distribution vraie.

Les distributions issues de l’apprentissage interviennent lors de l’approche d’inversion-classification bayésienne que nous avons proposée pour analyser les données LC-MS en associant quantification et analyse différentielle [3]. La méthode est trans-

posable à l’apprentissage des paramètres caractéristiques des lois de tous les paramètres intervenant.

Dans nos futurs travaux, nous continuons à compléter la hiérarchie en nous intéressant à d’autres phénomènes comme celui de la digestion où se pose la question de savoir si le découpage d’une protéine par la trypsine a été réalisé ou non. Nous avons ainsi déjà introduit la notion de rendement de digestion [10].

## Références

- [1] P. Grangeat et al. First demonstration on NSE biomarker of a computational environment dedicated to lab-on-chip based cancer diagnosis. Poster at 58th ASMS Conference, Salt Lake City, USA, 2010.
- [2] G. Strubel. *Reconstruction de profils moléculaires : modélisation et inversion d’une chaîne de mesure protéomique*. PhD thesis, École Polytechnique de Grenoble, France, 2008.
- [3] P. Szacherski, J.-F. Giovannelli, and P. Grangeat. Joint bayesian hierarchical inversion-classification and application in proteomics. In *IEEE Workshop Statistical Signal Processing*, Nice, France, June 2011.
- [4] L. Gerfault, G. Strubel, C. Paulus, J.-F. Giovannelli, and P. Grangeat. Évaluation statistique d’un algorithme bayésien pour la reconstruction de profils moléculaires par spectrométrie de masse. In *XXII<sup>ème</sup> Colloque GRETSI*, Dijon, France, 2009.
- [5] A. Gelman, J. B. Carlin, H. S. Stern, and D. B. Rubin. *Bayesian Data Analysis, Second Edition (Texts in Statistical Science)*. Chapman & Hall/CRC, 2 edition, July 2003.
- [6] V. Brun, C. Masselon, J. Garin, and A. Dupuis. Isotope dilution strategies for absolute quantitative proteomics. *Journal of Proteomics*, 72(5) :740–749, 2009.
- [7] C. Robert. *The Bayesian choice : from decision-theoretic foundations to computational implementation*. Springer, New York, USA, 2007.
- [8] C. Robert and G. Casella. *Monte Carlo statistical methods*. Springer, New York, 2nd ed. edition, 2004.
- [9] C. Vacar, J.-F. Giovannelli, and Y. Berthoumieu. Langevin and Hessian with Fisher approximation : Stochastic sampling for parameter estimation of structured covariance. In *ICASSP*, Prague, Czech Republic, May 2011.
- [10] P. Grangeat, P. Szacherski, L. Gerfault, and J.-F. Giovannelli. Bayesian hierarchical reconstruction of protein profiles including a digestion model. Presentation at 59th ASMS Conference, Denver, USA, June 2011.