Modélisation stochastique de processus pharmaco-cinétiques, application à la reconstruction tomographique par émission de positrons (TEP) spatio-temporelle

Mame Diarra Fall

UNIVERSITÉ PARIS SUD

Faculté des sciences d'Orsay

N° d'ordre: 10566

THÈSE

Présentée pour obtenir

LE GRADE DE DOCTEUR EN SCIENCES
DE L'UNIVERSITÉ PARIS-SUD

Spécialité: Traitement du signal et des images

École doctorale « Sciences et Technologies de l'Information, des Télécommunications et des Systèmes »

par

Mame Diarra FALL

Modélisation stochastique de processus pharmaco-cinétiques, application à la reconstruction tomographique par émission de positrons (TEP) spatio-temporelle

Soutenue le 09 Mars 2012 devant la Commission d'examen:

M.	Éric Barat	Encadrant de thèse
M.	Claude Comtat	Co-encadrant de thèse
M.	Michel Defrise	Rapporteur
M.	Jean-François Giovannelli	Président du jury
M.	Ali Mohammad-Djafari	Directeur de thèse
M.	Anthony Quinn	Examinateur
Mme.	Judith Rousseau	Rapporteur

Thèse préparée au
Laboratoire de Modélisation, Simulation et Systèmes
DRT/LIST/DCSI/LM2S, CEA Saclay,
91191 Gif-sur-Yvette, France.

Résumé

L'objectif de ce travail est de développer de nouvelles méthodes statistiques de reconstruction d'image spatiale (3D) et spatio-temporelle (3D+t) en Tomographie par Emission de Positons (TEP). Le but est de proposer des méthodes efficaces, capables de reconstruire des images dans un contexte de faibles doses injectées tout en préservant la qualité de l'interprétation. Ainsi, nous avons abordé la reconstruction sous la forme d'un problème inverse spatial et spatio-temporel (à observations ponctuelles) dans un cadre *bayésien non paramétrique*. La modélisation bayésienne fournit un cadre pour la régularisation du problème inverse mal posé au travers de l'introduction d'une information dite *a priori*. De plus, elle caractérise les grandeurs à estimer par leur distribution *a posteriori*, ce qui rend accessible la distribution de l'incertitude associée à la reconstruction. L'approche non paramétrique quant à elle pourvoit la modélisation d'une grande robustesse et d'une grande flexibilité. Notre méthodologie consiste à considérer l'image comme une densité de probabilité dans $\mathbb{R}^k$ (pour une reconstruction en k dimensions) et à chercher la solution parmi l'ensemble des densités de probabilité de $\mathbb{R}^k$.

La grande dimensionalité des données à manipuler conduit à des estimateurs n'ayant pas de forme explicite. Cela implique l'utilisation de techniques d'approximation pour l'inférence. La plupart de ces techniques sont basées sur les méthodes de Monte-Carlo par chaînes de Markov (MCMC). Dans l'approche bayésienne non paramétrique, nous sommes confrontés à la difficulté majeure de générer aléatoirement des objets de dimension infinie sur un calculateur. Nous avons donc développé une nouvelle méthode d'échantillonnage qui allie à la fois bonnes capacités de mélange et possibilité d'être parallélisé afin de traiter de gros volumes de données.

L'approche adoptée nous a permis d'obtenir des reconstructions spatiales 3D sans nécessiter de voxellisation de l'espace, et des reconstructions spatio-temporelles 4D sans discrétisation en amont ni dans l'espace ni dans le temps. De plus, on peut quantifier l'erreur associée à l'estimation statistique au travers des intervalles de crédibilité.

Mots-clés : mesures aléatoires, problèmes inverses, inférence bayésienne non paramétrique, MCMC, imagerie tomographique par émission de positrons.

Stochastic modeling of pharmaco-kinetic processes, applied to PET space-time reconstruction

Abstract

The aim of this work is to develop new statistical methods for spatial (3D) and space-time (3D+t) Positron Emission Tomography (PET) reconstruction. The objective is to propose efficient reconstruction methods in a context of low injected doses while maintaining the quality of the interpretation. We tackle the reconstruction problem as a spatial or a space-time inverse problem for point observations in a *Bayesian nonparametric* framework. The Bayesian modeling allows to regularize the ill-posed inverse problem via the introduction of a prior information. Furthermore, by characterizing the unknowns with their posterior distributions, the Bayesian context allows to handle the uncertainty associated to the reconstruction process. Being nonparametric offers a framework for robustness and flexibility to perform the modeling. In the proposed methodology, we view the image to reconstruct as a probability density in $\mathbb{R}^k$ (for reconstruction in k dimensions) and seek the solution in the space of whole probability densities in $\mathbb{R}^k$.

However, due to the size of the data, posterior estimators are intractable and approximation techniques are needed for posterior inference. Most of these techniques are based on Markov Chain Monte-Carlo methods (MCMC). In the Bayesian nonparametric approach, a major difficulty raises in randomly sampling infinite dimensional objects in a computer. We have developed a new sampling method which combines both good mixing properties and the possibility to be implemented on a parallel computer in order to deal with large data sets.

Thanks to the taken approach, we obtain 3D spatial reconstructions without any ad hoc space voxellization and 4D space-time reconstructions without any discretization, neither in space nor in time. Furthermore, one can quantify the error associated to the statistical estimation using the credibility intervals.

Keywords : random measures, inverse problems, Bayesian Nonparametric inference, MCMC, Positron Emission Tomography imaging.

Remerciements

La page la plus difficile à écrire car sans doute la plus personnelle. Celle où il faut remercier tous ceux qui ont contribué de près ou de loin au succès de cette thèse et ils sont nombreux. Parce qu'effectuer ma thèse en collaboration avec trois laboratoires m'a beaucoup apporté sur les plans scientifique et humain.

Mes premiers remerciements vont tout naturellement à celui qui a beaucoup contribué à la formation de la jeune chercheuse que je suis devenue à savoir mon encadrant Éric Barat. Merci Éric pour ta patience et ta grande disponibilité, pour ta passion de la recherche et ton enthousiasme sur le bayésien non paramétrique que tu m'as transmis. Merci aussi de m'avoir soutenue dans les moments difficiles et rassurée pendant les périodes de doute. Ces quelques lignes ne suffisent pas pour t'exprimer toute ma reconnaissance mais je voudrais juste te dire que tu as été un encadrant exemplaire, dont l'étendue des connaissances scientifiques, les qualités humaines et l'éthique m'ont beaucoup inspirée, et pour qui j'éprouve un immense respect et une grande estime.

Je remercie Claude Comtat qui a co-encadré cette thèse et à qui je dois ma compréhension de la physique de la TEP. Merci Claude pour ta rigueur, ta patience et tes explications détaillées. Je remercie aussi Ali Mohammad-Djafari d'avoir accepté d'être directeur de cette thèse. Merci pour tes conseils, tes questions et tes remarques sur mon travail. Je redis un grand merci collectif à vous trois (Éric, Claude et Ali) pour la grande confiance que vous m'avez accordée en me sélectionnant pour effectuer ce travail de recherche sur un sujet difficile, que je connaissais très peu au début. Je remercie Digiteo (Île-de-France) et Technologies pour la Santé (CEA) qui ont financé ces travaux.

J'exprime toute ma gratitude à Judith Rousseau et à Michel Defrise qui ont accepté la lourde charge d'être les rapporteurs de ma thèse. Je suis très honorée de l'intérêt et de l'enthousiasme que vous avez manifestés à ce travail. Je vous suis reconnaissante de la lecture attentive du manuscrit, de vos remarques et suggestions pertinentes.

Je suis heureuse qu'Anthony Quinn ait pu se libérer, malgré un emploi du temps chargé, pour faire partie de mon jury de thèse. Merci pour l'intérêt que tu as toujours porté à ce travail, pour la pertinence de tes questions et tes conseils bienveillants. Un grand merci à Jean-François Giovannelli qui m'a fait l'honneur d'être président du jury.

Cette thèse est avant tout un travail d'équipe et en plus d'Éric et Claude, mes autres "co-auteurs" y sont pour beaucoup. Merci à Thomas Dautremer, un programmeur hors du commun, toujours là pour répondre à mes nombreuses questions même en géométrie 3D ! Merci à Thierry Montagu, pour ton aide lors de mes débuts en Python et C++ mais aussi pour les moments de rigolade du matin. Je te souhaite que ça aille mieux les vendredis ! J'ai vraiment aimé travailler avec vous tous, on aura fait une belle équipe, sacrément efficace !

Merci à Boris, Antoine et Bertrand qui complétaient notre "équipe du déjeuner". J'ai beaucoup apprécié les moments passés au CEA de Saclay, les déjeuners où sciences et actualités se mêlaient avec tant d'humour. Ça me manque tellement ! Merci à Étienne pour ses encouragements pendant les moments restés tardivement à Saclay, à Han Wang, lui aussi compagnon des heures tardives et soutien mutuel lors de la rédaction de nos thèses respectives. J'en profite pour exprimer ma gratitude aux autres membres des laboratoires LM2S et LOAD du CEA, ainsi qu'à Olivier Gal. Merci à tous ceux avec qui j'ai partagé des moments de discussions scientifiques ou autres : Quentin, Marius, Michele, Caifang, Juan, Hamid etc. Merci à ma chère amie Tabea que j'y ai connue.

J'ai pu profiter d'une période d'immersion au SHFJ. Assister aux examens cliniques en TEP, avoir pu observer de près le déroulement du protocole, côtoyer médecins et biologistes a été une grande chance. Merci à tous ceux que j'y ai croisés. Je remercie par la même occasion Simon Stute pour sa connaissance de la TEP mais aussi pour sa jovialité.

Après le CEA et le SHFJ, cap à Supélec plus précisément au L2S où j'ai pu vivre là aussi d'agréables moments humains et scientifiques. Je remercie Thomas Rodet qui porte toute l'équipe du GPI de par sa disponibilité, ses conseils et son soutien. Merci également à Nicolas Gac et Aurélia pour les moments de papotage. Je tiens à remercier les membres de l'administration qui nous ont facilité la tâche au quotidien et lors de nos missions : Maryvonne, Helena, Myriam, Franck et Daniel. Sans oublier les thésards passés et présents pour les différents échanges. Un merci tout particulier à mes potes : François O., Nabil, Ziad, Sofiane, Boujemaa. Merci à Veronica pour ta présence féminine dans cette bande et pour tout ce que tu m'as apporté. Enfin, j'adresse la place charnière de ces remerciements à Hacheme, un ami comme on n'en rencontre pas souvent.

Je profite de cette thèse pour rendre un hommage à tous mes enseignants, de Dakar à Paris en passant par Amiens. Tous ceux qui exercent ce métier noble et magnifique et qui m'ont donné envie d'enseigner à mon tour. Merci à tous ceux qui ont cru en moi et m'ont encouragée.

J'ai la chance d'avoir des amis formidables. Merci pour votre présence physique ou morale dans les moments de doute (souvent) et de joie (parfois) qui jalonnent la vie de tout doctorant.

Je n'aurai jamais de mots assez forts pour exprimer ma gratitude à l'égard de ma famille, mon roc de stabilité. Merci à mes parents, mes frères et ma sœur, à qui je dois tant. Merci pour les valeurs que vous m'avez inculquées et qui me permettent d'aller au bout de ce que j'entreprends. Merci d'avoir su me porter pour que je réalise ce rêve. Merci à mes neveux et nièces qui rendent tout si léger ! Je vous dédie à tous cette thèse. Je la dédie aussi à la mémoire de Mame Coumba, mon amie d'enfance devenue ma belle-sœur, partie hélas beaucoup trop tôt en Février 2010. Tu te moquais gentiment de ce que tu appelais mon "goût immodéré pour les études". J'espère que de là-haut tu seras fière. Merci pour la magnifique filleule que tu m'as laissée.

Je m'excuse auprès de ceux que j'ai oublié de mentionner. Mais il faut savoir affronter les limites de sa pensée. Un dernier merci à tous ceux qui se sont déplacés pour assister à ma soutenance. Cela m'a vraiment fait chaud au cœur de voir autant de monde.

J'ai vécu une belle aventure, je remercie infiniment tous ceux qui y ont participé. Une page se ferme, en espérant en ouvrir une autre qui soit au moins tout aussi belle...

Table des matières

Page de garde **1**

Résumé **3**

Table des matières **13**

Notations **15**

Abbréviations **17**

1 Introduction générale **19**

1.1 Contexte 19

1.2 Vue d'ensemble du travail effectué 20

1.2.1 Reconstruction spatiale 3D 20

1.2.2 Reconstruction spatio-temporelle 4D 22

1.2.3 Nouvelle méthode d'échantillonnage 24

1.3 Organisation du manuscrit 24

2 Modélisation bayésienne non paramétrique **27**

2.1 Le processus de Dirichlet et ses représentations 28

2.1.1 Représentation par processus gamma 32

2.1.2 Échangeabilité 33

2.1.3 L'urne de Blackwell-MacQueen 34

2.1.4 Le processus du restaurant chinois 35

2.1.5 La construction « stick-breaking » 36

2.1.6 Mélange de processus de Dirichlet 39

2.1.7 Le mélange par processus de Dirichlet 41

2.2 Le modèle d'échantillonnage d'espèces 45

2.2.1 Séquence d'échantillonnage d'espèces 45
2.2.2 La fonction de probabilité des partitions échangeables 47
2.3 Le processus de Pitman-Yor 50
2.3.1 Définitions et propriétés 50
2.3.2 Le processus de Pitman-Yor 52
2.3.3 Le mélange par processus de Pitman-Yor 56
2.4 Les processus dépendants 57
2.4.1 Le processus de Dirichlet hiérarchique 57
2.4.2 Le processus de Dirichlet imbriqué 59
2.5 Les arbres de Pólya 60
2.5.1 Moments 61
2.5.2 Propriétés 62
2.5.3 Les arbres de Pólya canoniques 63
2.5.4 Paramétrisation de l'arbre de Pólya 63
2.5.5 Les arbres de Pólya finis 65
2.5.6 Processus dérivés d'arbres de Pólya 65
2.6 Conclusion 67

3 Méthodes MCMC **71**
3.1 Méthodes MCMC 72
3.1.1 Problématique 72
3.1.2 Algorithme de Metropolis-Hastings 73
3.1.3 Échantillonnage de Gibbs 75
3.1.4 Estimation des fonctionnelles 78
3.1.5 Diagnostic de convergence des algorithmes MCMC 78
3.2 Inférence dans les modèles de mélange par processus de Dirichlet 81
3.2.1 Méthodes marginales 82
3.2.2 Méthodes conditionnelles 93
3.2.3 Nouvelle méthode d'échantillonnage 101
3.2.4 Comparaisons des algorithmes 106
3.3 Conclusion 118

4 Reconstruction spatiale en TEP 3D **119**

4.1 Introduction à la TEP . . . 119
4.1.1 Principe . . . 119
4.1.2 Détection . . . 120
4.1.3 Acquisition des données . . . 121
4.2 Méthodes de reconstruction tomographique . . . 125
4.2.1 Méthodes analytiques . . . 126
4.2.2 La rétroprojection filtrée vue comme un modèle statistique . . . 130
4.2.3 Les méthodes du maximum de vraisemblance (ML) et du maximum *a posteriori* (MAP) . . . 133
4.2.4 Résumé et conclusion sur les méthodes de reconstruction . . . 139
4.3 Nouvelle méthode BNP de reconstruction 3D . . . 140
4.3.1 Formulation bayésienne du problème non paramétrique . . . 142
4.3.2 Modèles hiérarchiques génératifs pour les données TEP . . . 147
4.3.3 Inférence . . . 148
4.4 Comparaisons et résultats . . . 153
4.5 Conclusion . . . 167

5 Reconstruction spatio-temporelle en TEP 4D **169**
5.1 Introduction aux modèles compartimentaux . . . 169
5.2 Modèle compartimental dans le cas du FDG . . . 170
5.2.1 Le fluoro-déoxyglucose ($[^{18}F]$–FDG) . . . 170
5.2.2 Le modèle compartimental de Sokoloff . . . 171
5.3 Estimation des paramètres . . . 172
5.3.1 Les moindres carrés . . . 173
5.3.2 Analyse graphique . . . 173
5.3.3 Analyse spectrale . . . 174
5.3.4 La poursuite de base L_1 . . . 176
5.4 État de l'art de la reconstruction dynamique en TEP . . . 177
5.4.1 Reconstruction directe de cartes paramétriques . . . 178
5.4.2 Reconstructions spatio-temporelles . . . 180
5.4.3 Synthèse . . . 184
5.5 Nouvelle méthode de reconstruction en TEP 4D . . . 185
5.5.1 Formulation du problème . . . 185

5.5.2 Modèle hiérarchique pour les données TEP dynamiques 188
5.5.3 Modèle alternatif avec des variables de classification 189
5.5.4 Modèle alternatif avec des variables slices 190
5.5.5 Inférence 190
5.5.6 Estimées MCMC 194
5.5.7 Matériels de simulation et résultats 195
5.6 Conclusion 199

6 Conclusions et perspectives **209**
6.1 Conclusions 209
6.2 Perspectives 210
6.2.1 Estimation des hyperparamètres 210
6.2.2 Déconvolution des cinétiques 211
6.2.3 Accélération des calculs 212
6.2.4 Application aux données réelles 213

A Rappels sur la loi de Dirichlet **217**
A.1 La distribution de Dirichlet 217
A.1.1 Moments 217
A.1.2 Propriétés 218
A.2 Intégrale de Dirichlet 219

B Chaînes de Markov, convergence de MH et Gibbs **221**
B.1 Les chaînes de Markov 221
B.1.1 Homogénéité 221
B.1.2 Noyau de transition 222
B.1.3 Distribution invariante 222
B.1.4 Irréductibilité 224
B.1.5 Récurrence 224
B.1.6 Apériodicité 226
B.1.7 Récurrence au sens de Harris 226
B.2 Convergence de l'algorithme MH 227
B.2.1 Invariance 228
B.2.2 Irréductibilité 229

B.2.3 Récurrence 229
B.2.4 Apériodicité 229
B.3 Convergence de l'échantillonneur de Gibbs 230
B.3.1 Invariance 230
B.3.2 Irréductibilité 230
B.3.3 Récurrence 230

Notations

Les scalaires sont représentés par des minuscules, les vecteurs par des minuscules gras et les matrices par des majuscules gras.

Symbole	***Définition***
$\mathbb{R}$	ensemble des nombres réels.
$\mathbb{N}$	ensemble des entiers naturels.
$\mathbb{Z}$	ensemble des entiers relatifs.
$(\mathcal{X}, \mathcal{B})$	espace mesurable avec $\mathcal{X} \subset \mathcal{B}$ où $\mathcal{B}$ est une tribu sur $\mathcal{X}$. Typiquement, $\mathcal{X} \subseteq \mathbb{R}^d$ et $\mathcal{B}$ sera la tribu borélienne notée $\mathcal{B}(\mathcal{X})$ ou $\mathcal{B}$.
$(\mathcal{X}, \mathcal{B}, P)$	espace mesuré c'est-à-dire un espace mesurable $(\mathcal{X}, \mathcal{B})$ muni d'une mesure P. Si P est une mesure de probabilité, on parlera d'espace probabilisé.
$\mathbb{P}(A)$ ou $P(A)$	probabilité d'un évènement A.
$\delta_x(\cdot)$	mesure de Dirac sur un espace mesurable $(\mathcal{X}, \mathcal{B})$ telle que pour tout $B \in \mathcal{B}$, $\delta_x(B) = 1$ si $x \in B$, 0 sinon.
$\#B$	cardinalité d'un ensemble B.
ind	indépendant.
iid	indépendant et identiquement distribué.
$\propto$	proportionnel à.
$\sim$	distribué suivant.
$\mathbb{E}(X)$	Espérance de la variable aléatoire X.
$\mathbb{V}(X)$	Variance de la variable aléatoire X.
$\text{Bino}(n, p)$	Loi binomiale de paramètres n et p.
$\text{Mult}(\boldsymbol{p})$	Loi multinomiale de paramètre $\boldsymbol{p}$.
$\text{Beta}(a, b)$	Loi beta de paramètres a et b.
$\text{Dir}(\alpha_1, \ldots, \alpha_K)$	Distribution de Dirichlet de paramètres $\alpha_1, \ldots, \alpha_K$.
$\text{DP}(\alpha, G_0)$	Processus de Dirichlet de paramètres α et G_0.
$\text{GEM}(\alpha)$ ou $\text{GEM}(d, \alpha)$	Distribution des poids stick-breaking.
$\text{PD}(\alpha)$ ou $\text{PD}(d, \alpha)$	Distribution de Poisson-Dirichlet.
$\text{PY}(d, \alpha, G_0)$	Processus de Pitman-Yor de paramètres d, α et G_0.
$\mathcal{U}(\cdot\|a, b)$	Loi uniforme sur l'intervalle $]a, b]$.
$\mathcal{N}(\cdot\|\boldsymbol{m}, \boldsymbol{\Sigma})$	Loi normale multivariée de vecteur moyen $\boldsymbol{m}$ et de matrice de covariance $\boldsymbol{\Sigma}$.
$f_{\mathcal{N}}(\cdot\|\boldsymbol{m}, \boldsymbol{\Sigma})$	densité d'une loi $\mathcal{N}(\boldsymbol{m}, \boldsymbol{\Sigma})$.
$\text{Gamma}(\cdot\|a, b)$	Loi gamma de paramètres a et b.

$\Gamma(\cdot)$	Fonction gamma telle que $\Gamma(z) = \int_0^{+\infty} t^{z-1}e^{-t}\mathrm{d}t$ où z est un nombre complexe à partie réelle positive.
$\mathcal{W}(\cdot\|n, \mathbf{\Sigma})$	Loi de Wishart de paramètres n le degré de liberté et $\mathbf{\Sigma}$ la moyenne.
$\mathbf{1}_A(\cdot)$ ou $\mathbb{I}(A)$	fonction indicatrice telle que $\mathbf{1}_A(x) = 1$ si $x \in A$ et 0 sinon.

Abbréviations

Acronyme	***Définition***
3D	Trois dimensions de l'espace
4D	Trois dimensions de l'espace + dimension temporelle
BNP	Bayesian nonparametric (Bayésien non paramétrique)
CDF	Cumulative Distribution Function (Fonction de répartition)
CRP	Chinese Restaurant Process (processus du restaurant chinois)
DDP	Dependent Dirichlet Process (processus de Dirichlet dépendant)
DP	Dirichlet Process (processus de Dirichlet)
DPM	Dirichlet Process Mixture (mélange par processus de Dirichlet)
EM	Expectation-Maximization (algorithme de maximisation de l'espérance)
EPPF	Exchangeable Partition Probability Function (fonction de probablité des partitions échangeables)
EVP	Effets de Volume Partiel
FBP	Filtered BackProjection (rétroprojection filtrée)
FDG	Fluoro-Deoxy Glucose
HDP	Hierarchical Dirichlet Process (processus de Dirichlet hiérarchique)
IAT	Integrated Autocorrelation Time
IRM	Imagerie par Résonance Magnétique
ISBP	Invariance by Size-Biased Permutation (invariance par permutation size-biased)
LOR	Line Of Response (ligne de réponse)
MAP	Maximum *a posteriori*
MCMC	Markov Chain Monte-Carlo methods (méthodes de Monte-Carlo par chaînes de Markov)
MDP	Mixture of Dirichlet Processes (mélange de processus de Dirichlet)
MH	Algorithme de Metropolis-Hastings
ML	Maximum Likelihood (maximum de vraisemblance)
MPT	Mixture of Pólya Trees (mélange d'arbres de Pólya)
NDP	Nested Dirichlet Process (processus de Dirichlet imbriqué)
OSEM	Ordered Subsets Expectation-Maximization (méthode des sous-ensembles ordonnées pour la maximisation de l'espérance)
OSL	One Step Late (un algorithme de maximisation de la loi *a posteriori*)
PDF	Probability Density Function (fonction de densité de probabilité)
Pixel	PICture ELement (élément d'une image)
PT	Pólya Tree (arbre de Pólya)
PT_L	Arbre de Pólya tronqué au niveau L

PYP	Pitman-Yor Process (processus de Pitman-Yor)
RAM	Residual Allocation Model (modèle d'allocation des résidus)
ROI	Region Of Interest (région d'intérêt)
RPM	Random Probability Measure (mesure de probabilité aléatoire)
SNR	Signal to Noise Ratio (rapport signal sur bruit)
SPT	Shifted Pólya Trees (arbres de Pólya décalés)
SSM	Species Sampling Model (modèle d'échantillonnage d'espèces)
TAC	Time Activity Curve (courbe activité temps)
TEP	Tomographie par Émission de Positons
TOR	Tube Of Response (tube de réponse)
TR	Transformée de Radon
Voxel	VOlumetric piXEL (élement d'un volume)

1 Introduction générale

1.1 Contexte

Le but de cette thèse est de développer de nouvelles méthodes statistiques de reconstruction d'image spatiale (3D) et spatio-temporelle (3D+t) en Tomographie par Emission de Positons (TEP).

La TEP est une technique d'imagerie médicale fonctionnelle permettant de reconstruire chez l'Homme et chez le petit animal une image traduisant l'activité métabolique d'un organe. Elle diffère ainsi des techniques d'imagerie structurelle (radiologie, scanner, IRM) limitées à la production d'images purement anatomiques. La TEP repose sur l'injection au patient d'un traceur radiaoactif (appelé radiotraceur ou simplement traceur) émetteur de positons. Ces positons s'annihilent tout près de leurs lieux d'émission avec des électrons et cette annihilation produit deux photons gamma qui partent en direction opposée. Quand deux photons sont détectés dans un intervalle de temps assez proche, un évènement en *coïncidence* est enregistré, signifiant qu'une émission de positon a eu lieu sur la ligne virtuelle joignant les deux détecteurs en question (appelée LOR, pour line of response). L'objectif est, à partir des données mesurées, de reconstruire la distribution de l'activité du traceur dans l'organe cible.

D'un point de vue clinique, la TEP est utilisée en oncologie pour le diagnostic, le bilan d'extension et le suivi thérapeutique des cancers. Elle est aussi utilisée en neurologie pour l'étude des maladies neuro-dégénératives (Alzheimer, Parkinson) et pour l'épilepsie. On peut aussi citer la cardiologie comme exemple d'application de la TEP. L'une des contraintes majeures de la TEP réside dans l'injection au patient d'un traceur radioactif. Cependant, on observe une augmentation significative de l'exposition de la population aux rayonnements ionisants du fait de la multiplication des actes d'imagerie médicale. Ainsi, pour des questions de radioprotection, il serait utile d'appliquer un principe d'optimisation de la dose et qui consisterait à injecter la dose minimale possible qui soit compatible avec le but de l'examen. Pour cela, deux types de solutions seraient

envisageables. La première est de nature instrumentale et consisterait à agir au niveau de la surface de détection du tomographe mais son coût serait prohibitif. La deuxième voie est d'ordre méthodologique et serait de développer de nouvelles méthodes de reconstruction qui soient plus robustes dans un contexte de faibles doses et c'est l'objectif de ce travail. Ainsi, dans un contexte de diminution de la dose et sans nuire à la qualité de l'interprétation, nous nous proposons d'obtenir :

- des images avec de faibles doses de radiotraceurs,
- une estimation quantitative des volumes fonctionnels et de leur erreur associée.

1.2 Vue d'ensemble du travail effectué

Le travail effectué s'articule autour de trois parties principales : la reconstruction spatiale en trois dimensions, la reconstruction spatio-temporelle en quatre dimensions et le développement d'une méthode d'échantillonnage MCMC pour les modèles de mélange par processus de Pitman-Yor.

1.2.1 Reconstruction spatiale 3D

Nous avons, dans un premier temps, abordé la reconstruction statique en 3D. Les méthodes de reconstruction spatiale en tomographie d'émission sont usuellement divisées en deux grandes catégories : les méthodes analytiques et les méthodes statistiques.

1. Les premières méthodes de reconstruction développées en TEP sont dites *analytiques*. Elles utilisaient la rétroprojection simple puis une technique plus évoluée appelée rétroprojection filtrée (notée FBP pour Filtered Backprojection). Ces méthodes reposent sur l'inversion de la transformée en rayons X (qui coïncide avec la transformée de Radon en 2D) [Nat86]. Toutefois, les algorithmes de rétroprojection sont basés sur des versions idéalisées du système puisqu'il est supposé que la direction de la paire de photons gamma est une ligne droite : c'est le modèle de l'intégrale de ligne. Pour en simplifier la présentation, ce système idéal est illustré dans le cadre bidimensionnel dans la Figure 1.1. On note par f la fonction de densité de probabilité de la distribution radioactive du traceur, telle que $f(\boldsymbol{x})$ désigne l'activité présente au point de coordonnées (x, y). La densité de probabilité dans l'espace des détecteurs peut être modélisée par

$$g(x_r, \phi) = \frac{1}{2\rho} \int_{-\rho}^{\rho} f(x = x_r \cos\phi - y_r \sin\phi, y = x_r \sin\phi + y_r \cos\phi) \mathrm{d}y_r,$$

où la variable d'intégration y_r est la coordonnée le long de la ligne. Cette intégrale est appelée transformée de Radon (TR) de f. Elle est ici normalisée. La rétroprojection filtrée consiste en l'inversion de la TR pour reconstruire la densité f dans l'espace image. Les algorithmes de rétroprojection supposent donc que tous les (x_r, ϕ) sont exactement observés. Cela repose sur l'hypothèse implicite de l'existence d'un continuum de détecteurs. Cependant, en pratique nous sommes confrontés aux limitations du système. On peut citer entre autres le nombre fini restreint de détecteurs, la géométrie et l'efficacité des capteurs etc. (voir Figure 1.2). De plus, la rétroprojection filtrée est un problème inverse mal posé au sens où une

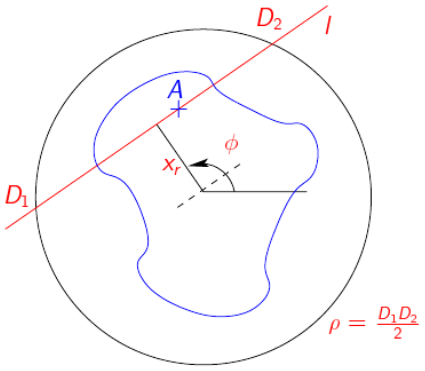


FIGURE 1.1 – Représentation schématique du problème idéalisé en 2D : patient, cercle de détection, LOR. Une émission ayant eu lieu au point A est détectée par la LOR D_1D_2. Cette dernière est paramétrée par deux variables définissant sa localisation et son orientation, le rayon x_r et l'angle ϕ avec l'axe vertical.

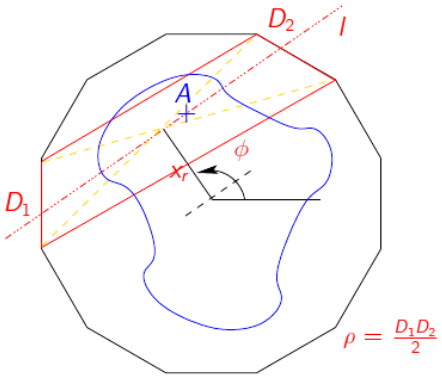


FIGURE 1.2 – Représentation schématique du problème réel en 2D : les directions des photons sont dans un volume parallélépipédique appelé TOR (pour tube of response) délimité par les ouvertures des détecteurs D_1 et D_2.

petite modification dans les données entraîne une grande variation dans l'image reconstruite. Cela se traduit par l'amplification du bruit dans les hautes fréquences. Tout cela combiné conduit à des images de reconstruction bruitées, présentant de gros artéfacts de reconstruction. La régularisation du problème est effectuée via le réglage du filtre passe-bas, nécessitant un compromis entre résolution spatiale et bruit.

2. Les méthodes *statistiques* basées sur l'approche du maximum de vraisemblance (ML) modélisent de façon plus réaliste le processus d'acquisition des données. Elles abordent le problème de reconstruction sous une forme discrétisée et minimisent l'anti-logarithme de la vraisemblance poissonienne des données de façon itérative ([SV82], [LC84] et [VSK85]). N'étant pas basées sur le modèle de l'intégrale de ligne, elles permettent ainsi d'incorporer dans le modèle plus de phénomènes physiques et de traiter des géométries non-standard. En conséquence, elles fournissent des images plus précises que les méthodes analytiques. Elles présentent toutefois un fort bruit de reconstruction dans chaque voxel dû au mauvais conditionnement de la matrice-système. Dans les plate-formes cliniques, la régularisation repose

sur des approches empiriques et consiste soit à post-filtrer l'image reconstruite, soit à initialiser l'algorithme avec une image uniforme et à le stopper avant la convergence.

Une meilleure approche consiste à régulariser pendant la reconstruction en pénalisant la vraisemblance poissonienne. Le critère à minimiser comporte ainsi un terme d'adéquation aux données et un terme de régularisation. Dans les méthodes *statistiques bayésiennes*, la régularisation est effectuée par l'information *a priori* sur les propriétés de l'image à reconstruire [Gre90], [GM87]. Lorsqu'elles sont utilisées en tomographie d'émission, les techniques bayésiennes sont basées sur l'approche du maximum *a posteriori* et reposent sur l'optimisation d'un critère consistant en la vraisemblance (terme d'attache aux données) et d'un *a priori* sur l'image (terme de régularisation).

Cependant, toutes les méthodes statistiques actuelles de reconstruction en TEP, qu'elles soient bayésiennes ou non, sont basées sur la formulation discrète de la distribution du radiotraceur dans l'espace sur lequel la reconstruction est à effecteur Cet espace est ainsi subdivisé en un nombre fini de voxels et la concentration du traceur est supposée constante dans chaque voxel. Cette discrétisation est source d'un certain nombre de biais puisque la taille du voxel doit être déterminée en amont sans une connaissance fine de la structure de l'objet à imager. La taille du voxel agit comme un paramètre de régularisation puisque plus elle est grande, plus l'image est lissée et moins bonne est la résolution spatiale. De plus, il n'apparaît pas naturel de procéder à une discrétisation en amont pour reconstruire la distribution du traceur qui est continue.

Dans ce contexte, nous avons abordé le problème de reconstruction comme un problème *indirect d'estimation de densité* dans un cadre *statistique bayésien non paramétrique*. La modélisation non paramétrique consiste à considérer l'image à reconstruire comme une densité de probabilité dans $\mathbb{R}^3$ et permet d'inférer directement dans l'espace, sans aucune discrétisation préalable. Pour cela, nous avons représenté la distribution d'intérêt (distribution volumique du traceur) sous forme d'une distribution de probabilité dans $\mathbb{R}^3$, dont la loi *a priori* est un mélange par processus de Pitman-Yor (généralisation à deux paramètres des processus de Dirichlet) [Pit96b]. La modélisation bayésienne adoptée permet de caractériser toute la loi de la distribution d'intérêt et ainsi de quantifier le degré de confiance associé à la reconstruction.

Le modèle génératif des données TEP que nous avons élaboré a permis de dériver l'algorithme de reconstruction MCMC associé. Le modèle proposé a été évalué sur données synthétiques issues d'un fantôme réaliste 3D. Les résultats ont mis en évidence la capacité du modèle à capturer la structure de l'activité spatiale cérébrale. Ce point revêt une importance particulière car l'approche non paramétrique que nous avons suivie substitue à la base finie de voxels classiquement utilisée en reconstruction, une somme potentiellement infinie de lois normales. Par ailleurs, nous avons illustré les capacités de notre approche à quantifier l'erreur au travers de la détermination des intervalles de crédibilité.

1.2.2 Reconstruction spatio-temporelle 4D

Après cette première étape, nous nous sommes attelés à la reconstruction d'images spatio-temporelles en 4D, toujours dans un contexte de faibles doses. La reconstruc-

tion d'images dynamiques est importante en TEP car elle apporte des informations supplémentaires sur la physiologie des tissus (taux de métabolisation du glucose, fonctionnement de neuro-récepteurs etc.).

En routine clinique, l'étude dynamique repose sur une segmentation du temps d'acquisition total en intervalles. Les données dans chaque tranche temporelle sont reconstruites indépendamment des autres. Cela induit une perte statistique en ignorant les données avant et après chaque tranche et la qualité de la reconstruction est d'autant plus dégradée que la durée de l'intervalle est courte. A contrario, des tranches plus longues biaisent le comportement dynamique.

Des techniques ont été développées et elles permettent d'incorporer la modélisation des cinétiques traduisant la dynamique du traceur dans la reconstruction des images et de traiter conjointement les deux processus. Ces méthodes décrivent l'activité dans chaque voxel et à chaque instant par une combinaison linéaire de fonctions de bases temporelles pondérées par des poids. Ces fonctions de base peuvent être prédéfinies et dans ce cas seules les pondérations sont à estimer. Plusieurs méthodes ont été développées et elles diffèrent principalement de par les fonctions temporelles utilisées ([MBPC97], [GGT+02], [NQAL02], [AL06]). Le modèle temporel peut aussi être ajusté pendant la reconstruction. Par exemple, [RSC+06] partant d'un nombre fini de fonctions temporelles (dictionnaire de fonctions) estiment les cinétiques présentes. Toutefois, ces cinétiques sont estimées à partir d'un nombre fini et fixé de fonctions de base temporelles. Il apparaît alors dans ces algorithmes, qu'en plus du problème qu'engendre le choix des voxels et que nous avons déjà évoqué, vient s'ajouter celui du choix du type et/ou du nombre des fonctions de base temporelles permettant de décrire de façon adéquate le comportement temporel de l'activité dans chaque voxel. Ainsi, tout comme la taille du voxel, le nombre des fonctions de base est aussi un paramètre de régularisation : plus le nombre de fonctions utilisées est faible, plus la régularisation est forte. Ce choix apparaît donc crucial mais est tout autant difficile : c'est le problème de la sélection de modèle dans les méthodes paramétriques.

Nous avons donc formulé la reconstruction 4D dans un cadre bayésien non paramétrique. La distribution d'activité spatio-temporelle est vue comme une distribution de probabilité dans $\mathbb{R}^4$ et reconstruite directement à partir des données. Cependant, l'aspect dynamique requiert une modélisation plus délicate car nous sommes en présence d'un mélange statistique de plusieurs contributions présentant des cinétiques différentes. L'abord de la reconstruction dans un cadre non paramétrique a pour conséquence que le nombre et la forme de ces cinétiques sont inconnues, en cela résident l'originalité mais aussi la difficulté du problème. Nous avons modélisé la densité d'émission spatio-temporelle par un processus dépendant basé sur un mélange par processus de Pitman-Yor (pour la partie spatiale) et d'arbres de Pólya (pour la partie temporelle). Les arbres de Pólya permettent de générer des mesures aléatoires presque sûrement continues [Lav92]. Le modèle hiérarchique résultant a permis d'inférer sur les volumes fonctionnels qui sont des zones du cerveau dont l'activité suit une certaine cinétique. Une des caractéristiques de l'approche proposée est que la découverte des cinétiques et des volumes fonctionnels est pilotée par les données. De plus, l'algorithme MCMC permet de générer des échantillons suivant la loi *a posteriori* de la distribution spatio-temporelle et l'on peut estimer l'erreur associée à la reconstruction ainsi que n'importe quelle fonctionnelle (moyenne, variance, intervalle de crédibilité etc.). Cette évaluation de l'erreur n'est que partiellement couverte par les techniques usuelles en TEP dynamique.

1.2.3 Nouvelle méthode d'échantillonnage

Dans l'approche bayésienne, la complexité analytique des lois *a posteriori* implique de faire appel à des techniques d'approximation pour l'inférence. La plupart de ces techniques d'échantillonnage sont basées sur les méthodes de Monte-Carlo par chaînes de Markov (MCMC). La principale difficulté dans notre approche est que les méthodes usuelles d'inférence ne s'appliquent plus puisque l'on travaille avec des objets aléatoires infini-dimensionnels. Deux classes de méthodes d'échantillonnage ont été développées dans la littérature, les méthodes marginales et les méthodes conditionnelles, chacune présentant des avantages et des inconvénients.

Afin de circonscrire le problème de la nature infinie de la distribution aléatoire, cette dernière est intégrée dans les approches marginales et on travaille seulement avec des échantillons générés suivant elle [Esc94], [MM98], [Nea00]. Dans les versions évoluées de ces méthodes, l'algorithme MCMC présente de bonnes capacités de mélange [Nea00]. Cependant, ces techniques sont basées sur des mises à jour incrémentales, ce qui les rend très séquentielles et difficilement parallélisables. Cette caractéristique est préjudiciable quand l'on travaille avec de gros volumes de données comme c'est le cas en TEP (une dizaine de millions d'évènements dans cette thèse).

Contrairement aux méthodes marginales, les approches conditionnelles représentent explicitement la distribution aléatoire. Afin de traîter le nombre infini de ses composantes, certains la tronquent de façon déterministe [IJ01]. D'autres techniques plus évoluées utilisent une troncature probabiliste par une stratégie dite de slice sampling. Elle permet, via l'introduction de variables auxiliaires, de travailler avec un nombre fini de composantes à chaque itération de l'algorithme tout en gardant la nature infinie de la distribution [Wal07], [KGW09]. Néanmoins, elles utilisent une représentation non-échangeable du processus qui contribue à diminuer ses performances de mélange.

Nous avons développé une nouvelle méthode d'échantillonnage des mélanges par processus de Pitman-Yor, donc applicable aux mélanges par processus de Dirichlet, et qui essaie d'allier les avantages de chacune des deux classes. Cet échantillonneur permet, par la stratégie du slice sampling, d'inférer sur ces mesures sans troncature et, de par l'espace où il est formulé, de conférer une propriété d'échangeabilité intrinsèque à la représentation. Cette combinaison fournit à la fois de bonnes capacités de mélange pour l'échantillonneur et la possibilité d'être parallélisé afin de traiter de gros volumes de données.

1.3 Organisation du manuscrit

Le reste du manuscrit est organisé comme suit :

Dans le chapitre 2, nous présentons les principaux aspects et propriétés de certains modèles bayésiens non paramétriques. L'accent sera mis sur ceux qui ont été utilisés dans ce travail à savoir le processus de Dirichlet, le processus de Pitman-Yor et l'arbre de Pólya.

Le chapitre 3 est consacré aux méthodes de Monte-Carlo par chaînes de Markov. Nous le débutons par un rappel sur les algorithmes de Metropolis-Hastings et de Gibbs,

puis nous abordons les méthodes d'inférence des mélanges par processus de Dirichlet. Enfin, nous présentons la nouvelle méthode d'échantillonnage que nous avons développée et la comparons avec trois autres techniques.

Dans le chapitre 4, on introduit tout d'abord les principes physiques liés à la TEP avant de procéder à un rappel des algorithmes classiques de reconstruction en TEP. Après cela, nous introduisons la méthode de reconstruction que nous avons développée pour la TEP 3D et la comparons avec deux approches classiques à savoir l'approche du maximum de vraisemblance (ML) et celle du maximum *a posteriori* (MAP).

Le chapitre 5 concerne la reconstruction spatio-temporelle en 4D. Nous le débutons par un rappel sur les modèles compartimentaux. Puis nous procédons à un état de l'art des méthodes de reconstruction d'images dynamiques en TEP et enfin présentons l'algorithme de reconstruction proposé.

Dans le chapitre 6, nous concluons ce travail et présentons ses différentes perspectives.

2

Modélisation bayésienne non paramétrique

Dans ce chapitre, nous passons en revue une partie des modèles bayésiens non paramétriques de la littérature. On décrit le processus de Dirichlet ainsi que quelques unes de ses diverses extensions et généralisations à savoir sa généralisation à deux paramètres (le processus de Pitman-Yor), son extension hiérarchique (le processus de Dirichlet hiérarchique et le processus de Dirichlet imbriqué) et l'arbre de Pólya dont le processus de Dirichlet est un cas particulier. La liste des modèles présentés ici est loin d'être exhaustive. L'objectif est de présenter ceux dont il sera question dans cette thèse et, par la même occasion, d'introduire les concepts clés qui seront mentionnés dans la suite du manuscrit. Une insistance particulière sera faite sur le processus de Dirichlet (DP), le processus de Pitman-Yor (PYP) ainsi que l'arbre de Pólya (PT) qui sont à la base des méthodes que nous avons développées pour la reconstruction TEP en 3D et en 4D. Dans ce chapitre, on se focalise sur la spécification des lois *a priori* dans des espaces de dimension infinie. Les problèmes liés au calcul seront abordés dans le chapitre 3 sur l'échantillonnage.

Introduction

Les modèles statistiques peuvent être utilisés afin de résumer l'information contenue dans des données observées et de pouvoir faire des prédictions futures. La modélisation bayésienne permet de prendre en compte la connaissance *a priori* sur le processus génératif des données. Cette connaissance est exprimée au travers de distributions *a priori* sur les paramètres du modèle.

Dans les modèles dits « paramétriques », la distribution des données est supposée appartenir à une certaine famille paramétrique. En conséquence, le nombre de paramètres (les inconnues) dans le modèle est fini. Par opposition, on désigne sous l'oxymore « modèle non paramétrique » un modèle ayant un nombre infini dénombrable de paramètres. Pour fixer les idées, considérons $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$, n observations qui sont des

réalisations de variables aléatoires réelles ou multivariées $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n$ toutes définies sur le même espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ et à valeurs dans $(\mathcal{X}, \mathcal{B})$. Soient $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n \overset{iid}{\sim} P$ où P est une distribution de probabilité inconnue. Supposons que la mesure P admette une densité $p(\cdot)$ et que l'on se donne pour objectif de déterminer cette densité à partir des observations $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$. C'est un exemple de problème inverse mal posé, bien connu en Statistique. On va donc chercher $p(\boldsymbol{x})$ sous la forme d'une fonction $f_{\boldsymbol{\theta}}(\boldsymbol{x})$ appartenant à une certaine famille $\mathcal{F} = \{p_{\boldsymbol{\theta}} : \boldsymbol{\theta} \in \boldsymbol{\Theta}\}$. Si les paramètres du modèle peuvent être décrits sous la forme d'un vecteur de taille finie, le modèle est dit *paramétrique*. La famille $\mathcal{F}$ peut alors s'écrire $\mathcal{F} = \{p_{\boldsymbol{\theta}} : \boldsymbol{\theta} \in \boldsymbol{\Theta} \subset \mathbb{R}^m\}$ où m est un nombre entier fini. Cependant, pour permettre plus de flexibilité et de robustesse vis-à-vis des données, on peut vouloir considérer les modèles où le paramètre $\boldsymbol{\theta}$ est infini dimensionnel. Dans ce cas, le modèle est appelé *non paramétrique*. L'utilisation des modèles non paramétriques est motivée par le fait que les modèles paramétriques peuvent sur-représenter ou sous-représenter les données s'il y'a inadéquation entre les données et la complexité du modèle. D'où la nécessité de choisir un modèle approprié. Cependant dans la pratique, ce choix est difficile à réaliser. L'approche bayésienne non paramétrique est donc une alternative au choix de modèle puisque le nombre de paramètres dans le modèle est potentiellement infini. Dans l'exemple d'estimation de densité, le paramètre sera la fonction de densité elle-même. L'espace considéré est alors $\mathcal{F} = \{p(\boldsymbol{x}) : \boldsymbol{x} \in \mathcal{X}, p \text{ est une densité de probabilité dans } \mathcal{X}, p \subset \mathcal{M}(\mathcal{X})\}$, où $\mathcal{M}(\mathcal{X})$ désigne l'ensemble des mesures de probabilité sur $\mathcal{X}$. Nous reviendrons plus en détails sur cet exemple dans la section 2.1.7. Mais on voit d'ores et déjà qu'une approche de la sorte implique de définir des mesures de probabilité sur une collection de mesures. De telles mesures sont appelées mesures de probabilité aléatoires (notées RPM pour random probability measures). La théorie des processus stochastiques traite de ce genre d'objets aléatoires à valeurs dans des espaces fonctionnels.

2.1 Le processus de Dirichlet et ses représentations

Bien que l'étude des fonctions aléatoires date d'assez longtemps (du temps de Bayes), son application aux statistiques bayésiennes non paramétriques est assez récente. En effet, ce n'est que dans les années 60 que David Blackwell et d'autres à Berkeley se sont posés la question de comment combiner les statistiques bayésiennes aux modèles non paramétriques. Thomas Ferguson [Fer73] formalisa alors une mesure de probabilité aléatoire qu'il appela processus de Dirichlet, qui en fait fût proposé plus tôt par David Freedman en 1963 [Fre63]. Les processus stochastiques, appelés aussi processus aléatoires, sont des distributions sur des espaces fonctionnels dont les réalisations (trajectoires du processus) sont des fonctions. Le processus de Dirichlet (DP) est un processus stochastique dont les réalisations sont des mesures de probabilité. C'est donc une distribution sur des distributions c'est-à-dire que chaque réalisation du processus de Dirichlet est elle-même une distribution de probabilité. À ce jour, le processus de Dirichlet est l'un des modèles bayésiens non paramétriques les plus connus et les plus utilisés et ce grâce à ses nombreuses propriétés.

Le processus de Dirichlet

Dans cette section, nous rappelons la définition formelle d'un processus de Dirichlet et quelques unes de ses propriétés. Puis, nous nous focaliserons sur les différentes représentations de ce processus.

Définition 1. *Un processus de Dirichlet est un processus stochastique* $\{H(A), A \in \mathcal{B}\}$ *dont les distributions fini-dimensionnelles sont distribuées suivant la loi de Dirichlet. Plus précisément, soient* $(\Theta, \mathcal{B})$ *un espace mesurable où* Θ *est un espace métrique séparable. Considérons* $(A_1, \ldots, A_K)$ *une partition mesurable finie de* Θ *:*

$$\bigcup_{j=1}^{K} A_j = \Theta, \qquad A_i \cap A_j = \emptyset \text{ si } i \neq j$$

Soient α *un nombre réel positif et* G_0 *une distribution de probabilité sur* $(\Theta, \mathcal{B})$*. On dit que* H *est distribuée suivant un processus de Dirichlet sur* $(\Theta, \mathcal{B})$ *de paramètres* α *et* G_0 *(on écrit* $H \sim DP(\alpha, G_0)$*) si,*

$$(H(A_1), \ldots, H(A_K)) \sim Dir(\alpha G_0(A_1), \ldots, \alpha G_0(A_K)). \tag{2.1}$$

On notera $\mathcal{P} \equiv \mathrm{DP}(\alpha G_0) \equiv \mathrm{DP}(\alpha, G_0)$ la loi du processus.

Démonstration. Ferguson [Fer73], utilisant les conditions de consistance de Kolmogorov, a montré l'existence du DP comme processus stochastique dont les distributions finies dimensionnelles sont données par la loi de Dirichlet.

Remarque 1. *Ferguson [Fer73] et d'autres auteurs définissent le DP avec un seul paramètre* $\nu(\cdot)$*. Dans ce cas,* $\nu(\cdot) = \alpha G_0(\cdot)$ *où* $\alpha = \nu(\Theta)$ *est la masse totale de* ν *et la distribution de base est donnée par* $G_0(\cdot) = \frac{\nu(\cdot)}{\nu(\Theta)}$*. Bien que cette paramétrisation puisse être utile en termes de notations, elle perd néanmoins en vue les rôles distincts de* α *et de* G_0 *dans la description d'un DP. En ce qui nous concerne, nous utiliserons donc la notation avec les deux paramètres.*

Le processus de Dirichlet est défini par ses distributions fini-dimensionnelles qui suivent une loi de Dirichlet. De ce fait, le DP possède plusieurs propriétés héritées de la distribution de Dirichlet (les propriétés de la distribution de Dirichlet sont rappelées dans l'annexe A). En particulier on a pour tout ensemble mesurable A,

$$\mathbb{E}[H(A)] = G_0(A),$$
$$\mathbb{V}[H(A)] = \frac{G_0(A)(1 - G_0(A))}{1 + \alpha}.$$

Démonstration. Pour montrer que $\mathbb{E}[H(A)] = G_0(A)$, il suffit de considérer la partition (A, A^c) de Θ. On a alors $(H(A), H(A^c)) \sim \mathrm{Beta}(\alpha G_0(A),\ \alpha(1 - G_0(A))$. D'où :

$$\mathbb{E}[H(A)] = \frac{\alpha G_0(A)}{\alpha G_0(A) + \alpha(1 - G_0(A))} = G_0(A).$$
$$\mathbb{V}[H(A)] = \frac{\alpha G_0(A)(\alpha - \alpha G_0(A))}{\alpha^2(1 + \alpha)} = \frac{G_0(A)(1 - G_0(A))}{1 + \alpha}.$$

□

$\mathbb{E}[H(A)] = G_0(A)$ signifie que si $\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_n | H \overset{iid}{\sim} H$ (on dira que $\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_n$ est un échantillon de H) avec $H \sim \mathrm{DP}(\alpha, G_0)$, alors G_0 est la distribution marginale des $\boldsymbol{\theta}_i$, appelée aussi distribution inconditionnelle. De plus, pour toute fonction ψ G_0-intégrable, $\mathbb{E}\left(\int \psi \mathrm{d}H\right) = \int \psi \mathrm{d}G_0$. Par ailleurs, quand $\alpha \to +\infty$, on a $H(A) \to G_0(A)$ pour tout ensemble mesurable A, c'est-à-dire que $H \rightharpoonup G_0$ (convergence faible ou point par point). Cependant, cela n'est pas équivalent à dire que $H \to G_0$. Comme on le verra, les mesures générées par un DP sont discrètes presque-sûrement. Le paramètre α peut être vu comme une variance-inverse : plus α est élevé, plus la variance est faible et plus les mesurées générées par un DP sont concentrées autour de la moyenne G_0, d'où son nom de *paramètre de concentration*.

Une propriété importante du processus de Dirichlet est sa conjugaison c'est-à-dire que sa distribution *a posteriori* est aussi un processus de Dirichlet [Fer73].

Théorème 1. *Soient $H \sim DP(\alpha G_0)$ et $\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_n | H \overset{iid}{\sim} H$. Alors,*

$$H | \boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_n \sim DP\left(\alpha G_0 + \sum_{i=1}^{n} \delta_{\boldsymbol{\theta}_i}\right). \tag{2.2}$$

Démonstration. Utilisant la notation avec les deux paramètres séparés, on peut aussi ré-écrire l'équation 2.2 sous la forme :

$$H | \boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_n \sim \mathrm{DP}\left(\alpha + n,\ \frac{1}{\alpha + n}\left(\alpha G_0 + \sum_{i=1}^{n} \delta_{\boldsymbol{\theta}_i}\right)\right).$$

On remarque d'abord que si $H \sim \mathrm{DP}(\alpha, G_0)$ et $\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_n | H \sim H$, alors par conjugaison entre la loi de Dirichlet et la loi multinomiale on a pour toute partition $A_1, \ldots, A_K$ de Θ,

$$(H(A_1), \ldots, H(A_K)) | \boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_n \sim \mathrm{Dir}(\alpha G_0(A_1) + n_1, \ldots, \alpha G_0(A_K) + n_K) \tag{2.3}$$

où $n_k = \#\{i : \boldsymbol{\theta}_i \in A_k\}$. Comme ce résultat est vrai pour toute partition finie mesurable de Θ, alors utilisant la condition de consistance de Kolmogorov, la loi *a posteriori* du DP est aussi un DP. Il reste à trouver ses paramètres. Soit $H' := H | \boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_n \sim \mathrm{DP}(\alpha', G_0')$ avec $\alpha' > 0$ et G_0' une mesure de probabilité sur Θ. On a alors

$$(H'(A_1), \ldots, H'(A_K)) \sim \mathrm{Dir}(\alpha' G_0'(A_1), \ldots, \alpha' G_0'(A_K)).$$

On a donc d'après l'équation 2.3,

$$(H'(A_1), \ldots, H'(A_K)) \sim \mathrm{Dir}(\alpha G_0(A_1) + n_1, \ldots, \alpha G_0(A_K) + n_K).$$

En sommant les paramètres de ces deux lois, on obtient

$$\alpha'(G_0'(A_1) + \ldots + G_0'(A_K)) = \alpha(G_0(A_1) + \ldots + G_0(A_K)) + n_1 + \cdots + n_K.$$

D'où $\alpha' = \alpha + n$.

Par ailleurs, pour tout ensemble mesurable A_k on a :

$$\mathbb{E}[H'(A_k)] = \frac{\alpha' G_0'(A_k)}{\sum_{j=1}^{K} \alpha' G_0'(A_j)} = G_0'(A_k) = \frac{\alpha G_0(A_k) + n_k}{\sum_{j=1}^{K} (\alpha G_0(A_j) + n_j)} = \frac{\alpha G_0(A_k) + n_k}{\alpha + n}.$$

D'où le résultat. □

Le paramètre de concentration α devient $\alpha + n$ après avoir observé n observations et la mesure de base $G_0(\cdot)$ devient $\frac{\alpha G_0(\cdot)+\sum_{i=1}^n \delta_{\boldsymbol{\theta}_i}(\cdot)}{\alpha+n}$.

L'espérance *a posteriori* est :

$$\widetilde{H}_n(\boldsymbol{\theta}) := \mathbb{E}[H(\boldsymbol{\theta})|\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_n] = \frac{\alpha}{\alpha+n} G_0(\boldsymbol{\theta}) + \frac{n}{\alpha+n}\hat{H}_n(\boldsymbol{\theta})$$

où $\hat{H}_n(\boldsymbol{\theta}) = \frac{1}{n}\sum_{i=1}^n \delta_{\boldsymbol{\theta}_i}(\boldsymbol{\theta})$.

La variance *a posteriori* est donnée par :

$$\mathbb{V}[H(\boldsymbol{\theta})|\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_n] = \frac{\widetilde{H}_n(\boldsymbol{\theta})(1-\widetilde{H}_n(\boldsymbol{\theta}))}{1+\alpha+n}.$$

Puisque l'estimateur bayésien optimal sous une perte quadratique est donnée par l'espérance, on a alors comme estimée bayésienne pour H :

$$\hat{H}(\cdot) = \widetilde{H}_n(\cdot).$$

L'espérance *a posteriori* est une combinaison convexe de la moyenne *a priori* et de la cdf empirique. Quand $\alpha \to 0$ ou $n \to +\infty$, l'estimateur $\hat{H}$ converge alors vers la distribution empirique $\hat{H}_n$, ce qui confère une propriété de consistance au processus de Dirichlet.

On peut tirer du théorème 1 le résultat suivant, souvent utilisé dans les algorithmes MCMC qu'on abordera dans le chapitre 3.

Distribution prédictive

Soit $H \sim \mathrm{DP}(\alpha, G_0)$ et considérons $\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \ldots | H \overset{iid}{\sim} H$. On s'intéresse à la distribution prédictive de $\boldsymbol{\theta}_{n+1}$ sachant $\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_n$. On a :

$$P(\boldsymbol{\theta}_{n+1}|\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_n) = \int P(\boldsymbol{\theta}_{n+1}|H, \boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_n)\mathrm{d}\mathcal{P}_{\boldsymbol{\theta}}(H)$$

où on désigne par $\mathcal{P}_{\boldsymbol{\theta}}$ la loi de $(H|\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_n)$. Puisque $\boldsymbol{\theta}_{n+1}|\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_n, H \sim H$, l'intégrale n'est alors autre que l'espérance de H sous la loi $\mathcal{P}_{\boldsymbol{\theta}}$. On a alors pour tout ensemble mesurable A,

$$\mathbb{P}(\boldsymbol{\theta}_{n+1} \in A|\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_n) = \mathbb{E}[H(A)|\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_n] = \frac{\alpha}{\alpha+n} G_0(A) + \frac{1}{\alpha+n}\sum_{i=1}^n \delta_{\boldsymbol{\theta}_i}(A).$$

Le fait que la distribution prédictive soit aussi la distribution de base dans la loi *a posteriori* du DP n'est pas surprenant puisque dans le DP, l'espérance *a posteriori* est aussi la mesure de base de la loi *a posteriori*.

Pour finir, signalons que les réalisations d'un DP sont presque-sûrement discrètes et ce, même si la mesure de base G_0 est continue. En effet, on a pour tout $A \subset \Theta$,

$$\mathbb{E}[H(A)|\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_n] = \frac{\alpha}{\alpha+n} G_0(A) + \frac{1}{\alpha+n}\sum_{i=1}^n \delta_{\boldsymbol{\theta}_i}(A) \underset{n\to+\infty}{\longrightarrow} \sum_{k=1}^{\infty} w_k \delta_{\boldsymbol{\theta}_k^*}(A) \quad p.s. \quad (2.4)$$

où les $\boldsymbol{\theta}_k^*$ sont les valeurs uniques des $\boldsymbol{\theta}_i$, $w_k = \lim_{n+\to\infty} \frac{n_k}{n}$ et $n_k = \{i : \boldsymbol{\theta}_i = \boldsymbol{\theta}_k^*\}$.
Si l'on conjecture que la loi *a posteriori* du DP est elle aussi concentrée autour de sa

moyenne comme la loi *a priori*, alors la formule 2.4 suggère que les mesures générées par un DP sont discrètes (presque-sûrement). Ferguson [Fer73] a montré cette propriété en utilisant la représentation du DP par un processus gamma. Une autre preuve de cette nature discrète est donnée dans [Bla73]. Sethuraman [Set94] en a fourni une preuve formelle basée sur une représentation dite « stick-breaking ».

Nous allons maintenant voir les différentes représentations du processus de Dirichlet.

2.1.1 Représentation par processus gamma

Ferguson [Fer73] fournit, en sus de la définition formelle du processus de Dirichlet, une construction basée sur le processus gamma normalisé. L'idée sous-jacente étant que puisque la distribution de Dirichlet peut être définie par la distribution jointe de variables aléatoires indépendantes suivant une loi gamma et divisées par leur somme (cf. annexe A), le processus de Dirichlet peut aussi être défini par un processus gamma à incréments indépendants divisé par la somme. Avant de décrire cette représentation, commençons par quelques rappels.

Processus de Lévy
Un processus stochastique $(Z_t, t \geq 0)$ avec $Z_0 = 0$ est un processus de Lévy si ses incréments sont stationnaires et indépendants. Autrement dit,

- pour tout $0 \leq t_0 \leq t_n$, les incréments $Z_{t_1} - Z_{t_0}, \ldots, Z_{t_n} - Z_{t_{n-1}}$ sont indépendants,
- pour tout $0 \leq s < t$, la loi de l'incrément $Z_t - Z_s$ dépend seulement de $t - s$.

Un processus de Lévy est *infiniment divisible*, c'est-à-dire que pour tout t et pour tout n, $Z_t - Z_0$ s'écrit comme une somme de n variables aléatoires *iid*. Les processus de Poisson, brownien, gamma sont des exemples de processus de Lévy.

Ferguson et Klass [FK72] ont montré qu'un processus à incréments indépendants peut s'écrire comme une somme d'un nombre dénombrable de hauteurs de sauts aléatoires en un nombre dénombrable de points aléatoires. Utilisant cette représentation on peut, en divisant par la somme des hauteurs de sauts, obtenir une mesure de probabilité discrète distribuée suivant un processus de Dirichlet.

Soient $\alpha > 0$, $t \in [0, 1]$ et $Q(t)$ une fonction de répartition sur $[0, 1]$. On définit les hauteurs de sauts aléatoires de la façon suivante :

$$\mathbb{P}(J_1 \leq x_1) = \exp(N(x_1)) \quad \text{pour } x_1 > 0$$

et pour $j = 2, 3 \ldots,$

$$\mathbb{P}(J_j \leq x_j | J_{j-1} = x_{j-1}, \ldots, J_1 = x_1) = \exp(N(x_j) - N(x_{j-1})) \text{ pour } 0 < x_j < x_{j-1}$$

où $N(\cdot)$ désigne la mesure de Lévy du processus gamma $\Gamma\mathrm{P}(\alpha Q(t), 1)$ et est donnée par :

$$N(x) = \alpha \int_x^\infty u^{-1} \exp(-u) \mathrm{d}u \text{ pour } x > 0.$$

On définit les variables aléatoires relatives aux temps de sauts $U_1, U_2, \ldots$ suivant *iid* une loi uniforme sur $[0, 1]$, indépendamment des $J_1, J_2, \ldots$. Alors,

$$Z_t = \sum_{j=1}^{\infty} J_j \mathbb{I}_{[0, Q(t)]}(U_j)$$

converge presque-sûrement pour tout $t \in [0,1]$. De plus, c'est un processus gamma à incréments indépendants, $Z_t \in \Gamma\mathrm{P}(\alpha Q(t), 1)$. En particulier, $Z_1 = \sum_{j=1}^{\infty} J_j$ converge presque-sûrement et $Z_1 \sim \mathrm{Gamma}(\alpha, 1)$. Si on pose $p_j = \frac{J_j}{Z_1}$, alors $p_j \geq 0$ et $\sum_{j=1}^{\infty} p_j = 1$ presque-sûrement.

On considère $\widetilde{J}_1 > \widetilde{J}_2 > \dots$ les tailles ordonnées des hauteurs de sauts $J_1, J_2, \dots$ et on pose

$$\widetilde{p_k} = \frac{\widetilde{J}_k}{\sum_{j=1}^{\infty} \widetilde{J}_j}. \tag{2.5}$$

Soient $\boldsymbol{\theta}_k \overset{iid}{\sim} G_0$ et indépendamment des $\widetilde{p_k}$, alors

$$H = \sum_{k=1}^{\infty} \widetilde{p}_k \delta_{\boldsymbol{\theta}_k} \sim \mathrm{DP}(\alpha, G_0). \tag{2.6}$$

Cette représentation montre explicitement que la mesure H générée par un DP est discrète. Cependant, l'utilisation pratique de cette construction est limitée par le fait qu'il faille évaluer une somme infinie pour évaluer la valeur de chacun des poids dans l'équation 2.5.

Avant d'aborder les autres représentations du processus de Dirichlet, nous allons introduire une notion très importante pour la suite.

2.1.2 Échangeabilité

Définition 2.

1. *Une séquence $\boldsymbol{\theta}_1, \dots, \boldsymbol{\theta}_n$ de variables aléatoires est dite échangeable si*

$$(\boldsymbol{\theta}_1, \dots, \boldsymbol{\theta}_n) \overset{d}{=} (\boldsymbol{\theta}_{\sigma(1)}, \dots, \boldsymbol{\theta}_{\sigma(n)})$$

   pour toute permutation $\sigma \in S(n)$ où $S(n)$ est le groupe des permutations de $\{1, \dots, n\}$.
2. *Une séquence $\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \dots$ est dite infiniment échangeable si pour tout $n \geq 1$, $\boldsymbol{\theta}_1, \dots, \boldsymbol{\theta}_n$ est échangeable.*

Dans la suite, nous ne distinguerons pas les termes « échangeable » et « infiniment échangeable ». On parlera d'échangeabilité indépendamment de la taille de la séquence.

L'hypothèse d'échangeabilité signifie que la loi jointe des variables ne dépend pas de leurs indices. C'est une hypothèse moins contraignante que celle de l'indépendance des variables aléatoires. En effet, de par la définition de l'échangeabilité, on voit bien que l'indépendance implique l'échangeabilité mais pas l'inverse. Nous allons l'illustrer dans l'exemple ci-après.

Exemple d'une séquence infiniment échangeable : l'urne de Pólya
Considérons une urne contenant b boules bleues et r boules rouges. On tire au hasard une boule dans l'urne et on la replace avec n nouvelles boules de la même couleur.

Supposons que ce tirage soit répété infiniment. Soit X_i la variable aléatoire telle que $X_i = 1$ si le $i^{\text{ème}}$ tirage de l'urne produit une boule bleue, 0 sinon. Alors :

$$\begin{aligned}\mathbb{P}(1,1,0,1) &= \frac{b}{(b+r)}\frac{b+n}{(b+r+n)}\frac{r}{(b+r+2n)}\frac{b+2n}{(b+r+3n)} \\ &= \frac{b}{(b+r)}\frac{r}{(b+r+2n)}\frac{b+n}{(b+r+n)}\frac{b+2n}{(b+r+3n)} \\ &= \mathbb{P}(1,0,1,1).\end{aligned}$$

Cette urne produit des variables échangeables mais elles ne sont pas indépendantes.

L'échangeabilité est une notion importante dans l'échantillonnage des DPM comme on le verra dans le chapitre 3. Elle permet d'ignorer l'ordre des données et de traiter chacune comme si c'était la dernière.

Un résultat important lié à l'hypothèse d'échangeabilité est le théorème de De Finetti. Il stipule qu'une séquence infinie de variables aléatoires est échangeable si et seulement la loi de probabilité jointe a une représentation sous forme de mélange.

Théorème 2. *De Finetti*
Soit $\boldsymbol{\theta}^\infty = (\boldsymbol{\theta}_n)_{n\geq 1}$ une séquence de variables aléatoires définies sur l'espace mesuré $(\Omega, \mathcal{F}, \mathbb{P})$ et telle que chaque $\boldsymbol{\theta}_i$ est à valeurs dans $(\Theta, \mathcal{B})$ où Θ est un espace métrique complet. Alors $\boldsymbol{\theta}^\infty$ est échangeable si et seulement s'il existe une mesure de probabilité Q sur $\mathcal{M}(\Theta)$, l'ensemble de toutes les mesures de probabilité sur Θ, telle que pour tout $n \geq 1$ et $A = A_1 \times \cdots \times A_n \times \Theta^\infty$

$$\mathbb{P}(\boldsymbol{\theta}^\infty \in A) = \int_{\mathcal{M}(\Theta)} \prod_{i=1}^{n} H(\boldsymbol{\theta}_i \in A_i) Q(\mathrm{d}H).$$

avec $A_i \in \mathcal{B}$ et $\Theta^\infty = \Theta \times \Theta \times \cdots$.
La mesure Q est appelée mesure de De Finetti de la séquence $\boldsymbol{\theta}^\infty$.

L'échangeabilité peut être formulée en termes de variables conditionnellement *iid*. D'après le théorème, $\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \ldots$ est une séquence échangeable si et seulement s'il existe une mesure de probabilité aléatoire H qui les rend conditionnellement *iid* :

$$\begin{aligned} H &\sim Q, \\ \boldsymbol{\theta}_i | H &\overset{iid}{\sim} H \text{ pour } i \geq 1. \end{aligned}$$

Si on suppose que les données sont échangeables (ici les $\boldsymbol{\theta}_i$), l'existence d'une distribution Q sur les paramètres du modèle (la loi H) n'est pas une hypothèse de modélisation mais une conséquence mathématique. Dans ce cas, le théorème de De Finetti est une justification à la modélisation bayésienne.

2.1.3 L'urne de Blackwell-MacQueen

Blackwell et MacQueen [BM73] décrirent une séquence d'échantillons générés par un DP en utilisant une extension du schéma de l'urne de Pólya afin de permettre un continuum de couleurs. Soient G_0 une distribution sur les couleurs et α un réel positif.

Considérons une urne ne contenant pas de balles au départ. Pour la première balle, on tire une couleur $\boldsymbol{\theta}_1 \sim G_0$, on peint la balle suivant cette couleur tirée et on la met dans l'urne. Après n tirages, soit on tire une nouvelle couleur $\boldsymbol{\theta}_{n+1} \sim G_0$ avec une probabilité $\frac{\alpha}{\alpha+n}$, on peint une balle avec la couleur tirée et on la met dans l'urne, soit on tire au hasard une balle dans l'urne (suivant la distribution empirique) avec une probabilité $\frac{n}{\alpha+n}$ et on remet les deux balles dans l'urne.

Ce schéma s'écrit :

$$\begin{aligned} \boldsymbol{\theta}_1 &\sim G_0, \\ \boldsymbol{\theta}_{n+1}|\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_n &\sim H_n := \frac{1}{\alpha+n}\left(\alpha G_0 + \sum_{i=1}^{n} \delta_{\boldsymbol{\theta}_i}\right). \end{aligned} \tag{2.7}$$

Blackwell et MacQueen ont montré que la séquence d'échantillons générée de cette façon converge vers un DP lorsque $n \to +\infty$.

Théorème 3. *[BM73]*
Soit une séquence de variables aléatoires $\{\boldsymbol{\theta}_i\}_{i=1}^n$ construite en utilisant le schéma 2.7. Alors :

1. *H_n converge presque-sûrement quand $n \to +\infty$ vers une mesure aléatoire discrète H,*
2. *$H \sim DP(\alpha, G_0)$,*
3. *les variables aléatoires $\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \ldots$ sont iid H.*

L'urne de Blackwell-MacQueen illustre la propriété de clustering du DP. Soient $\boldsymbol{\theta}_1^*, \ldots, \boldsymbol{\theta}_{K_n}^*$ les valeurs uniques de $\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_n$ et n_k le nombre de fois où $\boldsymbol{\theta}_k^*$ est répété. On peut donc réécrire la deuxième équation de (2.7) de la façon suivante :

$$\boldsymbol{\theta}_{n+1}|\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_n \sim \frac{1}{\alpha+n}\left(\alpha G_0 + \sum_{k=1}^{K_n} n_k \delta_{\boldsymbol{\theta}_k^*}\right). \tag{2.8}$$

On note que plus une couleur a déjà été tirée, plus elle a de chances d'être tirée à nouveau (probabilité proportionnelle à n_k).

L'avantage de la représentation de l'urne de Blackwell-MacQueen est qu'elle permet de générer des échantillons suivant H (où $H \sim \mathrm{DP}(\alpha, G_0)$) sans avoir à construire explicitement la mesure elle-même. Cette propriété est à la base des premières méthodes d'échantillonnage des modèles de mélange de processus de Dirichlet comme on le verra dans le chapitre d'après. De plus, cette séquence d'échantillons est *échangeable*.

2.1.4 Le processus du restaurant chinois

Le processus du restaurant chinois (CRP) est la représentation la plus naturelle du DP. C'est une distribution sur des partitions [Ald85], [Pit02]. On appelle partition un groupement de données indépendant des labels des clusters où elles sont affectées. Le CRP est un processus séquentiel dont le nom donné par Pitman et Dubins provient de l'allégorie suivante : imaginons un restaurant ayant un nombre infini de tables, chaque table pouvant accueillir une infinité de personnes. Au départ, le restaurant est vide et les

clients arrivent un à un. Le premier client x_1 entre et s'installe à la table 1. Le deuxième client entrant peut décider soit de s'installer avec le premier, soit de s'installer à une autre table auquel cas celle-ci est numérotée 2. Pour $n > 1$, supposons que n clients $x_1, \ldots, x_n$ sont déjà entrés dans le restaurant et se sont installés suivant une certaine configuration occupant K_n tables. Alors le client x_{n+1} s'installe soit à une table k déjà occupée avec une probabilité proportionnelle au nombre de clients n_k qui y sont déjà installés, soit s'assoit à une nouvelle table avec une probabilité proportionnelle à α. Désignant par c_i le numéro de la table où le client s'assoit, on a alors la distribution prédictive suivante sur les labels c_i :

$$\mathbb{P}(c_{n+1} = j | c_1, \ldots, c_n) = \frac{\alpha}{\alpha + n}\delta_{K_n+1}(j) + \sum_{k=1}^{K_n} \frac{n_k}{\alpha + n}\delta_k(j). \tag{2.9}$$

Si on identifie les clients par des entiers de 1 à n et les tables par les clusters, le CRP définit une distribution *a priori* sur la partition des données (façon dont les clients se sont assis) et le nombre de clusters (nombre de tables occupées). Si on ignore les valeurs des labels c_i et que l'on regarde seulement la partition résultante, les clients sont échangeables c'est-à-dire que leur ordre d'arrivée dans le restaurant n'importe pas.

Supposons maintenant qu'indépendamment de la séquence des x_i, on associe à chaque table occupée un repas tiré à partir d'un menu comportant un nombre infini de repas. Le premier client à s'asseoir à une table k choisit le repas $\boldsymbol{\theta}_k^*$ qui sera partagé par tous les autres qui s'y assiéront. Soit G_0 la distribution associée au menu (i.e $\boldsymbol{\theta}_k^* \sim G_0$ pour tout k). La distribution des repas est alors :

$$\mathbb{P}(\boldsymbol{\theta}_{n+1} \in \cdot | \boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_n) = \frac{\alpha}{\alpha + n}G_0(\cdot) + \sum_{k=1}^{K_n} \frac{n_k}{\alpha + n}\delta_{\boldsymbol{\theta}_k^*}(\cdot). \tag{2.10}$$

On retrouve la même loi que dans l'urne de Blackwell-MacQueen (équation (2.7)). La séquence $\{\boldsymbol{\theta}_i\}_{i=1}^n$ est donc aussi échangeable. Le CRP est illustré à la Figure 2.1.

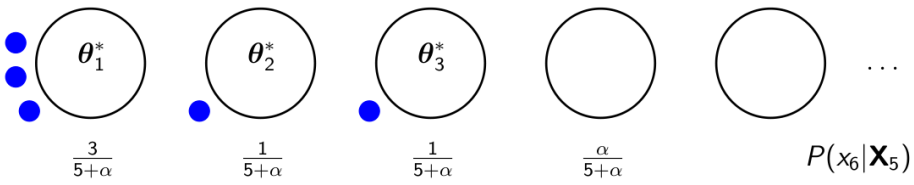


FIGURE 2.1 – Illustration du processus du restaurant chinois. Les tables sont représentées par les cercles et les clients par les pastilles bleues. Après que 5 clients se soient déjà installés, le 6ème client s'installe soit à la table 1 avec la probabilité $\frac{3}{5+\alpha}$, soit aux tables 2 et 3 avec la probabilité $\frac{1}{5+\alpha}$, soit s'asseoit à une nouvelle table avec la probabilité $\frac{\alpha}{5+\alpha}$.

2.1.5 La construction « stick-breaking »

La représentation à bâtons rompus dite « stick-breaking » est une définition constructive du processus de Dirichlet ([ST82], [Set94]). Elle est basée sur le modèle d'allocation des résidus (noté RAM pour Residual Allocation Model) appelé aussi stick-breaking.

Théorème 4. *[Set94]*
On considère une suite de variables aléatoires $\mathbf{v} = (v_1, v_2, \ldots)$ *telle que chaque* $v_i \sim Beta(1, \alpha)$. *On construit* $\mathbf{w} = (w_1, w_2, \ldots)$ *tels que* $w_1 = v_1$ *et* $\forall k \geq 2$, $w_k = v_k \prod_{j=1}^{k-1} (1 - v_j)$. *Enfin, on considère* $\mathbf{\Theta}^* = (\boldsymbol{\theta}_1^*, \boldsymbol{\theta}_2^*, \ldots) \overset{iid}{\sim} G_0$ *et indépendants des* v_j. *Alors la série,*

$$H = \sum_{k=1}^{\infty} w_k \, \delta_{\boldsymbol{\theta}_k^*} \tag{2.11}$$

converge presque-sûrement vers une mesure de probabilité aléatoire générée selon un $DP(\alpha, G_0)$.

La construction stick-breaking des poids w_k peut être comprise de la façon métaphorique suivante : considérons un bâton de longueur unité qui peut être cassé infiniment. A chaque itération k, on casse le bâton au point v_k et on affecte à w_k la longueur du morceau qu'on a cassé. La longueur des morceaux ainsi coupés correspond aux poids w_k dans dans l'équation 2.11. Cette construction est illustrée dans la Figure 2.2.

La distribution stick-breaking des poids est notée $\mathbf{w} \sim \text{GEM}(\alpha)$, dont le nom donné par Ewens ([Ewe72], [ET98]) représente Griffiths, Engen et McCloskey. Ces deux derniers ([McC65] et [Eng78]) ont introduit cette distribution dans le contexte de l'écologie tandis que Griffiths [Gri88] fut le premier à noter son importance en génétique.

La représentation stick-breaking est appelée définition *constructive* du processus de Dirichlet car elle représente explicitement la mesure aléatoire H générée suivant ce processus, contrairement à la représentation de l'urne de Blackwell-MacQueen et le processus du restaurant chinois qui marginalisent la distribution H et génèrent seulement des échantillons suivant elle.

La construction stick-breaking des poids a permis de faire la connexion entre le processus de Dirichlet et d'autres processus. En effet, les *a priori* stick-breaking sont assez généraux et incluent non seulement le processus de Dirichlet, mais aussi le processus de Dirichlet multinomial [MS95], le processus de Pitman-Yor [PY97], le processus Beta à deux paramètres [IZ00] etc. Des exemples supplémentaires peuvent être trouvés dans [IJ01]. La représentation stick-breaking est aussi le point de départ de l'extension du processus de Dirichlet à des mesures permettant d'introduire une certaine dépendance dans une collection de distributions comme le processus de Dirichlet dépendant (DDP) [Mac99], le processus de Dirichlet hiérarchique (HDP) [TJBB06], le processus de Dirichlet imbriqué (NDP) [RDG08], le π-DDP [GS06] etc. Enfin comme on le verra dans le chapitre 3, elle est à la source des premières méthodes MCMC qui représentent explicitement la mesure aléatoire en la tronquant. C'est-à-dire que la loi *a priori* $\sum_{k=1}^{\infty} w_k \delta_{\boldsymbol{\theta}_k^*}$ est remplacée par $\sum_{k=1}^{N} w_k \delta_{\boldsymbol{\theta}_k^*}$ pour une valeur de N convenablement choisie, par exemple de telle sorte que la variation totale entre la mesure infinie et la tronquée soit bornée par un certain ϵ ([IZ02], [MT98]).

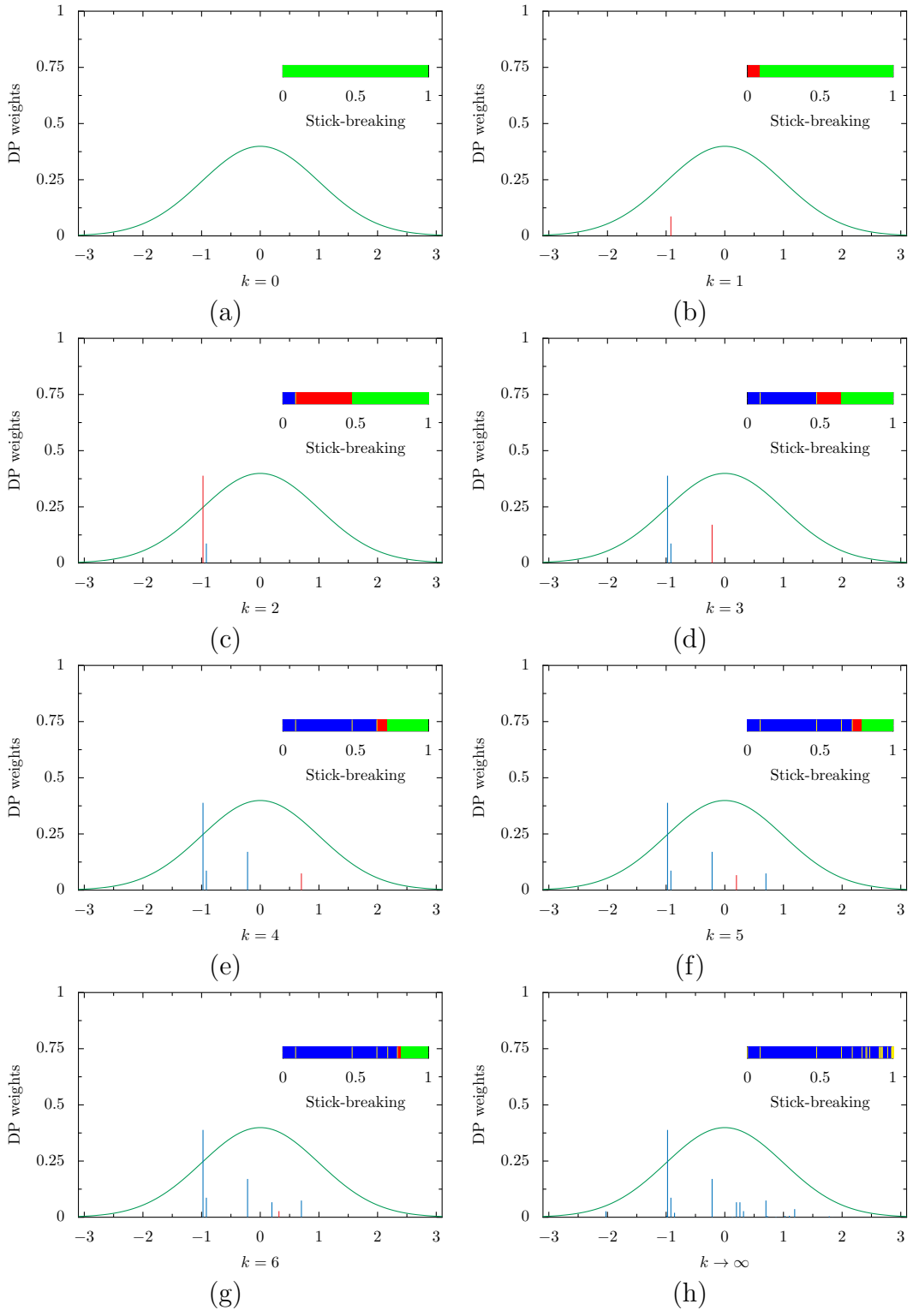


FIGURE 2.2 – Séquence de construction stick-breaking d'un DP($\alpha = 3, G_0 = \mathcal{N}(0,1)$). (a) itération $k = 0$, $G_0 = \mathcal{N}(0,1)$, (b) itération $k = 1$, on tire une localisation suivant G_0 et le poids associé suivant le stick-breaking (Dirac représenté en rouge), (c) idem pour l'itération $k = 2$, (d) itération $k = 3$, (e) itération $k = 4$, (f) itération $k = 5$, (g) itération $k = 6$, (h) à l'itération $k \to \infty$, on obtient une réalisation d'un DP($3, \mathcal{N}(0,1)$).

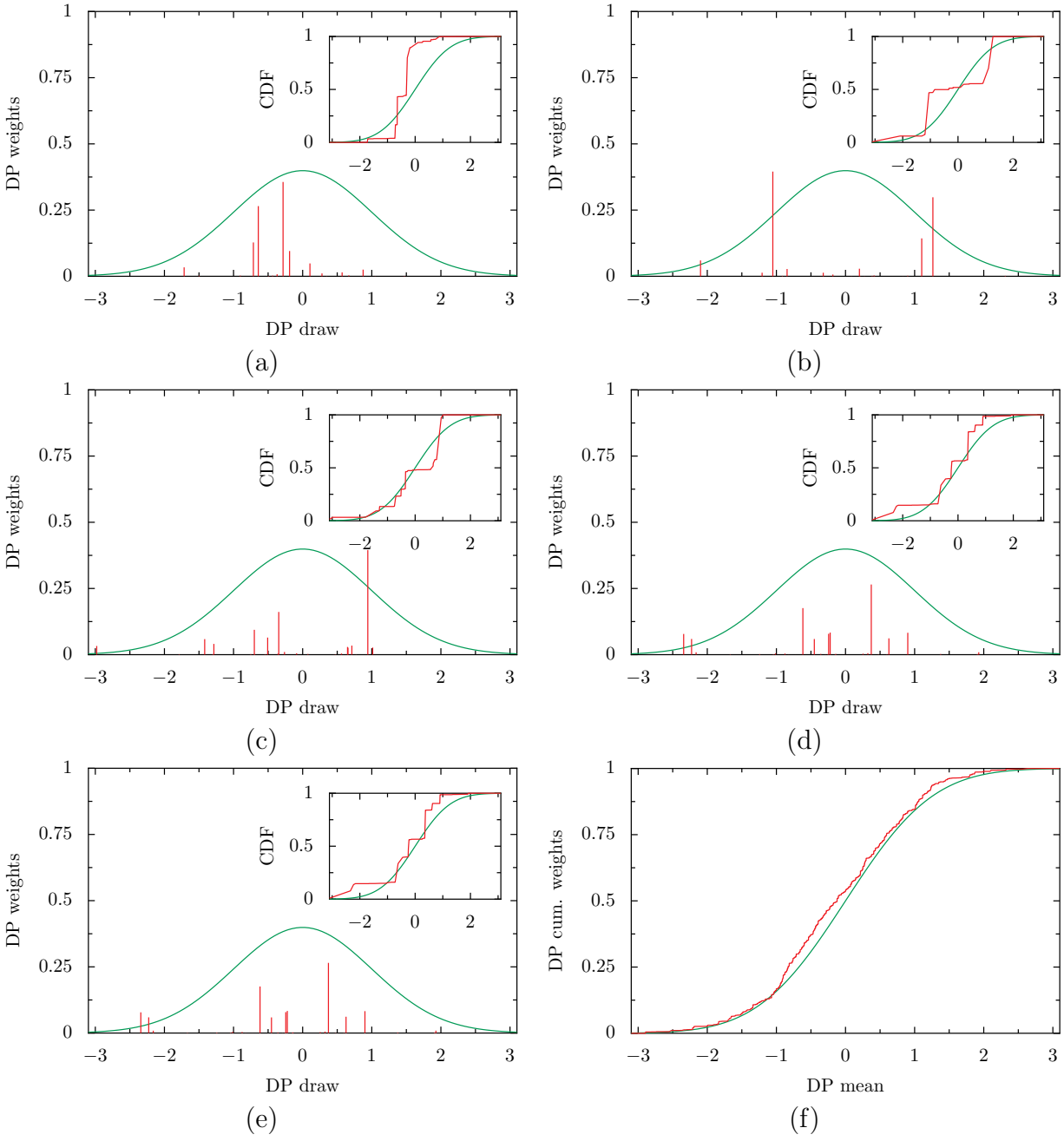


FIGURE 2.3 – (a) à (e) : exemples de réalisations d'un $\mathrm{DP}(3, \mathcal{N}(0,1))$ et des cdf associées. (f) : cdf pour la moyenne des tirages, on voit bien qu'elle tend vers celle de G_0.

2.1.6 Mélange de processus de Dirichlet

La définition du processus de Dirichlet avec un paramètre, la mesure $\nu = \alpha G_0(\cdot)$ (cf. Remarque 1), nécessite de spécifier un seul paramètre représentant à la fois l'espérance *a priori* G_0 et la force accordée à la croyance *a priori* α. L'idée du MDP (Mixture of Dirichlet Processes) est de permettre à ce paramètre ν d'être randomisé afin de permettre plus de flexibilité dans le modèle.

Définition 3. *[Ant74]*
Soit $(\Theta, \mathcal{B})$ *un espace mesurable. On considère un espace d'indices* U *muni d'une tribu*

$\mathcal{A}$, noté $(U, \mathcal{A})$, et μ une mesure de probabilité sur $(U, \mathcal{A})$. Soit $\nu(\cdot, \cdot)$ une mesure de transition[1] *de $U \times \mathcal{B} \to [0, +\infty[$ telle que :*

- *pour tout $u \in U, \nu(u, \cdot)$ est une mesure finie, non-nulle et non-négative sur $(\Theta, \mathcal{B})$,*
- *pour tout $B \in \mathcal{B}, \nu(\cdot, B)$ est mesurable sur $(U, \mathcal{A})$.*

On dit que H est un mélange de processus de Dirichlet sur $(\Theta, \mathcal{B})$, de distribution mélangeante μ, d'espace d'indices $(U, \mathcal{A})$ et de mesure de transition ν si pour toute partition mesurable $B_1, \ldots, B_K$ de Θ,

$$(H(B_1), \ldots, H(B_K)) \sim \int_U DP(\nu(u, A_1), \nu(u, A_2), \ldots, \nu(u, A_K)) \, \mathrm{d}\mu(u).$$

On note :

$$H \sim \int_U DP(\nu(u, \cdot)) \, \mathrm{d}\mu(u). \tag{2.12}$$

Le MDP revient à considérer le modèle hiérarchique suivant :

$$\begin{aligned} u &\sim \mu \\ H|u &\sim DP(\nu(u, \cdot)) \equiv \mathrm{DP}(\nu_u). \end{aligned}$$

En pratique, on pourrait choisir une famille paramétrique de mesures ν_u indexée par u et mettre une loi *a priori* sur l'hyperparamètre u.

Tout comme le DP, les mesures générées par un MDP sont discrètes. Le MDP bénéficie aussi d'une propriété de conjugaison : si $H \sim$ MDP, $\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_n | H \overset{iid}{\sim} H$ alors $H|\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_n$ suit aussi un MDP. La raison est que si H suit un MDP, alors conditionnellement à u, la distribution de H est un DP et donc la propriété de conjugaison est préservée : $H|u, \boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_n \sim \mathrm{DP}(\nu_u + \sum_{i=1}^n \delta_{\boldsymbol{\theta}_i})$. La distribution de $H|\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_n$ s'obtient en intégrant $\mathrm{DP}(\nu_u + \sum_{i=1}^n \delta_{\boldsymbol{\theta}_i})$ par rapport à la distribution conditionnelle de $u|\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_n$ et on a :

$$H|\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_n \sim \int \mathrm{DP}(\nu_u + \sum_{i=1}^n \delta_{\boldsymbol{\theta}_i}) \mathrm{d}\mu(u|\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_n).$$

Le DP est un cas particulier de MDP ([Ant74, Corollaire 3.1]). En effet, si H est générée suivant un MDP sur $(\Theta, \mathcal{B})$ dont l'espace d'indices est aussi $(\Theta, \mathcal{B})$ et si la mesure de transition est $\nu_u = \nu + \delta_u$ et la distribution mélangeante $P(\cdot) = \frac{\nu(\cdot)}{\nu(\Theta)}$, alors $H \sim DP(\nu(\cdot))$, ie

$$\int_\Theta \mathrm{DP}(\nu + \delta_u) \frac{\nu(\mathrm{d}u)}{\nu(\Theta)} = \mathrm{DP}(\nu).$$

Le mélange de processus de Dirichlet (MDP) est un mélange paramétrique de lois *a priori* non paramétriques : la distribution mélangeante μ est paramétrique et la distribution mélangée est un DP (équation 2.12). Si l'on considère un mélange non paramétrique de familles paramétriques, c'est-à-dire tel que la distribution mélangeante H est un DP et la densité mélangée f est paramétrique, on obtient les mélanges par processus de Dirichlet (DPM).

1. ν généralise la notion de probabilité de transition car $\nu(j, \Theta)$ n'est pas forcément égale à 1.

2.1.7 Le mélange par processus de Dirichlet

Comme nous l'avons déjà souligné, les mesures générées par un DP et par un MDP sont discrètes et, en conséquence, n'admettent pas de densité. Elles ne peuvent donc pas être utilisées comme loi *a priori* sur des densités. Les mélanges par processus de Dirichlet (notés DPM pour Dirichlet Process Mixtures) permettent de remédier à ce problème ([Fer83], [Lo84]).

Estimation de densité par DPM

Soient $\mathcal{X}$ et Θ deux espaces euclidiens munis respectivement des tribus $\mathcal{A}$ et $\mathcal{B}$ engendrées par les ouverts. Dans nos applications, ce seront typiquement des sous-ensembles de $\mathbb{R}^d$ muni de la tribu borélienne. Considérons $\boldsymbol{x}_1, \cdots, \boldsymbol{x}_n$ ($\boldsymbol{x}_i \in \mathcal{X}$), des réalisations d'une séquence de variables aléatoires indépendantes et identiquement distribuées suivant une distribution commune mais inconnue P, admettant p comme densité de probabilité. En estimation de densité, on cherche à inférer sur p. L'approche usuelle consiste à la représenter sous la forme d'un mélange de distributions,

$$p(\boldsymbol{x}) = \int_{\Theta} f(\boldsymbol{x}|\boldsymbol{\theta}) H(\mathrm{d}\boldsymbol{\theta})$$

où H est une distribution sur Θ, appelée distribution mélangeante et $\{f(\boldsymbol{x}|\boldsymbol{\theta}) : \boldsymbol{\theta} \in \Theta\}$ une famille de fonctions de densités, $f(\cdot|\boldsymbol{\theta})$ est appelée densité mélangée. La régularisation du problème inverse peut se faire dans un cadre bayésien. Une approche paramétrique à ce problème reviendrait à mettre une loi *a priori* sur les paramètres $\boldsymbol{\theta}$ de la densité p. Par exemple, dans le cas où la distribution P est supposée gaussienne, la loi *a priori* peut porter sur la moyenne et la matrice de covariance. Dans l'approche non paramétrique, on cherche à estimer la densité avec le moins d'hypothèses possibles sur sa forme. Les mélanges par processus de Dirichlet permettent de faire porter l'*a priori* directement sur p. De façon équivalente, cela revient à mettre la loi *a priori* sur la distribution mélangeante H en utilisant les processus de Dirichlet. Le modèle s'écrit ainsi :

$$\begin{aligned} p(\boldsymbol{x}) &= \int_{\Theta} f(\boldsymbol{x}|\boldsymbol{\theta}) H(\mathrm{d}\boldsymbol{\theta}), \\ H|\alpha, G_0 &\sim \mathrm{DP}(\alpha, G_0) \equiv \mathcal{P}. \end{aligned} \tag{2.13}$$

Ainsi, pour obtenir la loi *a priori* sur la densité p, l'idée des mélanges par processus de Dirichlet est de convoluer la mesure discrète H générée par un DP avec un noyau paramétrique $F(\cdot|\boldsymbol{\theta})$, de densité $f(\cdot|\boldsymbol{\theta})$ [2]. Puisque la distribution H est aléatoire, la densité p l'est aussi. Par exemple, si les $x_i \in \mathcal{X} = \mathbb{R}$, en considérant un noyau Gaussien paramétré par $\boldsymbol{\theta} = (\mu, \sigma) \in \Theta = \mathbb{R} \times \mathbb{R}^+$, $f(x|\mu, \sigma) = \frac{1}{\sigma}\phi(\frac{x-\mu}{\sigma})$ où ϕ est la densité de la loi normale standard $\mathcal{N}(\cdot|0, 1)$, la loi *a priori* $p(x)$ est alors un mélange de lois normales.

On peut réécrire le modèle 2.13 de la façon hiérarchique suivante

$$\begin{aligned} \boldsymbol{x}_i|\boldsymbol{\theta}_i &\sim F(\boldsymbol{x}_i|\boldsymbol{\theta}_i) \qquad \text{pour } i = 1, \ldots, n \\ \boldsymbol{\theta}_i|H &\sim H \qquad\qquad\quad \text{pour } i = 1, \ldots, n \\ H|\alpha, G_0 &\sim \mathrm{DP}(\alpha, G_0). \end{aligned} \tag{2.14}$$

2. f est à valeurs non-négatives et doit vérifier : 1) pour tout $\boldsymbol{\theta} \in \Theta$, $\int_{\mathcal{X}} f(\boldsymbol{x}|\boldsymbol{\theta})\mathrm{d}\boldsymbol{x} = 1$ et 2) pour tout $\boldsymbol{x} \in \mathcal{X}$, $\int_{\Theta} f(\boldsymbol{x}|\boldsymbol{\theta}) G_0(\mathrm{d}\boldsymbol{\theta}) < \infty$.

On remarque que le modèle 2.13 s'obtient par marginalisation du 2.14 par rapport à $\boldsymbol{\theta}_i$.

Pour l'inférence on a le résultat suivant qui dit que si le premier niveau du modèle hiérarchique 2.14 est modélisé par une distribution paramétrique, et le second niveau par un *a priori* DP, alors la distribution *a posteriori* du processus est un MDP.

Théorème 5. *[Ant74, Corollaire 3.1]*
Soient $\nu(\cdot) \equiv \alpha G_0(\cdot)$ et $H \sim DP(\nu)$. Soient $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$, n échantillons tirés d'une distribution P ayant pour densité p avec $p(\boldsymbol{x}) = \int f(\boldsymbol{x}|\boldsymbol{\theta})\mathrm{d}H(\boldsymbol{\theta})$. Alors, la loi a posteriori de la distribution mélangeante H est un MDP :

$$H|\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n \sim \int DP\left(\nu + \sum_{i=1}^{n} \delta_{\boldsymbol{\theta}_i}\right) \mathrm{d}P_{\boldsymbol{x}}(\boldsymbol{\theta})$$

où $P_{\boldsymbol{x}}(\boldsymbol{\theta})$ désigne la distribution a posteriori de $\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_n$ sachant $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$.

Démonstration. Soient $\mathcal{P}_{\boldsymbol{x}}(H) \equiv \mathcal{P}(H|\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n)$ la distribution *a posteriori* de $H|\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ et $P_{\boldsymbol{x}}(\boldsymbol{\theta}) \equiv P(\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_n|\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n)$, la distribution *a posteriori* de $\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_n|\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$. La distribution conditionnelle de $H|\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ est obtenue en intégrant la distribution conditionnelle de $H|\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_n, \boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ par rapport à la distribution conditionnelle de $\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_n|\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$:

$$\mathcal{P}_{\boldsymbol{x}}(H) = \int \mathcal{P}(H|\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_n, \boldsymbol{x}_1, \ldots, \boldsymbol{x}_n) P_{\boldsymbol{x}}(\mathrm{d}\boldsymbol{\theta}).$$

Puisque, sachant $\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_n$, H et $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ sont conditionnellement indépendants, on a $\mathcal{P}(H|\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_n, \boldsymbol{x}_1, \ldots, \boldsymbol{x}_n) = \mathcal{P}(H|\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_n) \equiv \mathrm{DP}(\nu + \sum_{i=1}^{n} \delta_{\boldsymbol{\theta}_i})$, d'où le résultat. □

Le théorème 5 donne la loi *a posteriori* de la distribution mélangeante du DPM sous la forme d'un MDP. De ce fait, les DPM sont souvent appelés mélanges de processus de Dirichlet dans la littérature, induisant une confusion entre les DPM et les MDP.

Distribution prédictive

L'inférence bayésienne permet de calculer la distribution prédictive *a posteriori* d'une quantité non-observée dont la distribution dépend des paramètres du modèle,

$$p(\boldsymbol{x}_{n+1}|\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n) = \int f(\boldsymbol{x}_{n+1}|\boldsymbol{\theta}_{n+1}) P(\mathrm{d}\boldsymbol{\theta}_{n+1}|\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n)$$

$$\text{où } P(\mathrm{d}\boldsymbol{\theta}_{n+1}|\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n) = \int P(\mathrm{d}\boldsymbol{\theta}_{n+1}|\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_n) P(\mathrm{d}\boldsymbol{\theta}_1, \ldots, \mathrm{d}\boldsymbol{\theta}_n|\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n)$$

$$\text{avec } P(\mathrm{d}\boldsymbol{\theta}_{n+1}|\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_n) = \frac{\alpha}{\alpha + n} G_0(\mathrm{d}\boldsymbol{\theta}_{n+1}) + \frac{1}{\alpha + n} \sum_{i=1}^{n} \delta_{\boldsymbol{\theta}_i}(\mathrm{d}\boldsymbol{\theta}_{n+1}).$$

La distribution prédictive *a posteriori* est aussi l'estimée bayésienne de p sous la loi *a posteriori* si la fonction de perte est quadratique.

Estimée bayésienne

Ayant observé $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$, on cherche une estimée pour p. Pour cela, on choisit une quantité $\hat{p}$ qui minimise une certaine fonction de perte. La quantité obtenue est appelée estimée bayésienne de p. La fonction perte quadratique

$$L(p, \hat{p}) = \int (p(\boldsymbol{x}) - \hat{p}(\boldsymbol{x}))^2 \mathrm{d}\boldsymbol{x}$$

est minimisée par l'espérance de la loi *a posteriori* qui est donc l'estimée bayésienne de p.

L'estimée bayésienne de p sous la loi *a posteriori* , notée $\hat{p}(\boldsymbol{x})$ ou $p_n(\boldsymbol{x})$, est donnée par :

$$\begin{aligned}
\hat{p}(\boldsymbol{x}) &= \mathbb{E}\left[p(\boldsymbol{x})|\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n\right] \\
&= \mathbb{E}\left[\int_{\Theta^n}\int_{\Theta} f(\boldsymbol{x}|\boldsymbol{\theta})[H(\mathrm{d}\boldsymbol{\theta})|\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_n)]P(\mathrm{d}\boldsymbol{\theta}_1, \ldots, \mathrm{d}\boldsymbol{\theta}_n|\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n)\right] \\
&= \int_{\Theta^n}\int_{\Theta} f(\boldsymbol{x}|\boldsymbol{\theta})\mathbb{E}\left([H(\mathrm{d}\boldsymbol{\theta})|\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_n)]\right) P(\mathrm{d}\boldsymbol{\theta}_1, \ldots, \mathrm{d}\boldsymbol{\theta}_n|\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n) \text{ par Fubini-Tonelli} \\
&= \int_{\Theta^n}\int_{\Theta} f(\boldsymbol{x}|\boldsymbol{\theta})\frac{\alpha G_0(\mathrm{d}\boldsymbol{\theta}) + \sum_{i=1}^n \delta_{\boldsymbol{\theta}_i}(\mathrm{d}\boldsymbol{\theta})}{\alpha + n} P(\mathrm{d}\boldsymbol{\theta}_1, \ldots, \mathrm{d}\boldsymbol{\theta}_n|\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n) \\
&= \frac{\alpha}{\alpha + n}\int_{\Theta} f(\boldsymbol{x}|\boldsymbol{\theta})G_0(\mathrm{d}\boldsymbol{\theta}) + \frac{1}{\alpha + n}\int_{\Theta^n}\sum_{i=1}^n f(\boldsymbol{x}|\boldsymbol{\theta}_i)P(\mathrm{d}\boldsymbol{\theta}_1, \ldots, \mathrm{d}\boldsymbol{\theta}_n|\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n) \\
&= \frac{\alpha}{\alpha + n}p_0(\boldsymbol{x}) + \frac{n}{\alpha + n}\hat{p}_n(\boldsymbol{x})
\end{aligned}$$

où

$$\hat{p}_n(\boldsymbol{x}) = \frac{1}{n}\sum_{i=1}^n \int_{\Theta^n} f(\boldsymbol{x}|\boldsymbol{\theta}_i)P(\mathrm{d}\boldsymbol{\theta}_1, \ldots, \mathrm{d}\boldsymbol{\theta}_n|\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n).$$

et

$$p_0(\boldsymbol{x}) = \int_{\mathcal{M}(\Theta)} p(\boldsymbol{x})\mathcal{P}(\mathrm{d}H) = \int_{\Theta} f(\boldsymbol{x}|\boldsymbol{\theta})G_0(\mathrm{d}\boldsymbol{\theta}).$$

$p_0(\boldsymbol{x})$ est la densité marginale de $\boldsymbol{x}$ et c'est aussi l'espérance de $p(\boldsymbol{x})$ sous la loi *a priori* . C'est donc l'estimée bayésienne sous le risque quadratique si l'on n'a pas d'observations. L'estimée bayésienne $\hat{p}(\boldsymbol{x})$ est donc composée d'un terme dû à l'*a priori* et d'un autre dû aux observations. Par des manipulations algébriques, Lo [Lo84] fournit une expression analytique de $\hat{p}(\boldsymbol{x})$. Cette formule est cependant difficile à utiliser en pratique car le nombre de termes augmente très rapidement avec la taille de l'échantillon. Cela est en grande partie dû à l'extrême complexité de $P(\mathrm{d}\boldsymbol{\theta}_1, \ldots, \mathrm{d}\boldsymbol{\theta}_n|\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n)$. En effet, en utilisant la règle de Bayes on a :

$$\begin{aligned}
P(\mathrm{d}\boldsymbol{\theta}_1, \ldots, \mathrm{d}\boldsymbol{\theta}_n|\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n) &= \frac{\prod_{i=1}^n f(\boldsymbol{x}_i|\boldsymbol{\theta}_i)\prod_{i=1}^n(\frac{\alpha G_0 + \sum_{j=1}^{i-1}\delta_{\boldsymbol{\theta}_j}}{\alpha+i-1})(\mathrm{d}\boldsymbol{\theta}_i)}{\int\prod_{i=1}^n f(\boldsymbol{x}_i|\boldsymbol{\theta}_i)\prod_{i=1}^n(\frac{\alpha G_0(\boldsymbol{\theta}_i) + \sum_{j=1}^{i-1}\delta_{\boldsymbol{\theta}_j}(\boldsymbol{\theta}_i)}{\alpha+i-1})\mathrm{d}\boldsymbol{\theta}_1, \ldots, \mathrm{d}\boldsymbol{\theta}_n} \\
&= \frac{\prod_{i=1}^n f(\boldsymbol{x}_i|\boldsymbol{\theta}_i)\prod_{i=1}^n(\alpha G_0 + \sum_{j=1}^{i-1}\delta_{\boldsymbol{\theta}_j})(\mathrm{d}\boldsymbol{\theta}_i)}{\int\prod_{i=1}^n f(\boldsymbol{x}_i|\boldsymbol{\theta}_i)\prod_{i=1}^n(\alpha G_0(\boldsymbol{\theta}_i) + \sum_{j=1}^{i-1}\delta_{\boldsymbol{\theta}_j}(\boldsymbol{\theta}_i))\mathrm{d}\boldsymbol{\theta}_1, \ldots, \mathrm{d}\boldsymbol{\theta}_n}.
\end{aligned}$$

Le produit $\prod_{i=1}^n(\alpha G_0(\mathrm{d}\boldsymbol{\theta}_i) + \sum_{j=1}^{i-1}\delta_{\boldsymbol{\theta}_j})(\mathrm{d}\boldsymbol{\theta}_i)$ comporte $n!$ termes. Le nombre de termes est réduit par l'effet de clustering du DP mais reste toutefois assez prohibitif, même pour un échantillon de taille modérée.

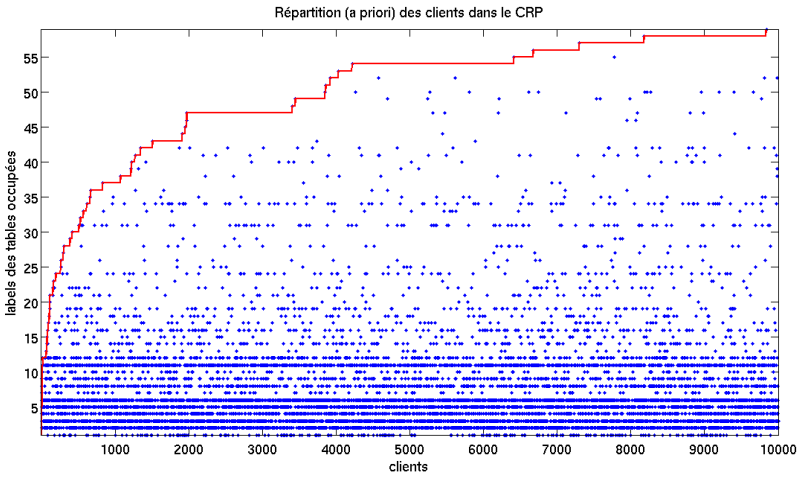


FIGURE 2.4 – Phénomène de renforcement du DP illustré avec le restaurant chinois. Pour $n = 10.000$ clients et $\alpha = 10$, seules 59 tables sont occupées.

Les méthodes MCMC que l'on verra dans le chapitre 3 permettent d'approximer la distribution prédictive.

Phénomène de renforcement statistique du DPM

Comme nous l'avons déjà souligné, la nature discrète de H implique que plusieurs $\boldsymbol{\theta}_i$ vont prendre la même valeur. Cela induit un clustering *a priori* des données $\boldsymbol{x}_i$ ayant la même valeur de paramètre. On appellera *cluster* k, $\{\boldsymbol{x}_i : \boldsymbol{\theta}_i = \boldsymbol{\theta}_k^*\}$. Il est alors évident que même si le nombre de composantes dans le DPM est potentiellement infini, le nombre de clusters non vides est fini et inférieur au nombre de données. Soit K_n le nombre de clusters non vides parmi n observations. On a vu que la probabilité *a priori* pour que $\boldsymbol{\theta}_i$ prenne une nouvelle valeur est $\frac{\alpha}{\alpha+i-1}$ (cf. schéma 2.7). L'espérance de K_n est alors :

$$\mathbb{E}[K_n|n] = \sum_{i=1}^{n} \frac{\alpha}{\alpha+i-1} = \alpha \sum_{i=1}^{n} \frac{1}{\alpha+i-1} \approx \alpha \log\left(1+\frac{n}{\alpha}\right).$$

L'apparition de nouveaux clusters dépend seulement de α et non de G_0. Il est alors clair que le paramètre α contrôle le nombre de clusters *a priori* de façon directe : plus il est grand, plus le nombre de composantes non vides l'est aussi. Toutefois, ce nombre n'explose pas puisque quand $n \to +\infty$ (il est en $O(\alpha \log n)$). Il augmente donc seulement de façon logarithmique avec le nombre d'observations comme on peut le noter dans la Figure 2.4. Cette croissance lente du nombre de clusters est lié au phénomène de renforcement statistique du processus de Dirichlet (le phénomène du "rich-gets-richer") qui induit un regroupement *a priori* des données dans de larges composantes et confère une propriété de *parcimonie* au modèle, d'où son utilisation en machine-learning.

Conclusion sur les DPM

Les modèles de mélange par processus de Dirichlet sont des outils de modélisation riches et flexibles, utilisés aussi bien pour l'estimation de densité que pour la classifi-

cation. Dans cette thèse, nous les utiliserons comme modèle d'estimation de densité, plus précisément pour l'estimation indirecte de densité en TEP. Dans la communauté machine-learning, la représentation 2.14 du DPM sert comme modèle de classification. Ici, la nature non paramétrique du processus de Dirichlet conduit à des modèles de mélange avec un nombre infini de composantes. Cela permet d'éviter l'étape du choix du modèle dans les modèles paramétriques, étape fastidieuse mais nécessaire afin de déterminer le nombre de composantes approprié. L'approche bayésienne non paramétrique est donc une bonne alternative au choix de modèle puisque le nombre de paramètres dans le modèle peut augmenter avec les données et la complexité de leur structure.

2.2 Le modèle d'échantillonnage d'espèces

Le modèle d'échantillonnage d'espèces peut être vu comme une généralisation du processus de Dirichlet (DP). Il est défini comme la mesure de probabilité aléatoire gouvernant une séquence échangeable d'échantillonnage d'espèces. Cette dernière est une généralisation de la séquence de l'urne de Pólya ([Pit95], [Pit96b], [HP98], [IJ03a], [Qui06], [LQMT08], [NQM08]).

2.2.1 Séquence d'échantillonnage d'espèces

On considère $(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \dots)$ une séquence de variables aléatoires définies sur $(\Omega, \mathcal{F}, \mathbb{P})$ et à valeurs dans $(\Theta, \mathcal{B})$. Supposons que $(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \dots)$ soit un tirage aléatoire à partir d'une population composée d'un très grand nombre d'espèces de divers types. Soit $\boldsymbol{\theta}_i$ l'espèce du $i^{\text{ème}}$ individu tiré. On suppose que chaque espèce différente a une certaine étiquette. On peut alors voir Θ comme un espace permettant d'étiqueter les espèces (spectre de couleurs, numéros, etc.). Soit M_j l'indice du premier individu observé appartenant à l'espèce j, i.e $M_1 := 1$ et pour tout $j > 1$, $M_j := \inf\{n : n > M_{j-1}, \boldsymbol{\theta}_n \notin \{\boldsymbol{\theta}_1, \dots, \boldsymbol{\theta}_{n-1}\}\}$ avec la convention $\inf \emptyset = \infty$. Soit $\boldsymbol{\theta}_j^* := \boldsymbol{\theta}_{M_j}$ la j$^{\text{ème}}$ espèce apparue. On définit par $\boldsymbol{n} = (n_1, n_2, \dots)$ où n_j est le nombre de fois où l'espèce $\boldsymbol{\theta}_j^*$ est apparue dans l'échantillon $\boldsymbol{\theta}_1, \dots, \boldsymbol{\theta}_n$:

$$n_j := \sum_{i=1}^{n} \mathbf{1}(\boldsymbol{\theta}_i = \boldsymbol{\theta}_j^*),$$

avec $\sum_j n_j = n$. Soit K_n le nombre d'espèces différentes apparues dans les n premières observations :

$$K_n := \max\{j : n_j > 0\}.$$

Définition 4. *Séquence d'échantillonnage d'espèces*
Soit G_0 une mesure de probabilité non-atomique sur $(\Theta, \mathcal{B})$. Une séquence échangeable $\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \dots$ est appelée séquence d'échantillonnage d'espèces si elle est générée via les distributions prédictives suivantes :

$$\boldsymbol{\theta}_1 \sim G_0,$$

$$\boldsymbol{\theta}_{n+1} | \boldsymbol{\theta}_1, \dots, \boldsymbol{\theta}_n \sim H_n := \sum_{j=1}^{K_n} p_j(\boldsymbol{n}) \delta_{\boldsymbol{\theta}_j^*} + p_{K_n+1}(\boldsymbol{n}) G_0 \tag{2.15}$$

où $\boldsymbol{\theta}_j^*, j = 1, \ldots, K_n$ *sont les valeurs distinctes parmi* $\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_n$ *dans leur ordre d'apparition et la séquence de fonctions de probabilité prédictives* $p_1, p_2, \ldots$ *est telle que*

$$p_j(\boldsymbol{n}) \geq 0 \text{ et } \sum_{j=1}^{K_n+1} p_j(\boldsymbol{n}) = 1 \text{ pour tout } \boldsymbol{n} \in \mathbf{N}^* = \bigcup_{k=1}^{\infty} \mathbb{N}^k,$$

avec $\mathbb{N}$ *désignant l'ensemble des entiers naturels et* $\mathbf{N}^*$ *l'ensemble des séquences d'entiers positifs.*

Remarques :

1) La séquence des $\boldsymbol{\theta}_i$ dans l'équation 2.15 est une séquence d'échantillonnage d'espèces que si elle est *échangeable*.

2) Les probabilités prédictives peuvent être comprises de la façon suivante : on affecte à l'espèce du premier individu tiré une étiquette aléatoire $\boldsymbol{\theta}_1 = \boldsymbol{\theta}_1^* \sim G_0$. Sachant les étiquettes $\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_n$ des n premiers individus tirés, on considère que le $(n+1)^{\text{ème}}$ individu sera de la $j^{\text{ème}}$ espèce déjà observée (auquel cas son étiquette sera $\boldsymbol{\theta}_j^*$) avec la probabilité $p_j(\boldsymbol{n})$, ou une nouvelle espèce (avec une étiquette tirée de G_0) avec la probabilité $p_{K_n+1}(\boldsymbol{n})$. Cela s'écrit

$$\begin{aligned} p_j(\boldsymbol{n}) &= \mathbb{P}(\boldsymbol{\theta}_{n+1} = \boldsymbol{\theta}_j^* | \boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_n, K_n) \text{ pour } j = 1, \ldots, K_n, \\ p_{K_n+1}(\boldsymbol{n}) &= \mathbb{P}(\boldsymbol{\theta}_{n+1} \notin \{\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_n\} | \boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_n, K_n). \end{aligned}$$

Une caractéristique importante est que ces probabilités dépendent indirectement des données via la taille des clusters $\boldsymbol{n} = (n_1, \ldots, n_{K_n})$. Par exemple, dans l'urne de Blackwell-MacQueen pour le DP (équation 2.7), la fonction de probabilité prédictive s'exprime suivant :

$$p_j(\boldsymbol{n}) = \frac{n_j}{n+\alpha} \mathbf{1}(1 \leq j \leq K_n) + \frac{\alpha}{\alpha+n} \mathbf{1}(j = K_n + 1) \tag{2.16}$$

avec $n = \sum_{j=1}^{K_n} n_j$.

Théorème 6. *[Pit96b, Proposition 11]*
Si $(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \ldots)$ *forme une séquence d'échantillonnage d'espèces alors*

1. *H_n, la distribution conditionnelle de $\boldsymbol{\theta}_{n+1} | \boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_n$ dans l'équation 2.15, converge presque-sûrement en norme de variation totale vers une distribution aléatoire H :*

$$H(\cdot) := \sum_{k=1}^{\infty} w_k \delta_{\boldsymbol{\theta}_k^*}(\cdot) + \left(1 - \sum_{k=1}^{\infty} w_k\right) G_0(\cdot) \tag{2.17}$$

   où $\sum_{k=1}^{\infty} w_k \leq 1$ p.s. avec $w_k := \lim_{n \to \infty} \frac{n_k}{n}$ p.s est la fréquence limite de la $k^{ème}$ espèce apparue dans la séquence $(\boldsymbol{\theta}_n)$,
2. *$\boldsymbol{\theta}_1^*, \boldsymbol{\theta}_2^*, \ldots \overset{iid}{\sim} G_0$ et sont indépendantes des w_k,*
3. *$\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \ldots | H \overset{iid}{\sim} H$.*

Pitman [Pit96b] souligne que ce théorème est défaillant par rapport au théorème de l'urne de Blackwell-MacQueen (Théorème 3) en ce sens qu'il suppose que la séquence est échangeable, ce qui découle de l'urne de Blackwell-MacQueen. Une variation à ce théorème est donnée par le corollaire suivant :

Corollaire 1. *[Pit96b], [HP98]*
Les variables $(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \dots)$ *forment une séquence d'échantillonnage d'espèces si et seulement* $\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \dots | H \overset{iid}{\sim} H$ *où* H *est une distribution de probabilité aléatoire sur* $(\Theta, \mathcal{B})$ *de la forme*

$$H(\cdot) = \sum_{k=1}^{\infty} \widetilde{w}_k \delta_{Z_k}(\cdot) + \left(1 - \sum_{k=1}^{\infty} \widetilde{w}_k\right) G_0(\cdot) \tag{2.18}$$

avec $(\widetilde{w}_k)_k$ *une séquence de poids aléatoires tels que*

$$\widetilde{w}_k \geq 0, \qquad \sum_{k=1}^{\infty} \widetilde{w}_k \leq 1 \ p.s.$$

et $Z_1, Z_2, \dots \overset{iid}{\sim} G_0$ *indépendamment des* $\widetilde{w}_k$*, et* G_0 *une mesure de probabilité diffuse sur* $(\Theta, \mathcal{B})$*.*

Définition 5. *Modèle d'échantillonnage d'espèces*
On appelle modèle d'échantillonnage d'espèces (noté SSM pour Species Sampling Model) la combinaison d'une mesure de probabilité aléatoire H *de la forme donnée dans l'équation 2.18 et d'un échantillon* $\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \dots$ *de* H*. En fait,* H *est la mesure de De Finetti associée à la séquence échangeable* $\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \dots$*.*

Si $\mathbb{P}(\sum_{k=1}^{\infty} \widetilde{w}_k = 1) = 1$, on dit que le SSM est *propre*. Si tel est le cas, on voit que la continuité de G_0 implique que H définit une distribution discrète et on a :

$$H(\cdot) = \sum_{k=1}^{\infty} \widetilde{w}_k \delta_{Z_k}(\cdot) = \sum_{k=1}^{\infty} w_k \delta_{\boldsymbol{\theta}_k^*}(\cdot) \tag{2.19}$$

où w_k et $\boldsymbol{\theta}_k^*$ sont définis dans l'équation 2.17.

Une séquence d'échantillonnage d'espèces est propre si et seulement si $\mathbb{P}(w_1 > 0) = 1$ ou, de façon équivalente, $\mathbb{P}(\lim_{n \to \infty} \frac{K_n}{n} = 0) = 1$ [Pit96b].

Le processus de Dirichlet (DP) est un cas particulier de SSM propre. Le terme $\sum_{k=1}^{\infty} w_k \delta_{\boldsymbol{\theta}_k^*}$ dans l'équation 2.19 correspond à la représentation stick-breaking de ce processus. Les SSM incluent donc le processus de Dirichlet mais aussi le processus de Pitman-Yor (PYP) que l'on verra par la suite. Les *a priori* stick-breaking généraux [IJ01] incluent le DP et le PYP. Ils s'écrivent sous la forme $\sum_{k=1}^{N} w_k \delta_{\boldsymbol{\theta}_k}^*$, où $1 \leq N \leq \infty$. Les poids sont définis comme $w_k = \prod_{j=1}^{k-1}(1 - v_j) v_k$ avec $v_k \sim \text{Beta}(a_k, b_k)$, pour des séquences $(a_1, a_2, \dots)$ et $(b_1, b_2, \dots)$ données. Cependant, les *a priori* stick-breaking généraux outre le DP et le PYP ne sont pas échangeables du fait de la non-symétrie de l'EPPF (cette fonction est définie dans le paragraphe qui suit). Ainsi, le processus de Pitman-Yor est le plus grand modèle d'allocation des résidus à définir un SSM échangeable.

2.2.2 La fonction de probabilité des partitions échangeables

On peut aussi définir un modèle d'échantillonnage d'espèces via une fonction de probabilité appelée EPPF (Exchangeable Partition Probability Function).

En effet, une séquence d'échantillonnage d'espèces $\boldsymbol{\theta} = (\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \dots)$ génère une partition aléatoire échangeable de l'ensemble $\{1, 2, \dots\}$ telle que deux entiers i et l appartiennent au même élément $\mathcal{A}_j$ de la partition si et seulement si $\boldsymbol{\theta}_i = \boldsymbol{\theta}_l = \boldsymbol{\theta}_j^*$. Soit Π_n la restriction de cette partition à $\{1, 2, \dots, n\}$ et K_n le nombre d'ensembles non-vides. On pose $\Pi_n := \{\mathcal{A}_1, \dots, \mathcal{A}_{K_n}\}$ où les $\mathcal{A}_j$ sont dans l'ordre d'apparition c'est-à-dire que $1 \in \mathcal{A}_1$ et pour tout $2 \leq j \leq K_n$, le premier élément de $\{1, \dots, n\} - (\mathcal{A}_1 \cup \dots \cup \mathcal{A}_{j-1})$ appartient à $\mathcal{A}_j$. Considérons $A_1, \dots, A_{K_n}$ une réalisation de Π_n. On a alors

$$\mathbb{P}\left(\Pi_n = \{A_1, \dots, A_{K_n}\}\right) = p(\#A_1, \dots, \#A_{K_n}) := p(n_1, \dots, n_{K_n}) \tag{2.20}$$

où p est une fonction non-négative et symétrique ie,

$$p(n_1, \dots, n_k) = p(n_{\sigma(1)}, \dots, n_{\sigma(k)})$$

pour toute permutation σ de $\{1, \dots, k\}$.

La fonction p est alors appelée EPPF dérivée de la séquence échangeable $(\boldsymbol{\theta}_n)$. Elle doit satisfaire :

$$\begin{aligned} p(1) &= 1, \\ p(\boldsymbol{n}) &= \sum_{j=1}^{K_n+1} p(\boldsymbol{n}^{j+}) \text{ pour tout } \boldsymbol{n} \in \mathbf{N}^*, \end{aligned} \tag{2.21}$$

où $\boldsymbol{n}^{j+}$ représente le vecteur $\boldsymbol{n}$ dont le j$^{\text{ème}}$ élément est incrémenté de 1 pour $1 \leq j \leq K_n$ et dont l'élément pour $j = K_n + 1$ est 1. La fonction p détermine donc la distribution de la partition aléatoire de $\{1, 2, \dots\}$ dont les classes $\mathcal{A}_j$ sont les classes d'équivalence pour la relation d'équivalence aléatoire $i \sim j$ si et seulement si $\boldsymbol{\theta}_i = \boldsymbol{\theta}_j$.

Inversement, toute fonction symétrique p, non-négative, définie sur $\mathbf{N}^* := \bigcup_{k=1}^{\infty} \mathbb{N}^k$, et satisfaisant l'équation 2.21 est l'EPPF d'une séquence échangeable $(\boldsymbol{\theta}_n)$ ([Pit95], [Pit96b]).

On peut donc, de façon alternative, caractériser un SSM par une EPPF p et une mesure de base diffuse G_0. Dans ce cas, les fonctions de probabilité prédictives dans l'équation 2.15 s'expriment en fonction de l'EPPF de la façon suivante :

$$p_j(\boldsymbol{n}) = \frac{p(\boldsymbol{n}^{j+})}{p(\boldsymbol{n})} \quad \text{pour } 1 \leq j \leq K_n + 1 \quad \text{si } p(\boldsymbol{n}) > 0.$$

Démonstration.

1. pour $j = 1, \dots, K_n$,

$$\begin{aligned} p_j(\boldsymbol{n}) &:= \mathbb{P}(\boldsymbol{\theta}_{n+1} = \boldsymbol{\theta}_j^* | \boldsymbol{\theta}_1, \dots, \boldsymbol{\theta}_n) \\ &= \frac{\mathbb{P}(\boldsymbol{\theta}_1, \dots, \boldsymbol{\theta}_n, \boldsymbol{\theta}_{n+1} = \boldsymbol{\theta}_j^*)}{\mathbb{P}(\boldsymbol{\theta}_1, \dots, \boldsymbol{\theta}_n, \boldsymbol{\theta}_{n+1} = \boldsymbol{\theta}_j^*) + \mathbb{P}(\boldsymbol{\theta}_1, \dots, \boldsymbol{\theta}_n, \boldsymbol{\theta}_{n+1} = \boldsymbol{\theta}_{new})} \text{ par la règle de Bayes} \\ &= \frac{p(n_1, \dots, n_j + 1, \dots, n_{K_n})}{\sum_{j=1}^{K_n} p(n_1, \dots, n_j + 1, \dots, n_{K_n}) + p(n_1, \dots, n_{K_n}, 1)} \text{ par la définition de l'EPPF } p. \end{aligned}$$

2. pour $j = K_{n+1}$,

$$\begin{aligned} p_{K_{n+1}}(\boldsymbol{n}) &:= \mathbb{P}(\boldsymbol{\theta}_{n+1} = \boldsymbol{\theta}_{new} | \boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_n) \\ &= \frac{\mathbb{P}(\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_n, \boldsymbol{\theta}_{n+1} = \boldsymbol{\theta}_{new})}{\mathbb{P}(\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_n, \boldsymbol{\theta}_{n+1} = \boldsymbol{\theta}_j^*) + \mathbb{P}(\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_n, \boldsymbol{\theta}_{n+1} = \boldsymbol{\theta}_{new})} \\ &= \frac{p(n_1, \ldots, n_{K_n}, 1)}{\sum_{j=1}^{K_n} p(n_1, \ldots, n_j + 1, \ldots, n_{K_n}) + p(n_1, \ldots, n_{K_n}, 1)}. \end{aligned}$$

En posant $p(\boldsymbol{n}) = p(n_1, \ldots, n_{K_n}, 1) + \sum_{j=1}^{K_n} p(n_1, \ldots, n_j + 1, \ldots, n_{K_n})$, on a le résultat.

□

Finalement, on voit que l'échangeabilité d'une séquence $(\boldsymbol{\theta}_n)$ se révèle être équivalente à l'existence d'une EPPF p qui détermine les probabilités prédictives p_j. Cette propriété est formalisée par le théorème suivant :

Théorème 7. *[Pit96b, Proposition 13 & Théorème 14], [HP98, Théorème 2]*
Soit $(\boldsymbol{\theta}_n)$ une séquence échangeable assujettie à une règle de prédiction de la forme

$$\begin{aligned} \boldsymbol{\theta}_1 &\sim G_0(\cdot), \\ \boldsymbol{\theta}_{n+1} | \boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_n &\sim \sum_{j=1}^{K_n} p_j(\boldsymbol{n}) \delta_{\boldsymbol{\theta}_j^*} + p_{K_n+1}(\boldsymbol{n}) G_0(\cdot). \end{aligned} \tag{2.22}$$

Alors les fonctions prédictives s'expriment en fonction de l'EPPF p associée à la partition aléatoire Π générée par la séquence $(\boldsymbol{\theta}_n)$:

$$p_j(\boldsymbol{n}) = \frac{p(\boldsymbol{n}^{j+})}{p(\boldsymbol{n})} \text{ pour } 1 \leq j \leq K_n + 1 \text{ si } p(\boldsymbol{n}) > 0. \tag{2.23}$$

Inversement, étant données une mesure de probabilité diffuse G_0 sur $(\Theta, \mathcal{B})$ et une fonction symétrique non-négative sur $\mathbf{N}^*$ satisfaisant $p(1) = 1$ et $p(\boldsymbol{n}) = \sum_{j=1}^{K_n+1} p(\boldsymbol{n}^{j+})$ pour tout $\boldsymbol{n} \in \mathbf{N}^*$, la règle de prédiction

$$\begin{aligned} \boldsymbol{\theta}_1 &\sim G_0(\cdot), \\ \boldsymbol{\theta}_{n+1} | \boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_n &\sim \sum_{j=1}^{K_n} p_j(\boldsymbol{n}) \delta_{\boldsymbol{\theta}_j^*} + p_{K_n+1}(\boldsymbol{n}) G_0(\cdot) \end{aligned} \tag{2.24}$$

(où les p_j sont donnés par l'équation 2.23) définit une séquence échangeable $(\boldsymbol{\theta}_n)$.

Une telle séquence $(\boldsymbol{\theta}_n)$ peut être construite en générant d'abord une partition aléatoire échangeable $\Pi = \{\mathcal{A}_1, \mathcal{A}_2, \ldots, \}$ dont l'EPPF est p, puis pour chaque ensemble k (cluster) de la partition, tirer $\boldsymbol{\theta}_k^* \overset{iid}{\sim} G_0$, indépendamment de Π. Et enfin, affecter $\boldsymbol{\theta}_k^*$ à chaque élément i du cluster k, i.e faire $\boldsymbol{\theta}_i = \boldsymbol{\theta}_k^*$ pour tout $i \in \mathcal{A}_k$.

Exemple :

L'EPPF correspondant aux fonctions prédictives dans l'urne de Blackwell-MacQueen (cf. équation 2.16) est donnée par :

$$p(n_1, \ldots, n_K) := p_\alpha(n_1, \ldots, n_K) = \frac{\alpha^{K-1} \prod_{j=1}^{K} (n_j - 1)!}{[1 + \alpha]_{n-1}} \tag{2.25}$$

avec $n = \sum_k n_k$ et $[x]_m = \frac{\Gamma(x+m)}{\Gamma(x)} = x(x+1)\dots(x+m-1)$ pour tout $m \geq 1$ et $[x]_0 = 1$. Comme on le verra, ce résultat s'étend à deux paramètres pour les processus de Pitman-Yor.

Remarque 2. *La terminologie « échantillonnage d'espèces » provient de l'application du modèle en écologie pour modéliser la diversité des espèces et en génétique des populations. Au lieu de regarder les réalisations des $\boldsymbol{\theta}_i$ qui représentent les labels des espèces et ne sont qu'arbitraires, on peut plutôt s'intéresser à la probabilité d'observer K_n espèces différentes, de tailles $(n_1, \dots, n_{K_n})$, dans un échantillon de taille n. Un exemple d'application en génétique des populations est la formule d'échantillonnage d'Ewens [ET98]. Cette formule provient de la théorie « non darwinienne » de l'évolution selon laquelle la variation génétique observée dans des populations ne serait pas due à la sélection naturelle mais plutôt serait un résultat de changements purement stochastiques. La formule d'Ewens donne la probabilité de la partition de n gênes en un nombre de types de gênes différents (allèles). Soit un échantillon aléatoire de n gamètes tiré d'une population et classé selon le gène à un locus particulier. On dénote par $m_j \geq 0, j = 1, \dots, n$ le nombre d'allèles apparaissant exactement j fois dans l'échantillon. Alors la probabilité qu'il y'ait m_1 allèles représentés une fois dans l'échantillon, m_2 allèles représentés deux fois, etc. est :*

$$p^*(m_1, \dots, m_n) = \frac{n!}{[\alpha]_n} \prod_{j=1}^{n} \left(\frac{\alpha}{j}\right)^{m_j} \frac{1}{m_j!}$$

où pour tout $n \geq 1$, $m_1, \dots, m_n$ est tel que $m_1 + 2m_2 + \dots + nm_n = n$.
Antoniak [Ant74] a donné une dérivée de cette formule dans le contexte des statistiques bayésiennes pour caractériser l'EPPF associée à l'échantillonnage a priori du processus de Dirichlet via l'urne de Blackwell-MacQueen (équation 2.25). Pitman généralise cette formule à deux paramètres pour caractériser l'EPPF associée au processus de Pitman-Yor([Pit92], [Pit95] et [Pit96b]).

2.3 Le processus de Pitman-Yor

Avant de présenter le processus de Pitman-Yor, nous commençons par introduire quelques définitions et résultats préalables.

2.3.1 Définitions et propriétés

La distribution GEM à deux paramètres

La distribution GEM des poids que nous avons vue dans la représentation stick-breaking du processus de Dirichlet s'étend à deux paramètres. Soit $d \in [0, 1]$, $\alpha > -d$, $v_1, v_2, \dots$ des variables aléatoires indépendantes telles que pour tout j, $v_j \sim \text{Beta}(1-d, \alpha + jd)$. On définit une séquence de poids (w_j) par le RAM c'est-à-dire $w_1 = v_1, w_2 = v_2(1-v_1), \dots, w_j = v_j \prod_{i=1}^{j-1}(1-v_i)$. La distribution de la séquence des poids $w_1, \dots, w_j$ est appelée distribution GEM de paramètres d et α. On note :

$$\mathbf{w} \sim \text{GEM}(d, \alpha).$$

La distribution de Poisson-Dirichlet à deux paramètres

Les statistiques d'ordre décroissant des poids de la distribution GEM à deux paramètres est la distribution de Poisson-Dirichlet à deux paramètres. Plus précisément, soit $\mathbf{w} \sim \mathrm{GEM}(d, \alpha)$. On considère la séquence réordonnée des poids $\widetilde{\mathbf{w}} = r(\mathbf{w})$ telle que $\widetilde{w}_1 \geq \widetilde{w}_2 \geq \dots$ et $r(\cdot)$ la fonction qui réordonne la séquence. Alors cette séquence réordonnée suit une distribution de Poisson-Dirichlet. On note :

$$\widetilde{\mathbf{w}} \sim \mathrm{PD}(d, \alpha).$$

Pour plus de discussions sur la distribution de Poisson-Dirichlet, on pourra se référer à Kingman [Kin75] qui a étudié cette distribution pour $d = 0$, $\mathrm{PD}(0, \alpha)$, et lui a donné son nom ; Perman [Per90] a étudié le cas $\alpha = 0$, $\mathrm{PD}(d, 0)$; enfin Perman *et al.* [PPY92] et Pitman et Yor [PY97] ont montré que ces deux séquences de probabilité étaient des cas particuliers d'une autre classe de familles de distributions, plus large et à deux paramètres, $\mathrm{PD}(d, \alpha)$.

Remarque 3. *La distribution de Poisson-Dirichlet à deux paramètres est aussi définie comme la hauteur des sauts d'un subordinateur gamma. Un subordinateur est un processus de Lévy stochastiquement non-décroissant. Nous ne détaillerons pas ici ces aspects relatifs aux subordinateurs. Pour plus d'informations, on pourra se référer à [PY97].*

La permutation biaisée par la taille

Le nom de la permutation biaisée par la taille (« size-biased ») provient de l'utilisation des distributions discrètes pour modéliser la division d'une très grande population (des animaux par exemple) en un grand nombre d'espèces différentes ([PY97], [Fen10], [Car99]).

Supposons que l'on ait une population composée d'un nombre infini dénombrable d'espèces différentes labellisées par $\mathbb{N}$ (l'ensemble des entiers naturels). Soit $\widetilde{p}_n$ la proportion dans cette population de la $n^{\text{ième}}$ espèce telle que $\widetilde{p}_k \geq 0$ et pour tout k, $\sum_{k=1}^{\infty} \widetilde{p}_k = 1$. Considérons maintenant un processus d'échantillonnage aléatoire à partir de cette population. La permutation biaisée par la taille de $(\widetilde{p}_n)$ est la séquence (p_n) réordonnée en fonction de l'ordre d'apparition des espèces dans le tirage. En d'autres termes, on construit la séquence (p_n) de la façon suivante : $p_1 = \widetilde{p}_n$ si le premier échantillon observé est de type n, $n \in \mathbb{N}$; $p_2 = \widetilde{p}_m$ si le second échantillon n'étant pas de type n est de type m, $m \in \mathbb{N} - n$; et ainsi de suite... On peut imaginer que l'on tire au hasard un premier individu et on note $\sigma(1)$ son espèce. Pour éviter de retomber sur la même espèce, on enlève de la population tous les individus de type $\sigma(1)$. Puis, on procède de la même façon pour le deuxième tirage. Comme on enlève les espèces tirées, on doit à chaque fois normaliser pour obtenir les poids des espèces restantes. Cette façon de tirer est appelé échantillonnage biaisé par la taille. Si on note par $\sigma(i)$ le type du $i^{\text{ème}}$ individu tiré, alors : $(p_1, p_2, \dots) := (\widetilde{p}_{\sigma(1)}, \widetilde{p}_{\sigma(2)}, \dots)$ est appelée permutation biaisée par la taille de $(\widetilde{p}_1, \widetilde{p}_2, \dots)$.

La distribution de la permutation biaisée par la taille s'écrit ainsi :

$$\mathbb{P}(p_1 = \widetilde{p}_n | \widetilde{p}_1, \widetilde{p}_2, \ldots) = \widetilde{p}_n, \text{ et pour tout } j \geq 1$$
$$\mathbb{P}(p_{j+1} = \widetilde{p}_n | p_1, \ldots, p_j, \widetilde{p}_1, \widetilde{p}_2, \ldots) = \frac{\widetilde{p}_n}{1 - p_1 - \cdots - p_j} \mathbb{I}\{\widetilde{p}_n \neq p_1, \ldots, p_j\}.$$

Définition 6. *Invariance par permutation biaisée par la taille*

Soit (p_n) une permutation biaisée par la taille d'une séquence de probabilité $(\widetilde{p}_n)$. On dit que $(\widetilde{p}_n)$ est invariant par permutation biaisée par la taille (notée ISBP pour Invariant by Size-Biazed Permutation) si $(\widetilde{p}_n)$ et (p_n) ont la même loi, i.e. $p_n \overset{d}{=} \widetilde{p}_n$.

Remarque 4. *La distribution des poids que nous avons vue dans la définition du DP comme processus gamma normalisé (équation 2.5) est une distribution de Poisson-Dirichlet $PD(0, \alpha)$. McCloskey [McC65] a montré que si $(\widetilde{p}_n) \sim PD(0, \alpha)$ et si on considère (p_n) une permutation biaisée par la taille de $(\widetilde{p}_n)$, alors $(p_n) \sim GEM(0, \alpha)$. Ce résultat montre pourquoi la réprésentation stick-breaking de Sethuraman (équation 2.11) et la construction du DP comme processus gamma normalisé de Ferguson (équation 2.6) conduisent au même processus qui est le processus de Dirichlet.*

Pitman [Pit96b] a montré que cette propriété d'invariance par permutation biaisée par la taille s'étend aux distributions à deux paramètres :

Théorème 8. *Soit $\boldsymbol{w} = w_1, w_2, \ldots$ une séquence de probabilité telle que $w_j > 0$ pour tout j et $\sum_j w_j = 1$. Supposons que $\boldsymbol{w} = (w_j)_j$ est définie suivant le RAM. Alors $\boldsymbol{w}$ est invariant par permutation biaisée par la taille si et seulement si $\boldsymbol{w} \sim GEM(d, \alpha)$ avec $d \in [0, 1]$, $\alpha > -d$.*

Pour résumer, nous avons vu que la permutation biaisée par la taille d'une distribution de Poisson-Dirichlet est une distribution GEM et, inversement, la distribution d'une séquence réordonnée de poids GEM est une distribution de Poisson-Dirichlet. De plus, la distribution GEM est le seul RAM qui soit invariant par permutation biaisée par la taille.

2.3.2 Le processus de Pitman-Yor

Définition 7. *Soient $d \in [0, 1]$, $\alpha > -d$ et $\widetilde{\mathbf{w}} \sim PD(d, \alpha)$ [3]. Soit G_0 une mesure de probabilité diffuse sur $(\Theta, \mathcal{B})$ (c'est-à-dire telle que $G_0(\boldsymbol{\theta}) = 0$ pour tout $\boldsymbol{\theta} \in \Theta$). On considère $Z_1, Z_2, \ldots$ des variables aléatoires iid suivant G_0, indépendantes de $\widetilde{\mathbf{w}}$. Alors, la mesure*

$$H(\cdot) = \sum_{k=1}^{\infty} \widetilde{w}_k \delta_{Z_k}(\cdot),$$

est dite distribuée suivant un processus de Pitman-Yor sur $(\Theta, \mathcal{B})$ de mesure de base G_0 et de paramètres d et α. On l'appelle aussi processus de Poisson-Dirichlet à deux paramètres. On note :

$$H \sim PY(d, \alpha, G_0).$$

3. Il existe aussi une autre paramétrisation du PYP, avec $d = -\kappa < 0$ et $\alpha = m\kappa$ pour un certain $\kappa > 0$ et $m = 2, 3, \ldots$

Quand $d = 0$, le processus de Pitman-Yor se ramène à un processus de Dirichlet de paramètre α et de mesure de base G_0.

Tout comme le processus de Dirichlet (DP), le processus de Pitman-Yor (PYP) possède plusieurs propriétés importantes et qui contribuent à son utilité pratique. En l'occurence, le PYP possède une représentation stick-breaking et une caractérisation en termes d'urne de Blackwell-MacQueen généralisée ([Pit96a], [Pit96b], [IJ01]).

Représentation stick-breaking du processus de Pitman-Yor

Ishwaran et James ([IJ01], [IJ03b]) ont étendu la représentation stick-breaking du processus de Dirichlet de Sethuraman [Set94] à d'autres mesures aléatoires plus générales comprenant entre autres le DP et le PYP.

Soient $d \in [0, 1]$, $\alpha > -d$ et $\mathbf{w} \sim$GEM(d, α). Soit G_0 une mesure de probabilité diffuse sur Θ et $\boldsymbol{\theta}_1^*, \boldsymbol{\theta}_2^*, \ldots$ des variables aléatoires *iid* suivant G_0, indépendantes de $\mathbf{w}$. Soit

$$H(\cdot) = \sum_{k=1}^{\infty} w_k \delta_{\boldsymbol{\theta}_k^*}(\cdot), \tag{2.26}$$

alors H est une mesure de probabilité aléatoire distribuée suivant un processus de Pitman-Yor de paramètres d, α et G_0.

Distribution *a posteriori* du PYP dans la représentation stick-breaking

Si la distribution *a priori* d'un PYP est caractérisée par une représentation stick-breaking, la loi *a posteriori* l'est aussi.

Plus précisément, soit H une mesure de probabilité aléatoire générée via le processus stick-breaking (équation 2.26). Soit $\boldsymbol{\Theta} = (\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_n)$ un échantillon tiré de H. Soit $\boldsymbol{c} = (c_1, \ldots, c_n)$ des variables de classification telles que $c_i = k$ ssi $\boldsymbol{\theta}_i = \boldsymbol{\theta}_k^*$. Soient $\mathbf{c}^* = (c_1^*, \ldots, c_{K_n}^*)$ les valeurs uniques de $c_1, \ldots, c_n$ et $\boldsymbol{\theta}^* = (\boldsymbol{\theta}_1^*, \ldots, \boldsymbol{\theta}_{K_n}^*)$ les valeurs uniques de $(\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_n)$. Alors,

$$H(\cdot)|\boldsymbol{\theta} = \sum_{k \in \mathbf{c}^*} w_k^* \delta_{\boldsymbol{\theta}_k^*}(\cdot) + \sum_{k \notin \mathbf{c}^*} w_k \delta_{Z_k}(\cdot) \tag{2.27}$$

avec $\forall\, k \notin \mathbf{c}^*$, $Z_k \overset{iid}{\sim} G_0$; les poids w_k^* suivent une distribution GEM avec des paramètres mis à jour c'est-à-dire $w_1^* = v_1^*, w_2^* = v_2^*(1 - v_1^*), \ldots, w_n^* = v_n^* \prod_{i=1}^{n-1}(1 - v_i^*)$ où $v_l^* \sim$ Beta$(1 - d + n_l, \alpha + ld + \sum_{m=l+1}^{\infty} n_m)$ avec $n_m = \sum_{i=1}^{n} \mathbf{1}(c_i = m)$. La mise à jour des poids qui s'effectue suivant une distribution GEM est une conséquence de la conjugaison entre la distribution de Dirichlet généralisée et l'échantillonnage multinomial dans le cas où la dimension de l'espace est finie. Ishwaran et James [IJ03b] ont démontré que cela reste vrai en dimension infinie.

Représentation généralisée de l'urne de Blackwell-MacQueen pour le PYP

Tout comme le DP, le PYP peut aussi être caractérisé par un mécanisme d'urne de Blackwell-MacQueen généralisée ([Pit95], [Pit96a]). Ce résultat provient en fait de la

caractérisation du PYP via le modèle d'échantillonnage d'espèces.

Soit $d \in [0,1]$, $\alpha > -d$ et la distribution prédictive suivante

$$p_j(\boldsymbol{n}) = \frac{n_j - d}{n+\alpha}\mathbf{1}(1 \leq j \leq K_n) + \frac{\alpha + dK_n}{n+\alpha}\mathbf{1}(j = K_{n+1}). \tag{2.28}$$

Cette règle de prédiction satisfait l'équation 2.23 pour la fonction EPPF $p := p_{d,\alpha}$ définie par

$$p_{d,\alpha}(n_1, \ldots, n_K) = \frac{\left(\prod_{l=1}^{K-1}(\alpha + ld)\right)\left(\prod_{j=1}^{K}[1-d]_{(n_j-1)}\right)}{[1+\alpha]_{n-1}}.$$

Puisque $p_{d,\alpha}$ est symétrique, $p_{d,\alpha}(1) = 1$, alors d'après le théorème 7, la séquence $(\boldsymbol{\theta}_n)$ générée par cette urne est échangeable et donc définit une séquence d'échantillonnage d'espèces. On peut alors définir le SSM associé à la séquence par la paire $(p_{d,\alpha}, G_0)$ où G_0 une mesure de probabilité diffuse sur Θ. La mesure caractérisant les $\boldsymbol{\theta}_i$ dans le SSM, i.e telle que $\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \ldots | H \sim H$ est donnée par

$$H(\cdot) = \sum_{k=1}^{+\infty} w_k \delta_{\boldsymbol{\theta}_k^*}(\cdot) = \sum_{k=1}^{+\infty} \widetilde{w}_k \delta_{Z_k}(\cdot)$$

où

1. $\widetilde{\mathbf{w}} \sim$PD$(d, \alpha)$,
2. $\mathbf{w}$ est une permutation biaisée par la taille de $\widetilde{\mathbf{w}}$ et $\mathbf{w} \sim$ GEM(d, α),
3. $Z_1, Z_2, \ldots$ sont *iid* suivant G_0 et indépendantes de $\widetilde{\mathbf{w}}$.

Dans cette représentation, les poids sont donnés par les fréquences empiriques, $w_k = \lim_{n\to\infty} \frac{n_k}{n}$, où les fréquences sont définies dans l'ordre d'apparition des espèces (i.e w_k est la fréquence de $\boldsymbol{\theta}_k^*$, la $k^{\text{ème}}$ espèce apparue. Ce sont donc des permutations biaisées par la taille des poids $\widetilde{w} \sim$ PD(d, α). Ce résultat montre que les distributions dans la représentation de Poisson-Dirichlet $(\sum_{k=1}^{+\infty} \widetilde{w}_k \delta_{Z_k})$ et dans la représentation stick-breaking $(\sum_{k=1}^{+\infty} w_k \delta_{\boldsymbol{\theta}_k^*})$ ont la même loi et c'est un processus de Pitman-Yor.

Soit $d \in [0,1]$, $\alpha > -d$ et G_0 une mesure de probabilité diffuse sur Θ. Soit $\boldsymbol{\theta} = (\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_n) \sim H$ où H est une mesure de probabilité sur Θ. Considérons une séquence de $\boldsymbol{\theta}_i$ générés via la distribution prédictive suivante

$$\mathbb{P}(\boldsymbol{\theta}_{n+1} \in \cdot | \boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_n) = \frac{\alpha + dK_n}{n+\alpha} G_0(\cdot) + \sum_{j=1}^{K_n} \frac{n_j - d}{n+\alpha} \delta_{\boldsymbol{\theta}_j^*}(\cdot) \tag{2.29}$$

où K_n désigne le nombre de $\boldsymbol{\theta}_i$ différents ; $\boldsymbol{\theta}_1^*, \ldots, \boldsymbol{\theta}_{K_n}^*$ les valeurs uniques de $\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_n$, chacune représentées n_j fois pour $j = 1, \ldots, K_n$.

Alors $H \sim$ PY(d, α, G_0) de distribution

$$H(\cdot) = \sum_{j=1}^{+\infty} w_j \delta_{\boldsymbol{\theta}_j^*}(\cdot)$$

où $\boldsymbol{w} \sim$ GEM(d, α), $\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_n \sim G_0$ indépendamment de $\boldsymbol{w}$.

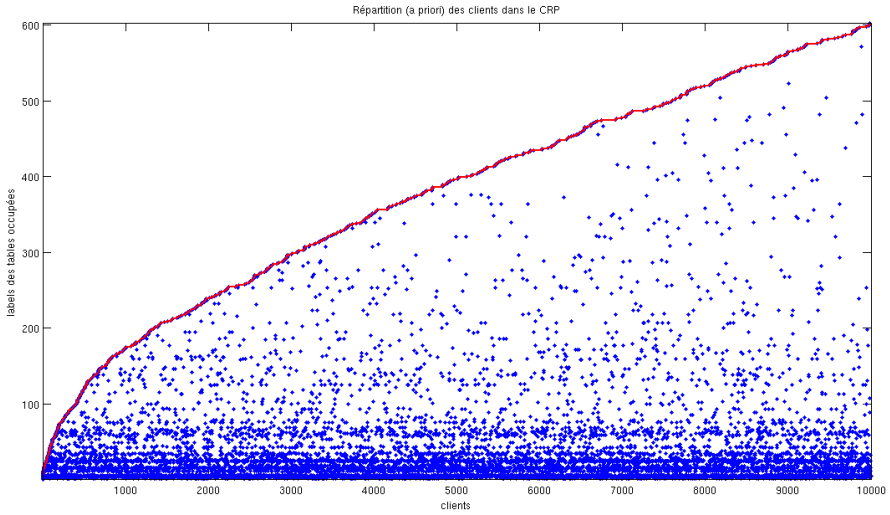


FIGURE 2.5 – Comportement de la loi de puissance (« power-law ») du PYP illustré avec le restaurant chinois. Pour $n = 10.000$, $d = 0.5$ et $\alpha = 10$, le nombre de clusters *a priori* (tables) est de 603.

Pour $d = 0$, on retrouve la distribution prédictive de l'urne de Blackwell-MacQueen pour le processus de Dirichlet de paramètre α.

Remarques :

1. *Contrairement au DP, dans le PYP les probabilités prédictives dépendent de K_n, le nombre de clusters observés dans les n premiers tirages. On peut donc utiliser la valeur de d pour régler la force sur la dépendance par rapport à K_n*
2. *On a vu, dans la représentation du restaurant chinois pour le DP, que le nombre de clusters augmente de façon logarithmique avec n, $K_n \sim \alpha \log(n)$ quand $n \to \infty$. Dans le PYP, pour $d \in]0, 1[$ et $\alpha > -d$, le nombre de clusters $K_n \sim S_{d,\alpha} n^d$ où $S_{d,\alpha}$ est une variable aléatoire positive dont la densité de probabilité ne dépend que de d et de α ([Pit03] et [Pit02], Théorème 31). L'augmentation de K_n est donc en $O(\alpha n^d)$ dans le PYP. Ce résultat est illustré dans la Figure 2.5.*

Remarque 5. *Dans l'équation 2.19, si $(\widetilde{w}_k)$ est une séquence de poids décroissante, alors (w_k) est une permutation biaisée par la taille de $(\widetilde{w}_k)$ [Pit96b]. Dans le DP et le PYP, les $(\widetilde{w}_k)$ suivent une distribution de Poisson-Dirichlet (qui par définition est une séquence de poids décroissants) et donc* $\mathbf{w}$ *est la permutation biaisée par la taille de* $\widetilde{\mathbf{w}}$ *et* $\mathbf{w}$ *suit une distribution GEM. Les SSM généralisent donc le DP et le PYP.*

Distribution *a posteriori* du PYP dans le SSM

Nous présentons ici un résultat important de Pitman ([Pit96b], Corollaire 20), dont on peut trouver une preuve détaillée dans la thèse de Carlton [Car99]. Ce résultat caractérise la distribution *a posteriori* dans un processus de Pitman-Yor, sous le modèle d'échantillonnage d'espèces.

Soient $H \sim \mathrm{PY}(d, \alpha, G_0)$ où G_0 une mesure de probabilité non-atomique telle que $\mathbb{E}(H) = G_0$. Soient $\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_n \sim H$ et K_n le nombre de valeurs distinctes de $\boldsymbol{\theta}_i$, soit

$\{\boldsymbol{\theta}_j^*\}_{j=1}^{K_n}$ l'ensemble des valeurs uniques de $\{\boldsymbol{\theta}_i\}_{i=1}^n$ et finalement soit n_j le nombre de fois où $\boldsymbol{\theta}_j^*$ est apparu. Alors,

$$H|\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_n \stackrel{d}{=} \sum_{j=1}^{K_n} w_j \delta_{\boldsymbol{\theta}_j^*} + r_{K_n} H_{K_n} \tag{2.30}$$

avec :

$$\begin{aligned} (w_1, \ldots, w_{K_n}, r_{K_n}) &\sim \mathrm{Dir}(n_1 - d, \ldots, n_{K_n} - d, \alpha + dK_n) \\ H_{K_n} &\sim \mathrm{PY}(d, \alpha + dK_n, G_0) \end{aligned}$$

avec H_{K_n} indépendant de $(w_1, \ldots, w_{K_n}, r_{K_n})$ et $\mathbb{E}(H_{K_n}) = G_0$.

Ce résultat est à la base de notre nouvelle méthode d'échantillonnage des mélanges par processus de Pitman-Yor. Nous reviendrons plus en détails sur ce résultat et ses conséquences dans le chapitre 3 sur l'échantillonnage.

2.3.3 Le mélange par processus de Pitman-Yor

Comme pour le processus de Dirichlet, les mesures générées par un processus de Pitman-Yor sont discrètes et ne peuvent être utilisées pour l'estimation de densités. Ici aussi, pour définir un *a priori* sur une mesure aléatoire dont les réalisations conduisent à des mesures de probabilité admettant des densités presque-sûrement, une solution consiste à convoluer le PYP avec un noyau paramétrique. C'est le modèle de mélange par processus de Pitman-Yor (Pitman-Yor process mixtures (PYM)). Il peut être vu comme la généralisation à deux paramètres du modèle de mélange par processus de Dirichlet (DPM) (cf. équation 2.13). La densité p dans un PYM est obtenue de la façon suivante :

$$\begin{aligned} p(\cdot) &= \int f(\cdot|\boldsymbol{\theta}) H(\mathrm{d}\boldsymbol{\theta}) \\ H|d, \alpha, G_0 &\sim \mathrm{PY}(d, \alpha, G_0). \end{aligned} \tag{2.31}$$

Exemple : mélange Pitman-Yor de Gaussiennes multivariées

Dans l'équation 2.31, on peut prendre pour noyau une distribution normale multivariée, de densité $f = f_{\mathcal{N}}$ ayant pour paramètres le vecteur moyen $\mathbf{m}$ et la matrice de covariance $\boldsymbol{\Sigma}$. La densité résultante p est alors

$$p(\boldsymbol{x}) = \sum_{k=1}^{\infty} w_k \, f_{\mathcal{N}}(\boldsymbol{x}|\mathbf{m}_k, \boldsymbol{\Sigma}_k) \tag{2.32}$$

avec $\boldsymbol{w} \sim \mathrm{GEM}(d, \alpha)$ et $\boldsymbol{\theta}_k^* = (\mathbf{m}_k, \boldsymbol{\Sigma}_k) \sim G_0$.

Puisque G_0 définit une distribution sur l'espace des paramètres des clusters, à savoir les $\boldsymbol{\theta}_k^* = (\mathbf{m}_k, \boldsymbol{\Sigma}_k)$, on peut la choisir suivant une loi Normal-Inverse Wishart $(\mathcal{NIW}_{\rho, n_0, \mu_0, \boldsymbol{\Sigma}_0})$, définie de la façon suivante :

$$\mathbf{m}|\boldsymbol{\Sigma} \sim \mathcal{N}(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}/\rho),$$

où ρ est le paramètre de précision, $\boldsymbol{\mu}_0$ la moyenne de la loi normale et

$$\boldsymbol{\Sigma}^{-1} \sim \mathcal{W}(n_0, (n_0\,\boldsymbol{\Sigma}_0)^{-1}),$$

avec $\mathcal{W}$ désignant la distribution de Wishart, n_0 le degré de liberté et $\boldsymbol{\Sigma}_0^{-1}$ la moyenne.

$$f_{\mathcal{W}}\left(\boldsymbol{\Sigma}^{-1}\right) = \frac{|n_0\boldsymbol{\Sigma}_0|^{\frac{n_0}{2}}}{2^{\frac{n_0 k}{2}}\,|\boldsymbol{\Sigma}|^{\frac{n_0-k-1}{2}}\,\Gamma_k(\frac{n_0}{2})} \exp\left(-\frac{n_0}{2} Tr\left(\boldsymbol{\Sigma}^{-1}\boldsymbol{\Sigma}_0\right)\right)$$

où $Tr\,(\mathbf{M})$ représente la trace de la matrice $\mathbf{M}$ et $\Gamma_k(\cdot)$ est la fonction gamma k-variée.

Remarque 6. *Comme nous l'avons souligné pour les mélanges par processus de Dirichlet (section 2.1.7), la distribution prédictive dans les mélanges par processus de Pitman-Yor n'a pas de forme utilisable en pratique. Les méthodes MCMC que l'on verra dans le chapitre 3 permettront d'obtenir de telles estimées.*

2.4 Les processus dépendants

Dans les modèles bayésiens non paramétriques vus jusqu'à présent, la loi *a priori* porte sur une seule distribution. Cependant, dans certaines applications, l'objectif est de modéliser une collections de distributions $\{G_s, s \in \mathcal{S}\}$ où $\mathcal{S}$ est un intervalle temporel, une région de l'espace etc.

Le processus de Dirichlet dépendant (noté DDP pour Dependent Dirichlet Process, [Mac99]) introduit une dépendance dans une collection de distributions en remplaçant les poids et les atomes dans la représentation stick-breaking par des processus stochastiques sur $\mathcal{S}$ (c'est-à-dire des collections indexées de variables aléatoires, l'ensemble d'indices étant typiquement l'espace ou le temps mais il peut être un espace plus général). L'idée est que pour tout s, $G_s \sim \mathrm{DP}(\alpha, G_0)$ et que si $s \approx s'$ alors G_s et $G_{s'}$ doivent être similaires. On a,

$$\forall s,\ G_s(\cdot) = \sum_{k=1}^{+\infty} \pi_{sk}\delta_{\boldsymbol{\theta}^*_{sk}}(\cdot) \tag{2.33}$$

c'est-à-dire que pour k fixé, les poids π_{sk} et les localisations $\boldsymbol{\theta}^*_{sk}$ sont des processus stochastiques généraux. On obtient des collections de processus de Dirichlet qui varient régulièrement comme fonction d'une certaine covariable. Un exemple est le DDP spatial [MKG01] où les poids π_k sont constants et chaque atome $\boldsymbol{\theta}^*_{sk}$ est un processus Gaussien (GP) indexé par $s \in \mathbb{R}^2$. Il existe plusieurs autres processus dépendants. On peut citer entre autres le π-DDP [GS06], les GSDP (Generalized Spatial Dirichlet Processes, [DGG07]), ou les KSB (Kernel Stick-Breaking processes [DP08]). Nous allons présenter brièvement deux cas particuliers de DDP à savoir le processus de Dirichlet hiérarchique (HDP) et le processus de Dirichlet imbriqué (NDP).

2.4.1 Le processus de Dirichlet hiérarchique

Le but du processus de Dirichlet hiérarchique (Hierarchical Dirichlet Process (HDP), [TJBB06]) est d'introduire une dépendance entre des distributions spécifiant des groupes

où les observations sont supposées êtres plus similaires dans un même groupe qu'entre les groupes. L'idée est de supposer que ces distributions sont tirées d'un DP commun dont la mesure de base est elle-même tirée d'un autre DP. Si on désigne par j l'indice des groupes et i celui des observations, le HDP est donné par le modèle suivant :

$$\begin{aligned} G_0 &\sim \mathrm{DP}(\alpha_0, H), \\ \forall j, G_j &\sim \mathrm{DP}(\alpha, G_0), \\ \forall i, \boldsymbol{\theta}_{ij} &\sim G_j. \end{aligned} \tag{2.34}$$

Ce qui implique que :

$$G_0(\cdot) = \sum_{k=1}^{+\infty} \omega_k \delta_{\boldsymbol{\theta}_k^*}(\cdot) \quad \text{avec} \quad \boldsymbol{\omega} \sim \mathrm{GEM}(\alpha_0) \quad \text{et} \quad \boldsymbol{\theta}_k^* \sim H.$$

Puisque G_0 a son support sur l'ensemble $\boldsymbol{\Theta}^* = (\boldsymbol{\theta}_k^*)_{k=1}^{\infty}$, alors les G_j ont nécessairement le même support. On a donc pour tout j,

$$G_j(\cdot) = \sum_{k=1}^{+\infty} \pi_{jk} \delta_{\boldsymbol{\theta}_k^*}(\cdot) \quad \text{avec} \quad \boldsymbol{\pi}_j = (\pi_{jk})_{k=1}^{\infty} \sim \mathrm{GEM}(\alpha). \tag{2.35}$$

La séquence des poids (w_k) est construite de la façon suivante (stick-breaking) :

$$\omega_k' \sim \mathrm{Beta}(1, \alpha_0) \qquad \omega_k = \omega_k' \prod_{l=1}^{k-1} (1 - w_l').$$

La relation entre $\boldsymbol{\omega}$ et $\boldsymbol{\pi}_j$ est donnée par :

$$\pi_{jk}' \sim \mathrm{Beta}\left(\alpha\omega_k, \alpha\left(1 - \sum_{l=1}^{k} \omega_l\right)\right) \qquad \pi_{jk} = \pi_{jk}' \prod_{l=1}^{k-1} (1 - \pi_{lj}').$$

En comparant les équations 2.33 et 2.35, on voit que le HDP est un DDP où les atomes sont constants entre les groupes, seuls les poids varient. Les atomes sont donc partagés par les groupes.

Dans le HDP, l'analogie du processus du restaurant chinois du DP est la franchise du restaurant chinois. Dans cette métaphore, on a un ensemble de restaurants, chaque restaurant comportant un nombre infini de tables. Les restaurants ont le même menu composé d'un nombre infini de repas. Le premier client à s'asseoir à une table sélectionne un repas (à partir du menu) qui sera partagé par tous les clients qui s'y installeront. Plusieurs tables dans plusieurs restaurants peuvent servir le même repas. Les restaurants correspondent aux groupes G_j, les clients aux $\boldsymbol{\theta}_{ji}$ (client i dans restaurant j) et les valeurs uniques $\boldsymbol{\theta}_1^*, \boldsymbol{\theta}_2^* \cdots \sim G_0$ composent le menu global. L'effet de renforcement statistique du processus se traduit par le fait que les clients préfèrent s'asseoir aux tables les plus occupées, et auront une préférence pour les repas les plus plébiscités par les autres clients.

Teh *et al.* [TJBB06] étendent aussi l'approche aux mélanges, conduisant aux modèles de mélange par processus de Dirichlet hiérarchiques (notés HDPM pour Hierarchical

Dirichlet Process Mixtures). Dans le HDPM, $\boldsymbol{x}_{ij}$ (l'observation i du groupe j) suit une distribution paramétrée par $\boldsymbol{\theta}_{ij}$. Le modèle hiérarchique s'écrit :

$$\begin{array}{rcll} \boldsymbol{x}_{ij} & \sim & F(\boldsymbol{x}_{ij}|\boldsymbol{\theta}_{ij}) & \text{pour } i = 1, 2, \ldots, n \\ \boldsymbol{\theta}_{ij} & \sim & G_j & \text{pour } i = 1, 2, \ldots, n \\ G_j & \sim & \mathrm{DP}(\alpha, G_0) & \text{pour } j = 1, 2, \ldots \\ G_0 & \sim & \mathrm{DP}(\alpha_0, H). & \end{array} \tag{2.36}$$

2.4.2 Le processus de Dirichlet imbriqué

L'idée du processus de Dirichlet imbriqué dit aussi emboité (Nested Dirichlet Process (NDP), [RDG08]) est de modéliser une collection de distributions aléatoires dépendantes et de pouvoir les clusteriser.

Un ensemble de distributions $\{G_1, \ldots, G_J\}$ est dit suivre un NDP de paramètres α, β et H si

$$\begin{array}{rcl} \forall j,\ G_j & \sim & Q \\ Q & \sim & \mathrm{DP}(\alpha, \mathrm{DP}(\beta, H)). \end{array} \tag{2.37}$$

Cette définition signifie que

$$Q(\cdot) \overset{d}{=} \sum_{k=1}^{\infty} \pi_k \delta_{G_k^*}(\cdot)$$

où

$$\boldsymbol{\pi} \sim \mathrm{GEM}(\alpha) \quad \text{et} \quad G_k^* \sim \mathrm{DP}(\beta, H).$$

Ce qui implique que

$$G_k^*(\cdot) \overset{d}{=} \sum_{l=1}^{\infty} \omega_{lk} \delta_{\boldsymbol{\theta}_{lk}^*}(\cdot) \quad \text{avec} \quad \omega_{lk} \sim \mathrm{GEM}(\beta) \quad \text{et} \quad \boldsymbol{\theta}_{lk}^* \sim H.$$

Le NDP peut être étendu aux modèles de mélange, appelés mélange par processus de Dirichlet imbriqués (Nested Dirichlet Process Mixtures (NDPM)). Ainsi, tandis que dans un DPM, la loi *a priori* sur la distribution mélangeante est un DP (cf. équation 2.13), dans un NDPM la distribution mélangeante est une collection de distributions $\{G_1, \ldots, G_J\}$ dont la loi *a priori* est un NDP. On a alors

$$\begin{aligned} p_j(\boldsymbol{x}) &= \int_{\Theta} f(\boldsymbol{x}|\boldsymbol{\theta}) G_j(\mathrm{d}\boldsymbol{\theta}), \\ G_j(\cdot) &\sim \sum_{k=1}^{\infty} \pi_k \delta_{G_k^*}(\cdot) \\ G_k^*(\cdot) &= \sum_{l=1}^{\infty} \omega_{lk} \delta_{\boldsymbol{\theta}_{lk}^*}(\cdot). \end{aligned} \tag{2.38}$$

On peut ré-écrire le modèle 2.38 de la façon hiérarchique suivante :

$$\begin{array}{rcll} \boldsymbol{x}_{ij} & \sim & F(\boldsymbol{x}_{ij}|\boldsymbol{\theta}_{ij}) & \text{pour } i = 1, 2, \ldots, n \\ \boldsymbol{\theta}_{ij} & \sim & G_j & \text{pour } j = 1, 2, \ldots, J \\ \{G_1, \ldots, G_J\} & \sim & NDP(\alpha, \beta, H). & \end{array} \tag{2.39}$$

Puisque $Q(\cdot) = \sum_{k=1}^{\infty} \pi_k \delta_{G_k^*}(\cdot)$ est discrète, plusieurs G_j prendront simultanément la même valeur G_k^* pour un certain k, induisant un effet de clustering sur ces distributions. De plus, comme chaque G_k^* est aussi discrète, on peut clusteriser les observations ayant la même distribution G_k^* et le même paramètre $\boldsymbol{\theta}_{lk}^*$ pour un certain l.

Remarque 7. *Le HDP et le NDP induisent tous les deux des dépendances mais de deux façons différentes. D'une part, dans le HDP les distributions G_j partagent les mêmes atomes (localisations des composantes) mais pas les mêmes poids. Il y'a donc une probabilité nulle pour que deux distributions G_j et $G_{j'}$ soient égales. En conséquence, le clustering est seulement effectué via les atomes. D'autre part, la construction du NDP implique qu'il y'a une probabilité non-nulle pour que deux distributions soient égales. Cela entraîne un clustering à la fois via les distributions et via les paramètres.*

Remarque 8. *Pour la reconstruction spatio-temporelle en TEP 4D (chapitre 5), nous modélisons les cinétiques par un processus emboité mais qui sera différent du NDP. Les cinétiques décrivent l'évolution dans le temps de l'activité dans une zone cérébrale donnée. Notre but sera de clusteriser ces cinétiques afin d'identifier des volumes fonctionnels (zones du cerveau dont l'activité a le même comportement cinétique). Mais contrairement au modèle 2.39, dans notre modélisation le nombre de groupes (les cinétiques) ne sera pas fixe et la mesure de base du NDP ne sera pas un processus de Dirichlet mais un processus d'arbres de Pólya que nous allons maintenant aborder.*

2.5 Les arbres de Pólya

Le concept de l'arbre de Pólya (Pólya tree (PT)) fut introduit par Ferguson [Fer74] et développé plus tard par Lavine ([Lav92], [Lav94]) et Mauldin *et al.* [MSW92]. Il repose sur une partition aléatoire, récursive et dyadique de l'espace. L'avantage de l'arbre de Pólya sur le processus de Dirichlet est que ses paramètres peuvent être choisis de telle sorte que les mesures générées soient absolument continues presque-sûrement [Fer74] tandis que son inconvénient majeur est la dépendance par rapport à la partition de l'espace, résultant en des discontinuités dans la loi *a posteriori*.

Nous adoptons la définition de l'arbre de Pólya suivant Lavine [Lav92], mais légèrement modifiée.

Définition 8. *Soit $E = \{0, 1\}$, E^m le m produit cartésien $E \times \cdots \times E$ avec $E^0 = \emptyset$. Posons $E^\star = \cup_{m=0}^{\infty} E^m$ et $\pi_m = \{B_{\boldsymbol{\epsilon}} : \boldsymbol{\epsilon} \in E^m\}$ une partition de $\mathcal{X}$ telle que $\pi_0 = \{\mathcal{X}\}$ et $\Pi = \cup_{m=0}^{\infty} \pi_m$. Une mesure de probabilité aléatoire G sur $\mathcal{X}$ est dite distribuée suivant un arbre de Pólya s'il existe des nombres positifs $\mathcal{A} = \{\alpha_{\boldsymbol{\epsilon}} : \boldsymbol{\epsilon} \in E^\star\}$ et des variables aléatoires $\mathcal{W} = \{W_{\boldsymbol{\epsilon}0} : \boldsymbol{\epsilon} \in E^\star\}$ tels que :*

- *$\mathcal{W}$ est une séquence de variables aléatoires toutes indépendantes,*
- *$\forall\ \boldsymbol{\epsilon} \in E^\star$, $W_{\boldsymbol{\epsilon}0} \sim Beta(\alpha_{\boldsymbol{\epsilon}0}, \alpha_{\boldsymbol{\epsilon}1})$,*
- *$\forall\ m \in \mathbb{N}$ et $\boldsymbol{\epsilon} = \epsilon_1 \cdots \epsilon_m \in E^m$,*

$$G(B_{\epsilon_1 \cdots \epsilon_m}) = \left(\prod_{\substack{j=1 \\ \epsilon_j = 0}}^{m} W_{\epsilon_1 \cdots \epsilon_{j-1} 0} \right) \left(\prod_{\substack{j=1 \\ \epsilon_j = 1}}^{m} (1 - W_{\epsilon_1 \cdots \epsilon_{j-1} 0}) \right)$$

où les termes pour $j = 1$ sont interprétés comme W_0 et $1 - W_0$.

On note $G \sim \mathrm{PT}(\mathcal{A}, \Pi)$ et pour $\boldsymbol{\epsilon} \in E^\star$, on a $W_{\boldsymbol{\epsilon}0} = G\left(B_{\boldsymbol{\epsilon}0} | B_{\boldsymbol{\epsilon}}\right)$.

On peut imaginer un arbre binaire partitionnant l'espace $\mathcal{X}$ et des nombres positifs α sur chaque branche de l'arbre. La distribution de l'arbre de Pólya est obtenue en tirant des variables aléatoires Beta divisant la masse à chaque point de la branche. Plus précisément, soit $B = \{B_{\boldsymbol{\epsilon}}\}$ la partition binaire de l'arbre tel que pour tout $\boldsymbol{\epsilon}, B_{\boldsymbol{\epsilon}}$ est scindé en deux partitions $B_{\boldsymbol{\epsilon}0}$ et $B_{\boldsymbol{\epsilon}1}$ (i.e $B_{\boldsymbol{\epsilon}} = B_{\boldsymbol{\epsilon}0} \cup B_{\boldsymbol{\epsilon}1}$ et $B_{\boldsymbol{\epsilon}0} \cap B_{\boldsymbol{\epsilon}1} = \emptyset$). Ainsi au niveau 1, (B_0, B_1) est la partition de $\mathcal{X}$, au niveau 2 les ensembles sont B_{00}, B_{01}, B_{10} et B_{11} où (B_{00}, B_{01}) partitionnent B_0 tandis que (B_{10}, B_{11}) partitionnent B_1, et ainsi de suite. Au niveau m, le nombre de partitions est de 2^m. Pour le cas où $\mathcal{X} =]0, 1]$, un choix canonique des partitions est donné par les intervalles dyadiques. Pour $j = 0, 1, \ldots, 2^m - 1$, $\Pi = \{]j/2^m, (j+1)/2^m]\}$. D'où, $\mathcal{X} =]0, 1]$ est partitionné en $B_0 =]0, 0.5]$ et $B_1 =]0.5, 1]$ au niveau 1 tandis qu'au niveau 2, B_0 est scindé en $]0, 0.25]$ et $]0.25, 0.5]$, et B_1 en $]0.5, 0.75]$ et $]0.75, 1]$ etc.

Walker et Mallick [WM97] donnent une description imagée de l'arbre de Pólya. On peut imaginer une particule cascadant entre ces partitions. Au début, la particule est en $\mathcal{X}$ et se déplace en B_0 avec une probabilité W_0 ou en B_1 avec une probabilité $W_1 = 1 - W_0$ et ainsi de suite. Pour tout $\boldsymbol{\epsilon}$, la particule entre en $B_{\boldsymbol{\epsilon}0}$ avec une probabilité $W_{\boldsymbol{\epsilon}0}$ ou en $B_{\boldsymbol{\epsilon}1}$ avec la probabilité $W_{\boldsymbol{\epsilon}1} = 1 - W_{\boldsymbol{\epsilon}0}$. Les probabilités des lieux de déplacements sont des variables aléatoires Beta.

La définition de l'arbre de Pólya, bien que paraissant compliquée, est en fait assez simple. Prenons un exemple, soit à calculer $G(B_{010})$. On est donc au niveau $m = 3$, $\epsilon_1 = 0, \epsilon_2 = 1, \epsilon_3 = 0$. Pour avoir le premier produit dans la définition, on parcourt l'arbre jusqu'au niveau $m = 3$ et on prend tous les j tels que $\epsilon_j = 0$: ce sont ϵ_1 et ϵ_3. Pour $j = 1$, on a $W_{\epsilon_1 \cdots \epsilon_{j-1} 0} := W_0$. Pour $j = 3$, on a $W_{\epsilon_1 \cdots \epsilon_{j-1} 0} = W_{\epsilon_1 \epsilon_2 0} = W_{010}$. Le premier produit vaut donc $W_0 W_{010}$. On procède de la même façon pour le deuxième produit, on parcourt l'arbre et on prend tous les j tels que $\epsilon_j = 1$. Il n'y en a qu'un et c'est ϵ_2. Pour $j = 2$, on a donc $W_{\epsilon_1 \cdots \epsilon_{j-1} 0} = W_{\epsilon_1 0} = W_{00}$. Le deuxième produit vaut donc $1 - W_{00}$ et on a $G(B_{010}) = W_0(1 - W_{00})W_{010}$.

Remarque 9. *Nous avons présenté les arbres de Pólya dans un cadre unidimensionnel. Il existe aussi des arbres de Pólya multidimensionnels mais l'implantation est difficile en dimension supérieure à deux (voir [PRLW03]).*

2.5.1 Moments

La loi Beta dans la définition de l'arbre de Pólya permet d'avoir les deux premiers moments facilement.

En effet,

$$\begin{aligned}
\mathbb{E}\left[G(B_{\epsilon_1\cdots\epsilon_m})\right] &= \mathbb{E}\left[\prod_{\substack{j=1\\ \epsilon_j=0}}^{m} W_{\epsilon_1\cdots\epsilon_{j-1}0}\prod_{\substack{j=1\\ \epsilon_j=1}}^{m}(1-W_{\epsilon_1\cdots\epsilon_{j-1}0})\right] \\
&= \prod_{\substack{j=1\\ \epsilon_j=0}}^{m}\mathbb{E}\left[W_{\epsilon_1\cdots\epsilon_{j-1}0}\right]\prod_{\substack{j=1\\ \epsilon_j=1}}^{m}\mathbb{E}\left[1-W_{\epsilon_1\cdots\epsilon_{j-1}0}\right] \text{ par indépendance des } W_{\boldsymbol{\epsilon}}, \\
&= \prod_{\substack{j=1\\ \epsilon_j=0}}^{m}\frac{\alpha_{\epsilon_1\cdots\epsilon_{j-1}0}}{\alpha_{\epsilon_1\cdots\epsilon_{j-1}0}+\alpha_{\epsilon_1\cdots\epsilon_{j-1}1}}\prod_{\substack{j=1\\ \epsilon_j=1}}^{m}\frac{\alpha_{\epsilon_1\cdots\epsilon_{j-1}1}}{\alpha_{\epsilon_1\cdots\epsilon_{j-1}0}+\alpha_{\epsilon_1\cdots\epsilon_{j-1}1}} \\
&= \prod_{j=1}^{m}\frac{\alpha_{\epsilon_1\cdots\epsilon_j}}{\alpha_{\epsilon_1\cdots\epsilon_{j-1}0}+\alpha_{\epsilon_1\cdots\epsilon_{j-1}1}}.
\end{aligned} \tag{2.40}$$

De la même façon, on peut obtenir la variance :

$$\begin{aligned}
\mathbb{V}\left[G(B_{\epsilon_1\cdots\epsilon_m})\right] = &\prod_{j=1}^{m}\frac{\alpha_{\epsilon_1\cdots\epsilon_j}(\alpha_{\epsilon_1\cdots\epsilon_j}+1)}{(\alpha_{\epsilon_1\cdots\epsilon_{j-1}0}+\alpha_{\epsilon_1\cdots\epsilon_{j-1}1})+(\alpha_{\epsilon_1\cdots\epsilon_{j-1}0}+\alpha_{\epsilon_1\cdots\epsilon_{j-1}1}+1)} \\
&-\prod_{j=1}^{m}\frac{\alpha^2_{\epsilon_1\cdots\epsilon_j}}{(\alpha_{\epsilon_1\cdots\epsilon_{j-1}0}+\alpha_{\epsilon_1\cdots\epsilon_{j-1}1})^2}.
\end{aligned}$$

2.5.2 Propriétés

1. **Tail-free :**
   L'arbre de Pólya et le processus de Dirichlet appartiennent à une classe de processus plus généraux, les processus dits « *tail-free* » . Une mesure aléatoire G est « tail-free » par rapport aux partitions $\Pi = \cup_{m=0}^{\infty}\pi_m$ (spécifiées dans la Définition 8) si les familles $\{W_\emptyset\}, \{W_0, W_1\}, \{W_{00}, W_{01}, W_{10}, W_{11}\}\ldots$ sont indépendantes [Fre63]. L'arbre de Pólya est donc un cas spécial de processus tail-free où les $W_{\boldsymbol{\epsilon}}$ suivent une loi Beta et sont mutuellement indépendantes. Le processus de Dirichlet est un arbre de Pólya tel que $\alpha_{\boldsymbol{\epsilon}0} = \alpha_{\boldsymbol{\epsilon}1}$ pour tout $\boldsymbol{\epsilon}$ [Fer74]. L'avantage du DP sur des PT plus généraux est que le DP est le seul processus tail-free tel que le choix de la partition n'affecte pas l'inférence. L'avantage des PT est que sous certaines conditions sur les $\{\alpha_{\boldsymbol{\epsilon}}, \boldsymbol{\epsilon} \in E^\star\}$, les mesures générées sont presque-sûrement absolument continues par rapport à la mesure de Lebesgue et admettent donc des densités.
2. **Conjugaison :**
   L'arbre de Pólya hérite de la propriété de conjugaison qui caractérise les processus tail-free [Fer74].

   Théorème 9. *Soit* $\boldsymbol{X} = (X_1, \ldots, X_n)\,|G \overset{iid}{\sim} G$ *où* $G \sim PT(\mathcal{A}, \Pi)$. *Alors* $G|\boldsymbol{X} \sim PT(\mathcal{A}^*, \Pi)$ *avec*

$$\begin{aligned}
\mathcal{A}^* &= \{\alpha^*_\epsilon = \alpha_\epsilon + n_\epsilon \ : \epsilon \in E^*\} \\
\text{avec } n_\epsilon &= \#\left\{i \in \{1, \ldots, n\} \ : \ X_i \in B_\epsilon\right\}.
\end{aligned}$$

Démonstration. Voir [MSW92], [Lav92] [Lav94].

Utilisant le théorème 9 et l'équation 2.40, on peut en déduire la distribution prédictive :

$$\begin{aligned}\mathbb{P}\left[X_{n+1} \in B_{\epsilon_1\cdots\epsilon_m} | \boldsymbol{X}\right] &= \mathbb{E}\left[G(B_{\epsilon_1\cdots\epsilon_m}) | \boldsymbol{X}\right] \\ &= \frac{\alpha_{\epsilon_1} + n_{\epsilon_1}}{\alpha_0 + \alpha_1 + n} \prod_{j=2}^{m} \frac{\alpha_{\epsilon_1\cdots\epsilon_j} + n_{\epsilon_1\cdots\epsilon_j}}{\alpha_{\epsilon_1\cdots\epsilon_{j-1}0} + \alpha_{\epsilon_1\cdots\epsilon_{j-1}1} + n_{\epsilon_1\cdots\epsilon_{j-1}}}.\end{aligned}$$

2.5.3 Les arbres de Pólya canoniques

En pratique, il peut être difficile de spécifier les paramètres $\mathcal{A}$ et la partition Π. On peut alors centrer l'arbre de Pólya autour d'une mesure de probabilité G_0. Considérons par exemple le cas où $\mathcal{X} = \mathbb{R}$, $\Omega = [\Omega^-,\ \Omega^+] \subset \mathbb{R}$ et G_0 une mesure de probabilité sur Ω, dont la fonction de répartition est $Q_0(t)$. A chaque niveau m de l'arbre, les partitions sont prises de telle sorte à coïncider avec les quantiles de la distribution et on prend $\alpha_{\boldsymbol{\epsilon}0} = \alpha_{\boldsymbol{\epsilon}1}$ pour tout $\boldsymbol{\epsilon} \in E^*$. Les éléments de la partition sont donc les intervalles

$$\left]Q_0^{-1}(j/2^m),\ Q_0^{-1}((j+1)/2^m)\right] \text{ pour } j = 0, 1, \ldots, 2^m - 1$$

avec $Q_0^{-1}(0) = \Omega^-$ et $Q_0^{-1}(1) = \Omega^+$. On a alors : au niveau 1, $B_0 =]Q_0^{-1}(0),\ Q_0^{-1}(1/2)]$ et $B_1 =]Q_0^{-1}(1/2),\ Q_0^{-1}(1)]$. Au niveau 2, $B_{00} =]Q_0^{-1}(0),\ Q_0^{-1}(1/4)]$, $B_{01} =]Q_0^{-1}(1/4),\ Q_0^{-1}(1/2)]$, $B_{10} =]Q_0^{-1}(1/2),\ Q_0^{-1}(3/4)]$ et $B_{11} =]Q_0^{-1}(3/4),\ Q_0^{-1}(1)]$. Et ainsi de suite... Cette construction est appelée *arbre de Pólya canonique* et est illustrée dans la Figure 2.6. Dans cette construction, les éléments de partition sont déterminés par la mesure G_0. On peut alors adopter une notation duale où on remplace Π par G_0 dans la paramétrisation. On notera alors $G \sim \mathrm{PT}(\mathcal{A}, G_0)$.
Dans ce contexte, si $G \sim \mathrm{PT}(\mathcal{A}, G_0)$ alors $\mathbb{E}(G) = G_0$. En effet pour tout $\boldsymbol{\epsilon} = \epsilon_1 \cdots \epsilon_m$,

$$\mathbb{E}\left[G(B_{\boldsymbol{\epsilon}})\right] = \prod_{j=1}^{m} \mathbb{E}\left[W_{\epsilon_1\cdots\epsilon_j}\right] = \prod_{j=1}^{m} \frac{\alpha_{\epsilon_1\cdots\epsilon_j}}{\alpha_{\epsilon_1\cdots\epsilon_{j-1}0} + \alpha_{\epsilon_1\cdots\epsilon_{j-1}1}} = 2^{-m} \text{ car } \alpha_{\boldsymbol{\epsilon}0} = \alpha_{\boldsymbol{\epsilon}1} \text{ pour tout } \boldsymbol{\epsilon}.$$

D'un autre côté, de par la définition de la fonction de répartition, on a

$$G_0(B_{\boldsymbol{\epsilon}}) = Q_0\left(Q_0^{-1}((j+1)/2^m)\right) - Q_0\left(Q_0^{-1}(j/2^m)\right) = 2^{-m}.$$

La mesure G_0 a donc un rôle similaire que dans le DP. Une fois que le PT est centré autour de G_0, ce sont les paramètres $\{\alpha_\epsilon\}$ qui déterminent les variations des mesures générées autour de G_0, tout comme le paramètre α du DP.

2.5.4 Paramétrisation de l'arbre de Pólya

Comme dans l'exemple qui précède, la plupart du temps dans la construction d'un arbre de Pólya on n'affecte pas une valeur différente de $\alpha_{\boldsymbol{\epsilon}}$ pour chaque $\boldsymbol{\epsilon}$. Pour éviter de spécifier beaucoup trop de paramètres, on prend plutôt $\alpha_{\boldsymbol{\epsilon}}$ dépendant de la longueur de la chaîne $\boldsymbol{\epsilon}$, i.e $\alpha_{\epsilon_1\ldots\epsilon_m} = a_m$ à chaque niveau m. Le choix de $\alpha_{\boldsymbol{\epsilon}}$ affecte la continuité des mesures générées et ces paramètres peuvent être choisis de façon à exprimer

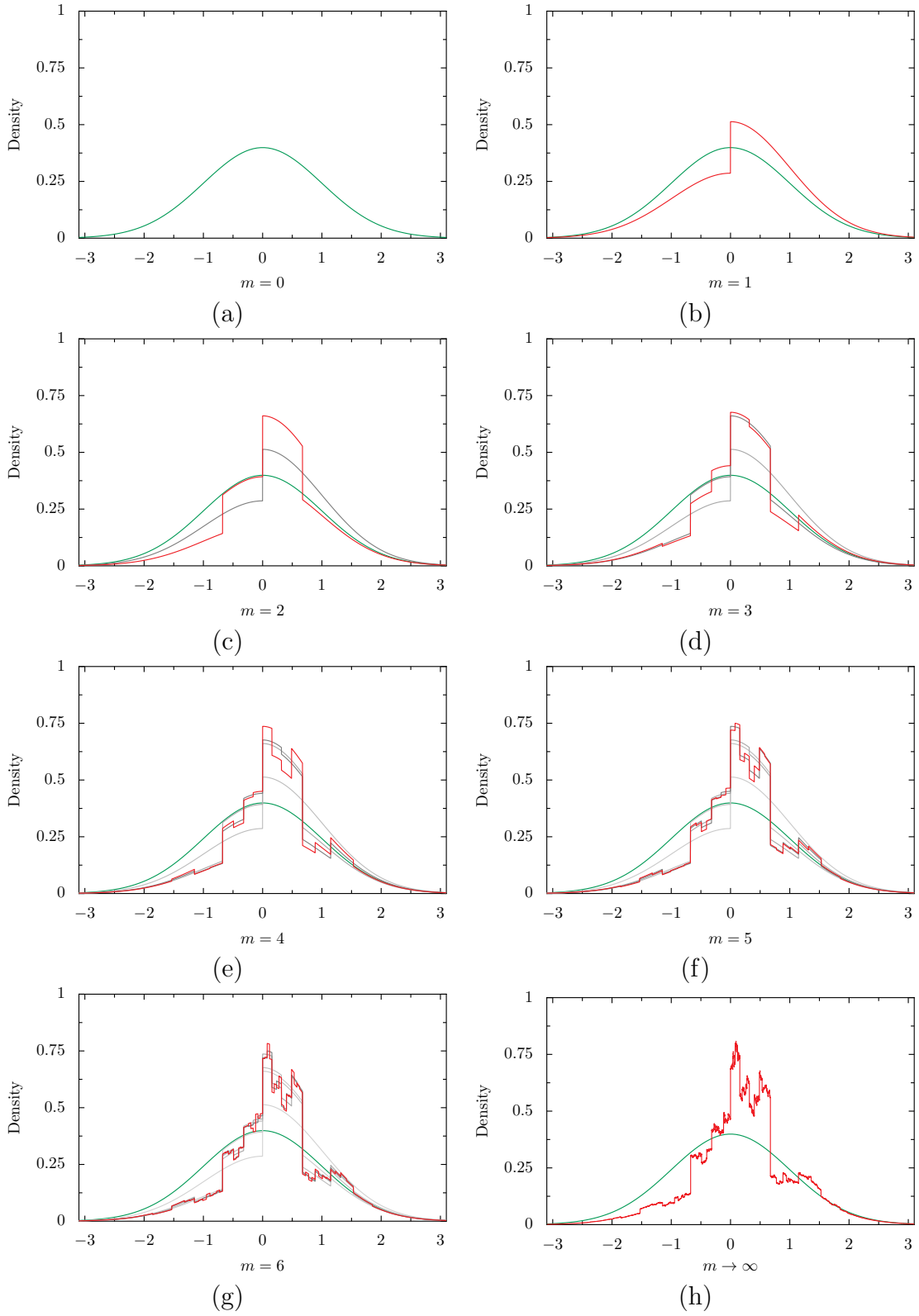


FIGURE 2.6 – Séquence de construction d'un arbre de Pólya avec $\mathcal{A} = \{\alpha_m = 3^m\}$ et $Q_0(t) = \phi(t)$, où $\phi(x)$ désigne la fonction de répartition de la loi $\mathcal{N}(0,1)$.) (a)-(g) : itérations $m = 1, \ldots, 6$, (h) à l'itération $m \to \infty$, on obtient une réalisation d'un $\mathrm{PT}(\{\alpha_m = 3^m\}, \mathcal{N}(0,1))$.

la force de l'*a priori* sur G. Par exemple, si l'on s'attend à ce que G soit continue, alors $G(B_{\epsilon 0})$ et $G(B_{\epsilon 1})$ doivent être proches pour des grandes valeurs de m. Ferguson [Fer74] a montré que prendre $a_m = m^2$ conduit à des mesures presque-sûrement absolument continues. C'est aussi le choix canonique opté par Lavine [Lav92]. Le processus de Dirichlet est lui obtenu en prenant $a_m = \frac{1}{2^m}$. Walker et Mallick [WM97] tout comme Paddock *et al.* [PRLW03] considèrent $a_m = cm^2$ où $c > 0$. Berger et Guglielmi [BG01] ont considéré cette famille et d'autres de la forme $a_m = c\rho(m)$, spécifiquement $\rho(m) = m^2, m^3, 2^m, 4^m, 8^m$. Le choix $\rho(m) = 8^m$ satisfait un théorème de consistance de l'arbre de Pólya d'après Barron *et al.* [BSW99]. En prenant différentes valeurs de c, Hanson et Johnson [HJ02] ont trouvé que la famille $a_m = c\, m^2$ est suffisamment riche pour capturer des caractéristiques intéressantes des distributions considérées. Ils ont prouvé deux résultats qui indiquent l'effet de c sur l'inférence. Si on définit $G(t|\boldsymbol{X}, G_0, c) := G\{]-\infty, t]\}|\boldsymbol{X}, G_0, c$, alors

$$G(t|\boldsymbol{X}, G_0, c) \xrightarrow{c\to\infty} G_0(t)$$
$$G(t|\boldsymbol{X}, G_0, c) \xrightarrow{c\to 0} \sum_{j=1}^{n} p_j \mathbb{I}_{]-\infty,t]}(X_j)$$

où $(p_1, \ldots, p_n) \sim \mathrm{Dir}(1, \ldots, 1)$. Ces résultats indiquent que l'on peut utiliser c pour exprimer la force sur la loi *a priori* : les grandes valeurs de c reflètent la croyance que la vraie distribution des observations F_x est similaire à G_0 tandis que les petites valeurs de c indiquent un manque de confort vis-à-vis de G_0. Ces résultats sont analogues à ceux obtenus dans le DP suivant les valeurs de α (cf. Théorème 1).

2.5.5 Les arbres de Pólya finis

De façon théorique, la construction d'un arbre de Pólya implique de partitionner l'espace indéfiniment. Cependant en pratique, puisque l'on ne peut pas mettre à jour un nombre infini de paramètres, on ne considère que ceux jusqu'à un certain niveau L. Cela conduit aux *arbres de Pólya finis*, appelés aussi *arbres de Pólya partiellement spécifiés* [Lav94]. On les note $G \sim \mathrm{PT}_L(\Pi, \mathcal{A})$ ou $G \sim \mathrm{PT}(\Pi_L, \mathcal{A}_L)$. Hanson et Johnson [HJ02] recommandent de choisir $L = \log_2 n$ où n est la taille de l'échantillon. En permettant à L d'augmenter avec la taille de l'échantillon, on permet à la mesure G de s'accommoder quand le nombre de données disponibles augmente. L'idée est la même que de décroître la largeur du bin dans un histogramme quand les données augmentent. En choisissant L de cette façon, il y'a *a priori* une observation en moyenne dans chaque ensemble B_ϵ au niveau L. Si l'on a suffisamment confiance dans la distribution de base G_0, Hanson [Han06] suggère $J \approx \log 2(n/N)$ où N est un nombre typique d'observations dans chaque ensemble au niveau J.

2.5.6 Processus dérivés d'arbres de Pólya

Comme nous l'avons déjà souligné, les mesures de probabilité générées par un PT dépendent des partitions pré-spécifiées (typiquement les partitions canoniques dyadiques). Ces partitions fixées contribuent aux limitations pratiques de l'arbre de Pólya. Ferguson [Fer74] souligne que l'arbre de Pólya permet de générer des mesures absolument continues par rapport à la mesure de Lebesgue mais que les densités présentent des

discontinuités à toutes les extrémités des ensembles de la partition (cela peut d'ailleurs être constaté sur la Figure 2.6). Cet inconvénient contribue à limiter l'utilisation des arbres de Pólya en estimation de densité.

Pour gommer ces effets, Paddock *et al.* [PRLW03] ont proposé les *arbres de Pólya randomisés* où la partition est supposée aléatoire. La construction est basée sur les partitions dyadiques comme dans l'arbre de Pólya simple mais plutôt que de choisir les milieux des ensembles comme points de coupure dans la partition, ils sont choisis aléatoirement à côté d'eux.

Une autre solution est basée sur les *mélanges d'arbres de Pólya* (notés MPT pour Mixtures of Pólya Trees, [Lav92], [HJ02]). Tout comme dans le mélange de processus de Dirichlet (cf. chapitre 2.1.6), dans le MPT les paramètres de la distribution moyenne G_0 et/ou de la famille $\mathcal{A}$ dépendent d'un paramètre aléatoire ψ. Cela permet plus de flexibilité puisque d'une part, dans un PT simple il peut être difficile de spécifier une seule distribution autour de laquelle centrer la mesure aléatoire. D'autre part, le choix d'une distribution G_0 non adaptée peut rendre très lente la convergence de G vers la vraie loi ([HJ02], [BSW99]). On peut donc considérer une mesure contenant des paramètres non spécifiés $G_0(\psi)$, et de mettre une loi *a priori* sur ces hyperparamètres. Le mélange d'arbres de Pólya revient à considérer le modèle hiérarchique suivant

$$\begin{aligned} G|\psi &\sim \mathrm{PT}(\Pi^{\psi}, \mathcal{A}^{\psi}) \\ \psi &\sim \mu(\psi). \end{aligned}$$

Plusieurs types de mélanges ont été proposés dans la littérature. On peut citer entre autres Lavine [Lav92], Berger et Guglielmi [BG01], Hanson et Johnson [HJ02].

Nous proposons les arbres de Pólya décalés (notés SPT pour Shifted Pólya Trees). Ce sont des arbres de Pólya perturbés, en déplaçant l'ensemble des partitions de l'arbre dans un espace plus grand afin d'atténuer les effets de la partition fixe dans l'inférence *a posteriori*. Cette technique fut introduite par Barat *et al.* [BDM07] pour l'application en spectrométrie nucléaire. Nous présentons ici une version simplifiée et plus générale.

Les arbres de Pólya décalés

Les arbres de Pólya décalés sont un mélange d'arbres de Pólya finis basé sur une construction canonique pour laquelle les quantiles sont décalés par un décalage (*shift*) aléatoire. Cette technique a pour effet de perturber la partition afin de tempérer les discontinuités aux frontières. Plus précisément, considérons un arbre de Pólya canonique fini d'ordre L sur $\Omega = [\Omega^-, \Omega^+] \subset \mathcal{X}$, noté $\mathrm{PT}_L(\mathcal{A}, G_0)$ défini comme dans les paragraphes 2.5.3 et 2.5.5. Soit Q_0 la fonction de répartition de G_0. Nous notons $\epsilon_{1:m} = \epsilon_1 \cdots \epsilon_m$ pour $m > 0$. La construction canonique permet de spécifier la partition

$$\Pi_L = \left\{ B_{\epsilon_{1:m}} = \left] Q_0^{-1}\left(\frac{j}{2^m}\right), Q_0^{-1}\left(\frac{j+1}{2^m}\right) \right] : 0 < m \leq L, 0 \leq j < 2^m \right\}$$

avec $j = \sum_{l=1}^{m} \epsilon_l 2^{m-l}$, ainsi que les coefficients d'échelle tels que $\alpha_{\epsilon_{1:m}} = a_m$ pour tout m.

L'arbre de Pólya décalé d'ordre $L+1$, noté $\mathrm{PT}^{\delta}_{L+1}(\mathcal{A}, G_0)$, est un mélange construit de la façon suivante.

Définition 9. *Soit δ une variable aléatoire entière uniforme sur $\{0, 2^L\}$,*

$$\delta \sim \mathcal{U}(\{0, 2^L\}).$$

On définit pour tout x, $[\![x]\!]_0^1 \triangleq \min(\max(x, 0), 1)$ et

$$\Pi_{L+1}^{\delta} = \left\{ B_{\epsilon_{1:m}}^{\delta} : 0 < m \leq L+1 \right\}$$

avec pour $m > 0$,

$$B_{\epsilon_{1:m}}^{\delta} = \left] Q_0^{-1}\left(\left[\!\left[\frac{j}{2^{m-1}} - \frac{\delta}{2^L} \right]\!\right]_0^1 \right), Q_0^{-1}\left(\left[\!\left[\frac{j+1}{2^{m-1}} - \frac{\delta}{2^L} \right]\!\right]_0^1 \right) \right]$$

avec $j = \sum_{l=1}^{m} \epsilon_l 2^{m-l}$ et $B_{\emptyset}^{\delta} = \Omega$.
On définit de même les coefficients d'échelle $\mathcal{A}^{\delta}$ pour $m > 0$

$$\alpha_{\epsilon_{1:m}}^{\delta} = 2^{m-1} a_{m-1} \left(\left[\!\left[\frac{j+1}{2^{m-1}} - \frac{\delta}{2^L} \right]\!\right]_0^1 - \left[\!\left[\frac{j}{2^{m-1}} - \frac{\delta}{2^L} \right]\!\right]_0^1 \right).$$

Remarque 10. *Si $G \sim PT_{L+1}^{\delta}(\mathcal{A}, G_0)$, alors $\mathbb{E}(G) = G_0$.*

Remarque 11. *Si l'on indice les blocs $B_{\epsilon_{1:L+1}}^{\delta}$ par j de 0 à $2^{L+1} - 1$ tel que $B_{\epsilon_{1:L+1}}^{\delta} = \mathcal{B}_{L+1,j}^{\delta}$, alors pour tout $0 \leq \delta \leq 2^L$ et pour tout $0 \leq l < 2^L$, $\mathcal{B}_{L+1,l+\delta}^{\delta} = \mathcal{B}_{L,l}$ où $\mathcal{B}_{L,l} = B_{\epsilon_{1:L}}$, le bloc l de la partition d'ordre L de l'arbre de Pólya canonique. Par ailleurs, les blocs $\mathcal{B}_{L+1,l}^{\delta}$ d'indice $l < \delta$ et $l \geq \delta + 2^L$ sont de masse nulle.*

Ces observations simplifient grandement l'interprétation du SPT d'ordre $L+1$. En effet, celui-ci peut ainsi être vu comme l'inclusion des partitions d'un arbre de Pólya canonique d'ordre L ($\mathrm{PT}_L(\mathcal{A}, G_0)$) dans des partitions où le nombre d'intervalles est doublé et dont le positionnement est aléatoire. De cette façon, la génération des arbres décalés peut aisément se faire à partir de la génération de δ et d'un tirage de PT d'ordre $L+1$ munis des paramètres garantissant la construction canonique. Ce mélange par décalage permet de mitiger les discontinuités en frontières de blocs. Notons que la propriété qui permet de retrouver pour tout δ la partition à l'ordre L de l'arbre canonique par simple sélection des blocs de la partition d'ordre $L+1$ du SPT, confère au modèle une simplicité d'inférence identique à celle de l'arbre de Pólya sans mélange.

2.6 Conclusion

Dans ce chapitre, nous avons rappelé les principales caractéristiques de certains modèles bayésiens non paramétriques qui peuvent être utilisés comme *a priori* dans des espaces infini-dimensionnels. Nous avons présenté le processus de Dirichlet, le processus de Pitman-Yor, les processus de Dirichlet hiérarchique et imbriqué ainsi que l'arbre de Pólya. Comme nous l'avons souligné en préambule, l'objectif n'était pas d'être exhaustif mais plutôt de présenter les modèles dont il sera question dans la suite de ce manuscrit. Par ailleurs, au vu de l'évolution rapide sur le sujet, il semble difficile voire impossible de passer en revue tous les modèles existants. En effet, le bayésien non paramétrique est un sujet actif de recherche aussi bien dans la communauté statistique qu'en machine

learning et plusieurs nouvelles distributions aléatoires ne cessent d'émerger. Parmi les modèles existants dans la littérature, on peut citer entre autres le processus gaussien, le processus Beta, les processus neutres à droite etc. Pour une revue plus complète, on peut citer par exemple les articles de [WDLS99], [MQ04] [CGR05], ainsi que deux livres sur le sujet [GR03] [HHMW10]. Toutes ces lois *a priori* ainsi que leurs combinaisons et extensions fournissent un large choix d'outils très riches pour la modélisation statistique.

Toutefois, soulignons que l'approche bayésienne non paramétrique pose certains problèmes aussi bien en théorie que dans l'application (cf. [BSW99], [GGR99], [CGR05]).

1. La première difficulté est le choix de la loi *a priori*. En effet, l'abord bayésien d'un problème implique de définir des distributions *a priori* sur l'espace des paramètres et d'inférer sur les lois *a posteriori*. Dans l'approche non paramétrique, l'espace des paramètres est infini dimensionnel. Par exemple, dans un problème d'estimation de densité, cela consiste en l'espace de toutes les densités de probabilité dans l'espace considéré. Dans un problème de régression linéaire, cela peut être l'espace de toutes les fonctions dans $\mathbb{R}^d$, etc. L'une des difficultés de l'approche bayésienne non paramétrique réside donc dans la spécification de la loi *a priori*. Dans une approche paramétrique, la loi *a priori* est choisie de façon subjective selon la connaissance qu'on a des paramètres. À défaut d'une connaissance concrète, on peut choisir des lois *a priori* non informatives. Dans l'approche non paramétrique, une connaissance subjective n'est pas possible car l'espace des paramètres est beaucoup trop vaste et un *a priori* non-informatif est difficile à construire à cause de l'absence de la mesure de Lebesgue. La subjectivité est donc souvent introduite au travers des hyperparamètres de la loi définissant par exemple les paramètres de localisation de la distribution.
2. Une autre difficulté est de générer sur un calculateur des objets aléatoires de dimension infinie. Les méthodes MCMC usuelles ne s'appliquent plus puisque l'espace des paramètres est infini-dimensionnel. Cela nécessite donc de développer des techniques d'échantillonnage spécifiques, dépendant à la fois du problème à traiter et de la loi *a priori* sur le paramètre fonctionnel (voir chapitre 3). Au final, le problème du calcul se retrouve lié au choix de l'*a priori*.
3. Enfin, l'une des difficultés théoriques de l'approche bayésienne non paramétrique est que les propriétés de consistance asymptotique et de convergence des estimateurs doit être étudiée avec attention. La consistance est une notion importante pour valider le choix d'un modèle *a priori*. En effet, considérons des observations $(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n) \overset{iid}{\sim} f(\cdot|\boldsymbol{\theta})$, $\boldsymbol{\theta} \in \Theta$ avec Θ désignant un espace de paramètres [4]. Dans les statistiques bayésiennes, on présume d'une loi *a priori* sur $\boldsymbol{\theta}$: $\boldsymbol{\theta} \sim \Pi$, et on considère la loi *a posteriori* $\Pi(\boldsymbol{\theta}|\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n)$. Soit $\boldsymbol{\theta}_0 \in \Theta$ la vraie valeur du paramètre. La consistance garantit que pour n assez grand, la loi *a posteriori* est concentrée autour de $\boldsymbol{\theta}_0$. Plus précisément, soient $\epsilon > 0$, $d(\cdot, \cdot)$ une distance (ou une pseudo-distance) sur Θ et $A_\epsilon = \{\boldsymbol{\theta} \in \Theta, \ d(\boldsymbol{\theta}_0, \boldsymbol{\theta}) < \epsilon\}$. La distribution *a posteriori* est *consistante* en $\boldsymbol{\theta}_0$ si

$$\Pi(A_\epsilon|\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n) \longrightarrow 1 \ \text{ p.s sous } P_{\boldsymbol{\theta}_0}.$$

   On peut aussi définir la vitesse de convergence asymptotique de la façon suivante. La loi *a posteriori* converge asymptotiquement à *vitesse* ϵ_n vers $\boldsymbol{\theta}_0$ s'il existe

4. Dans les modèles non paramétriques, le paramètre est infini-dimensionnel et peut être la fonction f elle-même.

$(\epsilon_n)_{n\geq 0} \to 0$ telle que :

$$\Pi(A_{\epsilon_n}|\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n) \longrightarrow 1 \quad \text{p.s ou en proba. sous } P_{\boldsymbol{\theta}_0}.$$

Dans les problèmes paramétriques (en dimension finie), la distribution *a posteriori* est consistante si la loi *a priori* n'exclut pas la vraie valeur du paramètre. Ceci est une conséquence du fait que la vraisemblance est très piquée autour de la vraie valeur du paramètre si la taille de l'échantillon devient grande. Cependant, dans les problèmes en dimension infinie, une telle conclusion est fausse et la consistance *a posteriori* doit être vérifiée avant d'utiliser un *a priori*. Les aspects sur la consistance et la convergence ne seront pas étudiés dans cette thèse. Soulignons quil existe cependant des résultats garantissant la convergence et la consistance pour certains modèles, suivant la distance utilisée puisque l'on est dans des espaces de dimension infinie et qu'on n'a plus l'équivalence des normes. Pour plus de détails sur le sujet, on pourra se référer par exemple au livre [GR03] et les références y figurant.

Les propriétés théoriques et pratiques des méthodes bayésiennes non paramétriques sont de mieux en mieux connues grâce la popularité grandissante de ces méthodes, ce qui en fait un sujet de recherche actif. Grâce aux développements pratiques, les applications des méthodes BNP sont devenues populaires pour modéliser des phénomènes complexes dans les domaines biomédical, environnemental, économétrique etc.

3

Méthodes MCMC

On appelle méthode Monte-Carlo toute procédure visant à calculer une valeur numérique en utilisant un processus aléatoire. Les méthodes de Monte Carlo par chaînes de Markov (notées MCMC pour Markov Chain Monte Carlo) sont des algorithmes de simulation qui permettent de générer des séquences d'échantillons, réalisations de chaînes de Markov, telles que ces échantillons soient distribués suivant une certaine loi d'intérêt. Les méthodes MCMC sont très utilisées dans le cadre de l'estimation bayésienne.

Cadre statistique bayésien

Considérons, par exemple, la relation simplifiée $\boldsymbol{y} = \mathcal{H}\boldsymbol{x}$ [1] où $\mathcal{H}$ est un opérateur linéaire, $\boldsymbol{y} = y_1, y_2, \cdots, y_M$ représentant le vecteur des observations et $\boldsymbol{x} = x_1, \cdots, x_N$ celui des inconnues. Le but de l'inversion est de retrouver $\boldsymbol{x}$ à partir de $\boldsymbol{y}$.

Dans les statistiques bayésiennes, on présume d'une distribution *a priori* sur les inconnues du modèle (ici le vecteur $\boldsymbol{x}$) qui sont donc vues comme des variables aléatoires. C'est la loi *a priori* c'est-à-dire que nous avons une idée de comment ces paramètres devraient être distribués. On note cette distribution *a priori* $p(\boldsymbol{x})$. Après avoir observé les données $\boldsymbol{y}$, nous modifions notre croyance sur la distribution de $\boldsymbol{x}$ et formons la distribution *a posteriori* de $\boldsymbol{x}$ sachant les observations en utilisant la règle de Bayes :

$$\pi(\boldsymbol{x}|\boldsymbol{y}) = \frac{p(\boldsymbol{y}|\boldsymbol{x})p(\boldsymbol{x})}{p(\boldsymbol{y})}. \tag{3.1}$$

où $p(\boldsymbol{y}) = \int p(\boldsymbol{y}|\boldsymbol{x})p(\boldsymbol{x})\mathrm{d}\boldsymbol{x}$. Cette intégrale ne dépendant pas des inconnues $\boldsymbol{x}$ est considérée comme une constante multiplicative et on écrit

$$\pi(\boldsymbol{x}|\boldsymbol{y}) \propto p(\boldsymbol{y}|\boldsymbol{x})p(\boldsymbol{x}). \tag{3.2}$$

1. Cette relation s'écrit souvent en prenant en compte les erreurs d'observations associées sous la forme $\boldsymbol{y} = \mathcal{H}\boldsymbol{x} + \boldsymbol{\epsilon}$ où $\boldsymbol{\epsilon}$ représente le M–vecteur des erreurs.

La loi $p(\boldsymbol{y}|\boldsymbol{x})$ est appelée vraisemblance. Elle représente la distribution des observations si $\boldsymbol{x}$ était connu.

Une fois la distribution *a posteriori* connue, on calcule un estimateur de $\boldsymbol{x}$ suivant cette loi. L'estimateur du maximum *a posteriori* est donné par :

$$\hat{\boldsymbol{x}}_{\mathrm{MAP}} = \underset{\boldsymbol{x}}{\operatorname{argmax}}\ \pi(\boldsymbol{x}|\boldsymbol{y}), \tag{3.3}$$

et l'estimateur de la moyenne *a posteriori* est :

$$\hat{\boldsymbol{x}}_{\mathrm{PM}} = \mathbb{E}(\boldsymbol{X}|\boldsymbol{Y}) = \int \boldsymbol{x}\,\pi(\boldsymbol{x}|\boldsymbol{y})\mathrm{d}\boldsymbol{x}. \tag{3.4}$$

Les estimateurs de type maximum *a posteriori* (MAP) et moyenne *a posteriori* (PM) nécessitent donc respectivement de faire appel à des techniques d'optimisation et d'intégration mais deux problèmes se posent. D'une part, la très grande dimensionalité des variables considérées (pixels d'une image, échantillons d'un signal etc.) rend très difficile les calculs d'intégration et l'optimisation. Une solution pour le calcul de l'équation 3.4 serait alors de l'approcher par la moyenne de N réalisations $\boldsymbol{x}^{(1)}, \boldsymbol{x}^{(2)}, \ldots, \boldsymbol{x}^{(N)}$ de la loi *a posteriori* si ces réalisations sont extraites d'une suite ergodique pour la moyenne,

$$\hat{\boldsymbol{x}}_{\mathrm{PM}} \approx \frac{1}{N}\sum_{k=1}^{N} \boldsymbol{x}^{(k)} \qquad \text{pour } N \text{ assez grand.}$$

Néanmoins, cela suppose de pouvoir générer des échantillons suivant la loi *a posteriori* π. Dans l'estimation bayésienne, il est souvent impossible d'avoir une expression analytique de cette loi car elle est souvent connue à une constante multiplicative près difficilement calculable (à cause de l'intégrale qui apparaît dans le dénominateur). On fait donc appel à des techniques d'approximation (approximation de Laplace, méthodes variationnelles, méthodes MCMC etc.). Les méthodes MCMC sont des méthodes de simulation numérique qui permettent de générer une chaîne de Markov convergeant en distribution vers la loi d'intérêt.

La théorie des chaînes de Markov et le cadre de convergence des algorithmes sont abordés dans l'annexe B. Nous allons maintenant présenter deux méthodes MCMC à savoir l'algorithme de Metropolis-Hastings et l'échantillonneur de Gibbs. Ces méthodes permettent de construire un noyau de transition (cf. annexe B) tel que les séquences d'échantillons générés forment une chaîne de Markov qui converge vers la loi cible. Les méthodes MCMC sont très utiles car elles permettent de simuler des distributions multivariées, complexes et non standard.

3.1 Méthodes MCMC

3.1.1 Problématique

La loi cible (loi d'intérêt) n'est pas connue ou est difficilement simulable. Dans le cadre bayésien, il s'agit de la distribution *a posteriori* connue à une constante multiplicative près. L'algorithme MCMC construit un noyau de transition dont la distribution invariante est la loi cible. Il convient alors de s'assurer de la convergence de la chaîne de Markov produite par l'algorithme vers la loi cible. Pour cela, il faut vérifier des conditions d'irréductibilité, de récurrence et d'apériodicité sur le noyau (cf. annexe B).

3.1.2 Algorithme de Metropolis-Hastings

L'algorithme de Metropolis-Hastings (MH) est un schéma de simulation permettant de générer des échantillons suivant une densité de probabilité $\pi(x)$ dont une forme analytique est disponible ou connue à une constante multiplicative près indépendante de x. Cette méthode fut d'abord utilisée par Metropolis *et al.* [MRR+53] pour la résolution de problèmes en mécanique statistique avant d'être élargie à un cadre plus général de simulation statistique par Hastings [Has70].

L'algorithme utilise un principe d'acceptation et de rejet de variables aléatoires générées par une loi instrumentale appelée aussi loi candidate. Le schéma de simulation est présenté dans le tableau 3.1.

Schéma d'échantillonnage de l'algorithme de Metropolis-Hastings

1. itération 1 : générer $X_0 \sim \pi_0$
2. itération t+1 :
   - Sachant $X^{(t)} = x^{(t)}$, proposer un candidat pour $x^{(t+1)}$ i.e générer $Y^{(t)} \sim q(y|x^{(t)})$. On note $y^{(t)}$ la valeur obtenue.
   - Calculer le rapport
   $$\alpha(x^{(t)}, y^{(t)}) = \min\{1, \frac{\pi(y^{(t)})q(x^{(t)}|y^{(t)})}{\pi(x^{(t)})q(y^{(t)}|x^{(t)})}\}.$$
   - tirer $U^{(t)} \sim \mathcal{U}[0,1]$.
     On note $u^{(t)}$ la valeur obtenue.
     Si $u^{(t)} < \alpha(x^{(t)}, y^{(t)})$ alors $X^{(t+1)} = y^{(t)}$, sinon $X^{(t+1)} = x^{(t)}$.
3. $t \leftarrow t+1$ et retour en 2.

TABLE 3.1 – Algorithme MH

Plusieurs remarques peuvent être faites sur l'algorithme MH.

- Le schéma de simulation montre qu'à chaque itération, un échantillon $y^{(t)}$ est simulé en fonction de l'échantillon précédent. Les variables aléatoires $(y^{(t)})_t$ ainsi produites par l'algorithme MH ne sont pas indépendantes.
- L'algorithme MH est spécifié par la loi candidate $q(\cdot)$ dont le choix est important. Le schéma de simulation est d'autant plus efficace que $q(y|x^{(t)})$ est simulable rapidement et est proche de π.
- La probabilité d'acceptation $\rho = \frac{\pi(y^{(t)})q(x^{(t)}|y^{(t)})}{\pi(x^{(t)})q(y^{(t)}|x^{(t)})}$ est définie que si $\pi(x^{(t)}) > 0$. Mais si on initialise la chaîne à une valeur $x^{(0)}$ telle que $\pi(x^{(0)}) > 0$ alors $\pi(x^{(t)}) > 0$ pour tout $t \in \mathbb{N}$ car les candidates $y^{(t)}$ telles que $\pi(y^{(t)}) = 0$ seront systématiquement rejetées ($\rho = 0$).
- La loi cible $\pi(\cdot)$ n'intervient que pour calculer la probabilité ρ et apparaît à la fois au numérateur et au dénominateur. La loi d'intérêt π peut alors juste être connue à une constante multiplicative près, ce qui est pratique pour les problèmes d'estimation bayésienne.

Dans l'algorithme MH, comme dans toutes les méthodes MCMC, les échantillons provenant des premières itérations de l'algorithme ne sont en général pas utilisées dans l'estimation des fonctionnelles. Elles correspondent à une période dite de *chauffage* ou *burn-in*, période jugée nécessaire pour que la chaîne s'extraie de ses conditions initiales.

L'algorithme de Metropolis-Hastings permet de simuler des échantillons suivant une loi cible π à partir d'une loi candidate q vérifiant des hypothèses assez faibles (cf. annexe B, paragraphe B.2). Grosso modo, une loi candidate q assez quelconque dans un support connexe permet de simuler n'importe quelle loi dans ce support. Toutefois, cette propriété d'universalité peut être dommageable en pratique. En effet, l'algorithme de simulation ne fonctionnera pas bien si, par exemple, les zones de forte probabilité de la densité q sont surtout localisées dans les queues de la distribution de π. Le choix de q est un problème complexe, faisant l'objet de plusieurs discussions dans la littérature ([Rob96], [GRS96]). Typiquement, q est choisie dans une famille de distributions qui permet de spécifier les paramètres de réglage (paramètres de location et d'échelle) de telle sorte à s'assurer du bon balayage du support de la loi cible.

Nous allons illustrer l'importance du réglage des paramètres de la loi candidate dans le cadre de l'algorithme de Metropolis-Hastings à marche aléatoire.

Algorithme de Metropolis-Hastings à marche aléatoire

C'est un cas particulier de l'algorithme MH où la variable candidate $y^{(t)}$ est générée dans un voisinage de la valeur courante $x^{(t)}$ suivant la relation :

$$y^{(t)} = x^{(t)} + \epsilon_t,$$

où ϵ_t est une perturbation aléatoire de loi q, indépendante de $x^{(t)}$. Puisque la valeur candidate est égale à la valeur courante à laquelle on rajoute du bruit, cet algorithme est appelé marche aléatoire. La loi candidate q s'écrit alors : $q(y|x) = q(y - x) = q(\epsilon)$. Si q est symétrique ($q(z) = q(-z)$), le rapport d'acceptation devient

$$\alpha(x^{(t)}, y^{(t)}) = \min\{1, \frac{\pi(y^{(t)})}{\pi(x^{(t)})}\} \tag{3.5}$$

et on tombe sur l'algorithme de Metropolis *et al.* [MRR+53].

Si q est positive dans un voisinage de 0, la chaîne est ergodique (cf. conclusion du théorème 17). Cela signifie qu'a priori, pour n'importe quelle loi candidate vérifiant cette condition, l'algorithme va converger et générer des échantillons suivant π. Mais nous allons voir que la vitesse de convergence dépend crucialement des paramètres de la loi candidate. Pour l'illustrer, nous considérons un exemple où la loi cible π est $\mathcal{N}(0, 1)$ et prenons trois lois de proposition :

1. $q(\cdot|x) = \mathcal{N}(x, 0.5)$,
2. $q(\cdot|x) = \mathcal{N}(x, 0.1)$,
3. $q(\cdot|x) = \mathcal{N}(x, 10)$.

Les résultats sont montrés à la Figure 3.1.

On peut constater que la chaîne 1 (en haut) « mélange bien », c'est-à-dire qu'elle bouge de façon fluide dans l'espace d'états. Bien qu'initialisé à une valeur extrême

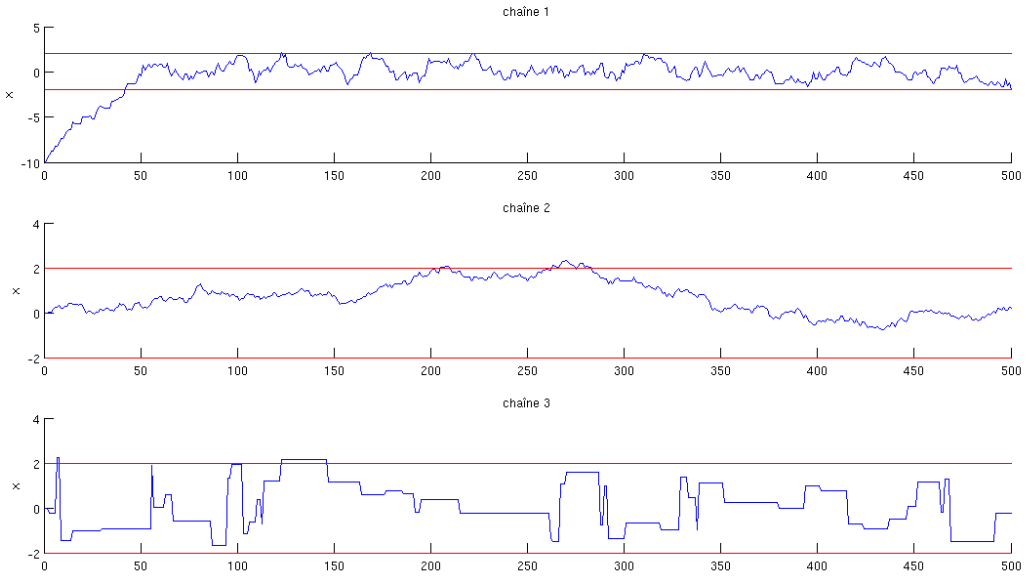


FIGURE 3.1 – 500 itérations de l'algorithme de MH à marche aléatoire. La loi cible est $\mathcal{N}(0,1)$ et les lois de proposition sont : 1) $q(\cdot|x) = \mathcal{N}(x, 0.5)$ (en haut) ; 2) $q(\cdot|x) = \mathcal{N}(x, 0.1)$ (au milieu) ; 3) $q(\cdot|x) = \mathcal{N}(x, 10)$ (en bas).

(-10), l'algorithme converge rapidement. Le support de la loi cible (délimité par les bandes rouges sur la figure) est balayé en 100 itérations. Au contraire, les chaînes 2 et 3 mélangent mal malgré une initialisation au mode de la loi cible.

Cet exemple simple illustre l'importance du choix du paramètre d'échelle de la loi candidate (ici la variance) dans l'exploration du support de la loi cible :

1. si la variance est trop faible, la loi de proposition génère des valeurs de $y^{(t)}$ proches de la valeur courante $x^{(t)}$. Le taux d'acceptation est élevé mais la chaîne est mal mélangeante. C'est le cas de la chaîne 2.
2. Si en revanche la variance est trop élevée, les candidates $y^{(t)}$ générées sont trop loin de la valeur courante $x^{(t)}$. Le taux d'acceptation est donc souvent faible et la chaîne mal mélangeante. La figure montre plusieurs itérations où la chaîne 3 ne bouge pas et recopie les valeurs précédentes.

La loi de proposition doit donc être réglée de façon à éviter ces deux cas extrêmes.

Nous allons maintenant aborder un deuxième algorithme d'échantillonnage MCMC à savoir l'échantillonneur de Gibbs.

3.1.3 Échantillonnage de Gibbs

L'algorithme de Metropolis-Hastings peut présenter des difficultés à explorer convenablement le support de la loi cible. Ces difficultés sont d'autant plus marquées lorsque la loi est multidimensionnelle car le nombre d'échantillons nécessaires pour avoir une couverture suffisante du support est alors très important. L'algorithme d'échantillonnage

de Gibbs est particulièrement adapté à la simulation de lois multidimensionnelles. Pour simuler suivant la loi cible, l'échantillonneur de Gibbs exploite ses distributions conditionnelles si elles existent. Pour résumer, l'algorithme d'échantillonnage de Gibbs est utilisé pour simuler $\pi(\boldsymbol{x})$ quand :

1. $\boldsymbol{x}$ admet une décomposition de la forme :

$$\boldsymbol{x} = (x_1, \ldots, x_N), \tag{3.6}$$

2. les lois conditionnelles

$$\pi_i(x_i|x_{-i}) \tag{3.7}$$

sont simulables facilement, avec la notation $x_{-i} = (x_1, x_2, \ldots, x_{i-1}, x_{i+1}, \ldots x_N)$ et $\pi_i(x_i|x_{-i})$ désigne la densité conditionnelle de x_i sachant x_{-i}.

L'algorithme est présenté dans le tableau 3.2.

Schéma de l'algorithme d'échantillonnage de Gibbs

1. itération $t = 0$, initialiser $x^{(0)} \sim \pi_0(x)$,
2. itération t, tirer

$$\begin{aligned}
x_1^{(t+1)} &\sim \pi_1(x_1|x_2^{(t)}, \ldots, x_N^{(t)}) \\
x_2^{(t+1)} &\sim \pi_2(x_2|x_1^{(t+1)}, x_3^{(t)} \ldots, x_N^{(t)}) \\
&\vdots \\
x_i^{(t+1)} &\sim \pi_i(x_i|x_1^{(t+1)}, \ldots, x_{i-1}^{(t+1)}, x_{i+1}^{(t)} \ldots, x_N^{(t)}) \\
&\vdots \\
x_N^{(t+1)} &\sim \pi_N(x_N|x_1^{(t+1)}, \ldots, x_{N-1}^{(t+1)})
\end{aligned}$$

3. $t \leftarrow t + 1$ et aller en (2).

TABLE 3.2 – Échantillonneur de Gibbs.

L'échantillonneur de Gibbs fut introduit par Geman et Geman [GG84] pour le traitement d'image, puis généralisé à une variété de problèmes conventionnels en Statistique par Gelfand et Smith [GS90]. L'algorithme de Gibbs est décrit en détails dans le livre de Robert [Rob96, paragraphe 5]. Pour une description simple et rapide, on pourra se referer à Casella et George [CG92].

Remarques

– L'échantillonneur de Gibbs nécessite la connaissance et la possibilité d'échantillonner suivant les lois conditionnelles au contraire de l'algorithme MH.

- Seules les lois conditionnelles sont utilisées pour la simulation. Donc pour un problème de grande dimension, toutes les simulations sont univariées. En fait, on verra dans le paragraphe qui suit que les composantes x_i ne sont pas forcément scalaires.
- À l'inverse de l'algorithme MH, ici tous les échantillons simulés sont acceptés (taux d'acceptation=1). L'échantillonneur de Gibbs est la composition de N algorithmes MH avec des probabilités d'acceptation uniformément égales à 1.
- Le schéma de simulation de l'algorithme de Gibbs est, par construction, multidimensionnel. Cet algorithme est donc seulement applicable aux modèles comportant au moins deux variables aléatoires. Dans certains cas, il est nécessaire de considérer et de simuler des variables artificielles pour l'implantation.
- Comme nous le verrons plus tard, l'échantillonneur de Gibbs est particulièrement bien adapté aux modèles hiérarchiques.

Échantillonnage par blocs, complétion, marginalisation

Nous avons présenté l'échantillonneur de Gibbs comme un algorithme consistant en la simulation de N lois univariées pour générer un paramètre N-dimensionnel. Dans certains cas, l'algorithme peut être rendu plus efficace en échantillonnant conjointement plusieurs variables, c'est l'*échantillonnage par blocs*. Par exemple, quand deux composantes sont fortement corrélées dans la loi cible π, le mélange peut être lent en utilisant une mise à jour s'effectuant coordonnée par coordonnée. Échantillonner ces deux composantes corrélées de façon jointe dans un seul bloc peut alors améliorer le mélange. De façon générale, l'échantillonnage par blocs permet d'accélérer la convergence de l'algorithme, surtout quand l'on traite des variables de grandes dimensions.

La simulation peut aussi être facilitée par la *complétion* du modèle qui consiste à introduire des variables supplémentaires. On dit que la densité g est une complétion de f si f est une loi marginale de g. L'intérêt d'une complétion du modèle est que les lois conditionnelles de g sont parfois plus simples à simuler que celles de f. C'est souvent le cas dans les modèles bayésiens hiérarchiques. Cette complétion peut être naturelle, cela signifie que les variables introduites ont un sens physique et on parle de *variables de complétion*. Un exemple est un modèle avec des données manquantes. La complétion peut aussi être artificielle, au sens où les variables rajoutées ne sont qu'utilitaires et n'ont pas de signification physique. Elles servent alors uniquement à simplifier la simulation des variables d'intérêt : on parle de *variables auxiliaires*. De tels algorithmes sont développés dans cette thèse comme on le verra dans la partie sur l'inférence des modèles de mélange de processus de Dirichlet (paragraphe 3.2).

Enfin, dans certaines situations, le mélange dans l'échantillonneur de Gibbs peut être rendu plus efficace en marginalisant certaines variables. Ce sont les méthodes *marginales* dites aussi algorithmes de Gibbs *collapsés*. Elles reposent sur la marginalisation analytique d'un ou de plusieurs paramètres du modèle. Ces paramètres pouvant être des paramètres de nuisance ou des variables d'intérêt. Si les variables marginalisées sont de grande dimension, le mélange peut être plus efficace puisque l'espace des paramètres est alors réduit de façon drastique.

Nous verrons des exemples de ces trois types d'échantillonneurs dans le paragraphe 3.2 sur l'inférence des DPM.

Remarque 12. *On souligne qu'il existe des algorithmes MCMC dits hybrides. Ce sont des algorithmes utilisant simultanément des étapes d'échantillonnage de Gibbs et des étapes de l'algorithme de Metropolis-Hastings. Leur utilisation est motivée par le fait que dans l'échantillonnage de Gibbs, certaines lois conditionnelles peuvent être impossibles à simuler. On peut donc remplacer chaque étape i où une simulation suivant la loi conditionnelle $\pi_i(x_i|\boldsymbol{x}_{-i})$ est impossible par une étape MH. Les méthodes que nous avons développées pour la reconstruction des images TEP utilisent ce principe (chapitres 4 et 5).*

3.1.4 Estimation des fonctionnelles

Les échantillons générés après convergence de la chaîne permettent de calculer des fonctions des quantiles. Dans l'inférence bayésienne, on résume souvent la distribution *a posteriori* des paramètres d'intérêt par des fonctionnelles tels que la moyenne *a posteriori*, la médiane, la variance et les corrélations, les intervalles de crédibilité etc. Si $\boldsymbol{X} = (X_1, \ldots, X_N)$ est la variable d'intérêt alors,

1. l'espérance marginale de X_i est estimée par :
$$\mathbb{E}(X_i) = \bar{X}_i \approx \frac{1}{n - n_0 - 1} \sum_{t=n_0+1}^{n} X_i^{(t)}$$
où n_0 est le nombre d'itérations burn-in et n le nombre total d'itérations.
2. La variance marginale de X_i s'obtient par :
$$\mathbb{V}(X_i) \approx \frac{1}{n - n_0 - 1} \sum_{t=n_0+1}^{n} (X_i^{(t)} - \bar{X}_i)^2.$$
3. Un intervalle de crédibilité à $100(1-2p)\%$, noté $[c_p, c_{1-p}]$, pour X_i peut être estimé en prenant respectivement pour c_p et c_{1-p} les quantiles d'ordre p et $1-p$ de $\{X_i^{(t)}, t = n_0 + 1, \ldots, n\}$.

L'estimation correcte de ces quantités demande que les échantillons générés après le burn-in représentent de façon adéquate la distribution d'intérêt. Cela nécessite donc de s'assurer que la chaîne ait bien mis à jour les spécificités de la loi cible π. Cela pose néanmoins plusieurs questions en pratique que nous allons évoquer maintenant.

3.1.5 Diagnostic de convergence des algorithmes MCMC

Sous certaines conditions que nous avons évoquées dans l'annexe B, les algorithmes MCMC convergent vers la loi cible : c'est le théorème ergodique. Autrement dit, si on laisse tourner l'algorithme « suffisamment » longtemps, les échantillons générés convergent vers la loi cible π. Mais le théorème ergodique ne précise pas le nombre d'itérations nécessaires pour que la chaîne génère des échantillons distribués suivant π, ni ne précise une estimée de l'erreur commise. La mise en œuvre d'une méthode MCMC nécessite donc de répondre à certaines questions pratiques pour s'assurer de la convergence de l'algorithme. Ces questions sont discutées par exemple dans [Rob96, chapitre 6], [GRS96] et [Tie94] et nous allons en récapituler les principales.

1. Détermination du burn-in :
   La question qui se pose en premier lieu est celle du nombre d'itérations à effectuer avant que la chaîne n'atteigne le régime stationnaire. Si la chaîne est bien mélangeante, la période de burn-in est assez courte. En revanche, une chaîne mal mélangeante conduit à un burn-in beaucoup plus long. Dans l'exemple 3.1, un burn-in de 100 itérations suffit pour la chaîne 1 tandis que pour la chaîne 3, le burn-in doit être supérieur à 500. Toutefois, la période de burn-in est assez difficile à déterminer en pratique et ce, particulièrement en grande dimension.
   – La méthode la plus évidente et la plus intuitive pour déterminer le burn-in consiste en l'inspection de la chaîne produite sur plusieurs initialisations. Cela consiste à visualiser sur plusieurs replicats et avec des poins initiaux différents, l'évolution de $\{X(t)\}$ en fonction du temps. L'itération n_0 à partir de laquelle la chaîne ne bouge plus (ou très peu) est prise comme burn-in. Toutefois cette approche permet, au mieux, de détecter des non-stationnarités fortes.
   – Une approche plus efficace repose sur les moyennes empiriques. À l'analyse de la série brute, on préfère généralement substituer celle des moyennes cumulées,

$$\bar{X}_i = \frac{1}{T}\sum_{t=1}^{T} X_i^{(t)}.$$

   Une condition nécessaire de convergence est alors la stationnarité de $\bar{X}_i$.

2. Initialisation :
   Une méthode MCMC nécessite de choisir un point de départ de la chaîne. En principe, si la chaîne est irréductible, le choix de l'initialisation n'affecte pas la distribution stationnaire. Mais la vitesse de convergence de l'algorithme peut être fortement affectée par cette initialisation. La convergence sera lente si la simulation reste pendant plusieurs itérations dans une région fortement influencée par la distribution initiale. Une chaîne dont la vitesse de mélangeance est élevée (chaîne 1 dans l'exemple 3.1) va rapidement s'extraire des conditions initiales même si le point de départ est choisi à une valeur extrême. En revanche si la chaîne mélange lentement, l'initialisation doit être choisie de façon plus attentive afin d'éviter un burn-in trop long. Dans les problèmes multidimensionnels, le choix de l'initialisation est crucial. Une loi cible multimodale dont les modes sont séparés par des vallées profondes (zones de faible probabilité) peut conduire à des chaînes qui mélangent mal et qui restent bloquées dans une région de l'espace d'états pendant de longues périodes. Il est recommandé d'effectuer un certain nombre de lancements de l'algorithme, avec des valeurs initiales très dispersées, et de contrôler que les estimations ne sont pas sensibles au choix des valeurs initiales. Cependant, des valeurs initiales extrêmes peuvent aussi conduire à un burn-in très long. Dans les méthodes bayésiennes, l'initialisation est typiquement choisie aléatoirement à partir de la loi *a priori* ou effectuée près d'un mode de la distribution *a posteriori*. Mais initialiser la simulation près du mode de la loi *a posteriori* n'est pas une garantie de succès si la chaîne ne bouge pas de façon fluide autour du support de la loi *a posteriori*.

3. Nombre de chaînes :

   En pratique, se pose aussi la question du nombre de chaînes à considérer dans l'algorithme MCMC. Ce choix est sujet à des recommandations non nécessairement concordantes dans la littérature.

- Plusieurs chaînes : cela consiste à lancer plusieurs chaînes en parallèle et de contrôler la convergence vers la loi stationnaire en comparant les estimations des quantités d'intérêt sur les différentes chaînes. Pour cela, on considère plusieurs réalisations de l'algorithme dont chacune est simulée avec un faible nombre d'itérations. Ce point de vue part du principe qu'il est pratiquement impossible de diagnostiquer la convergence d'une chaîne de Markov à partir d'une seule trajectoire et qu'une chaîne peut avoir convergé alors qu'en réalité ce n'est pas le cas. Cela se produit par exemple si la chaîne est restée très longtemps dans le voisinage de son point de départ. L'idée de simuler plusieurs chaînes suivant des valeurs initiales variées est de permettre de réduire la dépendance aux conditions initiales et de contrôler plus facilement la convergence vers la loi stationnaire en comparant les estimations des quantités d'intérêt sur les différentes chaînes.

  La reproche faite à cette approche est que la comparaison de plusieurs courtes chaînes ne prouve pas la convergence car il existe des distributions dont certaines spécificités ne sont visibles qu'avec un très grand nombre d'itérations. Cela est d'autant plus vrai dans le cas d'une distribution multimodale où les trajectoires souvent nécessitent un nombre d'itérations assez important pour passer d'un mode de la distribution à un autre. Dans ce cas, on préfère utiliser une seule chaîne puisqu'alors une chaîne unique de taille MT à faible taux de mélangeance aura probablement une plus grande proximité avec la loi stationnaire que M chaînes de taille T, qui auront tendance à demeurer dans le voisinage de leur point de départ.

- Une seule chaîne très longue : ce sont les méthodes à chaîne unique. Dans cette approche, une seule très longue simulation de l'algorithme est utilisée pour calculer les estimateurs Monte-Carlo. Avec cette chaîne, on a plus de chances de trouver de nouveaux modes avec et de visiter les zones de faible probabilité.

Au final si on a plusieurs processeurs, la « solution » serait de lancer plusieurs longues chaînes sur chaque processeur.

4. D'autres questions supplémentaires se posent comme par exemple la bonne exploration du support de la loi cible. L'idée la plus simple pour s'en assurer consiste à visualiser plusieurs trajectoires pour chaque composante X_i. Quant à la détermination du temps d'arrêt n de l'algorithme, il est conseillé de lancer plusieurs chaînes en parallèle avec différentes initialisations et de comparer les estimées. Si elles diffèrent trop, on peut augmenter n.

Pour finir, rappelons que le problème de la vitesse de convergence n'est pas spécifique aux algorithmes stochastiques. C'est aussi un problème dans les méthodes déterministes, par exemple dans l'algorithme EM (Expectation-Maximization). Dans EM, on peut diagnostiquer la convergence en surveillant par exemple l'augmentation de la vraisemblance ou effectuer plusieurs lancements de l'algorithme avec des initialisations différentes et vérifier s'ils convergent au même point ou à des solutions multiples. De façon générale, cette approche s'applique aux techniques Monte-Carlo mais avec des difficultés supplémentaires. L'algorithme est stochastique et on ne peut s'attendre à d'aucune quantité statistique qui accroît ou décroît de façon monotone. De plus, la convergence a lieu vers une distribution et non vers un point. Enfin, la convergence lente peut aussi être due à un modèle inapproprié.

Après avoir présenté les méthodes MCMC dans un cadre paramétrique, nous allons maintenant aborder l'échantillonnage dans un cadre non paramétrique.

3.2 Inférence dans les modèles de mélange par processus de Dirichlet

L'avantage de l'approche bayésienne non paramétrique est de pourvoir la modélisation d'une grande flexibilité en permettant d'incorporer l'incertitude sur des fonctions. Cependant, cette flexibilité est aux dépens de la complexité calculatoire puisque les distributions *a posteriori* sur des espaces fonctionnels sont plus complexes. L'inférence nécessite donc des outils spécifiques. De ce fait, bien qu'ils furent formalisés dans les années 60-70, l'utilisation des processus de Dirichlet (DP) en pratique n'a été possible qu'au début des années 90, après le développement de techniques d'échantillonnage pour l'inférence basées sur les méthodes Monte-Carlo par chaînes de Markov (MCMC). En effet, l'inférence analytique étant impossible dans les modèles de mélange par processus de Dirichlet (DPM), on a recours à des techniques d'approximation basées sur les MCMC, en particulier l'échantillonneur de Gibbs.

Nous avons vu dans le chapitre 2 qu'il existait plusieurs façons de définir le processus de Dirichlet : processus du restaurant chinois (CRP), représentation en urne de Pólya (urne de Blackwell-MacQueen), représentation stick-breaking etc. Ces différentes façons de définir le même processus ont conduit à différentes techniques d'échantillonnage des DPM. Dans ce chapitre, nous présentons les principales méthodes d'échantillonnage des DPM. La liste des algorithmes qui seront décrits ici est loin d'être exhaustive. Nous espérons juste que cela permettra d'avoir une vue générale de ces techniques d'échantillonnage.

Les méthodes d'échantillonnage de Gibbs pour les modèles de mélange par processus de Dirichlet peuvent être divisées en deux classses :

1. **les méthodes marginales :** comme leur nom l'indique, elles marginalisent la distribution aléatoire H du DPM dans la loi *a posteriori* (dans le modèle hiérarchique) et utilisent soit la représentation de l'urne de Blackwell-MacQueen pour échantillonner les paramètres $\boldsymbol{\theta}$, soit le processus du restaurant chinois pour l'échantillonnage des variables indicatrices des classes.
2. **les méthodes conditionnelles :** contrairement aux méthodes marginales, les méthodes conditionnelles représentent la distribution H explicitement en utilisant par exemple la construction stick-breaking. Parmi elles, on trouve les méthodes qui approximent le DP en tronquant le nombre de composantes et d'autres qui n'utilisent pas de troncature sèche.

Nous présentons ici quelques unes des principales méthodes d'échantillonnage dans chacune de ces deux classes. Nous commençons d'abord par décrire les méthodes marginales que nous divisons en deux catégories selon qu'on utilise des lois *a priori* conjuguées ou pas. Pour finir, nous présentons les méthodes conditionnelles. Nous discutons des limites de ces différents types d'échantillonneurs et de l'utilité de développer une nouvelle méthode d'échantillonnage.

NB : Pour plus de simplicité dans la présentation, on présente les algorithmes en

considérant le paramètre α et les hyperparamètres dans la mesure de base G_0 fixés. Mais ces paramètres peuvent aussi être à estimer comme on le verra plus tard.

Notations : Nous commençons par rappeler les notations qui seront utilisées tout au long de cette partie. Les observations sont $\boldsymbol{X} = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n\}$. À ces observations, sont associés des paramètres latents $\boldsymbol{\Theta} = \{\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_n\}$. La propriété de clustering des DP implique que certains $\boldsymbol{\theta}_i$ seront identiques. On note $\boldsymbol{\Theta}^* = \{\boldsymbol{\theta}_1^*, \boldsymbol{\theta}_2^* \ldots\}$ les valeurs uniques parmi $\boldsymbol{\Theta} = \{\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_n\}$. Le terme « cluster » désignera un ensemble de données ayant la même valeur de paramètre. Les variables indicatrices c_i dénotent l'appartenance des données aux clusters et sont telles que $c_i = k$ si et seulement si $\boldsymbol{\theta}_i = \boldsymbol{\theta}_k^* = \boldsymbol{\theta}_{c_i}^*$. L'entier n_k désignera le nombre d'observations affectées au cluster k. Comme le nombre de composantes est potentiellement infini, il est impossible de toutes les représenter pendant l'inférence. On désignera par « composantes actives » les composantes qui ont des données d'affectées et « composantes dormantes » celles qui sont vides. Le terme « composante singleton » sera utilisé pour désigner une composante comportant une seule observation. Le nombre de composantes actives sera noté K_n et le nombre de composantes représentées K^*.

3.2.1 Méthodes marginales

On les appelle aussi les méthodes de Gibbs collapsées. Ce sont les premières techniques qui ont été développées pour l'inférence dans les modèles de mélange par processus de Dirichlet. Le premier algorithme fut introduit dans le cadre de lois conjuguées dans la thèse non publiée d'Escobar en 1988, puis sera publié plus tard dans [Esc94] et [EW95]. Dans ces méthodes, la mesure aléatoire H est marginalisée et on utilise la représentation en urne de Pólya pour tirer des échantillons suivant H (les paramètres $\boldsymbol{\theta}$). La convergence de l'algorithme d'Escobar étant lente, différentes contributions eurent lieu pour l'accélérer ([Nea91], [WME94], [Mac94], [BM96]). Elles reposent sur le processus du restaurant chinois et génèrent les variables de classification. Toutes ces méthodes précédemment citées ont été développées dans le cadre de lois conjuguées, c'est-à-dire que la vraisemblance des données et la distribution de base G_0 sont conjuguées [2]. Dans ce cas, les calculs des lois *a posteriori* conditionnelles sont plus simples et peuvent être faits analytiquement. Cependant dans le cas non-conjugué, ces calculs sont difficiles à effectuer, l'échantillonnage est plus difficile et nécessite des techniques plus élaborées [MM98], [WD98], [Nea00] etc.

Échantillonnage dans le cadre de lois conjuguées

Le premier algorithme développé par Escobar pour l'échantillonnage des DPM est aussi le plus simple ([Esc94], [EW95]). Si on suppose que les clusters sont numérotés consécutivement suivant leur *ordre d'arrivée*, la loi *a priori* des $\boldsymbol{\theta}_i$ est donnée par l'urne de Blackwell-MacQueen. Cette représentation garantissant l'échangeabilité des observations et des clusters, on peut traiter chaque observation i comme si c'était la dernière des n échantillons. La distribution conditionnelle *a priori* de $\boldsymbol{\theta}_i$ sachant tous les autres

2. On parle de conjugaison quand la loi *a priori* et la loi *a posteriori* appartiennent à la même famille de distributions.

paramètres $\boldsymbol{\theta}_j, j \neq i$ (qu'on notera $\boldsymbol{\theta}_{-i}$) s'écrit alors :

$$\boldsymbol{\theta}_i|\boldsymbol{\theta}_{-i} \sim \frac{\alpha}{\alpha+n-1}G_0 + \frac{1}{\alpha+n-1}\sum_{j=1,j\neq i}^{n}\delta_{\boldsymbol{\theta}_j}.$$

La loi *a posteriori* conditionnelle (qui sera utilisée par l'échantillonneur de Gibbs) s'obtient avec la règle de Bayes, en combinant cette loi *a priori* avec la vraisemblance de $\boldsymbol{x}_i$, $f(\boldsymbol{x}_i|\boldsymbol{\theta}_i)$:

$$\begin{aligned} H(\boldsymbol{\theta}_i|\boldsymbol{x}_i,\boldsymbol{\theta}_{-i}) &\propto f(\boldsymbol{x}_i|\boldsymbol{\theta}_i)\left(\frac{\alpha}{\alpha+n-1}G_0(\boldsymbol{\theta}_i) + \frac{1}{\alpha+n-1}\sum_{j\neq i}\delta_{\boldsymbol{\theta}_j}(\boldsymbol{\theta}_i)\right) \\ &\propto \alpha \int f(\boldsymbol{x}_i|\boldsymbol{\theta}_i)G_0(\mathrm{d}\boldsymbol{\theta}_i)\,\frac{f(\boldsymbol{x}_i|\boldsymbol{\theta}_i)G_0(\boldsymbol{\theta}_i)}{\int f(\boldsymbol{x}_i|\boldsymbol{\theta}_i)G_0(\mathrm{d}\boldsymbol{\theta}_i)} + \sum_{j\neq i} f(\boldsymbol{x}_i|\boldsymbol{\theta}_i)\delta_{\boldsymbol{\theta}_j}(\boldsymbol{\theta}_i) \\ &= \frac{\alpha}{Z}\int f(\boldsymbol{x}_i|\boldsymbol{\theta}_i)G_0(\mathrm{d}\boldsymbol{\theta}_i)\,\frac{f(\boldsymbol{x}_i|\boldsymbol{\theta}_i)G_0(\boldsymbol{\theta}_i)}{\int f(\boldsymbol{x}_i|\boldsymbol{\theta}_i)G_0(\mathrm{d}\boldsymbol{\theta}_i)} + \frac{1}{Z}\sum_{j\neq i} f(\boldsymbol{x}_i|\boldsymbol{\theta}_i)\delta_{\boldsymbol{\theta}_j}(\boldsymbol{\theta}_i), \end{aligned} \tag{3.8}$$

où Z est le facteur de normalisation tel que la somme des probabilités vaille 1,

$$Z = \alpha\int f(\boldsymbol{x}_i|\boldsymbol{\theta}_i)G_0(\mathrm{d}\boldsymbol{\theta}_i) + \sum_{j\neq i} f(\boldsymbol{x}_i|\boldsymbol{\theta}_j).$$

Soient

$$H_i = \frac{f(\boldsymbol{x}_i|\boldsymbol{\theta}_i)G_0(\boldsymbol{\theta}_i)}{\int f(\boldsymbol{x}_i|\boldsymbol{\theta}_i)G_0(\mathrm{d}\boldsymbol{\theta}_i)}, q_0 = \frac{\alpha}{Z}\int f(\boldsymbol{x}_i|\boldsymbol{\theta}_i)G_0(\mathrm{d}\boldsymbol{\theta}_i) \text{ et } q_j = \frac{f(\boldsymbol{x}_i|\boldsymbol{\theta}_j)}{Z}. \tag{3.9}$$

On peut donc ré-écrire la loi *a posteriori* conditionnelle de $\boldsymbol{\theta}_i$:

$$\boldsymbol{\theta}_i \begin{cases} \sim H_i \text{ avec la probabilité } q_0 \\ = \boldsymbol{\theta}_j \text{ avec la probabilité } q_j. \end{cases}$$

La mise à jour des paramètres s'effectue en échantillonnant de façon répétée suivant la distribution *a posteriori* conditionnelle de $\boldsymbol{\theta}_i$. Le schéma global de l'algorithme est présenté dans le tableau 3.3.

L'avantage de l'algorithme 1 est sa simplicité dans la mise en œuvre. Il produit une chaîne de Markov ergodique, cependant la convergence vers la distribution *a posteriori* peut être lente. Cela se produit lorsque q_j est supérieur à q_0 et plusieurs observations sont alors associées au même paramètre. Dans ce cas, l'échantillonneur peut rester bloqué sur les valeurs uniques et il peut prendre plusieurs itérations avant qu'une nouvelle valeur ne soit générée. La chaîne de Markov est alors mal mélangeante.

Pour pallier cela, une solution proposée dans [BM96] et dans [WME94] est de ré-échantillonner les valeurs uniques $\boldsymbol{\theta}_1^*, \ldots, \boldsymbol{\theta}_{K_n}^*$ à la fin de chaque itération de l'échantillonneur. On exploite donc la propriété de clustering du DP en regroupant les observations appartenant au même cluster et on met à jour directement les paramètres

Algorithme 1

État de la chaîne de Markov : $\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_n$.

Itérations $t = 1, \ldots, T$,

- pour $i = 1, \ldots, n$,
mettre à jour $\boldsymbol{\theta}_i^{(t)} | \boldsymbol{\theta}_{-i}^{(t)}, \boldsymbol{x}_i$ en tirant suivant la loi *a posteriori* conditionnelle $\boldsymbol{\theta}_i | \boldsymbol{x}_i, \theta_{-i}$ donnée dans l'équation 3.8.

TABLE 3.3 – Algorithme de [Esc94] et de [EW95].

des clusters. Pour ce faire, on a besoin d'introduire des variables latentes de classification indiquant quelle composante est associée à l'observation i. Au lieu de mettre à jour les paramètre $\boldsymbol{\theta}_i$, on met à jour les variables indicatrices c_i et les paramètres des composantes $\boldsymbol{\theta}_k^*$ associées à au moins une donnée, ce qui mettra à jour simultanément tous les $\boldsymbol{\theta}_i$ identiques. L'échantillonneur de Gibbs utilise les distributions conditionnelles suivantes pour les c_i :

– pour les composantes non-vides :

$$P(c_i = k | \boldsymbol{x}_i, \boldsymbol{c}_{-i}, \boldsymbol{\Theta}^*) = Z \frac{n_{-i,k}}{n - 1 + \alpha} f(\boldsymbol{x}_i | \boldsymbol{\theta}_k^*), \tag{3.10}$$

– pour les nouvelles composantes :

$$P(c_i \neq c_j \ \forall j \neq i | \boldsymbol{x}_i, \boldsymbol{c}_{-i}, \boldsymbol{\Theta}^*) = Z \frac{\alpha}{n - 1 + \alpha} \int f(\boldsymbol{x}_i | \boldsymbol{\theta}^*) \mathrm{d}G_0(\boldsymbol{\theta}^*), \tag{3.11}$$

où $n_{-i,k}$ est le nombre d'observations appartenant au cluster k, l'observation i n'étant pas prise en compte et Z la constante de normalisation appropriée.

La mise à jour des paramètres des composantes non-vides se fait suivant la loi *a posteriori*, basée sur la loi G_0 et la vraisemblance de toutes les données d'affectées :

$$p(\boldsymbol{\theta}_k^* | \boldsymbol{X}, \boldsymbol{c}) \propto G_0(\mathrm{d}\boldsymbol{\theta}_k^*) \prod_{i:c_i=k} f(\boldsymbol{x}_i | \boldsymbol{\theta}_k^*). \tag{3.12}$$

Il peut arriver que $n_{-i,j}$ soit égal à 0 pour un certain j. Cela se produit lorsque la configuration précédente a mis $\boldsymbol{x}_i$ dans un cluster de taille 1, i.e $n_j = 1$ pour $j = c_i$. Dans ce cas, on enlève de la chaîne le paramètre $\boldsymbol{\theta}_j^*$ associé et on affecte $\boldsymbol{x}_i$ à un autre cluster. Les paramètres des composantes nouvellement créées sont tirés suivant la distribution *a posteriori* de $\boldsymbol{\theta}_j^*$, en combinant G_0 et la vraisemblance basée sur l'observation $\boldsymbol{x}_i$.

Le schéma de simulation est présenté dans le tableau 3.4.

On peut simplifier cet algorithme en marginalisant les $\boldsymbol{\theta}_k^*$ et échantillonner seulement les c_i. C'est la méthode employée par MacEachern [Mac94] et Neal [Nea91]. L'algorithme échantillonne suivant les probabilités conditionnelles suivantes,

– pour les composantes actives

$$P(c_i = k | \boldsymbol{x}_i, \boldsymbol{c}_{-i}, \boldsymbol{\Theta}^*) = Z \frac{n_{-i,k}}{n - 1 + \alpha} \int f(\boldsymbol{x}_i | \boldsymbol{\theta}^*) \mathrm{d}H_{-i,k}(\boldsymbol{\theta}^*), \tag{3.13}$$

Algorithme 2

État de la chaîne de Markov : $\boldsymbol{c} = (c_1, \ldots, c_n)$, $\boldsymbol{\Theta}^* = (\boldsymbol{\theta}^*_1, \ldots, \boldsymbol{\theta}^*_{K_n})$.

Itérations $t = 1, \ldots, T$,

- pour $i = 1, \ldots, n$,
  1. Si $n_{-i,c_i} = 0$, enlever $\boldsymbol{\theta}^*_{c_i}$ de l'état de la chaîne.
  2. Mettre à jour c_i suivant les équations 3.10 et 3.11.
     Si la valeur c_i tirée est différente des autres c_j, tirer $\boldsymbol{\theta}^*_{c_i}$ à partir de $H_i \propto G_0(\boldsymbol{\theta}^*_{c_i}) \times f(\boldsymbol{x}_i|\boldsymbol{\theta}^*_{c_i})$ et l'ajouter à l'état de la chaîne.
- pour $k = 1, \ldots, K_n$ (ré-échantillonnage des paramètres des composantes actives), mettre à jour les $\boldsymbol{\theta}^*_k$ à partir de la loi *a posteriori* de l'équation 3.12.

TABLE 3.4 – Algorithme de [BM96] et [WME94].

– pour les composantes dormantes :

$$P(c_i \neq c_j \ \forall j \neq i | \boldsymbol{x}_i, \boldsymbol{c}_{-i}, \boldsymbol{\Theta}^*) = Z \frac{\alpha}{n-1+\alpha} \int f(\boldsymbol{x}_i|\boldsymbol{\theta}^*) \mathrm{d}G_0(\boldsymbol{\theta}^*) \tag{3.14}$$

où $H_{-i,k}$ est la distribution *a posteriori* de $\boldsymbol{\theta}^*$ basée sur la loi *a priori* G_0 et toutes les observations, outre $\boldsymbol{x}_i$, qui sont associées à la composante k,

$$H_{-i,k}(\boldsymbol{\theta}^*) = \frac{G_0(\boldsymbol{\theta}^*) \prod_{j \neq i, c_j = k} f(\boldsymbol{x}_j|\boldsymbol{\theta}^*_k)}{\int G_0(\boldsymbol{\theta}^*) \prod_{j \neq i, c_j = k} f(\boldsymbol{x}_j|\boldsymbol{\theta}^*) \mathrm{d}\boldsymbol{\theta}^*}.$$

Ce dernier algorithme est résumé dans le tableau 3.5.

Algorithme 3

État de la chaîne de Markov : $\boldsymbol{c} = (c_1, \ldots, c_n)$

Itérations $t = 1, \ldots, T$,

- pour $i = 1, \ldots, n$, mettre à jour c_i en tirant suivant les probabilités conditionnelles données dans les équations 3.13 et 3.14.

TABLE 3.5 – Algorithme de [Nea91] et [Mac94] .

Les trois algorithmes d'échantillonnage décrits ci-dessus produisent des chaînes de Markov ergodiques et qui convergent. Ils nécessitent toutefois que la distribution de base G_0 soit conjuguée à la densité mélangée $f(\boldsymbol{x}|\boldsymbol{\theta})$ (par exemple dans le cas où elles sont toutes deux gaussiennes) pour que les intégrales dans les équations 3.9, 3.11 et 3.14

puissent être calculables analytiquement. Par ailleurs, l'échantillonnage suivant la loi *a posteriori* $H_i \propto G_0(\boldsymbol{\theta}^*_{c_i}) \times f(\boldsymbol{x}_i|\boldsymbol{\theta}^*_{c_i})$ peut être très difficile dans le cas non conjugué. Par conséquent, pour ne pas se limiter aux modèles restrictifs du cas conjugué, d'autres techniques ont été proposées et nous allons maintenant les aborder.

Échantillonnage dans le cadre de lois non conjuguées

Algorithme no-gaps

MacEachern et Müller [MM98] proposent une augmentation du modèle, basée sur l'introduction de composantes auxiliaires de paramètres tirés suivant G_0. Leur algorithme appelé « no-gaps » est paramétré de telle sorte qu'il n'y ait pas « de trou » dans les c_i : les composantes actives sont numérotées de 1 à K_n puis le modèle est étendu de $K_n + 1$ à n afin d'inclure des composantes vides qui peuvent être vues comme des composantes potentielles non encore affectées,

$$\underbrace{\boldsymbol{\theta}^*_1, \ldots, \boldsymbol{\theta}^*_{K_n}}_{\text{actives}} \underbrace{\boldsymbol{\theta}^*_{K_n+1}, \ldots, \boldsymbol{\theta}^*_n}_{\text{potentielles}}.$$

L'augmentation du modèle permet de remplacer l'évaluation des intégrales par de simples calculs de vraisemblances.

MacEachern et Müller proposent aussi une étape de permutation des labels pour les composantes singletons ($n_{c_i} = 1$). Si on désigne par K_n^- le nombre de composantes actives pour $j \neq i$ c'est-à-dire n'incluant pas la $i^{\text{ième}}$ observation (ces composantes sont labellisées de 1 à K_n^-), cette permutation se traduit pour la classe singleton par $P(c_i = K_n) = \frac{1}{K_n}$ et $P(c_i < K_n) = 1 - \frac{1}{K_n} = \frac{K_n - 1}{K_n} = \frac{K_n^-}{K_n^- + 1}$.

L'algorithme peut être décrit ainsi : on labellise les composantes de 1 à K_n^-. Si $c_i \neq c_j$ pour tout $j \neq i$ alors soit le label c_i est laissé inchangé avec la probabilité $\frac{K_n^-}{K_n^- + 1}$, soit $c_i = K_n^- + 1$ auquel cas on affecte à $\boldsymbol{\theta}_{K_n^- + 1}$ la valeur courante de $\boldsymbol{\theta}_{c_i}$. Ainsi seul le label change, la valeur du paramètre reste inchangée. L'algorithme utilise les probabilités conditionnelles suivantes,

$$P(c_i = k|\boldsymbol{x}_i, \boldsymbol{c}_{-i}, \boldsymbol{\Theta}) \quad = Z\, n_{-i,k} f(\boldsymbol{x}_i|\boldsymbol{\theta}^*_k) \text{ pour } k = 1, \ldots, K_n^- \tag{3.15}$$

$$P(c_i = K_n^- + 1|\boldsymbol{x}_i, \boldsymbol{c}_{-i}, \boldsymbol{\Theta}) \quad = Z \frac{\alpha}{K_n^- + 1} f(\boldsymbol{x}_i|\boldsymbol{\theta}_{K_n^-}) \text{ si } k = K_n^- + 1. \tag{3.16}$$

Une fois la mise à jour des variables indicatrices effectuée, on procède à celle des paramètres des composantes suivant la loi *a posteriori* conditionnelle donnée dans l'équation 3.12. Le schéma de simulation est résumé dans le tableau 3.6. Bien que théoriquement l'état de la chaîne comporte n composantes, on a juste besoin d'en représenter $K_n^- + 1$. On peut toujours inclure de nouvelles composantes si besoin, en tirant leurs paramètres suivant la distribution *a priori* G_0.

Cet algorithme peut s'appliquer pour n'importe quel modèle qui est tel que l'on puisse échantillonner à partir de G_0 et que l'on puisse calculer $f(x_i|\boldsymbol{\theta})$, sans avoir besoin que G_0 soit conjuguée à f. Cependant, il mélange lentement à cause de la probabilité réduite d'affecter une observation à une composante nouvellement créée. Par ailleurs,

Algorithme 4

État de la chaîne de Markov : $\boldsymbol{c} = (c_1, \ldots, c_n), \boldsymbol{\Theta}^* = (\boldsymbol{\theta}_1^*, \ldots, \boldsymbol{\theta}_n^*)$

Itérations $t = 1, \ldots, T$,

- pour $i = 1, \ldots, n$,
  1. si $n_{c_i} = 1$ ($K_n^- = K_n - 1$) alors
     - avec la probabilité $\frac{K_n^-}{K_n^- + 1}$, laisser c_i inchangé.
     - sinon, labelliser c_i comme étant le cluster $K_n^- + 1$.
  2. mettre à jour c_i suivant les équations 3.15 et 3.16.
  3. Réarranger les clusters actifs de 1 à K_n.
- Mise à jour des paramètres des composantes
  1. pour $k = 1, \ldots, K_n$, mettre à jour les $\boldsymbol{\theta}_k^*$ suivant la loi *a posteriori* donnée dans l'équation 3.12,
  2. pour $k = K_n + 1, \ldots, n$, tirer les $\boldsymbol{\theta}_k^*$ suivant G_0.

TABLE 3.6 – Algorithme de [MM98].

l'augmentation du modèle de sorte à comporter n composantes peut être inefficace surtout si $K_n << n$.

Algorithmes de Neal

Pour traiter le cas non-conjugué, Neal [Nea00] propose plusieurs approches :

1. [Nea00, Algorithme 5] consiste à mettre à jour les c_i en utilisant un algorithme de Metropolis-Hastings (MH) dont la loi candidate est donnée par la loi *a priori* conditionnelle. Une variante plus efficace de cet algorithme [Nea00, Algorithme 7] combine l'algorithme MH et l'échantillonneur de Gibbs.
2. [Nea00, Algorithme 8] est un peu similaire à celle de MacEachern et Müller [MM98] que nous venons de voir. Elle consiste aussi à introduire des composantes auxiliaires mais qui n'existent que de façon temporaire.

Ces deux approches utilisent la représentation du modèle de mélange par processus de Dirichlet comme limite d'un modèle fini Dirichlet-multinomial et nous allons maintenant l'expliciter. Considérons un modèle de mélange fini sous la forme du modèle hiérarchique suivant :

$$\begin{aligned} \boldsymbol{x}_i|\boldsymbol{c}, \boldsymbol{\Theta} &\sim f(\boldsymbol{x}_i|\boldsymbol{\theta}_{c_i}^*) \\ c_i|\boldsymbol{w} &\sim \text{Mult}(w_1, \ldots, w_k), i.e\ \mathbb{P}(c_i = k) = w_k \\ \boldsymbol{\theta}_k^*|G_0 &\sim G_0 \\ \boldsymbol{w} = w_1, \ldots, w_K|\alpha &\sim \text{Dir}(\frac{\alpha}{K}, \ldots, \frac{\alpha}{K}). \end{aligned} \tag{3.17}$$

Les variables indicatrices $\boldsymbol{c} = \{c_1, \ldots, c_n\}$ servent à identifier à quelle composante du mélange les observations $\boldsymbol{X} = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n\}$ sont affectées, $c_i = k$ si $\boldsymbol{\theta}_i = \boldsymbol{\theta}_{c_i}^* = \boldsymbol{\theta}_k^*$. Les

poids suivent une distribution de Dirichlet symétrique de paramètre $\frac{\alpha}{K}$ qui tend vers 0 quand $K \to \infty$. L'espérance a priori de n'importe quelle composante k est $1/K$. Les valeurs que prennent les labels c_i n'ont donc pas de signification particulière.

La distribution *a priori* des poids est :

$$p(w_1, \ldots, w_K|\alpha) = \frac{\Gamma(\alpha)}{\Gamma(\alpha/K)^K} \prod_{j=1}^{K} w_j^{\alpha/K-1}.$$

Sachant les poids, le nombre d'observations dans chaque classe est multinomiale

$$p(n_1, \ldots, n_K|w_1, \ldots, w_K) = \frac{n!}{n_1!n_2!\ldots n_K!} \prod_{j=1}^{K} w_j^{n_j},$$

et la loi des variables indicatrices :

$$p(c_1, \ldots, c_n|w_1, \ldots, w_K) = \prod_{j=1}^{K} w_j^{n_j}.$$

On peut marginaliser les poids et utiliser l'intégrale de Dirichlet (annexe A, équation A.7) pour avoir la loi *a priori* des $c_i|\alpha$,

$$\begin{aligned} p(c_1, \ldots, c_n|\alpha) &= \int p(c_1, \ldots, c_n|w_1, \ldots, w_K) p(w_1, \ldots, w_K|\alpha) \mathrm{d}w_1, \ldots, \mathrm{d}w_K \\ &= \frac{\Gamma(\alpha)}{\Gamma(\alpha/K)^K} \int \prod_{j=1}^{K} w_j^{n_j+\alpha/K-1} \mathrm{d}w_j = \frac{\Gamma(\alpha)}{\Gamma(n+\alpha)} \prod_{j=1}^{K} \frac{\Gamma(n_j+\alpha/K)}{\Gamma(\alpha/K)}. \end{aligned}$$

La distribution *a priori* conditionnelle de c_i est alors

$$\begin{aligned} P(c_i = k|c_1, \ldots, c_{i-1}) &= \frac{P(c_1, \ldots, c_{i-1}, c_i = k)}{P(c_1, \ldots, c_{i-1})} \\ &= \frac{\Gamma(\alpha)\Gamma(\alpha/K)^{-K} \int w_{c_1} \ldots w_{c_{i-1}} w_k \; w_1^{(\alpha/K)-1} \ldots w_K^{(\alpha/K)-1} \mathrm{d}\boldsymbol{w}}{\Gamma(\alpha)\Gamma(\alpha/K)^{-K} \int w_{c_1} \ldots w_{c_{i-1}} w_1^{(\alpha/K)-1} \ldots w_K^{(\alpha/K)-1} \mathrm{d}\boldsymbol{w}} \\ &= \frac{n_{-i,k} + \frac{\alpha}{K}}{\alpha + i - 1} \end{aligned}$$

où $n_{-i,k}$ désigne le nombre d'observations $\boldsymbol{x}_i$, $i < k$, qui sont affectées à la classe k.

Quand $K \to +\infty$, le paramètre de concentration α/K de la distribution de Dirichlet tend vers 0 et les distributions *a priori* des c_i sont données par les limites suivantes,

$$P(c_i = k|c_1, \ldots, c_{i-1}) = \frac{n_{-i,k}}{\alpha + n - 1} \tag{3.18}$$

$$P(c_i \neq c_{i'} \text{ pour tout } i \neq i'|c_1, \ldots, c_{i-1}) = \frac{\alpha}{\alpha + n - 1}. \tag{3.19}$$

En effet pour les composantes actives, $P(c_i = k|c_1, \ldots, c_{i-1}) \propto n_{-i,k} + \frac{\alpha}{K} \overset{K\to\infty}{\longrightarrow} n_{-i,k}$ et pour les pour les composantes dormantes, $P(c_i = k|c_1, \ldots, c_{i-1}) \propto \frac{\alpha(K-K_n)}{K} \overset{K\to\infty}{\longrightarrow} \alpha$ $(K >> K_n)$.

Puisque les c_i n'ont pas de signification particulière, on peut considérer que les c_i ne sont significatifs que s'ils sont égaux ou différents des autres c_j. Dans ce cas, les équations 3.18 et 3.19 sont tout ce dont on a besoin pour définir le modèle. Par ailleurs, si on prend $\boldsymbol{\theta}_i = \boldsymbol{\theta}^*_{c_i}$, la limite du modèle de mélange 3.17 quand $K \to \infty$ est un modèle de mélange par processus de Dirichlet (DPM),

$$\begin{aligned} \boldsymbol{x}_i|\boldsymbol{\theta}_i &\sim f(\boldsymbol{x}_i|\boldsymbol{\theta}_i) \\ \boldsymbol{\theta}_i|H &\sim H \\ H|\alpha, G_0 &\sim DP(\alpha, G_0). \end{aligned} \tag{3.20}$$

En effet, si on marginalise H dans le modèle 3.20, la distribution conditionnelle des $\boldsymbol{\theta}_i$ sachant α et G_0 est donnée par l'urne de Blackwell-MacQueen (modèle 2.7). On voit alors la correspondance entre les probabilités conditionnelles des $\boldsymbol{\theta}_i$ dans l'urne de Blackwell-MacQueen et celle des c_i dans les équations 3.18 et 3.19.

Utilisant ce modèle, Neal [Nea00] propose deux algorithmes d'échantillonnage pour l'inférence des DPM dans le cas non-conjugué.

Metropolis Hastings et échantillonnage de Gibbs partiel

Le principe du cinquième algorithme que nous allons présenter [Nea00, Algorithme 7] est de considérer un algorithme de Metropolis-Hastings (MH) pour la mise à jour des c_i. On note par g la loi de proposition et π et la loi cible (loi *a posteriori* des c_i).

[Nea00, Algorithme 5] considère d'abord une loi de proposition choisie comme étant la loi *a priori* des c_i (équations 3.18 et 3.19). La probabilité d'acceptation est alors égale au rapport des vraisemblances,

$$\rho(c'_i, c_i) = \frac{g(c_i|c'_i)\pi(c'_i)}{g(c'_i|c_i)\pi(c_i)} = \frac{f(\boldsymbol{x}_i|\boldsymbol{\theta}^*_{c'_i})}{f(\boldsymbol{x}_i|\boldsymbol{\theta}^*_{c_i})}$$

où c'_i est la variable candidate et c_i la valeur courante.

La mise à jour des variables indicatrices s'écrit ainsi :

1. on tire une candidate c'_i utilisant les lois *a priori* conditionnelles données dans les équations 3.18 et 3.19. C'est-à-dire que
   - soit $c'_i = k \quad (k \in \{1, \dots, K_n\})$ avec la probabilité $\frac{n_{-i,k}}{n-1+\alpha}$,
   - soit $c'_i = k \quad (k \notin \{1, \dots, K_n\})$ avec la probabilité $\frac{\alpha}{n-1+\alpha}$.
     Si tel est le cas, i.e on a créé une nouvelle composante singleton, alors on tire $\boldsymbol{\theta}^*_{c'_i} \sim G_0$.
2. Une fois c'_i tirée, la probabilité de l'accepter comme nouvelle valeur pour c_i est donnée par
$$r(c'_i, c_i) = \min[1, \rho(c'_i, c_i)].$$

On peut noter que si la valeur courante de c_i est une composante singleton, elle ne peut pas être choisie comme candidate c'_i. Dans ce cas, si la candidate choisie n'est pas une des composantes actives, c'est forcément une nouvelle composante. Par ailleurs dans cet algorithme MH, de par la loi de proposition, la décision d'une nouvelle valeur pour c_i favorise plus les composantes ayant beaucoup de données.

Le problème de cet algorithme [Nea00, Algorithme 5] est que les candidates correspondantes à des composantes nouvellement créées sont générées avec une probabilité proportionnelle à α . Si α est petit alors cette probabilité est faible. [Nea00, Algorithme 7] propose de modifier la loi candidate de telle sorte que :

1. si c_i n'est pas singleton, qu'on puisse proposer de la changer en une composante nouvellement créée c_i', de paramètre $\boldsymbol{\theta}^*_{c_i'} \sim G_0$. Cette nouvelle composante est acceptée comme nouvelle valeur de c_i avec la probabilité

$$r(c_i', c_i) = \min \left[1, \ \frac{\alpha}{n-1} \frac{f(\boldsymbol{x}_i|\boldsymbol{\theta}^*_{c_i'})}{f(\boldsymbol{x}_i|\boldsymbol{\theta}^*_{c_i})} \right]. \tag{3.21}$$

2. Si c_i est singleton, la candidate c_i' est systématiquement choisie parmi les composantes non-vides telle que $c_i' = k$ avec la probabilité $\frac{n_{-i,k}}{n-1}$. Cette candidate est acceptée avec la probabilité

$$r(c_i', c_i) = \min \left[1, \ \frac{n-1}{\alpha} \frac{f(\boldsymbol{x}_i|\boldsymbol{\theta}^*_{c_i'})}{f(\boldsymbol{x}_i|\boldsymbol{\theta}^*_{c_i})} \right]. \tag{3.22}$$

L'algorithme MH ainsi modifié produit une chaîne de Markov ergodique qui converge mais la chaîne est mal-mélangeante. Le problème est que le changement d'une observation d'une composante existante à une autre ne peut se faire qu'en passant par un état non-probable où l'observation en question est dans une composante singleton. Pour améliorer le mélange de l'algorithme MH, Neal propose d'y introduire une étape supplémentaire d'échantillonnage de Gibbs appliquée seulement aux observations non singletons et qui ne peuvent changer que pour des composantes actives. Le schéma global de cet algorithme, avec la loi de proposition modifiée et la mise à jour partielle par échantillonnage de Gibbs, est présenté dans le tableau 3.7.

Échantillonnage de Gibbs avec des paramètres auxiliaires

Le deuxième algorithme proposé par Neal [Nea00, Algorithme 8] utilise aussi des composantes auxiliaires tout comme MacEachern et Müller [MM98]. Mais contrairement à celles de MacEachern et Müller, les composantes utilisées par Neal ne sont que temporaires. L'algorithme repose sur la méthode des variables auxiliaires qui consiste à introduire des variables à un certain niveau de l'échantillonnage MCMC puis à les écarter par la suite.

Méthode des variables auxiliaires :

L'idée de cette méthode est que l'on peut échantillonner suivant une distribution d'intérêt $\pi(x)$ en échantillonnant suivant une loi jointe $\pi(x, u)$ qui laisse invariante $\pi(x)$. C'est le cas si $\pi(x)$ est la distribution marginale de la v.a. X sous la loi jointe $\pi(x, u)$. [3]

On peut étendre cette idée de façon à utiliser des variables auxiliaires qui sont ajoutées à un moment dans l'échantillonnage, puis écartées par la suite. Cela conduit

3. Dans le contexte bayésien, $\pi(x)$ conditionne implicitement sur les données observées, donc $\pi(u|x)$ peut aussi conditionner sur les observations.

Algorithme 5

État de la chaîne de Markov : $\boldsymbol{c} = (c_1, \ldots, c_n), \boldsymbol{\Theta}^* = (\boldsymbol{\theta}^*_1, \ldots, \boldsymbol{\theta}^*_{K_n})$

Itérations $t = 1, \ldots, T$,

- pour $i = 1, \ldots, n$, (mise à jour des c_i)

  1. si $\boldsymbol{x}_i$ appartient à une composante non-singleton, alors la candidate c'_i est une composante nouvellement créée de paramètre $\boldsymbol{\theta}^*_{c'_i} \sim G_0$. Elle est acceptée comme nouvelle valeur pour c_i avec la probabilité donnée dans l'équation 3.21,
  2. sinon, si $\boldsymbol{x}_i$ appartient à une composante singleton, tirer une candidate c'_i à partir des composantes non-vides, telle que $c'_i = k$ avec la probabilité $\frac{n_{-i,k}}{n-1}$. Cette candidate est acceptée avec la probabilité donnée dans l'équation 3.22. Si elle n'est pas acceptée, garder l'ancienne valeur de c_i.

- pour $i = 1, \ldots, n$, (échantillonnage de Gibbs)
  - pour toutes les composantes non-singletons, mettre à jour c_i suivant la probabilité conditionnelle
  $$P(c_i = k | \boldsymbol{c}_{-i}, \boldsymbol{x}_i, \boldsymbol{\Theta}^*) = Z \, \frac{n_{-i,k}}{n-1} f(\boldsymbol{x}_i | \boldsymbol{\theta}^*_k)$$
  avec Z désignant une constante de normalisation.
- pour $k = 1, \ldots, K_n$ (mise à jour des paramètres des composantes), mettre à jour les $\boldsymbol{\theta}^*_k$ suivant la loi *a posteriori* donnée dans l'équation 3.12.

TABLE 3.7 – [Nea00, Algorithme 7].

à une chaîne de Markov dont l'état permanent est constitué des variables x mais on introduit de façon temporaire des variables u pendant la mise à jour des x. La méthode des variables auxiliaires procède de la façon suivante :

1. tirer une valeur pour u suivant la distribution conditionnelle $\pi(u|x)$,
2. mettre à jour (x, u) d'une façon qui laisse $\pi(x, u)$ invariante,
3. écarter u et garder la valeur de x.

On peut donc introduire autant de variables auxiliaires $u_1, u_2 \ldots$ jugées nécessaires, tant que $\pi(x)$ est la distribution marginale de x sous la loi jointe $\pi(x, u_1, u_2, \ldots)$, pour construire une chaîne de Markov qui converge vers $\pi(x)$. Après convergence de l'algorithme, les échantillons $\{x^{(t)}\}$ peuvent être utilisés pour inférer sur $\pi(x)$ et les échantillons $\{u^{(t)}\}$ peuvent être ignorés.

[Nea00, Algorithme 8] utilise cette technique pour la mise à jour des c_i. Comme dans l'algorithme 2, l'état permanent de la chaîne de Markov est constitué des c_i et des paramètres $\boldsymbol{\theta}^*_k$ des composantes actives mais pendant la mise à jour des c_i, on utilise m variables auxiliaires ($m \geq 1$) représentant les composantes potentielles non encore représentées. Le choix du nombre de composantes auxiliaires dépend de l'utilisateur et est guidé par un compromis entre coût de calcul et propriétés de mélange.

Comme les données sont échangeables et les labels c_i des composantes complètement

arbitraires, on peut traiter chaque donnée i comme si c'était la dernière et l'affecter soit à une composantes active, soit à une composante auxiliare. Utilisant les mêmes notations que précédemment, K_n^- désigne le nombre de composantes actives sans tenir compte de l'observation i. On labellise les composantes de sorte que les composantes 1 à K_n^- sont actives et le reste dormantes. La probabilité *a priori* d'affecter $\boldsymbol{x}_i$ à l'une des $k \leq K_n^-$ composantes actives est $\frac{n_{-i,k}}{\alpha+n-1}$ et la probabilité de créer une nouvelle composante est $\frac{\alpha}{\alpha+n-1}$, qui sera équitablement répartie entre les m composantes auxiliaires.
La première étape de l'échantillonneur consiste à tirer suivant la distribution conditionnelle de ces paramètres auxiliaires, connaissant la valeur courante de c_i et le reste des variables. Si la valeur courante c_i est une composante non-singleton, les paramètres auxiliaires sont tirés *iid* suivant G_0. Sinon, si c_i est une composante singleton, alors elle doit alors être associée à l'une des m composantes auxiliaires. Pour cela, on devrait techniquement choisir aléatoirement une des m composantes auxiliaires mais grâce à l'échangeabilité, on peut labelliser c_i comme étant la première composante auxiliaire (i.e $c_i = K_n^- + 1$) qui reçoit comme valeur de paramètre $\boldsymbol{\theta}^*_{c_i}$. Les paramètres pour les $m-1$ composantes auxiliaires restantes (si $m > 1$) sont tirées *iid* suivant G_0. La mise à jour des c_i s'effectue suivant les probabilités conditionnelles :

$$P(c_i = k|\boldsymbol{x}_i, \boldsymbol{c}_{-i}, \boldsymbol{\Theta}^*) = \begin{cases} Z \ \dfrac{n_{-i,k}}{n-1+\alpha} \ f(\boldsymbol{x}_i|\boldsymbol{\theta}^*_k) \text{ pour } k = 1, \ldots, K_n^- \\ Z \ \dfrac{\alpha}{m(n-1+\alpha)} \ f(\boldsymbol{x}_i|\boldsymbol{\theta}^*_k) \text{ pour } k = K_n^- + 1, \ldots, K_n^- + m \end{cases} \tag{3.23}$$

où Z est la constante de normalisation appropriée.

Une fois la valeur de c_i choisie, on écarte toutes les composantes vides. Le schéma de l'algorithme est résumé dans le tableau 3.8.

Cet algorithme ressemble à celui de MacEachern et Müller pour $m = 1$. Mais ici, la probabilité de créer une nouvelle composante singleton dans la mise à jour de la $i^{\text{ième}}$ observation est beaucoup plus élevée que dans l'algorithme de MacEachern et Müller et l'inverse est aussi vrai. Ceci peut être intéressant pour les petites valeurs de α puisque dans ce cas de tels changements ont lieu que très rarement. D'un autre côté, quand m tend vers l'infini, cet algorithme se comporte comme le deuxième algorithme que nous avons vu dans le cas conjugué. En effet, les m valeurs pour $\boldsymbol{\theta}^*_k$ tirées de G_0 peuvent être utilisées pour fournir une approximation Monte-Carlo de l'intégrale $\int f(x_i|\boldsymbol{\theta}^*)\mathrm{d}G_0(\boldsymbol{\theta}^*)$.

Nous venons de résumer quelques uns des algorithmes marginaux pour l'inférence dans les modèles de mélange par processus de Dirichlet. Ces méthodes utilisent la représentation du processus de Dirichlet en urne de Blackwell-MacQueen ou le processus du restaurant chinois. L'avantage de ces représentations est que la règle de prédiction, i.e la probabilité marginale d'affecter la donnée $\boldsymbol{x}_i$ à une des composantes existantes ou à une nouvelle composante, définit une partition aléatoire échangeable sur les entiers $\{1, \ldots, n\}$. La probabilité jointe d'une affectation $c_1, \ldots, c_n$ est alors indépendante de l'ordre dans lequel on a tiré les observations. Ce qui donne en général aux méthodes marginales de bonnes propriétés de mélange. Cependant, une propriété commune à tous ces algorithmes est qu'ils reposent sur des mises à jour incrémentales c'est-à-dire que pour affecter la $i^{\text{ième}}$ observation à une composante, on a besoin de connaître la configuration pour toutes les autres observations. Ce qui fait que les méthodes marginales sont très difficilement parallélisables. Nous discuterons plus en détails de ces propriétés dans

Algorithme 6

État de la chaîne de Markov : $\boldsymbol{c} = (c_1, \ldots, c_n), \boldsymbol{\Theta}^* = (\boldsymbol{\theta}^*_1, \ldots, \boldsymbol{\theta}^*_{K_n})$

Itérations $t = 1, \ldots, T$

- pour $i = 1, \ldots, n$

  1. Si $c_i \neq c_j$ pour tout $j \neq i$ alors,
     - $c_i \longleftarrow K_n^- + 1$ et $\boldsymbol{\theta}^*_{K_n^-+1} \longleftarrow \boldsymbol{\theta}^*_{c_i}$.
     - tirer les paramètres $\boldsymbol{\theta}^*_k$ $(K_n^- + 1 < k \leq K_n^- + m)$ *iid* suivant G_0.
  2. Sinon, s'il existe un $j \neq i$ tel que $c_i = c_j$ alors,
     - tirer les paramètres $\boldsymbol{\theta}^*_k$ $(K_n^- + 1 \leq k \leq K_n^- + m)$ *iid* suivant G_0.
  3. Mettre à jour c_i selon l'équation 3.23.
  4. Écarter toutes les composantes dormantes.

- pour $k = 1, \ldots, K_n$ (mise à jour des paramètres des composantes), mettre à jour les $\boldsymbol{\theta}^*_k$ suivant la loi *a posteriori* donnée dans l'équation 3.12.

TABLE 3.8 – [Nea00, Algorithme 8].

la comparaison entre les méthodes marginales et conditionnelles. Enfin, une autre limite des méthodes marginales est qu'elles ne représentent pas explicitement la distribution H réalisation d'un processus de Dirichlet mais plutôt des échantillons de H. L'alternative à cette approche est de représenter explicitement cette distribution. C'est l'objet des méthodes conditionnelles que nous allons maintenant aborder.

3.2.2 Méthodes conditionnelles

Les algorithmes conditionnels représentent explicitement la mesure générée par le DP en utilisant la représentation stick-breaking du processus. La difficulté est alors la façon de traiter la distribution infinie H. Ishwaran et James [IJ01] ont recours à une approximation et tronquent la mesure après une valeur N choisie. Une alternative qui permet d'éviter la troncature a été introduite dans l'échantillonneur rétrospectif de Papaspiliopoulos et Roberts [PR07]. Walker [Wal07] propose une méthode d'échantillonnage basée sur la technique du « slice », qui ne tronque pas le nombre de composantes mais qui nécessite seulement d'échantillonner un nombre fini de variables à chaque itération. Toujours dans cette même perspective, Kalli *et al.* [KGW09] proposent une version améliorée de l'algorithme de [Wal07] appelée le « slice efficient ».

Toutes ces méthodes conditionnelles sont basées sur la représentation stick-breaking du mélange par processus de Dirichlet. Nous rappelons que dans cette représentation le

modèle s'écrit :

$$\begin{aligned} \boldsymbol{x}_i|c_i, \boldsymbol{\Theta}^* &\sim f(\boldsymbol{x}_i|\boldsymbol{\theta}^*_{c_i}) \\ c_i|\boldsymbol{w} &\sim \sum_{k=1}^{\infty} w_k \delta_k(\cdot) \\ \boldsymbol{w}|\alpha &\sim \text{GEM}(\alpha) \\ \boldsymbol{\theta}^*_k|G_0 &\sim G_0. \end{aligned} \tag{3.24}$$

On voit qu'ici, contrairement aux méthodes marginales où ils sont intégrés, les poids du mélange sont représentés explicitement en même temps que les paramètres des composantes et les variables indicatrices. Cependant dans les algorithmes marginaux, la marginalisation des poids permet de travailler directement avec les variables indicatrices qui sont en nombre fini plutôt qu'avec l'infinité de poids. Le problème dans les méthodes conditionnelles est de traiter le nombre infini de composantes. Les méthodes par troncature utilisent une approximation en tronquant le nombre de composantes.

Échantillonnage par troncature du processus

Ishwaran et James ([IJ01], [IJ03b]), utilisant le fait que les poids dans la distribution GEM décroissent de façon exponentielle en loi, ont montré qu'on pouvait tronquer le DPM à une valeur entière N choisie et ainsi écarter les termes $N+1$, $N+2\ldots$ La condition $v_N = 1$ est nécessaire afin que la mesure tronquée H_N définisse une mesure de probabilité. Elle a alors pour distribution :

$$H_N(\cdot) = \sum_{k=1}^{N} w_k \delta_{\boldsymbol{\theta}^*_k}(\cdot).$$

On peut ré-écrire le modèle hiérarchique 3.24 sous cette troncature :

$$\begin{aligned} \boldsymbol{x}_i|c_i, \boldsymbol{\Theta}^* &\sim f(\boldsymbol{x}_i|\boldsymbol{\theta}^*_{c_i}) \\ c_i|\boldsymbol{w} &\sim \sum_{k=1}^{N} w_k \delta_k(\cdot) \\ \boldsymbol{w}|\alpha &\sim \text{GEM}(\alpha) \\ \boldsymbol{\theta}^*_k|G_0 &\sim G_0 \end{aligned} \tag{3.25}$$

L'échantillonneur de Gibbs utilise les distributions *a posteriori* conditionnelles suivantes.

1. Distribution conditionnelle des variables indicatrices :
   Connaissant les poids, les affectations sont indépendantes entre-elles. On peut alors mettre à jour les variables indicatrices de façon conjointe et sans avoir à conditionner sur toutes les autres indicatrices. Ceci diffère des algorithmes vus jusqu'à présent dans les méthodes marginales. La loi *a posteriori* conditionnelle des c_i est donnée par :

$$c_i|\boldsymbol{w}, \boldsymbol{\Theta}^*, \boldsymbol{X} \sim \sum_{k=1}^{N} w_{k,i}\delta_k \quad \text{pour } i = 1, \ldots, n \tag{3.26}$$
$$\text{avec } (w_{1,i}, \ldots, w_{N,i}) \propto (w_1\ f(\boldsymbol{x}_i|\boldsymbol{\theta}^*_1), \ldots, w_N\ f(\boldsymbol{x}_i|\boldsymbol{\theta}^*_N)).$$

2. Distribution conditionnelle des poids :
   La loi *a posteriori* conditionnelle des poids est, tout comme la distribution *a priori*, une distribution GEM grâce à la conjugaison entre la distribution de Dirichlet généralisée et l'échantillonnage multinomial :

$$w_1 = v_1^* \text{ et } w_k = v_k^* \prod_{l=1}^{k-1}(1 - v_l^*) \quad \text{pour } k = 2, \ldots, N-1 \tag{3.27}$$

   avec

$$v_k^* \sim \text{Beta}(1 + n_k,\ \alpha + \sum_{l=k+1}^{N} n_l) \quad \text{pour } k = 1, \ldots, N-1 \tag{3.28}$$

   où $n_k = \#\{i : c_i = k\}$.

Le schéma de l'algorithme est résumé dans le tableau 3.9.

Algorithme 7

État de la chaîne de Markov : $\boldsymbol{c} = (c_1, \ldots, c_n)$, $\boldsymbol{w} = (w_1, \ldots, w_N)$, $\boldsymbol{\Theta}^* = (\boldsymbol{\theta}_1^*, \ldots, \boldsymbol{\theta}_N^*)$

Itérations $t = 1, \ldots, T$

- pour $i = 1, \ldots, n$ (mise à jour des variables indicatrices)
  affecter $\boldsymbol{x}_i$ à une des N composantes suivant l'équation 3.26.
- pour $k = 1, \ldots, N$ (mise à jour des poids des composantes)
  tirer v_k suivant l'équation 3.28 et calculer w_k comme dans l'équation 3.27.
- pour $k = 1, \ldots, N$ (mise à jour des paramètres des composantes),
  tirer $\boldsymbol{\theta}_k^*$ suivant la densité proportionnelle à

$$G_0(\mathrm{d}\boldsymbol{\theta}_k^*) \prod_{\{i:c_i=k\}} f(\boldsymbol{x}_i|\boldsymbol{\theta}_k^*) \quad \text{pour } k = 1, \ldots, K_n.$$

TABLE 3.9 – Algorithme de [IJ01].

Cet algorithme est simple à implanter car la troncature fait qu'il est similaire aux échantillonneurs standards dans les modèles fini-dimensionnels. Cependant, même si des méthodes pour contrôler la précision de la troncature existent ([IJ01], [IJ03b]), il serait préférable de garder la nature infinie du modèle et d'éviter une troncature déterministe. Ainsi, des méthodes permettant d'échantillonner suivant la distribution *a posteriori* exacte tout en nécessitant un nombre fini de composantes à chaque itération ont été proposées par Papaspiliopoulos et Roberts [PR07] et Walker [Wal07]. Cette dernière utilise une stratégie dite de *slice sampling*, basée sur l'introduction de variables auxiliaires. Le slice sampling de Walker fut amélioré par Kalli *et al.* [KGW09]. Ce dernier, appelé « slice efficient », échantillonne par blocs les poids et les variables auxiliaires pendant les itérations. Ce qui simplifie la génération des poids et améliore le mélange comparé à l'algorithme de Walker.

Échantillonnage sans troncature

Idée du slice sampling :

Nous avons vu que la méthode des variables auxiliaires permettait d'introduire des variables qui rendent plus facile l'échantillonnage des variables d'intérêt. Le slice sampling est une méthode à variables auxiliaires qui exploite le fait que l'on peut échantillonner suivant une distribution en tirant des variables uniformément dans la région sous le tracé de sa fonction de densité de probabilité [Nea03]. Le problème de l'échantillonnage suivant une distribution arbitraire devient ainsi un problème d'échantillonnage suivant des distributions uniformes.

Plus précisément, supposons que l'on veuille générer suivant la distribution d'une v.a à valeurs dans un sous-ensemble de $\mathbb{R}^n$, dont la densité p est connue à une constante près. Soit f la densité telle que $p(\boldsymbol{x}) \propto f(\boldsymbol{x})$. L'échantillonnage suivant la loi d'intérêt p peut se faire en échantillonnant uniformément dans la région $n+1$ dimensionnelle sous la courbe de f. Formellement, on introduit une variable auxiliaire réelle u et on définit une distribution jointe pour $\boldsymbol{x}$ et u qui est uniforme dans la région $U = \{(\boldsymbol{x}, u) : \ 0 < u < f(\boldsymbol{x})\}$ sous la courbe ou la surface définie par f. La densité jointe de $\boldsymbol{x}$ et de u est donnée par

$$p(\boldsymbol{x}, u) = \begin{cases} 1/Z & \text{si } 0 < u < f(\boldsymbol{x}) \\ 0 & \text{sinon} \end{cases}$$

où $Z = \int f(\boldsymbol{x})\mathrm{d}\boldsymbol{x}$. La densité marginale de $\boldsymbol{x}$ est alors

$$p(\boldsymbol{x}) = \int_0^{f(\boldsymbol{x})} (1/Z)\mathrm{d}u = f(\boldsymbol{x})/Z$$

comme souhaitée. Pour échantillonner suivant $\boldsymbol{x}$, on peut échantillonner de façon conjointe suivant $(\boldsymbol{x}, u)$ et ignorer u. Cependant, il peut ne pas être facile de générer des points uniformément dans U. On peut alors définir une chaîne de Markov qui converge vers cette distribution uniforme. L'échantillonneur de Gibbs peut être utilisé. Auquel cas, on tire alternativement suivant la distribution conditionnelle de u sachant la valeur courante de $\boldsymbol{x}$–qui est uniforme dans l'intervalle $[0, f(\boldsymbol{x})]$– et suivant la distribution conditionnelle de $\boldsymbol{x}$ sachant la valeur courante de u–qui est uniforme dans la région $S = \{\boldsymbol{x} : \ u < f(\boldsymbol{x})\}$.

Pour illustrer cette technique, considérons le cas où la loi cible est la distribution normale standard de densité $p(x) = \frac{1}{\sqrt{2\pi}} \exp(-\frac{x^2}{2})$. On suppose qu'elle est connue à la constante près, i.e on connaît $f(x) = \exp(-\frac{x^2}{2})$. On utilise la technique du slice pour générer suivant p. On utilise un échantillonneur de Gibbs qui tire alternativement suivant les distributions conditionnelles suivantes :

$$u|x \sim \mathcal{U}(\cdot|0, f(x)) \tag{3.29}$$
$$x|u \sim \mathcal{U}(\cdot|S) \tag{3.30}$$

où $\mathcal{U}(\cdot|\cdot)$ désigne la loi uniforme et $S = \{x : u < f(x)\} =]-\sqrt{2\ln(u)}, \sqrt{2\ln(u)}[$. Le résultat est montré à la Figure 3.2.

Des variations à cette technique du slice ont été développées. Dans l'inférence des DPM, Walker [Wal07] utilise des variables « slices » pendant la mise à jour des variables indicatrices et les écarte par la suite.

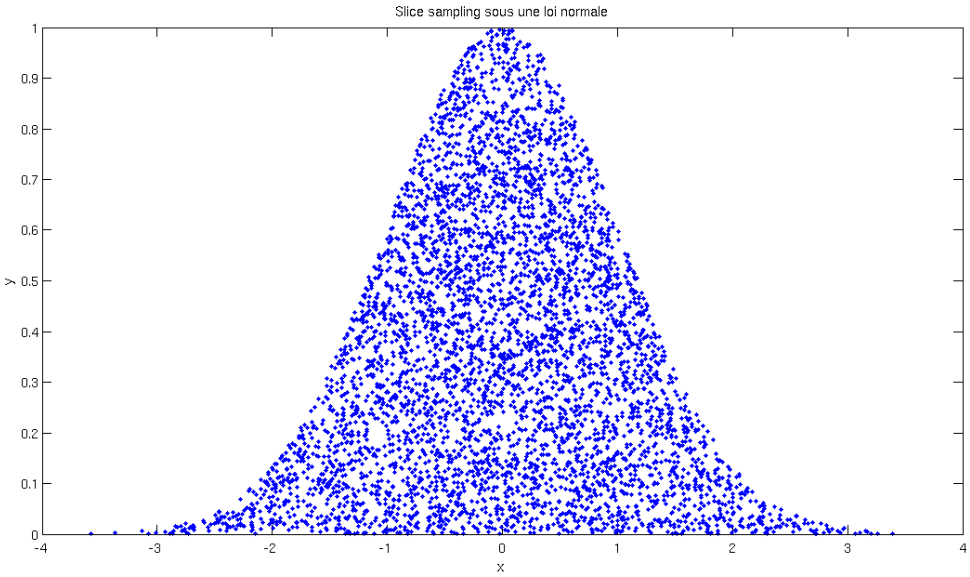


FIGURE 3.2 – 5000 itérations de l'échantillonnage de Gibbs dans le slice sampling. La loi cible est $\mathcal{N}(0,1)$. L'algorithme tire successivement suivant les distributions conditionnelles des équations 3.29 et 3.30.

Le slice sampler

L'idée du slice sampler de Walker [Wal07] est d'introduire des variables auxiliaires qui rendent le modèle conditionnellement fini. Considérons le modèle de mélange :

$$p(\boldsymbol{x}) = \int f(\boldsymbol{x}|\boldsymbol{\theta}) H(\mathrm{d}\boldsymbol{\theta})$$

où H suit un *a priori* stick-breaking (équation 2.11). La densité d'une observation $\boldsymbol{x}_i$, sachant $\boldsymbol{w}$ et $\boldsymbol{\Theta}^*$, est

$$p\left(\boldsymbol{x}_i\right) = \sum_{k=1}^{\infty} w_k \; f\left(\boldsymbol{x}_i|\boldsymbol{\theta}_k^*\right).$$

On introduit $\mathbf{u} = (u_1, u_2, \ldots, u_n)$ des variables auxiliaires suivant une loi uniforme telles que la densité jointe de $(\boldsymbol{x}_i, u_i)$ est

$$p\left(\boldsymbol{x}_i, u_i\right) = \sum_{k} \mathbf{1}(u_i < w_k) \; f\left(\boldsymbol{x}_i|\boldsymbol{\theta}_k^*\right) = \sum_{k=1}^{\infty} w_k \; f\left(\boldsymbol{x}_i|\boldsymbol{\theta}_k^*\right) \mathcal{U}\left(u_i|0, w_k\right).$$

Si on intègre cette loi jointe sur u_i par rapport à la mesure de Lebesgue, on retrouve bien la loi de $p\left(\boldsymbol{x}_i\right)$. La loi jointe est donc bien définie ainsi que la loi marginale de u_i. La densité conditionnelle de $\boldsymbol{x}_i$ sachant u_i est

$$p\left(\boldsymbol{x}_i|u_i\right) = \frac{1}{\sum_k \mathbf{1}(u_i < w_k)} \sum_{k \in \{j : w_j > u_i\}} f\left(\boldsymbol{x}_i|\boldsymbol{\theta}_k^*\right).$$

Sachant u_i, le nombre de composantes dans le mélange est fini puisque $\{j : w_j > u_i\}$ est un ensemble de cardinal fini pour tout $u_i > 0$. En conséquence, le modèle de mélange

infini est transformé en un modèle de mélange fini, de poids tous égaux à $\frac{1}{\sum_k \mathbf{1}(u_i<w_k)}$. On peut compléter le modèle par l'introduction d'une variable d'affectation c_i et considérer

$$p\left(\boldsymbol{x}_i, u_i, c_i\right) = \mathbf{1}(u_i < w_{c_i})\ f\left(\boldsymbol{x}_i|\boldsymbol{\theta}^*_{c_i}\right) = w_{c_i}\ \mathcal{U}\left(u_i|0, w_{c_i}\right)\ f\left(\boldsymbol{x}_i|\boldsymbol{\theta}^*_{c_i}\right).$$

Les distributions conditionnelles nécessaires pour implanter l'échantillonneur de Gibbs sont celles des variables slices u_i, des indicatrices c_i, des paramètres des composantes $\boldsymbol{\theta}^*_k$ et des poids stick-breaking v_k (utilisés pour construire les poids des composantes w_k, voir équation 2.11). Grâce aux variables slices u_i, on a juste besoin d'échantillonner un nombre fini de v_k et $\boldsymbol{\theta}^*_k$ à chaque étape de l'échantillonneur. Dans la première version du slice sampler de Walker [Wal07], chacune de ces variables est tirée indépendamment des autres et l'échantillonnage des v_k est assez compliqué.

Dans la version « efficace » du slice proposée par Kalli *et al.* [KGW09], les poids des composantes $\boldsymbol{w}$ et les variables auxiliaires $\boldsymbol{u}$ sont regroupés dans un *bloc* pendant les itérations. Ce qui, en marginalisant les variables auxiliaires, simplifie largement la génération des poids et résulte en un algorithme qui a de meilleures propriétés de mélange que celui de Walker [Wal07]. Nous allons donc décrire en détails la méthode de Kalli *et al.* [KGW09]. Les variables à échantillonner à chaque itération sont :

$$\{(\boldsymbol{\theta}^*_k, v_k), k = 1, 2, \ldots; (c_i, u_i), i = 1, \ldots, n\}.$$

1. Distribution conditionnelle de u_i :
   La loi de u_i sachant $\boldsymbol{w}$ et c_i est une loi uniforme et est donnée par :
   $$(u_i|\boldsymbol{w}, c_i) \sim \mathcal{U}(0, w_{c_i}) = \frac{\mathbf{1}\{u_i < w_{c_i}\}}{w_{c_i}}. \tag{3.31}$$

2. Distribution conditionnelle de $\boldsymbol{\theta}^*_k$:
   $$G_0(\mathrm{d}\boldsymbol{\theta}^*_j) \prod_{\{i:c_i=j\}} f(\boldsymbol{x}_i|\boldsymbol{\theta}^*_j). \tag{3.32}$$

3. Distribution conditionnelle de w_k :
   Cette distribution conditionnelle de w_k sachant $\boldsymbol{c}$ est la même que dans l'échantillonneur tronqué de Ishwaran et James [IJ01] que nous avons déjà présenté. Il s'agit d'une distribution GEM avec des paramètres mis à jour. C'est-à-dire :
   $$\text{on tire } v^*_k \sim \text{Beta}\left(1 + \sum_{i=1}^{n} \mathbf{1}(c_i = k),\ \alpha + \sum_{i=1}^{n} \mathbf{1}(c_i > k)\right) \tag{3.33}$$
   $$\text{et on construit } w_k = v^*_k \prod_{l=1}^{k-1} (1 - v^*_l).$$

4. Distribution conditionnelle de c_i :
   $$\mathbb{P}(c_i = k) \propto \mathbf{1}(k : w_k > u_i)\ f(\boldsymbol{x}_i|\boldsymbol{\theta}^*_k). \tag{3.34}$$

Pour tirer suivant cette fonction de masse, on a besoin de connaître le nombre de composantes nécessaires à chaque itération. Soit K^* ce nombre. On sait qu'avec le slice, on a juste besoin de représenter les composantes dont le poids est supérieur à u_i. Comme la somme des poids est égale à un, si K^* composantes sont représentées alors une des composantes non représentées a un poids maximal de $1 - \sum_{k=1}^{K^*} w_k$. Si cette valeur est plus grande que u_i, alors on continue la construction stick-breaking jusqu'à ce que le résidu (poids des composantes non encore affectées) soit inférieure à u_i. On est alors sûr qu'il n'y a pas de $w_k > u_i$ pour $k > K^*$. En l'occurrence, K^* est le plus petit K tel que

$$\sum_{k=1}^{K} w_k > 1 - u^*$$

où $u^* = \min\{u_1, \ldots, u_n\}$.

Le schéma d'échantillonnage est décrit dans le tableau 3.10.

Algorithme 8

État permanent de la chaîne de Markov : $\boldsymbol{c} = (c_1, \ldots, c_n)$, $\boldsymbol{w} = (w_1, \ldots, w_{K^*})$, $\boldsymbol{\Theta}^* = (\boldsymbol{\theta}_1^*, \ldots, \boldsymbol{\theta}_{K^*}^*)$

Itérations $t = 1, \ldots, T$

- pour $i = 1, \ldots, n$ (mise à jour des variables indicatrices)
  - Tirer une variable u_i suivant la loi conditionnelle donnée dans l'équation 3.31. Prendre $u^* = \min\{u_1, \ldots, u_n\}$.
  - Tant que $r_{k-1} > u^*$ (poids des composantes pour $k > K_n$),

$$\begin{aligned} v_k &= \text{Beta}\,(1,\ \alpha) \\ w_k &= v_k r_{k-1} \\ r_k &= r_{k-1}\,(1 - v_k)\,. \end{aligned}$$

    Prendre $K^* = \min\left(\{k :\ r_k < u^*\}\right)$.
  - Tirer les paramètres des nouvelles composantes $\boldsymbol{\theta}_k^* \sim G_0$ (pour $k = K_n + 1, \ldots, K^*$).
  - Affecter $\boldsymbol{x}_i$ à une des composantes suivant la loi donnée dans l'équation 3.34.
- Pour $k = 1, \ldots, K_n$ (mise à jour des poids des composantes), tirer w_k suivant l'équation 3.33.
- Pour $k = 1, \ldots, K_n$ (mise à jour des paramètres des composantes), tirer $\boldsymbol{\theta}_k^*$ suivant la densité donnée dans l'équation 3.32.

TABLE 3.10 – Algorithme de [KGW09].

Nous avons présenté plusieurs méthodes d'échantillonnage pour l'inférence dans les modèles de mélange par processus de Dirichlet, aussi bien des approches marginales que conditionnelles. Les méthodes conditionnelles utilisent la construction stick-breaking du DP. Dans cette construction, les poids des composantes sont explicitement représentés.

En conséquence, les variables indicatrices sont mises à jour sans conditionnement sur les autres. Cette caractéristique laisse présager de bonnes propriétés de mélange sur les variables indicatrices. Toutefois, l'ordre biaisé par la taille entraîne une non-échangeabilité des labels. Nous allons maintenant détailler les avantages et les limitations de chaque classe d'algorithmes.

Comparaisons entre les méthodes marginales et conditionnelles

Le principal avantage des méthodes conditionnelles utilisant la représentation stick-breaking est de représenter explicitement la mesure aléatoire générée par le processus. Cela rend donc possible l'inférence directe sur la mesure, contrairement aux méthodes marginales. De plus, par construction, elle ne souffrent pas du problème de conjugaison rencontré dans les algorithmes marginaux. Enfin, comme les poids des composantes sont représentés explicitement, la mise à jour des variables indicatrices s'effectue sans avoir à conditionner sur les autres. Cette propriété rend ces algorithmes capables de mettre à jour des paramètres par blocs et faciles à implanter dans un ordinateur parallèle.

D'un autre côté, en marginalisant les poids des composantes, les méthodes marginales rendent l'allocation très séquentielle puisque l'affectation d'une donnée nécessite de conditionner sur toutes les autres variables indicatrices. Cette mise à jour incrémentale est très préjudiciable surtout quand on travaille avec de gros volumes de données. Un autre inconvénient de la marginalisation est que le calcul des fonctionnelles *a posteriori* nécessite d'effectuer des tirages supplémentaires. Cependant, une caractéristique importante est que les poids aléatoires sont collapsés par la marginalisation et cela résulte en une réduction cruciale de la dimension de l'espace des paramètres.

Une des différences majeures entre les méthodes marginales et conditionnelles provient de l'espace de formulation de l'échantillonnage. Le processus du restaurant chinois (CRP) et l'urne de Pólya d'un côté, la représentation stick-breaking de l'autre sont des représentations du processus de Dirichlet mais définis dans des espaces différents. Le CRP est défini dans l'espace des partitions tandis que le stick-breaking l'est dans l'espace des labels des clusters. On appelle partition un regroupement de données indépendamment des labels des clusters. C'est-à-dire que les allocations $c_1 = 1, c_2 = 1, c_3 = 2$ représentent la même partition $(1,2)(3)$ que $c_1 = 3, c_2 = 3, c_3 = 2$. La règle de prédiction définie par l'urne de Blackwell et MacQueen opère dans l'espace des classes d'équivalence des labels des clusters $(c_1 = c_2 \neq c_3 = c_4 = c_5 \dots)$, tandis que la représentation stick-breaking est définie dans l'espace explicite des labels $(c_1 = 2, c_2 = 2, c_3 = 8, c_4 = 8, c_5 = 8 \dots)$ [PISW06]. Dans la représentation en urne de Pólya et dans le CRP, l'échantillonneur ignore les labels et considère toutes les instances où $(c_1 = c_2 \neq c_3 = c_4 = c_5 \dots)$ comme une seule classe d'équivalence. La numération des variables d'allocation et des clusters n'a donc pas d'importance et les labels des clusters sont échangeables. Par contre, la représentation stick-breaking est explicitement définie dans l'espace des labels des clusters et les labels ne sont pas échangeables. En effet, dans la construction stick-breaking, les poids sont stochastiquement ordonnés : $\mathbb{E}(w_1) < \mathbb{E}(w_2) < \dots$ Chaque composante du modèle a un poids *a priori* différent et est explicitement représentée avec un ordre biaisé par la taille sur les labels des clusters. Par conséquent, les composantes ne sont pas interchangeables et l'étiquetage *a priori* des clusters contribue à l'échantillonnage *a posteriori*. Ces différences théoriques ont

	Propriétés	Conséquences
Marginales	Marginalisation des poids	– Allocation séquentielle – Difficilement parallélisable – Nécessité d'effectuer d'autres tirages pour le calcul des fonctionnelles de la RPM
	– Échangeabilité des labels – Réduction de l'espace des paramètres	Bonnes propriétés de mélange
Conditionnelles	Représentation explicite de la RPM	– Possibilité de paralléliser – Accès direct à la RPM et à ses fonctionnelles
	Non-échangeabilité des labels	Propriétés de mélange moindres

TABLE 3.11 – Tableau récapitulatif des forces et faiblesses des méthodes marginales et conditionnelles.

de réelles conséquences pratiques. Les méthodes marginales ont souvent de meilleures propriétés de mélange que les algorithmes conditionnels et nos comparaisons dans la section 3.2.4 le corroborent.

Quand l'on travaille avec des algorithmes conditionnels non-échangeables, l'une des solutions pour améliorer le mélange est d'introduire des étapes supplémentaire d'échange des labels. En effet, étant donné que dans la représentation stick-breaking les clusters ayant les labels les plus faibles ont des probabilités *a priori* plus grandes que ceux avec des labels plus élevés, l'échantillonneur doit pouvoir mélanger efficacement sur les labels afin d'éviter de biaiser le clustering. Dans [PISW06], les auteurs recommandent d'utiliser systématiquement dans l'échantillonnage de Gibbs deux étapes additionnelles d'algorithme Metropolis-Hastings (« label-swap » et « label-permute ») afin d'améliorer le mélange sur les clusters. Quand on travaille avec des *a priori* stick-breaking, cela semble être la seule façon d'améliorer le mélange sur les labels des clusters (voir aussi [PR07]).

Nous résumons dans le tableau 3.2.2 les forces et faiblesses des méthodes marginales et conditionnelles.

3.2.3 Nouvelle méthode d'échantillonnage

Nous proposons un algorithme conditionnel qui est différent de ceux vus jusqu'à présent. Notre échantillonneur opère dans l'espace des classes d'équivalence sur les labels

des clusters et les identités des clusters sont arbitraires et non significatives. Ces labels sont donc échangeables et aucun mélange n'est nécessaire sur les variables indicatrices.

Le but de ce nouvel échantillonneur est de pouvoir combiner les propriétés qui font défaut aux méthodes marginales et conditionnelles. On voudrait donc un échantillonneur qui puisse combiner les caractéristiques suivantes :

1. approche conditionnelle en retenant la distribution aléatoire générée par le processus (pas de marginalisation des poids),
2. échantillonnage dans l'espace des classes d'équivalence où les clusters sont échangeables (et non pas dans l'espace des labels des clusters) afin d'obtenir une chaîne bien mélangeante,
3. échantillonnage par blocs afin que l'on puisse paralléliser les calculs (surtout l'étape d'affectation). Cette propriété de parallélisation est cruciale pour les gros volumes de données que nous manipulons en reconstruction TEP.

L'algorithme d'échantillonnage est développé pour les modèles de mélange par processus de Pitman-Yor (PYM) et est donc applicable aux modèles de mélange par processus de Dirichlet (DPM). On rappelle que ce dernier est obtenu en prenant $d = 0$ dans le PYM. Nous proposons deux variantes à la méthode afin de traiter la partie infinie du processus. Notre échantillonneur repose sur un résultat de Pitman [Pit96b, Corollaire 20] que nous avons déjà abordé dans la section 2.3.2. Il exprime la distribution *a posteriori* des limites des fréquences relatives des atomes dans le modèle d'échantillonnage d'espèces (SSM)(cf. section 2.2). Ce modèle assure l'échangeabilité de la partition aléatoire sous-jacente, caractérisée par la symétrie de l'EPPF (section 2.2.2). Le modèle s'exprime ainsi. Soit $H \sim \mathrm{PY}(d, \alpha, G_0)$ où G_0 est une mesure de probabilité non-atomique. Considérons un échantillon $\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_n$ tiré de H. Soit K_n le nombre de valeurs distinctes des $\boldsymbol{\theta}_i$, soit $\{\boldsymbol{\theta}_j^*\}_{j=1}^{K_n}$ l'ensemble des valeurs uniques des $\{\boldsymbol{\theta}_i\}_{i=1}^n$ et enfin soit n_j la fréquence de $\boldsymbol{\theta}_j^*$. Alors,

$$H|\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_n \stackrel{d}{=} \sum_{j=1}^{K_n} w_j \delta_{\boldsymbol{\theta}_j^*} + r_{K_n} H_{K_n} \tag{3.35}$$

où :

$$\begin{aligned} (w_1, \ldots, w_{K_n}, r_{K_n}) &\sim \mathrm{Dir}(n_1 - d, \ldots, n_{K_n} - d, \alpha + dK_n) \\ H_{K_n} &\sim \mathrm{PY}(d, \alpha + dK_n, G_0) \end{aligned}$$

et H_{K_n} est indépendante de $(w_1, \ldots, w_{K_n}, r_{K_n})$. On peut trouver une preuve détaillée de ce corollaire dans [Car99]. On vérifie facilement que l'échangeabilité de la distribution *a posteriori* est assurée dans l'équation 3.35 puisqu'elle est composée d'une distribution de Dirichlet symétrique et d'un PYP inconditionnel (indépendant des données observées).

Remarque 13. *Nous avons vu que la distribution GEM à deux paramètres associée à la distribution a priori des fréquences limites est invariante par permutation size-biased (ISBP) (cf. Théorème 8) et qu'il y'a équivalence entre échangeabilité de la partition aléatoire sous-jacente et la propriété d'ISBP de la distribution a priori des poids. L'échangeabilité est perdue dans la mise à jour conditionnelle dans la représentation stick-breaking. Dans cette représentation, les labels sont explicites et définis par la loi a priori avant n'importe quelle séquence d'échantillonnage et la loi a posteriori de $H|\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_n$ dépend explicitement de quels labels les $\boldsymbol{\theta}_i$ sont affectés (cf. équation 2.27). Au contraire, dans le modèle d'échantillonnage d'espèces, les atomes correspondent*

aux classes d'équivalence et leurs labels sont définis dans l'ordre d'apparition des nouvelles espèces. Les labels ne sont qu'arbitraires et ne préexistent pas dans le processus d'échantillonnage.

Variantes proposées

Pour échantillonner suivant le processus indépendant infini H_{K_n} dans l'équation 3.35, nous proposons deux variantes :

- la première utilise une version seuillée de l'algorithme du slice sampler de [KGW09] ;
- la seconde est basée sur une troncature de la loi H_{K_n} du PYM. Cette troncature fut à l'origine suggérée dans [IJ01].

1. **Échantillonneur échangeable et seuillé basé sur le « slice »**

   Nous proposons d'abord un échantillonneur utilisant la technique du « slice » comme [Wal07, KGW09] pour échantillonner suivant la distribution *a posteriori* du PYP. Les principales étapes sont maintenant résumées.

   On introduit $\mathbf{u} = (u_1, u_2, \ldots, u_n)$ des variables auxiliaires uniformes telles que la densité jointe de $(\boldsymbol{x}_i, u_i)$ sachant $\boldsymbol{w}$ et $\boldsymbol{\Theta}^*$ est

   $$p(\boldsymbol{x}_i, u_i) = \sum_{k=1}^{\infty} w_k \; f(\boldsymbol{x}_i|\boldsymbol{\theta}_k^*) \, \mathcal{U}(u_i|0, \xi_k) \tag{3.36}$$

   où ξ_k est une variable dépendante telle que pour tout k,

   $$\xi_k = \min(w_k, \zeta) \tag{3.37}$$

   où $\zeta \in]0, 1]$ est indépendante de w_k. Ici, ζ est un seuil que l'on propose afin d'améliorer les propriétés de mélange de l'échantillonneur comparé à Walker [Wal07] et Kalli *et al* [KGW09]. Le seuil ζ peut être une variables aléatoire ou déterministe. Par exemple, une valeur déterministe typique de ζ et qui est un bon compromis entre propriétés de mélange et coût de calcul est le poids moyen du premier atome dans la partie de la loi *a posteriori* indépendante des données, $r_{K_n} H_{K_n}$. On peut l'exprimer par

   $$\zeta = \frac{(\alpha + d\,\mathbb{E}_{\alpha,d}(K_n))(1-d)}{(\alpha+1)(\alpha+n)}$$

   avec $\mathbb{E}_{\alpha,d}(K_n) = \sum_{i=1}^{n} \frac{(\alpha+d)_{i-1\uparrow}}{(\alpha+1)_{i-1\uparrow}}$ (où $(x)_{a\uparrow} = \Gamma(x+a)/\Gamma(x)$), que l'on peut approximer pour n suffisamment grand par $\mathbb{E}_{\alpha,d}(K_n) \approx \frac{\Gamma(\alpha+1)}{d\,\Gamma(\alpha+d)} n^d$ (voir [Pit02]).
   Utilisant l'équation 3.37, on peut réécrire l'équation 3.36 de la façon suivante :

   $$p(\boldsymbol{x}_i, u_i) = \mathbf{1}(\zeta > u_i)\, \zeta^{-1} \sum_{w_k > \zeta} w_k \; f(\boldsymbol{x}_i|\boldsymbol{\theta}_k^*) + \sum_{w_k \leq \zeta} \mathbf{1}(w_k > u_i)\, f(\boldsymbol{x}_i|\boldsymbol{\theta}_k^*)$$

où les deux sommes sont finies car $\#\{j : w_j > \varepsilon\} < \infty$, pour tout $\varepsilon > 0$. L'utilisation de u permet d'échantillonner un nombre fini K^* de poids et de paramètres pour le PYP inconditionnel.

L'échantillonneur de Gibbs permet de générer des variables suivant la distribution *a posteriori* jointe $f(\boldsymbol{\Theta}^*, \boldsymbol{c}, \boldsymbol{w}, \boldsymbol{u}|\boldsymbol{X})$, en tirant suivant chacune des distributions conditionnelles. Comme dans [KGW09], on échantillonne de façon jointe $\mathbf{w}, \mathbf{u}$. Les distributions conditionnelles impliquées sont :

- $f(c|\boldsymbol{\theta}^*, w, u)$,
- $f(\boldsymbol{\theta}^*|c, w, u)$,
- $f(w, u|c, \boldsymbol{\theta}^*) = f(u|w, c, \boldsymbol{\theta}^*) f(w|c, \boldsymbol{\theta}^*)$.

Nous allons maintenant détailler chacune de ces distributions.

(a) Distribution conditionnelle de u :
Il s'agit d'une loi uniforme,

$$u_i|w_1, w_2, \ldots, w_{K_n}, \mathbf{c} \overset{\text{ind.}}{\sim} \mathcal{U}\left(u_i|0, \min\left(w_{c_i}, \zeta\right)\right). \tag{3.38}$$

(b) Distribution conditionnelle de w :
On rappelle que K_n désigne le nombre de clusters distincts sur les n observations et n_k le nombre d'observations appartenant au cluster k.
– Pour les composantes non-vides, la distribution conditionnelle est une loi de Dirichlet :

$$w_1, \ldots, w_{K_n}, r_{K_n}|\mathbf{c} \sim \text{Dir}\left(n_1 - d, \ldots, n_{K_n} - d, \alpha + K_n\, d\right). \tag{3.39}$$

– Pour les composantes vides, il s'agit de la distribution *a priori* dans le GEM :

$$\begin{aligned} v_k &\sim \text{Beta}\left(1 - d, \alpha + k\, d\right) \\ w_k &= v_k\, r_{k-1} \\ r_k &= r_{k-1}\left(1 - v_k\right). \end{aligned} \tag{3.40}$$

Il est important de souligner qu'à chaque itération, les clusters non-vides sont labellisés en fonction de leur *ordre d'apparition* dans l'échantillonnage. On opère ainsi dans l'espace des *classes d'équivalence sur les labels des clusters non-vides* qui sont alors *échangeables* [Pit96b]. L'*a priori* stick-breaking (ISBP pour les PYP) concerne seulement les clusters non-vides à une itération donnée de l'échantillonneur de Gibbs. Comme nous l'avons déjà souligné, cela laisse présager de bonnes propriétés de mélange comparé à l'échantillonnage stick-breaking de [IJ01]. Par ailleurs, l'échantillonneur garde la distribution aléatoire.

(c) Distribution conditionnelle de c :
L'échantillonnage des variables de classification nécessite de calculer une constante de normalisation qui devient faisable par l'utilisation des variables auxiliaires puisque le choix des c_i est dans un ensemble fini.

$$c_i|\mathbf{w}, \mathbf{u}, \boldsymbol{\Theta}^*, \mathbf{X} \overset{\text{ind}}{\sim} \sum_{k=1}^{K^*} w_{k,i}\, \delta_k\left(\cdot\right) \tag{3.41}$$

où

$$w_{k,i} \propto \mathbf{1}\left(w_k > u_i\right) \max\left(w_k, \zeta\right) f\left(\boldsymbol{x}_i | \boldsymbol{\theta}_k^*\right)$$

et $\sum_{j=1}^{K^*} w_{k,i} = 1$.

À noter que dans un souci d'accélérer les calculs, il peut être utile de ranger les poids w_k, $k > K_n$ en ordre décroissant. De ce fait, on peut éviter de faire les tests pour tout $k > \kappa$ dès lors que $w_\kappa < u_i$.

(d) Distribution conditionnelle de $\boldsymbol{\theta}^*$:

$$G_0(\mathrm{d}\boldsymbol{\theta}_k^*) \prod_{i:c_i=k} f(\boldsymbol{x}_i | \boldsymbol{\theta}_k^*). \tag{3.42}$$

La structure de l'échantillonneur par blocs permet d'implanter l'algorithme facilement sur une machine parallèle puisque, à chaque itération, on retient la distribution aléatoire pour l'ensemble des données.

Le schéma d'échantillonnage est décrit dans le tableau 3.12.

Algorithme 9

État permanent de la chaîne de Markov : $\boldsymbol{c} = (c_1, \ldots, c_n)$, $\boldsymbol{w} = (w_1, \ldots, w_{K^*})$, $\boldsymbol{\Theta}^* = (\boldsymbol{\theta}_1^*, \ldots, \boldsymbol{\theta}_{K^*}^*)$

Itérations $t = 1, \ldots, T$

- pour $i = 1, \ldots, n$
  - Tirer u_i suivant l'équation 3.38.
    Faire $u^* = \min\{u_1, \ldots, u_n\}$.
  - Tant que $r_{k-1} > u^*$ (tirage des poids pour $k > K_n$),
    tirer suivant l'équation 3.40.
    Faire $K^* = \min\left(\{k : r_k < u^*\}\right)$.
  - Tirer les paramètres des nouvelles composantes suivant la loi *a priori* :

$$\boldsymbol{\theta}_k^* \overset{\text{i.i.d.}}{\sim} G_0, \text{ pour } K_n < k \leq K^*. \tag{3.43}$$

  - Affecter $\boldsymbol{x}_i$ à l'une des composantes suivant la loi donnée dans l'équation 3.41.
- Pour $k = 1, \ldots, K_n$ (mise à jour des poids),
  tirer $w_1, \ldots, w_{K_n}, r_{K_n}$ suivant l'équation 3.39.
- Pour $k = 1, \ldots, K_n$ (mise à jour des paramètres des composantes),
  tirer $\boldsymbol{\theta}_k^*$ suivant la densité proportionnelle à celle de l'équation 3.42.

TABLE 3.12 – Première variante de notre algorithme.

2. **Échantillonneur de Gibbs échangeable et tronqué**

La deuxième variante de l'algorithme que l'on propose est une alternative à la première. Elle est aussi basée sur la distribution *a posteriori* du PYP donnée

dans l'équation 3.35. Mais au lieu d'utiliser la stratégie du slice sampling pour échantillonner la partie infinie du PYP inconditionnel, nous utilisons une approximation en le tronquant au niveau L. Cette troncature élimine la nécessité d'utiliser des variables auxiliaires. Cela conduit à un PYP tronqué pour la partie indépendante du processus. Ce schéma de troncature a été suggéré par Ishwaran et James [IJ01]. Il consiste à approximer l'équation 3.35 par

$$\sum_{j=1}^{K_n} w_j \delta_{\boldsymbol{\theta}_j^*} + r_{K_n} H_{K_n}^*$$

où $H_{K_n}^*$ est une troncature presque-sûre de H_{K_n}, i.e une troncature de H_{K_n} au niveau L. Le nombre total de composantes représentées est alors $K^* = K_n + L$. Les distributions conditionnelles utilisées par notre échantillonneur de Gibbs sont les suivantes :

(a) Distribution conditionnelle des poids :
- pour les composantes affectées il s'agit de la distribution de Dirichlet,

$$w_1, w_2, \ldots, w_{K_n}, r_{K_n} | \mathbf{c} \sim \mathrm{Dir}\left(n_1 - d, n_2 - d, \ldots, n_{K_n} - d, \alpha + K_n\, d\right). \tag{3.44}$$

- pour les composantes non-affectées ($K_n < k \leq K^*$) c'est la loi GEM,

$$\begin{aligned} v_k &\sim \mathrm{Beta}\left(1 - d, \alpha + k\, d\right) \\ w_k &= v_k\, r_{k-1} \\ r_k &= r_{k-1}\left(1 - v_k\right). \end{aligned} \tag{3.45}$$

- Distribution conditionnelle des variables de classification,

$$c_i | \mathbf{w}, \boldsymbol{\Theta}^*, \mathbf{X} \overset{\text{ind}}{\sim} \sum_{k=1}^{K^*} w_{k,i}\, \delta_k\left(\cdot\right) \tag{3.46}$$

avec

$$w_{k,i} \propto w_k\ f\left(\boldsymbol{x}_i | \boldsymbol{\theta}_k^*\right) \quad \text{et} \quad \sum_{k=1}^{K^*} w_{k,i} = 1.$$

- Distribution conditionnelle des paramètres
  - pour les composantes non-vides, c'est la densité proportionnelle à

$$G_0(\mathrm{d}\boldsymbol{\theta}_k^*) \prod_{i:c_i=k} f(\boldsymbol{x}_i | \boldsymbol{\theta}_k^*) \text{ pour tout } k \leq K_n. \tag{3.47}$$

Le schéma d'échantillonnage de cette deuxième variante est décrit dans le tableau 3.13.

3.2.4 Comparaisons des algorithmes

Dans cette section nous évaluons, sur plusieurs jeux de données, les performances d'un algorithme marginal, de deux algorithmes conditionnels ainsi que des deux variantes de l'échantillonneur que nous proposons. Les algorithmes que nous comparons sont les suivants :

Algorithme 10

État permanent de la chaîne de Markov : $\boldsymbol{c} = (c_1, \ldots, c_n)$, $\boldsymbol{w} = (w_1, \ldots, w_{K^*})$, $\boldsymbol{\Theta}^* = (\boldsymbol{\theta}^*_1, \ldots, \boldsymbol{\theta}^*_{K^*})$

Itérations $t = 1, \ldots, T$

- Mise à jour des variables d'affectation :
  pour $i = 1, \ldots, n$, affecter $\boldsymbol{x}_i$ à l'une des composantes suivant la loi donnée dans l'équation 3.46.
- Mise à jour des poids :
  - pour les composantes non-vides
    tirer $w_1, w_2, \ldots, w_{K_n}, r_{K_n}$ suivant l'équation 3.44.
  - pour les composantes vides
    - tirer w_k pour $K_n < k \leq K^*$ suivant l'équation 3.45.
    - Faire $w_{K^*} = r_{K^*-1}$ tel que $v_{K^*} = 1$.
- Mise à jour des paramètres des composantes non-vides,
  pour $k = 1, \ldots, K_n$, tirer $\boldsymbol{\theta}^*_k$ suivant la densité proportionnelle à celle de l'équation 3.47.

TABLE 3.13 – Deuxième variante de notre algorithme.

- algorithme 8 de Neal [Nea00] (Algo. 8),
- le slice sampler « efficace » de Kalli *et al.* [KGW09] (Slice efficient),
- l'échantillonneur de Gibbs tronqué de Ishwaran et James [IJ01] (Trunc.),
- les deux variantes de notre échantillonneur échangeable, à savoir la version slice (Slice exch. thres) et la version tronquée (Trunc. exch.).

Nous allons aussi examiner de plus près l'apport du seuil que nous proposons. Pour cela nous implantons, en sus, la version slice de notre échantillonneur échangeable mais sans le seuil (Slice exch. without thres.) et le « Slice efficient » de Kalli *et al.* modifié par l'introduction de notre seuil (Slice eff. thres.).

Spécification des données :

Nous avons testé les algorithmes avec $f(\cdot|\boldsymbol{\theta})$ un noyau Gaussien de paramètres $\boldsymbol{\theta}^* = (\mu, \sigma^{-2})$ et G_0 une distribution Normal-Inverse Gamma i.e, $G_0(\mu, \sigma) = \mathcal{N}(\mu|\eta, \kappa^{-1}) \times \text{Gamma}(\sigma^{-2}|\gamma, \beta)$ où $\text{Gamma}(\cdot|\gamma, \beta)$ désigne la distribution Gamma de densité proportionnelle à $x^{\gamma-1}e^{-x/\beta}$. Dans un but de comparaison avec les autres méthodes, nous avons utilisé les mêmes jeux de données qu'elles, à savoir deux jeux de données réelles et deux de données simulées.

Les données simulées ont été générées suivant les deux modèles de mélanges de Gaussiennes suivants :

- un mélange bimodal (bimod) : $0.5\mathcal{N}(-1, 0.5^2) + 0.5\mathcal{N}(1, 0.5^2)$,
- un mélange unimodal lepto-kurtic (lepto) : $0.67\mathcal{N}(0, 1) + 0.33\mathcal{N}(0.3, 0.25^2)$.

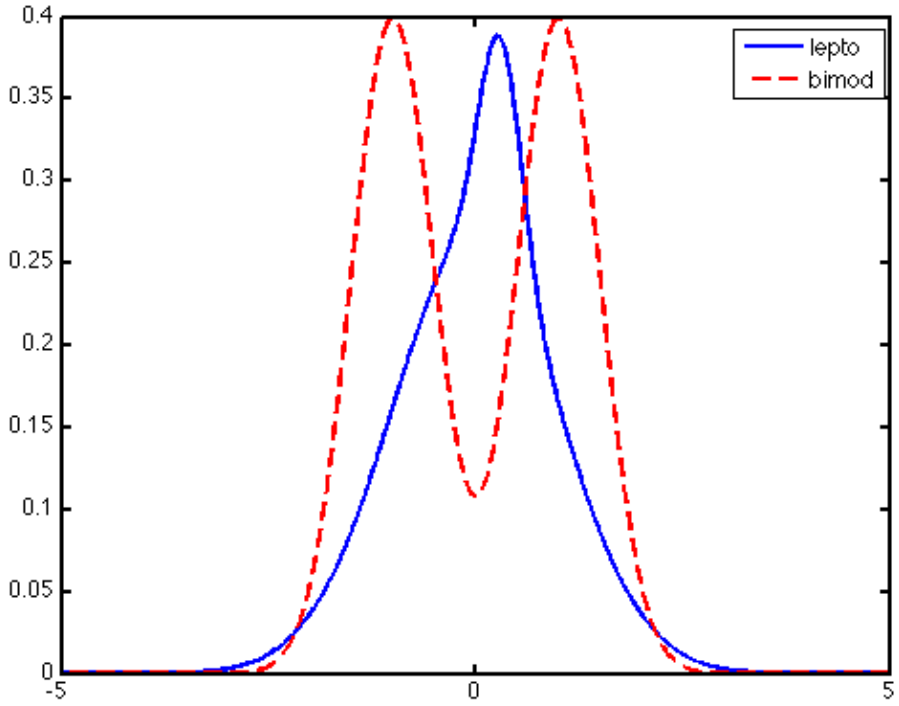


FIGURE 3.3 – Mélange bimodal (bimod) et mélange unimodal lepto-kurtic.

Ces données sont tracées à la Figure 3.3.

Afin de juger les performances des algorithmes sur des données de petite et grande taille, nous avons généré $n = 100$, $n = 1.000$ et $n = 10.000$ observations pour chacun des mélanges.

Les données réelles sont les suivantes :

- Galaxy : ce sont des vélocités (en 10^3 km/s) de 82 galaxies mesurées par rapport à la nôtre. Il s'agit d'un jeu de données souvent utilisé dans les problèmes d'estimation de densité. On peut citer par exemple Escobar et West [EW95], Green et Richardson [GR01].
- S & P 500 : cela consiste en 2023 indices de rendement journaliers. Ces données ont été utilisées dans la première version de l'échantillonneur de Kalli *et al.* [KGW09].

Performances des algorithmes :

- Pour contrôler la convergence des algorithmes, nous surveillons deux quantités : la déviance de la densité estimée et le nombre de clusters.
  La déviance est vue comme une fonction globale de tous les paramètres du modèle et est définie par :

$$D = -2\sum_{i=1}^{n} \log\left(\sum_{j} \frac{n_j}{n} f(x_i|\boldsymbol{\theta}_j^*)\right)$$

  où n_j est la taille du cluster j.

Cette fonctionnelle a aussi été utilisée dans les études comparatives des algorithmes par Neal [Nea00], Green et Richardson [GR01], Papaspiliopoulos et Roberts [PR07] et Kalli *et al.* [KGW09].

- Les performances des algorithmes dans leur régime stationnaire sont jugées en regardant les deux fonctionnelles suivantes : l'IAT et le temps de calcul par itération.

Integrated Autocorrelation Time (IAT)
Dans un algorithme MCMC, les états successifs de la chaîne sont corrélés. De ce fait, la variance des estimées produites est plus grande que dans un tirage statique (échantillonnage indépendant). Les chaînes de Markov lentement mélangeantes exhibent parfois des dépendances après plusieurs centaines d'échantillons. Il est alors possible que la chaîne puisse être considérée comme convergente mais que cette dépendance ne permette pas d'obtenir des estimateurs Monte-Carlo fiables. On peut prendre en compte des impératifs d'*indépendance* entre les valeurs simulées et chercher à fournir des variables quasi-indépendantes. La dépendance des échantillons simulés peut se mesurer par des corrélations. L'IAT est défini par Sokal [Sok97] :

$$\tau = 1 + 2\sum_{j=1}^{\infty} \rho_j$$

où ρ_j désigne la fonction d'autocorrélation de l'échantillon au lag j.
L'IAT est un indicateur d'une bonne ou mauvaise mélangeance et une mesure de l'*efficacité* d'un algorithme MCMC. Il contrôle aussi l'erreur statistique sur les mesures Monte-Carlo. En effet, les échantillons corrélés générés par une chaîne de Markov à l'équilibre entraînent une variance 2τ fois plus grande que dans un tirage indépendant [Sok97]. Si on note par τ_j l'IAT produit par l'algorithme j pour une quantité donnée, alors $\tau_1/\tau_2 = k > 1$ signifie que l'algorithme 1 nécessite k fois plus d'itérations que l'algorithme 2 pour produire la même erreur Monte-Carlo [PR07]. Quand on compare donc deux algorithmes Monte-Carlo alternatifs pour un même problème, le plus efficace est celui qui a le plus petit IAT puisqu'il fournira de meilleures estimées. Signalons aussi que l'IAT a été utilisé par d'autres auteurs pour comparer des méthodes MCMC. On peut citer entre autres [Nea00], [GR01], [PR07], [KGW09].
Cependant, le calcul de l'IAT est difficile en pratique. Suivant [Sok97], un estimateur de τ peut être obtenu en sommant les estimées des autocorrélations jusqu'à un lag L fixé

$$\hat{\tau} = 1 + 2\sum_{j=1}^{L} \hat{\rho}_j, \tag{3.48}$$

où $\hat{\rho}_j$ une estimée de l'autocorrélation normalisée, i.e

$$\hat{\rho}_j = \frac{\hat{C}(j)}{\hat{C}(0)}$$

avec

$$\hat{C}(j) = \frac{1}{N-j}\sum_{i=1}^{N-j}(f_i - \bar{f})(f_{i+j} - \bar{f})$$

pour une fonctionnelle f.

Le choix du point de coupure L dépend de l'utilisateur. C'est un compromis entre d'un côté le biais qu'il introduit

$$\text{biais}(\hat{\tau}) = \frac{1}{2}\sum_{|l|>L}\rho_l + o\left(\frac{1}{N}\right), \tag{3.49}$$

et de l'autre la variance sur $\hat{\tau}$ qu'il contrôle

$$\text{var}(\hat{\tau}) \approx \frac{2(2L+1)}{N}\tau^2 \tag{3.50}$$

où N est la taille de l'échantillon Monte-Carlo.

Toutefois, l'IAT n'est pas suffisant pour juger les performances relatives des algorithmes. Il est aussi important de regarder le temps mis par l'échantillonneur pour effectuer une itération. Nous pensons qu'un « bon » algorithme doit réaliser un bon compromis entre l'IAT et la durée d'une itération.

Paramétrisation des algorithmes :

On prend $d = 0$ dans le PYM, ce qui le réduit ainsi à un DPM. Le paramètre de forme du PYM (paramètre de précision du DPM) est donné par $\alpha = \{1, 0.2, 5\}$.

Les hyperparamètres sont fixés en s'appuyant sur les données comme Green et Richardson [GR01], de la façon suivante : si R est l'étendue des données, on prend $\eta = R/2$, $\kappa = 1/R^2$, $\gamma = 2$ et $\beta = 0.02R^2$.

L'échantillonneur par blocs de Ishwaran et James [IJ01] a été tronqué au niveau $N = 3\alpha\log(n)$, où n est la taille des données. Cela induit une erreur de troncature représentant la distance L_1 entre la densité marginale des données dans le modèle avec et sans troncature (voir [IJ01]). Les erreurs de troncature pour les différents jeux de données sont indiquées dans le tableau 3.14.

Données	ϵ
Galaxy ($n = 82$)	7.4139e-04
Lepto/bimod ($n = 100$)	9.0413e-04
Lepto/bimod ($n = 1000$)	8.2446e-06
S&P 500 ($n = 2023$)	2.2572e-06

TABLE 3.14 – Erreur de troncature de l'algorithme Trunc.

La deuxième variante de l'algorithme que nous proposons a été tronquée au niveau $L = 2\alpha\log(n)$ (pour échantillonner la partie infinie du PYP). L'algorithme 8 de Neal [Nea00] a été testé avec $m = 2$ composantes auxiliaires.

Pour le calcul de l'IAT, nous suivons les instructions de Sokal [Sok97] qui recommande d'utiliser un nombre suffisant d'itérations. Pour chaque jeu de données, nous

avons fait tourner chacun des échantillonneurs sur 2.000.000 itérations dont les 200.000 premières sont écartées pour la période de burn-in. Nous pensons que cela est suffisant pour obtenir des résultats fiables.

Résultats en termes d'IAT :
Nous reportons dans les tableaux 3.15 à 3.20 une partie des résultats, sur les données réelles (Galaxy et S&P 500) et des données simulées (Lepto et Bimod pour $n = 100$ et $n = 1000$). Ici, nous montrons seulement les résultats obtenus avec $\alpha = 1$. Les résultats pour le processus à deux paramètres (Pitman-Yor) ne sont pas non plus présentés. Chaque tableau contient le temps par itération (en millisecondes) et l'estimée de l'IAT (cf équation 3.48) sur le nombre de clusters et sur la déviance. À noter que les algorithmes ont été implantés sans parallélisation. L_D et L_M désignent respectivement les points de coupure pour le calcul de l'IAT sur la déviance et sur le nombre moyen de clusters noté $\bar{K}_n$. Les estimées des erreurs sont mises entre parenthèses (formule 3.50). Les courbes d'autocorrélation sont affichées dans les Figures 3.5 et 3.6 pour une comparaison visuelle

Estimation de densité :
L'échantillonneur de Gibbs permet d'obtenir une approximation de la densité prédictive qui, rappelons le, n'a pas de forme explicite utilisable dans les DPM. Pour cela, on utilise les échantillons générés après convergence. Chaque itération t post burn-in fournit un échantillon de la densité :

$$p_t(x) \approx \sum_{k=1}^{K} w_k^{(t)} f_{\mathcal{N}}\left(x|\boldsymbol{\theta}_k^{*(t)}\right)$$

avec K désignant respectivement le nombre de composantes actives pour la méthode marginale, représentées pour les méthodes tronquées, retenues à chaque itération pour les algorithmes slices. w_k représente le poids de la composante k dans les algorithmes conditionnels, obtenu en prenant la fréquence empirique n_k/n dans l'algorithme marginal. L'estimateur pour la densité est la moyenne sur les tirages :

$$\hat{p} \approx \frac{1}{N} \sum_{t=1}^{N} p_t,$$

où N est le nombre d'itérations MCMC post burn-in.

Nous montrons dans la Figure 3.4 l'histogramme des données et la densité estimée par chacun des algorithmes sur les données Galaxy.

Résultats :
Les courbes des densités estimées, les valeurs des déviances estimées et le nombre moyen de clusters confirment que tous les algorithmes ont effectué une estimation correcte des inconnues et peuvent donc être jugés au travers de leur performance de mélange.

Sur toutes les simulations que nous avons effectuées, il s'est avéré que :
– comme on s'y attendait, l'algorithme 8 marche mieux que tous les algorithmes conditionnels grâce à la structure non-identifiable de l'affectation. De plus, la marginalisation des composantes du mélange accélère la convergence de l'algorithme grâce à la réduction drastique de l'espace. On pourra se référer à Papaspiliopoulos et Roberts [PR07] et Porteous *et al.* [PISW06] pour avoir plus de détails sur les raisons pour lesquelles les approches marginales surpassent les conditionnelles.

- les deux variantes de la méthode que nous proposons marchent mieux que tous les autres algorithmes conditionnels, grâce à l'échangeabilité et l'introduction du seuil. L'algorithme « Slice Efficient » est le moins performant.
- Une différence entre les deux variantes proposées est que dans la version tronquée, la longueur de l'approximation doit être décidée avant l'échantillonnage. La plupart du temps, cela ne constitue pas un problème dans un mélange par processus de Dirichlet. Mais dans un mélange par processus de Pitman-Yor, l'approximation peut donner lieu à des estimations biaisées pour des longueurs de troncature modérées. La version slice échangeable seuillée que nous avons proposée réalise une troncature adaptative à chaque itération et constitue un bon compromis entre IAT et temps de calcul.

Interprétation :
Nous pensons que la mélangeance lente dûe à la non-échangeabilité dans la représentation stick-breaking *a posteriori* est accentuée par l'absence des poids dans l'étape d'affectation des algorithmes utilisant les variables slices. Cela pourrait freiner l'échantillonneur de Gibbs dans l'étape d'affectation, pour changer par exemple une observation d'une composante ayant peu d'observations à une qui en a beaucoup. L'introduction du seuil que nous proposons devrait faciliter un tel changement. Pour valider cette conjecture, nous avons examiné l'effet de ce seuil sur l'algorithme « Slice Efficient ». Cette introduction décroit l'autocorrélation et est visuellement notable sur la Figure 3.5, sur les différences entre les courbes bleue (le « Slice Efficient » original) et verte (la version seuillée). De plus, le seuil rend la différence en termes d'IAT petite entre le « Slice Efficient » seuillé et l'échantillonneur tronqué qui, lui, prend en compte les poids lors de la mise à jour des variables d'affectation. D'un autre côté, quand on enlève le seuil dans la première variante de notre algorithme, l'IAT augmente (on peut voir les différences entre les courbes rouge et noire). D'une façon générale, nous avons pu noter sur tous les jeux de données et de paramètres testés, que le seuil fait décroître l'autocorrélation plus vite, particulièrement dans les premiers lags. Cependant, cela augmente légèrement le temps de calcul par itération.

Nous nous intéressons maintenant au gain apporté par la propriété d'échangeabilité. Cela est illustré par les différences entre la courbe rouge (Slice exch. thres.) et la courbe verte (Slice eff. thres.) et dans les différences entre les courbes violette (Trunc. exch.) et vert-forêt (Trunc.). On peut aussi remarquer, sur la courbe noire, que la courbe d'autocorrélation obtenue par notre algorithme slice sans seuil décroît et atteint 0 plus vite que les algorithmes non échangeables (Trunc., Slice eff. thres et Slice efficient). Ce comportement a aussi été observé sur les autres jeux de données, même pour un grand nombre de données. Cette propriété est due à l'échangeabilité.

Enfin, il est très important de souligner que les deux variantes de notre algorithme ainsi que l'algorithme de Neal [Nea00] sont les seuls à avoir été stables pour toutes les expériences : sur des simulations variées, nous avons toujours obtenu les mêmes résultats sur chaque type de données. En revanche, pour les échantillonneurs non-échangeables ([IJ01] et [KGW09]) nous avons noté une convergence erratique, particulièrement avec de grandes données (par exemple lepto avec $n = 10,000$).

Tables de comparaison ($d = 0$, $\alpha = 1$)

	IAT sur $\bar{K}_n$	*IAT sur la déviance*	$\bar{K}_n$ *estimé*	*Déviance estimée*
Slice exch. thres	14.48(0.37)	2.88(0.05)	3.986(0.93)	1561.14(21.61)
Trunc. exch.	14.42(0.37)	2.94(0.05)	3.989(0.93)	1561.16(21.69)
SE without thres.	35.52(0.92)	4.77(0.09)	3.989(0.93)	1561.15(21.61)
Truncated	38.65(1.00)	3.63(0.07)	3.996(0.94)	1561.15(21.66)
Slice efficient	60.65(1.57)	5.28(0.10)	3.991(0.93)	1561.15(21.62)
Slice eff. thres.	37.82(0.98)	3.61(0.07)	3.986(0.93)	1561.08(22.17)
Algo 8 ($m = 2$)	8.25(0.21)	2.57(0.05)	3.987(0.93)	1561.16(21.62)

TABLE 3.15 – Galaxy $n = 82$, $L_D = 150$, $L_M = 300$.

	IAT sur $\bar{K}_n$	*IAT sur la déviance*	$\bar{K}_n$ *estimé*	*Déviance estimée*
Slice exch. thres	22.58(0.75)	145.07(3.06)	4.977(0.82)	14990.47(57.56)
Trunc. exch.	21.59(0.72)	148.69(3.14)	4.978(0.82)	14990.50(59.09)
SE without thres.	93.93(3.13)	194.26(4.10)	4.976(0.82)	14990.35(57.80)
Truncated	32.75(1.09)	148.66(3.14)	4.975(0.81)	14990.94(59.44)
Slice efficient	105.92(3.53)	204.63(4.32)	4.965(0.81)	14991.21(61.14)
Slice eff. thres.	34.87(1.16)	145.46(3.07)	4.969(0.82)	14990.56(60.53)
Algo 8 ($m = 2$)	13.55(0.45)	106.28(2.24)	4.980(0.82)	14990.38(59.09)

TABLE 3.16 – S&P 500 $n = 2023$, $L_D = 200$, $L_M = 500$.

	IAT sur $\bar{K}_n$	*IAT sur la déviance*	$\bar{K}_n$ *estimé*	*Déviance estimée*
Slice exch. thres	28.76(0.74)	5.85(0.11)	3.801(1.66)	287.59(8.46)
Trunc. exch.	28.51(0.74)	6.00(0.11)	3.808(1.67)	287.58(8.48)
SE without thres.	70.56(1.82)	9.54(0.17)	3.799(1.67)	287.58(8.43)
Truncated	54.38(1.40)	5.89(0.11)	3.789(1.66)	287.58(8.42)
Slice efficient	99.92(2.58)	8.76(0.16)	3.784(1.65)	287.58(8.42)
Slice eff. thres.	55.41(1.43)	5.65(0.10)	3.794(1.66)	287.61(8.58)
Algo 8 ($m = 2$)	15.59(0.40)	5.20(0.09)	3.794(1.66)	287.59(8.52)

TABLE 3.17 – Bimod $n = 100$, $L_D = 150$, $L_M = 300$.

	IAT sur $\bar{K}_n$	*IAT sur la déviance*	$\bar{K}_n$ *estimé*	*Déviance estimée*
Slice exch. thres	93.28(3.93)	3.65(0.07)	3.806(1.73)	2735.14(8.66)
Trunc. exch.	91.20(3.85)	3.69(0.07)	3.795(1.72)	2735.14(8.66)
SE without thres.	228.64(9.64)	5.44(0.10)	3.809(1.73)	2735.15(8.67)
Truncated	156.27(6.60)	3.71(0.07)	3.777(1.71)	2735.13(8.62)
Slice efficient	257.25(10.85)	5.13(0.09)	3.766(1.68)	2735.12(8.61)
Slice eff. thres.	150.13(6.33)	3.81(0.07)	3.798(1.71)	2735.15(8.72)
Algo 8 ($m = 2$)	47.25(1.99)	3.06(0.06)	3.798(1.72)	2735.14(8.65)

TABLE 3.18 – Bimod $n = 1000$, $L_D = 150$, $L_M = 800$.

	IAT sur $\bar{K}_n$	*IAT sur la déviance*	$\bar{K}_n$ *estimé*	*Déviance estimée*
Slice exch. thres	25.61(0.85)	13.53(0.29)	3.991(1.64)	239.74(11.38)
Trunc. exch.	24.78(0.83)	13.46(0.28)	3.983(1.63)	239.75(11.31)
SE without thres.	90.56(3.02)	42.74(0.90)	3.991(1.63)	239.73(11.26)
Truncated	41.22(1.37)	17.03(0.36)	4.001(1.64)	239.72(11.31)
Slice efficient	120.71(4.03)	46.28(0.98)	3.979(1.64)	239.77(11.29)
Slice eff. thres.	44.73(1.49)	16.98(0.36)	3.989(1.64)	239.77(11.82)
Algo 8 ($m = 2$)	14.79(0.49)	9.83(0.28)	3.994(1.63)	239.72(11.36)

TABLE 3.19 – Lepto $n = 100$, $L_D = 200$, $L_M = 500$.

	IAT sur $\bar{K}_n$	*IAT sur la déviance*	$\bar{K}_n$ *estimé*	*Déviance estimée*
Slice exch. thres	235.49(9.93)	13.17(0.24)	4.006(2.05)	2400.51(18.68)
Trunc. exch.	237.70(10.02)	13.66(0.25)	4.022(2.08)	2400.48(18.72)
SE without thres.	462.09(19.49)	18.75(0.34)	3.973(1.99)	2400.53(18.73)
Truncated	294.24(12.41)	12.67(0.23)	3.958(2.01)	2400.47(18.49)
Slice efficient	472.95(19.95)	16.91(0.31)	3.864(1.92)	2400.45(18.26)
Slice eff. thres.	302.50(12.76)	13.80(0.25)	3.978(2.07)	2400.53(18.93)
Algo 8 ($m = 2$)	148.81(6.28)	11.55(0.21)	4.018(2.07)	2400.48(18.69)

TABLE 3.20 – Lepto $n = 1000$, $L_D = 150$, $L_M = 800$.

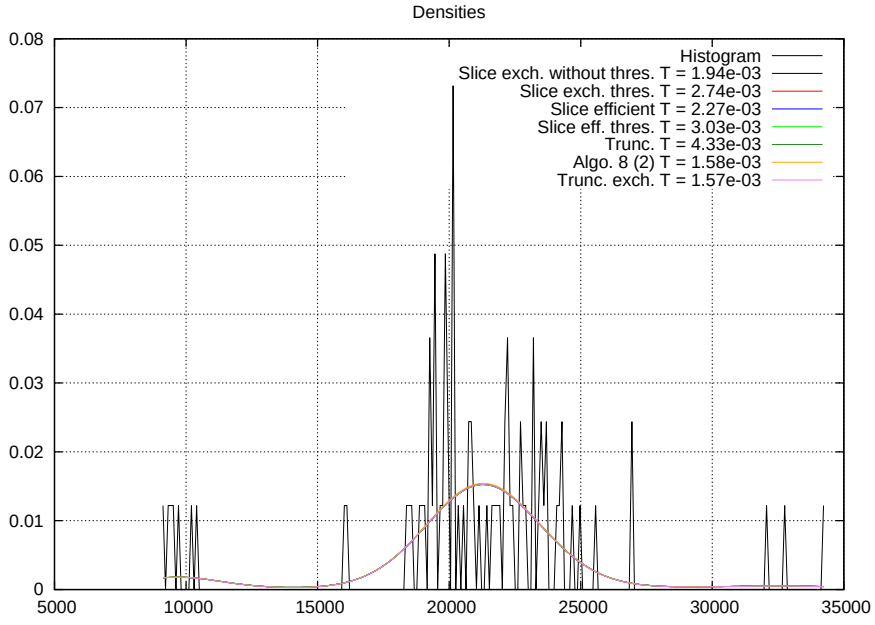


FIGURE 3.4 – Histogramme des données et densités estimées (Galaxy).

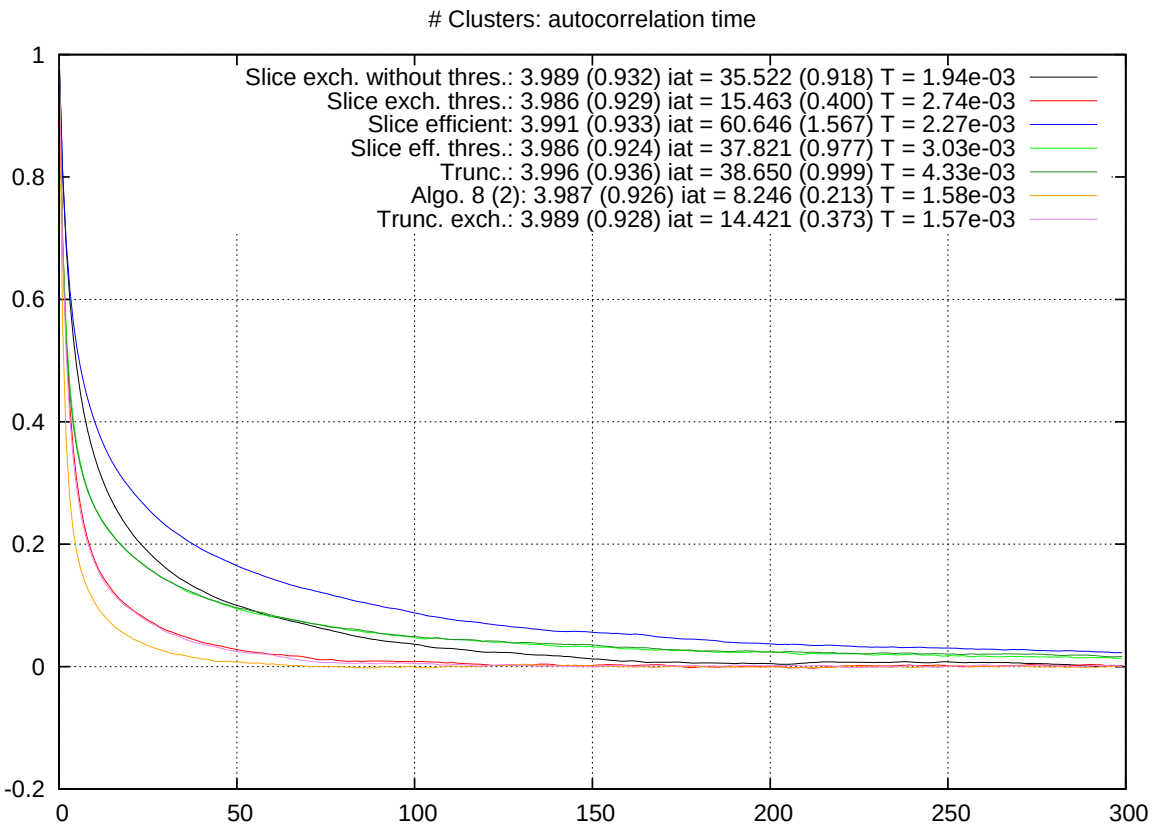


FIGURE 3.5 – Courbes d'autocorrélation utilisées pour estimer l'IAT sur le nombre de clusters (Galaxy).

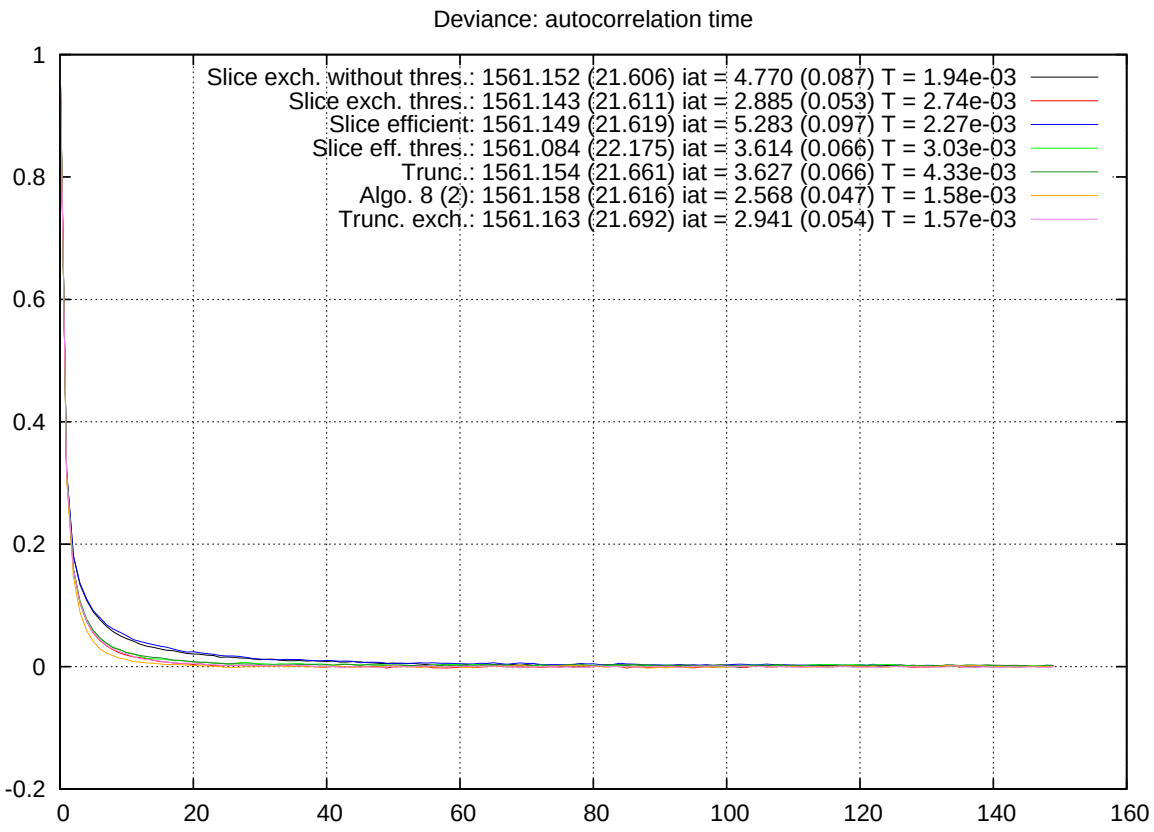


FIGURE 3.6 – Courbes d'autocorrélation utilisées pour estimer l'IAT sur la déviance (Galaxy).

3.3 Conclusion

Quand les modèles deviennent de plus en plus complexes, la mélangeance lente d'un algorithme peut être sévèrement inhibant. Il semble alors être utile voire primordial de développer des modèles permettant d'améliorer le mélange et d'expérimenter des stratégies pour réduire le temps de calcul. Dans un souci de satisfaire une contrainte de capacité de parallélisation tout en maintenant des propriétés de mélange proches des approches type urne de Pólya, nous avons développé deux schémas d'échantillonnage pour les modèles de mélange par processus de Pitman-Yor. Ces échantillonneurs sont ainsi applicables aux modèles de mélange par processus de Dirichlet. Par ailleurs, l'introduction du seuil que nous proposons permet d'améliorer le mélange dans l'algorithme du « Slice Efficient » qui peut être utilisé pour des modèles de mélanges basés sur des *a priori* stick-breaking plus généraux, autres que le processus de Dirichlet et le processus de Pitman-Yor.

Nous avons tenté, dans les méthodes que nous avons proposées, de combiner les propriétés d'échantillonnage par blocs des approches conditionnelles qui retiennent la distribution aléatoire avec l'échangeabilité dans le modèle tout comme les algorithmes basés sur l'urne de Pólya. Les échantillonneurs résultants fournissent des propriétés de mélange très intéressantes. En particulier, notre algorithme slice échangeable et seuillé (Slice exch. thres) conduit à un bon compromis entre IAT et temps de calcul, tout en évitant une troncature sèche du modèle. Dans notre étude expérimentale, il est apparu que les algorithmes conditionnels standards peuvent présenter des résultats inattendus de défaut de convergence. Cet inconvénient est surtout renforcé pour les données de grande taille. Au contraire, l'algorithme marginal de Neal et les techniques d'échantillonnage que nous avons proposées présentent un comportement stable pour toutes les situations.

4

Reconstruction spatiale en TEP 3D

Dans ce chapitre, nous procédons à une description succincte des principes de formation, de détection et d'acquisition des données en TEP. Par la suite, les méthodes analytiques et statistiques de reconstruction des images en TEP sont abordées. Pour finir nous présentons la nouvelle méthode de reconstruction que nous avons proposée et procédons à une étude comparative avec les approches de maximum de vraisemblance et de maximum *a posteriori*.

4.1 Introduction à la TEP

La tomographie par émission de positons (TEP) ou PET en anglais pour Positron Emission Tomography est une technique d'imagerie médicale fonctionnelle permettant de mesurer *in vivo* chez l'Homme ou chez l'animal, et avec une précision de quelques millimètres, la distribution d'un paramètre physiologique (par exemple le métabolisme du glucose, le débit sanguin etc.). Cette cartographie est obtenue à partir de la distribution volumique et temporelle d'un radiotraceur spécifique injecté au patient. La TEP diffère des techniques d'imagerie structurelle ou anatomique (rayons X, imagerie par résonance magnétique) disposant d'une meilleure résolution spatiale mais limitées à la production d'images anatomiques.

4.1.1 Principe

La TEP repose sur le principe physique de l'émission de positons. Un traceur radioactif est administré au patient. Ce dernier est obtenu en incorporant un radioisotope dans une molécule biologique traceuse. Un exemple typique de molécule vectrice utilisée pour l'étude du métabolisme du glucose est le déoxyglucose (DG) qui est un analogue du glucose. Lorsque cette molécule est combinée au ^{18}F (radioisotope), le résultat est appelé

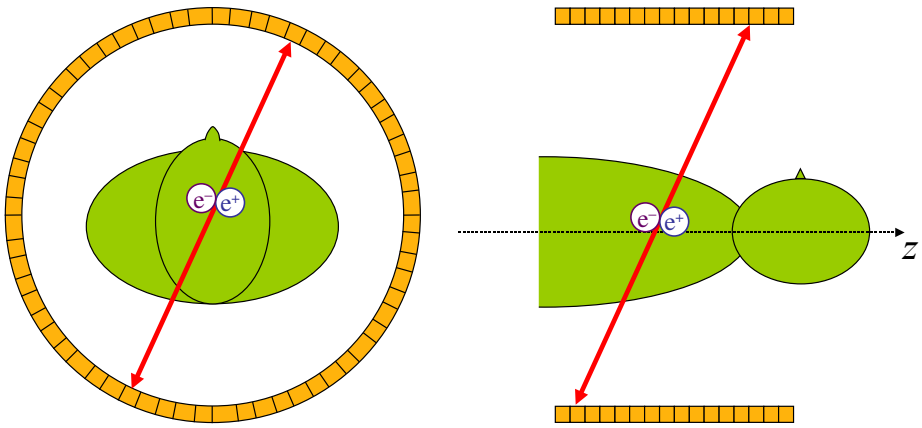


FIGURE 4.1 – Un positon e^+ s'annihile avec un électron e^- conduisant à l'émission de deux photons gamma partant en sens opposé (flèche rouge). Les photons sont enregistrés par les détecteurs tout autour du patient (couronne orange).

fluoro-déoxyglucose ($[^{18}F]$–FDG). On le désigne sous le nom de radiopharmaceutique, radiotraceur, traceur radioactif ou simplement traceur.

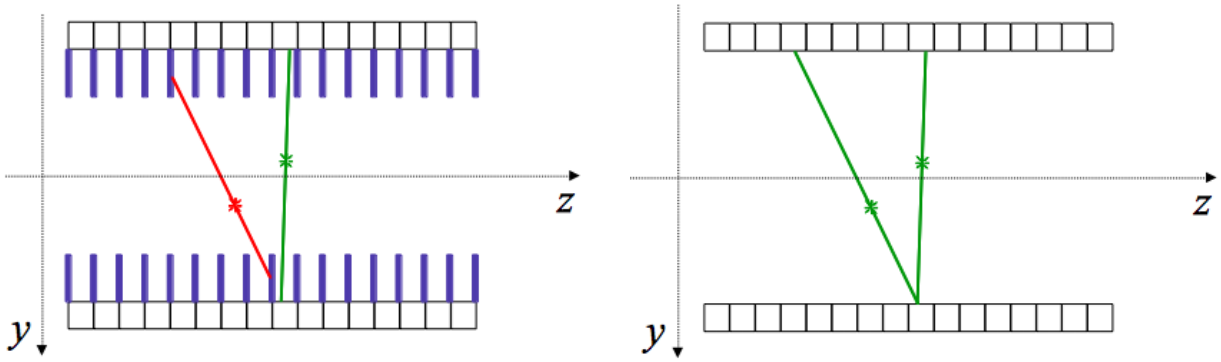
Le radioisotope utilisé en TEP présente une instabilité due à un surplus de protons dans son noyau. Il cherche donc à revenir à un état stable par transformation d'un proton p en neutron n ; on parle de désintégration $\beta+$. Cette réaction est à l'origine de l'émission d'un neutrino (ν) et d'un positron (antiparticule de l'électron, noté e^+) et elle s'écrit :

$$p \rightarrow n + \nu + e^+.$$

Le positron parcourt une courte distance dans les tissus (libre parcours du positron) en dissipant son énergie cinétique par des collisions avec des électrons périphériques. Au repos, le positron s'annihile avec un électron et cela produit deux photons gamma de 511 keV qui partent simultanément en sens opposé (voir Figure 4.1). C'est cette émission à 180° qui est à la base de la TEP.

4.1.2 Détection

Les photons issus de l'annihilation interagissent avec les électrons du milieu par effet photoélectrique, diffusion Compton ou Rayleigh et peuvent être absorbés ou diffusés. Seule une fraction de ces photons s'échappe sans avoir interagi. Ils sont détectés par la couronne de détecteurs entourant le patient. Le tomographe est souvent constitué de plusieurs couronnes circulaires (appelées aussi anneaux), chacune comprenant plusieurs détecteurs. Des circuits électroniques à son intérieur permettent de comparer les temps d'arrivée des photons. La détection simultanée (dite en coïncidence) de deux photons, c'est-à-dire dans un intervalle de temps de quelques nanosecondes au maximum, suppose



(a) (b)

FIGURE 4.2 – Dans le mode d'acquisition 2D (Figure (a)), on place des septas matérialisés ici en bleu. Seules les coïncidences issues des plans directs i.e $\Delta z = 0$ (ligne verte) ou adjacents i.e $|\Delta z| = 1$ sont enregistrées. Les coïncidences symbolisées par la ligne rouge ne sont pas enregistrées. En acquisition 3D (Figure (b)), toutes les coïncidences sont enregistrées.

qu'ils proviennent de la même annihilation. On réalise ce que l'on appelle une collimation électronique. Cette détection en coïncidence suppose que l'annihilation a eu lieu à l'intérieur du volume joignant les deux détecteurs. Ce volume a typiquement la forme d'un parallélépipède allongé et est appelé tube de réponse (noté TOR pour tube of response). Ce tube est souvent assimilé à la ligne joignant les centres des deux détecteurs, appelée ligne de réponse (LOR).

4.1.3 Acquisition des données

Modes d'acquisition

L'acquisition consiste à enregistrer le nombre d'évènements dans chaque LOR. On distingue deux modes d'acquisition : le bidimensionnel et le tridimensionnel.

L'acquisition bidimensionnelle (2D) appelée aussi coupe-par-coupe était en vigueur dans les premiers tomographes. Elle consiste à placer des collimateurs annulaires (septas) devant les détecteurs et à n'enregistrer que les LOR orthogonales à l'axe du tomographe (cf. Figure 4.2). Toutefois, cela diminue la sensibilité du système en éliminant une grande partie des évènements.

Le mode d'acquisition 3D (volumique) exploite pleinement la collimation électronique et enregistre tous les évènements entre toutes les couronnes de détecteurs. Cela augmente la sensibilité du système mais s'accompagne aussi d'une augmentation de la contamination par les évènements fortuits et diffusés (définis plus loin à la page 124).

Paramétrisation des LOR

Lorsqu'une coïncidence a lieu, les identifiants des cristaux impliqués sont sauvegardés. Cela équivaut à repérer la position de la LOR dans l'espace des détecteurs.

En 2D, une LOR est repérée par l'angle azimuthal définissant son orientation, $\phi \in [0, \pi]$, et la position radiale x_r, avec $x_r \in [-R, R]$ où R est le rayon de l'anneau de détection (cf Figure 4.3). Ainsi, à chaque LOR, est associée un repère mobile (x_r, y_r) tel que le changement de coordonnées de (x, y) à (x_r, y_r) s'écrive :

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix} \begin{pmatrix} x_r \\ y_r \end{pmatrix}.$$

On a alors $x_r = x\cos\phi + y\sin\phi$.

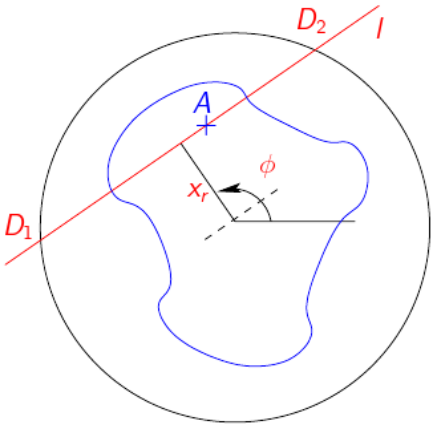


FIGURE 4.3 – Une émission ayant eu lieu au point A est détectée par la LOR l joignant les centres des deux détecteurs D_1 et D_2. La LOR est paramétrée par la position radiale x_r et l'angle azimuthal ϕ.

En 3D, une LOR l est

1. soit repérée par les positions D_1 et D_2 des deux détecteurs dans leur couronne et l'identification a_1 et a_2 des couronnes,
2. soit repérée par son orientation (θ, ϕ) et sa position dans le plan 2D orthogonal à la LOR (x_r, y_r). Les paramètres sont alors la position radiale x_r, y_r, l'angle azimuthal ϕ et l'angle co-polaire θ (cf. Figure 4.4).

Le changement de coordonnées est donné par :

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -\sin\phi & -\cos\phi\sin\theta & \cos\phi\cos\theta \\ -\cos\phi & -\sin\phi\sin\theta & \sin\phi\cos\theta \\ 0 & \cos\theta & \sin\theta \end{pmatrix} \begin{pmatrix} x_r \\ y_r \\ z_r \end{pmatrix}. \tag{4.1}$$

L'axe z_r de la LOR s'écrit $z_r = z_r(\theta, \phi) = [\cos\phi\cos\theta, \sin\phi\cos\theta, \sin\theta]$.

Sinogramme ou list-mode

Les évènements sont enregistrés sous deux formes : le mode sinogramme et le mode liste (list-mode).

1. Dans le mode liste, on enregistre séquentiellement les coïncidences détectées en précisant pour chacune le temps d'arrivée ainsi que la position du couple de détecteurs. Chaque évènement est ainsi repéré spatialement (D_1, D_2, a_1, a_2) et temporellement. Dans nos travaux de reconstruction, nous opérons en list-mode.

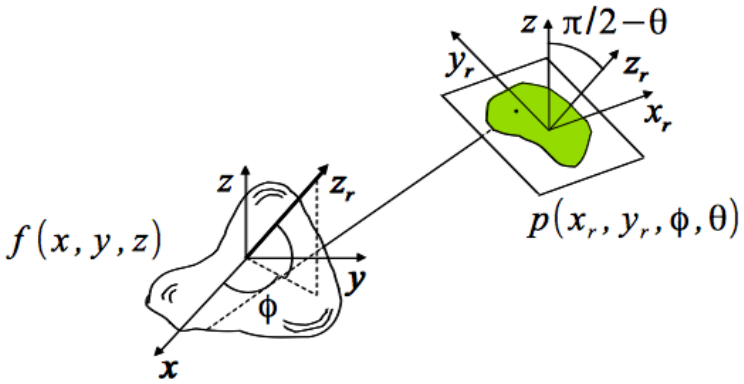


FIGURE 4.4 – Projections en 3D

2. Le sinogramme est une matrice contenant, pour une durée donnée, l'ensemble des projections de l'objet pour toutes les distances et tous les angles. Les projections sont reliées aux mesures par le modèle de l'intégrale de ligne qui suppose que la somme des évènements sur une LOR est égale à la projection de f le long de la LOR. En 2D, la projection 1D de l'objet est donnée par :

$$p(x_r,\phi) = \int_{-\infty}^{+\infty} f(x,y)dy_r = \int_{-\infty}^{+\infty} f(x_r\cos\phi - y_r\sin\phi, x_r\sin\phi + y_r\cos\phi)dy_r$$

où $f(x,y)$ est la distribution de l'activité au point (x,y) du champ de vue, pour un z_0 donné. Le sinogramme est alors une matrice 2D contenant tous les éléments de projection indexés par ϕ et r. Chaque ligne de la matrice correspond à la projection mono-dimensionnelle de l'objet pour un angle azimuthal donné. Le nombre de lignes est égal au nombre d'angles de mesure ϕ et le nombre de colonnes est donné par le nombre d'éléments de projection pour une position angulaire. Par symétrie, $p(x_r,\phi+\pi) = p(-x_r,\phi)$, et on ne stocke que les projections pour $\phi \in [0,\pi[$.
En 3D, les projections 2D de la distribution radioactive $f(x,y,z)$ s'écrivent suivant la transformée en rayons X 3D,

$$p(x_r,y_r,\phi,\theta) = \int f(x,y,z)dz_r.$$

(cf. Figure 4.4 et équation 4.1). Le sinogramme est une matrice 4D décrivant l'espace des projections 2D.

Chaque élément de la matrice sinogramme est appelé bin. Si une coïncidence est détectée sur une LOR, le bin correspondant dans le sinogramme est alors incrémenté. La taille du sinogramme peut être réduite en regroupant les LOR d'angles de projection adjacents en une seule projection. Un bin du sinogramme peut donc correspondre à une LOR entre deux ou plusieurs paires de détecteurs.

L'avantage du sinogramme est de réduire la taille du fichier de sauvegarde par une représentation compacte des données. Cependant, l'information temporelle est perdue et le ré-échantillonnage spatial induit une perte en résolution spatiale. Le format liste permet de conserver l'information temporelle et se fait sans perte d'information. L'inconvénient est que le fichier contenant les évènements est très volumineux et sa taille croît en fonction de l'activité dans le champ de vue du

tomographe et de la durée de l'examen (cf. [Def07]).

Une fois les données acquises, l'étape suivante consiste à la reconstruction des images TEP. Cependant l'obtention d'images quantitatives nécessite de corriger au préalable, ou d'inclure dans la modélisation, les phénomènes venant perturber la mesure.

Correction des phénomènes physiques

Le modèle physique simple que nous avons décrit auparavant et sur lequel repose la TEP est perturbé par des phénomènes biaisant la mesure. Ces phénomènes dépendent aussi bien du système d'imageur que du patient. En effet, le système de détection recueille trois types de signaux de photons issus de coïncidences dites vraies, fortuites et diffusées. Lorsque deux photons détectés proviennent d'une même annihilation et n'ont subi aucune interaction avec l'organisme entre leurs points d'émission et de détection, on dit que c'est une *coïncidence vraie*. Le système enregistre aussi des coïncidences accidentelles quand deux photons provenant de l'annihilation de deux positons différents sont détectés dans la même fenêtre temporelle. On parle de *coïncidences fortuites ou aléatoires*. Enfin, l'un des deux photons d'annihilation peut diffuser sur des électrons lors de son parcours dans l'organisme ou dans le système (diffusion Compton) : ce sont les *coïncidences diffusées*. Cette diffusion s'accompagne souvent d'une modification de la direction de propagation et l'annihilation est assignée à une position incorrecte. Ces deux phénomènes induisent un biais dans la mesure en ce sens que l'annihilation est localisée à tort sur la LOR supposée. Seules les coïncidences vraies constituent le signal permettant de reconstruire la distribution volumique du traceur. Les coïncidences fortuites et diffusées doivent donc être corrigées directement sur les mesures, ou inclues dans l'étape de reconstruction.

La correction des coïncidences fortuites peut se faire en soustrayant l'estimation des coïncidences fortuites aux coïncidences acquises. Pour cela, une façon de faire est d'estimer le taux de coïncidences fortuites à partir du nombre de photons détectés par chaque détecteur suivant la relation :

$$F_{D_1,D_2} = 2\tau S_{D_1} S_{D_2},$$

où

- F_{D_1,D_2} désigne le taux de coïncidences fortuites sur la LOR joignant les détecteurs D_1 et D_2,
- S_{D_1} (respectivement S_{D_2}) est le taux de photons détectés par le détecteur D_1 (respectivement D_2),
- 2τ est la largeur de la fenêtre de coïncidence.

La discrimination des photons diffusés s'effectue en partie dans le module de détection car la validation d'un évènement en coïncidence est basée sur un critère temporel (évènements détectés dans la même fenêtre temporelle) mais aussi sur un critère énergétique (350-650 keV). Le but de la fenêtre énergétique est de chercher à enregistrer les évènements non diffusés. En effet, les photons ayant subi au moins une diffusion perdent une partie de leur énergie. Cependant, la faible résolution énergétique des détecteurs ne permet pas une discrimination efficace. Il existe des méthodes de calcul

basées sur la section efficace de Klein-Nishina permettant d'estimer le taux de photons diffusés et de les soustraire aux données.

L'absorption dans l'organisme d'une grande partie des photons nécessite de compenser la perte de signal. Ce phénomène dit d'atténuation photonique doit être corrigé car il entraîne des inhomogénéités dans l'image reconstruite, en sous-évaluant l'activité en profondeur et en amplifiant celle en périphérie. L'atténuation est un phénomène non-isotrope, qui varie en fonction de la nature des milieux traversés, de l'énergie des photons etc. En TEP, l'atténuation est une fonction globale de la LOR. La correction peut donc s'effectuer en multipliant les mesures dans chaque LOR par les coefficients d'atténuation de la LOR en question. Ces coefficients s'obtiennent en réalisant deux examens de transmission utilisant des sources externes de positons. Le coefficient d'atténuation de la LOR est alors obtenu en divisant le comptage par LOR dans la mesure à vide (sans le patient) par le comptage obtenu en présence du patient. La correction de l'atténuation peut aussi être faite en utilisant une carte d'atténuation obtenue par tomodensitométrie (appelé couramment scanner ou tomographie de transmission par rayons X).

Les données TEP doivent aussi être normalisées à cause de la non-uniformité de la sensibilité des LOR. En effet, pour une même activité, la sensibilité est variable d'une LOR à une autre. Cette sensibilité est fonction de termes physiques (efficacité intrinsèque des détecteurs) mais aussi de termes géométriques dûs à la courbure du tomographe (effets d'arc).

Il est à noter aussi que d'autres phénomènes perturbent la mesure et doivent être corrigés ou modélisés afin d'obtenir une image quantitative. On peut citer entre autres le libre parcours du positron avant l'annihilation, l'acolinéarité des trajectoires des photons, le temps mort etc.

Enfin, comme l'activité décroit dans le temps à cause de la décroissance radioactive, il est nécessaire de corriger cette décroissance pour que les données reconstruites ne rendent compte que des variations biologiques.

4.2 Méthodes de reconstruction tomographique

Une fois l'acquisition effectuée, la distribution volumique du traceur doit être reconstruite à partir des coïncidences mesurées. Les méthodes de reconstruction spatiale en tomographie d'émission sont usuellement divisées en deux grandes catégories : les méthodes analytiques et les méthodes statistiques.

1. L'algorithme de reconstruction des premiers systèmes d'imagerie TEP s'appuie sur l'inversion de la transformée en rayons X, ce sont les méthodes dites analytiques. Elles sont basées sur une modélisation des données et de l'image dans un espace continu et de l'inversion de la transformée en rayons X en 3D. Cette dernière est un opérateur mathématique reliant la fonction f à ses intégrales de lignes. La transformée de Radon relie une fonction à ses intégrales sur des hyperplans. La transformée en rayons X coïncide avec la transformée de Radon en 2D (cf. [Nat86]).

   Les méthodes analytiques sont considérées en TEP comme des algorithmes déterministes ne prenant pas en compte la statistique de l'émission et de la détection. Elles conduisent à des images présentant des artéfacts de reconstruction.

2. Par opposition, on appelle méthodes statistiques celles basées sur l'approche du maximum de vraisemblance (ML) ou sur les approches bayésiennes comme la méthode du maximum *a posteriori* (MAP). Elles abordent le problème de reconstruction sous une forme discrétisée et minimisent de façon itérative l'anti-logarithme de la vraisemblance poissonienne des données (ML) ou une version pénalisée de cet anti-logarithme (MAP). Comparées aux méthodes analytiques, elles modélisent de façon plus réaliste le processus d'acquisition des données.

4.2.1 Méthodes analytiques

Elle sont basées sur le modèle de l'intégrale de ligne. L'hypothèse qui est faite est que le nombre d'évènements enregistrés par deux détecteurs en coïncidence est proportionnel au nombre de désintégrations ayant eu lieu sur la LOR joignant les deux détecteurs. Le but est de reconstruire la distribution d'activité f à partir des ses projections mesurées pour $|x_r| \leq R$ et $0 \leq \phi \leq \pi$. Cela nécessite donc au préalable de supprimer tous les évènements dits « parasites », à savoir les coïncidences fortuites et diffusées, et de corriger de l'atténuation ainsi que de la non-uniformité de l'efficacité des détecteurs.

Reconstructions analytiques en 2D

La rétroprojection simple ou épandage

Une première approche pour reconstituer f à partir des projections consiste à considérer l'opérateur de rétroprojection simple. Pour cela, on inverse le processus de projection en propageant la valeur mesurée sur une LOR à l'ensemble de ses points. L'estimateur pour f s'obtient en sommant les contributions de toutes les LOR ainsi propagées,

$$\hat{f}(x,y) = b(x,y) = \int_0^\pi p(x_r = x\cos\phi + y\sin\phi, \phi)\mathrm{d}\phi.$$

Cependant, cette façon d'épandre induit des artéfacts dits en étoile qui sont des zones d'activité fictive. Une méthode permettant d'éliminer ces artéfacts en étoile consiste par exemple à utiliser la transformée de Fourier inverse (notée DIFT pour Direct Inverse Fourier Transform). Cela repose sur le théorème de la coupe centrale que nous allons maintenant voir.

Théorème de la section centrale

Soit $F(\nu_x, \nu_y)$ la transformée de Fourier bidimensionnelle de $f(x,y)$ définie par,

$$F(\nu_x, \nu_y) = \int_{\mathbb{R}^2} f(x,y)e^{-2i\pi\langle\boldsymbol{\nu},\boldsymbol{x}\rangle}\mathrm{d}x\mathrm{d}y$$

où $\boldsymbol{\nu} = (\nu_x, \nu_y)'$ sont les fréquences spatiales associées à $\boldsymbol{x} = (x,y)'$; $\langle\boldsymbol{\nu}, \boldsymbol{x}\rangle$ est le produit scalaire entre $\boldsymbol{\nu}$ et $\boldsymbol{x}$.
Soit $P(\nu, \phi)$ la transformée de Fourier de la projection $p(x_r, \phi)$ donnée par

$$P(\nu,\phi) = \int_{\mathbb{R}} p(x_r,\phi)e^{-2i\pi\nu x_r}\mathrm{d}x_r$$

où ν est la fréquence associée à x_r.

Le théorème de la section centrale relie la transformée de Fourier 1D d'une projection d'angle ϕ à la transformée de Fourier 2D de l'image suivant

$$P(\nu, \phi) = F(\nu_x, \nu_y)$$

où $\nu_x = \nu \cos\phi$ et $\nu_y = \nu \sin\phi$.

Démonstration.

$$\begin{aligned} P(\nu,\phi) &= \int_{\mathbb{R}} p(x_r,\phi) e^{-2i\pi\nu x_r} \mathrm{d}x_r \\ &= \int_{\mathbb{R}} \left[\int_{\mathbb{R}} f(x,y) \mathrm{d}y_r \right] e^{-2i\pi\nu x_r} \mathrm{d}x_r \\ &= \int_{\mathbb{R}} \int_{\mathbb{R}} f(x,y) e^{-2i\pi\nu(x\cos\phi + y\sin\phi)} \mathrm{d}x\mathrm{d}y \\ &= \int_{\mathbb{R}} \int_{\mathbb{R}} f(x,y) e^{-2i\pi(x\nu_x + y\nu_y)} \mathrm{d}x\mathrm{d}y \end{aligned}$$

où $\nu_x := \nu\cos\phi$ et $\nu_y := \nu\sin\phi$. □

Reconstruction par transformée de Fourier inverse (DIFT)

D'après le théorème de la section centrale, si on dispose des projections pour tous les angles $\phi \in [0, \pi]$, on peut retrouver la transformée de Fourier F de f et reconstruire l'image en utilisant la formule d'inversion de la transformée de Fourier,

$$f(x,y) = \int_{\mathbb{R}^2} F(\nu_x,\nu_y) e^{2i\pi(x\nu_x + y\nu_y)} \mathrm{d}\nu_x \mathrm{d}\nu_y.$$

Le problème est qu'en pratique, on dispose d'une nombre limité de projections. Par ailleurs, le calcul de $F(\nu_x, \nu_y)$ nécessite de passer de la grille polaire des projections $P(\nu, \phi)$ à la grille rectangulaire de $F(\nu_x, \nu_y)$. L'implantation de cette technique nécessite donc une étape d'interpolation des données de projection. On lui préfère plutôt une autre technique basée sur la rétroprojection filtrée.

La rétroprojection filtrée (FBP)

On considère la transformée de Fourier inverse qui donne $f(x, y)$ à partir de l'espace fréquentiel,

$$f(x,y) = \int_{\mathbb{R}^2} F(\nu_x,\nu_y) e^{2i\pi(x\nu_x + y\nu_y)} \mathrm{d}\nu_x \mathrm{d}\nu_y.$$

Utilisant le théorème de la coupe centrale, on peut remplacer la transformée de Fourier de l'image par celle des projections,

$$f(x,y) = \int_{\mathbb{R}} \int_{\mathbb{R}} P(\nu,\phi) e^{2i\pi(x\nu_x + y\nu_y)} \mathrm{d}\nu_x \mathrm{d}\nu_y.$$

On peut appliquer le théorème de changement de variable en considérant l'application $F : (\nu, \phi) \to (\nu_x = \nu\cos\phi,\ \ \nu_y = \nu\sin\phi)$. Le jacobien de la transformation étant ν, on a

$$\begin{aligned} f(x,y) &= \int_{\mathbb{R}} \int_{\mathbb{R}} P(\nu,\phi) e^{2i\pi(x\nu_x + y\nu_y)} \mathrm{d}\nu_x \mathrm{d}\nu_y. \\ &= \int_0^{2\pi} \left[\int_{\mathbb{R}} P(\nu,\phi) \nu e^{2i\pi\nu x_r} \mathrm{d}\nu \right] \mathrm{d}\phi. \end{aligned}$$

où $x_r = x\cos\phi + y\sin\phi$.
La symétrie implique que le point (ν, ϕ) a la même valeur que point $(-\nu, \phi + \pi)$. On peut donc utiliser la valeur absolue de ν pour parcourir le plan fréquentiel et faire varier ϕ de 0 à π. L'équation devient alors,

$$f(x,y) = \int_0^{\pi} \left[\int_{\mathbb{R}} P(\nu,\phi)|\nu| e^{2i\pi\nu x_r} \mathrm{d}\nu \right] \mathrm{d}\phi.$$

L'équation interne $\int_{\mathbb{R}} P(\nu,\phi)|\nu| e^{2i\pi\nu x_r}\mathrm{d}\nu$ est la transformée de Fourier inverse de la transformée de Fourier de la projection multipliée par la valeur absolue de ν. Cette multiplication de la transformée de Fourier des projections par la valeur absolue de ν correspond à un filtrage par le filtre rampe $H(\nu) = |\nu|$.
La méthode de rétroprojection filtrée (FBP, pour filtered backprojection) consiste à reconstruire $f(x,y)$ par la rétroprojection des projections filtrées $\hat{p}(x_r,\phi) = \int_{\mathbb{R}} P(\nu,\phi)|\nu| e^{2i\pi\nu x_r}\mathrm{d}\nu$:

$$f(x,y) = \int_0^{\pi} \hat{p}(x_r,\phi).$$

L'algorithme FBP consiste donc d'abord à filtrer les projections par le filtre rampe, puis appliquer une transformée de Fourier inverse et enfin rétroprojeter le résultat. Le filtre rampe annule la composante continue représentant la moyenne du signal et introduit des valeurs négatives [Bru02]. Cela permet d'effacer les artéfacts en étoile laissés par les autres projections durant la rétroprojection. L'inconvénient est qu'il amplifie les hautes fréquences, correspondant aux détails de l'image mais aussi au bruit. Ce problème peut être partiellement résolu en introduisant un filtrage lissant du type un filtre passe-bas dont le rôle est d'éliminer les hautes fréquences. Les projections sont donc lissées par un filtre global combinant le filtre rampe et le filtre passe-bas, c'est la fenêtre dite d'apodisation. La fréquence de coupure du filtre passe-bas détermine la force du lissage : plus elle est basse, plus l'image est lissée. Cela nécessite donc de réaliser un compromis entre la réduction du bruit et la perte de résolution spatiale. Le choix de la fréquence doit être déterminé par deux facteurs. D'une part, elle doit dépendre du rapport signal sur bruit, i.e du nombre de coïncidences détectées. Plus la statistique de comptage est élevée, plus la fréquence de coupure doit être grande, et inversement pour une faible statistique. D'autre part, la fréquence maximale qui peut être recouvrée à partir des données est donnée par la fréquence de Nyquist ν_N, reliée au pas d'échantillonnage Δx_r du sinogramme par

$$\nu_N = \frac{1}{2\Delta x_r}.$$

La fréquence de coupure maximale est donc ν_N.

La rétroprojection filtrée est formulée théoriquement dans l'espace continu des données et suppose un échantillonnage continu des projections. Son implantation pratique nécessite de discrétiser l'axe x_r suivant $x_r = k.\Delta x_r, k = 0, \ldots, N_{x_r} - 1$ et les angles ϕ_j suivant $\phi_j = j.\Delta\phi, j = 0, \ldots, N_\phi - 1$. Après échantillonnage, la transformée de Fourier ainsi que son inverse sont remplacées par leurs versions discrètes. Au final, l'algorithme FBP s'écrit,
Pour un angle ϕ fixé :

1. calculer la transformée de Fourier discrète des projections,
2. la multiplier par le filtre rampe et le filtre passe-bas,

3. calculer la transformée de Fourier inverse pour chaque projection filtrée,
4. rétroprojeter les projections filtrées,
5. répétition des étapes 1 à 4 pour chaque angle ϕ.

Reconstruction analytique en 3D

Nous avons présenté l'algorithme de rétroprojection filtrée dans le cas bidimensionnel. La rétroprojection filtrée peut également s'écrire en 3D mais nécessite de prendre en compte les spécificités de l'acquisition 3D comme la redondance et la troncature des données. D'une part, en 3D le filtre de reconstruction n'est pas unique à cause de la redondance des données et il existe une famille de filtres satisfaisant à une condition énoncée par Defrise *et al.* [DCT95]. Le filtre de reconstruction le plus utilisé est celui de Colsher [Col80].

D'autre part, la rétroprojection filtrée nécessite des projections parallèles et non tronquées. Or dans le mode d'acquisition 3D, du fait de la longueur axiale finie (sur l'axe z), les projections ne peuvent pas être mesurées pour tous les angles. Les projections sont dites alors tronquées ou incomplètes [Nat86]. Cette troncature est illustrée dans la Figure 4.5 et nous y reviendrons dans la section 4.3.

Une des solutions proposées pour la reconstruction analytique en 3D consiste à réarranger les projections et à appliquer les techniques de reconstruction du 2D [DKT+97]. Une autre solution consiste à compléter d'abord les données manquantes puis à reconstruire les images avec un algorithme 3D [KR89]. Ces données sont complétées en utilisant une reconstruction par FBP 2D des données acquises pour les projections droites ($\theta = 0$) puis à projeter l'image reconstruite dans l'espace des projections. Nous ne détaillerons pas plus ces aspects sur la théorie de la rétroprojection 3D appliquée à la TEP puisqu'elle n'est pas étudiée dans cette thèse. Plus de détails peuvent être trouvés dans les livres [Tow98] et [VB03].

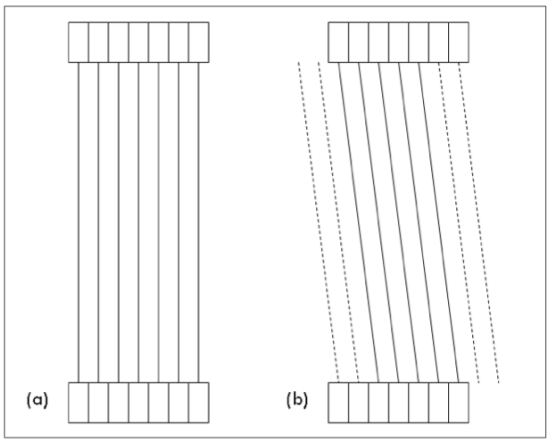


FIGURE 4.5 – Illustration de la troncature des projections en 3D. La figure (a) correspond aux projections mesurées pour $\theta = 0$ et la figure (b) à celles pour $\theta \neq 0$. Les lignes en pointillés représentent les projections non mesurées. Les projections sont complètes uniquement pour $\theta = 0$.

4.2.2 La rétroprojection filtrée vue comme un modèle statistique

Nous avons vu que les algorithmes de rétroprojection filtrée reposent sur l'inversion directe de la transformée en rayons X. Bien que qualifiées de méthodes analytiques dans la reconstruction en tomographie d'émission (par opposition aux méthodes statistiques), le FBP peut être vu comme un problème d'estimation indirecte de densité basé sur les méthodes à noyaux. Le noyau K_{δ_n} de l'estimateur est un filtre b-bande, $b = 1/\delta_n$ et le pas δ_n représente la largeur de bande du noyau. C'est-à-dire que K_{δ_n} est une fonction dont la transformée de Fourier est nulle à l'extérieur de la boule de rayon $1/\delta_n$. Nous allons l'illustrer dans le cadre 2D pour en simplifier la présentation.

Soit f la fonction de densité de probabilité de la distribution radioactive du traceur, telle que $f(\boldsymbol{x})$ désigne l'activité présente au point $\boldsymbol{x}$ de coordonnées (x, y) (voir Figure 4.6). La densité de probabilité dans l'espace des détecteurs peut être modélisée par

$$g(x_r, \phi) = \frac{1}{2\rho} \int_{-\rho}^{\rho} f(x = x_r \cos\phi - y_r \sin\phi, y = x_r \sin\phi + y_r \cos\phi) \mathrm{d}y_r,$$

où la variable d'intégration y_r est la coordonnée le long de la ligne. Cette intégrale est la transformée de Radon (TR) normalisée de f. En posant $\vec{\phi} = (\cos(\phi), \sin(\phi))'$ et $\boldsymbol{x} = (x, y)'$, on peut réécrire :

$$g(x_r, \phi) = \frac{1}{2\rho} \int_{\boldsymbol{x}:\langle \boldsymbol{x}, \vec{\phi} \rangle = x_r} f(\boldsymbol{x}) \mathrm{d}\boldsymbol{x}.$$

La reconstruction consiste en l'inversion de la TR g pour reconstruire la densité f dans l'espace image. Plus précisément, supposons qu'on ait n observations (x_{r_i}, ϕ_i) tirées *iid* de la densité g. Le but de l'inversion est de construire un estimateur $\hat{f}$ pour la densité inconnue f, à partir de $\{(x_{r_i}, \phi_i,), i = 1, \ldots, n, x_{r_i} \in [-R, +R], \phi_i \in [0, \pi]\}$ où R est le rayon de l'anneau de détection.
On définit une fonction « noyau » pour l'estimateur :

$$K_{\delta_n}(u) \propto \int_0^{\frac{1}{\delta_n}} r \cos(ur) \, \mathrm{d}r \text{ pour tout } u \in \mathbb{R}.$$

La fonction K_{δ_n} a la forme des noyaux usuels mais n'est pas un noyau à proprement parler. Il ne satisfait pas le lemme de Bochner puisque la convolution de K_{δ_n} avec g ne tend pas vers g mais plutôt vers f [Cav00]. En d'autres termes, ce noyau approxime l'inverse de la TR. La fonction K_{δ_n} correspond en fait à un filtre b-bande à bande limitée, avec $b = 1/\delta_n$ où δ_n représente la largeur de bande (fenêtre) du noyau. La transformée de Fourier de K_{δ_n} est nulle à l'extérieur de la boule de rayon b. En effet, la transformée de Fourier de K_{δ_n} est donnée par

$$F_{K_{\delta_n}}(\nu) \propto |\nu| \, \mathbf{1}_{|\nu| \leq \frac{1}{\delta_n}}.$$

On utilise la fonction K_{δ_n} pour construire un estimateur pour f

$$\hat{f}(\boldsymbol{x}) = \frac{1}{n} \sum_{i=1}^{n} K_{\delta_n} \left(\langle \boldsymbol{x}, \vec{\phi}_i \rangle - x_{r_i} \right),$$

où $n = \#$ évènements, $\vec{\phi} = (\cos(\phi), \sin(\phi))'$, $\boldsymbol{x} = (x, y)'$.
L'estimateur $\hat{f}$ peut être vu comme une rétroprojection évènement par évènement. Le choix du paramètre δ_n détermine la résolution spatiale de $\hat{f}$.

Cette approche est intéressante d'un point de vue théorique car elle permet d'obtenir des résultats sur la vitesse de convergence et l'efficacité de ces estimateurs linéaires ([JS90], [Cav00]). Ces estimateurs sont asymptotiquement efficaces dans le cadre minimax. Toutefois, l'estimateur est sous-optimal d'un point de vue pratique en ce sens qu'il traite tous les évènements de la même façon. Par ailleurs, comme les algorithmes de rétroprojection, cette approche repose sur une version idéalisée du problème de reconstruction. En effet, il est supposé que la direction de propagation des photons est une ligne qui peut être observée, c'est le modèle de l'intégrale de ligne. Cela repose sur l'hypothèse implicite de l'existence d'un continuum de détecteurs. Cependant en pratique, nous sommes confrontés aux limitations du système comme le nombre restreint de détecteurs. L'ouverture des détecteurs doit être prise en compte dans la modélisation puisque l'on n'a pas une ligne mais un tube (voir Figure 4.7). Enfin, la limitation principale est le nombre relativement restreint d'évènements détectés. Parce qu'une substance radioactive est injectée au patient, le nombre total de molécules radiomarquées est limité. Cela résulte donc en un nombre restreint d'évènements détectés comparé à la tomographie en rayons X par exemple et conduit à un faible rapport signal sur bruit dans les données. De plus, la reconstruction tomographique est un problème inverse mal posé au sens où une petite modification dans les données entraîne une grande variation dans l'image reconstruite. Cela se traduit dans l'algorithme de rétroprojection filtrée par l'amplification du bruit dans les hautes fréquences. Tout cela combiné (problème inverse mal posé et problème idéalisé) conduit à des images bruitées présentant de gros artéfacts de reconstruction. La régularisation est effectuée via la largeur du filtre passe-bas ou, de façon équivalente, via la fenêtre du noyau dans l'estimation indirecte de densité. La fréquence de coupure du filtre définit la force de lissage. Plus elle est basse et plus les hautes fréquences, en particulier le bruit, seront atténués. Mais cela a également pour conséquence de dégrader le contraste, c'est-à-dire d'empêcher de récupérer les détails de l'image. Le choix du filtre est donc important dans la reconstruction par rétroprojection filtrée et nécessite un compromis entre résolution spatiale et bruit. La taille de la la fenêtre de coupure est souvent un choix empirique, basé sur la qualité subjective de l'image.

Une meilleure approche consiste à régulariser pendant la reconstruction en pénalisant la vraisemblance poissonienne des observations. Ces approches permettent de modéliser les processus aléatoires régissant l'émission et la détection. Les techniques de maximum de vraisemblance (ML pour Maximum Likelihood) minimisent l'antilog de la vraisemblance poissonienne des données ([SV82], [LC84], [VSK85]). Pour limiter le bruit dans les images ML, les techniques de régularisation bayésienne ont été développées. Elles maximisent la loi *a posteriori*, composée d'un terme d'adéquation aux données (vraisemblance) et d'un terme de régularisation (loi *a priori*) [GM87], [Gre90], [BS96]. Contrairement aux méthodes de rétroprojection filtrée (FBP), ces approches ne sont pas basées sur la méthode de l'intégrale de ligne et permettent ainsi d'incorporer dans la modélisation plus d'effets physiques et de traiter des géométries non-standard. Elles ont en général un meilleur SNR que les méthodes FBP.

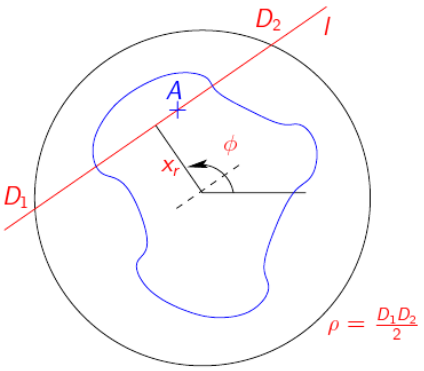


FIGURE 4.6 – Représentation schématique du problème idéalisé en 2D : patient, cercle de détection, LOR. Une émission ayant eu lieu au point A est détectée par la LOR D_1D_2. Cette dernière est paramétrée par deux variables définissant sa localisation et son orientation, le rayon x_r et l'angle ϕ.

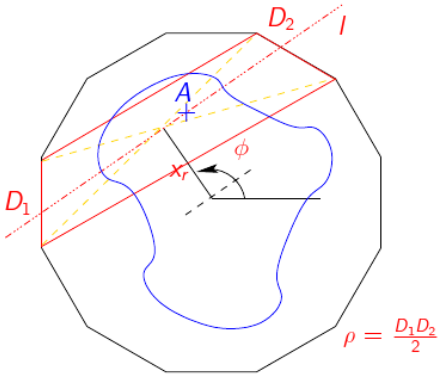


FIGURE 4.7 – Représentation schématique du problème réel en 2D : les directions des photons sont dans un volume parallélépipédique appelé TOR (pour tube of response) délimité par les ouvertures des détecteurs D_1 et D_2.

4.2.3 Les méthodes du maximum de vraisemblance (ML) et du maximum *a posteriori* (MAP)

Les méthodes analytiques que nous avons vues dans la section précédente modélisent l'image comme une fonction continue $f(\boldsymbol{x}) = f(x, y)$ (ou $f(\boldsymbol{x}) = f(x, y, z)$ en 3D) et les projections sous la forme d'une fonction continue $p(x_r, \phi)$ (ou $p(x_r, y_r, \phi, \theta)$ pour le 3D). Leur implantation pratique nécessite de discrétiser les projections et l'image à reconstruire. Les méthodes itératives que nous allons maintenant aborder reposent sur une représentation discrète des données et de l'image. Ces méthodes sont indépendantes du mode d'acquisition utilisé. L'image est représentée comme un vecteur de J pixels (en 2D) ou de voxels (3D), $\boldsymbol{f} = f_1, \ldots, f_J$ où f_j est l'intensité dans le pixel ou le voxel j. Les données sont représentées sous la forme d'un vecteur $\boldsymbol{y} = y_1, \ldots, y_L$. Les méthodes nécessitent d'abord de définir une base de représentation de l'image, puis de décrire un modèle de formation des données, et enfin de définir une fonction de coût et un algorithme permettant son optimisation. Ces différentes étapes sont maintenant décrites.

Modélisation de l'image

Soit $f(\boldsymbol{x})$ la concentration du traceur au point $\boldsymbol{x}$. On fait l'hypothèse que la fonction objet peut être décrite par une combinaison linéaire de fonctions de bases b_j représentant une décomposition du champ de vue :

$$f(\boldsymbol{x}) \approx \sum_{j=1}^{J} f_j b_j(\boldsymbol{r}),$$

pour $\boldsymbol{x} \in \mathbb{R}^k$. La représentation la plus utilisée est le pixel en 2D ou le voxel en 3D (cf. [VB03], [Tow98]). La base des pixels est définie par :

$$\begin{aligned} b_j(x, y) &= 1 \quad \text{si } |x - x_j| < \Delta x/2 \text{ et } |y - y_j| < \Delta y/2 \\ &= 0 \quad \text{si } |x - x_j| \geq \Delta x/2 \text{ ou } |y - y_j| \geq \Delta y/2 \end{aligned}$$

où $j = (j_x, j_y)$ est le numéro du pixel, le centre du pixel j est $(x_j = j_x \Delta x, y_j = j_y \Delta y)$.

Cette représentation s'étend au 3D pour définir le voxel comme un élément de volume parallélépipédique.

Cette paramétrisation du problème réduit le problème de reconstruction à la détermination du vecteur $\{f_j, j = 1, \ldots, J\}$ où f_j représente l'activité dans le pixel ou le voxel j.

Modélisation des données

Les données sont représentées par le vecteur contenant l'ensemble des projections, $\boldsymbol{y} = \{y_1, \ldots, y_L\}$ où y_i désigne le nombre d'évènements enregistrés dans la LOR i. Il est possible de se ramener au cas du mode liste en choisissant $y_i = 1$.

Dans les techniques dites algébriques, y_i est modélisée par

$$y_i = \sum_{j=1}^{J} a_{ij} f_j$$

a_{ij} représente la contribution du voxel j à la mesure dans le TOR i. L'élément a_{ij} représente la sensibilité s du détecteur i au point j du champ de vue (FOV),

$$a_{ij} = \int_{\text{FOV}} b_j(\boldsymbol{x})\, s(i|\boldsymbol{x})\, \mathrm{d}\boldsymbol{x}.$$

Elle est typiquement prise comme valant 1 à l'intersection du TOR i et du voxel j, et 0 ailleurs. Tous les éléments a_{ij} sont regroupés dans une matrice $\boldsymbol{A}$ appelée matrice système. On peut donc écrire le processus d'acquisition sous la forme suivante

$$\boldsymbol{y} = \boldsymbol{A}\boldsymbol{f}.$$

Tous les phénomènes linéaires intervenant lors de l'acquisition statique (atténuation, inefficacité des détecteurs, résolution non-uniforme des détecteurs etc.) peuvent être inclus dans la modélisation de $\boldsymbol{A}$. Il s'agit d'une matrice de grande taille, $\boldsymbol{A} \in \mathbb{R}^{L \times J}$ où L est le nombre de LORs possibles et J le nombre d'éléments de base (nombre de pixels ou de voxels). Elle peut être factorisée sous la forme [LQ00], [QLC+98] :

$$\boldsymbol{A} = \boldsymbol{A}_{sens}\boldsymbol{A}_{atten}\boldsymbol{A}_{reso}\boldsymbol{A}_{geom}\boldsymbol{A}_{resoImage}$$

où

- $\boldsymbol{A}_{sens}$ est une matrice diagonale contenant les facteurs d'efficacité de chaque paire de détecteurs. Elle peut être mesurée par des procédures de calibration et fournie à l'utilisateur.
- $\boldsymbol{A}_{atten}$ est la matrice d'atténuation. C'est une matrice diagonale contenant les facteurs d'atténuation (probabilités de survie des photons). Elle peut être obtenue par les procédures de correction d'atténuation déjà décrites.
- $\boldsymbol{A}_{reso}$ appelée aussi matrice de flou du sinogramme. Elle modélise les phénomènes créant une incertitude sur la position réelle du point d'impact jouant ainsi sur la résolution. Ces phénomènes sont dûs à l'acolinéarité, la diffusion inter-cristaux etc. Elle est obtenue, par exemple, par simulation Monte-Carlo ou mesurée.
- $\boldsymbol{A}_{geom}$ est la matrice de projection géométrique appelée aussi projecteur. Chaque élément de la matrice est égale à la probabilité géométrique qu'une paire de photons de photons émis au niveau du pixel ou du voxel j atteigne le détecteur i en l'absence de milieu atténuant. Il s'agit d'une matrice de grande taille mais creuse car elle comporte de nombreuses valeurs nulles puisque la probabilité de détection dans une LOR de coïncidences issues d'un voxel éloigné de cette LOR est nulle.
  Une modélisation grossière de cette matrice considère a_{ij} comme la longueur du segment i traversant le pixel j. Ce modèle correspond à la transformée de Radon en 2D ou en rayons X en 3D et ne prend pas en compte l'ouverture des détecteurs. Pour cela, on doit calculer la surface d'intersection du TOR avec le voxel.
  Une modélisation plus correcte repose sur le calcul de l'angle solide formé par les couples de détecteurs, vu depuis chacun des voxels [QLC+98].
- $\boldsymbol{A}_{resoImage}$ est la matrice de flou de l'image. Elle modélise les phénomènes non pris en compte dans $\boldsymbol{A}_{reso}$ comme par exemple le libre parcours du positron.

Contrairement aux méthodes algébriques, dans les techniques statistiques l'hypothèse qui est faite est que y_i est une variable aléatoire, dont l'espérance s'écrit

$$\mathbb{E}(y_i) = \sum_{j=1}^{J} a_{ij} f_j + s_i + r_i \tag{4.2}$$

où

- a_{ij} représente la probabilité qu'une coïncidence produite au voxel j soit détectéé par la LOR i. Ces éléments sont regroupés dans la matrice système qui est la même que dans les techniques algébriques,
- s_i est le nombre de coïncidences diffusées et
- r_i le nombre de coïncidences fortuites.

Le vecteur des projections est donc un vecteur aléatoire dont l'espérance est $\mathbb{E}(\boldsymbol{y}) = \{\mathbb{E}(y_1), \ldots, \mathbb{E}(y_L)\}$. On peut l'écrire sous la forme matricielle suivante :

$$\bar{\boldsymbol{y}} = \mathbb{E}(\boldsymbol{y}) = \boldsymbol{A}\boldsymbol{f} + \boldsymbol{s} + \boldsymbol{r}. \tag{4.3}$$

Si on néglige le temps mort, les données sont les mesures de comptage dans chaque LOR i et sont modélisées comme des variables aléatoires indépendantes suivant la loi de Poisson. La fonction de vraisemblance est alors

$$p(\boldsymbol{y}|\boldsymbol{f}) = \prod_{i=1}^{L} \frac{\exp(-\bar{y}_i)\bar{y}_i^{y_i}}{y_i!}.$$

Remarque : Dans l'algorithme FBP, les données sont pré-corrigées avant d'appliquer l'algorithme. Dans les méthodes ML et MAP, aucune correction ne doit être effectuée au risque de violer l'hypothèse poissonienne des données et fausser ainsi la reconstruction. L'algorithme doit donc être appliqué aux données brutes et tous les effets physiques (atténuation etc.) doivent être inclus dans la matrice système.

Définition de la solution

Utilisant l'équation 4.3, le problème de reconstruction équivaut à déterminer $\boldsymbol{f}$ à partir de $\boldsymbol{A}$ et $\boldsymbol{y}$. Une inversion directe de la matrice $\boldsymbol{A}$ n'est pas possible d'abord parce qu'elle n'est pas forcément carrée et aussi du fait de sa très grande taille (le nombre de LOR $L \approx 10^6$ et peut aller jusqu'à 10^9 en 3D). De plus, la matrice est mal conditionnée. On préfère donc utiliser des techniques itératives reposant sur l'optimisation d'un critère. Pour cela, on fait appel à une fonction de coût définissant la proximité entre l'objet à reconstruire et les données mesurées, $\Phi(\boldsymbol{f}, \boldsymbol{y})$. L'estimateur est alors choisi comme la valeur qui maximise la fonction de coût,

$$\boldsymbol{f}^* = \arg\max_{\boldsymbol{f}} \Phi(\boldsymbol{f}, \boldsymbol{y})$$

sous la contrainte $f_j \geq 0$. On va présenter deux approches possibles : la méthode du maximum de vraisemblance (ML) et la méthode du maximum *a posteriori* (MAP).

1) La méthode du maximum de vraisemblance (ML)

L'image est reconstruite en maximisant la vraisemblance Poisson ou, de façon équivalente, son logarithme vu comme une fonction de $\boldsymbol{f}$:

$$L(\boldsymbol{y}|\boldsymbol{f}) = L(\boldsymbol{f}) = \sum_{i=1}^{L} (-\bar{y}_i + y_i \log(\bar{y}_i) - \log(y_i!)) = \Phi(\boldsymbol{f}, \boldsymbol{y}).$$

La solution du maximum de vraisemblance est

$$\hat{\boldsymbol{f}}_{ML} = \arg\max_{\boldsymbol{f}} L(\boldsymbol{f}).$$

Pour calculer cette solution, l'un des algorithmes les plus souvent utilisés en TEP est la méthode du maximum de vraisemblance basée sur l'algorithme EM (notée ML-EM pour maximum-likelihood expectation-maximization). Il fut introduit par Dempster *et al.* [DLR77] et appliqué en TEP par Shepp et Vardi [SV82] et Lange et Carson [LC84]. Sa version accélérée, OSEM, fut proposée par Hudson et Larkin [HL94].

Pour simplifier la présentation de l'algorithme EM, nous en présentons une dérivation où les coïncidences diffusées et fortuites ne sont pas prises en compte. L'algorithme EM est basé sur l'introduction de variables cachées complètes mais non-observées $\boldsymbol{x}$, reliant les variables incomplètes observées $\boldsymbol{y}$ à $\boldsymbol{f}$. L'algorithme utilise alternativement deux étapes, entre le calcul de l'espérance conditionnelle (étape E) de la log-vraisemblance des variables complètes, et l'estimée courante de $\boldsymbol{f}$ maximisant cette quantité (étape M). L'algorithme est de la forme :

$$\begin{aligned} \text{Étape E} : \Phi(\boldsymbol{f}|\boldsymbol{f}^{(k)}) &= \mathbb{E}\left(\ln L(\boldsymbol{x}|\boldsymbol{f})|\boldsymbol{y}, \boldsymbol{f}^{(k)}\right) \\ \text{Étape M} : \boldsymbol{f}^{(k+1)} &= \arg\max_{\boldsymbol{f}} \Phi(\boldsymbol{f}|\boldsymbol{f}^{(k)}). \end{aligned}$$

En TEP, ces variables cachées complètes non-observées sont $\boldsymbol{x} = \{x_{ij}, i = 1, \ldots, L; j = 1, \ldots, J\}$ où x_{ij} représentant le nombre de photons émis par le voxel j et détectés par la LOR i, et telle que

$$y_i = \sum_{j=1}^{J} x_{ij}.$$

Ces données complètes non-observées obéissent à une loi de Poisson de moyenne $\bar{x}_{ij} = a_{ij} f_j$. On peut alors construire leur log-vraisemblance $L(\boldsymbol{x}|\boldsymbol{f})$. L'étape E de l'algorithme consiste à estimer l'espérance conditionnelle des variables non-observées.

Si l'on ne prend pas en compte les évènements diffusés et fortuits dans l'équation 4.3, l'algorithme ML-EM en TEP est de la forme :

$$\text{Étape E} : \Phi(\boldsymbol{f}|\boldsymbol{f}^{(k)}) = \sum_{j=1}^{J} \left(\boldsymbol{f}_j^{(k)} \sum_{i=1}^{L} \frac{a_{ij} y_i}{\sum_{l=1}^{J} a_{il} \boldsymbol{f}_l^{(k)}} \log(a_{ij} f_j) - f_j \sum_{i=1}^{L} a_{ij} \right) \tag{4.4}$$

$$\text{Étape M} : \boldsymbol{f}_j^{(k+1)} = \frac{\boldsymbol{f}_j^{(k)}}{\sum_{i=1}^{L} a_{ij}} \sum_{i=1}^{L} \frac{a_{ij} y_i}{\sum_{l=1}^{J} a_{il} \boldsymbol{f}_l^{(k)}}. \tag{4.5}$$

L'algorithme est initialisé avec une image uniforme ($f_i^{(0)} = 1,\ i = 1, \ldots, J$) ou une image obtenue par FBP (auquel cas, les valeurs négatives doivent être mises à 0).

L'algorithme ML-EM peut être vu comme une succession d'étapes de projections et de rétroprojections où la rétroprojection du rapport entre les valeurs mesurées et calculées produit une image de correction qui, multipliée à l'estimée précédente, fournit l'estimée de l'image reconstruite à l'itération suivante. En effet, dans l'équation 4.5 donnant l'étape (M) on a :

- $\sum_l a_{il} \boldsymbol{f}_l^{(k)}$ représente les projections de l'image courante $\boldsymbol{f}^{(k)}$. Ce facteur peut être vu comme les données qui seraient mesurées si l'estimée courante $\boldsymbol{f}^{(k)}$ était la vraie image.
- $\sum_i (\frac{a_{ij} y_i}{\sum_l a_{il} \boldsymbol{f}_l^{(k)}})$ est une image de correction représentant la rétroprojection du rapport entre les données mesurées et estimées.
- $\sum_i a_{ij}$ est un facteur de normalisation.

Cette étape qui est à appliquer pixel par pixel, peut être étendue de sorte à s'appliquer à toute l'image. Le schéma itératif donnant l'estimée de l'image à l'itération $k+1$ en fonction de la précédente et des données mesurées et de la valeur courante se résume ainsi :

$$\text{Image}^{(k+1)} = \text{Image}^{(k)} \times \text{Rétroprojection normalisée de} \left(\frac{\text{Projections mesurées}}{\text{Projections de l'image}} \right).$$

L'algorithme ML-EM présente plusieurs propriétés contribuant à sa popularité en TEP.

- La fonction de coût augmente de façon monotone à chaque itération augmentant ainsi la vraisemblance, $\Phi(\boldsymbol{f}^{(k+1)}, \boldsymbol{y}) \geq \Phi(\boldsymbol{f}^{(k)}, \boldsymbol{y})$.
- Les itérées $\boldsymbol{f}^k$ convergent pour $k \to +\infty$ vers une image maximisant localement la vraisemblance.
- Si l'estimée initiale est positive, les estimées des itérations d'après le sont aussi.
- Par ailleurs, les images reconstruites par ML-EM présentent moins d'artéfacts que celles obtenues par FBP.

Cependant, cette méthode présente un certain nombre d'inconvénients :

- la convergence est relativement lente c'est-à-dire qu'il faut plusieurs dizaines d'itérations avant d'obtenir une image satisfaisante. De plus, le temps mis pour effectuer une itération de l'algorithme EM correspond à celui d'une reconstruction complète par FBP. Des versions accélérées de l'algorithme ML-EM ont été développées et implantées en routine clinique (algorithme OSEM par exemple [HL94], RAMLA [BDP96]).
- Le bruit augmente avec le nombre d'itérations. En effet, l'algorithme converge vers une image qui essaie de « coller » au mieux aux données bruitées d'un problème mal-conditionné. Cela induit une instabilité qui dégrade la qualité des images estimées lorsque le nombre d'itérations dépasse un certain seuil.
  En pratique les solutions adoptées consistent soit à stopper l'algorithme au bout d'un certain nombre d'itérations avant sa convergence, soit à post-filtrer l'image solution (typiquement par un filtre Gaussien 3D).
  Il peut alors sembler plus pertinent d'inclure la régularisation dans l'étape de reconstruction en pénalisant la vraisemblance ou en incluant une information dite *a priori* basée sur les techniques bayésiennes ([Gre90], [GM87]).

2) La méthode du maximum *a posteriori* (MAP)

Dans les méthodes bayésiennes, on cherche à limiter l'apparition du bruit au travers d'un terme définissant les propriétés *a priori* de l'image à reconstruire. Pour ce faire, l'image est considérée comme un vecteur aléatoire suivant une certaine loi *a priori*. Puis, on utilise la règle de Bayes pour former la loi *a posteriori* de $\boldsymbol{f}$ sachant les mesures :

$$p(\boldsymbol{f}|\boldsymbol{y}) = \frac{p(\boldsymbol{y}|\boldsymbol{f})p(\boldsymbol{f})}{p(\boldsymbol{y})}$$

où $p(\boldsymbol{y}|\boldsymbol{f})$ est la vraisemblance Poisson des données, $p(\boldsymbol{f})$ la loi *a priori* sur l'image, et $p(\boldsymbol{y}) = \int p(\boldsymbol{y}|\boldsymbol{f})p(\boldsymbol{f})\mathrm{d}\boldsymbol{f}$ est un facteur normalisation.

Comme leur nom l'indique, dans les algorithmes MAP l'estimée de $\boldsymbol{f}$ est choisie comme étant la valeur qui maximise la loi *a posteriori*. La fonction objective est donc le logarithme de la loi *a posteriori*. Comme le dénominateur est indépendant de $\boldsymbol{f}$, la fonction objective est donnée par :

$$\Phi(\boldsymbol{f}, \boldsymbol{y}) = L(\boldsymbol{y}|\boldsymbol{f}) + \log p(\boldsymbol{f})$$

où $L(\boldsymbol{y}|\boldsymbol{f}) = L(\boldsymbol{f})$ désigne le logarithme de la vraisemblance. La fonction à maximiser comporte donc

- un terme d'attache aux données (la vraisemblance),
- un terme de régularisation (la loi *a priori*).

La loi *a priori* a pour but de réduire la classe de fonctions où l'on cherche la solution. Typiquement, elle est choisie de façon à favoriser une certaine homogénéité dans l'image, en affectant par exemple une faible probabilité aux images présentant de grandes disparités entre voxels voisins et une probabilité plus grande aux images homogènes par morceaux tout en permettant des transitions rapides. L'un des modèles les plus utilisés repose sur la distribution de Gibbs. Elle s'exprime sous la forme :

$$p(\boldsymbol{f}) = \frac{1}{Z}\exp(-\beta U(\boldsymbol{f}))$$

où Z est une constante de normalisation, $\beta > 0$ est un paramètre contrôlant le compromis entre la force accordée aux données mesurées et le terme de régularisation. De grandes valeurs de β favorisent des images lisses, peu bruitées. Le choix de ce paramètre pose un problème même si des méthodes automatiques ou basées sur les données ont été proposées. Mais le problème demeure et est comparable à celui du choix de la taille de la fenêtre d'apodisation dans les méthodes FBP, ou du filtre dans les méthodes ML-EM [OF97]. La fonction U, appelée fonction d'énergie de Gibbs, contrôle le lissage de l'image. C'est une somme de fonctions potentielles qui s'exprime sous la forme :

$$U(\boldsymbol{f}) = \sum_j \sum_{k \in V_j, k>j} \psi\left(f_j - f_k\right)$$

où V_j désigne un voisinage défini autour du voxel j, par exemple les 26 voisins dans un cube $3 \times 3 \times 3$ centré en j. La fonction ψ doit favoriser une similarité de concentration entre voxels voisins tout en préservant les bords et les frontières dans l'image. Plusieurs fonctions ont été proposées (fonctions quadratiques, non-quadratiques, convexes, non-convexes), selon qu'elles pénalisent plus ou moins les discontinuités ([Hub81], [GM87], [Gre90], [QLC+98]). Ces fonctions sont en général caractérisées par le fait qu'elles sont croissantes avec les valeurs des différences des valeurs absolues $|f_j - f_k|$.

Au final, l'estimateur du maximum *a posteriori* est donné par

$$\hat{\boldsymbol{f}}_{\text{MAP}} = \arg\max_{\boldsymbol{f}\geq 0}(L(\boldsymbol{f}) - \beta U(\boldsymbol{f})).$$

Pour l'implantation de l'algorithme, on cherche à maximiser le critère Φ en annulant ses dérivées partielles

$$\partial\Phi(\boldsymbol{f}) = \frac{\partial L(\boldsymbol{f})}{\partial f_j} - \beta\frac{\partial U(\boldsymbol{f})}{\partial f_j}$$

pour $j = 1, \ldots, J$. Mais contrairement à l'algorithme ML-EM, l'introduction du terme de pénalité entraîne que l'on ne peut en général pas aboutir à une formulation explicite. La solution proposée par Green [Gre90] consiste à évaluer $\frac{\partial U(f)}{\partial f_j}$ à partir de l'estimée image calculée à l'itération précédente. La mise à jour s'effectue ainsi

$$\boldsymbol{f}_j^{(k+1)} = \frac{\boldsymbol{f}_j^{(k)}}{\sum_{i=1}^{L} a_{ij} + \beta \left.\frac{\partial U(f)}{\partial f_j}\right|_{\boldsymbol{f}=\boldsymbol{f}^{(k)}}} \sum_{i=1}^{L} \frac{a_{ij} y_i}{\sum_{l=1}^{J} a_{il}\boldsymbol{f}_l^{(k)}}.$$

C'est l'algorithme One-Step-Late (OSL). Cet algorithme sera utilisé dans nos études comparatives. Il est à souligner que d'autres techniques ont aussi été proposées. On peut citer entre autres les algorithmes GEM (Generalized EM, de Hebert et Leahy [HL89]), SAGE (Space-Alternating Generalized EM, de Fessler [FH95]) ou encore la méthode de Bouman et Sauer [BS96].

4.2.4 Résumé et conclusion sur les méthodes de reconstruction

Nous venons de résumer les principales méthodes de reconstruction en tomographie d'émission. Nous avons vu que ces méthodes sont généralement divisées en deux catégories : les méthodes analytiques et les méthodes statistiques.

Les méthodes analytiques sont basées sur une inversion directe de la transformée en rayons X. Cette modélisation est dite « continu-continu » en ce sens que le problème de reconstruction est formulé dans les espaces continus des données et de l'image (au moins de façon théorique). Leur implantation pratique nécessite une étape de discrétisation puisque les données sont discrètes. Les méthodes analytiques opèrent linéairement sur les données et permettent d'obtenir des reconstructions rapides. Par ailleurs, les algorithmes sont non biaisés, ce qui est un avantage pour certaines études. Les méthodes analytiques sont cependant basées sur des modèles simplifiés des processus physiques et ne prennent pas en compte la nature stochastique de l'émission et de la détection. Par conséquent, les images reconstruites présentent beaucoup d'artéfacts.

Pour pallier cette limitation, les méthodes statistiques ont été proposées. Comme nous l'avons vu, elles sont basées sur des techniques d'optimisation de la vraisemblance (ML) ou la vraisemblance pénalisée (PML) dite aussi maximum *a posteriori* (MAP). Les méthodes reposant sur l'approche du maximum de vraisemblance (ML) modélisent de façon plus réaliste les processus d'acquisition que les méthodes analytiques, par la prise en compte la nature aléatoire du phénomène. Elles fournissent en général des images présentant un meilleur rapport signal-sur-bruit que les méthodes analytiques. Néanmoins, ces images sont bruitées à cause du mauvais conditionnement du problème. Pour limiter l'apparition du bruit, les techniques de reconstruction dites bayésiennes

ou de vraisemblance pénalisée ont été proposées. Elles reposent sur l'introduction d'un terme de régularisation sous la forme d'une loi *a priori* sur l'image à reconstruire. Lorsqu'elles sont utilisées en TEP, ces méthodes sont basées sur l'approche du maximum *a posteriori* MAP. Les méthodes statistiques sont dites « discret-discret » car elles requièrent une formulation discrète de l'image et des données. Ainsi, la distribution spatiale de radioactivité est représentée sous forme vectorielle dans une base de fonctions prédéfinies et fixées. En principe, le choix des fonctions de base détermine le modèle l'image. Le plus souvent en 3D, ces fonctions de base sont des fonctions indicatrices de volumes élémentaires (les voxels). Un voxel est un élément volumique, représentant une valeur sur une grille régulière dans un espace 3D. La concentration du traceur est supposée constante à l'intérieur de chaque voxel. La discrétisation réduit le problème de reconstruction à un problème d'estimation paramétrique, le paramètre étant un vecteur de taille fini mais comportant un très grand nombre de coefficients inconnus qui sont à estimer à partir des observations. Se pose alors le problème de la détermination du nombre et des dimensions adéquats des voxels pour représenter la fonction. En effet, la taille du voxel doit être déterminée en amont sans une connaissance fine de la structure de l'objet à imager. La taille du voxel agit ainsi comme un paramètre de régularisation puisque plus elle est grande, plus l'image est lissée et moins bonne est la résolution spatiale. La discrétisation est donc source d'un certain nombre de biais. Par ailleurs, s'il paraît naturel de formuler le problème dans l'espace discret des détecteurs puisque le nombre de mesures collectées est fini, en revanche la discrétisation de l'image représentant la distribution du radiotraceur n'a rien de naturel puisque la distribution d'émission est une quantité continue.

Dans ce contexte, nous proposons une troisième catégorie de reconstruction reposant sur un modèle « discret-continu ». Plus précisément, le but est de reconstruire la fonction représentant la distribution de radioactivité à partir du nombre fini de mesures (à savoir les projections discrètes des émissions aléatoires détectées). L'abord du problème de reconstruction dans le cadre proposé permet d'éviter n'importe quelle forme de grille (grille de fonctions ou de voxels). La régularisation du problème inverse se fait dans un cadre complètement bayésien. En effet, l'approche bayésienne en TEP est souvent réduite à l'estimée ponctuelle fournie par l'approche du maximum *a posteriori*. Notre approche est *complètement bayésienne* en ce sens qu'elle permet d'estimer la distribution entière d'activité *a posteriori*.

4.3 Nouvelle méthode BNP de reconstruction 3D

Les méthodes statistiques de reconstruction d'images en tomographie d'émission permettent la prise en compte du bruit inhérent aux données acquises. Afin de traiter un nombre fini de variables, ces approches procèdent à une paramétrisation finie du problème. L'approche paramétrique nécessite donc de formuler les données et l'image à reconstruire dans des espaces discrets. Bien que les mesures soient de nature discrète, il n'en est pas de même pour l'image puisqu'elle traduit une distribution d'émission. Il semble alors plus pertinent d'aborder le problème de reconstruction dans un cadre continu. Fessler [Fes04] soulignait qu'il n'était pas illusoire d'envisager la reconstruction d'un objet continu à partir d'un ensemble fini de mesures, comme c'est le cas en régression non paramétrique, mais qu'il s'agissait d'une tâche compliquée pour la tomo-

graphie. Nous nous donnons pour objectif d'effectuer cette tâche.

Nous proposons d'aborder la reconstruction dans un cadre statistique fondamentalement différent de ceux vus jusqu'à présent. Aussi, nous estimons que le cadre méthodologique dans lequel est abordé la reconstruction doit
- être unifié,
- être mathématiquement consistent,
- permettre de reconstruire directement l'image (vue comme une fonction continue) à partir des données discrètes,
- permettre la quantification de l'incertitude associée à la reconstruction.

Nous avons présenté dans le chapitre 2 des outils de modélisation statistique non paramétrique, en particulier les mélanges par processus de Dirichlet (DPM) et par processus de Pitman-Yor (PYM). Ces mélanges permettent l'estimation d'une densité à partir de mesures discrètes issues d'elles. Ainsi, l'abord du problème de reconstruction dans un cadre d'estimation de densité permettrait d'envisager l'utilisation de tels outils. Ce qui d'une part, pourvoirait la modélisation d'une grande flexibilité. D'autre part, l'approche bayésienne non paramétrique en mettant une loi *a priori* directement sur la mesure en question, permet de caractériser toute sa distribution permettant ainsi une estimation directe de l'erreur. Par ailleurs, les processus de Dirichlet et de Pitman-Yor sont caractérisés par une propriété de clustering sous-jacente. Cette propriété est intéressante et serait particulièrement utile pour l'estimation de volumes fonctionnels pour la reconstruction 4D qui sera abordée dans le chapitre 5.

Cependant, la difficulté en tomographie est que contrairement à l'estimation de densité non paramétrique, les mesures observées ne sont pas directement issues de la densité à estimer. On doit alors pouvoir introduire l'information manquante à savoir les lieux d'émission sachant les observations. Sitek [Sit11] adopte cette idée en utilisant la méthode des ensembles d'origines. Pour cela, il définit l'image comme un ensemble d'origines des mesures détectées et estime les localisations des origines des évènements. Ainsi, sachant les mesures, les évènements sont replacés stochastiquement dans des ensembles. Sitek utilise cependant les voxels pour replacer les évènements, en calculant le nombre moyen d'émissions dans chaque voxel, afin de décider de celui contenant l'émission en question. L'abord du problème de reconstruction sous la forme d'une estimation indirecte de densité permet de s'affranchir de cela.

Nous abordons la reconstruction comme un problème d'estimation non paramétrique d'une fonction de densité de probabilité à partir d'observations issues indirectement d'elle. L'objectif étant de reconstruire la densité d'émission spatiale directement à partir des données, sans aucune discrétisation. Il s'agit donc d'une rupture méthodologique comparée à toutes les approches statistiques de reconstruction en TEP. Pour la régularisation du problème inverse mal posé, nous suivons une approche bayésienne, en construisant une loi *a priori* pour la distribution d'intérêt dans $\mathcal{M}(\mathbb{R}^3)$, l'espace de toutes les mesures de probabilité dans $\mathbb{R}^3$. Cette approche est alors dite *bayésienne non paramétrique*. L'avantage de l'approche bayésienne non paramétrique est donc de fournir un contexte de régularisation bayésienne pour un problème d'estimation non paramétrique. Par ailleurs, on a accès à la distribution entière de la distribution d'activité *a posteriori* et non pas une estimée ponctuelle comme dans les approches ML et MAP. En conséquence, on peut estimer n'importe quelle fonctionnelle permettant de quantifier le degré de confiance associée à la reconstruction (variance, intervalles de crédibilité etc.)

Ce qui est bien approprié pour l'imagerie quantitative, dans les reconstructions avec un faible nombre de données, appelées « faibles doses » en TEP.

Préambule : On distinguera deux espaces : l'espace objet et l'espace d'observation. Le premier désigne la région du cerveau où l'on veut effectuer la reconstruction et le second celui des détecteurs.

Notations : Dans la suite, nous adoptons les notations suivantes : $\boldsymbol{x} \in \mathbb{R}^3$ désigne les coordonnées d'un point de l'espace et on le notera $\boldsymbol{x} = (x_1, x_2, x_3)$, F est la distribution de probabilité des observations et G la distribution spatiale d'émission. La densité de G sera notée f_G. La distribution de projection est notée $\mathcal{P}(\cdot|\boldsymbol{x})$. Enfin l'espace objet est noté $\mathcal{X}$ et celui des observations $\mathcal{Y}$.

4.3.1 Formulation bayésienne du problème non paramétrique

L'information physique d'intérêt est l'intensité d'émission spatiale $\boldsymbol{f}$ du processus de Poisson. Nous considérons qu'une version normalisée de cette intensité, à savoir la densité d'émission spatiale, est une quantité d'intérêt suffisante en tomographie d'émission.

Comme nous l'avons déjà souligné, en TEP chaque évènement en coïncidence est affectée à une ligne virtuelle (LOR) joignant les centres des deux détecteurs. Nous avons vu qu'une LOR l est caractérisée par les paramètres définissant sa position dans l'espace des détecteurs. Nous construisons une correspondance entre les paramètres des LOR et l'indice l de telle sorte que la variable y_i corresponde à l'indice de la LOR ayant enregistré la $i^{\text{ème}}$ coïncidence observée. Plus précisément, l'observation y_i désigne les coordonnées de la LOR qui joint les deux détecteurs ayant enregistré les photons en coïncidence issus de l'annihilation des positons ayant eu lieu en $\boldsymbol{x}_i$. Le problème de reconstruction se formule alors ainsi dans le cadre bayésien non paramétrique pour les problèmes inverses :

$$\begin{aligned} G &\sim \mathcal{G} \\ F(\cdot) &= \int_{\mathcal{X}} \mathcal{P}(\cdot|\boldsymbol{x})\, G(\mathrm{d}\boldsymbol{x}) \\ y_i|F &\overset{\text{iid}}{\sim} F, \text{ pour } i = 1, \ldots, n \end{aligned} \tag{4.6}$$

- $\mathcal{X}$ désigne l'espace-objet,
- $\boldsymbol{y} = \{y_1, \ldots, y_n\}$ est l'ensemble des observations distribuées suivant F,
- G est une mesure de probabilité aléatoire distribuée suivant $\mathcal{G}$ et définie sur $(\mathcal{X}, \sigma(\mathcal{X}))$,
- $\mathcal{P}(\cdot|\boldsymbol{x})$ est une loi de probabilité connue, définie sur $(\mathcal{Y}, \sigma(\mathcal{Y}))$.

Dans le contexte de la TEP,

- $\mathcal{X}$ est une région bornée de $\mathbb{R}^3$. C'est l'espace dans lequel la reconstruction doit être effectuée,
- $\boldsymbol{y} = \{y_1, \ldots, y_n\}$ désigne l'ensemble des indices des LORs,
- G est la distribution spatiale dont la densité f_G (densité d'émission spatiale) est à estimer,

– $\mathcal{P}(\cdot|\boldsymbol{x})$ est la distribution de projection. C'est la distribution donnant la probabilité de détecter une paire de photons dans une LOR l sachant une émission ayant eu lieu en $\boldsymbol{x}$. Nous reviendrons plus loin sur sa modélisation mais signalons d'ores et déjà que dans les approches statistiques basées sur la discrétisation de l'espace (ML et MAP), cette loi de probabilité est discrétisée et représentée sous la forme de la matrice appelée matrice-système.

Dans la formulation 4.6 du problème inverse, $G(\cdot)$ représente la distribution spatiale correspondant aux *évènements enregistrés*. Nous avons déjà souligné dans la section 4.2.1 qu'en 3D on est en présence de données tronquées et toutes les LOR ne sont pas observables. La sensibilité du système est alors dépendante de la position du point d'annihilation dans le champ de vue. Si l'on suppose que l'on a des détecteurs parfaits et qu'il n'y a pas de milieu atténuant alors la probabilité de détecter la paire de photons résultante dans un couple de détecteurs est purement géométrique. Elle est donnée en 3D par l'angle solide formé par les deux détecteurs vus depuis le point d'annihilation. Lorsque la distance radiale augmente, la distance entre les deux détecteurs va diminuer (et par là, la surface délimitée par les détecteurs) entrainant une diminution de l'angle solide. De ce fait, l'enregistrement d'un évènement en coïncidence est plus probable lorsque l'annihilation a eu lieu près du centre du système. Afin de prendre en compte ces effets, nous introduisons une LOR virtuelle indicée par 0 représentant les évènements non observés et la variable aléatoire $y^* = (y > 0)$ pour désigner une LOR réelle (ayant enregistré une coïncidence). Partant de ces remarques, on note que dans l'équation 4.6, le conditionnement sur y^* est implicite, c'est-à-dire que $G(\mathrm{d}\boldsymbol{x}) = G^*(\mathrm{d}\boldsymbol{x}|y^*)$ où $G^*(\mathrm{d}\boldsymbol{x}|y^*)$ est la distribution de $\boldsymbol{x}$ conditionnée au fait que la coïncidence a été enregistrée. On peut l'exprimer par règle de Bayes de la façon suivante

$$G^*\left(\mathrm{d}\boldsymbol{x}|y^*\right) = \frac{\mathbb{P}\left(y^*|\boldsymbol{x}\right) G^\dagger\left(\mathrm{d}\boldsymbol{x}\right)}{\mathbb{P}(y^*)}. \tag{4.7}$$

Ce qui équivaut à

$$G^\dagger\left(\mathrm{d}\boldsymbol{x}\right) = \frac{G^*\left(\mathrm{d}\boldsymbol{x}|y^*\right)\mathbb{P}(y^*)}{\mathbb{P}\left(y^*|\boldsymbol{x}\right)} \tag{4.8}$$

où $G^\dagger(\cdot)$ est la distribution spatiale de tous les évènements (enregistrés et non enregistrés) et pour tout $\boldsymbol{x} \in \mathcal{X}$, $\mathbb{P}\left(y^*|\boldsymbol{x}\right) = \sum_{l=1}^{L} \mathbb{P}(y = l|\boldsymbol{x})$ est la probabilité qu'une émission localisée en $\boldsymbol{x}$ soit enregistrée par le système, en prenant en compte la géométrie du système et ses propriétés physiques. Son expression est explicitée dans le paragraphe qui suit. Puisque $G^\dagger(\mathrm{d}\boldsymbol{x})$ est une version normalisée de $\frac{G^*(\mathrm{d}\boldsymbol{x}|y^*)}{\mathbb{P}(y^*|\boldsymbol{x})}$ (cf. équation 4.8), l'inférence sur la loi *a posteriori* $G^*(\mathrm{d}\boldsymbol{x}|y^*)|\boldsymbol{Y}$ suffit pour estimer la loi *a posteriori* $G^\dagger(\mathrm{d}\boldsymbol{x})|\boldsymbol{Y}$ dès lors que $\mathbb{P}\left(y^*|\boldsymbol{x}\right) > 0$ pour tout $x \in \mathcal{X}$ (c'est-à-dire que tout lieu d'émission dans l'espace-objet est potentiellement observable).

Pour des raisons de simplicité dans les notations, dans le reste de cette section, nous omettons le conditionnement sur y^* puisque dans l'inférence nous ne traitons que les évènements enregistrés. On utilisera donc $G(\mathrm{d}\boldsymbol{x})$ pour représenter $G^*(\mathrm{d}\boldsymbol{x}|y^*)$, la distribution correspondant à un évènement enregistré.

Modélisation de la distribution de projection

On modélise le projecteur par une loi de probabilité $\mathcal{P}(y|\boldsymbol{x})$ que l'on appellera distribution de projection. Cette loi de probabilité est entièrement déterminée par la

donnée des probabilités $\mathbb{P}(y = l|\boldsymbol{x})$ pour tout $\boldsymbol{x}$, de coordonnées spatiales (x_1, x_2, x_3), et appartenant au champ de vue du tomographe. Soit donc $\mathbb{P}(y = l|\boldsymbol{x})$ la probabilité conditionnelle qu'une paire de photons soit détectée dans le TOR l sachant que l'annihilation a eu lieu en $\boldsymbol{x}$. La probabilité d'enregistrer une émission par le système est $\mathbb{P}(y^*|\boldsymbol{x}) = \sum_{l=1}^{L} \mathbb{P}(y = l|\boldsymbol{x}) \leq 1$ et la probabilité que l'émission ne soit pas enregistrée $\mathbb{P}(y = 0|\boldsymbol{x}) = 1 - \mathbb{P}(y^*|\boldsymbol{x})$. On rappelle que la notation $y = 0$ représente un évènement non observé.

Dans les travaux liés à cette thèse, le projecteur est déterminé par la probabilité géométrique qu'une paire de photons soit enregistrée dans le TOR l sachant que l'annihilation est localisée en $\boldsymbol{x}$. Cette probabilité est obtenue en calculant l'angle solide délimité par le TOR, vu à partir du point $\boldsymbol{x}$. Si le TOR l met en jeu le couple de capteurs i et j, cette probabilité s'obtient par intégration dans les domaines angulaires des coordonnées sphériques :

$$\mathbb{P}(y = l|\boldsymbol{x}) = \frac{1}{4\pi} \int_0^{2\pi} \int_{-\pi}^{\pi} \mathbf{1}\left(D(\boldsymbol{x}, \phi, \theta) \in \mathcal{C}_i\right) \mathbf{1}\left(D(\boldsymbol{x}, \phi, \theta) \in \mathcal{C}_j\right) \sin\theta \, \mathrm{d}\phi \, d\theta,$$

où $D(\boldsymbol{x}, \phi, \theta) \in \mathcal{C}_i$ signifie que la droite passant par $\boldsymbol{x}$, d'azimuth ϕ et de zénith θ intercepte le capteur i.
Dans le cas général, cette expression n'a pas de forme analytique. On y remédie en assimilant chacun des capteurs à un petit rectangle vertical. Ce modèle est légitime pour tout instrument réel présentant une forme cylindrique, sans nécessairement une symétrie de révolution. Chaque capteur i doit pouvoir simplement être décrit par quatre points dans l'espace. Soient A_i, B_i, C_i et D_i ces quatre points, de coordonnées :

- $A_i(xi_1, yi_1, zi_1)^T$,
- $B_i(xi_2, yi_2, zi_1)^T$,
- $C_i(xi_1, yi_1, zi_2)^T$,
- $D_i(xi_2, yi_2, zi_2)^T$.

Les capteurs sont suffisamment petits pour que l'on puisse approximer, avec très peu d'erreurs, cet angle solide par le produit de deux angles de vue correspondant à la projection sur deux plans orthogonaux. L'un, $\mathcal{H}_1$, est un plan normal à l'axe x_3 ; le second, $\mathcal{H}_2$, est le plan passant par les axes verticaux de chacun des capteurs. Autrement dit, si on note respectivement par V_i et V_j les axes verticaux des capteurs i et j, alors V_i est la droite passant par les points M_i (milieu du segment $[A_i, B_i]$) et N_i (milieu de $[C_i, D_i]$) (voir Figure 4.8).

On définit pour tout point p du plan et pour tous segments s_1 et s_2 de ce même plan, la fonction suivante :

$$PV(p, s_1, s_2) = \frac{1}{\pi} \int_0^{\pi} \mathbf{1}(D(p, \theta) \in s_1) \, \mathbf{1}(D(p, \theta) \in s_2) \, \mathrm{d}\theta$$

où $D(p, \theta) \in s$ signifie que la droite passant par p et d'angle θ coupe le segment s.
Si on désigne par $|_{\mathcal{H}_1}$ la projection sur $\mathcal{H}_1$, on a alors

$$\mathbb{P}(y = l|\boldsymbol{x}) \approx PV(\boldsymbol{x}|_{\mathcal{H}_1}, \mathcal{C}_i|_{\mathcal{H}_1}, \mathcal{C}_j|_{\mathcal{H}_1}) \, PV(\boldsymbol{x}|_{\mathcal{H}_2}, \mathcal{C}_i|_{\mathcal{H}_2}, \mathcal{C}_j|_{\mathcal{H}_2})$$

où :

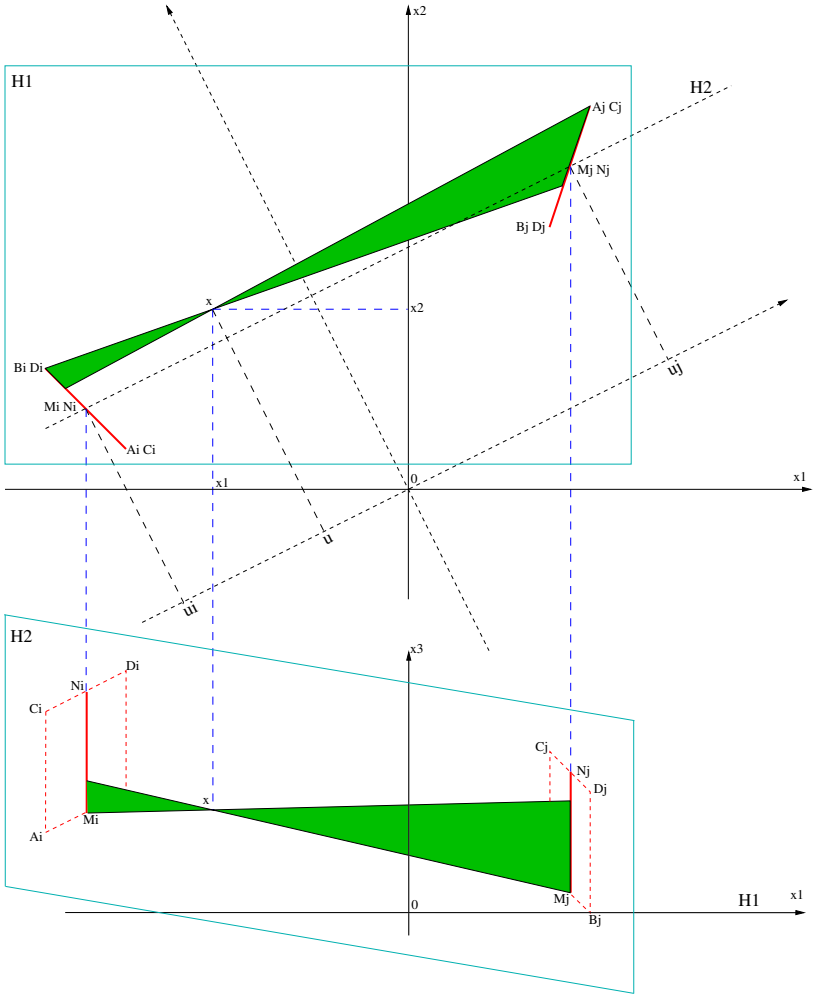


FIGURE 4.8 – Schéma de deux capteurs i et j et des deux angles de vue à partir du point $\boldsymbol{x}$.

$$\boldsymbol{x}|_{\mathcal{H}_1} = (x_1,\, x_2)^T$$
$$\mathcal{C}_i|_{\mathcal{H}_1} = [(xi_1, yi_1)^T,\, (xi_2, yi_2)^T]$$
$$\mathcal{C}_j|_{\mathcal{H}_1} = [(xj_1, yj_1)^T,\, (xj_2, yj_2)^T]$$

et

$$\boldsymbol{x}|_{\mathcal{H}_2} = (u,\, x_3)^T$$
$$\mathcal{C}_i|_{\mathcal{H}_2} = [(u_i, zi_1)^T,\, (u_i, zi_2)^T]$$
$$\mathcal{C}_j|_{\mathcal{H}_2} = [(u_j, zj_1)^T,\, (u_j, zj_2)^T]$$

avec u, u_i et u_j désignant les abscisses des points $\boldsymbol{x}$, M_i et M_j dans un repère du plan $\mathcal{H}_2$.

La fonction PV se calcule ensuite de la manière suivante. On repère parmi les quatre points constituant les extrémités des segments s_1 et s_2, les deux points $p_1(p_{1x}, p_{1y})$ et $p_2(p_{2x}, p_{2y})$ qui encadrent l'angle de vue (cf. Figure 4.9). L'expression de la fonction est

alors

$$PV\,(p(p_x,p_y),s_1,s_2) = \begin{cases} \frac{1}{\pi}|\arctan(\frac{dy2dx1-dy1dx2}{dx1dx2+dy1dy2})| & \text{si } dx1dx2+dy1dy2 \neq 0 \\ \frac{1}{2} & \text{sinon.} \end{cases}$$

avec

$dx1 = p_{1x} - p_x$,

$dy1 = p_{1y} - p_y$,

$dx2 = p_{2x} - p_x$

et $dy2 = p_{2y} - p_y$.

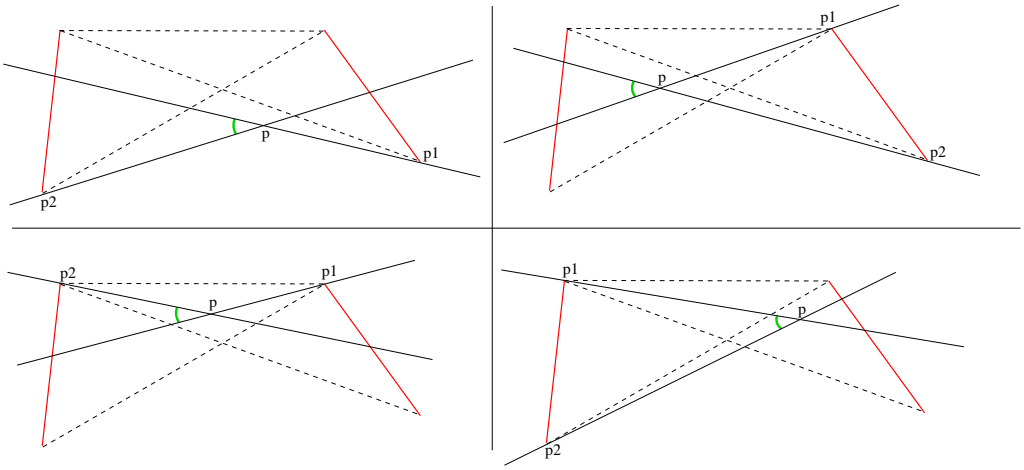


FIGURE 4.9 – Illustration de l'angle de vue en fonction de la position du point p dans le tube de réponse. L'angle de vue en un point $\boldsymbol{x}$ est donné par l'angle maximal que peuvent former les rayons passant par ce point et qui sont tels que la détection en coïncidence puisse avoir lieu dans le tube.

Après une caractérisation probabiliste du système physique et de ses propriétés géométriques, nous allons maintenant nous intéresser à la modélisation des données.

Loi *a priori* sur la distribution d'activité

Dans la régularisation bayésienne, on munit les inconnues du modèle de lois *a priori*. Dans le contexte bayésien non paramétrique, cette loi *a priori* porte directement sur la loi G et s'exprime par $G \sim \mathcal{G}$, où $\mathcal{G}$ est une distribution sur des distributions, *i.e*, chaque tirage suivant $\mathcal{G}$ est une mesure de probabilité sur $\mathcal{X}$. On dit alors que G est une mesure de probabilité aléatoire. La loi *a priori* que nous avons utilisée pour la densité f_G de G est un mélange par processus de Pitman-Yor (PYM) (cf. chapitre 2, paragraphe 2.3.3). Nous rappelons que l'idée des PYM est de convoluer une mesure discrète générée par un processus de Pitman-Yor (PYP) avec une fonction continue paramétrique afin d'obtenir un *a priori* sur une densité. Plus précisément, soit H une mesure de probabilité générée

suivant un processus de Pitman-Yor,

$$H(\cdot) = \sum_{k=1}^{\infty} w_k \delta_{\boldsymbol{\theta}_k^*}(\cdot)$$

Pour obtenir un PYM, la fonction H est convoluée avec un noyau continu $\phi(\cdot|\boldsymbol{\theta})$. Dans notre cas, il s'agit d'une gaussienne 3D de paramètres $\boldsymbol{\theta} = (\mathbf{m}, \boldsymbol{\Sigma})$ où $\mathbf{m}$ représente le vecteur moyen et $\boldsymbol{\Sigma}$ la matrice de covariance. Ceci conduit à la loi *a priori* suivante sur la densité de G,

$$f_G(\boldsymbol{x}) = \int \phi(\boldsymbol{x}|\boldsymbol{\theta}) H(\boldsymbol{\theta}) \mathrm{d}\boldsymbol{\theta} = \sum_{k=1}^{\infty} w_k f_{\mathcal{N}}(\boldsymbol{x}|\boldsymbol{\theta}_k^*).$$

4.3.2 Modèles hiérarchiques génératifs pour les données TEP

Nous décrivons ici le processus génératif des données. On suppose que l'on veut générer n observations c'est-à-dire que n paires de photons en coïncidence soient enregistrées. Soient $\boldsymbol{X} = (\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n)$, n émissions localisées dans l'espace objet. Si tous les $\boldsymbol{x}_i$ étaient observés, le problème de reconstruction serait celle de l'estimation d'une densité à partir de ses observations et pourrait être résolu par les méthodes d'estimation de densité non paramétrique que l'on a vu dans le chapitre 2. Néanmoins, en TEP le problème est plus compliqué puisqu'ici, on n'observe pas directement ces émissions mais plutôt leurs projections $\boldsymbol{Y} = (\boldsymbol{y}_1, \ldots, \boldsymbol{y}_n)$ dans l'espace des détecteurs. Notre problème de reconstruction est alors celui d'une estimation indirecte de densité. On parle d'estimation indirecte de densité lorsque les observations ne sont pas distribuées suivant la loi d'intérêt G mais plutôt suivant une loi F reliée à G par un opérateur $\mathcal{H}$ sous la forme $G = \mathcal{H}F$. Dans l'équation 4.6, l'opérateur $\mathcal{H}$ est l'opérateur intégrale. Afin de compléter le modèle, nous introduisons des variables cachées comme dans ML-EM. Ici, les variables cachées correspondent aux lieux d'annihilation $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$. L'augmentation du modèle permet d'expliciter la distribution jointe de toutes les variables,

$$\begin{aligned} f(\boldsymbol{y}, \boldsymbol{X}, \boldsymbol{c}, \boldsymbol{\Theta}, \boldsymbol{w}) =& P(\boldsymbol{y}|\boldsymbol{X}) \times f(\boldsymbol{X}|\boldsymbol{c}, \boldsymbol{\Theta}) \\ & \times P(\boldsymbol{c}|\boldsymbol{w}) \times f(\boldsymbol{w}) \times f(\boldsymbol{\Theta}). \end{aligned} \tag{4.9}$$

Les distributions sont données dans le modèle hiérarchique qui suit :

pour $i = 1, \ldots, n$,

$$\begin{aligned} y_i|\boldsymbol{x}_i &\overset{\text{ind}}{\sim} \mathcal{P}(y_i|\boldsymbol{x}_i) \\ \boldsymbol{x}_i|\boldsymbol{\theta}_i &\overset{\text{ind}}{\sim} \mathcal{N}(\boldsymbol{x}_i|\boldsymbol{\theta}_i) \\ \boldsymbol{\theta}_i|H &\overset{\text{iid}}{\sim} H \\ H &\sim \mathrm{PY}(d, \alpha, G_0) \end{aligned} \tag{4.10}$$

où $\mathcal{P}$ est la distribution de la projection, $\boldsymbol{\theta}_i$ désigne les paramètres de la composante à laquelle $\boldsymbol{x}_i$ est affectée, H est un processus de Pitman-Yor de distribution

$$H(\cdot) = \sum_{k=1}^{\infty} w_k \, \delta_{\boldsymbol{\theta}_k^*}(\cdot),$$

où

- $\mathbf{w} = w_1, w_2, \ldots \sim \text{GEM}(d, \alpha)$ sont les poids des composantes,
- $\boldsymbol{\theta}^* = \boldsymbol{\theta}_1^*, \boldsymbol{\theta}_2^*, \ldots$ sont les valeurs distinctes de $\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \ldots$, et sont i.i.d. G_0. Puisque G_0 définit une distribution sur l'espace des paramètres des clusters, à savoir les $\boldsymbol{\theta}_k^* = (\mathbf{m}_k, \boldsymbol{\Sigma}_k)$, nous l'avons choisie suivant une loi Normal-Inverse Wishart ($\mathcal{NIW}_{\rho,n_0,\mu_0,\boldsymbol{\Sigma}_0}$), définie de la façon suivante :

$$\mathbf{m}_k|\boldsymbol{\Sigma}_k \sim \mathcal{N}\left(\boldsymbol{\mu}_0,\, \boldsymbol{\Sigma}_k/\rho\right),$$

où ρ est le paramètre de précision, $\boldsymbol{\mu}_0$ la moyenne de la loi normale et

$$\boldsymbol{\Sigma}_k^{-1} \sim \mathcal{W}(n_0, (n_0\,\boldsymbol{\Sigma}_0)^{-1}),$$

avec $\mathcal{W}$ désignant la distribution de Wishart, n_0 le degré de liberté et $\boldsymbol{\Sigma}_0^{-1}$ la moyenne.

Ceci conduit à un modèle avec une infinité potentielle de composantes. Puisque H est discrète, plusieurs $\boldsymbol{\theta}_i$ seront identiques ; cela induit un regroupement *a priori* des données $\boldsymbol{x}_i$ suivant les valeurs des paramètres des composantes, c'est l'effet de *renforcement statistique* de ce processus. Le modèle 4.10 peut donc être ré-écrit de façon équivalente en introduisant des variables de classification c_i permettant d'identifier la composante Gaussienne à partir de laquelle l'émission $\boldsymbol{x}_i$ sera tirée, c'est-à-dire $c_i = k$ ssi $\boldsymbol{\theta}_i = \boldsymbol{\theta}_k^*$. Cela s'exprime ainsi :

$$\begin{aligned}
y_i|\boldsymbol{x}_i &\overset{\text{ind}}{\sim} \mathcal{P}\left(y_i|\boldsymbol{x}_i\right) \\
\boldsymbol{x}_i|c_i, \boldsymbol{\Theta}^* &\overset{\text{ind}}{\sim} \mathcal{N}\left(\boldsymbol{x}_i|\boldsymbol{\theta}_{c_i}^*\right) \\
c_i|\boldsymbol{w} &\overset{\text{iid}}{\sim} \sum_{k=1}^{\infty} w_k \delta_k(\cdot) \\
\boldsymbol{w} &\sim \text{GEM}(d, \alpha) \\
\boldsymbol{\theta}_k^* &\overset{\text{iid}}{\sim} \mathcal{NIW}_{\rho,n_0,\bar{m}\mu_0,\boldsymbol{\Sigma}_0}
\end{aligned} \tag{4.11}$$

4.3.3 Inférence

Nous avons vu dans le chapitre 3 les méthodes MCMC pour l'inférence *a posteriori* dans les modèles de mélange par processus de Dirichlet (DPM). Nous y avons aussi présenté la nouvelle méthode d'échantillonnage que nous avons proposée pour les modèles de mélange par processus de Pitman-Yor, généralisation à deux paramètres des mélanges par processus de Dirichlet (cf. chapitre 3, section 3.2.3). La méthode d'échantillonnage proposée utilise la stratégie dite de « slice sampling ». Cette stratégie est basée sur l'introduction de variables auxiliaires uniformes qui rendent le modèle conditionnellement fini et permet ainsi d'éviter une troncature sèche du modèle. Nous avons aussi proposé l'introduction d'un seuil ainsi que la formulation de l'échantillonnage dans l'espace des classes d'équivalence des clusters. Cela permet d'améliorer les propriétés de mélange de l'algorithme comparé aux autres méthodes conditionnelles (cf. chapitre 3, section 3.2.4). C'est cet algorithme qui a été utilisé pour la reconstruction en TEP. Nous ne détaillerons donc pas ici toutes les étapes mais présentons les principales.

On commence par ré-écrire le modèle hiérarchique 4.11 en introduisant les variables slices.

Modèle hiérarchique génératif avec les variables slices

Soient $\mathbf{u} = (u_1, u_2, \ldots, u_n)$ des variables auxiliaires suivant une distribution uniforme telles que la densité jointe pour tout $(\boldsymbol{x}_i, u_i)$ sachant $\boldsymbol{w}$ et $\boldsymbol{\Theta}^*$ soit

$$f_G(\boldsymbol{x}_i, u_i) = \sum_{k=1}^{\infty} w_k \, \mathcal{N}(\boldsymbol{x}_i | \boldsymbol{\theta}_k^*) \; \mathcal{U}(u_i | 0, \xi_k) \tag{4.12}$$

où $\mathcal{U}(\cdot|a, b)$ désigne une distribution uniforme sur $]a, b]$ et ξ_k est une variable dépendante telle que k,

$$\xi_k = \min(w_k, \zeta) \tag{4.13}$$

avec $\zeta \in]0, 1]$, indépendante de w_k. Ici, ζ est un seuil que l'on introduit afin d'améliorer les propriétés de mélange (cf. chapitre 3, section 3.2.3). Utilisant l'équation 4.13, on peut ré-écrire l'équation 4.12 de la façon suivante

$$f_G(\boldsymbol{x}_i, u_i) = \mathbf{1}(\zeta > u_i) \, \zeta^{-1} \sum_{w_k > \zeta} w_k \, \mathcal{N}(\boldsymbol{x}_i | \boldsymbol{\theta}_k^*) + \sum_{w_k \leq \zeta} \mathbf{1}(w_k > u_i) \; \mathcal{N}(\boldsymbol{x}_i | \theta_k^*)$$

où toutes les deux sommes sont finies puisque $\#\{j : w_j > \varepsilon\} < \infty$, pour tout $\varepsilon > 0$.

On peut donc ré-écrire le modèle hiérarchique 4.11 en considérant ces variables uniformes,

$$\begin{aligned}
y_i | \boldsymbol{x}_i &\overset{\text{ind}}{\sim} \mathcal{P}(y_i | \boldsymbol{x}_i) \\
\boldsymbol{x}_i | c_i, \boldsymbol{\Theta}^* &\overset{\text{ind}}{\sim} \mathcal{N}\left(\boldsymbol{x}_i | \boldsymbol{\theta}_{c_i}^*\right) \\
c_i, u_i | \boldsymbol{w} &\overset{\text{iid}}{\sim} \sum_{k=1}^{\infty} \mathcal{U}(u_i | 0, \min(w_k, \zeta)) \, w_k \delta_k(c_i) \\
\boldsymbol{w} &\sim \text{GEM}(d, \alpha) \\
\boldsymbol{\theta}_k^* &\overset{\text{iid}}{\sim} \mathcal{NIW}_{\rho, n_0, \boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0}.
\end{aligned} \tag{4.14}$$

Algorithme MCMC pour la reconstruction en TEP

L'inférence sur la distribution *a posteriori* de l'activité $G(\cdot)|\boldsymbol{y}$ consiste en l'inversion du processus génératif des données dans le modèle hiérarchique 4.14. Pour cela, on a besoin d'effectuer des tirages suivant les lois *a posteriori* de chacune des variables du modèle. Le schéma de simulation peut se résumer ainsi : à l'itération 0, les poids et les paramètres des composantes sont initialisés suivant leurs lois *a priori* respectives. Lors des itérations suivantes, l'algorithme va mettre à jour chacun des paramètres du modèle suivant sa distribution conditionnelle *a posteriori*. Pour cela, l'échantillonneur va générer successivement des blocs de variables suivant les conditionnelles :

1. Proposition de localisation des annihilations : $\boldsymbol{x}|\boldsymbol{w}, \boldsymbol{\Theta}^*, \boldsymbol{y}$.
2. Affectation des annihilations aux composantes du mélange : $\mathbf{c}|\boldsymbol{w}, \boldsymbol{\Theta}^*, \boldsymbol{x}$.
3. Mise à jour des paramètres des composantes (moyennes et matrices de covariance) : $\boldsymbol{\Theta}^*|\mathbf{c}, \boldsymbol{x}$.
4. Mise à jour des poids et des variables slices : $\boldsymbol{w}, \boldsymbol{u}|\mathbf{c}$.

Grâce aux propriétés de conjugaison des lois *a priori* choisies, toutes ces lois conditionnelles sont simulables avec un échantillonneur de Gibbs sauf celle de $(\boldsymbol{x}|\boldsymbol{w}, \boldsymbol{\Theta}^*, \boldsymbol{y})$ pour laquelle nous avons eu recours à un algorithme Metropolis-Hastings (MH). En effet, la distribution de projection $\mathcal{P}$ prend en compte les propriétés géométriques du système de détection et cette distribution est impliquée dans la loi *a posteriori* des c_i. Cela exclut toute tentative d'échantillonner directement suivant la conditionnelle $(\boldsymbol{x}|\boldsymbol{w}, \boldsymbol{\Theta}^*, \boldsymbol{y})$.

Les différentes lois conditionnelles impliquées dans l'algorithme MCMC vont maintenant être explicitées. Les distributions conditionnelles pour les poids et pour les variables de classification ont déjà été détaillées dans la première variante du nouvel algorithme que nous avons proposé pour l'inférence des DPM (chapitre 3, section 3.2.3). Nous donnons ici juste leurs expressions.

Distribution conditionnelle de $(\boldsymbol{w}, \boldsymbol{u}|\boldsymbol{c})$

Soit K_n le nombre de clusters distincts après n observations, n_k le nombre de données affectées au cluster k et K^* le nombre de clusters représentés par le slice. On échantillonne de façon jointe $\mathbf{w}, \mathbf{u}|\mathbf{c}$ en tirant d'abord $w_1, w_2, \ldots, w_{K_n}|\mathbf{c}$, puis $\mathbf{u}|w_1, w_2, \ldots, w_{K_n}, \mathbf{c}$, et enfin $w_{K_n+1}, w_{K_n+2}, \ldots|\mathbf{u}$.

- Tirer w_k pour $k \leq K_n$
$$w_1, \ldots, w_{K_n}, r_{K_n}|\mathbf{c} \sim \mathrm{Dir}\left(n_1 - d, \ldots, n_{K_n} - d, \alpha + K_n\, d\right).$$
- Tirer $u_i|w_1, w_2, \ldots, w_{K_n}, \mathbf{c}$
$$u_i|w_1, w_2, \ldots, w_{K_n}, \mathbf{c} \overset{\text{ind.}}{\sim} \mathcal{U}\left(u_i|0, \min\left(w_{c_i}, \zeta\right)\right).$$
- Faire $u^* = \min(\mathbf{u})$.
- Tirer w_k pour $k > K_n$. Tant que $r_{k-1} > u^*$,
$$\begin{aligned} V_k &\sim \mathrm{Beta}\left(1 - d, \alpha + k\, d\right) \\ w_k &= V_k\, r_{k-1} \\ r_k &= r_{k-1}\left(1 - V_k\right). \end{aligned}$$
- Faire $K^* = \min\left(\{k : r_k < u^*\}\right)$.

À la fin de chaque itération, on enlève tous les w_k tels que $w_k < u^*$ pour $k \leq K_n$ puisque la probabilité de leur allouer des données est nulle.

Distribution conditionnelle de $(\mathbf{c}|\mathbf{w}, \mathbf{u}, \boldsymbol{\Theta}^*, \mathbf{X})$

On a
$$c_i|\mathbf{w}, \mathbf{u}, \boldsymbol{\Theta}^*, \mathbf{X} \overset{\text{ind}}{\sim} \sum_{k=1}^{K^*} w_{k,i}\, \delta_k\left(\cdot\right) \tag{4.15}$$
où
$$w_{k,i} \propto \mathbf{1}\left(w_k > u_i\right) \max\left(w_k, \zeta\right) f_{\mathcal{N}}\left(\boldsymbol{x}_i|\boldsymbol{\theta}_k^*\right)$$
et $\sum_{j=1}^{K^*} w_{k,i} = 1$.

On rappelle qu'à chaque itération, les clusters non-vides sont labellisés en fonction de leur apparition dans le processus d'échantillonnage, ce qui garantit l'échangeabilité.

Distribution conditionnelle de $(\boldsymbol{\Theta}^*|\mathbf{c}, \mathbf{X})$

L'échantillonnage suivant cette distribution est facile grâce aux propriétés de conjugaison de la loi Normal-Inverse Wishart.

On a pour tout $k \leq K_n$ (mise à jour des moyennes et matrices de covariance pour les composantes non-vides),

$$\mathbf{m}_k|\boldsymbol{\Sigma}_k, \mathbf{c}, \mathbf{X} \overset{\text{ind}}{\sim} \mathcal{N}\left(\boldsymbol{\mu}_k, \frac{\boldsymbol{\Sigma}_k}{\rho_k}\right)$$

où

$$\begin{aligned}
\rho_k &= \rho + n_k \\
\boldsymbol{\mu}_k &= \frac{\rho\,\boldsymbol{\mu}_0 + n_k\,\widehat{\boldsymbol{\mu}_k}}{\rho + n_k}, \\
\widehat{\boldsymbol{\mu}_k} &= \frac{1}{n_k}\sum_{\{i:c_i=k\}} \boldsymbol{x}_i
\end{aligned}$$

et

$$\boldsymbol{\Sigma}_k^{-1}|\mathbf{c}, \mathbf{X} \overset{\text{ind}}{\sim} \mathcal{W}(\eta_k, (\eta_k\,\mathbf{S}_k)^{-1})$$

où

$$\begin{aligned}
\eta_k &= n_0 + n_k, \\
\mathbf{S}_k &= \frac{1}{\eta_k}\left[n_0\,\boldsymbol{\Sigma}_0 + n_k\,\widehat{\boldsymbol{\Sigma}_k} + \frac{(\boldsymbol{\mu}_0 - \widehat{\boldsymbol{\mu}_k})(\boldsymbol{\mu}_0 - \widehat{\boldsymbol{\mu}_k})^\mathsf{T}}{\frac{1}{\rho} + \frac{1}{n_k}}\right], \\
\widehat{\boldsymbol{\Sigma}_k} &= \frac{1}{n_k}\sum_{\{i:c_i=k\}} (\boldsymbol{x}_i - \widehat{\boldsymbol{\mu}_k})(\boldsymbol{x}_i - \widehat{\boldsymbol{\mu}_k})^\mathsf{T}.
\end{aligned}$$

Le tirage des paramètres pour les composantes non affectées s'effectue suivant la loi *a priori* :

$$\boldsymbol{\theta}_k^* \overset{\text{i.i.d.}}{\sim} \mathcal{NIW}_{\rho,n_0,\boldsymbol{\mu}_0,\boldsymbol{\Sigma}_0}, \text{ pour } K_n < k \leq K^*.$$

Remarques : La moyenne $\boldsymbol{\mu}_k$ d'une composante est une somme pondérée de la moyenne *a priori* $\boldsymbol{\mu}_0$ et la moyenne empirique $\widehat{\boldsymbol{\mu}_k}$. Si n_k est grand, la loi *a priori* devient non-informative. On remarque aussi de même l'influence de la variance empirique $\widehat{\boldsymbol{\Sigma}_k}$ sur la variance $\boldsymbol{\Sigma}_k$.

Distribution conditionnelle de $(\mathbf{X}|\mathbf{y}, \mathbf{w}, \mathbf{u}, \boldsymbol{\Theta}^*)$

Comme nous l'avons souligné, l'échantillonnage direct suivant cette distribution conditionnelle est impossible à cause de la distribution de projection. On utilise une étape Metropolis-Hastings et une loi candidate (loi de proposition) pour le tirage conditionnel du lieu d'émission. La loi de proposition est choisie comme étant le produit entre le mélange Pitman-Yor et une loi normale dont les paramètres sont choisis de telle sorte à approximer la distribution de la projection $\mathcal{P}$ définie dans l'équation (4.6). Le résultat est un mélange de gaussiennes dans la direction de la ligne de réponse considérée.

Plus précisément, utilisant la règle de Bayes, on peut écrire

$$\boldsymbol{x}_i|y_i, u_i, \mathbf{w}, \boldsymbol{\Theta}^* \overset{\text{ind}}{\sim} \mathcal{P}(y_i|\boldsymbol{x}_i) \times \widetilde{G}_i(\boldsymbol{x}_i)$$

avec

$$f_{\widetilde{G}_i}(\boldsymbol{x}_i) \propto \sum_{k=1}^{K^*} \mathbf{1}\left(w_k > u_i\right) \max\left(w_k, \zeta\right) f_{\mathcal{N}}\left(\boldsymbol{x}_i|\boldsymbol{\theta}_k^*\right)$$

et $\mathcal{P}(y_i|\boldsymbol{x}_i)$ est la loi de projection des données observées, dont les probabilités sont définies dans le paragraphe 4.3.1. Puisque ce projecteur n'a pas une forme standarde, on utilise une étape Metropolis-Hastings (MH) pour générer des échantillons suivant la loi conditionnelle de $\boldsymbol{x}_i$. Soit G_i^* la distribution candidate pour $i = 1, \ldots, n$. À chaque itération $t+1$, on doit générer $\boldsymbol{x}_i^* \sim G_i^*$, l'accepter ou le refuser suivant la probabilité d'acceptation. On propose un schéma MH indépendant où

$$f_{G_i^*}(\boldsymbol{x}_i^*) \propto f_{\mathcal{N}}(\boldsymbol{x}_i^*|\mathbf{m}_{y_i}^*, \boldsymbol{\Sigma}_{y_i}^*) \times f_{\widetilde{G}_i}(\boldsymbol{x}_i^*).$$

Cela signifie que pour construire la distribution candidate, on remplace $\mathcal{P}(y_i|\boldsymbol{x}_i)$ par une Gaussienne $\mathcal{N}(\boldsymbol{x}_i^*|\mathbf{m}_{Y_i}^*, \boldsymbol{\Sigma}_{y_i}^*)$ dont les paramètres $\mathbf{m}_{y_i}^*$ et $\boldsymbol{\Sigma}_{y_i}^*$ sont choisis de telle sorte à approximer le projecteur $\mathcal{P}(y_i|\boldsymbol{x}_i)$. Nous avons choisi une Gaussienne très allongée dans la direction de la LOR et dont l'extension dans le plan orthogonal correspond à la taille du tube de réponse ($\approx$ dimensions du capteur). On peut alors écrire $f_{G_i^*}$ comme un mélange re-pondéré de densités normales,

$$f_{G_i^*}(\boldsymbol{x}_i^*) \propto \sum_{k=1}^{K^*} \mathbf{1}\left(w_k > u_i\right) \max\left(w_k, \zeta\right) \widetilde{\eta}_{k,y_i} \times f_{\mathcal{N}}\left(\boldsymbol{x}_i^*|\widetilde{\mathbf{m}}_{k,y_i}, \widetilde{\boldsymbol{\Sigma}}_{k,y_i}\right)$$

avec pour tout $k \leq K^*$ et pour tout $l \leq L$,

$$\begin{aligned}
\widetilde{\eta}_{k,l} &= f_{\mathcal{N}}\left(\mathbf{m}_l^*|\mathbf{m}_k,\ \boldsymbol{\Sigma}_l^* + \boldsymbol{\Sigma}_k\right) \\
\widetilde{\mathbf{m}}_{k,l} &= \left((\boldsymbol{\Sigma}_l^*)^{-1} + \boldsymbol{\Sigma}_k^{-1}\right)^{-1} \left((\boldsymbol{\Sigma}_l^*)^{-1}\mathbf{m}_l^* + \boldsymbol{\Sigma}_k^{-1}\mathbf{m}_k\right) \\
\widetilde{\boldsymbol{\Sigma}}_{k,l} &= \left((\boldsymbol{\Sigma}_l^*)^{-1} + \boldsymbol{\Sigma}_k^{-1}\right)^{-1}.
\end{aligned}$$

On peut maintenant procéder à l'échantillonnage de $\boldsymbol{x}_i$, pour tout $i \leq n$, à l'itération $t+1$.

– Générer $\widetilde{c}_i$ tel que

$$\widetilde{c}_i \overset{\text{ind}}{\sim} \sum_{k=1}^{K^*} \widetilde{w}_{k,i}\delta_k(\cdot)$$

avec

$$\widetilde{w}_{k,i} \propto \mathbf{1}\left(w_k > u_i\right) \max\left(w_k, \zeta\right) \widetilde{\eta}_{k,y_i}$$

et $\sum_{k=1}^{K^*} \widetilde{w}_{k,i} = 1$.

– Générer $\boldsymbol{x}_i^*$ à partir de la composante choisie

$$\boldsymbol{x}_i^* \overset{\text{ind}}{\sim} \mathcal{N}\left(\boldsymbol{x}_i^*|\widetilde{\mathbf{m}}_{\widetilde{c}_i,y_i}, \widetilde{\boldsymbol{\Sigma}}_{\widetilde{c}_i,y_i}\right).$$

– Calculer le rapport d'acceptation $\omega(\boldsymbol{x}_i^*, \boldsymbol{x}_i^{(t)})$

$$\omega(\boldsymbol{x}_i^*, \boldsymbol{x}_i^{(t)}) = \frac{\mathbb{P}(y_i|\boldsymbol{x}_i^*)}{\mathbb{P}(y_i|\boldsymbol{x}_i^{(t)})} \cdot \frac{f_{\mathcal{N}}\left(\boldsymbol{x}_i^{(t)}|\mathbf{m}_{y_i}^*, \boldsymbol{\Sigma}_{y_i}^*\right)}{f_{\mathcal{N}}\left(\boldsymbol{x}_i^*|\mathbf{m}_{y_i}^*, \boldsymbol{\Sigma}_{y_i}^*\right)}.$$

On peut noter que $\widetilde{G}_i(\boldsymbol{x}_i)$ disparaît du rapport d'acceptation car il apparaît à la fois au numérateur et au dénominateur.

– Accepter $\boldsymbol{x}_i^*$ (i.e prendre $\boldsymbol{x}_i^{(t+1)} = \boldsymbol{x}_i^*$, sinon faire $\boldsymbol{x}_i^{(t+1)} = \boldsymbol{x}_i^{(t)}$) avec la probabilité

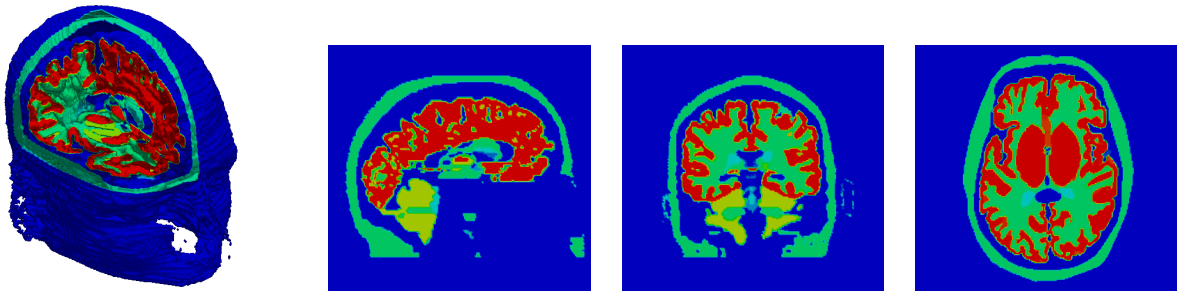
$$\min(1,\ \omega(\boldsymbol{x}_i^*, \boldsymbol{x}_i^{(t)})).$$

4.4 Comparaisons et résultats

Dans cette partie, nous décrivons les résultats des études comparatives. Pour évaluer les performances de la méthode proposée, nous l'avons testée et comparée avec les méthodes de maximum de vraisemblance (ML) et de maximum *a posteriori* (MAP).

Matériel

Un algorithme Monte-Carlo permettant de simuler des examens TEP a été utilisé. L'algorithme combine le modèle de la caméra ECAT EXACT HR$^+$ (Siemens Healthcare) [BZA+97] avec un fantôme numérique cérébral 3D discrétisé en voxels et basé sur une version modifiée du fantôme de Zubal ([ZHS+94]). Le fantôme est composé de $256 \times 256 \times 128$ voxels, de taille de voxel $1.1 \times 1.1 \times 1.1$ mm. Il est présenté à la Figure 4.10. La



fantôme 3D | Vue sagittale | Vue coronale | Vue axiale

FIGURE 4.10 – Fantôme utilisé pour la simulation d'examens TEP.

simulation Monte-Carlo effectuée correspond à un examen de type FDG et est de telle sorte que $n = 10^7$ soient enregistrés, en prenant en compte le champ de vue du système. Les photons sont détectés par un scanner à 32 couronnes, de rayon de couronne 412mm, de champ de vue axial 155mm et de section 4×4 mm^2. Chaque couronne est composée de 576 détecteurs individuels. Ceci correspond à la géométrie du tomographe HR$^+$. Les effets d'atténuation et de coïncidences fortuites et diffusées n'ont pas été simulés.

Nous allons maintenant détailler la paramétrisation de chacun des algorithmes.

Méthode proposée : approche bayésienne non paramétrique (BNP)

Pour la mise en œuvre de notre algorithme, les paramètres du PYM et du $\mathcal{NIW}$ ont été choisis comme suit : $\alpha = 1000$, $d = 0.02$, $\boldsymbol{\Sigma}_0 = 6.25 \times \mathbb{I}_3$ (où $\mathbb{I}_3$ désigne la matrice identité de $\mathbb{R}^3$), $n_0 = 4, \rho = 10^{-2}$, $\boldsymbol{\mu}_0 = (0,0,0)^\mathsf{T}$. Ces choix spécifiques pour ces hyperparamètres ont été trouvés convenables d'après les expérimentations. Pour permettre plus de flexibilité dans le modèle, on peut aussi estimer ces hyperparamètres. Pour cela, on doit placer une distribution *a priori* sur un ou plusieurs de ces paramètres et inférer sur leur loi *a posteriori* . Cette procédure sera discutée dans les perspectives de la thèse, au chapitre 6.

Pour l'algorithme MCMC, 15.000 itérations ont été effectuées dont les 5.000 premières ont été écartées pour la période de burn-in. Après convergence, chaque itération complète de l'algorithme fournit des échantillons suivant la loi *a posteriori* jointe de $(\boldsymbol{x}, c, \boldsymbol{\theta}^*, w|y)$. Les échantillons obtenus à l'itération t, $(\boldsymbol{x}^{(t)}, c^{(t)}, \boldsymbol{\theta}^{*(t)}, w^{(t)})$, peuvent être utilisés pour calculer une estimée de la loi *a posteriori* d'activité $G(\boldsymbol{x})|\boldsymbol{y}$ des évènements enregistrés, de densité

$$f_{G^{(t)}}(\boldsymbol{x}) \approx \sum_{k=1}^{K^*} w_k^{(t)} f_{\mathcal{N}}\left(\boldsymbol{x}|\boldsymbol{\theta}_k^{*(t)}\right)$$

où K^* désigne le nombre de composantes retenues à chaque itération par l'algorithme du slice. Cette densité comporte donc deux parties à savoir les composantes représentées et les composantes non-représentées. Utilisant l'équation 4.8, la mesure de probabilité de tous les évènements s'exprime suivant

$$f_{G^{\dagger(t)}}(\boldsymbol{x}) \propto \frac{f_{G^{(t)}}(\boldsymbol{x})}{\sum_{l=1}^{L} \mathbb{P}(y=l|\boldsymbol{x})}. \tag{4.16}$$

L'estimateur que nous avons choisi pour la distribution d'activité est l'espérance conditionnelle de f_G sur les $N = 10.000$ tirages retenus,

$$\mathbb{E}(f_{G^\dagger}|\boldsymbol{Y}) \approx \frac{1}{N} \sum_{t=1}^{N} f_{G^{\dagger(t)}}.$$

Cet estimateur est aussi la densité prédictive de l'activité comme nous l'avons vu dans le chapitre 2, $\mathbb{E}(G^\dagger(\cdot)|\boldsymbol{Y}) = \mathbb{P}(\boldsymbol{x}_{n+1} \in \cdot|\boldsymbol{Y})$. Elle donne la probabilité de la localisation d'une annihilation non encore observée sachant n émissions déjà observées.

La méthode proposée a été comparée avec les approches du maximum de vraisemblance (ML) et du maximum *a posteriori* (MAP).

Approches du maximum de vraisemblance (ML) et du maximum *a posteriori* (MAP)

Ces méthodes ont déjà été décrites dans la section 4.2.3. On rappelle que dans ces algorithmes, l'image à reconstruire est discrétisée et représentée sous forme vectorielle $\boldsymbol{f} = (f_j, j = 1, \ldots, J)$, f_j désignant l'intensité dans le voxel j.

Dans ML, on cherche l'image $\boldsymbol{f}$ qui maximise la log-vraisemblance des observations, vue comme une fonction de $\boldsymbol{f}$ et notée $L(\boldsymbol{f})$ en utilisant l'algorithme EM. Les images ont été post-filtrées après convergence par un filtre gaussien dont la largeur totale à mi-hauteur (LTMH) est de 4 mm. La largeur du filtre est fixée de façon à minimiser l'erreur moyenne quadratique par rapport au fantôme.

Dans la méthode du MAP, on cherche l'image $\boldsymbol{f}$ qui maximise la loi *a posteriori*

$$p(\boldsymbol{f}|\boldsymbol{y}) = \frac{p(\boldsymbol{y}|\boldsymbol{f})p(\boldsymbol{f})}{p(\boldsymbol{y})}$$

avec $p(\boldsymbol{y}|\boldsymbol{f})$ désignant la vraisemblance poissonnienne des données et $p(\boldsymbol{f})$ la loi *a priori* sur l'image. Nous avons modélisé cette loi *a priori* par un champ de Gibbs dans le but de favoriser des intensités similaires dans les voxels voisins tout en préservant les contours. On a alors,

$$p(\boldsymbol{f}) \propto \exp\left(-\beta U(\boldsymbol{f})\right) \propto \exp\left(-\beta \sum_{r,s} w_{rs}\ \psi\left(\frac{f_r - f_s}{\delta}\right)\right)$$

où β et δ sont des paramètres, w_{rs} est un poids déterminant le voisinage entre les voxels r et s et $\psi(u)$ est une fonction paire, minimisée à $u = 0$. La fonction potentielle ψ que nous avons choisie est la fonction log cosh suggérée par [Gre90]. Avec cette fonction, la fonction d'énergie U est convexe. Comme la log-vraisemblance est concave, le critère Φ l'est aussi. De plus, la fonction log cosh permet de moins pénaliser les discontinuités comparée à d'autres fonctions de pénalité. Les paramètres β et δ sont choisis de telle sorte à minimiser l'erreur moyenne quadratique par rapport au fantôme. Les valeurs $\beta = .25$ et $\delta = 10$ ont été obtenues. Le voisinage d'un voxel est composé par l'ensemble des cinq voisins sur chacune de ses faces. Enfin, le poids $w_{r,s}$ est donné par l'inverse de la distance quadratique entre les deux voxels r et s. L'image est reconstruite en maximisant la fonction de coût suivante,

$$\Phi(\boldsymbol{f}, \boldsymbol{y}) = L(\boldsymbol{y}|\boldsymbol{f}) - \beta \sum_{r,s} w_{rs}\ \psi\left(\frac{f_r - f_s}{\delta}\right)$$

où L désigne le logarithme de la vraisemblance. L'algorithme utilisé pour maximiser cette fonction est le One-Step-Late (OSL) de Green et présenté dans la section 4.2.3. Avec la fonction log cosh, l'algorithme OSL converge vers la solution MAP pour des valeurs de β peu élevées.

Résultats et interprétations

Les Figures 4.11 -4.13 montrent en 3D et sur des vues axiale et coronale le fantôme utilisé pour générer les données, ainsi que les cartes d'activité reconstruites par l'approche proposée ainsi que celles obtenues par les approches ML et MAP. Il est à signaler que dans notre approche, la discrétisation est seulement effectuée pour visualiser l'image et a été choisie *a posteriori* suivant la taille du fantôme, soit $256 \times 256 \times 128$. La même discrétisation a été utilisée pour les algorithmes MAP et ML.

Tout d'abord, on peut visuellement constater sur les courbes 3D de la Figure 4.11 que notre reconstruction a lieu dans un espace continu tout en préservant les bords, et

ce même dans les régions froides. Les approches ML et MAP fournissent quant à elles des images très bruitées. On note aussi sur les coupes 2D que l'estimateur BNP exhibe des régions lisses (matière blanche) et des régions plus sinueuses (matière grise, crâne). Cela est dû à la capacité du modèle $\mathcal{NIW}$ d'estimer une matrice de covariance pour chaque composante du mélange.

De façon plus détaillée, on remarque de manière générale dans la Figure 4.12 une dégradation de l'activité dans les structures de faible épaisseur, en particulier dans le crâne et les ventricules. Cette dégradation est probablement accentuée par les effets de volume partiel (EVP) non liés à l'échantillonnage (effets de spill-in et de spill-out) [1] puisque la région crânienne est soumise sur toute sa surface aux effets de spill-over de la part du cortex et les ventricules à ceux moins importants de la matière blanche. La dégradation est beaucoup plus importante dans les algorithmes ML et MAP. Dans la méthode BNP, la délimitation du crâne est assez nette et l'activité dans les ventricules mieux retrouvée. Par ailleurs, on remarque aussi que l'effet de spill-over du putamen et des noyaux caudés sur les structures environnantes impacte beaucoup moins l'approche BNP. Ces résultats confirment le fait que l'absence d'une discrétisation ad hoc dans la méthode proposée permet d'éviter la dégradation supplémentaire due à la discrétisation de l'espace.

Ces résultats illustrent de la capacité de l'approche proposée à améliorer le rapport signal sur bruit et la résolution de l'image lorsque l'on travaille avec un nombre relativement faible de données. Le nombre potentiellement infini de composantes participe à la régularisation du problème puisque le nombre de composantes allouées est adapté en fonction de l'information disponible dans les données (nombre de composantes actives ≈ 4000). Le caractère non paramétrique se traduit ainsi par le fait que ce nombre peut augmenter avec les observations et ainsi améliorer la résolution de la distribution d'activité. La stratégie du « slice sampling » que nous avons adoptée permet d'éviter une troncature du modèle infini tout en imposant un nombre fini de composantes du mélange gaussien à chaque itération (≈ 10.000). L'algorithme MCMC présente aussi de bonnes propriétés de mélange, grâce à la propriété d'échangeabilité intrinsèque des composantes ainsi qu'au seuil choisi.

1. De par la limitation de la résolution spatiale du tomographe, on observe dans les petites structures un « échappement » de l'activité résultant en une sous-évaluation de l'activité réelle, c'est l'effet dit de « spill-out ». Cet « échappement » de l'activité contamine les structures environnantes, conduisant à une sur-estimation de leur activité, c'est l'effet de « spill-in ». Ces effets sont accentués par la discrétisation de l'espace en voxels où ces effets sont particulièrement visibles au niveau des frontières des régions fonctionnelles. La combinaison de ces effets est dite effets de volume partiel (EVP). Dans certains cas, la sous-estimation de l'activité due au spill-out est compensée par la contamination provenant des régions chaudes « spill-over ». Cependant l'EVP se traduit la plupart du temps par une sous-estimation des concentrations radioactives mesurées pour toutes les régions dont la taille est inférieure à deux ou trois fois la résolution spatiale.

Enfin, avec l'approche adoptée on caractérise toute la distribution et on ne se limite pas à une estimée ponctuelle comme dans les approches ML et MAP. La distribution *a posteriori* de l'incertitude est directement accessible pour l'analyse quantitative de toute région d'intérêt puisqu'on peut estimer n'importe quelle fonctionnelle *a posteriori*. On montre dans la Figure 4.14, l'écart-type conditionnel de notre estimée par rapport au fantôme. L'inférence bayésienne permet de calculer un intervalle de crédibilité pour le paramètre d'intérêt. Étant données les observations, un intervalle crédible à $100(1-2p)\%$, doit inclure *au moins* $100(1-2p)\%$ de la vraie valeur du paramètre. L'intervalle de crédibilité pour un paramètre θ_i, noté $[c_p, c_{1-p}]$, peut être estimé en prenant respectivement pour c_p et c_{1-p} les quantiles d'ordre p et $1-p$ des n valeurs simulées par l'échantillonnage Monte-Carlo, $\{\theta_i^{(t)}, t = 1, \ldots, n\}$. Nous utilisons cette technique pour calculer un intervalle crédible à 95% de la concentration d'activité sur un profil horizontal au niveau du putamen. Pour cela, on prend les quantiles d'ordre 2.5% et 97.5%. Les résultats sont montrés à la Figure 4.16. La concentration d'activité sur ce même profil estimée par ML et BNP est montrée dans la Figure 4.15.

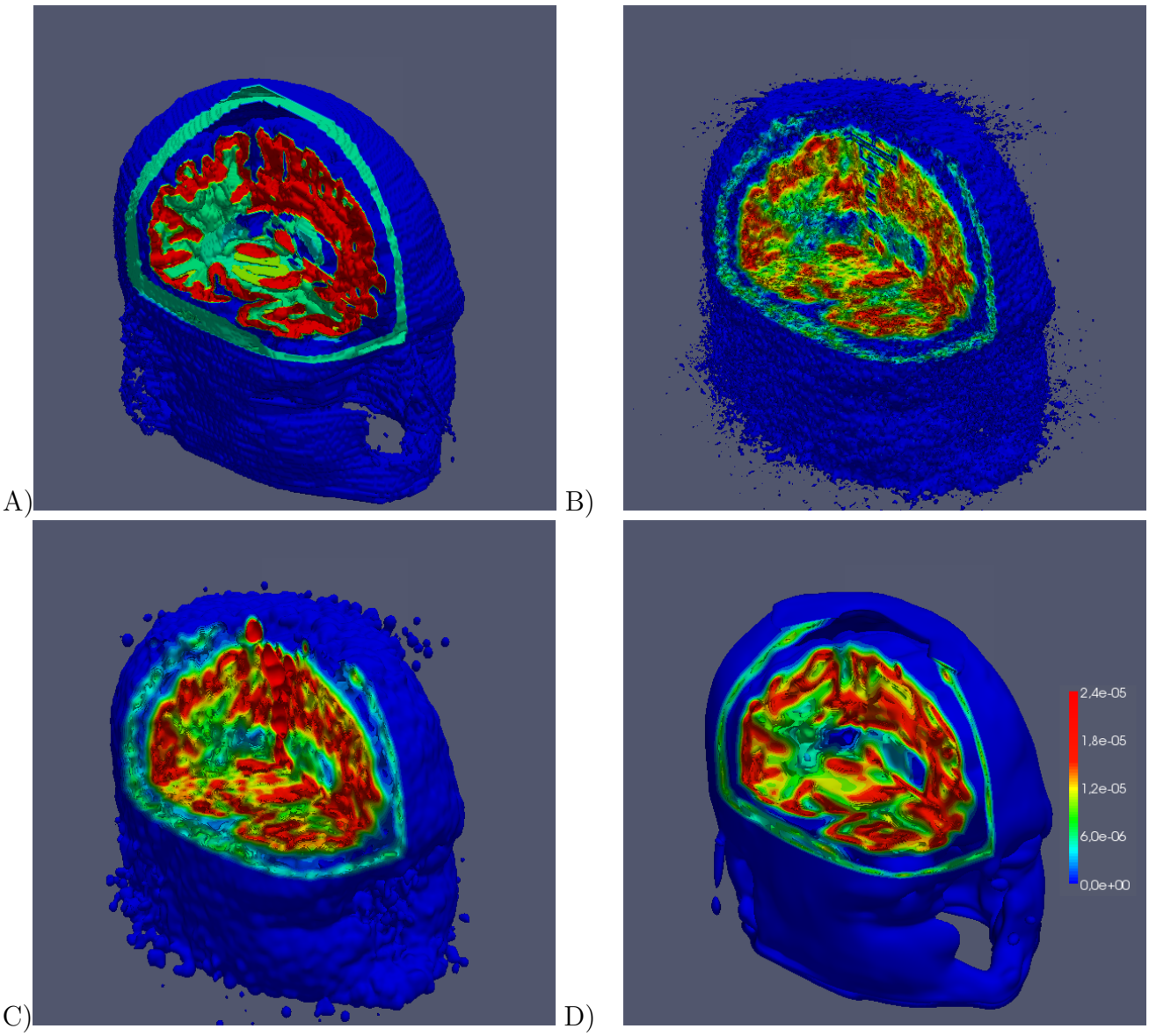


FIGURE 4.11 – Isosurfaces : A) fantôme 3D ; B) Reconstruction MAP ; C) Reconstruction ML et D) Reconstruction BNP

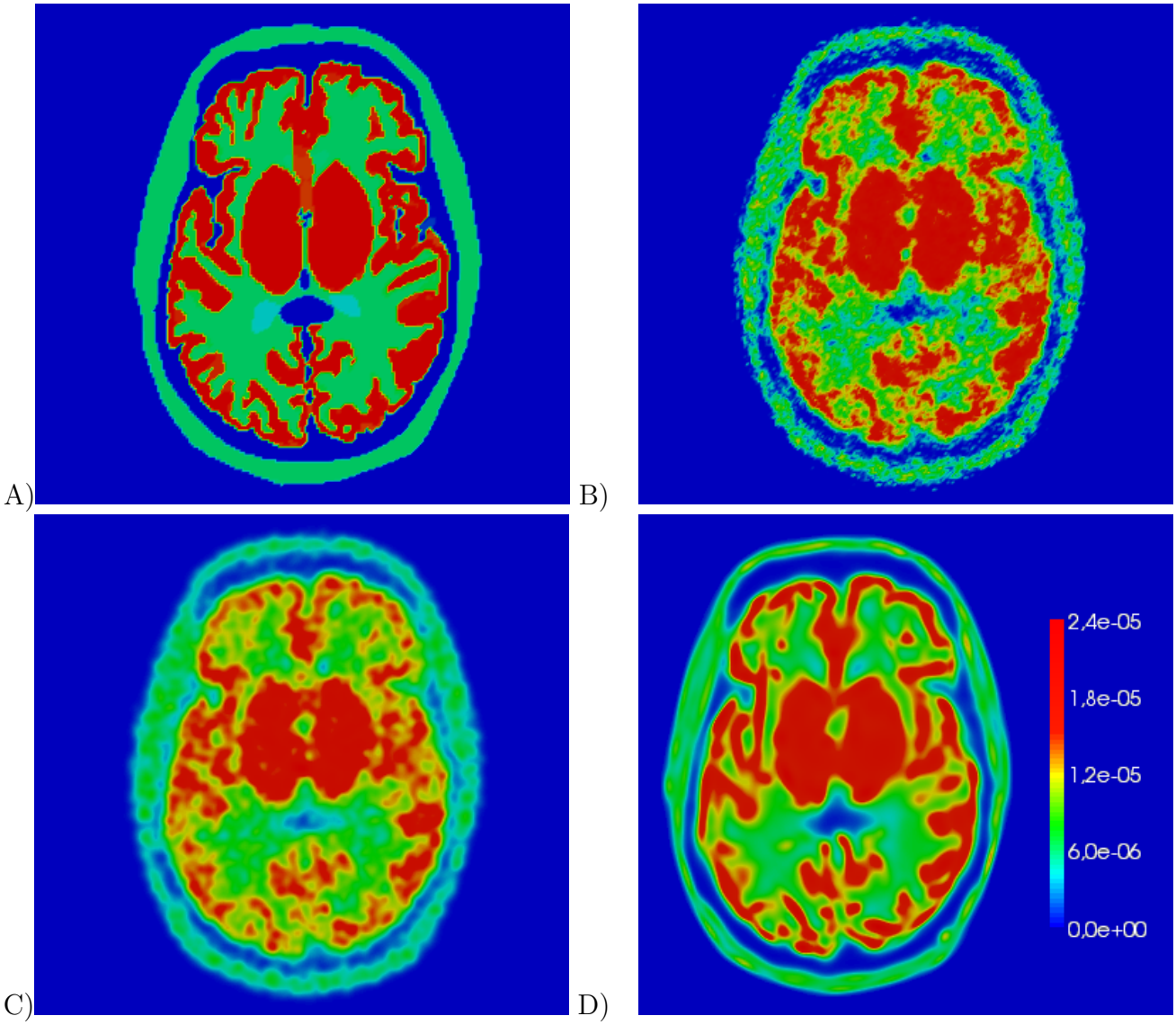


FIGURE 4.12 – Vues axiales : A) fantôme 3D ; B) Reconstruction MAP ; C) Reconstruction ML et D) Reconstruction BNP

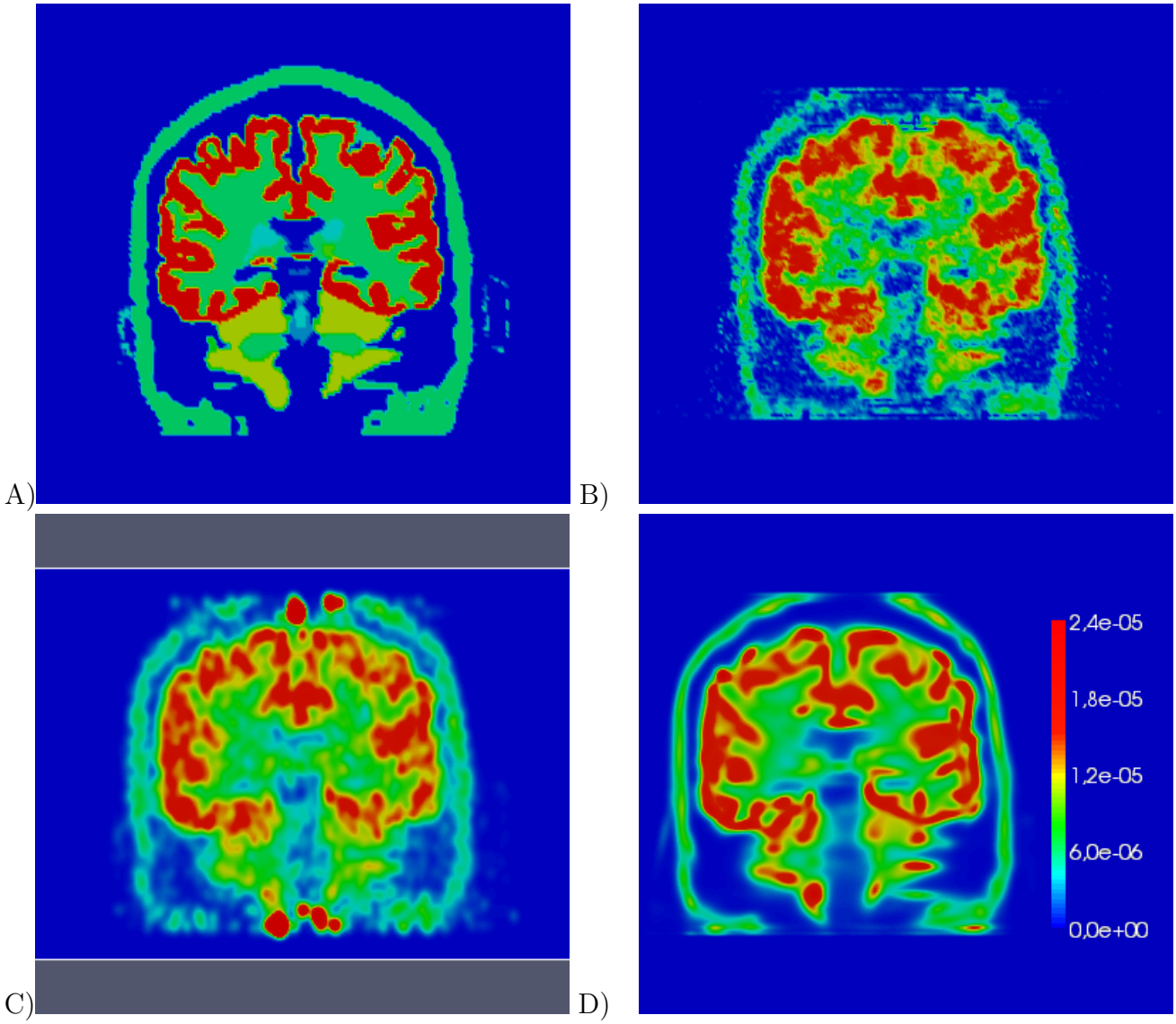


FIGURE 4.13 – Vues coronales : A) fantôme 3D ; B) Reconstruction MAP ; C) Reconstruction ML et D) Reconstruction BNP

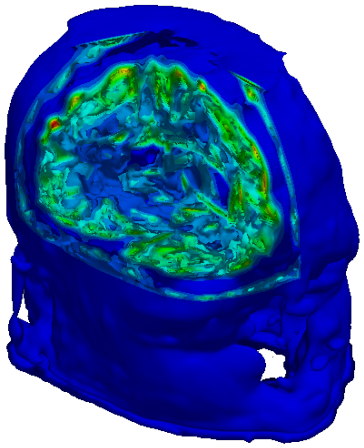


FIGURE 4.14 – Écart-type conditionnel de la méthode BNP par rapport au fantôme.

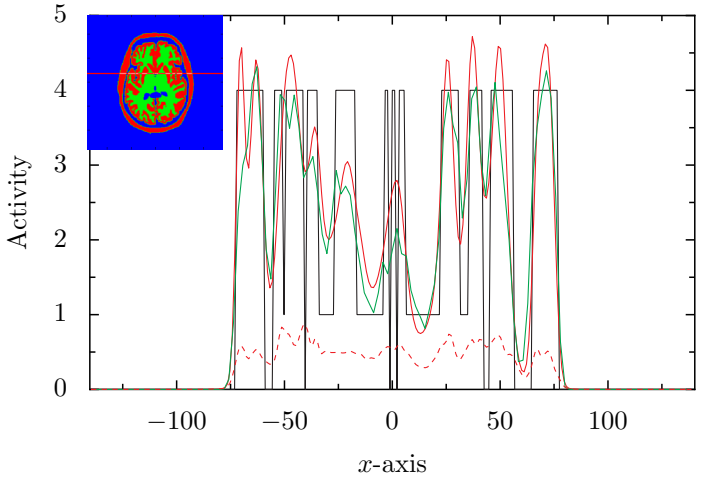


FIGURE 4.15 – Estimées de la concentration d'activité sur un profil horizontal au niveau du putamen : ML post-filtré (vert), BNP (rouge). L'activité du fantôme est tracée en noir et l'écart-type conditionnel estimé par l'algorithme BNP en pointillés rouges.

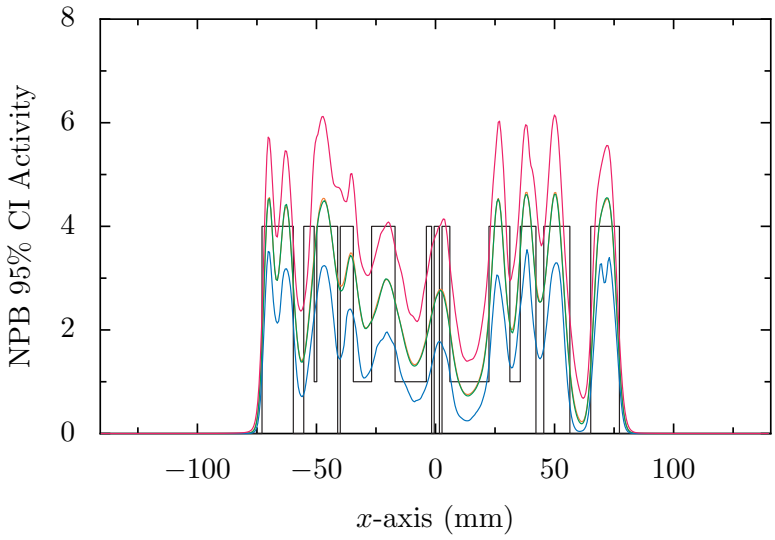


FIGURE 4.16 – Intervalle crédible à 95% sur le même profil. On trace les quantiles d'ordre 97.5% (rouge) et 2.5% (bleu). La médiane de l'activité est dessinée en vert et l'activité du fantôme en noir.

Afin d'évaluer la qualité des reconstructions, nous avons comparé sur une coupe axiale 2D la racine carrée de l'erreur quadratique moyenne normalisée (EMQN) (norme L_2) obtenue avec l'approche proposée ainsi que les techniques MAP et ML post-filtré. La référence est le fantôme. L'EMQN est donnée par :

$$\sqrt{\frac{\sum_i (p_i - q_i)^2}{\sum_i p_i^2}}$$

où les p_i sont les valeurs réelles (fantôme) et les q_i les valeurs estimées. Les résultats sont donnés dans le tableau 4.1.

MAP	**ML post-filtré**	**BNP**
45.6%	41.7%	41.8%

TABLE 4.1 – EMQN (%) pour reconstructions ML post-filtré, MAP et BNP

D'après ce premier tableau, les résultats obtenus par ML post-filtré et BNP sont similaires. La méthode MAP fournit la plus grande erreur. Toutefois l'EMQN n'est pas un critère optimal. En effet, il s'agit d'une mesure globale de dissimilarité, qui pénalise de la même façon la sur-estimation ou la sous-estimation de l'activité dans les zones froides et les zones chaudes. Des critères plus pertinents sont basés sur les mesures de disparité entre des mesures de probabilité, comme la divergence de Kullback ou la distance de Hellinger. Nous avons donc procédé à une comparaison plus poussée entre ML post- filtré et BNP sur plusieurs critères définis comme suit :

1. Divergence de Kullback :
$$\sum_i p_i \frac{\ln p_i}{q_i}$$
2. Distance de Hellinger :
$$\sqrt{\frac{1}{2}\sum_i \left(\sqrt{p_i} - \sqrt{q_i}\right)^2}$$
3. Nous avons aussi calculé la norme L_1 c'est-à-dire l'erreur absolue moyenne normalisée (EAMN) :
$$\frac{\sum_i |p_i - q_i|}{\sum_i |p_i|}.$$

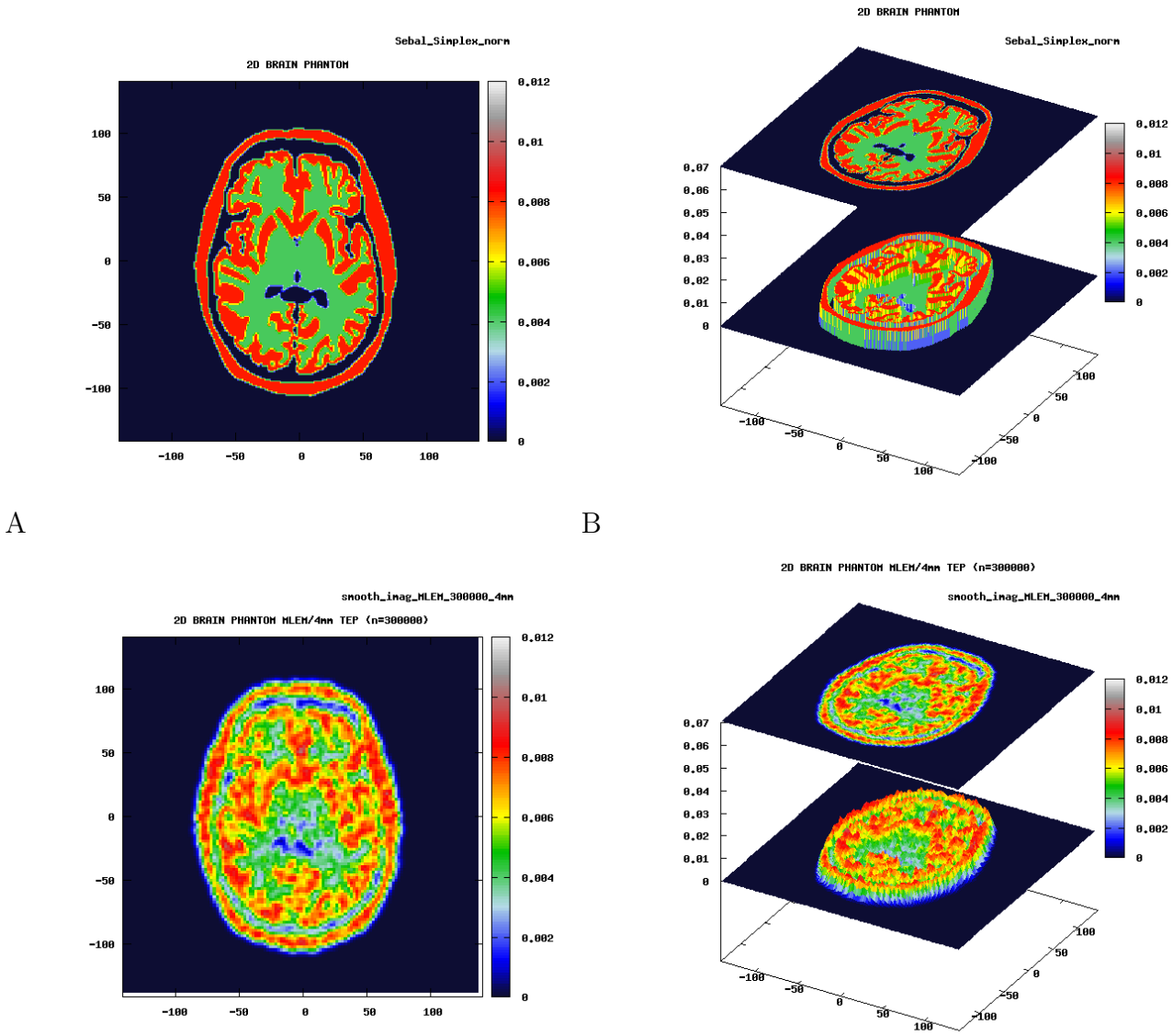
Les résultats de comparaison entre l'approche proposée (BNP) et l'algorithme ML-EM pour différentes LTMH de filtres sont donnés dans le tableau 4.2. Les images du fantôme 2D et des reconstructions BNP et ML-EM (pour un filtre de 4mm) sont montrées dans les Figures 4.17 et 4.18.

Algorithme	DH	KL	EMQN	EAMN
BNP	26.23	19.41	41.88	41.47
MLEM/3mm	28.56	24.06	48.03	47.65
MLEM/4mm	28.22	20.99	41.89	43.17
MLEM/5mm	29.31	21.75	41.75	44.58
MLEM/6mm	30.69	23.50	43.18	47.55
MLEM/7mm	32.09	25.57	45.08	50.97
MLEM/8mm	33.40	27.69	47.06	54.37

TABLE 4.2 – Performances (en %) pour les reconstructions ML et BNP : (1er et 2ème).

Ces résultats confirment que l'approche proposée permet d'obtenir des meilleurs résultats en termes d'erreurs minimales que les approches ML et MAP.

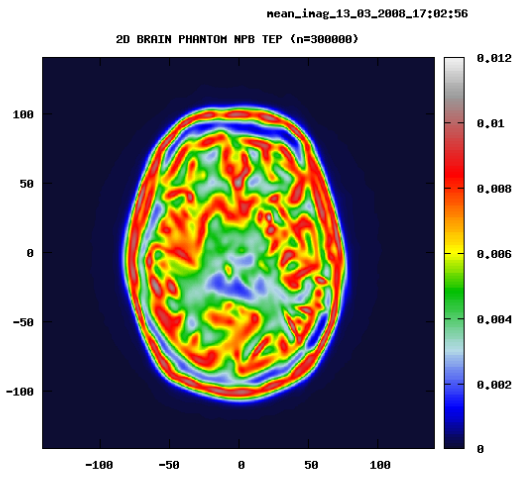
Toutefois une étape de validation statistique de l'estimateur proposé est nécessaire. Le but est de pouvoir comparer, sur n'importe quelle région cérébrale pré-définie, *les intervalles crédibles bayésiens* obtenus par notre méthode avec des *intervalles de confiance* estimés à partir des réplicats.



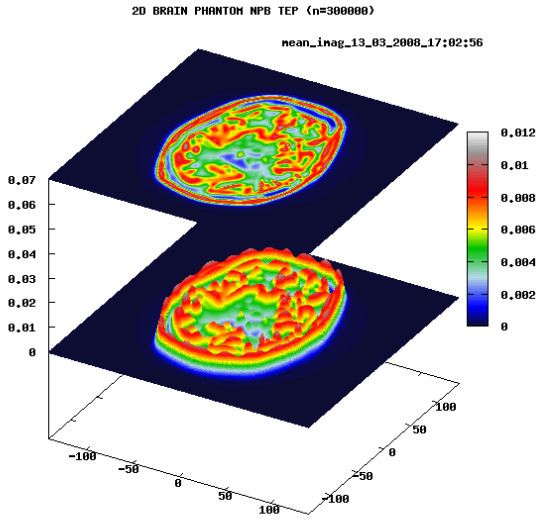
A B

C D

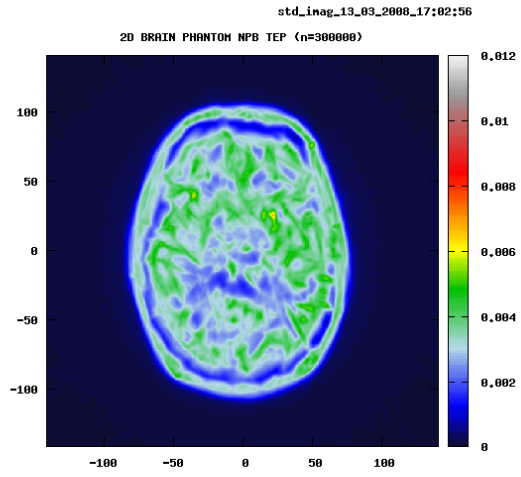
FIGURE 4.17 – A-B : fantôme 2D ; C-D : Reconstruction ML-EM post-filtré (4mm)



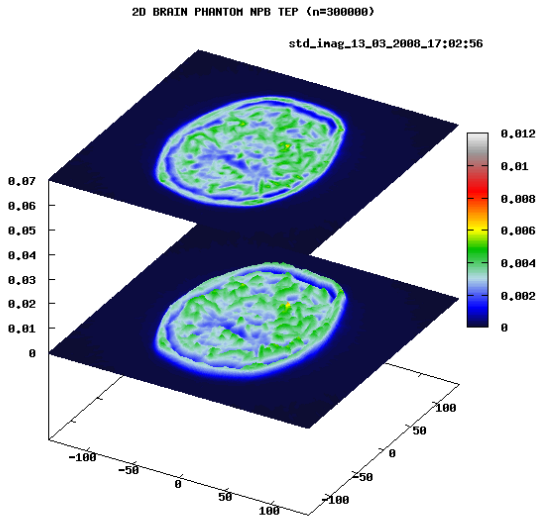
A



B



C



D

FIGURE 4.18 – A-B Reconstruction BNP ; C-D Écart-type BNP

Validation statistique de l'estimateur proposé

L'approche bayésienne permet de calculer l'incertitude de reconstruction *a posteriori* se basant sur le seul jeu de données disponible. Cependant, l'estimée bayésienne de la distribution spatiale d'activité *a posteriori* ne peut être utilisée en pratique que si l'on a confiance en elle d'un point de vue fréquentiste. Ce test de validité fréquentiste peut être effectué par une validation statistique empirique, où l'on cherche à établir sur un nombre significatif de réplicats, le nombre d'intervalles crédibles bayésiens contenant la vraie valeur. Un intervalle crédible bayésien à 95% représente l'intervalle dont la probabilité pour qu'il contienne la vraie valeur, se basant sur les observations, est de 0.95. Cela diffère de l'intervalle de confiance fréquentiste représentant l'intervalle qui contiendrait la vraie valeur, avec une fréquence de 0.95, pour un nombre infini de réplicats.

Nous avons généré 20 réplicats pour respectivement $n = 10^7$ et $n = 10^6$ évènements enregistrés. Pour chaque réplicat et pour chaque jeu de données, nous avons effectué une reconstruction BNP où tous les échantillons MCMC sont conservés. Ce travail a été effectué en utilisant les ressources de calcul du Grand Equipement National de Calcul Intensif (GENCI-[CCRT/CINES/IDRIS] Grant 2011-t2011036706). De plus, on voudrait évaluer la capacité de la méthode proposée à détecter de petites différences de concentrations dans un contexte de faible dose. Pour ce faire, un fantôme a été créé de telle sorte que la concentration d'activité dans le putamen droit est diminuée de 20% par rapport à celle du putamen gauche. Le fantôme est représenté à la Figure 4.19.

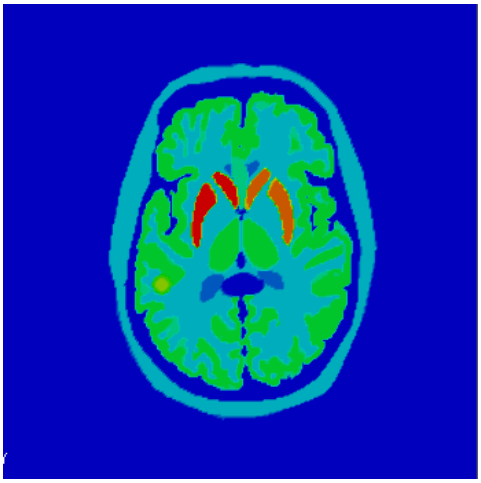


FIGURE 4.19 – Coupe 2D du fantôme cérébral 3D voxellisé : hypofixation de 20% sur le putamen droit.

L'estimateur choisi est toujours l'espérance conditionnelle sur tirages MCMC post burn-in. Se basant sur les volumes labellisés définis à partir du fantôme, on peut calculer la concentration d'activité dans le putamen droit et gauche. À noter qu'ici, nous ne prenons pas en compte l'effet de volume partiel. Ainsi, la valeur de référence pour la vraie valeur de concentration dans un volume donné est choisie comme étant la moyenne de toutes les concentrations d'activité estimées dans tous les réplicats. La Figure 4.20 présente les résultats obtenus pas les réplicats avec $n = 10^7$ et $n = 10^6$ évènements représentant respectivement une dose conventionnelle injectée et $\frac{1}{10}$ de cette dose. Il s'est avéré que les intervalles crédibles sont bien séparés entre le putamen gauche et

droit et que tous les intervalles obtenus avec les réplicats contiennent la vraie valeur (aussi bien pour le putamen gauche que le droit). Enfin, les histogrammes *a posteriori* montrent deux populations bien séparées (cf. Figure 4.20), même avec un faible nombre d'évènements. Ce qui est aussi un indicateur de la bonne mélangeance de la méthode d'échantillonnage MCMC proposée.

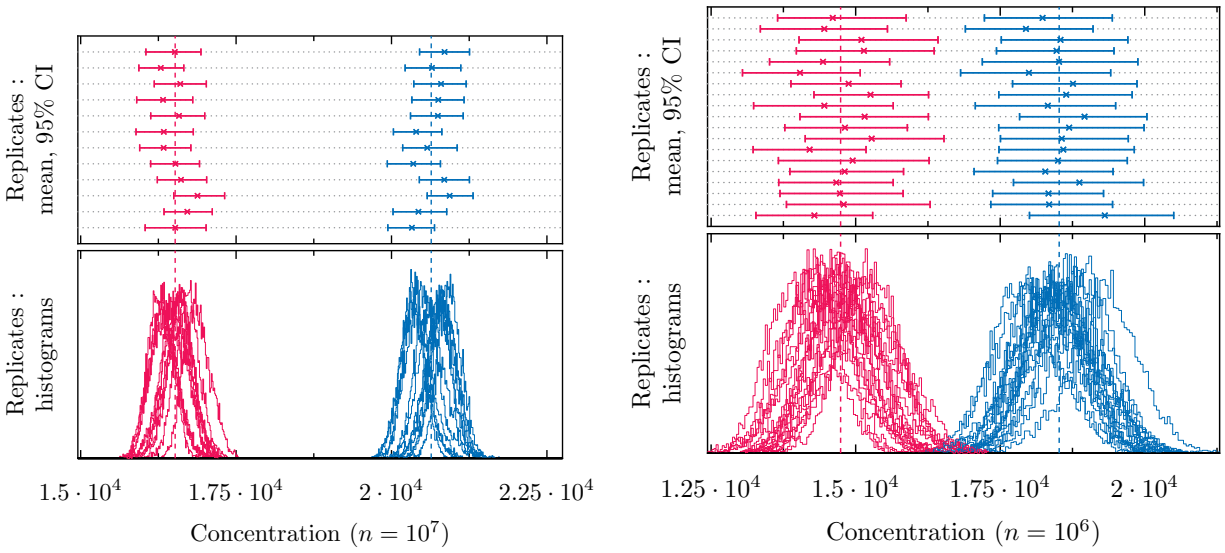


FIGURE 4.20 – $n = 10^7$ (colonne de gauche) et $n = 10^6$ (colonne de droite) désintégrations. Concentration d'activité du putamen droit (rouge) *vs.* putamen gauche (bleu). *Ligne du haut :* Moyenne *a posteriori* estimée (croix), vraie valeur attendue (pointillés) et intervalles crédibles 95% (barres horizontales) pour chaque réplicat. *Ligne du bas :* histogrammes *a posteriori*.

4.5 Conclusion

Lorsque le nombre d'évènements est relativement faible, il est important d'utiliser un algorithme permettant une meilleure restitution de la résolution spatiale. Dans cette optique, nous avons proposé une modélisation flexible et numériquement accessible pour la reconstruction d'images 3D en TEP. Afin de s'affranchir de la base figée des voxels, nous avons adopté une approche plus robuste dite non paramétrique, où le nombre de paramètres dans le modèle s'adapte automatiquement aux données et à leur structure. La régularisation du problème a été effectuée dans un cadre bayésien. L'introduction des variables cachées que sont les lieux d'emission permet d'effectuer le clustering directement dans l'espace-objet et ainsi d'éviter toute discrétisation *a priori*. Cette propriété de clustering de l'approche proposée sera d'ailleurs plus explicite dans la reconstruction 4D qui sera abordée au prochain chapitre. L'absence de discrétisation implique qu'il n'y ait pas de matrice système à calculer. Mais la distribution de la projection $\mathcal{P}(y_i|\boldsymbol{x}_i)$, qui dépend entre autres de la géométrie du tomographe, est explicitement inclue dans l'étape Metropolis-Hastings utilisée pour proposer les lieux des émissions.

Les résultats de simulation présentés ont montré la capacité de l'approche proposée à reconstruire de bonnes images dans un contexte de faibles doses injectées. Les struc-

tures sont bien mieux reconstruites en utilisant l'approche proposée, on a une meilleure récupération d'activité dans les zones froides et une meilleure délinéation des structures. De plus, cette approche permet de caractériser toute la loi de la distribution aléatoire d'activité et l'on peut ainsi évaluer des paramètres importants pour la quantification en imagerie médicale tels que les intervalles de crédibilité.

Cependant, la méthode proposée est assez coûteuse en temps de calcul comparée à l'approche EM. Puisqu'il n'y a pas de voxellisation dans cette approche, le coût de calcul n'est pas donné par la matrice-système. En pratique, les boucles des propositions des emissions et les affectations donnent approximativement le coût de calcul (coût proportionnel à # évènements $\times$ # composantes). Ce résultat indique que la méthode est attractive quand on travaille avec de faibles doses injectées. Des approches pour la réduction du temps de calcul seront discutées dans le chapitre 6 sur les perspectives.

Dans les résultats présentés, nous n'avons pas simulé des coïncidences fortuites, diffusées et atténuées. Mais cela n'est pas limitant car la méthode proposée peut être étendue de façon à inclure ces effets au travers de la loi de projection $\mathcal{P}(y_i|\boldsymbol{x}_i)$. Cette approche sera discutée plus en détails dans le chapitre 6.

5

Reconstruction spatio-temporelle en TEP 4D

Dans ce chapitre, nous présentons la méthode proposée pour la reconstruction d'images dynamiques en TEP. Avant cela, nous débutons par une présentation du modèle compartimental utilisé et des méthodes d'estimation des paramètres physiologiques d'intérêt. Ensuite, nous procédons à un bref état de l'art de la reconstruction dynamique en TEP. Le but n'est pas d'être exhaustif mais plutôt de pointer le doigt sur les limites des principales méthodes de reconstruction dynamique en TEP et, par là, proposer une nouvelle méthode originale de reconstruction. Cette méthode est par la suite détaillée et ses performances seront testées sur un fantôme cérébral réaliste.

Avant d'entrer dans le vif du sujet, commençons par un petit glossaire de termes qui seront utilisés tout au long de ce chapitre.

Glossaire :

- La courbe décrivant l'évolution de l'activité au cours du temps (la dynamique du traceur) sera appelée cinétique et notée TAC (pour Time Activity Curve).
- Un volume fonctionnel représente un ensemble de localisations spatiales présentant un comportement cinétique similaire.
- Un compartiment est un ensemble d'entités caractérisées par une indiscernabilité de devenir.

5.1 Introduction aux modèles compartimentaux

Nous avons vu dans le chapitre 4 que la reconstruction d'images statiques permettait d'obtenir une cartographie 3D de la distribution de radioactivité. Cependant, l'extraction de paramètres caractéristiques du processus biologique sondé nécessite de connaître

l'évolution au cours du temps de la distribution volumique du traceur dans les organes. En effet, le devenir d'une molécule dans l'organisme dépend de l'état dans lequel elle se trouve (libre ou liée). Mais l'entité marquée étant la molécule d'intérêt, qu'elle soit liée ou non, on ne peut pas distinguer ces deux formes par la simple détection du marquage. C'est pourquoi l'analyse temporelle est nécessaire pour caractériser les paramètres biologiques. Pour ce faire, on utilise un modèle mathématique basé sur des hypothèses physiologiques pour décrire la pharmaco-cinétique du traceur et de ses dérivés, c'est-à-dire les différentes étapes de leur transformation dans l'organisme cible. Ce modèle, appelé modèle compartimental, permet de synthétiser les informations relatives au processus biologique étudié par l'utilisation de compartiments et de taux d'échange entre ces compartiments. Un compartiment est un ensemble d'entités biologiques indiscernables l'une de l'autre et vouées au même devenir. Il est donc constitué d'un ensemble homogène sur le plan cinétique. Il existe des échanges ou des transferts entre les compartiments. Ces transferts peuvent être réversibles ou irréversibles et peuvent se faire soit par des transports physiques (via le débit sanguin par exemple), soit par des réactions chimiques (transformation en un métabolite). Les taux d'échange entre compartiments sont souvent supposés constants. Ils représentent la probabilité, pour une molécule donnée, d'entrer ou de sortir d'un compartiment. La cinétique du traceur et de ses dérivés est alors décrite par un système d'équations différentielles du premier ordre dont la résolution permettra d'identifier les taux d'échanges entre compartiments qui permettront *in fine* d'extraire les paramètres physiologiques d'intérêt.

On peut résumer la modélisation compartimentale comme étant une représentation simplifiée d'un système biologique au moyen d'un ensemble de compartiments présentant des interactions entre-eux. Les échanges entre compartiments sont formulés par un système d'équations différentielles dont la résolution analytique ou numérique permet d'estimer des paramètres ayant une interprétation biologique. Chaque radiotraceur est caractérisé par des propriétés de fixation et de contraintes d'utilisation qui lui sont propres. Nous nous intéressons ici à la modélisation pharmaco-cinétique en TEP pour l'imagerie cérébrale, après injection du fluoro-déoxyglucose ([^{18}F]–FDG) dans la circulation sanguine.

5.2 Modèle compartimental dans le cas du FDG

5.2.1 Le fluoro-déoxyglucose ($\left[^{18}\mathbf{F}\right]$–FDG)

Le fluoro-déoxyglucose ([^{18}F]–FDG, noté aussi FDG) est l'un des traceurs les plus utilisés en TEP. Il permet d'étudier le métabolisme du glucose. En oncologie, le FDG permet de différentier les tumeurs malignes et bénignes, d'évaluer la réponse à des chimiothérapies. En effet, en raison de leur activité accrue, la consommation de glucose est plus importante dans les cellules tumorales que dans les cellules saines. Le FDG permet ainsi la détection de la tumeur primaire et des métastases et le suivi thérapeutique par la caractérisation de l'évolution de l'atteinte tumorale sous thérapie. En recherche neurologique, le FDG permet d'évaluer l'activité cérébrale pour l'étude de maladies neuro-dégénératives comme Alzheimer par exemple.

Le [^{18}F]–FDG est obtenu en incorporant l'isotope ^{18}F dans le déoxyglucose (DG).

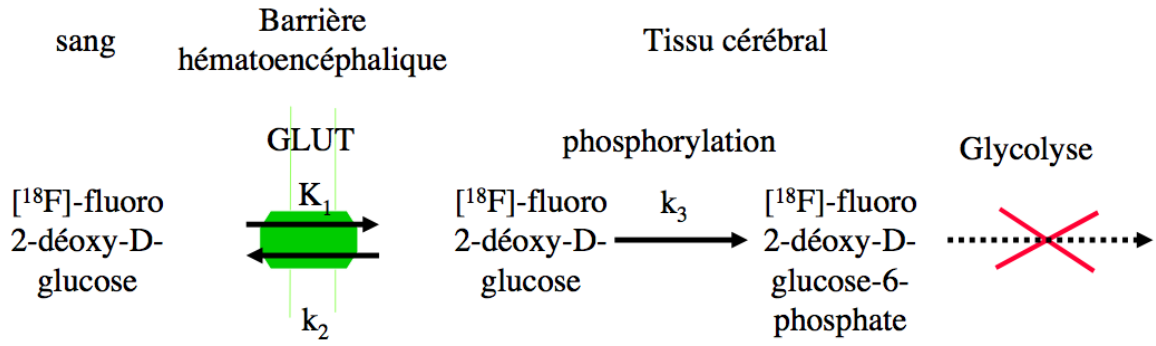


FIGURE 5.1 – Le FDG se fixe sur des enzymes et traverse la barrière hémato-encéphalique (BBB). À l'intérieur du tissu cérébral, il est phosphorylé par les enzymes hexokynase et transformé en FDG-6-P. Ce dernier ne subit pas les autres étapes de la glycolyse et s'accumule dans les cellules. Les constantes définissent la vitesse du transfert des molécules entre les compartiments : K_1 définit le nombre de molécules de FDG traversant la BBB par unité de temps (du sang artériel vers le tissu cérébral), k_2 le taux de transfert dans le sens inverse et k_3 désigne le taux de phosphorylation du FDG.

Ce dernier est un analogue du glucose, c'est-à-dire une composante semblable mais dont les propriétés chimiques sont légèrement différentes de celles du glucose. En effet, tout comme le glucose, le déoxyglucose pénètre la barrière hémato-encéphalique pour rentrer dans les cellules à l'intérieur desquelles il est dégradé par des enzymes appelées hexokynase, c'est la phosphorylation. Mais contrairement à la molécule issue du glucose (Glucose-6-Phosphate), la molécule dérivée du FDG (FDG-6-Phosphate) ne subit pas les autres étapes de la glycolyse et reste piégée dans la cellule sous sa forme phosphorylée. Alors qu'il existe une voie de réaction inverse pour le Glucose-6-Phosphate (la déphosphorylation), la réversibilité de la réaction pour le FDG-6-Phosphate est faible et souvent négligée pour la durée de l'examen. Le FDG-6-Phosphate s'accumule dans la cellule de façon proportionnelle à sa consommation de glucose et reflète ainsi la consommation énergétique des cellules. Ce processus est résumé dans la Figure 5.1.

5.2.2 Le modèle compartimental de Sokoloff

Ce modèle fut initialement décrit pour du déoxyglucose marqué au ^{14}C et sur des animaux [SRK+77]. Le modèle est décrit par un système d'équations différentielles régissant les passages d'un compartiment à un autre et dont la résolution permettra d'extraire les paramètres physiologiques, en particulier le taux de métabolisme régional du glucose dans le cerveau et par unité de temps ($rCMR_{glu}$).

Le modèle est composé de trois compartiments :

1. plasmatique artériel (modélisant le FDG dans les artères),
2. tissulaire précurseur (pour le FDG libre dans les tissus),
3. tissulaire (pour le FDG phosphorylé dans les tissus sous forme de FDG-6-P).

Quand la concentration artérielle est supposée connue, le compartiment plasmatique artériel n'est pas considéré comme compartiment dans la mesure où sa concentration

d'activité n'est pas à prédire par le modèle. Ce modèle compartimental est alors classé comme modèle à deux compartiments.

Pour obtenir les équations différentielles, l'hypothèse qui est faite est que le flux net dans chaque compartiment est donné par la somme des entrées à laquelle on soustrait la somme des sorties. On désigne par :

- C_a la quantité de traceurs dans le compartiment sanguin artériel (concentration de FDG disponible dans les artères),
- C_1 la concentration dans le compartiment tissulaire précurseur (concentration du FDG libre dans le tissu) et
- C_2 la concentration dans le compartiment tissulaire (concentration de FDG métabolisé).

On considère les constantes suivantes :

- K_1 décrivant le passage par le flux artériel du FDG du compartiment sanguin vers les tissus,
- k_2 pour le transfert inverse, par le reflux veineux, des tissus vers le sang et
- k_3 pour le taux de métabolisation à l'intérieur des tissus. Il caractérise la phosphorylation du FDG en FDG-6-P.

Si la durée de l'examen n'excède pas 1h, la déphosphorylation du FDG-6-P n'est pas observée, ce qui correspond à $k_4 = 0$.

La fonction $C_a(t)$ décrit l'évolution de la concentration de traceurs dans le plasma artériel et qui est susceptible de rentrer dans le tissu cérébral. Elle est appelée fonction d'entrée artérielle et est supposée connue. Elle peut être obtenue à partir de prélèvements sanguins lors de l'examen clinique ou estimée par des mesures sur les images reconstruites (par exemple, par la moyenne de l'activité dans une région ne contenant que du sang artériel). On peut alors décrire l'évolution temporelle de l'activité dans les compartiments tissulaire précurseur et tissulaire par le système d'équations suivant :

$$\begin{aligned}\frac{\mathrm{d}C_1}{\mathrm{d}t} &= K_1 C_a(t) - (k_2 + k_3) C_1(t) + k_4 C_2(t) \\ \frac{\mathrm{d}C_2}{\mathrm{d}t} &= k_3 C_1(t) - k_4 C_2(t).\end{aligned}$$

Cependant, les compartiments tissulaire précurseur et tissulaire sont indistinguables à partir des mesures TEP. L'activité mesurée dans les tissus représente donc la somme de l'activité dans les deux compartiments [1], i.e $C(t) = C_1(t) + C_2(t)$. En utilisant l'hypothèse d'irréversibilité ($k_4 = 0$), la solution est donnée par [VB03, chapitre 6] :

$$C(t) = C_a(t) \otimes \left(\frac{K_1 k_2}{k_2 k_3} \exp(-(k_2 + k_3)t) + \frac{K_1 k_3}{k_2 k_3} \right) \tag{5.1}$$

où $\otimes$ désigne l'opérateur de convolution.

5.3 Estimation des paramètres

Le modèle décrit par l'équation 5.1 est un *modèle direct* : sachant la fonction d'entrée $C_a(t)$, la configuration du modèle et ses constantes d'échange, on peut prédire l'évolution

1. En réalité, l'activité mesurée en TEP comprend aussi une partie de la fraction vasculaire $C_a(t)$.

de la concentration d'activité dans les tissus. Cependant, les seuls paramètres mesurables *in vivo* sont la cinétique artérielle $C_a(t)$ et la concentration $C(t)$ mesurée par l'image TEP. L'estimation des paramètres physiologiques est donc un *problème inverse* : sachant les mesures de l'activité dans les tissus, la fonction d'entrée et une configuration du modèle, on cherche à produire des estimées des constantes de transfert. En général, on ne peut pas estimer de façon robuste tous les paramètres du modèle. On estime plutôt des combinaisons de ces paramètres (macro-paramètres). Deux types de méthodes sont généralement utilisées [GGT+02] :

1. les méthodes guidées par le modèle (« model-driven ») : elles nécessitent un modèle compartimental pré-spécifié et ajustent les données mesurées aux prédictions du modèle en utilisant par exemple la méthode des moindres carrés,
2. les méthodes guidées par les données (« data-driven ») : elles n'utilisent pas de modèle compartimental *a priori*. Elles sont de trois types : l'analyse graphique, l'analyse spectrale et la poursuite de base.

5.3.1 Les moindres carrés

Quand il y'a beaucoup de données collectées, l'une des méthodes les plus utilisées en pratique est celle des moindres carrées. Se basant sur modèle compartimental, les paramètres sont obtenus en ajustant les mesures d'activité dans tissus à celles prédites par le modèle. Pour cela, on minimise le critère suivant

$$\sum_{i=1}^{F}(C_i - C(t_i))^2$$

où F est le nombre de mesures ; $C_i, i = 1, \ldots, F$, sont les mesures d'activité dans le tissu aux temps t_i. Elles sont obtenues en subdivisant l'acquisition en intervalles temporels de durée $t_1, \ldots, t_F$ appelés frames et C_i est la mesure de la concentration moyenne d'activité dans un voxel ou une région d'intérêt pendant la durée t_i. Enfin, $C(t_i)$ est la prédiction par le modèle de l'activité dans le tissu au temps t_i.
Du fait de la variabilité du bruit dans les mesures, on préfère la méthode des moindres carrés pondérés où le terme à minimiser est donné par

$$\sum_{i=1}^{F} w_i(C_i - C(t_i))^2$$

où w_i désigne le poids associé à la mesure i.

Les méthodes guidées par les données ne nécessitent pas de spécifier la structure du modèle compartimental et sont plus généralement applicables.

5.3.2 Analyse graphique

Les méthodes graphiques [PBF+83], [Gje82] consistent en une transformation des données mesurées. L'analyse graphique la plus utilisée lorsque $k_4 = 0$ est celle de Gjedde-Patlak. Nous allons la décrire ici à dans le cas spécifique du modèle compartimental pour

le FDG. Utilisée dans le cas du FDG, elle permet d'estimer la constante de captation du glucose dans chaque élement de voxel ou dans une ROI.

La méthode graphique de Gjedde-Patlak décrit le comportement du modèle FDG décrit à la Figure 5.1 lorsque le FDG libre a atteint l'état d'équilibre de telle sorte que le ratio entre les concentrations $\frac{C(t)}{C_a}$ soit linéaire. Pour cela, admettons que la fonction d'entrée ait atteint l'équilibre et devienne une constante (notée (C_a)). Dans ce cas, l'équation 5.1 s'écrit

$$C(t) = C_a \left(\frac{K_1 k_2}{(k_2+k_3)^2} \left(1 - \exp(-(k_2+k_3)t\right) + \frac{K_1 k_3}{k_2+k_3} t \right). \tag{5.2}$$

Après un certain temps d'examen t^* où le terme exponentiel dans l'équation 5.2 est jugé suffisamment petit, le ratio entre les concentrations d'activité dans le tissu et le sang est

$$\text{pour tout } t > 0, \quad \frac{C(t)}{C_a} = \frac{K_1 k_2}{(k_2+k_3)^2} + K_i t \tag{5.3}$$

où

$$K_i = \frac{K_1 k_2}{k_2 + k_3}.$$

Si la fonction d'entrée n'est pas constante (ce qui est le cas), le temps t dans l'équation 5.3 est remplacé par $\frac{\int_0^t C_a(s)\mathrm{d}s}{C_a(t)}$. Ce terme est appelé « stretched time » ou encore « temps normalisé ».

Ainsi, si on trace le volume apparent de distribution $\frac{C(t)}{C_a(t)}$ par rapport au « stretched time », on obtient, à partir d'un certain temps, un tracé linéaire de pente K_i. En pratique, $C(\cdot)$ est donnée par la TAC des mesures $\{C_i, i = 1, \ldots, F\}$ où F est le nombre de frames. C_a est la fonction d'entrée artérielle obtenue par prélèvements sanguins ou estimée. Le taux de métabolisme cérébral régional du glucose (noté rCMRglu pour regional Cerebral Metabolic Rate of GLUcose) dans un élément de volume homogène est donné par

$$\text{rCMR}_{glu} = K_i \frac{C_p}{LC} \tag{5.4}$$

où C_p désigne la concentration de glucose dans le plasma et LC (pour « lumped constant ») est une constante qui exprime le rapport entre la phosphorylation du FDG et celle du glucose. Elle vaudrait un si le FDG était phosphorilé de la même façon que le glucose une fois dans la cellule. Le taux rCMR_{glu} exprime la part du glucose qui est métabolisé en une minute et constitue le paramètre d'intérêt pour le médecin.

La méthode graphique est une méthode simple, permettant d'estimer la constante K_i mais pas les constantes individuelles. Par ailleurs, elle repose sur une hypothèse de linéarité qui n'est pas toujours vérifiée. Ce qui peut biaiser l'estimation, surtout lorsque le RSB diminue.

5.3.3 Analyse spectrale

La méthode d'analyse spectrale de Cunningham et Jones [CJ93] calcule les paramètres d'intérêt dans le modèle, vus comme une fonction de la réponse impulsionnelle (IRF) [2]. Pour cela, les TAC sont modélisées par une combinaison linéaire de fonctions

2. La réponse impulsionnelle du tissu notée IRF (pour Impulse Response Function) décrit l'évolution de l'activité si la cinétique d'entrée était un bolus d'amplitude 1.

de base. Ces dernières sont obtenues par la convolution de la cinétique d'entrée avec des exponentielles décroissantes et sont choisies de telle sorte à couvrir le spectre des cinétiques possibles. Puis, les poids sont estimés par adéquation aux données d'activité mesurées sur les images, utilisant la méthode des moindres carrés .

L'analyse spectrale repose sur la caractéristique générale suivante du système. L'évolution de l'activité dans le tissu peut être décrite par la convolution entre la cinétique d'entrée et la réponse impulsionnelle du tissu. De façon générale, l'IRF $h(t)$ est donnée par une somme d'exponentielles décroissantes (une exponentielle par compartiment tissulaire) et s'écrit

$$h(t) = \sum_{i=1}^{n} \phi_i \exp(-\theta_i t)$$

où n est le nombre de compartiments dans le tissu cible, ϕ_i et θ_i sont des fonctions algébriques des constantes de transfert.
Dans un modèle compartimental où la cinétique d'entrée est la cinétique artérielle, on peut décrire la cinétique du tissu par

$$C(t) = C_a(t) \otimes h(t).$$

Par exemple, dans le modèle FDG pour $k_4 = 0$, on a $n = 2$ et l'évolution de l'activité dans le tissu cérébral donnée dans l'équation 5.1 s'écrit

$$C(t) = C_a(t) \otimes h(K_1, k_2, k_3, t) \tag{5.5}$$

où

$$h(K_1, k_2, k_3, t) = K_1 \exp\left(-(k_2 + k_3)t\right) + \frac{K_1 k_3}{k_2 + k_3}\left(1 - \exp\left(-(k_2 + k_3)t\right)\right).$$

Partant de cela, l'analyse spectrale considère la décomposition suivante :

$$C(t) = \sum_{j=1}^{N} \phi_j \psi_j(t)$$

où $\phi_1, \ldots, \phi_N$ sont des pondérations et $\psi_1, \ldots, \psi_N$ une base de fonctions fixées (dictionnaire). Cette base est donnée par la convolution entre la cinétique d'entrée (on considère ici qu'il s'agit de la cinétique artérielle) avec des exponentielles décroissantes,

$$\psi_j(t) = \exp(-\theta_j t) \otimes C_a(t).$$

La base est pré-calculée en choisissant N valeurs pour les termes de clairance exponentielle θ_j parmi la gamme de valeurs physiologiques possibles. Dans ce cas, les inconnues du système sont donc les paramètres ϕ_j. Si on désigne par C_t la concentration d'activité mesurée dans le tissu au temps t et par $\boldsymbol{C}$ le vecteur contenant l'ensemble de ces mesures, l'estimation des paramètres ϕ_j s'obtient en minimisant le critère

$$\|\boldsymbol{C} - \boldsymbol{\psi}\boldsymbol{\phi}\|_2^2.$$

En pratique (routine clinique), la valeur mesurée dans l'image correspondante à la frame numérotée k représente l'activité moyenne du tissu sur la durée considérée, c'est-à-dire que pour tout $k = 1, \ldots, F$ (où F est le nombre de frames),

$$C_k = \frac{1}{t_k^2 - t_k^1} \int_{t_k^1}^{t_k^2} C(t)\mathrm{d}t$$

avec t_k^1 et t_k^2 désignant respectivement les temps de début et de fin de la $k^{\text{ème}}$ frame temporelle. Le vecteur des observations est $\boldsymbol{C} = (C_1, \ldots, C_F)^T$. La matrice des cinétiques de base doit alors être en adéquation avec les données mesurée. Elle s'écrit

$$\boldsymbol{\Psi} = \begin{pmatrix} \psi_{11} & \psi_{12} & \cdots & \psi_{1N} \\ \psi_{2,1} & \psi_{2,2} & \cdots & \psi_{2,F} \\ \vdots & \vdots & \ddots & \vdots \\ \psi_{F1} & \psi_{F2} & \cdots & \psi_{FN} \end{pmatrix}$$

où pour tout $j = 1, \ldots, N$ et $k = 1, \ldots, F$, les ψ_{jk} sont calculés de la façon suivante

$$\psi_{jk} = \frac{1}{t_k^2 - t_k^1} \int_{t_k^1}^{t_k^2} \exp(-\theta_j t) \otimes C_a(t) \mathrm{d}t.$$

Le nombre N de fonctions de base du dictionnaire doit être choisi assez grand pour contenir toutes les cinétiques possibles. Comme N est supérieur au nombre F de frames temporelles, le système est sous-déterminé et présente une infinité de solutions. La stabilisation de l'estimation nécessite donc d'ajouter un terme de régularisation pour contraindre l'espace des solutions possibles. Dans la méthode spectrale, la régularisation se fait par l'introduction d'une contrainte de positivité sur les pondérations ϕ_j. La méthode utilise l'algorithme des moindres carrés non-négatifs (NNLS) pour minimiser le critère

$$\min_{\phi_j \geq 0} \|\boldsymbol{C} - \boldsymbol{\psi}\boldsymbol{\phi}\|_2^2.$$

Pour tenir compte de la variabilité temporelle dans la variance des mesures, on utilise les moindres carrés pondérés et le critère à minimiser est alors

$$\min_{\phi_j \geq 0} \|\boldsymbol{W}^{\frac{1}{2}}(\boldsymbol{C} - \boldsymbol{\psi}\boldsymbol{\phi})\|_2^2$$

où $\boldsymbol{W}$ est une matrice $F \times F$, souvent diagonale et contenant les termes de pondération. Les termes diagonaux de $\boldsymbol{W}$ peuvent être obtenus en prenant par exemple pour $k = 1, \ldots, F$

$$w_k = \frac{(t_k^2 - t_k^1)^2}{n_k}$$

où n_k est le nombre de comptages dans la frame k [GLHC97]. Après résolution, on obtient un petit nombre de valeurs de ϕ_j estimées qui peuvent être utilisées, avec les θ_j correspondantes, pour estimer l'IRF et par là les paramètres.

5.3.4 La poursuite de base L_1

La méthode de débruitage par poursuite de base de Gunn *et al.* [GGT+02] est basée sur le même principe que l'analyse spectrale c'est-à-dire en un ajustement non-linéaire du modèle temporel de l'activité aux données mesurées. Elle consiste aussi en une décomposition de la courbe d'activité sur une base pré-fixée de fonctions (dictionnaire) obtenues par convolution de la cinétique d'entrée avec des exponentielles décroissantes. La différence avec l'analyse spectrale réside dans le terme de régularisation choisi. La

poursuite de base utilise une pénalité par une norme L_1 pour favoriser la parcimonie sur les coefficients estimés. Le critère à minimiser est

$$\min_{\phi_j \geq 0} \|\boldsymbol{W}^{\frac{1}{2}}(\boldsymbol{y} - \boldsymbol{\psi}\boldsymbol{\phi})\|_2^2 + \mu\|\boldsymbol{\phi}\|_1$$

où μ est un paramètre de régularisation fixé ou à estimer. Il contrôle le compromis entre erreurs d'approximation et parcimonie du modèle. Si la valeur de μ est trop petite, les données sont sur-estimées et si elle est trop grande, les données sont sous-estimées.

L'analyse spectrale et la poursuite de base ne nécessitent pas de description *a priori* des traceurs et sont, en principe, applicables à plusieurs types de radiotraceurs. Elles permettent de reconstruire des images paramétriques pour les macro-paramètres d'intérêt. Les images paramétriques sont des représentations de la distribution spatiale d'un paramètre (la valeur dans un pixel/voxel correspond à la valeur d'un paramètre issu du modèle). Cependant, ce sont des modèles d'ajustement des prédictions aux données mesurées (les courbes activité-temps notées TAC pour Time Activity Curve). Elles nécessitent donc que les TAC soient peu bruitées. Par ailleurs, l'estimation robuste des paramètres nécessite beaucoup de données et donc un échantillonnage temporel fin. Ce qui est contradictoire en TEP. Un autre problème de ces méthodes est qu'elles nécessitent que le dictionnaire des fonctions puisse couvrir le spectre des cinétiques possibles. Le choix du nombre de fonctions de base est un choix difficile, guidé par un compromis entre précision et temps de calcul. Un autre problème dans la poursuite de base L_1 est le choix du paramètre μ. Dans [GGT+02], μ est choisi de sorte à minimiser l'erreur du LOOCV (leave-one-out cross-validation) sur un grand ensemble de valeurs.

5.4 État de l'art de la reconstruction dynamique en TEP

Comme nous l'avons déjà vu, en routine clinique la reconstruction dynamique en TEP se fait dans un cadre statique, où la variable temps est ignorée. Cela s'appuie sur la subdivision du temps d'acquisition total en tranches temporelles appelées frames. Afin de s'adapter aux variations de la fixation du traceur, ces frames sont de durée variable et sont plus courtes au début de l'acquisition (quelques secondes) qu'à la fin (dizaines de minutes). Les données dans chaque intervalle temporel sont reconstruites indépendamment les unes des autres par les techniques ML-EM ou FBP vues dans le chapitre 4. Cela résulte en une série d'images 3D, une par frame. Bien que permettant un suivi dynamique de la distribution du traceur, cette approche n'est de loin pas optimale. En effet, elle présente l'inconvénient que seule une partie des observations est utilisée pour reconstruire une image. Cette perte statistique induit fatalement une détérioration de la qualité de la reconstruction d'autant plus marquée que la tranche temporelle est courte. Cette approche nécessite donc un compromis entre des protocoles d'acquisition de durée plus longue (ce qui augmente la statistique de comptage mais dégrade la résolution temporelle) et des courts intervalles temporels (améliorant la résolution temporelle mais aux dépens de la qualité de l'image). Par ailleurs, l'imagerie TEP est par essence dynamique dans la mesure où la distribution du radiopharmaceutique injecté dans l'organisme évolue dans le temps en fonction des interactions moléculaires avec les tissus.

Ainsi, d'autres stratégies ont été proposées dans la littérature et elles permettent d'in-

corporer la modélisation des cinétiques dans le processus de reconstruction. Deux types d'approches existent et consistent soit à estimer les cartes paramétriques directement à partir des observations, soit à procéder d'abord à une reconstruction spatio-temporelle puis à estimer par la suite la carte paramétrique à partir de l'activité mesurée. Afin de mieux comprendre leurs limites, nous allons présenter les grandes lignes de chacune d'entre-elles.

5.4.1 Reconstruction directe de cartes paramétriques

Plutôt que de traiter la reconstruction et la modélisation des cinétiques en deux tâches séparées, certaines approches proposent de les traiter conjointement et de reconstruire directement les cartes paramétriques à partir des données, sans une reconstruction préalable de l'image. Cela consiste en l'inversion du modèle compartimental : partant des données mesurées on reconstruit l'image paramétrique.

Méthode spectrale

Meikle *et al.* [MMC+98] appliquent la méthode d'analyse spectrale de Gunn *et al.* (décrite précédemment) directement aux données de projections dans un cadre 2D+t. Pour cela, ils modélisent d'abord les données en utilisant la transformée de Radon dépendante du temps de la façon suivante

$$y_i(t) = \sum_{j \in J} a_{ij} f_j(t)$$

où $y_i(t)$ est le nombre d'évènements dans le bin de sinogramme i ($\approx$ la LOR i) au temps t, J l'ensemble des pixels contribuant à la LOR i, a_{ij} la probabilité qu'une paire de photons émis au pixel j soit détectée dans la LOR i et $f_j(t)$ la TAC correspondant au pixel j. Puis, l'activité dans le pixel j est modélisée par convolution de la fonction d'entrée avec des exponentielles décroissantes (tout comme dans la méthode spectrale) et on a alors

$$y_i(t) = \sum_{j \in J} \left(a_{ij} \sum_{l=1}^{N} \phi_{jl} \exp(-\theta_l t) \otimes C_a(t) \right),$$

où N est le nombre de fonctions de base du dictionnaire. On peut donc appliquer les moindres carrés pondérés pour minimiser pour chaque LOR i, la différences entre les projections mesurées et les $y_i(t)$ du modèle. Les éléments de la matrice de pondération $\boldsymbol{W}$ sont donnés par l'inverse de la variance Poisson des observations. En effectuant l'adéquation du modèle aux données de cette sorte, les paramètres estimés fournissent une estimée de l'IRF pour la LOR en question. Cela peut être vu comme l'IRF du mélange de tissus le long de la LOR i,

$$\hat{H}_i(t) = \sum_{j \in J} a_{ij} \hat{h}_j(t).$$

Cette IRF projetée correspond à une forme générale de la transformée de Radon et peut être inversée par la méthode FBP ou EM afin de reconstruire les cartes paramétriques de l'IRF $\hat{h}$ en chaque j et t.

Cette méthode produit des résultats assez similaires que la méthode d'analyse spectrale appliquée à l'image reconstruite.

Approche MAP

Kamasak *et al.* [KBMS05] proposent aussi d'estimer les cartes paramétriques directement à partir des données du sinogramme mais contrairement à l'approche précédente, ils n'utilisent pas de fonctions temporelles pré-fixées. Ils proposent des termes de pénalité quadratique pour contraindre spatialement les paramètres. L'approche proposée a été développée pour un modèle compartimental avec deux compartiments tissulaires, de paramètres (k_1, k_2, k_3, k_4) (par exemple le modèle du FDG présenté précédemment avec $k_4 \neq 0$).

Pour construire des images paramétriques, les paramètres sont exprimés au niveau des voxels $(k_{1j}, k_{2j}, k_{3j}, k_{4j}$, pour $j = 1, \ldots, J)$ où J désigne le nombre de voxels dans l'image. Ils procèdent d'abord à une re-paramétrisation du problème pour faciliter l'optimisation du problème et les nouveaux paramètres sont $\boldsymbol{\phi}_j = (a_j, b_j, c_j, d_j)$ obtenus par des combinaisons non linéaires des anciens paramètres. Le problème de l'estimation revient en celui d'une matrice $4 \times J$ notée $\boldsymbol{\Phi} = (\boldsymbol{\phi}_1, \ldots, \boldsymbol{\phi}_J)$ contenant l'ensemble des paramètres pour tous les voxels de l'image.

Utilisant le modèle compartimental, ils expriment la concentration d'activité dans le voxel j au temps t comme une fonction des paramètres a_j, b_j, c_j, d_j, notée $C(\boldsymbol{\phi}_j, t)$. Puis ils l'étendent à une fonction faisant correspondre l'image paramétrique $\boldsymbol{\Phi}$ à $\mathcal{C}(\boldsymbol{\Phi}) = (C(\boldsymbol{\phi}_1), \ldots, C(\boldsymbol{\phi}_J))^T$, une matrice $J \times F$ où F est le nombre de frames. Elle contient la concentration d'activité dans chaque voxel et à chaque temps. Si on note par $y_i(t_k)$ le nombre de mesures dans le bin de sinogramme i pour la frame k (ce sont des variables de Poisson indépendantes) et par $\mathcal{C}(\boldsymbol{\Phi}, t_k)$ la $k^{\text{ème}}$ colonne de $\mathcal{C}(\boldsymbol{\Phi})$ contenant l'activité pour tous les voxels au temps t_k, on a

$$\mathbb{E}\left(y_i(t_k) | \mathcal{C}(\boldsymbol{\Phi}, t_k)\right) = \sum_{j=1}^{J} A_{ij} C(\boldsymbol{\phi}_j, t_k) + r$$

que l'on peut écrire sous forme matricielle

$$\mathbb{E}(\boldsymbol{y} | \mathcal{C}(\boldsymbol{\Phi})) = \boldsymbol{A}\mathcal{C}(\boldsymbol{\Phi}) + \boldsymbol{r}$$

où $\boldsymbol{y}$ est le vecteur des mesures, $\boldsymbol{A}$ la matrice de projection et $\boldsymbol{r}$ le vecteur des coïncidences fortuites. On peut noter la ressemblance de cette équation avec l'équation 4.2. Pour calculer les paramètres $\boldsymbol{\Phi}$, la fonction de coût est donnée par un critère MAP et consiste en l'antilog de la vraisemblance $L(\boldsymbol{y}|\boldsymbol{\Phi})$ auquel ils rajoutent un terme *a priori* choisi comme étant un champ de Markov Gaussien.

Ils montrent sur des simulations d'examens TEP 4D que les erreurs quadratiques moyennes pour l'estimation des quatre paramètres sont inférieures à celles obtenues en passant avec une étape supplémentaire de reconstruction consistant d'abord en une reconstruction (par FBP ou MAP) puis en l'estimation de ces paramètres.

La difficulté dans cette approche est que l'algorithme itératif utilisé pour minimiser la fonction de coût est un algorithme de descente par coordonnées et ils ont recours à

une série de Taylor du second ordre pour approximer les variations de la fonction de coût. Par ailleurs, la méthode utilise aussi une optimisation imbriquée pour réduire le temps de calcul et assurer la robustesse de la convergence.

L'avantage d'une estimation directe des paramètres d'intérêt est de réduire théoriquement le temps de calcul d'un facteur n_f/n_p où n_f est le nombre de frames temporelles et n_p le nombre de paramètres cinétiques dans le modèle compartimental. L'avantage d'utiliser l'approche image pour estimer les cartes paramétriques est que la régularisation spatiale permet de réduire la grande variance dans les images paramétriques.

Nous allons maintenant aborder les reconstructions permettant d'exploiter toutes les données spatio-temporelles pour reconstruire des images dynamiques.

5.4.2 Reconstructions spatio-temporelles

Au lieu de subdiviser l'acquisition en tranches temporelles indépendamment reconstruites, des techniques ont été développées pour reconstruire directement la distribution spatio-temporelle d'activité. Cela permet d'exploiter la corrélation temporelle entre les frames et d'améliorer le rapport signal-sur-bruit des images reconstruites. Pour cela, l'activité est décomposée sur une base de fonctions spatiales (typiquement les voxels) et temporelles. L'idée est, à partir d'un certain nombre de fonctions de base temporelles, de reconstituer les cinétiques de l'image (les courbes d'activité temporelle (TAC)) par combinaison linéaire de ces fonctions de base. La courbe d'activité s'exprime de la façon suivante :

$$f(j,t) = \sum_{n=1}^{N} w_{jn} b_n(t)$$

où $f(j,t)$ représente l'activité dans le voxel j et au temps t ; w_{jn} la pondération spatiale associée à la fonction temporelle n dans le voxel j, $b_n(t)$ la valeur de la fonction temporelle de base n évaluée au temps t . Dans chaque voxel, les contributions des cinétiques de base sont données par les pondérations w_{jn}. Plusieurs méthodes utilisant cette approche ont été développées et elles diffèrent principalement de par les fonctions temporelles utilisées.

Base temporelle prédéfinie et fixée

Les fonctions de base temporelles peuvent être prédéfinies et fixées. Dans ce cas, le problème de la reconstruction revient à estimer les pondérations spatiales (sous une contrainte de positivité). Nous allons résumer les approches de Matthews *et al.* [MBPC97] et Nichols *et al.* [NQAL02].

1. Matthews *et al.* [MBPC97] utilisent une base de fonctions temporelles expérimentales obtenues par une décomposition en valeurs singulières d'images reconstruites par rétroprojection filtrée ou des données issues de populations. Ces fonctions étant ainsi fixées, ils calculent les pondérations spatiales associées en utilisant l'algorithme EM.

   L'idée est la suivante. L'activité dans le pixel j de la $k^{\text{ème}}$ frame temporelle (de

temps t_k) s'exprime par :

$$f(j,t_k) = \sum_{n=1}^{N} w_{jn} b_n(t_k)$$

où $b_n(t_k)$ est la valeur de la cinétique n évaluée au temps t_k de la frame k et w_{jn} son poids dans le pixel j.

Pour décrire plus en détails cette approche, on définit par $\boldsymbol{W} = \{W_{jn}, j = 1,\dots,J, n = 1,\dots,N\}$ où W_{jn} représente le nombre de comptages émis par le pixel j, dû à la cinétique n et détectés pendant toute la durée du scan. Comme dans ML-EM, ils supposent que

$$W_{jn} \sim \text{Poisson}(\psi_{jn})$$

avec $\psi_{jn} = \mathbb{E}(W_{jn})$.

Si on désigne par

- $B_{t_k n}$ la probabilité qu'une émission ait lieu pendant la $k^{\text{ème}}$ frame temporelle sachant qu'elle provient de la cinétique n. Cette probabilité est obtenue en calculant l'aire sous la courbe temporelle n pour la durée de la frame k, divisée par l'aire totale de la courbe (sur toute la durée du scan). Cela équivaut à définir $B_{t_k n} = \frac{b_{t_k n}}{\sum_k b_{t_k n}}$
- A_{ij} la probabilité qu'une émission soit détectée par la LOR i sachant qu'elle a eu lieu au pixel j.

Si les mesures sont l'ensemble des y_{it_k} (nombre d'évènements détectés dans la LOR i pendant la durée t_k) alors $A_{ij}B_{t_k n}$ représente la probabilité de détecter un évènement dans la LOR i pendant la frame k sachant que l'émission a eu lieu dans le pixel j et la cinétique n. L'algorithme ML-EM peut alors être utilisé (cf. section 4.2.3) pour estimer les ψ_{jn} en substituant f par ψ, y_i par y_{it_k} et a_{ij} par $A_{ij}B_{t_k n}$.

La méthode nécessite donc de choisir les fonctions de base temporelles $b_{t_k n}$ (et par là les probabilités $B_{t_k n}$). Matthews *et al* l'obtiennent

- par décomposition en valeurs singulières SVD sur des images préalablement reconstruites par FBP : l'idée est d'extraire les cinétiques extrémales telles que toutes les autres s'obtiennent par combinaison linéaire de ces cinétiques.
- en utilisant des TAC issues d'une base de données collectées sur des reconstructions provenant d'autres patients et où les cinétiques sont extraites en considérant les concentrations moyennes d'activité dans les régions d'intérêt.

Dans les résultats qu'ils produisent, on observe une amélioration visuelle dans les images paramétriques reconstruites avec cependant des biais lors d'une expérience avec deux isotopes. Par ailleurs, le succès de cette approche est complètement dépendante des fonctions dynamiques choisies.

2. Nichols *et al.* [NQAL02] ont recours à une base de fonctions B-splines cubiques pour modéliser les courbes temporelles et seules les pondérations sont à estimer.

   Ils utilisent le stockage des données en mode liste et modélisent la fonction d'intensité de chaque voxel j par une combinaison linéaire de B-splines cubiques,

$$\eta(j,t) = \sum_{n=1}^{N} w_{jn} b_n(t) \tag{5.6}$$

avec $\eta(j,t) \geq 0$ pour tout t où $\eta(j,\cdot)$ est la fonction d'intensité du voxel j, w_{jn} la contribution de la $i^{\text{ème}}$ fonction de base sur le voxel j, b_n la $i^{\text{ème}}$ fonction de la base des B-splines.

Soit a_{ij} la probabilité de détecter sur la paire de détecteurs i une émission ayant eu lieu dans le voxel j. En supposant que les probabilités de détection sont indépendantes et invariantes dans le temps, alors la détection en coïncidence au niveau de la LOR i est un processus de Poisson non-homogène dont la fonction d'intensité s'exprime par

$$\lambda(i,t) = \sum_{j=1}^{J} a_{ij}\eta(j,t) \tag{5.7}$$

où $\eta(j,t)$ est donnée dans l'équation 5.6. Ils expriment alors la log-vraisemblance des données du comptage list-mode, d'intensité $\lambda(i,t)$, vue comme une fonction des poids $\boldsymbol{w} = \{w_{jn}, j = 1,\ldots,J, n = 1,\ldots,N\}$. Dans l'approche proposée, cette vraisemblance est modifiée par l'introduction de trois termes de pénalité. Puis un algorithme de gradient conjugué est utilisé pour l'optimisation de la fonction de coût.

Sur des examens en TEP cérébrale, les auteurs ont montré que cette approche réalise un meilleur compromis entre le biais et la variance que les reconstructions statiques multi-frames effectuées avec l'approche MAP. Cependant, cette méthode utilise plusieurs termes de pénalité pour imposer des contraintes de douceur et de positivité. En effet, le critère à maximiser est donné par

$$L(\boldsymbol{w}|D) - \alpha\rho(\boldsymbol{w}) - \beta\phi(\boldsymbol{w}) - \gamma\nu(\boldsymbol{w})$$

où

- $L(\boldsymbol{w}|D)$ est la vraisemblance des données,
- $\rho(\boldsymbol{w})$ est un terme de régularité temporelle, donné dans chaque voxel j par l'intégrale de la courbure de η_j au carré,
- $\phi(\boldsymbol{w})$ est un terme de régularité spatiale et choisie comme étant une contraint quadratique
- $\nu(\boldsymbol{w})$ terme de pénalité de la négativité.
- α, β et γ sont des paramètres de réglage.

Cela pose aussi le problème du choix des paramètres de réglage. Dans [NQAL02], les paramètres α et β sont ajustés de telle sorte à obtenir la même résolution que dans les examens cliniques.

Base temporelle ajustée pendant la reconstruction

Plutôt que de fixer la base temporelle, cette dernière peut être ajustée pendant la reconstruction. Dans ce cas on procède à une estimation conjointe des fonctions de base temporelles et des pondérations spatiales. Reader *et al.* [RSC+06] adoptent ce principe et cherchent la base fonctionnelle qui ajuste au mieux le jeu de données. Pour cela, l'algorithme EM en 4D est utilisé pour estimer alternativement les fonctions de base temporelles et leurs pondérations.

Dans un premier temps, la distribution spatio-temporelle est décomposée suivant

$$f(j,t) = \sum_{n=1}^{N} w_{jn} b_n(t)$$

où b_n désigne les fonctions temporelles, N le nombre de ces fonctions temporelles, w_{jn} les pondérations pour $j = 1, \ldots, J$ (où J est le nombre de voxels) et $t = 1, \ldots, F$ (F est le nombre d'indices temporels). La relation peut s'écrire sous forme matricielle

$$\boldsymbol{f} = \boldsymbol{B}\boldsymbol{w}$$

où $\boldsymbol{f}$ est un vecteur de taille JF contenant la concentration d'activité dans chaque voxel du champ de vue et à chaque temps, $\boldsymbol{B}$ est une matrice de taille $JF \times JN$ contenant les fonctions de base temporelles et $\boldsymbol{w}$ un vecteur de taille JN contenant les pondérations de chacune des fonctions temporelles dans chacun des voxels. Dans cette première étape, les fonctions spatio-temporelles sont fixées et les pondérations spatiales sont à estimer. En considérant les projections spatio-temporelles, on peut écrire

$$\bar{\boldsymbol{y}} = \boldsymbol{A}\boldsymbol{f} = \boldsymbol{A}\boldsymbol{B}\boldsymbol{w}$$

où $\boldsymbol{A}$ est la matrice de projection spatio-temporelle, de taille $IF \times JF$ (où I est le nombre de LOR) contenant les probabilités qu'une émission provenant du voxel j soit détectée par la LOR i pour chaque temps t. Elle est considérée comme invariante dans le temps et consiste en F réplicats de la matrice statique. En modélisant les données y_i par des variables de Poisson de paramètres $\bar{y}_i$, l'algorithme ML-EM (cf. section 4.2.3) peut être utilisé pour estimer les pondérations en considérant comme matrice système $\boldsymbol{A}\boldsymbol{B}$. La mise à jour des poids s'effectue alors suivant :

$$\boldsymbol{w}^{(k+1)} = \frac{\boldsymbol{w}^{(k)}}{\boldsymbol{B}^T\boldsymbol{A}^T\boldsymbol{I}}\boldsymbol{B}^T\boldsymbol{A}^T\frac{\boldsymbol{y}}{\boldsymbol{A}\boldsymbol{B}\boldsymbol{w}^{(k)}}.$$

La deuxième étape de l'algorithme consiste à mettre à jour les fonctions de base temporelles en considérant les poids $\boldsymbol{w}$ fixés. Pour cela, ils procèdent à un réarrangement entre les coefficients connus et inconnus. La décomposition de $\boldsymbol{f}$ considérée est la suivante

$$\boldsymbol{f} = \boldsymbol{M}\boldsymbol{t}$$

où $\boldsymbol{M}$ est une matrice $JF \times FC$ contenant les pondérations spatiales et $\boldsymbol{t}$ un vecteur de taille FN contenant les N fonctions temporelles pour chaque temps. Les projections s'écrivent alors

$$\bar{\boldsymbol{y}} = \boldsymbol{A}\boldsymbol{M}\boldsymbol{t}.$$

De la même manière que pour les pondérations, l'algorithme ML-EM est utilisé avec cette fois-ci comme matrice-système $\boldsymbol{A}\boldsymbol{M}$ et la mise à jour des fonctions temporelles s'effectue suivant,

$$\boldsymbol{t}^{(k+1)} = \frac{\boldsymbol{t}^{(k)}}{\boldsymbol{M}^T\boldsymbol{A}^T\boldsymbol{I}}\boldsymbol{M}^T\boldsymbol{A}^T\frac{\boldsymbol{y}}{\boldsymbol{A}\boldsymbol{M}\boldsymbol{t}^{(k)}}.$$

Cette méthode de reconstruction utilise toute la dynamique temporelle disponible et permet d'améliorer le rapport signal-sur-bruit dans les images reconstruites, en comparaison avec des reconstructions statiques multi-frames réalisées avec l'algorithme ML-EM. L'inconvénient de cette technique réside dans le fait que la régularisation du problème dépend du nombre de fonctions choisies. En effet, cette approche requiert de spécifier à l'avance le nombre de fonctions temporelles. Ce choix est critique car ce nombre doit être assez grand pour pouvoir représenter toutes les cinétiques présentes mais pas trop au risque de dégrader la reconstruction. Par exemple sur des données bruitées, il a été noté que les cinétiques présentes ne sont bien représentées que si le nombre de fonctions temporelles est inférieur au nombre de frames.

5.4.3 Synthèse

Nous avons présenté un état de l'art non-exhaustif, mais résumant les points principaux des méthodes de reconstruction d'image en TEP dynamique. Dans cette synthèse, nous allons résumer ce qui nous paraît comme limitations dans ces techniques. Puis, nous présenterons les principales caractéristiques de la nouvelle méthode de reconstruction proposée.

- En routine clinique, l'estimation des paramètres cinétiques s'appuie sur une subdivision du temps d'acquisition et procède à une première étape de reconstruction d'images 3D discrétisées. Puis, les paramètres sont estimés à partir de cette série d'images. Cette procédure n'est de loin pas optimale pour plusieurs raisons.
  - Ignorer la variable temps entraîne fatalement une perte statistique, ce qui nuit à la qualité des images reconstruites.
  - L'analyse compartimentale pour le calcul des paramètres physiologiques d'intérêt nécessite une segmentation préalable des images par un expert pour définir les régions d'intérêt (ROI). La plupart du temps, cela se fait à partir de l'image anatomique. Ensuite, la concentration d'activité moyenne est évaluée dans chaque ROI et pour chaque tranche temporelle. Cette méthode de tracé des ROI pose aussi problème car non seulement elle est subjective et fastidieuse (surtout en 3D), mais elle suppose aussi la localisation anatomique des régions fonctionnelles.

- Les reconstructions directes d'images spatio-temporelles qui ont été proposées dans la littérature permettent de diminuer le bruit en exploitant la corrélation temporelle dans l'évolution de l'activité dans les voxels. Mais ces méthodes présentent aussi un certain nombre de limitations.

  1. *Sélection de modèle :*
     Les cinétiques sont estimées à partir d'un nombre fixé et/ou fini de fonctions de base temporelles (dictionnaire). Le choix du nombre et de la forme des fonctions du dictionnaire est crucial pour le succès de la méthode de reconstruction. En effet, ces fonctions doivent être à même de décrire de façon adéquate toutes les dynamiques présentes dans le champ de vue du tomographe. Le nombre et/ou le type des fonctions de base temporelles constitue donc un paramètre de régularisation du problème. Leur choix est important mais tout aussi difficile, c'est le problème de la sélection de modèle dans les méthodes paramétriques. Par ailleurs, contraindre la base de fonctions à appartenir à une certaine famille paramétrique est limitant pour la flexibilité du modèle.

  2. *Discrétisation spatio-temporelle :*
     Ces techniques requièrent la discrétisation de tout l'espace-temps relatif à l'objet à reconstruire entraînant d'une part la manipulation et le stockage de gros volumes de données. D'autre part, la réduction significative de la dose injectée, donc une dégradation du rapport signal-sur-bruit dans les données, conduirait à préférer éviter toute discrétisation *a priori* qui induirait une perte statistique. En effet, la discrétisation de toutes les dimensions spatiale et temporelle, exhibe fréquemment un grand nombre de voxels non informatifs. De surcroît, les effets de volume partiel (EVP) que nous avons déjà

évoqués dans la reconstruction 3D et dûs en grande partie à la résolution spatiale limitée du tomographe, mais aussi à la discrétisation spatiale, peuvent conduire à des analyses biaisées. En effet, un voxel peut regrouper plusieurs tissus lesquels peuvent avoir des cinétiques différentes. La cinétique mesurée dans un voxel peut donc correspondre à un mélange de cinétiques issues de tissus différents. Le problème sera de retrouver les cinétiques pures, c'est-à-dire caractérisant une structure fonctionnelle spécifique, à partir des cinétiques observées.

- Dans ce contexte, l'approche que nous visons doit permettre de :
  1. substituer aux méthodes de reconstruction classiques, une reconstruction directe des volumes fonctionnels dans un espace spatio-temporel continu, s'affranchissant ainsi de toute discrétisation spatiale ou dans le temps.
  2. déterminer quantitativement et objectivement les volumes fonctionnels. Cette détermination est quantitative dans la mesure où l'on doit se prononcer sur l'incertitude associée à la reconstruction (ce qui est important en faible dose) et objective puisque le nombre, le positionnement, et l'extension des volumes fonctionnels ne sont pas définis au préalable.

  Nous proposons une technique originale de reconstruction cherchant à atteindre ces objectifs, tout en restant robuste lorsque le rapport signal-sur-bruit des données acquises est dégradé, permettant ainsi de diminuer la dose injectée au patient sans nuire à la qualité de l'interprétation de l'examen. Pour ce faire, nous abordons le problème de reconstruction 4D et d'extraction des volumes fonctionnels dans un cadre bayésien non paramétrique. Cela revient à appréhender la question sous la forme d'un problème inverse de classification, segmentation et estimation spatio-temporel où les nombres de composantes spatiales mais aussi temporelles doivent être inférées. En effet, l'identification des volumes fonctionnels nécessite de pouvoir identifier les régions spatiales présentant un fonctionnement cinétique similaire. L'abord de la reconstruction spatio-temporelle dans un cadre non paramétrique a pour conséquence que le nombre et la forme de ces cinétiques sont inconnues, en cela résident l'originalité et aussi la difficulté du problème.

5.5 Nouvelle méthode de reconstruction en TEP 4D

Pour la reconstruction spatio-temporelle, nous procédons comme dans le cadre statique 3D c'est-à-dire que la concentration de traceurs au cours du temps est considérée comme une distribution de probabilité spatio-temporelle dans $\mathbb{R}^4$.

5.5.1 Formulation du problème

On s'intéresse à l'estimation de la densité de la distribution spatio-temporelle d'émission. On suppose que les données sont stockées sous le format liste. Si on considère les données brutes non-corrigées et que l'on néglige le temps mort, le processus de comptage est un processus de Poisson non-homogène (non-homogène car la fonction d'intensité du processus varie dans le temps à cause de la décroissance radioactive). La

fonction d'intensité du processus de Poisson est proportionnelle à la densité d'émission spatio-temporelle. Nous nous intéressons à cette densité.

Dans le format liste, les données sont paramétrées par quatre paramètres spatiaux définissant la position de la LOR impliquée dans la détection et le temps d'arrivée des photons. Les données spatio-temporelles mesurées sont donc définies par cinq paramètres. En construisant une correspondance entre les paramètres des LOR, on note par y_i l'indice de la LOR correspondant à la $i^{\text{ème}}$ coïncidence observée et par t_i le temps d'arrivée enregistré. Le problème de reconstruction se formule ainsi dans le contexte bayésien non paramétrique

$$\begin{aligned} G &\sim \mathcal{G} \\ F(\cdot, t) &= \int_{\mathcal{X}} \mathcal{P}(\cdot|\boldsymbol{x})\, G(\mathrm{d}\boldsymbol{x}, t) \\ y_i, t_i|F &\overset{\text{iid}}{\sim} F, \text{ pour } i = 1, \ldots, n \end{aligned} \tag{5.8}$$

où $G(\cdot, \cdot)$ est une mesure de probabilité aléatoire distribuée suivant $\mathcal{G}$, définie sur $(\mathcal{X} \times \mathcal{T}, \sigma(\mathcal{X}) \otimes \sigma(\mathcal{T}))$ avec $\mathcal{X} \subset \mathbb{R}^3$, $\mathcal{T} \subset \mathbb{R}^+$; $G(\cdot, \cdot)$ est la distribution d'émission spatio-temporelle des évènements enregistrés qui est à estimer à partir des observations F-distribuées $(\mathbf{Y}, \mathbf{T}) = \{(y_1, t_1), \ldots, (y_n, t_n)\}$. Enfin, $\mathcal{P}(\cdot|\boldsymbol{x})$ est une distribution de probabilité connue, définie sur $(\mathcal{Y}, \sigma(\mathcal{Y}))$ et appelée distribution de la projection.

La nature incomplète des données induit des pertes d'évènements (cf. section 4.3.1). On définit la densité d'émission spatio-temporelle de tous les évènements par

$$G^{\dagger}(\boldsymbol{x}, t) \propto \frac{G(\boldsymbol{x}, t)}{\mathbb{P}(y^*|\boldsymbol{x})} \tag{5.9}$$

où $y^* = (y > 0)$ représente une donnée enregistrée, $\mathbb{P}(y^*|\boldsymbol{x}) = \sum_{l=1}^{L} \mathbb{P}(y = l|\boldsymbol{x})$ est la probabilité d'enregistrer par le système une émission localisée en $\boldsymbol{x}$, prenant en compte la géométrie du système et ses propriétés physiques. Comme en 3D, nous inférons sur la distribution aléatoire d'activité $G(\cdot, \cdot)$ correspondant aux évènements enregistrés.

On se concentre maintenant sur la modélisation des données.

Loi *a priori* sur la distribution d'activité

Dans l'approche bayésienne, on munit les inconnues du modèle de lois *a priori*. Dans le contexte bayésien non paramétrique, cette loi *a priori* porte directement sur la loi G et s'exprime par $G \sim \mathcal{G}$, où $\mathcal{G}$ est une distribution sur des distributions *i.e*, chaque tirage suivant $\mathcal{G}$ est une mesure de probabilité sur $\mathcal{X} \times \mathcal{T}$.

On rappelle que la loi *a priori* que nous avons utilisée pour la reconstruction statique en 3D est un mélange par processus de Pitman-Yor (PYM) de Gaussiennes et s'écrit

$$f_G(\boldsymbol{x}_i) = \sum_{k=1}^{\infty} w_k \mathcal{N}(\boldsymbol{x}_i|\boldsymbol{\theta}_k^*). \tag{5.10}$$

En 4D, la loi *a priori* pour la distribution aléatoire de l'activité est un mélange qui s'écrit sous la forme

$$f_G(\boldsymbol{x}_i, t_i) = \sum_{k=1}^{\infty} w_k C(\boldsymbol{x}_i, t_i|\boldsymbol{\lambda}_k) \tag{5.11}$$

où C est une densité de probabilité sur $\mathcal{X} \times \mathcal{T}$, paramétrée par $\boldsymbol{\lambda}$.
Pour l'application à l'imagerie cérébrale fonctionnelle, on considère la distribution spatio-temporelle d'activité comme étant séparable en espace et en temps pour chaque composante du mélange. Cette hypothèse est justifiée par le fait que les paramètres spatiaux ne varient pas dans le temps. Cela permet d'écrire le modèle 5.11 de la façon suivante

$$f_G(\boldsymbol{x}_i, t_i) = \sum_{k=1}^{\infty} w_k \, \mathcal{N}(\boldsymbol{x}_i | \widetilde{\boldsymbol{\theta}}_k) \widetilde{Q}_k(t_i) \tag{5.12}$$

où $\widetilde{Q}_k$ est une densité de probabilité modélisant l'évolution temporelle de l'activité dans la composante k. On peut d'ores et déjà remarquer qu'en marginalisant le modèle spatio-temporel 5.12 par rapport au temps, on retrouve le modèle utilisé pour la reconstruction spatiale (équation 5.10).
Toujours dans la même perspective qu'en 3D, nous suivons une approche bayésienne non paramétrique pour la modélisation des cinétiques, ceci afin de permettre la flexibilité et l'adéquation du modèle aux données. Étant données que les cinétiques sont des fonctions continues et à support compact dans $\mathbb{R}^+$ (troncature à droite) du fait de la durée limitée de l'examen , nous proposons de les modéliser par des arbres de Pólya définis à la section 2.5. On peut écrire le modèle 5.12 sous la forme hiérarchique suivante

$$\begin{aligned} \boldsymbol{x}_i, t_i | \boldsymbol{\theta}_i, Q_i &\overset{\text{ind}}{\sim} \mathcal{N}\left(\boldsymbol{x}_i | \boldsymbol{\theta}_i\right) \times Q_i(t_i) \\ \boldsymbol{\theta}_i, Q_i | H &\overset{\text{iid}}{\sim} H \\ H &\sim PY\left(\alpha, \beta, G_0 \times \ \mathrm{PT}\left(\mathcal{A}, Q_0\right)\right). \end{aligned} \tag{5.13}$$

H est distribuée suivant un processus de Pitman-Yor et on a

$$H(\cdot) = \sum_{k=1}^{\infty} w_k \, \delta_{\widetilde{\boldsymbol{\theta}}_k, \widetilde{Q}_k}(\cdot)$$

où

- $\mathbf{w} = w_1, w_2, \ldots \sim \mathrm{GEM}(\alpha, \beta)$,
- $\widetilde{\boldsymbol{\theta}} = \widetilde{\boldsymbol{\theta}}_1, \widetilde{\boldsymbol{\theta}}_2, \ldots$ sont les valeurs distinctes de $\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \ldots$, et sont $i.i.d$ suivant G_0 un modèle Normal-Inverse Wishart,
- $\widetilde{\mathbf{Q}} = \widetilde{Q}_1, \widetilde{Q}_2, \ldots$ sont les valeurs distinctes de $Q_1, Q_2, \ldots$ et sont $i.i.d$ suivant $\mathrm{PT}(\mathcal{A}, Q_0)$, un processus d'arbres de Pólya de paramètres $\mathcal{A}$ et de distribution de base Q_0.

La nature discrète de H implique que plusieurs $(\boldsymbol{\theta}_i, Q_i)$ seront identiques et donc, plusieurs observations $(\boldsymbol{x}_i, t_i)$ vont partager la même composante k, c'est-à-dire la même Gaussienne $\mathcal{N}(\cdot|\widetilde{\boldsymbol{\theta}}_k)$ et la même cinétique $\widetilde{Q}_k(\cdot)$. Cependant, puisque nous avons pour vocation d'identifier des volumes fonctionnels qui sont des zones spatiales ayant la même cinétique, cela voudrait dire qu'intuitivement on devrait chercher à avoir moins de composantes cinétiques $\widetilde{Q}_k$ que de clusters Gaussiens (définis par le paramètre $\widetilde{\boldsymbol{\theta}}_k$). On devrait donc pouvoir permettre à des clusters spatiaux de partager la même cinétique. Cela se traduit par l'introduction d'un deuxième niveau de hiérarchie dans le modèle 5.13 et qui permettrait de clusteriser les distributions $\widetilde{Q}_k$ elles-mêmes.

Nous avons vu dans la section 2.4.2 que le processus de Dirichlet imbriqué (NDP) de Rodriguez *et al.* [RDG08] permettait de modéliser des distributions de probabilité et de

les clusteriser. On rappelle qu'un ensemble de distributions $\{G_1, \ldots, G_J\}$ suit un NDP de paramètres α, β et H (où H est la mesure de base du processus) si

$$\begin{aligned} \{G_1, \ldots, G_J\} &\sim \mathcal{Q} \\ \mathcal{Q} &\sim \mathrm{DP}(\alpha, \mathrm{DP}(\beta, H)). \end{aligned} \tag{5.14}$$

Cette définition signifie que

$$\begin{aligned} \{G_1, \ldots, G_J\} &\sim \mathcal{Q} \text{ où } \mathcal{Q} \overset{d}{=} \sum_{k=1}^{\infty} \pi_k \delta_{G_k^*} \text{ et } G_k^*(\cdot) \overset{d}{=} \sum_{l=1}^{\infty} \omega_{lk} \delta_{\boldsymbol{\theta}_{lk}^*(\cdot)} \\ \text{avec } \boldsymbol{\theta}_{lk}^* &\sim H \qquad \pi_k \sim \mathrm{GEM}(\alpha) \qquad \omega_{lk} \sim \mathrm{GEM}(\beta). \end{aligned}$$

Nous adaptons cette définition de façon à l'appliquer aux processus de Pitman-Yor par l'ajout d'un deuxième paramètre dans le DP et définir ainsi les processus de Pitman-Yor imbriqués. Néanmoins dans notre problème, contrairement à celui de Rodriguez *et al.*, le groupe de distributions que nous cherchons à clusteriser n'est pas fini (ce sont les cinétiques $\widetilde{Q}_k$). Par ailleurs dans la modélisation que nous proposons, la mesure de base de la distribution $\mathcal{Q}$ n'est pas un DP mais un arbre de Pólya. Nous définissons alors un nouveau processus nommé *processus de Pitman-Yor imbriqué d'arbres de Pólya*, noté NPY-PT (pour « Nested Pitman-Yor process with a Pólya Tree as base distribution »).

Définition 10. *Un ensemble de distributions $\{\widetilde{Q}_1, \widetilde{Q}_2, \ldots\}$ suit un NPY-PT de paramètres γ, ν, $\mathcal{A}$ et Q_0 si*

$$\begin{aligned} \{\widetilde{Q}_1, \widetilde{Q}_2, \ldots\} &\sim \mathcal{K}_0 \\ \mathcal{K}_0 &\sim PY(\gamma, \nu, PT(\mathcal{A}, Q_0)). \end{aligned} \tag{5.15}$$

Partant de cette définition, nous pouvons maintenant décrire le modèle hiérarchique génératif que nous proposons pour les données TEP spatio-temporelles.

5.5.2 Modèle hiérarchique pour les données TEP dynamiques

On rappelle que seulement les projections $\boldsymbol{Y} = (y_1, \ldots, y_n)$ sont observées en tomographie d'émission et donc que les lieux d'annihilation $\boldsymbol{X} = (\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n)$ sont introduits comme des variables cachées. Le modèle hiérarchique génératif s'écrit :

$$\begin{aligned} y_i|\boldsymbol{x}_i &\overset{\text{ind}}{\sim} \mathcal{P}(y_i|\boldsymbol{x}_i) \\ \boldsymbol{x}_i, t_i|\boldsymbol{\theta}_i, Q_i &\overset{\text{ind}}{\sim} \mathcal{N}(\boldsymbol{x}_i|\boldsymbol{\theta}_i) \times Q_i(t_i) \\ \boldsymbol{\theta}_i, Q_i|H &\overset{\text{iid}}{\sim} H \\ H &\sim PY(\alpha, \beta, G_0 \times \mathcal{K}_0) \\ \mathcal{K}_0 &\sim PY(\gamma, \nu, \mathrm{PT}(\mathcal{A}, Q_0)) \end{aligned} \tag{5.16}$$

H est un processus de Pitman-Yor de distribution

$$H(\cdot) = \sum_{k=1}^{\infty} w_k \, \delta_{\widetilde{\boldsymbol{\theta}}_k, \widetilde{Q}_k}(\cdot)$$

avec

- $\mathbf{w} = w_1, w_2, \ldots \sim \text{GEM}(\alpha, \beta)$,
- $\widetilde{\boldsymbol{\theta}} = \widetilde{\boldsymbol{\theta}}_1, \widetilde{\boldsymbol{\theta}}_2, \ldots$ sont les valeurs distinctes de $\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \ldots$, et sont $i.i.d$ suivant G_0 un modèle Normal-Inverse Wishart,
- $\widetilde{\mathbf{Q}} = \widetilde{Q}_1, \widetilde{Q}_2, \ldots$ sont $i.i.d$ suivant $\mathcal{K}_0$ distribuée suivant

$$\mathcal{K}_0(\cdot) = \sum_{j=1}^{\infty} \pi_j \, \delta_{Q_j^\star}(\cdot)$$

où

- $\boldsymbol{\pi} = \pi_1, \pi_2, \ldots \sim \text{GEM}(\gamma, \nu)$,
- $\mathbf{Q}^\star = Q_1^\star, Q_2^\star, \ldots$ sont les valeurs uniques de $\widetilde{Q}_1, \widetilde{Q}_2, \ldots$ et sont indépendantes et identiquement distribuées suivant $\text{PT}(\mathcal{A}, Q_0)$, un processus d'arbres de Pólya de paramètres $\mathcal{A}$ et de distribution de base Q_0.

Ce modèle hiérarchique peut être compris de la façon suivante : chaque $(\boldsymbol{x}_i, t_i)$ est généré suivant une Gaussienne $\mathcal{N}(\boldsymbol{x}_i|\boldsymbol{\theta}_i)$ et une fonction temporelle $Q_i(t_i)$. Puisque $H(\cdot)$ est discrète, cela induit un premier clustering des observations partageant les mêmes paramètres spatiaux $\widetilde{\boldsymbol{\theta}}$ et les mêmes cinétiques $\widetilde{Q}$. Le second niveau de clustering sur les cinétiques est dû à la nature discrète de $\mathcal{K}_0(\cdot)$, ce qui permet à plusieurs $\widetilde{Q}_k$ de prendre simultanément la même valeur Q_j^* pour un certain j.

Le modèle 5.16 peut être ré-écrit en introduisant des variables de classification latentes $\mathbf{c}$ et $\mathbf{d}$ représentant respectivement l'appartenance aux clusters observationnels et distributionnels (cinétiques).

5.5.3 Modèle alternatif avec des variables de classification

Soit $\mathbf{c} = (c_1, c_2, \ldots, c_n)$ le vecteur des variables de classification des émissions aux composantes spatio-temporelles telles que $(\boldsymbol{\theta}_i, Q_i) = (\widetilde{\boldsymbol{\theta}}_{c_i}, \widetilde{Q}_{c_i})$ pour tout $i \leq n$. Soit $\mathbf{d} = (d_1, d_2, \ldots)$, les variables d'allocation des cinétiques telles que $\widetilde{Q}_k = Q^\star_{d_k}$ pour tout k. On ré-écrit le modèle 5.16 de la façon suivante :

$$\begin{aligned}
y_i|\boldsymbol{x}_i &\overset{\text{ind}}{\sim} \mathcal{P}(y_i|\boldsymbol{x}_i) \\
\boldsymbol{x}_i, t_i|c_i, \mathbf{d}, \widetilde{\boldsymbol{\Theta}}, \widetilde{\boldsymbol{Q}} &\overset{\text{ind}}{\sim} \mathcal{N}\left(\boldsymbol{x}_i|\widetilde{\boldsymbol{\theta}}_{c_i}\right) \times Q^*_{d_{c_i}}(t_i) \\
c_i|\boldsymbol{w} &\overset{\text{iid}}{\sim} \sum_{k=1}^{\infty} w_k \delta_k(\cdot) \\
d_k|\boldsymbol{\pi} &\overset{\text{iid}}{\sim} \sum_{j=1}^{\infty} \pi_j \delta_j(\cdot) \\
\boldsymbol{w} &\sim \text{GEM}(\alpha, \beta) \\
\widetilde{\boldsymbol{\theta}}_k &\overset{\text{iid}}{\sim} \mathcal{NIW}_{\rho, n_0, \boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0} \\
\boldsymbol{\pi} &\sim \text{GEM}(\gamma, \nu) \\
Q_j^* &\sim \text{PT}_L(\mathcal{A}, Q_0)
\end{aligned} \tag{5.17}$$

5.5.4 Modèle alternatif avec des variables slices

Un point clé dans l'inférence est de traiter les distributions infinies qui apparaissent dans le modèle 5.17, tout en évitant de les tronquer. Tout comme dans la reconstruction spatiale, nous utilisons la stratégie de slice sampling. Soit $\mathbf{u} = (u_1, , u_2, \ldots, , u_n)$ des variables auxiliaires uniformes telles que pour tout $(\boldsymbol{x}_i, t_i, u_i)$,

$$\left(\boldsymbol{x}_i, t_i, u_i | \boldsymbol{w}, \widetilde{\boldsymbol{\Theta}}, \widetilde{\boldsymbol{Q}}\right) \sim \sum_{k=1}^{\infty} \mathcal{U}\left(u_i | 0, \zeta_k\right) w_k \mathcal{N}\left(\boldsymbol{x}_i | \widetilde{\boldsymbol{\theta}}_k\right) \widetilde{Q}_k\left(t_i\right) \tag{5.18}$$

où $\mathcal{U}\left(\cdot | a, b\right)$ désigne la loi uniforme sur l'intervalle $]a, b]$ et pour tout k, $\xi_k = \min\left(w_k, \zeta\right)$ où $\zeta \in]0, 1]$ est un seuil que l'on introduit pour améliorer les propriétés de mélange de l'algorithme MCMC (cf. chapitre 3). On peut, alternativement ré-écrire la distribution 5.18

$$\left(\boldsymbol{x}_i, t_i, u_i | \boldsymbol{w}, \widetilde{\boldsymbol{\Theta}}, \widetilde{\boldsymbol{Q}}\right) \sim \zeta^{-1} \sum_{u_i < \zeta < w_k} w_k \mathcal{N}\left(\boldsymbol{x}_i | \widetilde{\boldsymbol{\theta}}_k\right) \widetilde{Q}_k\left(t_i\right) + \sum_{u_i < w_k \leq \zeta} \mathcal{N}\left(\boldsymbol{x}_i | \widetilde{\boldsymbol{\theta}}_k\right) \widetilde{Q}_k\left(t_i\right).$$

De la même manière, nous introduisons des variables auxiliaires uniformes $\boldsymbol{v} = (v_1, v_2, \ldots)$ pour le NPY-PT

$$\left(\widetilde{Q}_k, v_k | \boldsymbol{\pi}, \boldsymbol{Q}^*\right) \sim \zeta^{-1} \sum_{v_k < \zeta < \pi_j} \pi_j \, \delta_{Q_j^\star}\left(\widetilde{Q}_k\right) + \sum_{v_k < \pi_j \leq \zeta} \delta_{Q_j^\star}\left(\widetilde{Q}_k\right).$$

Le modèle hiérarchique 5.17 peut être reformulé en introduisant ces variables slices uniformes :

$$\begin{aligned}
y_i | \boldsymbol{x}_i &\overset{\text{ind}}{\sim} \mathcal{P}\left(y_i | \boldsymbol{x}_i\right) \\
\boldsymbol{x}_i, t_i | c_i, \mathbf{d} &\overset{\text{ind}}{\sim} \mathcal{N}\left(\boldsymbol{x}_i | \widetilde{\boldsymbol{\theta}}_{c_i}\right) \times Q^*_{d_{c_i}}(t_i) \\
c_i, u_i | \boldsymbol{w} &\overset{\text{iid}}{\sim} \sum_{k=1}^{\infty} \mathcal{U}\left(u_i | 0, \min(w_k, \zeta)\right) w_k \delta_k(c_i) \\
d_k, v_k | \boldsymbol{\pi} &\overset{\text{iid}}{\sim} \sum_{j=1}^{\infty} \mathcal{U}\left(u_i | 0, \min(\pi_j, \zeta)\right) \pi_j \delta_j(d_k) \\
\boldsymbol{w} &\sim \text{GEM}(\alpha, \beta) \\
\widetilde{\boldsymbol{\theta}}_k &\overset{\text{iid}}{\sim} \mathcal{NIW}_{\rho, n_0, \boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0} \\
\boldsymbol{\pi} &\sim \text{GEM}(\gamma, \nu) \\
Q_j^* &\sim \text{PT}_L(\mathcal{A}, Q_0)
\end{aligned} \tag{5.19}$$

5.5.5 Inférence

L'inférence s'effectue en considérant l'inversion du modèle hiérarchique 5.19. Nous proposons un algorithme MCMC pour générer des échantillons suivant $G(\cdot)|\boldsymbol{y}, \boldsymbol{t}$. L'échantillonneur tire successivement des blocs de variables suivant les distributions

conditionnelles suivantes :

$$\begin{array}{ll}
\text{Proposition de localisation de l'émission :} & (\boldsymbol{X}|\mathbf{w},\mathbf{u},\widetilde{\boldsymbol{\Theta}},\mathbf{Q}^{\star},\mathbf{d},\mathbf{y},\mathbf{t}) \\
\text{Allocation de l'émission aux atomes de } H : & (\mathbf{c}|\mathbf{w},\mathbf{u},\widetilde{\boldsymbol{\Theta}},\mathbf{Q}^{\star},\mathbf{d},\boldsymbol{x},\mathbf{t}) \\
\text{Mise à jour des paramètres spatiaux :} & (\widetilde{\boldsymbol{\Theta}}|\mathbf{c},\boldsymbol{x}) \\
\text{Tirage des poids de } H \text{ et variables slices associées :} & (\mathbf{w},\mathbf{u}|\mathbf{c}) \\
\text{Allocation des cinétiques aux atomes de } \mathcal{K}_0 : & (\mathbf{d}|\boldsymbol{\pi},\mathbf{v},\mathbf{Q}^{\star},\mathbf{c},\mathbf{t}) \\
\text{Mise à jour des cinétiques} & (\mathbf{Q}^{\star}|\mathbf{d},\mathbf{c},\mathbf{t}) \\
\text{Tirage des poids de } \mathcal{K}_0 \text{ et variables slices associées :} & (\boldsymbol{\pi},\mathbf{v}|\mathbf{d})
\end{array} \tag{5.20}$$

L'algorithme MCMC est un échantillonneur de Gibbs excepté l'échantillonnage de $(\boldsymbol{x}_i|\mathbf{w},\,\widetilde{\boldsymbol{\Theta}},\mathbf{Q}^{\star},\mathbf{d},y_i,t_i)$ qui requiert une étape Metropolis-Hastings.

Les distributions conditionnelles sont maintenant explicitées.

Échantillonnage de $(\boldsymbol{w},\boldsymbol{u}|\boldsymbol{c})$

Cette distribution est la même que dans le chapitre 4 pour la reconstruction spatiale 3D. On rappelle que si K_n est le nombre de valeurs uniques parmi n observations et $n_k = \#\{i : c_i = k\}$ pour $k \leq K_n$ alors l'échantillonnage s'effectue de la façon suivante.

- Tirer w_k pour $k \leq K_n$

$$w_1,\ldots,w_{K_n},r_{K_n}|\mathbf{c} \sim \mathrm{Dir}\left(n_1-\alpha,\ldots,\,n_{K_n}-\alpha,\,\beta+K_n\,\alpha\right).$$

- Tirer $u_i|w_1,w_2,\ldots,w_{K_n},\mathbf{c}$

$$u_i|w_1,w_2,\ldots,w_{K_n},\mathbf{c} \overset{\text{ind.}}{\sim} \mathcal{U}\left(u_i|0,\min\left(w_{c_i},\zeta\right)\right).$$

- Prendre $u^{\star} = \min(\mathbf{u})$.
- Tirer w_k pour $k > K_n$. Tant que $r_{k-1} > u^{\star}$

$$\begin{aligned} V_k &\sim \mathrm{Beta}\left(1-\alpha,\beta+k\,\alpha\right) \\ w_k &= V_k\,r_{k-1} \\ r_k &= r_{k-1}\left(1-V_k\right). \end{aligned}$$

- Prendre $K^{*} = \min\left(\{k : r_k < u^{\star}\}\right)$.

Échantillonnage de $(\mathbf{c}|\mathbf{w},\mathbf{u},\widetilde{\boldsymbol{\Theta}},\mathbf{Q}^{\star},\mathbf{d},\boldsymbol{x},\mathbf{t})$

$$c_i|\mathbf{w},\mathbf{u},\widetilde{\boldsymbol{\Theta}},\mathbf{Q}^{\star},\mathbf{d},\boldsymbol{x},\mathbf{t} \overset{\text{ind}}{\sim} \sum_{k=1}^{K^{*}} w_{k,i}\,\delta_k\left(\cdot\right) \tag{5.21}$$

où

$$w_{k,i} \propto \mathbf{1}\left(w_k > u_i\right)\max\left(w_k,\zeta\right) f_{\mathcal{N}}\left(\boldsymbol{x}_i|\widetilde{\boldsymbol{\theta}}_k\right) f_{Q^{*}_{d_k}}\left(t_i\right)$$

et $\sum_{j=1}^{K^{*}} w_{k,i} = 1$.

Échantillonnage de $(\widetilde{\boldsymbol{\Theta}}|\mathbf{c}, \boldsymbol{x})$

Cette conditionnelle est aussi un modèle Normal-Inverse Wishart avec des paramètres mis à jour

– Mise à jour des vecteurs moyens et matrices de covariance pour les composantes non-vides : $\forall\ k \leq K_n$,

$$\mathbf{m}_k|\boldsymbol{\Sigma}_k, \mathbf{C}, \boldsymbol{x} \overset{\text{ind}}{\sim} \mathcal{N}\left(\mu_k,\ \frac{\boldsymbol{\Sigma}_k}{\rho_k}\right)$$

avec

$$\begin{aligned}
\rho_k &= \rho + n_k \\
\mu_k &= \frac{\rho\,\mu_0 + n_k\,\widehat{\mu_k}}{\rho + n_k}, \\
\widehat{\mu_k} &= \frac{1}{n_k} \sum_{\{i:c_i=k\}} \boldsymbol{x}_i
\end{aligned}$$

$$\text{and } {\boldsymbol{\Sigma}_k}^{-1}|\mathbf{C}, \boldsymbol{x} \overset{\text{ind}}{\sim} \mathcal{W}(\eta_k, (\eta_k\,\mathbf{S}_k)^{-1})$$

avec

$$\begin{aligned}
\eta_k &= n_0 + n_k, \\
\mathbf{S}_k &= \frac{1}{\eta_k}\left[n_0\,\boldsymbol{\Sigma}_0 + n_k\,\widehat{\boldsymbol{\Sigma}_k} + \frac{(\mu_0 - \widehat{\mu_k})(\mu_0 - \widehat{\mu_k})^{\mathsf{T}}}{\frac{1}{\rho} + \frac{1}{n_k}}\right], \\
\widehat{\boldsymbol{\Sigma}_k} &= \frac{1}{n_k} \sum_{\{i:c_i=k\}} (\boldsymbol{x}_i - \widehat{\mu_k})(\boldsymbol{x}_i - \widehat{\mu_k})^{\mathsf{T}}.
\end{aligned}$$

– Tirer les paramètres des composantes vides suivant leurs lois *a priori*

$$\widetilde{\boldsymbol{\theta}}_k \overset{\text{i.i.d.}}{\sim} \mathcal{NIW}_{\rho,n_0,\mu_0,\boldsymbol{\Sigma}_0}, \text{ pour } K_n < k \leq K^*.$$

Échantillonnage de $(\boldsymbol{\pi}, \mathbf{v}|\mathbf{d})$

Nous adoptons le même schéma que pour l'échantillonnage des $(\mathbf{w}, \mathbf{u}|\mathbf{c})$, excepté qu'ici les variables à affecter ne sont pas les observations indexées par i mais plutôt les clusters non-vides indexés par $k,\ k \leq K_n$. Soit $m_j = \#\{k \leq K_n : d_k = j\}$, le nombre d'atomes de H alloués au $j^{\text{ème}}$ atome de $\mathcal{K}_0$ et J_{K_n} le plus grand indice des atomes actifs de $\mathcal{K}_0$. On échantillonne de façon jointe $\pi, \mathbf{v}|\mathbf{d}$ en tirant d'abord $\pi_1, \pi_2, \ldots, \pi_{J_{K_n}}|\mathbf{d}$, puis $\mathbf{v}|\pi_1, \pi_2, \ldots, \pi_{J_{K_n}}, \mathbf{d}$, et enfin $\pi_{J_{K_n+1}}, \pi_{J_{K_n+2}}, \ldots|\mathbf{v}$.

– Tirer π_j pour $j \leq J_{K_n}$

$$\pi_1, \ldots, \pi_{J_{K_n}}, r_{J_{K_n}}|\mathbf{d} \sim \text{Dir}\left(m_1 - \gamma, \ldots, m_{J_{K_n}} - \gamma,\ \nu + J_{K_n}\,\gamma\right).$$

– Tirer $v_j|\pi_1, \pi_2, \ldots, \pi_{J_{K_n}}, \mathbf{d}$

$$v_j|\pi_1, \pi_2, \ldots, \pi_{J_{K_n}}, \mathbf{d} \overset{\text{ind}}{\sim} \mathcal{U}\left(v_j|0, \min\left(\pi_{d_j}, \zeta'\right)\right).$$

– Faire $v^\star = \min(\mathbf{v})$.

– Tirer π_j pour $j > J_{K_n}$. Tant que $\rho_{j-1} > v^\star$

$$U_j \sim \text{Beta}\,(1-\gamma, \nu + j\,\gamma)$$
$$\pi_j = U_j\,\rho_{j-1}$$
$$\rho_j = \rho_{j-1}\,(1-U_j)\,.$$

– Faire ,$J^\star = \min\left(\{j :\ \rho_j < v^\star\}\right)$.

Échantillonnage de $(\mathbf{d}|\boldsymbol{\pi}, \mathbf{v}, \mathbf{Q}^\star, \mathbf{c}, \mathbf{t})$

– Échantillonner indépendamment pour tout $k \leq K_n$

$$d_k|\boldsymbol{\pi}, \mathbf{v}, \mathbf{Q}^\star, \mathbf{c}, \mathbf{t} \overset{\text{ind}}{\sim} \sum_{j=1}^{J^\star} \pi_{j,k}\,\delta_j\,(\cdot)$$

où

$$\pi_{j,k} \propto \mathbf{1}\,(\pi_j > v_k)\max\left(\pi_j, \zeta'\right) \prod_{i:c_i=k} f_{Q_j^*}\,(t_i)\,.$$

– Échantillonner indépendamment pour tout $K_n < k \leq K^*$

$$d_k|\boldsymbol{\pi}, \mathbf{v}, \mathbf{Q}^\star, \mathbf{c}, \mathbf{t} \overset{\text{ind}}{\sim} \sum_{j=1}^{J^\star} \pi_{j,k}\,\delta_j\,(\cdot)$$

avec

$$\pi_{j,k} \propto \mathbf{1}\,(\pi_j > v_k)\max\left(\pi_j, \zeta'\right).$$

La vraisemblance de l'arbre de Pólya intervient dans l'affectation d'une cinétique $\widetilde{Q}$ à une cinétique « principale » $Q^\star$. Cela permet d'augmenter la probabilité de choisir une composante principale avec un poids *a priori* faible mais ayant beaucoup de données d'affectées.

Échantillonnage de $(\mathbf{Q}^\star|\mathbf{d}, \mathbf{c}, \mathbf{t})$

– Grâce aux propriétés de conjugaison des arbres de Pólya, la distribution conditionnelle de $\mathbf{Q}^\star|\mathbf{d}, \mathbf{c}, \mathbf{t}$ est aussi un arbre de Pólya avec des paramètres mis à jour

$$Q_j^\star|\mathbf{d}, \mathbf{c}, \mathbf{t} \overset{\text{ind}}{\sim} \text{PT}_L(\mathcal{A}'_j, Q_0)$$

avec

$$\mathcal{A}'_j = \{\alpha'_{j,\epsilon} : \epsilon \in E^*\}$$
$$\alpha'_{j,\epsilon} = \alpha_\epsilon + n_{j,\epsilon}$$

et

$$n_{j,\epsilon} = \#\{i \in \{1,\ldots,n\} : d_{c_i} = j \text{ and } t_i \in B_\epsilon\}.$$

– Pour tout $J_{K_n} < j \leq J^*$, tirer indépendamment

$$Q_j^\star|\mathbf{d}, \mathbf{c}, \mathbf{t} \overset{\text{iid}}{\sim} \text{PT}_L(\mathcal{A}, Q_0).$$

Échantillonnage de $(\boldsymbol{x}|\boldsymbol{y}, \mathbf{t}, \mathbf{u}, \mathbf{w}, \widetilde{\boldsymbol{\Theta}}, \mathbf{Q}^\star, \mathbf{d})$

La génération directe suivant cette distribution n'est pas faisable à cause du projecteur et une étape Metropolis-Hastings est nécessaire. Utilisant la règle Bayes, on a

$$\boldsymbol{x}_i|y_i, t_i, u_i, \mathbf{w}, \widetilde{\boldsymbol{\Theta}}, \mathbf{Q}^\star, \mathbf{d} \overset{\text{ind}}{\sim} \mathcal{P}(y_i|\boldsymbol{x}_i) \times \widetilde{G}_i(\boldsymbol{x}_i)$$

avec

$$f_{\widetilde{G}_i}(\boldsymbol{x}_i) \propto \sum_{k=1}^{K^*} \mathbf{1}\left(w_k > u_i\right) \max\left(w_k, \zeta\right) f_{\mathcal{N}}\left(\boldsymbol{x}_i|\widetilde{\boldsymbol{\theta}}_k\right) f_{Q^*_{d_k}}\left(t_i\right)$$

et $\mathcal{P}(y_i|\boldsymbol{x}_i)$ la distribution de projection.
Soit G_i^* la distribution de proposition pour $i = 1, \ldots, n$,

$$f_{G_i^*}(\boldsymbol{x}_i^*) \propto f_{\mathcal{N}}(\boldsymbol{x}_i^*|\mathbf{m}_{y_i}^*, \boldsymbol{\Sigma}_{y_i}^*) \times f_{\widetilde{G}_i}(\boldsymbol{x}_i^*).$$

La loi candidate est un produit entre le mélange Gaussien multivarié et une loi normale choisie de façon à approximer la distribution de projection dans la LOR. Il en résulte un mélange re-pondéré de lois normales dans la direction de la LOR. Ainsi pour générer $\boldsymbol{x}_i$ pour tout $i \leq n$, à l'itération $s+1$,

- on choisit d'abord une des Gaussiennes du mélange en tirant indépendamment

$$\widetilde{c}_i \overset{\text{ind}}{\sim} \sum_{k=1}^{K^*} \widetilde{w}_{k,i} \delta_k(\cdot)$$

  où

$$\widetilde{w}_{k,i} \propto \mathbf{1}\left(w_k > u_i\right) \max\left(w_k, \zeta\right) \widetilde{\eta}_{k,y_i} f_{Q^*_{d_k}}\left(t_i\right)$$

  et $\sum_{k=1}^{K^*} \widetilde{w}_{k,i} = 1$.
- puis on tire $\boldsymbol{x}_i^*$ suivant la Gaussienne choisie

$$\boldsymbol{x}_i^* \overset{\text{ind}}{\sim} \mathcal{N}\left(\boldsymbol{x}_i^*|\widetilde{\mathbf{m}}_{\widetilde{c}_i,y_i}, \widetilde{\boldsymbol{\Sigma}}_{\widetilde{c}_i,y_i}\right).$$

- on calcule le taux d'acceptation $\omega(\boldsymbol{x}_i^*, \boldsymbol{x}_i^{(t)})$

$$\omega(\boldsymbol{x}_i^*, \boldsymbol{x}_i^{(s)}) = \frac{\mathbb{P}(y_i|\boldsymbol{x}_i^*)}{\mathbb{P}(y_i|\boldsymbol{x}_i^{(s)})} \cdot \frac{f_{\mathcal{N}}\left(\boldsymbol{x}_i^{(s)}|\mathbf{m}_{y_i}^*, \boldsymbol{\Sigma}_{y_i}^*\right)}{f_{\mathcal{N}}\left(\boldsymbol{x}_i^*|\mathbf{m}_{y_i}^*, \boldsymbol{\Sigma}_{y_i}^*\right)}.$$

- et on accepte $\boldsymbol{x}_i^*$ avec la probabilité

$$\min(1,\ \omega(\boldsymbol{x}_i^*, \boldsymbol{x}_i^{(s)})).$$

5.5.6 Estimées MCMC

À la convergence de l'algorithme, on obtient des échantillons suivant la distribution *a posteriori* jointe de $(\boldsymbol{X}, \boldsymbol{c}, \widetilde{\boldsymbol{\Theta}}, \boldsymbol{w}, \mathbf{d}, \mathbf{Q}^*, \boldsymbol{\pi}|\boldsymbol{y}, \boldsymbol{t})$. À partir de chaque tirage s retenu, on peut construire une estimée de la densité spatio-temporelle pour les évènements enregistrés

$$f_{G^{(s)}}\left(\boldsymbol{x}, t\right) \approx \sum_{k=1}^{K^*} w_k^{(s)} f_{\mathcal{N}}\left(\boldsymbol{x}|\widetilde{\boldsymbol{\theta}}_k^{(s)}\right) f_{\widetilde{Q}_k^{(s)}}\left(t\right) \tag{5.22}$$

où K^* est le nombre de composantes retenues par le slice.

La densité spatio-temporelle de tous les évènements est obtenue par normalisation avec la distribution de projection

$$f_{G^{\dagger(s)}}(\boldsymbol{x}, t) \propto \frac{f_{G^{(s)}}(\boldsymbol{x}, t)}{\sum_{l=1}^{L} \mathbb{P}(y = l|\boldsymbol{x})}. \tag{5.23}$$

L'estimateur que nous avons choisi pour la distribution d'activité est l'espérance conditionnelle de $f_{G^\dagger}$ sur les $N = 10.000$ tirages retenus,

$$\mathbb{E}(f_{G^\dagger}|\boldsymbol{Y}) \approx \frac{1}{N} \sum_{s=1}^{N} f_{G^{\dagger(s)}}.$$

Cet estimateur est aussi la densité prédictive de l'activité spatio-temporelle, $\mathbb{E}(G^\dagger(\cdot, \cdot)|\boldsymbol{Y}) = \mathbb{P}(\boldsymbol{x}_{n+1} \in \cdot, t_{n+1} \in \cdot|\boldsymbol{Y})$. Elle donne la probabilité de la localisation d'une annihilation non encore observée et le temps d'observation correspondant sachant n émissions déjà observées.

5.5.7 Matériels de simulation et résultats

Matériels

Les désintégrations ont été générées suivant un fantôme cérébral 3D voxellisé tel que $n = 10^7$ coïncidences ont été enregistrées en prenant en compte le champ de vue du scanner (cf. reconstruction 3D, chapitre 4). Le fantôme est composé de plusieurs volumes fonctionnels, chacun identifié par une TAC spécifique. Les TACs correspondent à l'évolution temporelle de l'activité dans le compartiment vasculaire (veines et artères) et dans les compartiments tissulaires que sont la matière grise, la matière blanche, le cervelet et une tumeur. L'activité générée dans le fantôme correspond à un examen FDG avec $k_4 = 0$. La fonction d'entrée est considérée comme étant l'activité dans le compartiment sanguin. Le modèle compartimental utilisé est un peu différent de celui de Sokoloff (présenté au paragraphe 5.2.2) en ce sens qu'il tient compte de la fraction du volume sanguin dans les tissus. En conséquence, la TAC de chaque volume fonctionnel est supposé être un mélange de la cinétique propre du tissu et de la cinétique sanguine, en tenant compte de la fraction vasculaire volumique. Le modèle décrivant l'évolution de l'activité s'écrit alors

$$C(t) = V_b C_b(t) + (1 - V_b)\left[C_b(t) \otimes h(K_1, K_2, K_3, t)\right] \tag{5.24}$$

où V_b est la fraction volumique, h est l'IRF du tissu cérébral, $C(t)$ et C_b désignent respectivement l'évolution de l'activité dans le tissu et dans le sang. On retrouve le modèle 5.5 pour $V_b = 0$ et $C_a = C_b$. Ce modèle est plus général car en réalité la radioactivité mesurée par l'image TEP comprend la somme de l'activité des compartiments tissulaires mais aussi une fraction du volume sanguin. La constante de demi-vie du ^{18}F a aussi été prise en compte dans la simulation. On représente dans la Figure 5.2 les TAC simulées.

Paramétrisation de l'algorithme

Pour la marginale spatiale les principaux paramètres du processus de Pitman-Yor ont été choisis de la façon suivante $\alpha = .02$, $\beta = 1000$, $\mathbf{\Sigma}_0 = 6.25 \times \mathbb{I}_3$, $n_0 = 4$. Les paramètres pour la marginale temporelle du processus imbriqué de cinétiques sont $\gamma = 0$, $\nu = 1$. La mesure de base Q_0 est une loi uniforme. Nous avons calculé 5.000 itérations burn-in pour l'échantillonneur, suivies de 10.000 retenues pour calculer les fonctionnelles de la distribution d'émission spatio-temporelle.

Estimées

Un avantage de l'approche bayésienne non paramétrique est que la distribution entière du volume fonctionnel est accessible. À partir de l'équation 5.22, on peut marginaliser par rapport au temps pour définir la densité spatiale du volume fonctionnel j associé à la cinétique $Q_j^\star$ de la façon suivante : pour tout j (labels des atomes de $\mathcal{K}_0$),

$$f_{VF_j}(\boldsymbol{x}) = \sum_{k:\,\widetilde{Q}_k = Q_j^\star}^{K^*} w_k \, f_{\mathcal{N}}\left(\boldsymbol{x}|\widetilde{\boldsymbol{\theta}}_k\right). \tag{5.25}$$

De la même façon, on obtient la marginale temporelle (somme pondérée de toutes les cinétiques) :

$$f(t) = \sum_k w_k f_{\widetilde{Q}_k}(t) = \sum_j \left(\sum_{k:\,\widetilde{Q}_k = Q_j^\star} w_k \right) f_{Q_j^*}(t). \tag{5.26}$$

Dans l'approche bayésienne non paramétrique, le nombre de cinétiques à chaque itération est illimité. Une façon de représenter les cinétiques est d'utiliser un « histogramme temporel ponctuel » effectué sur les tirages MCMC. Ce *spectre des cinétiques* est obtenu de la façon suivante : on calcule, par marginalisation de la distribution $H(\widetilde{\boldsymbol{\theta}}, \widetilde{Q}|\mathbf{Y})$ sur l'espace des paramètres $\widetilde{\boldsymbol{\theta}}$, la distribution de probabilité des valeurs des cinétiques pour tout t,

$$\mathbb{P}(\mathcal{K}(t)|\mathbf{Y}) = \int_\Theta H\left(\widetilde{\boldsymbol{\theta}}, \widetilde{Q}|\mathbf{Y}\right) \mathrm{d}\widetilde{\boldsymbol{\theta}} = \sum_{j=1}^{J^*} \Big(\sum_{k:\,\widetilde{Q}_k = Q_j^*}^{K^*} w_k \Big) \, \delta_{Q_j^*(t)}$$

où J^* est le nombre de cinétiques retenues par l'algorithme du slice (par analogie à K^* pour les composantes spatio-temporelles). En d'autres termes, pour calculer ce spectre, on affecte à chaque cinétique « principale » Q_j^* un poids donné par la somme des poids des composantes qui lui sont affectées. En incrémentant l'histogramme des valeurs des cinétiques par le poids correspondant à chaque itération de l'échantillonneur, on obtient l'espérance *a posteriori* du spectre des cinétiques comme tracée à la Figure 5.3. Cet histogramme peut être comparé avec les cinétiques simulées (Figure 5.2).

Enfin, une autre caractéristique de l'approche non paramétrique est qu'en sus du nombre illimité de cinétiques, les labels des composantes ne sont pas significatifs et les

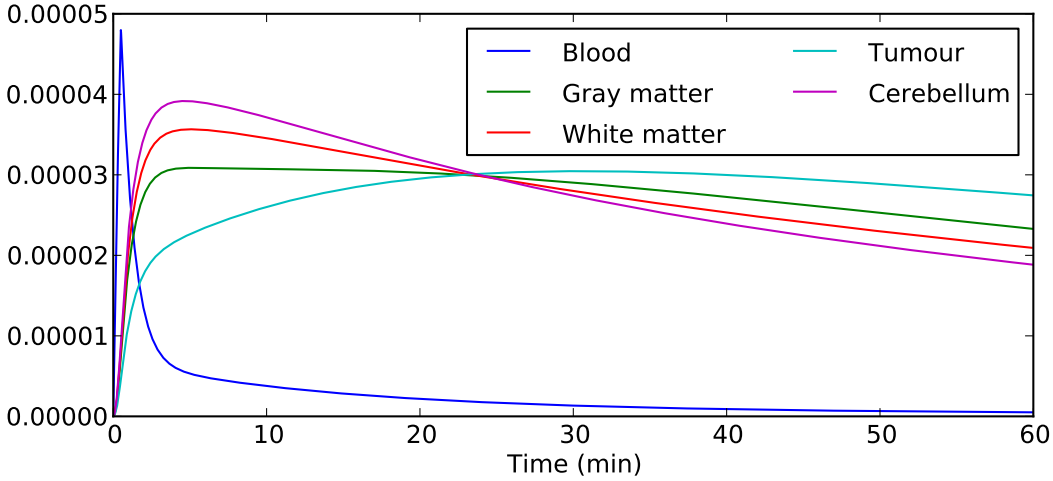


FIGURE 5.2 – Cinétiques utilisées pour la simulation des données.

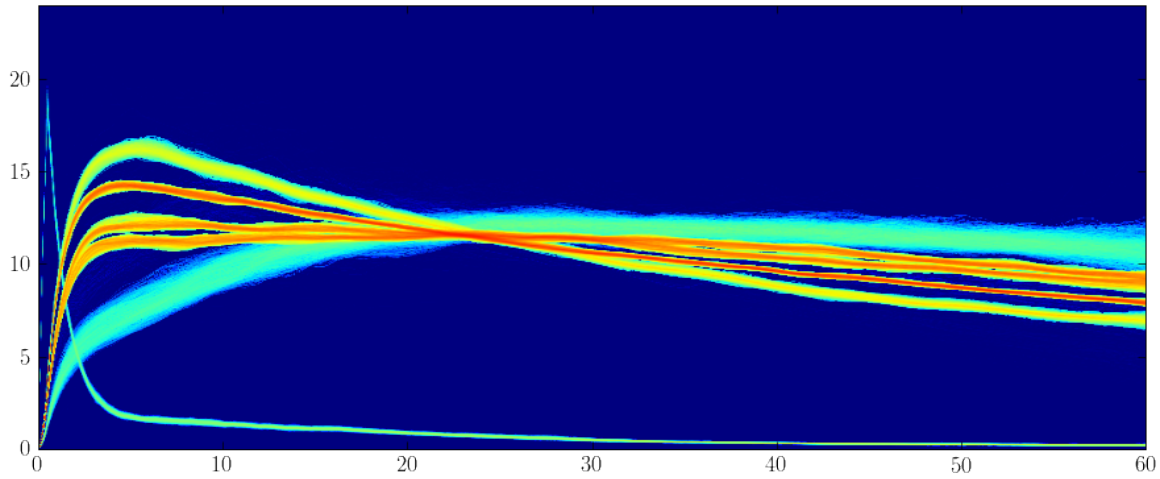
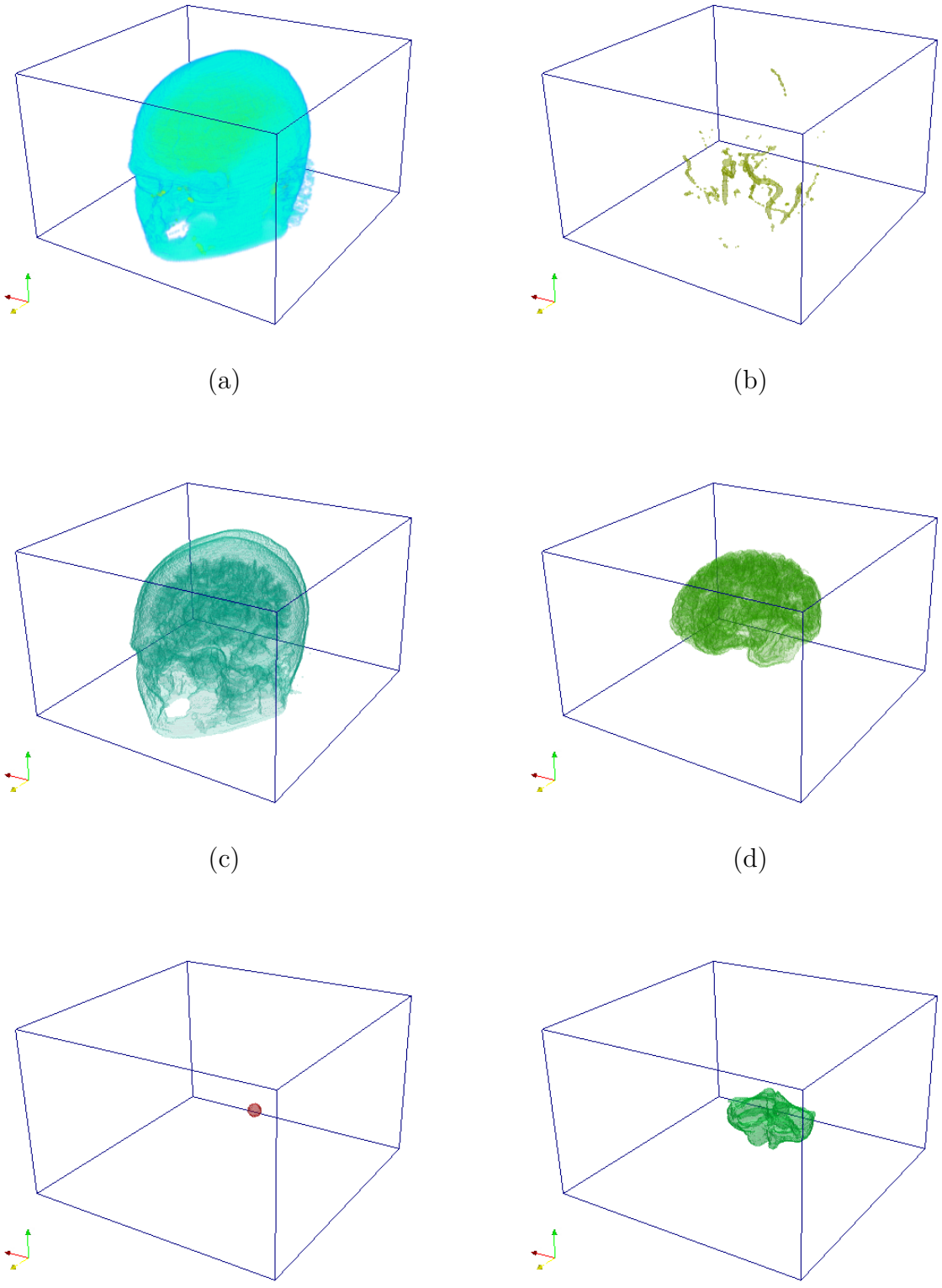


FIGURE 5.3 – Spectre des cinétiques estimées sur les tirages Monte-Carlo. La couleur de la courbe indique le poids de la TAC dans la distribution d'activité reconstruite.

composantes ne sont donc pas identifiables. En effet, les labels peuvent changer d'une composante à une autre au cours des itérations et, par conséquent, on ne peut pas assurer qu'un label particulier représentera toujours la même cinétique durant l'échantillonnage. Ce phénomène, connu sous le terme de *label-switching*, n'est pas spécifique au non paramétrique et est aussi observé dans les modèles paramétriques. Il est donc nécessaire de procéder à un post-traitement du spectre des TACs si l'on veut reconstruire des volumes fonctionnels. Ce post-traitement consiste d'abord à sélectionner un éventail de courbes dans le spectre. Utilisant cet éventail, on peut reconstruire les volumes fonctionnels les plus significatifs. En appliquant ce procédé aux TACs estimées (Figure 5.3), on peut tracer les volumes fonctionnels (Figures 5.5 à 5.11). La Figure 5.4 montre en perspective les volumes fonctionnels utilisés pour la simulation.



(a) (b) (c) (d) (e) (f)

FIGURE 5.4 – Volumes fonctionnels utilisés : (a) total, (b) sang, (c) matière blanche, (d) matière grise, (e) tumeurs et (f) cervelet.

Résultats de reconstruction

Le spectre des cinétiques estimées (Figure 5.3) est en accord avec celui des cinétiques utilisées (Figure 5.2). On peut remarquer qu'au début de l'acquisition la concentration de FDG radioactif est très forte dans le compartiment vasculaire puis diminue très vite, au fur et à mesure que les tissus fixent le traceur (phase de captation). On peut noter aussi que l'évolution de l'activité décroit par la suite dans les autres tissus sauf pour la tumeur.

La Figure 5.3 illustre de la capacité du modèle proposé à capturer la structure des cinétiques. De par le côté non paramétrique, le nombre de composantes est potentiellement infini. La technique du slice sampling adoptée permet de traiter un nombre fini de composantes à chaque itération (nombre de composantes retenues est $K^* \approx 10000$ pour les composantes spatio-temporelles et $J^* \approx 15$ pour les cinétiques). Le nombre de composantes actives est adapté par le modèle, se basant sur les données disponibles ($K_n \approx 4000, J_{K_n} \approx 7$). Par ailleurs, la modélisation proposée n'est basée sur aucune structure de modèle compartimental. Nous ne présumons pas de forme de cinétiques ou de volumes fonctionnels, ni ne faisons d'hypothèses sur l'ordre du modèle. Ce qui permet de s'adapter à des structures de cinétiques complexes et rendant la modélisation appropriée quand le nombre de compartiments tissulaires est inconnu. Signalons tout de même que les fonctions de base temporelles n'ont pas de sens physiologique puisqu'elles ne sont pas identifiables.

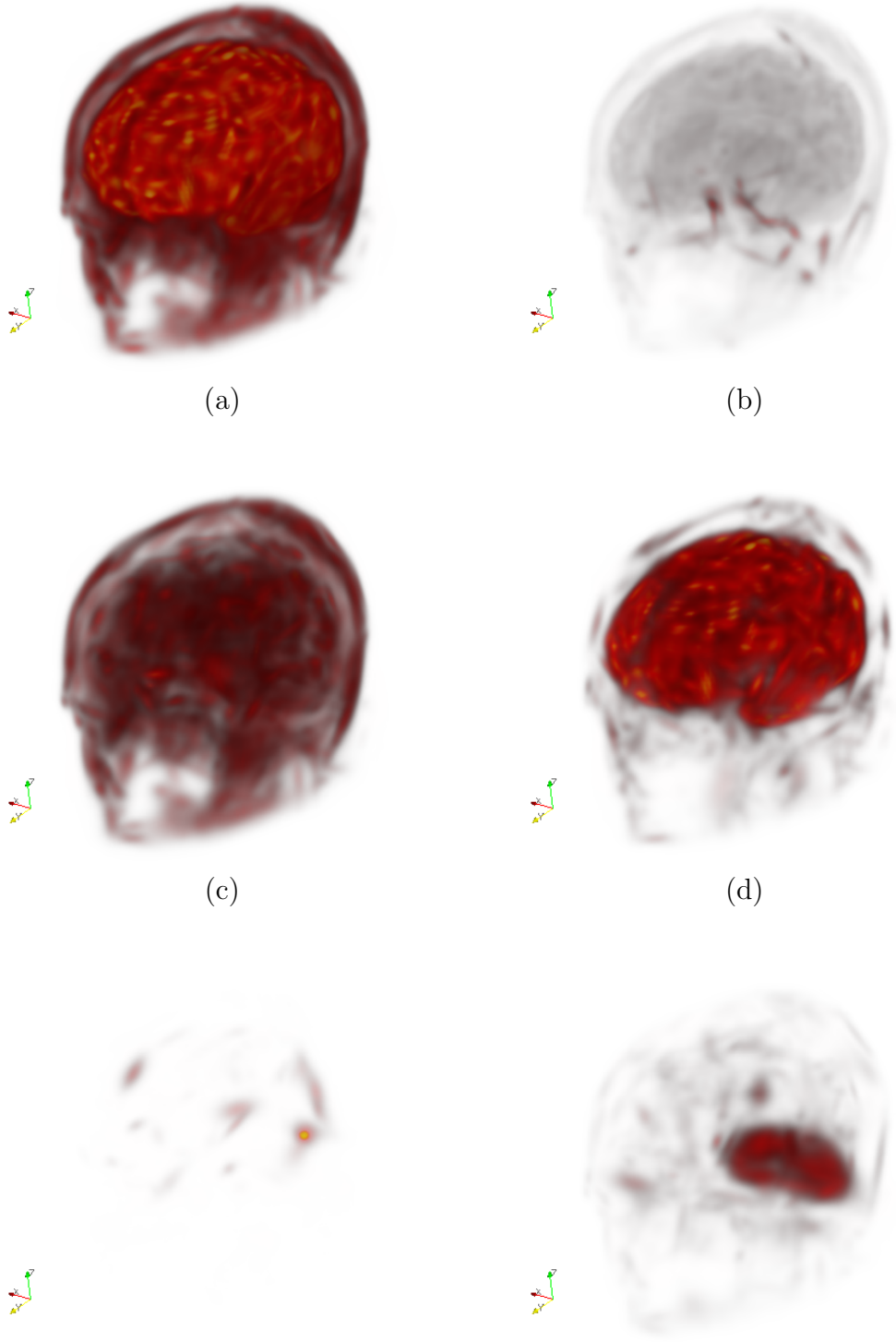
On peut aussi noter sur les volumes fonctionnels obtenus (Figures 5.6 à 5.11) que le rapport signal-bruit est mieux car on exploite les corrélations spatio-temporelles. Toutefois, le contraste pourrait être amélioré, notamment en ce qui concerne la matière blanche et la tumeur. En effet, le clustering des volumes fonctionnels est actuellement basé sur la seule densité temporelle des cinétiques. On pense qu'introduire la variable de concentration temporelle (i.e l'amplitude des cinétiques) comme information discriminante pourrait permettre d'améliorer ce contraste entre les volumes fonctionnels. On pourrait l'envisager par l'utilisation de processus de Dirichlet locaux proposés par Chung et Dunson [CD11].

5.6 Conclusion

Nous avons abordé le problème de reconstruction et de détermination des volumes fonctionnels dans un cadre spatio-temporel continu. Cette technique procède directement à un clustering spatio-temporel des désintégrations dans des domaines continus, sans passer par une phase de discrétisation, et en estime la distribution a posteriori. Cette segmentation fonctionnelle a pour avantage d'être plus reproductible, ce qui peut permettre de pallier à la faible résolution en médecine nucléaire en général et en TEP en particulier.

Une des caractéristiques de l'approche proposée est que la découverte des cinétiques et des volumes fonctionnels est pilotée par les données et l'on ne fait aucune hypothèse sur le nombre et la forme des cinétiques. Par ailleurs, la substitution à la représentation dans la base traditionnelle (voxels et tranches temporelles) par un modèle fonctionnel probabiliste génère un volume de données dépendant uniquement de la quantité d'infor-

mation à représenter. Cette propriété, en plus de la détermination robuste des volumes fonctionnels, fait que l'approche proposée est bien adaptée surtout lorsque le nombre d'observations diminue et par là le rapport signal-sur-bruit des données.

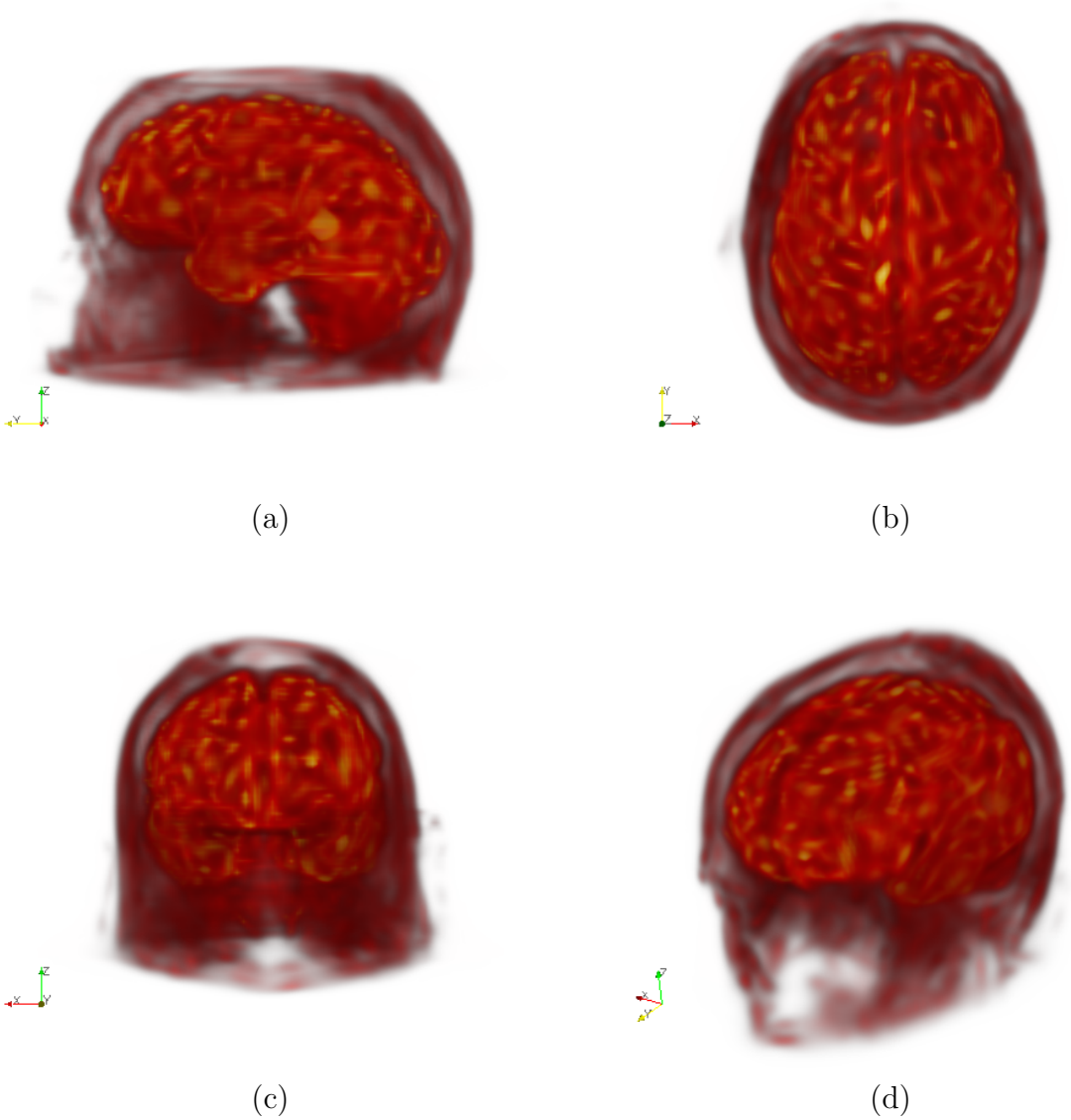


(a) (b)

(c) (d)

(e) (f)

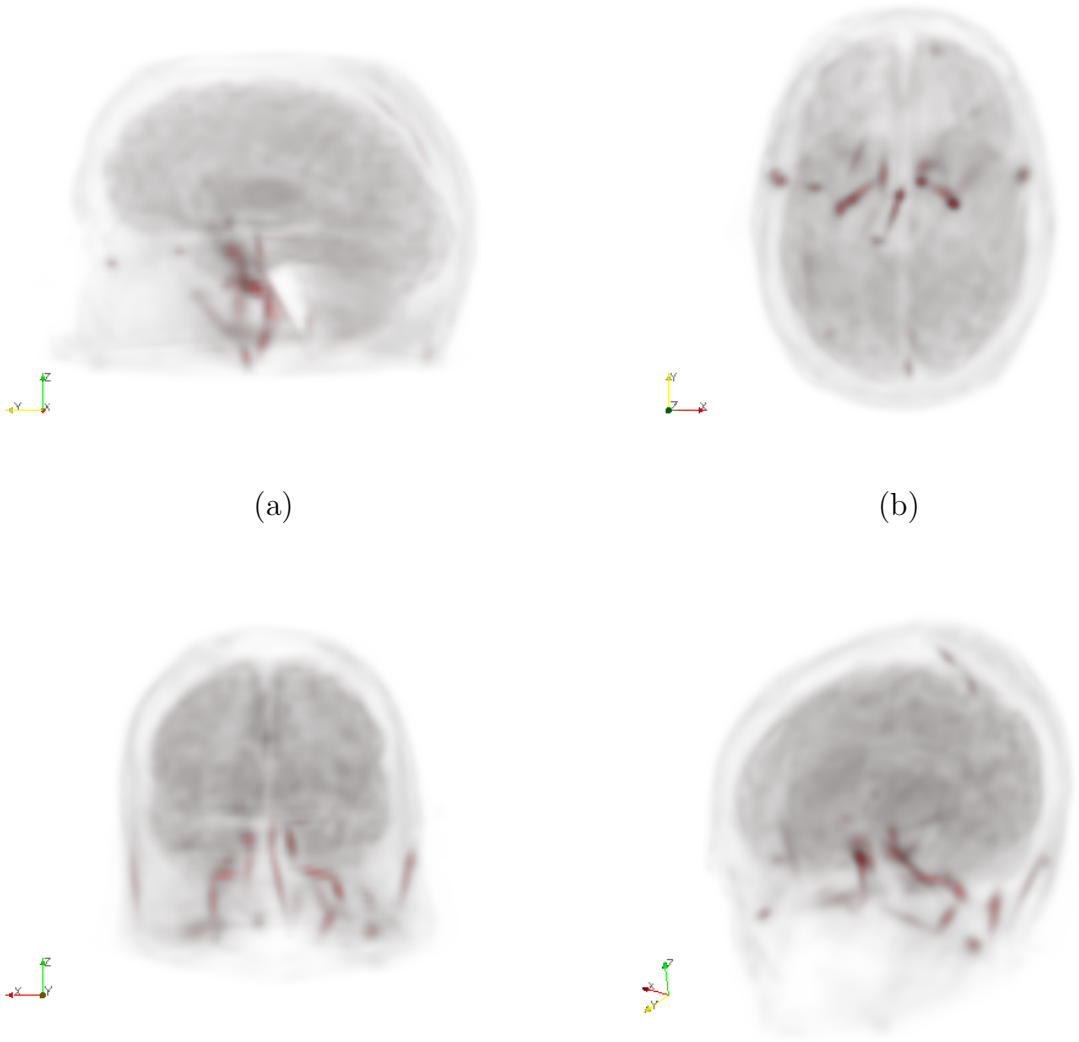
FIGURE 5.5 – Marginales spatiales (volumes fonctionnels estimés) : (a) total, (b) sang, (c)matière blanche, (d) matière grise, (e) tumeurs, (f) cervelet.



(a) (b)

(c) (d)

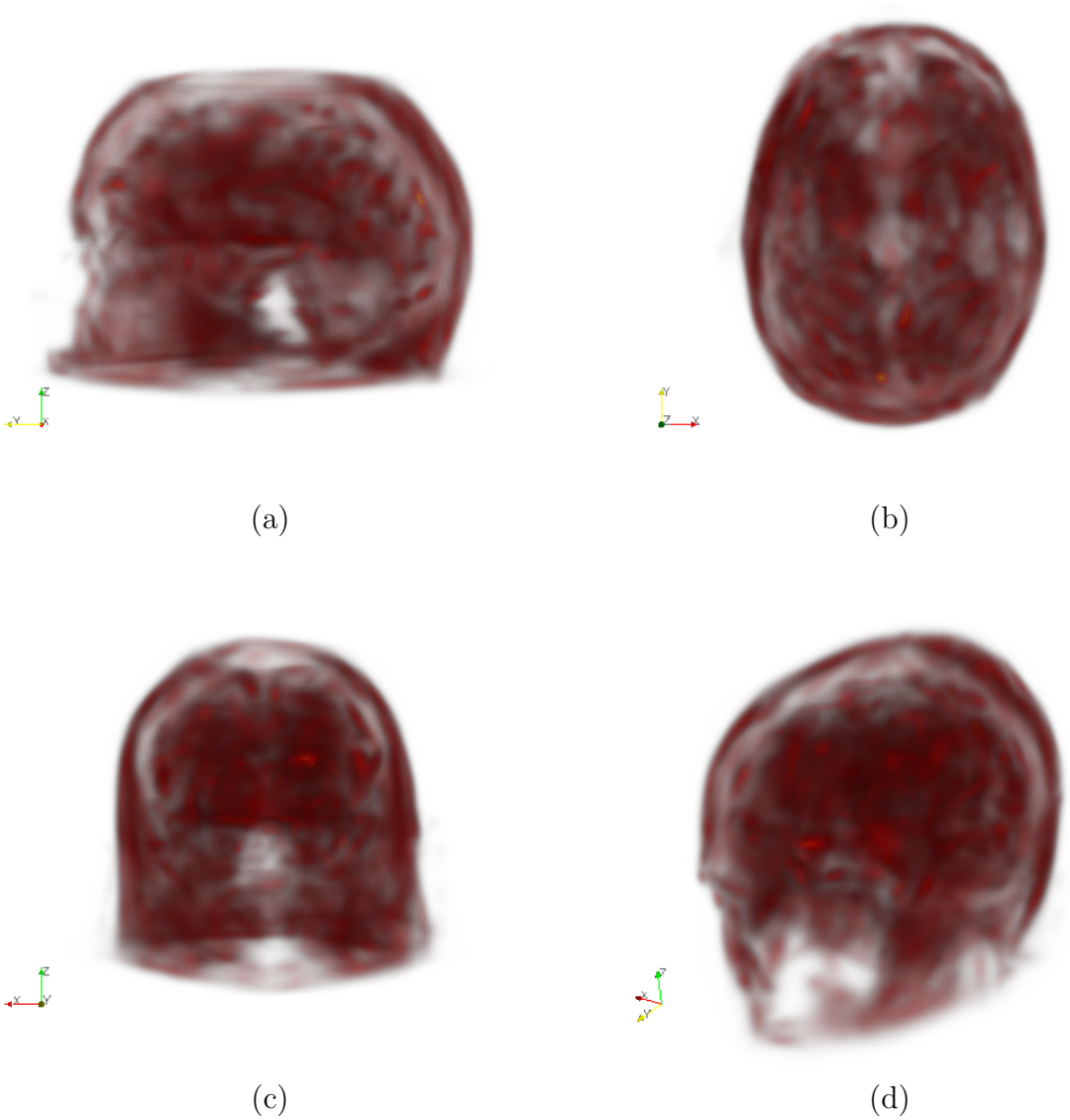
FIGURE 5.6 – Marginale spatiale totale (densité spatiale).



(a) (b)

(c) (d)

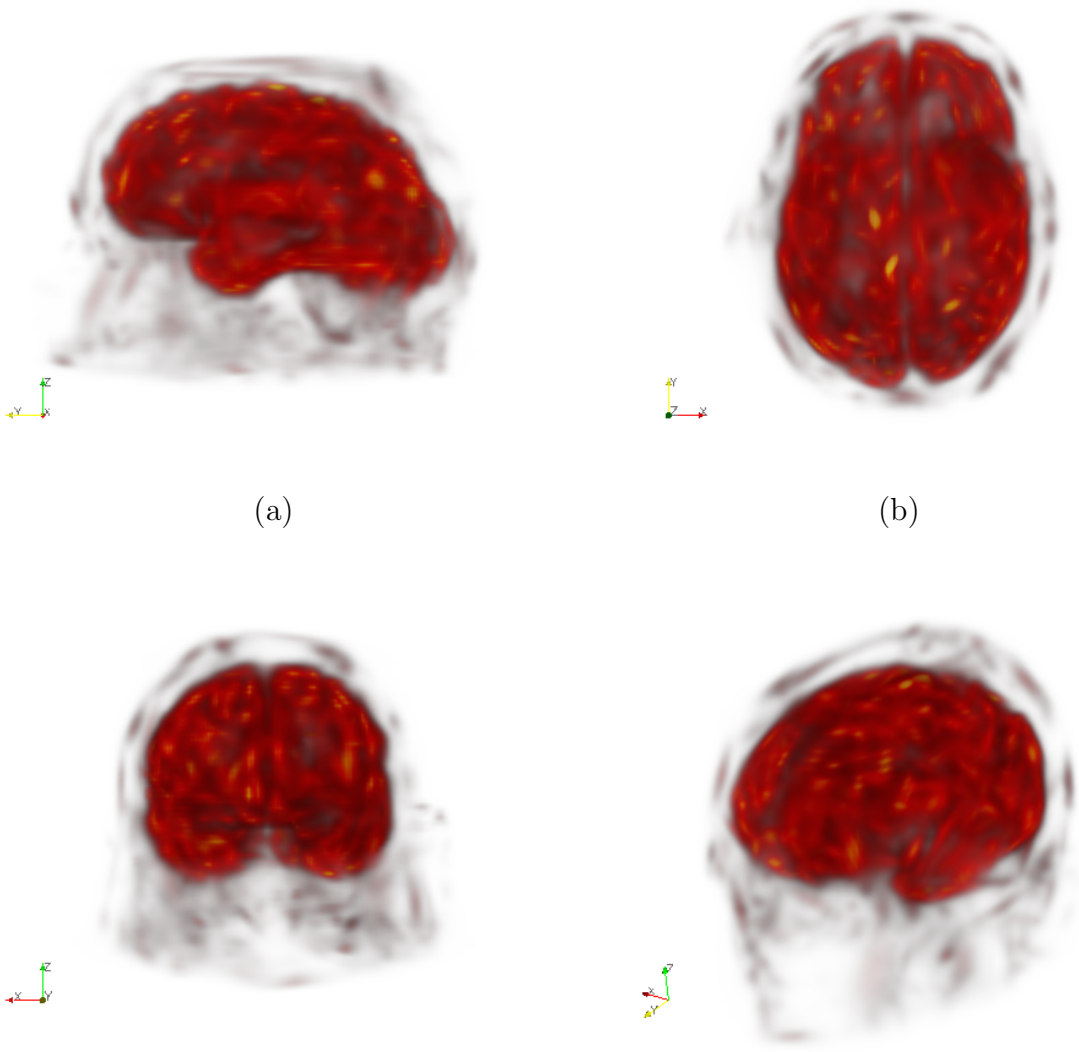
FIGURE 5.7 – Volume fonctionnel estimé pour le sang.



(a) (b)

(c) (d)

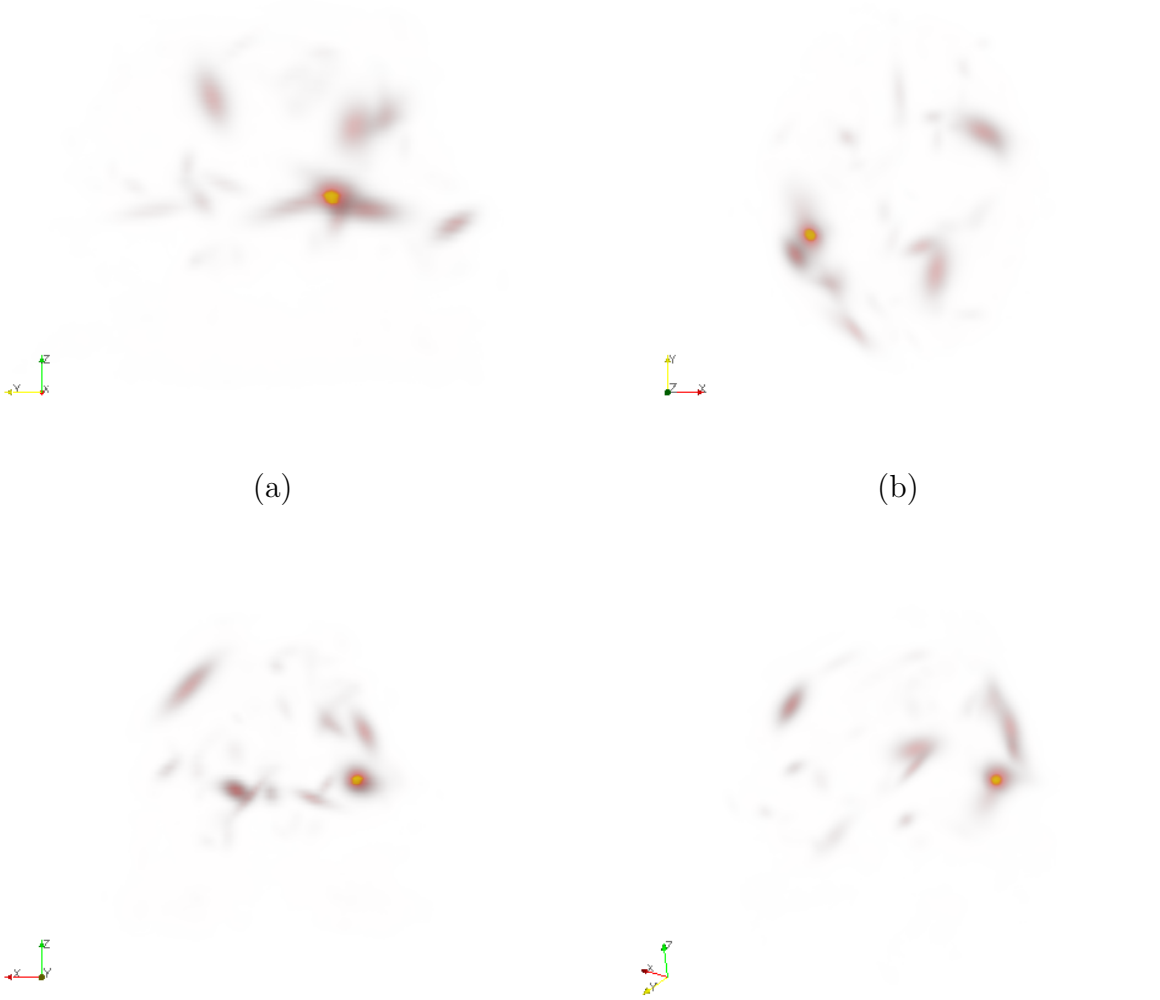
FIGURE 5.8 – Volume fonctionnel estimé pour la matière blanche.



(a) (b)

(c) (d)

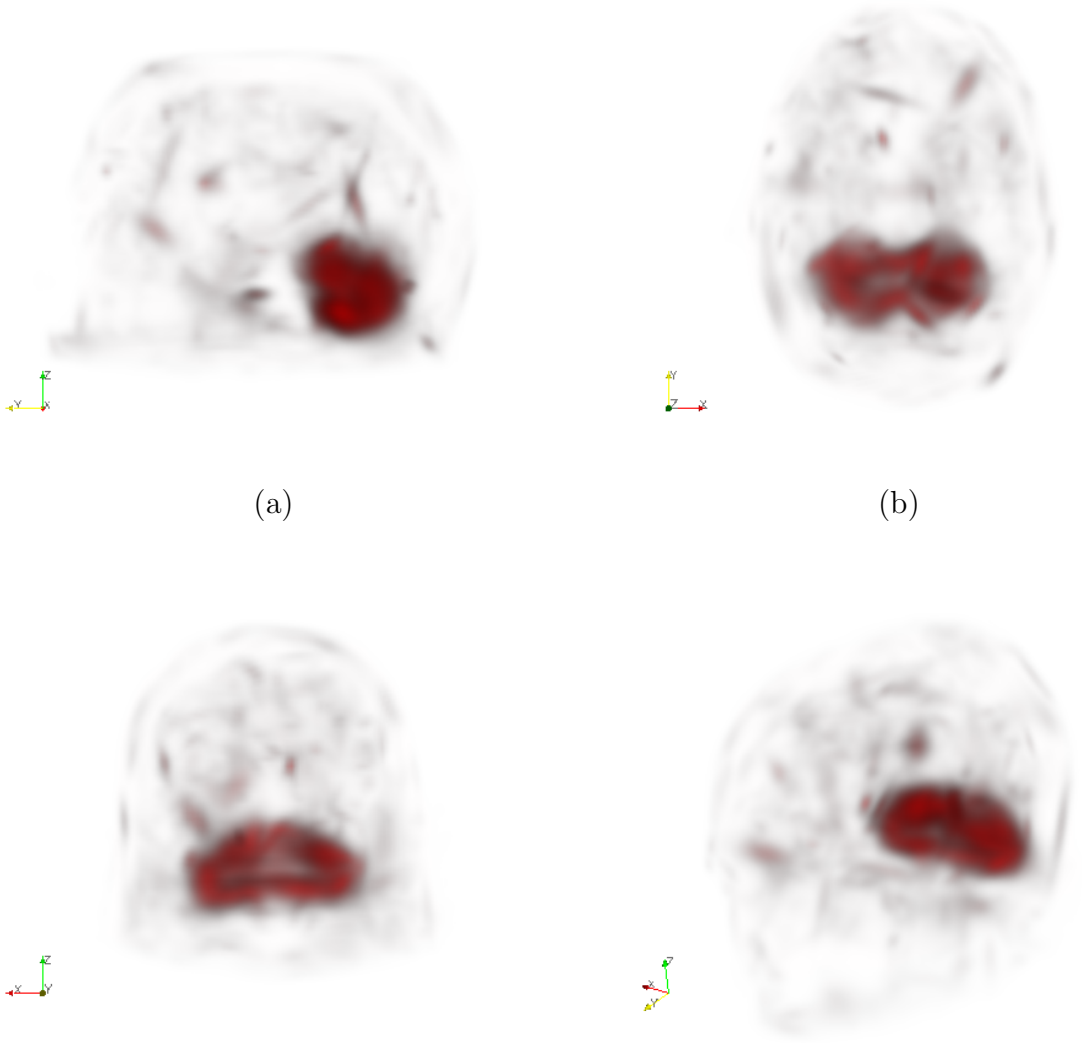
FIGURE 5.9 – Volume fonctionnel estimé pour la matière grise.



(a) (b)

(c) (d)

FIGURE 5.10 – Volume fonctionnel estimé pour la tumeur.



(a) (b)

(c) (d)

FIGURE 5.11 – Volume fonctionnel estimé pour le cervelet.

6

Conclusions et perspectives

6.1 Conclusions

Nous avons proposé, implanté et évalué :

1. une nouvelle méthode d'échantillonnage pour les mélanges par processus de Poisson-Dirichlet à deux paramètres (plus connus sous le nom de mélanges par processus de Pitman-Yor). Cet échantillonneur a été comparé avec trois autres méthodes d'échantillonnage parmi les plus populaires dans la littérature bayésienne non paramétrique.

2. une méthode originale de reconstruction d'images spatiales en trois dimensions (3D) en Tomographie par Emission de Positrons (TEP). L'approche proposée a permis d'adresser le problème de la détermination quantitative de l'erreur de reconstruction. La méthode a été testée sur un fantôme cérébral réaliste et comparée avec les algorithmes de maximum de vraisemblance (implanté dans les machines actuelles en clinique) et de maximum *a posteriori*.

3. une méthode pour la reconstruction d'images spatio-temporelles (3D+t) en TEP proposant un clustering spatio-temporel dans un espace continu. Ce qui est adéquat dans un but de diminution de la dose injectée au patient.

La nouvelle méthode d'échantillonnage présentée dans le chapitre 3 a été développée dans le cadre général du mélange par processus de Pitman-Yor (PYM) et est ainsi applicable à l'un des modèles les plus utilisés dans les statistiques bayésiennes non paramétriques, à savoir les mélanges par processus de Dirichlet (DPM). Notre échantillonneur exhibe à la fois de bonnes propriétés de mélange et offre la possibilité d'accélérer les calculs par une parallélisation sur divers processeurs. Cette propriété est essentielle lorsque l'on travaille avec de gros volumes de données. Ces deux propriétés combinées confèrent de bonnes propriétés théoriques et pratiques à la méthode d'échantillonnage proposée.

En ce qui concerne la TEP (chapitres 4 et 5) nous avons démontré, sur une simulation numérique des examens par Monte Carlo, l'intérêt de l'approche bayésienne non paramétrique pour extraire automatiquement une information pertinente et objective des examens, tout en restant robuste lorsque le nombre de données est relativement faible. Cette diminution de la dose de traceurs injectée est très importante pour des questions de radioprotection car cela induirait une diminution de l'exposition aux rayonnements ionisants pour le patient et son entourage ainsi que pour le personnel médical. Par ailleurs, la TEP est la technique de référence en imagerie moléculaire pour l'évaluation de la réponse thérapeutique en cancérologie, mais aussi en neurologie pour l'étude des maladies d'Alzheimer et de Parkinson, et en épilepsie. Ainsi, si la preuve est démontrée sur données réelles, cela ouvre de bonnes perspectives pour l'application de la méthode proposée. Notons également que l'approche proposée fournit des intervalles de crédibilité qui sont à même de fournir des indicateurs quantitatifs utiles pour le médecin.

6.2 Perspectives

- Les deux variantes de la nouvelle méthode d'échantillonnage que nous avons proposée au chapitre 3 pour les processus de Pitman-Yor fournissent des propriétés de mélange bien meilleures que les algorithmes conditionnels basées sur des *a priori* stick-breaking non-échangeables. Nous avons montré également que l'introduction du seuil que nous proposons permettait d'améliorer le mélange dans la méthode conditionnelle de Kalli *et al.* [KGW09] et que cela serait utile pour l'application à des *a priori* stick-breaking plus généraux que les mélanges par processus de Pitman-Yor ou de Dirichlet. L'une des perspectives concerne la mise en place d'une étape de permutation des labels pour les algorithmes stick-breaking non-échangeables comme suggérée par Porteous *et al.* [PISW06] et Papaspiliopoulos *et al.*[PR07]. Cele devrait permettre d'améliorer la mélangeance de la chaîne MCMC.

- Les perspectives concernant la TEP s'articulent autour de quatre points principaux :

  1. l'estimation des hyperparamètres,
  2. la déconvolution des cinétiques,
  3. l'accélération des calculs,
  4. l'application aux données réelles.

  Nous allons maintenant détailler chacun de ces points.

6.2.1 Estimation des hyperparamètres

On rappelle que le processus de Pitman-Yor $\mathrm{PY}(d, \alpha, G_0)$ est paramétré par deux réels $d \in [0, 1]$, $\alpha > -d$ et une mesure de base G_0 définissant la localisations des clusters. Dans les résultats concernant la reconstruction spatiale (chapitre 4) et la reconstruction spatio-temporelle (chapitre 5), les hyperparamètres α et d ainsi que ceux de la mesure de base $G_0 = \mathcal{NIW}$ ont été fixés. Pour permettre plus de flexibilité dans le modèle, on peut aussi mettre des distributions *a priori* sur ces hyperparamètres et inférer sur leurs

lois *a posteriori*. Dans les mélanges par processus de Pitman-Yor, d règle la croissance asymptotique *a priori* du nombre unique de composantes et α le nombre global de clusters différents. Ces hyperparamètres peuvent être estimés en mettant une loi *a priori* Gamma sur α et une loi uniforme sur d (cf. [Teh06] pour le processus de Pitman-Yor et [Wes92] pour le processus de Dirichlet). Cependant, dans nos applications en TEP, les résultats convenables ont été obtenus pour $\alpha \geq 500$ et des petites valeurs de d. Nous pensons que cela est dû à la trop grande variabilité du nombre de clusters dans le PYP (on a vu à la section 2.3.2 du chapitre 2 que l'augmentation du nombre de clusters *a priori* dans le processus de Pitman-Yor est en $O(\alpha n^d)$ (cf figure 2.5)). Nous pensons donc qu'il est préférable pour la reconstruction tomographique de se restreindre au cas d'un mélange par processus de Dirichlet (i.e $d = 0$ dans le PYM). Nous proposons une loi *a priori* Gamma sur α, de paramètre de forme $k_\alpha = 10$ et d'inverse de paramètre d'intensité $\theta = 100$ par exemple.

Les paramètres du modèle Normal-Inverse Wishart ($\mathcal{NIW}$) contrôlent son comportement et contribuent à la régularisation du problème inverse. Par exemple, $\mathbf{\Sigma}_0$ contrôle la variabilité *a priori* des clusters. Pour les hyperparamètres de ce modèle, deux sont pertinents à estimer, $\mathbf{\Sigma}_0$ et n_0. Si on met une loi *a priori* non-informative sur n_0 ($n_0 > 4$) et que l'on utilise par exemple le modèle suivant pour $\mathbf{\Sigma}_0$,

$$\mathbf{\Sigma}_0 = \frac{n_0 - 4}{n_0} \tau \times \mathbb{I}_3$$

où $\mathbb{I}_3$ désigne la matrice identité de $\mathbb{R}^3$ et τ suit une distribution Gamma *a priori*, alors la loi *a posteriori* conditionnelle de $\mathbf{\Sigma}_0$ est aussi une distribution Gamma et on peut inférer sur elle à partir des données. La loi *a posteriori* conditionnelle de n_0 n'a pas de forme standard et peut être estimée via une étape Metropolis-Hastings.

6.2.2 Déconvolution des cinétiques

Nous avons souligné dans le chapitre 5 que, de par la modélisation non paramétrique proposée, la découverte des cinétiques et des volumes fonctionnels était guidée par les données. Cela permettait une grande flexibilité dans le modèle et aussi une robustesse vis-à-vis des données. Cependant les cinétiques obtenues (TAC) ne sont pas interprétables. D'un autre côté, afin de permettre une interprétation des volumes fonctionnels, il est d'usage d'utiliser des modèles paramétriques compartimentaux traduisant les équations d'évolution des concentrations radioactives. À dessein de joindre la souplesse de l'approche non paramétrique aux possibilités d'interprétation, nous proposons un modèle semi-paramétrique. Cela revient à « envelopper » le modèle compartimental par une couche non paramétrique. L'idée est d'introduire la connaissance *a priori* sur le modèle compartimental sans bien sûr fixer un nombre de compartiments tissulaires ni même d'utiliser une base d'exponentielles décroissantes. Nous proposons d'utiliser la connaissance selon laquelle la cinétique dans un tissu est un mélange de la cinétique sanguine (en tenant compte de la fraction volumique) et de la convolution de la cinétique d'entrée sanguine avec la réponse du tissu (cf. équation 5.24). Les cinétiques principales dans le modèle hiérarchique que nous avons proposé pour les données dynamiques (section 5.5.2) ne sont plus distribuées suivant un arbre de Pólya simple mais suivant un

mélange de la façon suivante :

$$Q^*(t) = \pi_0 Q_0^*(t) + (1-\pi_j)\sum_{j>0}\pi_j \delta_{Q_j^*(t)}$$
$$Q_j^*(t) = Q_0^*(t) \otimes T_j(t)$$

avec

$$Q_0^* \sim \mathrm{PT}\left(\mathcal{A}, Q_0\right)$$
$$T_j^* \sim \mathrm{PT}\left(\mathcal{A}, Q_0\right)$$

où Q_0^* est la cinétique sanguine, T_j est la réponse du $j^{\text{ème}}$ compartiment dans le tissu cible.
Ce modèle permettrait d'isoler la cinétique sanguine et ainsi d'obtenir une estimation non invasive de la fonction d'entrée c'est-à-dire sans prélèvements sanguins et un calcul des paramètres physiologiques d'intérêt.

6.2.3 Accélération des calculs

Un inconvénient de l'approche proposée est le temps de calcul même s'il a été implanté sur une machine parallèle ($\approx$ un jour sur un ordinateur haut-performant). Notons tout de même que contrairement aux approches usuelles, l'absence de discrétisation dans notre méthode a pour conséquence que la complexité numérique ne dépend pas du nombre de voxels mais demeure toujours proportionnelle au nombres de coïncidences observées, quelle que soit la dimension de l'espace de reconstruction. Cette caractéristique est favorable dans le cas d'une faible dose injectée et ne conduit pas à une explosion de la complexité en passant du 3D au 4D contrairement à l'algorithme ML-EM par exemple.

Cependant le temps de calcul, même s'il n'est pas rédhibitoire, demeure une limitation. Cet écueil est dû en grande partie à l'échantillonnage MCMC qui est très coûteux en termes de complexité numérique. Une alternative aux MCMC serait les méthodes bayésiennes variationnelles développées pour le mélange par processus de Dirichlet. Elles consistent à approximer analytiquement les distributions *a posteriori* [BJ06]. Un autre chemin d'investigation serait d'implanter l'algorithme proposé sur des processeurs GPU.

Enfin, une solution que nous avons testée consiste en l'optimisation de l'algorithme MCMC de reconstruction proposé, par l'échantillonnage en un bloc de l'étape de proposition de localisation des annihilations et de l'affectation de l'annihilation à une des composants du mélange (cf. section 4.3.3 par exemple). En effet, en pratique ces deux étapes constituent la partie charnière du temps de calcul qui est ainsi proportionnel au nombre d'évènements multiplié par le nombre de composantes. Nous proposons donc l'échantillonnage dans un bloc regroupant les étapes de proposition de localisation et de l'affectation. D'après la règle de Bayes, on peut écrire la loi *a posteriori* conditionnelle de $(\boldsymbol{c}, \mathbf{X}|\mathbf{w}, \mathbf{u}, \boldsymbol{\Theta}^*, \mathbf{y})$ de la façon suivante :

$$c_i, \boldsymbol{x}_i | y_i, u_i, \mathbf{w}, \boldsymbol{\Theta}^* \overset{\text{ind}}{\sim} \mathcal{P}(y_i|\boldsymbol{x}_i) \times \widetilde{G}_i(c_i, \boldsymbol{x}_i)$$

avec

$$f_{\widetilde{G}_i}(c_i, \boldsymbol{x}_i) \propto \sum_{k=1}^{K^*} \mathbf{1}\left(w_k > u_i\right) \max\left(w_k, \zeta\right) f_{\mathcal{N}}\left(\boldsymbol{x}_i | \boldsymbol{\theta}_k^*\right) \delta_k\left(c_i\right)$$

où $\mathcal{P}(y_i|\boldsymbol{x}_i)$ est la distribution de projection. Soit G_i^* la distribution candidate pour $i = 1, \ldots, n$. À chaque itération $t+1$, on génère $(c_i^*, \boldsymbol{x}_i^*) \sim G_i^*$ avec

$$f_{G_i^*}(c_i^*, \boldsymbol{x}_i^*) \propto f_{\mathcal{N}}(\boldsymbol{x}_i^*|\mathbf{m}_{y_i}^*, \boldsymbol{\Sigma}_{y_i}^*) \times f_{\widetilde{G}_i}(c_i^*, \boldsymbol{x}_i^*).$$

où $\mathcal{N}(\boldsymbol{x}_i^*|\mathbf{m}_{Y_i}^*, \boldsymbol{\Sigma}_{y_i}^*)$ est la Gaussienne dont les paramètres sont choisis pour approximer le projecteur dans la LOR. On peut alors ré-écrire $f_{G_i^*}$ comme un mélange re-pondéré de lois normales

$$f_{G_i^*}(c_i^*, \boldsymbol{x}_i^*) \propto \sum_{k=1}^{K^*} \mathbf{1}\left(w_k > u_i\right) \max\left(w_k, \zeta\right) \widetilde{\eta}_{k,y_i} \times f_{\mathcal{N}}\left(\boldsymbol{x}_i^*|\widetilde{\mathbf{m}}_{k,y_i}, \widetilde{\boldsymbol{\Sigma}}_{k,y_i}\right) \delta_k\left(c_i^*\right)$$

L'échantillonnage s'effectue de la façon suivante pour tout $i \leq n$, à l'itération $t+1$.

– Échantillonner indépendamment c_i^* tel que

$$c_i^* \overset{\text{ind}}{\sim} \sum_{k=1}^{K^*} \widetilde{w}_{k,i} \delta_k(\cdot)$$

avec

$$\widetilde{w}_{k,i} \propto \mathbf{1}\left(w_k > u_i\right) \max\left(w_k, \zeta\right) \widetilde{\eta}_{k,y_i}$$

et $\sum_{k=1}^{K^*} \widetilde{w}_{k,i} = 1$.

– Échantillonner indépendamment $\boldsymbol{x}_i^*$ tel que

$$\boldsymbol{x}_i^* \overset{\text{ind}}{\sim} \mathcal{N}\left(\boldsymbol{x}_i^*|\widetilde{\mathbf{m}}_{c_i^*,y_i}, \widetilde{\boldsymbol{\Sigma}}_{c_i^*,y_i}\right).$$

– Calculer le taux d'acceptation $\omega((c_i^*, \boldsymbol{x}_i^*)'; (c_i^{(t)}, \boldsymbol{x}_i^{(t)})')$

$$\omega((c_i^*, \boldsymbol{x}_i^*)'; (c_i^{(t)}, \boldsymbol{x}_i^{(t)})') = \frac{\Pr(y_i|\boldsymbol{x}_i^*)}{\Pr(y_i|\boldsymbol{x}_i^{(t)})} \cdot \frac{f_{\mathcal{N}}\left(\boldsymbol{x}_i^{(t)}|\mathbf{m}_{y_i}^*, \boldsymbol{\Sigma}_{y_i}^*\right)}{f_{\mathcal{N}}\left(\boldsymbol{x}_i^*|\mathbf{m}_{y_i}^*, \boldsymbol{\Sigma}_{y_i}^*\right)} = \omega(\boldsymbol{x}_i^*; \boldsymbol{x}_i^{(t)}).$$

– Accepter $(c_i^*, \boldsymbol{x}_i^*)$ (i.e faire $(c_i^{(t+1)}, \boldsymbol{x}_i^{(t+1)}) = (c_i^*, \boldsymbol{x}_i^*)$, sinon faire $(c_i^{(t+1)}, \boldsymbol{x}_i^{(t+1)}) = (c_i^{(t)}, \boldsymbol{x}_i^{(t)})$ avec la probabilité

$$\min(1,\ \omega(\boldsymbol{x}_i^*; \boldsymbol{x}_i^{(t)})).$$

Cette étape d'échantillonnage par blocs permet de réduire le temps de calcul d'au moins 30%.

6.2.4 Application aux données réelles

Le travail en cours consiste en l'application aux données réelles. Cela nécessite au préalable une prise en compte des phénomènes physiques alourdissant la modélisation statistique comme les évènements fortuits et diffusés (cf. section 4.1.3 sur la correction des phénomènes physiques). L'approche proposée demande à ce qu'aucun pré-traitement ne soit effectué sur les données d'acquisition. Ce bruit de fond doit alors être inclus dans la modélisation. Cependant, il est d'usage en TEP de considérer la distribution des diffusés comme connue après une (ou plusieurs) reconstructions préliminaires. De même, on peut estimer la distribution des fortuits (par la méthode des fenêtres décalées

par exemple). On peut alors supposer que la distribution des coïncidences diffusés et fortuites est connue et est donnée par $\mathcal{B}(y,t)$. La distribution F dans le modèle 5.8 est remplacée par $\mathcal{F}$ qui s'écrit alors

$$\mathcal{F}(y,t) = \pi F(y,t) + (1-\pi)\mathcal{B}(y,t)$$

où

$$F(y,t) = \int_{\mathcal{X}} \mathcal{P}(\cdot|\boldsymbol{x})\; G(\mathrm{d}\boldsymbol{x},t)$$

et $\pi \in]0,1]$. On peut définir pour $i = 1,\ldots,n$ une variable indicatrice B telle que $B_i = 0$ si la coïncidence est vraie (i.e $Y_i \sim F$) et $B_i = 1$ sinon (i.e $Y_i \sim \mathcal{B}$). Dans l'inférence, on doit effectuer à chaque itération de l'algorithme MCMC un tirage Bernouilli pour tester si l'évènement est vrai ou non. La distribution conditionnelle *a posteriori* de B_i s'écrit

$$B_i \sim \text{Bernouilli}\left(\frac{(1-\pi)\mathcal{B}(y,t)}{\pi\widetilde{F}(y,t) + (1-\pi)\mathcal{B}(y,t)}\right)$$

où

$$\widetilde{F}(y,t) = \int_{\mathcal{X}} \mathcal{P}(\cdot|\boldsymbol{x})\; \widetilde{G}(\mathrm{d}\boldsymbol{x})$$

avec

$$f_{\widetilde{G}_i}(\boldsymbol{x}_i) \propto \sum_{k=1}^{K^*} \mathbf{1}(w_k > u_i) \max(w_k,\zeta)\, f_{\mathcal{N}}\left(\boldsymbol{x}_i|\widetilde{\boldsymbol{\theta}}_k\right) f_{Q^*_{d_k}}(t_i)\,.$$

Le calcul de l'intégrale $\widetilde{F}(y,t)$ n'a pas de forme explicite. On peut l'approximer par échantillonnage d'importance en prenant comme loi d'importance le mélange re-pondéré de densités normales utilisé pour la proposition de localisation de l'annihilation (cf. sections 5.5.5 et 4.3.3),

$$f_{G^*_i}(\boldsymbol{x}^*_i) \propto \sum_{k=1}^{K^*} \mathbf{1}(w_k > u_i) \max(w_k,\zeta)\, \widetilde{\eta}_{k,y_i}\, f_{\mathcal{N}}\left(\boldsymbol{x}^*_i|\widetilde{\mathbf{m}}_{k,y_i}, \widetilde{\boldsymbol{\Sigma}}_{k,y_i}\right) f_{Q^*_{d_k}}(t_i).$$

Résultats préliminaires sur données réelles

Nous présentons ici des résultats de reconstruction préliminaires sur données réelles, avec prise en compte de la correction des coïncidences diffusées et fortuites. Les données sont acquises sous le format liste par la caméra HRRT (Siemens Healthcare, Knoxville, TN, USA), pour une durée d'examen de 90 min. Le traceur utilisé est le PIB (Pittsburgh Compound-B), un analogue de la thioflavine T, utilisé en recherche clinique en TEP cérébrale pour l'étude de la maladie d'Alzheimer. Il est marqué au ^{11}C et noté [^{11}C]−PIB.

Afin d'évaluer la technique de reconstruction proposée (BNP) dans un contexte spécifique de faible dose, les données du fichier list-mode ont été aléatoirement décimées d'un facteur 5, conduisant à une dose artificiellement réduite de 5. Pour la reconstruction BNP, 7000 itérations ont été effectuées dont 1000 écartées pour la période de burn-in. L'estimateur est l'espérance *a posteriori* sur les tirages post burn-in. Quant à la reconstruction ML-EM avec le code constructeur, la dose n'a pas été réduite et les images obtenues après convergence ont été post filtrées par un filtre Gaussien (LTMH=3 mm). La figure 6.1 montre les résultats pour la méthode BNP (dose ÷5) et la reconstruction ML-EM (dose ÷1).

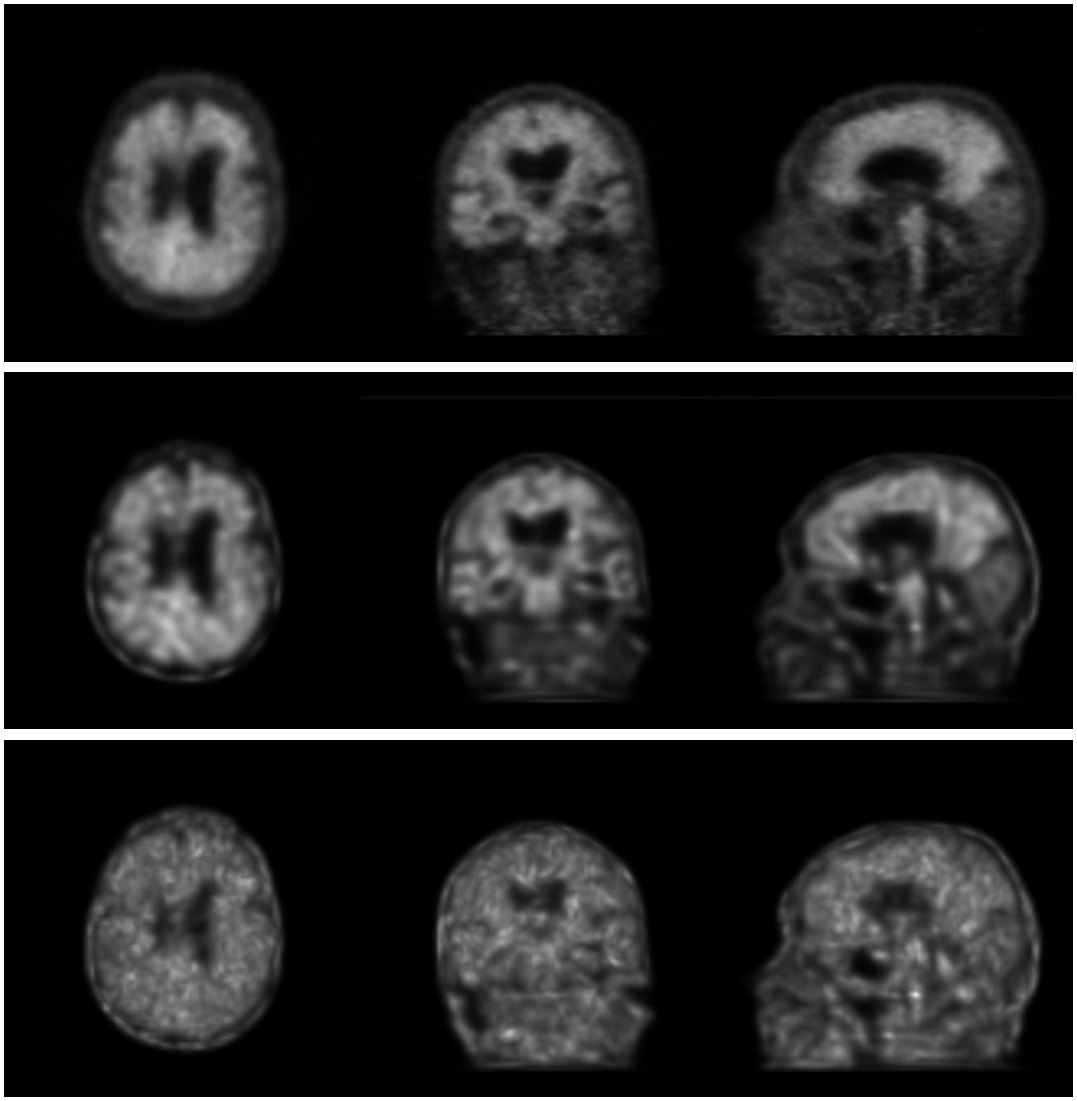


FIGURE 6.1 – *Lignes :* haut : reconstruction constructeur ML-EM, (dose ÷1), milieu : reconstruction BNP (dose ÷5), bas : écart-type BNP (dose ÷5).
Colonnes : vues axiale, coronale et sagittale.

A

Rappels sur la loi de Dirichlet

A.1 La distribution de Dirichlet

La loi de Dirichlet est une distribution sur le simplexe

$$S_K = \{(\pi_1, \ldots, \pi_K) : \ \pi_k \geq 0 \text{ et } \sum_k \pi_k = 1\}.$$

Soient K variables aléaloires $\boldsymbol{\pi} = (\pi_1, \ldots, \pi_K) \sim \text{Dir}(\alpha_1, \ldots, \alpha_K)$ avec ($\alpha_k > 0$ pour tout k). Alors la densité est donnée par

$$p(\boldsymbol{\pi}|\boldsymbol{\alpha}) = \frac{\Gamma(\sum_{k=1}^K \alpha_k)}{\prod_{k=1}^K \Gamma(\alpha_k)} \prod_{k=1}^K \pi_k^{\alpha_{k-1}}$$

où $\Gamma(\cdot)$ réprésente la fonction gamma. Pour $K = 2$, on retrouve la loi Beta.

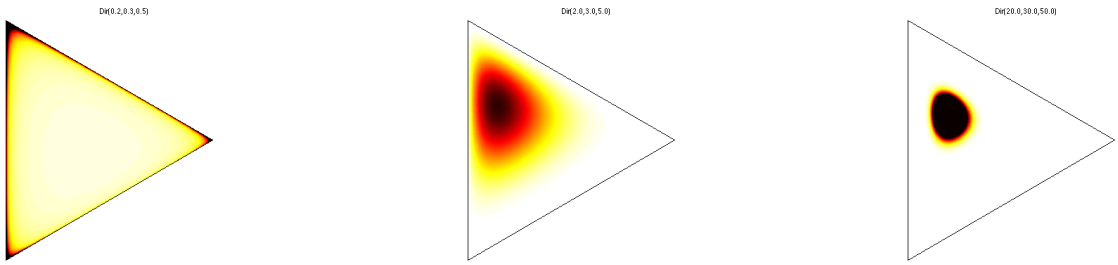
A.1.1 Moments

Espérances

$$\mathbb{E}(\pi_k) = \frac{\alpha_k}{\sum_{j=1}^K \alpha_j} \tag{A.1}$$

$$\mathbb{E}(\pi_k^2) = \frac{\alpha_k(1+\alpha_k)}{\left(\sum_{j=1}^K \alpha_j\right)\left(1+\sum_{j=1}^K \alpha_j\right)} \tag{A.2}$$

$$\mathbb{E}(\pi_j, \pi_k) = \frac{\alpha_j \alpha_k}{\left(\sum_{j=1}^K \alpha_j\right)\left(1+\sum_{j=1}^K \alpha_j\right)} \tag{A.3}$$



1) 2) 3)

FIGURE A.1 – Cartes de couleur de la densité de Dirichlet pour $K = 3$ et différents jeux de paramètres. De 1) à 3) : $\boldsymbol{\alpha} = (0.2, 0.3, 0.5), (2, 3, 5), (20, 30, 50)$. Plus le paramètre d'échelle est grand, plus les échantillons sont concentrés autour de la moyenne.

Variance, covariance

$$\mathbb{V}(\pi_k) = \frac{\alpha_k \left((\sum_{j=1}^K \alpha_j) - \alpha_k \right)}{\left(\sum_{j=1}^K \alpha_j \right)^2 \left((\sum_{j=1}^K \alpha_j) + 1 \right)} = \frac{\mathbb{E}[\pi_k]\,(1 - \mathbb{E}[\pi_k])}{\left(\sum_{j=1}^K \alpha_j \right) + 1}. \tag{A.4}$$

$$\mathrm{Cov}(\pi_j, \pi_k) = \frac{-\alpha_j \alpha_k}{\left(\sum_{j=1}^K \alpha_j \right)^2 \left((\sum_{j=1}^K \alpha_j) + 1 \right)}. \tag{A.5}$$

La distribution de Dirichlet est illustrée à la Figure A.1 pour $K = 3$ et pour trois jeux de paramètres. Les échantillons sont générés dans le simplexe définis par le triangle. Les paramètres ont la même amplitude relative, le paramètre d'échelle $\alpha_0 = \sum_{k=1}^K \alpha_k$ contrôle la dispersion des échantillons générés autour de la moyenne.

A.1.2 Propriétés

Agrégation

Si $(\pi_1, \ldots, \pi_K) \sim \mathrm{Dir}(\alpha_1, \ldots, \alpha_K)$ alors $(\pi_1, \ldots, \pi_i + \pi_j, \ldots, \pi_K) \sim \mathrm{Dir}(\alpha_1, \ldots, \alpha_i + \alpha_j, \ldots, \alpha_K)$. On peut en tirer la loi marginale de chaque π_k,

$$\pi_k \sim \mathrm{Beta}(\alpha_k, \sum_{j=1}^K \alpha_j). \tag{A.6}$$

Décimation

Soit

$$(\pi_1, \ldots, \pi_K) \sim \mathrm{Dir}(\alpha_1, \ldots, \alpha_K)$$
$$(\tau_1, \tau_2) \sim \mathrm{Dir}(\alpha_1\beta_1, \alpha_1\beta_2) \text{ avec } \beta_1 + \beta_2 = 1$$
$$\text{alors } (\pi_1\tau_1, \pi_1\tau_2, \pi_2, \ldots, \pi_K) \sim \mathrm{Dir}(\alpha_1\beta_1, \alpha_1\beta_2, \alpha_2, \ldots, \alpha_K).$$

Construction par normalisation de variables gamma

Soit Gamma(α, β) désignant la loi gamma de paramètre de forme $\alpha \geq 0$ et de paramètre d'échelle $\beta > 0$. Si $\alpha = 0$, la loi est dégénérée en 0 et si $\alpha > 0$ alors la densité par rapport à la mesure de Lebesgue sur $\mathbb{R}$ est

$$f(z|\alpha, \beta) = \frac{1}{\Gamma(\alpha)\beta^\alpha} e^{-z/\beta} z^{\alpha-1} \mathbf{1}_{[0,+\infty]}(z).$$

Considèrons K variables aléatoires indépendantes $Z_1, \ldots, Z_k$ telles que pour tout j, $Z_j \sim$ Gamma$(\alpha_j, 1)$ (avec $\alpha_j > 0$ pour tout j). On a alors $\sum_{i=1}^K Z_i \sim$ Gamma$(\sum_{i=1}^K \alpha_i, 1)$. Soit la séquence $(\pi_1, \ldots, \pi_K)$ construite de la façon suivante : pour $j = 1, \ldots, K$, $\pi_j = \frac{Z_j}{\sum_{i=1}^K Z_i}$. Alors

$$(\pi_1, \ldots, \pi_K) \sim \text{Dir}(\alpha_1, \ldots, \alpha_K).$$

Conjugaison

La distribution de Dirichlet est conjuguée à la loi multinomiale

$$\begin{aligned} \boldsymbol{\pi} = (\pi_1, \ldots, \pi_K) &\sim \text{Dir}(\alpha_1, \ldots, \alpha_K) \\ c_1, \ldots, c_n | \boldsymbol{\pi} &\sim \text{Mult}(\boldsymbol{\pi}) \\ \text{alors } (\pi_1, \pi_2, \ldots, \pi_K) | \boldsymbol{c} &\sim \text{Dir}(\alpha_1 + n_1, \ldots, \alpha_K + n_K) \text{ où } n_k = \#\{i : c_i = k\}. \end{aligned}$$

On peut alors marginaliser les poids en utilisant l'intégrale de Dirichlet (équation A.7) pour définir la probabilité de l'observation $n + 1$ sachant les n précédentes, c'est la distribution prédictive.

Distribution prédictive

$$\mathbb{P}(c_{n+1} = k | c_1, \ldots, c_n) = \frac{\alpha_k + n_k}{\sum_{j=1}^K \alpha_j + n}.$$

A.2 Intégrale de Dirichlet

Si $\boldsymbol{\pi} = (\pi_1, \ldots, \pi_K) \sim \text{Dir}(\alpha_1, \ldots, \alpha_K)$ alors

$$\int \pi_1^{\alpha_1 - 1}, \ldots, \pi_K^{\alpha_K - 1} \mathrm{d}\pi_1 \ldots \mathrm{d}\pi_K = \frac{\prod_{j=1}^K \Gamma(\alpha_j)}{\Gamma\left(\sum_{j=1}^K \alpha_j\right)}. \tag{A.7}$$

B

Chaînes de Markov, convergence de MH et Gibbs

Dans cette annexe, nous rappelons d'abord quelques notions théoriques et résultats de convergence sur les chaînes de Markov. Puis, nous abordons la convergence de l'algorithme de Metropolis-Hastings (MH) et de l'échantillonneur de Gibbs.

B.1 Les chaînes de Markov

Définition 11. *Soit $(\mathcal{X}, \mathcal{B}, P)$ un espace probabilisé et $(X^{(n)})$ une séquence de variables aléatoires à valeurs dans $\mathcal{X}$. Les éléments de la séquence $(X^{(n)})$ forment une chaîne de Markov ayant pour espace d'états $\mathcal{X}$ si les transitions successives sont statistiquement indépendantes. Autrement dit, la loi conditionnelle de $X^{(n)}|X^{(n-1)}, X^{(n-2)} \ldots, X^{(0)}$ est la même que celle de $X^{(n)}|X^{(n-1)}$:*

$$P(X^{(n)} \in A_n | X^{(n-1)} \in A_{n-1}, \ldots, X^{(0)} \in A_0) = P(X^{(n)} \in A_n | X^{(n-1)} \in A_{n-1}) \quad \text{(B.1)}$$

pour tous éléments $A_0, \ldots, A_n$ de la tribu $\mathcal{B}$.

L'espace d'états $\mathcal{X}$ considéré sera l'espace euclidien $\mathbb{R}^d$ (ou un sous-ensemble de $\mathbb{R}^d$) muni de la tribu borélienne $\mathcal{B} = \mathcal{B}(\mathbb{R}^d)$.

B.1.1 Homogénéité

Les chaînes de Markov vérifient souvent l'hypothèse d'homogénéité c'est-à-dire que la loi de $X^{(n)}|X^{(n-1)}$ ne dépend pas de n :

$$P(X^{(n)} \in B | X^{(n-1)} \in A) = P(X^{(m)} \in B | X^{(m-1)} \in A) \text{ pour tous } m, n \in \mathbb{N}^*. \quad \text{(B.2)}$$

B.1.2 Noyau de transition

Définition 12. *Une fonction $P : \mathcal{X} \times \mathcal{B} \to [0,1]$ est appelée noyau de transition sur $\mathcal{X}$ si les deux propriétés suivantes sont vérifiées :*
- *pour tout $x \in \mathcal{X}$, $A \in \mathcal{B} \to P(x,A)$ est une mesure de probabilité sur $(\mathcal{X},\mathcal{B})$.*
- *pour tout $A \in \mathcal{B}$, $x \to P(x,A)$ est mesurable.*

Le noyau de transition est une distribution conditionnelle qui représente la probabilité de passer du point x à un ensemble A, $P(x,A) \equiv P(A|x)$. Il régit donc la génération de $X^{(n+1)}$ sachant $X^{(n)}$. Étant une fonction de probabilité sur $\mathcal{X}$, le noyau vérifie $P(x,\mathcal{X}) = 1$. Par ailleurs, $P(x,\{x\})$ n'est pas nécessairement nulle. En d'autres termes, il est permis que la chaîne reste au point x.

Si la mesure $P(x,\cdot)$ est à densité, on appelle souvent, par abus de langage, noyau de transition la densité $K(x,y)$ du noyau $P(x,\cdot)$ tel que :

$$\forall A \in \mathcal{B},\ P(x,A) = \int_A K(x,y)\mathrm{d}y. \tag{B.3}$$

Nous nous intéresserons aux chaînes continues, c'est-à-dire telles que le noyau de transition admette une densité par rapport à la mesure de Lebesgue.

Noyau de transition d'une chaîne de Markov homogène
Une chaîne de Markov homogène $(X^{(n)})$ est entièrement déterminée par la donnée d'une distribution initiale π_0 (ou d'un point de départ $X^{(0)}$) et du noyau de transition

$$P(x,A) = P(X^{(n+1)} \in A | X^{(n)} = x). \tag{B.4}$$

Noyau de n transitions
On appelle noyau de n transitions ($n > 1$) de la chaîne $(X^{(k)})$,

$$P^n(x,A) = \int_{\mathcal{X}} P^{n-1}(y,A)P(x,\mathrm{d}y) = P(X^{(n)} \in A | X^{(0)} = x) \tag{B.5}$$

où $P^1(x,A) = P(x,A)$.

Proposition 1. *Équations de Chapman-Kolmogorov*
Soit $P(\cdot,\cdot)$ un noyau de transition sur $(\mathcal{X},\mathcal{B})$ et $P^n(\cdot,\cdot)$ le noyau de n transitions associé. Alors pour tous $(m,n) \in \mathbb{N}^2$, $x \in \mathcal{X}$ et $A \in \mathcal{B}$,

$$P^{m+n}(x,A) = \int_{\mathcal{X}} P^n(y,A)P^m(x,\mathrm{d}y).$$

B.1.3 Distribution invariante

Si une chaîne de Markov $(X^{(n)})$ converge, la loi de $(X^{(n)})$ doit être indépendante de n à partir d'un entier n assez grand. Il doit alors exister une distribution π telle que si $X^{(n)} \sim \pi$ alors $X^{(n+1)} \sim \pi$.

Définition 13. *Soit π^* une mesure σ-finie sur $(\mathcal{X}, \mathcal{B})$. On dit que la chaîne de noyau de transition $P(\cdot,\cdot)$ est π^*-invariante si pour tout sous-ensemble mesurable* dy *on a :*

$$\pi^*(\mathrm{d}y) = \int_{\mathcal{X}} P(x, \mathrm{d}y)\pi^*(\mathrm{d}x). \text{ [1]} \tag{B.6}$$

Si cette condition est satisfaite, π^* est dite distribution invariante de la chaîne.

Si la mesure π^* a pour densité $\pi(y)$ par rapport à la mesure de Lebesgue sur $\mathcal{X} = \mathbb{R}^d$, alors

$$\pi^*(\mathrm{d}y) = \pi(y)\mathrm{d}y = \int_{\mathcal{X}} P(x, \mathrm{d}y)\pi(x)\mathrm{d}x. \tag{B.7}$$

Dans la suite, on considèrera que π^* est une mesure de probabilité admettant une densité π et on parlera de π-invariance de la chaîne. En fait, la distribution invariante sera donnée par la loi cible dans les algorithmes MCMC.

La π-invariance indique que si $X^{(0)} \sim \pi$ alors $X^{(n)} \sim \pi$ pour tout $n \geq 1$, d'où le nom de distribution stationnaire.

Il est souvent difficile de vérifier directement la propriété d'invariance (équation B.7) pendant la construction d'un algorithme MCMC. On peut alors vérifier une condition suffisante : la condition de réversibilité appelée aussi condition de balance détaillée ou condition d'équilibre.

Proposition 2. *Soit $K(x, y)$ la densité du noyau de transition $P(x,\cdot)$ d'une chaîne de Markov donnée. Si $K(x, y)$ satisfait la condition de balance détaillée, i.e si*

$$K(x, y)\pi(x) = K(y, x)\pi(y)$$

alors la chaîne admet π comme distribution invariante.

Démonstration. Soit $K(x, y)$ satisfaisant la condition de balance détaillée. Alors,

$$\int_{\mathcal{X}} K(x, y)\pi(x)\mathrm{d}x = \int_{\mathcal{X}} K(y, x)\pi(y)\mathrm{d}x = \pi(y)\int_{\mathcal{X}} K(y, x)\mathrm{d}x = \pi(y)$$

π est donc bien la distribution invariante de la chaîne. □

Définition 14. *Réversibilité*
Une chaîne de Markov satisfaisant la condition de balance détaillée est dite π-réversible ou réversible par rapport à π.

Définition 15. *Distribution d'équilibre*
La distribution invariante π est dite distribution d'équilibre (ou distribution limite) de la chaîne si pour π-presque tout x et tout $A \in \mathcal{B}$,

$$\lim_{n\to+\infty} P^n(x, A) = \pi(A)$$

où P^n est le noyau de la chaîne après n itérations.

L'existence d'une distribution invariante $\pi(\cdot)$ ne garantit pas que les échantillons générés par la chaîne converge vers cette mesure $\pi(\cdot)$. Pour ce faire, le noyau de transition de la chaîne doit satisfaire trois propriétés principales : l'irréductibilité, l'apériodicité et la récurrence.

1. $P(x, \mathrm{d}y)$ est la probabilité d'aller dans un petit sous-ensemble mesurable $\mathrm{d}y \in \mathcal{X}$ sachant que la chaîne était en x.

B.1.4 Irréductibilité

L'irréductibilité permet de s'assurer que la chaîne de Markov est capable d'explorer toutes les parties intéressantes de l'espace d'états quelle que soit la distribution initiale. Dans les espaces d'états continus, elle est définie au travers d'une mesure σ-finie.

On désigne par τ_A le temps de premier retour en $A \in \mathcal{B}$ (appelé aussi temps de premier passage ou temps d'arrêt en A),

$$\tau_A = \inf\{n \geq 1 : X^{(n)} \in A\},$$

avec la convention $\inf \emptyset = \infty$, i.e. $\tau_A = \infty$ s'il n'existe pas un $n \geq 1$ tel que $X^{(n)} \in A$.

Définition 16. *Soit ϕ une mesure non-nulle et σ-finie sur $(\mathcal{X}, \mathcal{B})$. On dit que la chaîne de Markov $(X^{(n)})$ (ou, de façon équivalente, son noyau de transition) est ϕ-irréductible si pour tout $A \in \mathcal{X}$ et pour tout point de départ $x \in \mathcal{X}$, la propriété suivante est vérifiée :*

$$\phi(A) > 0 \Rightarrow P(\tau_A < +\infty | X^{(0)} = x) > 0. \tag{B.8}$$

Cette équation signifie que tout ensemble de mesure non nulle par rapport à ϕ peut être atteint en un temps fini, quel que soit l'état initial de la chaîne.

Tout comme pour l'invariance, on peut vérifier une condition suffisante pour l'irréductibilité.

Proposition 3. *Une chaîne de Markov $(X^{(n)})$ de noyau $P(\cdot, \cdot)$ est ϕ-irréductible si pour un certain $n \geq 1$, la distribution $P^n(\cdot, \cdot)$ admet une densité strictement positive (notée f) par rapport à ϕ. En d'autres termes, s'il existe une fonction positive f telle que*

$$P^n(x, A) = \int_A f(x, y)\phi(\mathrm{d}y)$$

pour tous $x \in \mathcal{X}$ et $A \in \mathcal{B}$.

Les algorithmes de simulation MCMC décrits dans cette thèse (Metropolis-Hastings et Gibbs notamment) vérifient cette propriété pour $\phi = \pi$ où π est la distribution d'intérêt. Dans la suite, nous prendrons $\phi = \pi$ pour parler d'irréductibilité.

L'irréductibilité est une notion importante puisqu'elle garantit que tous les ensembles intéressants seront explorés par la chaîne $(X^{(n)})$. Cependant, elle n'est pas suffisante pour assurer de bonnes propriétés de convergence puisqu'elle ne garantit pas une fréquence suffisante de passage des trajectoires de la chaîne dans une telle région. D'où la notion de récurrence qui est plus forte que l'irréductibilité puisqu'elle impose que les trajectoires passent en moyenne une infinité de fois dans de tels ensembles.

B.1.5 Récurrence

Définition 17. *[Tie94]*
Une chaîne de Markov $(X^{(n)})$ est récurrente si elle est π-irréductible et si pour tout $A \in \mathcal{B}$ tel que $\pi(A) > 0$, les deux conditions suivantes sont vérifiées :

1) $P(X^{(n)} \in A$ *infiniment souvent* $|X^{(0)} = x) > 0$ *pour tout point de départ x,*

2) $P(X^{(n)} \in A$ *infiniment souvent* $|X^{(0)} = x) = 1$ *pour π-presque tout point de départ x.*

La notation $\{A_n$ infiniment souvent$\}$ signifie que la séquence A_n se produit infiniment souvent.

Définition 18. *Positivité*
Une chaîne de Markov est dite positive si elle est irréductible et possède une mesure invariante σ-finie.

Proposition 4. *Lien entre positivité et récurrence*
Si la chaîne $(X^{(n)})$ est positive, alors elle est récurrente [Rob96, page 108].
On parlera alors de chaînes positives pour les chaînes récurrentes et positives.

Théorème 10. *[GRS96, page 64], [Tie94, page 1713]*
Si $(X^{(n)})$ est une chaîne π-irréductible et π-invariante, alors π est l'unique distribution invariante de la chaîne et la chaîne est récurrente positive.

Dans les algorithmes MCMC, la loi cible sera construite de telle sorte qu'elle soit la distribution invariante π. Si on montre que la chaîne est π-irréductible, elle sera alors récurrente positive.

Nous avons le premier résultat de convergence suivant, et qui donne la convergence en moyenne des itérés de la chaîne vers π.

Théorème 11. *Soit $(X^{(n)})$ une chaîne de Markov de noyau de transition $P(\cdot,\cdot)$, irréductible et ayant $\pi(\cdot)$ comme mesure de probabilité invariante. On pose*

$$\mu_{n,x}(\cdot) = \frac{1}{n}\sum_{k=1}^{n} P^k(x,\cdot) \equiv \frac{1}{n}\sum_{k=1}^{n} P(X^{(k)} \in \cdot | X^{(0)} = x).$$

Alors pour π-presque tout point de départ $x \in \mathcal{X}$,

$$\lim_{n\to\infty} \|\mu_{n,x} - \pi\|_{TV} = 0 \tag{B.9}$$

où $\|\cdot\|_{TV}$ désigne la norme de variation totale, définies pour deux distributions μ_1 et μ_2 par

$$\|\mu_1 - \mu_2\|_{TV} = \sup_A \|\mu_1(A) - \mu_2(A)\|.$$

On peut aussi, sous les mêmes hypothèses, obtenir une loi forte des grands nombres.

Théorème 12. *Soit $(X^{(n)})$ une chaîne de Markov de noyau de transition $P(\cdot,\cdot)$, π-irréductible et π-invariante. Soit h une fonction π-intégrable, définie sur $\mathcal{X}$ et à valeurs réelles. Alors pour π-presque tout point de départ x,*

$$\frac{1}{n}\sum_{k=1}^{n} h(X^{(k)}) \overset{n\to\infty}{\longrightarrow} \int_{\mathcal{X}} h(x)\pi(x)\mathrm{d}x \text{ presque-sûrement.} \tag{B.10}$$

Le résultat de convergence du théorème 11 peut être amélioré grâce à la notion d'apériodicité.

B.1.6 Apériodicité

L'apériodicité est un concept plus technique permettant d'écarter d'éventuels comportements périodiques ou cycliques de la chaîne. Une chaîne est périodique s'il y'a des parties de l'espace d'états qu'elle peut seulement visiter à certains temps régulièrement espacés. L'apériodicité permet donc d'éviter que le comportement de la chaîne ne soit dirigé par des comportements déterministes dans la manière d'engendrer l'échantillon $X^{(n)}$ à partir de $X^{(n-1)}$.

Définition 19. *[Tie94]*
Soit $(X^{(n)})$ une chaîne de Markov dont le noyau de transition P est π-irréductible. On dit que le noyau (ou la chaîne) est périodique s'il existe un entier $d \geq 2$ et une séquence $\{E_0, \ldots, E_{d-1}\}$ formant une partition de $\mathcal{X}$ telle que pour tout $i = 0, \ldots, d-1$ et pour tout $x \in E_i$,

$$P(x, E_j) = 1 \text{ pour } j = i + 1 \text{ mod } d.$$

Sinon, le noyau est dit apériodique.

La propriété d'apériodicité est souvent difficile à vérifier directement et il est plus aisé de vérifier une condition suffisante dite d'apériodicité forte (cf. [GRS96, page 65]). Dans le cas des chaînes continues, une autre condition suffisante d'apériodicité est que la densité $f(\cdot|x_n)$ soit strictement positive dans un voisinage de x_n. Ainsi, la chaîne peut rester un nombre arbitraire d'instants dans ce voisinage avant de visiter un ensemble A quelconque [Rob96, page 101].

En rajoutant la condition d'apériodicité au théorème 11, on obtient une convergence plus forte portant directement sur le noyau P^n et non plus sur les moyennes.

Théorème 13. *Soit $(X^{(n)})$ une chaîne de Markov de noyau $P(\cdot, \cdot)$, π-irréductible, π-invariante et apériodique. Alors pour π-presque tout point de départ x,*

$$P^n(x, \cdot) \to \pi(\cdot) \text{ en variation totale.} \tag{B.11}$$

B.1.7 Récurrence au sens de Harris

La convergence dans les théorèmes 11, 12 et 13 n'est pas vérifiée pour les points de départ x qui sont de mesure nulle par rapport à π. En fait, dans la définition 17 de la récurrence, la condition 2) n'est pas vérifiée pour de tels points. La récurrence au sens de Harris permet de corriger cela.

Définition 20. *[Tie94]*
Une chaîne de Markov $(X^{(n)})$ est récurrente au sens de Harris si elle est π-irréductible et si pour tout $A \in \mathcal{B}$ tel que $\pi(A) > 0$,

$$P(X^{(n)} \in A \text{ infiniment souvent } | X^{(0)} = x) = 1 \text{ pour tout point de départ } x.$$

La récurrence au sens de Harris impose donc que presque toute trajectoire de la chaîne passe une infinité de fois dans tout ensemble A tel que $\pi(A) > 0$.

Signalons que contrairement à la conclusion du théorème 10, l'invariance et l'irréductibilité ne suffisent pas à conclure que la chaîne est Harris-récurrente. Il existe cependant une condition suffisante.

Théorème 14. *[Tie94, Corollaire 1]*
Soit $P(\cdot,\cdot)$ un noyau π-irréductible, de distribution invariante π. Si $P(x,\cdot)$ est absolument continu par rapport à π pour tout x, alors P est récurrent au sens de Harris.

Cette condition est très souvent satisfaite par les algorithmes de simulation que nous avons abordés dans cette thèse. Ce qui garantit une propriété de récurrence au sens de Harris aux chaînes produites.

Définition 21. *Ergodicité*
Une chaîne de Markov irréductible, récurrente au sens de Harris et apériodique est dite ergodique.

Théorème 15. *Si $(X^{(n)})$ est une chaîne de Markov ergodique dont la distribution invariante est π, alors les théorèmes 11, 12 et 13 sont vérifiés pour tout point de départ $x \in \mathcal{X}$ (cf. [GRS96]).*

Les résultats de convergence que nous venons de voir sont importants pour s'assurer de la convergence des itérés produits par la chaîne vers la loi d'intérêt, et ce quelle que soit la distribution initiale. Néanmoins, ces résultats ne précisent pas la vitesse de convergence vers la loi cible. Pour cela, il faut introduire les notions d'ergodicité géométrique et uniforme.

Définition 22. *Ergodicité géométrique*
Une chaîne de Markov ergodique de distribution invariante π est géométriquement ergodique s'il existe une fonction M positive, π-intégrable et une constante $r \in [0,1[$ tels que pour tout point de départ x et tout entier n,

$$\|P^n(x,\cdot) - \pi(\cdot)\|_{TV} \leq M(x) r^n.$$

Le plus petit r pour lequel il existe une fonction M satisfaisant cette condition est dite *vitesse de convergence*.

Une condition plus forte est l'ergodicité uniforme de la chaîne.

Définition 23. *Ergodicité uniforme*
Une chaîne de Markov ergodique de distribution invariante π est uniformément ergodique s'il existe une constante $0 < C < \infty$ et une constante $r \in [0,1[$ tels que pour tout point de départ x et tout entier n,

$$\|P^n(x,\cdot) - \pi(\cdot)\|_{TV} \leq C r^n.$$

Les chaînes géométriquement ergodiques et uniformément ergodiques permettent l'existence de théorèmes centraux limites [Tie94].

B.2 Convergence de l'algorithme MH

Comme nous venons de le voir, il est nécessaire d'imposer des conditions minimales sur le noyau de transition d'un algorithme MCMC pour que la loi cible π soit la distribution limite d'une séquence d'échantillons générés par l'algorithme. Ces conditions de régularité sont l'invariance, l'irréductibilité, la récurrence et l'apériodicité.

Le noyau de transition de l'algorithme MH est donné par :

$$P_{MH}(x, dy) = \alpha(x, y)q(x, y)dy + \left[1 - \int_{\mathcal{X}} \alpha(x, z)q(x, z)dz\right] \delta_x(dy)$$

avec $P_{MH}(x, dy) := P_{MH}(dy|x)$ et $q(x, y) := q(y|x)$.
Le noyau de transition est donc un mélange de la loi candidate et d'un Dirac. Le premier terme définit les transitions de y en x avec la probabilité $\alpha(x, y)$, et le deuxième les transitions de x en x avec la probabilité $r(x) = 1 - \int_{\mathcal{X}} \alpha(x, z)q(x, z)dz$.

B.2.1 Invariance

On a le théorème suivant :

Théorème 16. *Soit supp* $\pi = \{x : \pi(x) > 0\}$ *désignant le support de* π. *On suppose que ce support est connexe. Alors pour toute loi instrumentale* q *telle que*

$$supp\ \pi \subset \bigcup_{x \in\ supp\ \pi} supp\ q(\cdot|x),$$

π *est la distribution invariante de la chaîne produite par l'algorithme MH.*

Démonstration. Il suffit de montrer que

$$\int P_{MH}(x, A)\pi(\mathrm{d}x) = \int_A \pi(y)\mathrm{d}y.$$

Cette démonstration est faite dans [Rob96, page 129].

Alternativement, on peut aussi aisément montrer que la densité K_{MH} du noyau P_{MH} vérifie la condition de balance détaillée. Il suffit de considérer séparément les deux termes du noyau. Pour le deuxième terme, on a trivialement :

$$\left(1 - \int \alpha(x, z)q(x, z)dz\right) \delta_x(y)\pi(x) = \left(1 - \int \alpha(y, z)q(y, z)dz\right) \delta_y(x)\pi(y)$$

donc on a bien $r(x)\delta_x(y)\pi(x) = r(y)\delta_y(x)\pi(y)$.

Il reste à vérifier que $\alpha(x, y)q(x, y)$ vérifie la condition de réversibilité. On a :

$$\begin{aligned}
\alpha(x, y)q(x, y)\pi(x) &= \begin{cases} q(x, y)\pi(x) & \text{si } q(y, x)\pi(y) > q(x, y)\pi(x) \\ q(y, x)\pi(y) & \text{sinon} \end{cases} \\
&= \min\left(q(x, y)\pi(x), q(y, x)\pi(y)\right) \\
&= \min\left(q(y, x)\pi(y), q(x, y)\pi(x)\right) \\
&= \alpha(y, x)q(y, x)\pi(y)
\end{aligned}$$

L'invariance de la loi cible π est donc assurée quelle soit la loi candidate $q(\cdot, \cdot)$ dont le support contient celui de la loi $\pi(\cdot)$. □

Comme nous l'avons déjà signalé, l'existence d'une loi stationnaire π ne suffit pas. Il faut aussi s'assurer de la convergence vers π, c'est-à-dire que la loi stationnaire puisse être atteinte quel que soit le point de départ de la chaîne. Pour cela, il faut montrer la π-irréductibilité et l'apériodicité de la chaîne puisqu'alors l'existence de π comme loi stationnaire implique l'ergodicité de la chaîne.

B.2.2 Irréductibilité

Pour l'irréductibilité, une condition suffisante est la positivité de la distribution candidate q sur le support de π :

$q(x, y) = q(y|x) > 0$ pour tous $(x, y) \in \mathcal{E} \times \mathcal{E}$ (où $\mathcal{E}$ est le support connexe de π).

Dans ce cas, on peut atteindre tout ouvert du support en une seule étape.

Il existe aussi une condition moins restrictive sur q et qui garantit l'irréductibilité et l'apériodicité :

Théorème 17. *Soit π bornée et positive sur tout compact de $\mathcal{X}$. S'il existe des constantes $\epsilon > 0, \delta > 0$ tels que*

$$|x - y| < \delta \Rightarrow q(y|x) > \epsilon$$

alors la chaîne est π-irréductible et apériodique.

Ce résultat est démontré dans [Rob96, page 131].

Ce résultat s'interprète ainsi : si la loi candidate autorise un déplacement dans un voisinage $V(x^{(t)})$ de $x^{(t)}$ de diamètre borné et que la la loi cible π est telle que le taux d'acceptation $\alpha(x^{(t)}, y^{(t)})$ soit strictement positif dans ce voisinage, on peut atteindre n'importe quelle région de $\mathcal{E}$ en k étapes pour k assez grand (à condition que le support $\mathcal{E}$ soit connexe). Ce théorème est utile pour les algorithmes MH à marche aléatoire car il suffit que q soit symétrique dans un voisinage de 0 pour assurer l'ergodicité de la chaîne [Rob96].

B.2.3 Récurrence

Si on a la π-irréductibilité et la π-invariance, alors π est l'unique distribution invariante et la chaîne est *récurrente positive* d'après le théorème 10. Les théorèmes de convergence 11, 12 et 13 s'appliquent alors pour π-presque tout point départ. En fait, il s'avère qu'ils s'appliquent pour tout point de départ car si le noyau de l'algorithme MH est π-irréductible, alors il est récurrent au sens de Harris [Tie94, Corollaire 2].

B.2.4 Apériodicité

L'apériodicité de la chaîne dans l'algorithme MH est vérifiée si $q(\cdot)$ n'est pas le noyau de transition d'une chaîne de Markov réversible et π-invariante, c'est-à-dire si l'on n'a pas

$$\pi(x)q(y|x) = \pi(y)q(x|y) \text{ presque sûrement.}$$

En d'autres termes, les évènements $\{x^{(t+1)} = x^{(t)}\}$ sont autorisés avec une probabilité non-nulle.

B.3 Convergence de l'échantillonneur de Gibbs

Tout comme dans l'algorithme MH, il faut vérifier les propriétés d'invariance, d'irréductibilité et de récurrence de la chaîne produite par l'échantillonneur de Gibbs.

B.3.1 Invariance

Le noyau de transition de l'algorithme de Gibbs est donné par :

$$\begin{aligned} K_{Gibbs}(\boldsymbol{x}, \boldsymbol{y}) &= K_{Gibbs}(\boldsymbol{y}|\boldsymbol{x}) \\ &= \pi_1(y_1|x_2, \ldots, x_N) \times \pi_2(y_2|y_1, x_3, \ldots, x_N) \times \cdots \times \pi_N(y_N|y_1, y_2, \ldots, y_{N-1}). \end{aligned}$$

Utilisant ce noyau, on peut montrer la π-invariance de façon directe, i.e

$$P(\boldsymbol{y} \in A) = \int \mathbf{1}_A(y) K(\boldsymbol{y}|\boldsymbol{x}) \pi(\boldsymbol{x}) \mathrm{d}\boldsymbol{y} \mathrm{d}\boldsymbol{x} = \int_A \pi(\boldsymbol{y}) \mathrm{d}\boldsymbol{y}.$$

(cf. [Rob96, page 172]).

Une autre façon de montrer l'invariance est d'utiliser la propriété selon laquelle l'algorithme d'échantillonnage de Gibbs est la composition de N algorithmes de MH où tous les échantillons proposés sont acceptés [Rob96, page 173]. Cependant, les propriétés de convergence ne s'appliquent pas directement car les N noyaux, pris individuellement, ne sont jamais irréductibles puisque contraints à des sous-espaces de dimension inférieure.

B.3.2 Irréductibilité

Une condition suffisante pour l'irréductibilité de la chaîne est la contrainte de positivité sur π, c'est-à-dire que $\pi^i(x_i) > 0$ pour $i = 1, \ldots, n$, où $\pi^i(x_i) = \int_{x_{-i}} \pi(\boldsymbol{x})$ la distribution marginale de x_i. Cette contrainte signifie que le support de la loi cible $\pi(\cdot)$ est le produit cartésien des supports des lois conditionnelles $\pi^i(\cdot)$.

Toutefois, la contrainte de positivité peut être relativement contraignante. Une condition moins contraignante pour garantir l'irréductibilité existe :

s'il existe $\delta > 0$ telle que pour tout $i = 1, \ldots, N$,

$$\pi_i(x_i|x_1, \ldots, x_{i-1}, x'_{i+1}, \ldots, x'_N) > 0 \text{ si } |x - x'| < \delta$$

pour x et $x' \in \mathrm{supp}(\pi)$ et si le support de π est connexe, alors la chaîne produite par l'échantillonneur de Gibbs est irréductible et apériodique.

B.3.3 Récurrence

Si le noyau est irréductible, la récurrence au sens de Harris est vérifiée si le noyau de transition est absolument continu par rapport à π. L'absolue continuité est souvent vérifiée par le noyau de Gibbs, sauf si une des étapes de l'échantillonnage est remplacée par une étape MH (dans l'échantillonnage hybride).

Théorème 18. *Si la condition de positivité est vérifiée et si le noyau P_{Gibbs} est absolument continu par rapport à π, alors pour tout point de départ x,*

$$\lim_{N\to\infty} \|P^n(x,\cdot) - \pi(\cdot)\|_{TV} = 0. \tag{B.12}$$

On a aussi pour toute fonction h π-intégrable,

$$\frac{1}{N}\sum_{k=1}^{N} h(x^{(k)}) \overset{N\to\infty}{\longrightarrow} \int_{\mathcal{X}} h(x)\pi(x)\mathrm{d}x \text{ presque-sûrement.} \tag{B.13}$$

Bibliographie

[AL06] E. Asma and R.M. Leahy. Mean and covariance properties of dynamic PET reconstructions from list-mode data. *IEEE Transactions on Medical Imaging*, 25(1) :42–54, 2006.

[Ald85] D. Aldous. Exchangeability and related topics. *École d'Été de Probabilités de Saint-Flour XIII-1983*, pages 1–198, 1985.

[Ant74] C.E. Antoniak. Mixtures of Dirichlet processes with applications to Bayesian nonparametric problems. *Ann. Statist.*, 2 :1152–1174, 1974.

[BDM07] É. Barat, T. Dautremer, and T. Montagu. Nonparametric Bayesian inference in nuclear spectrometry. In *IEEE Nuclear Science Symposium Record, NSS/MIC '07.*, 2007.

[BDP96] J. Browne and A.B. De Pierro. A row-action alternative to the EM algorithm for maximizing likelihood in emission tomography. *IEEE Transactions on Medical Imaging*, 15(5) :687–699, 1996.

[BG01] J.O. Berger and A. Guglielmi. Bayesian and conditional frequentist testing of a parametric model versus nonparametric alternatives. *Journal of the American Statistical Association*, 96 :174–184, 2001.

[BJ06] D. M. Blei and M.I. Jordan. Variational inference for Dirichlet process mixtures. *Bayesian Analysis*, 1(1) :121–144, 2006.

[Bla73] D. Blackwell. Discreteness of Ferguson selection. *The Annals of Statistics*, 1 :356–358, 1973.

[BM73] D. Blackwell and J. B. MacQueen. Ferguson distributions via Pólya urn schemes. *Ann. Statist.*, 1 :353–355, 1973.

[BM96] C. A. Bush and S. N. MacEachern. A semiparametric Bayesian model for randomised block designs. *Biometrika*, 83(2) :275–285, 1996.

[Bru02] P.P. Bruyant. Analytic and iterative reconstruction algorithms in SPECT. *Journal of Nuclear Medicine*, 43(10) :1343–1358, 2002.

[BS96] C.A. Bouman and K. Sauer. A unified approach to statistical tomography using coordinate descent optimization. *IEEE Transactions on Image Processing*, 5(3) :480–492, 1996.

[BSW99] A. Barron, M. J. Schervish, and L. Wasserman. The consistency of posterior distributions in nonparametric problems. *Ann. Statist.*, 27 :536–561, 1999.

[BZA+97] G. Brix, J. Zaers, L.-E. Adam, M. E. Bellemann, H. Ostertag, H. Trojan, U. Haberkorn, J. Doll, F. Oberdorfer, and W. Lorenz. Performance evaluation of a whole-body PET scanner using the NEMA protocol. *Journal of Nuclear Medicine*, 38(10) :1614–1623, 1997.

[Car99] M. Carlton. *Applications of the Two-Parameter Poisson-Dirichlet Distribution*. PhD thesis, UCLA Department of Statistics, 1999.

[Cav00] L. Cavalier. Efficient estimation of a density in a problem of tomography. *Ann. Statist.*, 28(2) :630–647, 2000.

[CD11] Y. Chung and D.B. Dunson. The local Dirichlet process. *Annals of the Institute of Statistical Mathematics*, 63(1) :59–80, 2011.

[CG92] G. Casella and E. I. George. Explaining the Gibbs sampler. *American Statistician*, 46(3) :167–174, 1992.

[CGR05] N. Choudhuri, S. Ghosal, and A. Roy. Bayesian methods for function estimation. *Handbook of statistics*, 25 :373–414, 2005.

[CJ93] V.J Cunningham and T. Jones. Spectral analysis of dynamic PET studies. *Journal of Cerebral Blood Flow and Metabolism*, 13(1) :15–23, 1993.

[Col80] J.G. Colsher. Fully-three-dimensional Positron Emission Tomography. *Physics in medicine and biology*, 25 :103–115, 1980.

[DCT95] M. Defrise, R. Clack, and DW Townsend. Image reconstruction from truncated, two-dimensional, parallel projections. *Inverse Problems*, 11 :287–313, 1995.

[Def07] M. Defrise. Reconstruction d'image en tomographie par émission. *Médecine nucléaire*, 31(4) :142–152, 2007.

[DGG07] J.A. Duan, M. Guindani, and A.E. Gelfand. Generalized spatial Dirichlet process models. *Biometrika*, 94(4) :809–825, 2007.

[DKT+97] M. Defrise, P.E. Kinahan, D.W. Townsend, C. Michel, M. Sibomana, and D.F. Newport. Exact and approximate rebinning algorithms for 3-D PET data. *IEEE Transactions on Medical Imaging*, 16(2) :145–158, 1997.

[DLR77] A.P. Dempster, N.M. Laird, and D.B. Rubin. Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 1–38, 1977.

[DP08] D.B. Dunson and J.H. Park. Kernel stick-breaking processes. *Biometrika*, 95(2) :307–323, 2008.

[Eng78] S. Engen. *Stochastic abundance models, with emphasis on biological communities and species diversity*. Chapman and Hall, 1978.

[Esc94] M.D. Escobar. Estimating normal means with a Dirichlet process prior. *J. Am. Stat. Assoc.*, 89 :268–277, 1994.

[ET98] W.J. Ewens and S. Tavaré. The Ewens Sampling formula. *Encyclopedia of Statistical Sciences, Update*, 2 :230–234, 1998.

[EW95] M.D. Escobar and M. West. Bayesian density estimation and inference using mixtures. *J. Am. Stat. Assoc.*, 90 :577–588, 1995.

[Ewe72] W.J. Ewens. The sampling theory of selectively neutral alleles. *Theoretical Population Biology*, 3(1) :87–112, 1972.

[Fen10] S. Feng. *The Poisson-Dirichlet Distribution and Related Topics*. Springer Verlag, 2010.

[Fer73] T. S. Ferguson. A Bayesian analysis of some nonparametric problems. *Ann. Statist.*, 1 :209–230, 1973.

[Fer74] T. S. Ferguson. Prior distributions on spaces of probability measures. *Ann. Statist.*, 2 :615–629, 1974.

[Fer83] T.S. Ferguson. Bayesian density estimation by mixtures of Normal distributions. *Recent advances in Statistics : papers in honor of Herman Chernoff on his sixtieth birthday*, pages 287–302, 1983.

[Fes04] J.A. Fessler. Statistical methods for image reconstruction. In *Course on statistical image reconstruction methods, NSS-MIC*, 2004.

[FH95] J.A. Fessler and A.O. Hero. Penalized maximum-likelihood image reconstruction using space-alternating generalized EM algorithms. *IEEE Transactions on Image Processing*, 4(10) :1417–1429, 1995.

[FK72] T.S. Ferguson and M.J. Klass. A representation of independent increment processes without Gaussian components. *The Annals of Mathematical Statistics*, 43(5) :1634–1643, 1972.

[Fre63] D. A. Freedman. On the asymptotic behavior of Bayes estimates in the discrete case. *Ann. Math. Statist.*, 34 :1386–1403, 1963.

[GG84] S. Geman and D. Geman. Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images. *IEEE Trans. Pattern. Anal. Mach. Intell.*, 6 :721–741, 1984.

[GGR99] S. Ghosal, J. K. Ghosh, and R.V. Ramamoorthi. Posterior consistency of Dirichlet mixtures in density estimation. *Ann. Statist.*, 27 :143–158, 1999.

[GGT+02] R.N. Gunn, S.R. Gunn, F.E. Turkheimer, J.A.D. Aston, and V.J. Cunningham. Positron emission tomography compartmental models : a basis pursuit strategy for kinetic modelling. *Journal of Cerebral Blood Flow and Metabolism*, 22 :1425–1439, 2002.

[Gje82] A. Gjedde. Calculation of cerebral glucose phosphorylation from brain uptake of glucose analogs in vivo : a re-examination. *Brain Research Reviews*, 4(2) :237–274, 1982.

[GLHC97] R.N. Gunn, A.A. Lammertsma, S.P. Hume, and V.J. Cunningham. Parametric imaging of ligand-receptor binding in PET using a simplified reference region model. *Neuroimage*, 6(4) :279–287, 1997.

[GM87] S. Geman and D. E. McClure. Statistical methods for tomographic image reconstruction. In Bulletin of the ISI, editor, *Proceedings of the 46th Session of the International Statistical Institute*, volume 52, pages 5–21, 1987.

[GR01] P. J. Green and S. Richardson. Modelling heterogeneity with and without the Dirichlet process. *Scandinavian Journal of Statistics*, 28 :355–375, 2001.

[GR03] J. K. Ghosh and R. V. Ramamoorthi. *Bayesian Nonparametrics*. Springer, 2003.

[Gre90] P. J. Green. Bayesian reconstructions from emission tomography data using a modified EM algorithm. *IEEE Trans. Med. Imaging*, 9(1) :84–93, 1990.

[Gri88] R. C. Griffiths. On the distribution of points in a Poisson-Dirichlet process. *Journal of applied probability*, 25(2) :336–345, 1988.

[GRS96] W. R. Gilks, S. Richardson, and D. J. Spiegelhalter. *Markov Chain Monte Carlo in practice*. Chapman and Hall, 1996.

[GS90] A.E. Gelfand and A.F.M. Smith. Sampling-based approaches to calculating marginal densities. *Journal of the American statistical association*, 85 :398–409, 1990.

[GS06] J. E. Griffin and M. F. J. Steel. Order-based dependent Dirichlet processes. *Journal of the American Statistical Association*, 101(473) :179–194, 2006.

[Han06] T. Hanson. Inference for mixtures of finite Pólya tree models. *Journal of the American Statistical Association*, 101 :1548–1565, 2006.

[Has70] W.K Hastings. Monte carlo sampling methods using Markov chains and their applications. *Biometrika*, 57(1) :97–109, 1970.

[HHMW10] N. L. Hjort, C. Holmes, P. Müller, and S. G. Walker. *Bayesian Nonparametrics*. Cambridge University Press, April 2010.

[HJ02] T. Hanson and W. O. Johnson. Modeling regression error with a mixture of Pólya trees. *J. Am. Stat. Assoc.*, 97 :1020–1033, 2002.

[HL89] T. Hebert and R. Leahy. A generalized EM algorithm for 3D Bayesian reconstruction from poisson data using Gibbs priors. *IEEE Transactions on Medical Imaging*, 8(2) :194–202, 1989.

[HL94] H.M. Hudson and R.S. Larkin. Accelerated image reconstruction using ordered subsets of projection data. *IEEE Transactions on Medical Imaging*, 13(4) :601–609, 1994.

[HP98] B. Hansen and J. Pitman. Prediction rules for exchangeable sequences related to species sampling. Technical Report 520, Department of Statistics, University of California, Berkeley, 1998.

[Hub81] P.J. Huber. *Robust statistics*, volume 1. Wiley, New-York, 1981.

[IJ01] H. Ishwaran and L. F. James. Gibbs sampling methods for stick-breaking priors. *J. Am. Stat. Assoc.*, 96 :161–173, 2001.

[IJ03a] H. Ishwaran and L. F. James. Generalized weighted chinese restaurant processes for species sampling mixture models. *Statistica Sinica*, 13 :1211–1235, 2003.

[IJ03b] H. Ishwaran and L. F. James. Some further developments for stick-breaking priors : Finite and infinite clustering and classification. *Sankhya Series A*, 65 :577–592, 2003.

[IZ00] H. Ishwaran and M. Zarepour. Markov chain Monte Carlo in approximate Dirichlet and beta two-parameter process hierarchical models. *Biometrika*, 87 :371–390, 2000.

[IZ02] H. Ishwaran and M. Zarepour. Exact and approximate sum representation for the Dirichlet process. *Canadian Journal of Statistics*, 26 :269–283, 2002.

[JS90] I. M. Johnstone and B. W. Silverman. Speed of estimation in positron emission tomography and related inverse problems. *Ann. Statist.*, 18(1) :251–280, 1990.

[KBMS05] M.E. Kamasak, C.A. Bouman, E.D. Morris, and K. Sauer. Direct reconstruction of kinetic parameter images from dynamic PET data. *IEEE Transactions on Medical Imaging*, 24(5) :636–650, 2005.

[KGW09] M. Kalli, J. E. Griffin, and S. G. Walker. Slice sampling mixture models. Technical Report UKC/IMS/08/024, revised, IMSAS, University of Kent, 2009.

[Kin75] J.F.C. Kingman. Random discrete distributions. *Journal of the Royal Statistical Society. Series B (Methodological)*, 37(1) :1–22, 1975.

[KR89] P.E. Kinahan and J.G. Rogers. Analytic 3D image reconstruction using all detected events. *IEEE Transactions on Nuclear Science*, 36(1) :964–968, 1989.

[Lav92] M. Lavine. Some aspects of Pólya tree distributions for statistical modelling. *Ann. Statist.*, 20 :1222–1235, 1992.

[Lav94] M. Lavine. More aspects of Pólya tree distributions for statistical modelling. *The Annals of Statistics*, 22(3) :1161–1176, 1994.

[LC84] K. Lange and R. Carson. EM reconstruction algorithm for emission and transmission tomography. *J. Comp. Assist. Tomo.*, 8 :306–316, 1984.

[Lo84] A. Y. Lo. On a class of Bayesian Nonparametric estimates : I. density estimates. *The Annals of Statistics*, 12 :351–357, 1984.

[LQ00] R.M. Leahy and J. Qi. Statistical approaches in quantitative positron emission tomography. *Statistics and Computing*, 10(2) :147–165, 2000.

[LQMT08] J. Lee, F. A. Quintana, P. Müller, and L. Trippa. Defining predictive probability functions for species sampling models. Technical report, Working paper, 2008.

[Mac94] S. N. MacEachern. Estimating normal means with a conjugate style Dirichlet process prior. *Communications in Statistics : Simulation and Computation*, 23(3) :727–741, 1994.

[Mac99] S. N. MacEachern. Dependent nonparametric processes. In *ASA Proceedings on the Section on Bayesian Statistical Science*, pages 50–55, 1999.

[MBPC97] J. Matthews, D. Bailey, P. Price, and V. Cunningham. The direct calculation of parametric images from dynamic PET data using maximum-likelihood iterative reconstruction. *Phys. Med. Biol*, 42(6) :1155–1173, 1997.

[McC65] J.W. McCloskey. *A model for the distribution of individuals by species in an environment*. PhD thesis, Michigan State University, 1965.

[MKG01] S.N. MacEachern, A. Kottas, and A. Gelfand. Spatial nonparametric Bayesian models. In *Proceedings of the Section on Bayesian Statistics of the American Statistical Association*, 2001.

[MM98] S. N. MacEachern and P. Müller. Estimating mixture of Dirichlet process models. *J. Comput. Graph. Stat.*, 7 :223–238, 1998.

[MMC+98] S.R. Meikle, J.C. Matthews, V .J. Cunningham, D.L. Bailey, L Livieratos, T Jones, and P Price. Parametric image reconstruction using spectral analysis of PET projection data. *Phys. Med. Biol*, 43 :651–666, 1998.

[MQ04] P. Müller and F. Quintana. Nonparametric Bayesian data analysis. *Statistical Science*, 19(1) :95–110, 2004.

[MRR+53] N. Metropolis, A.W. Rosenbluth, M.N. Rosenbluth, A.H. Teller, and E. Teller. Equation of state calculations by fast computing machines. *The journal of Chemical Physics*, 21(6) :1087–1092, 1953.

[MS95] P. Muliere and P. Secchi. A note on a proper Bayesian bootstrap. Technical Report 18, Universita degli Sudi di Pavia, Dipartamento di Economia Politica e Metodi Quantitativi, 1995.

[MSW92] R. D. Mauldin, W. D. Sudderth, and S. C. Williams. Pólya trees and random distributions. *Ann. Statist.*, 20 :1203–1221, 1992.

[MT98] P. Muliere and L. Tardella. Approximating distributions of random functionals of Ferguson-Dirichlet priors. *Canadian Journal of Statistics*, 26(2) :283–297, 1998.

[Nat86] F. Natterer. *The Mathematics of Computerized Tomography*. Society for Industrial and Applied Mathematics, 1986.

[Nea91] R. M. Neal. Bayesian mixture modeling. In *Maximum entropy and Bayesian Methods : Proceedings of the 11th International Workshop on Maximum Entropy and Bayesian Methods of Statistical Analysis, Seattle*, pages 197–211, 1991.

[Nea00] R. M. Neal. Markov chain sampling methods for Dirichlet process mixture models. *Journal of Computational and Graphical Statistics*, 9(2) :249–265, 2000.

[Nea03] R.M. Neal. Slice sampling. *The Annals of Statistics*, 31(3) :705–741, 2003.

[NQAL02] T.E. Nichols, J. Qi, E. Asma, and R.M. Leahy. Spatiotemporal reconstruction of list-mode PET data. *IEEE Trans. Med. Imaging*, 21 :396–404, 2002.

[NQM08] C. Navarrete, F. A. Quintana, and P. Müller. Some issues in nonparametric Bayesian modeling using species sampling models. *Statistical Modelling*, 8 :3–21, 2008.

[OF97] J.M. Ollinger and J.A. Fessler. Positron-emission tomography. *Signal Processing Magazine, IEEE*, 14(1) :43–55, 1997.

[PBF+83] C.S. Patlak, R.G. Blasberg, J.D. Fenstermacher, et al. Graphical evaluation of blood-to-brain transfer constants from multiple-time uptake data. *Journal of Cerebral Blood Flow and Metabolism*, 3(1) :1–7, 1983.

[Per90] M. Perman. *Random discrete distributions derived from subordinators*. PhD thesis, University of California, Berkeley, 1990.

[PISW06] I. R. Porteous, A. Ihler, P. Smyth, and M. Welling. Gibbs sampling for (coupled) infinite mixture models in the stick breaking representation. In *UAI*. AUAI Press, 2006.

[Pit92] J. Pitman. The two-parameter generalization of Ewens' random partition structure. Technical Report 345, Dept. Statistics, U.C. Berkeley, 1992.

[Pit95] J. Pitman. Exchangeable and partially exchangeable random partitions. *Probab. Th. Rel. Fields*, 102 :145–158, 1995.

[Pit96a] J. Pitman. Random discrete distributions invariant under size-biased permutation. *Adv. Appl. Prob*, 28 :525–539, 1996.

[Pit96b] J. Pitman. Some developments of the Blackwell-MacQueen urn scheme. In T.S. Ferguson et al., editor, *Statistics, Probability and Game Theory ; Papers in honor of David Blackwell*, volume 30 of *Lecture Notes-Monograph Series*, pages 245–267. Institute of Mathematical Statistics, 1996.

[Pit02] J. Pitman. Combinatorial stochastic processes. Technical Report 621, Dept. Statistics, U.C. Berkeley, 2002.

[Pit03] J. Pitman. Poisson-Kingman partitions. *Statistics and Science : A Festschrift for Terry Speed, IMS Lectures Notes Monograph*, 40 :1–34, 2003.

[PPY92] M. Perman, J. Pitman, and M. Yor. Size-biased sampling of Poisson point processes and excursions. *Probability Theory and Related Fields*, 92(1) :21–39, 1992.

[PR07] O. Papaspiliopoulos and G. O. Roberts. Retrospective Markov chain sampling Monte Carlo methods for Dirichlet process hierarchical models. *Biometrika*, 95 :169–186, 2007.

[PRLW03] S. M. Paddock, F. Ruggeri, M. Lavine, and M. West. Randomized Pólya tree models for nonparametric Bayesian inference. *Stat. Sinica*, 13 :443–460, 2003.

[PY97] J. Pitman and M. Yor. The two-parameter Poisson-Dirichlet distribution derived from a stable subordinator. *Ann. Proba.*, 25 :855–900, 1997.

[QLC+98] J. Qi, R.M. Leahy, S.R. Cherry, A. Chatziioannou, and T.H. Farquhar. High-resolution 3D Bayesian image reconstruction using the microPET small-animal scanner. *Physics in Medicine and Biology*, 43(4) :1001–1013, 1998.

[Qui06] F. A. Quintana. A predictive view of Bayesian clustering. *Journal of Statistical Planning and Inference*, 136(8) :2407–2429, 2006.

[RDG08] A. Rodriguez, D.B. Dunson, and A.E. Gelfand. The nested Dirichlet process. *Journal of the American Statistical Association*, 103(483) :1131–1154, 2008.

[Rob96] C. P. Robert. *Méthodes de Monte Carlo par chaînes de Markov*. Economica, 1996.

[RSC+06] A.J. Reader, F.C. Sureau, C. Comtat, R. Trébossen, and I. Buvat. Joint estimation of dynamic PET images and temporal basis functions using fully 4D ML-EM. *Physics in Medicine and Biology*, 51(21) :5455–5474, 2006.

[Set94] J. Sethuraman. A constructive definition of Dirichlet priors. *Stat. Sinica*, 4 :639–650, 1994.

[Sit11] A. Sitek. Reconstruction of emission tomography data using origin ensembles. *IEEE Transactions on Medical Imaging*, 30 :946–956, 2011.

[Sok97] A. D. Sokal. Monte carlo methods in statistical mechanics : Foundations and new algorithms. *NATO Adv. Sci. Inst. Ser. B Phys.*, 361 :131–192, 1997.

[SRK+77] L. Sokoloff, M. Reivich, C. Kennedy, M.H.D. Rosiers, CS Patlak, K.D. Pettigrew, O. Sakurada, and M. Shinohara. The [14C] deoxyglucose method for the measurement of local cerebral glucose utilization : Theory, procedure, and normal values in the conscious and anesthetized albino rat. *Journal of Neurochemistry*, 28(5) :897–916, 1977.

[ST82] J. Sethuraman and R.C. Tiwari. Convergence of Dirichlet measures and the interpretation of their parameters. *Statistical decision theory and related topics III*, 2 :305–315, 1982.

[SV82] L.A. Shepp and Y. Vardi. Maximum likelihood reconstruction in positron emission tomography. *IEEE Trans. Med. Imag.*, 1 :113–122, 1982.

[Teh06] Y. W. Teh. A hierarchical Bayesian language model based on Pitman-yor processes. In *Proceedings of the 21st International Conference on Computational Linguistics and the 44th annual meeting of the Association for Computational Linguistics*, pages 985–992. Association for Computational Linguistics, 2006.

[Tie94] L. Tierney. Markov chains for exploring posterior distributions. *The Annals of Statistics*, 22(4) :1701–1728, 1994.

[TJBB06] Y. W. Teh, M. I. Jordan, M. J. Beal, and D. M. Blei. Hierarchical Dirichlet processes. *J. Am. Stat. Assoc.*, 101(476) :1566–1581, 2006.

[Tow98] D.W. Townsend. *The theory and practice of 3D PET*, volume 32. Kluwer Academic Publishers, 1998.

[VB03] P.E. Valk and D.L. Bailey. *Positron Emission Tomography : basic science and clinical practice*. Springer Verlag, 2003.

[VSK85] Y. Vardi, L.A. Shepp, and L. Kaufman. A statistical model for Positron Emission Tomography. *J. Am. Stat. Assoc.*, 80 :8–20, 1985.

[Wal07] S. G. Walker. Sampling the Dirichlet mixture model with slices. *Comm. Statist.*, 36 :45–54, 2007.

[WD98] S.G. Walker and P. Damien. Sampling methods for Bayesian nonparametric inference involving stochastic processes. *Practical Nonparametric and Semiparametric Bayesian Statistics*, 133 :243–254, 1998.

[WDLS99] S.G. Walker, P. Damien, P.W. Laud, and A.F.M. Smith. Bayesian nonparametric inference for random distributions and related functions. *Journal of the Royal Statistical Society : Series B (Statistical Methodology)*, 61(3) :485–527, 1999.

[Wes92] M. West. Hyperparameter estimation in Dirichlet process mixture models. Technical report, Institute of Statistics and Decision Sciences, Duke University, 1992.

[WM97] S. G. Walker and B. K. Mallick. Hierarchical generalized linear models and frailty models with Bayesian nonparametric mixing. *Journal of the Royal Statistical Society, Series B*, 59 :845–860, 1997.

[WME94] M. West, P. Müller, and M. D. Escobar. Hierarchical priors and mixture models, with application in regression and density estimation. *Aspects of Uncertainty*, pages 363–386, 1994.

[ZHS+94] I. G. Zubal, C. R. Harrell, E. O. Smith, Z. Rattner, G. Gindi, and P. B. Hoffer. Computerized three-dimensional segmented human anatomy. *Medical Physics*, 21(2) :299–302, 1994.