HAL
open science

Développement et optimisation d'une caméra Compton miniature à masque codé : méthode d'analyse d'un environnement radiatif par spectro-identification et localisation 3D de sources gamma

Geoffrey Daniel

▶ **To cite this version:**

Geoffrey Daniel. Développement et optimisation d'une caméra Compton miniature à masque codé : méthode d'analyse d'un environnement radiatif par spectro-identification et localisation 3D de sources gamma. Autre [cond-mat.other]. Université Paris Cité, 2020. Français. NNT : 2020UNIP7057 . tel-03628810

HAL Id: tel-03628810

https://theses.hal.science/tel-03628810

Submitted on 3 Apr 2022

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

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

Université de Paris

THÈSE DE DOCTORAT

Spécialité Astronomie et Astrophysique – Instrumentation Nucléaire

Développement et optimisation d'une caméra Compton miniature à masque codé : Méthode d'analyse d'un environnement radiatif par spectro-identification et localisation 3D de sources gamma

Présentée par Geoffrey DANIEL

ÉCOLE DOCTORALE

Astronomie et Astrophysique d'Île de France (ED 127)

THÈSE PRÉPARÉE AU

CEA Saclay – DRF / IRFU / Département d'Astrophysique (DAp)

Dirigée par Olivier LIMOUSIN

Présentée et soutenue publiquement le 2 octobre 2020

Devant le jury composé de :

Prof. Thomas PATZAK	Université de Paris, Président
Prof. Jean-Claude ANGÉLIQUE	ENSICAEN, Rapporteur
Dr. Mathieu AUBRY	ENPC, Rapporteur
Dr. Denis BERNARD	LLR/Polytechnique, Examinateur
Dr. Voichita MAXIM	INSA Lyon, Examinatrice
Dr. Françoise PRÊTEUX	ENPC, Examinatrice
Dr. Olivier LIMOUSIN	CEA Saclay, Directeur de thèse

Except where otherwise noted, this is work licensed under https://creativecommons.org/licenses/by-nc-nd/3.0/fr/

Résumé

La caractérisation d'une scène radiologique est un enjeu majeur dans de nombreux domaines, comme la sûreté et la sécurité nucléaire, les problématiques d'assainissement-démantèlement ou encore dans le domaine médical. Une telle analyse consiste à répondre à deux questions principales. D'une part, quels sont les radioéléments présents dans la scène et en quelles quantités ? D'autre part, où se situent ces sources radio-émettrices ? Les systèmes conçus pour répondre à ces questions utilisent l'information portée par les photons X et gamma issus de ces sources radioactives, d'où le nom de « caméra gamma » donné à ces systèmes. Leurs performances dépendent naturellement de celles de l'instrument utilisé, mais aussi des performances des algorithmes de traitement appliqués aux données acquises par l'instrument.

Dans le cadre de cette thèse, j'étudie le développement d'un système de caméra gamma, basé sur la technologie Caliste, détecteur CdTe de photons de haute-énergie, issu de plus de quinze années de développement, initialement pour des applications en observation astrophysique. Les performances spectrométriques et en imagerie de ce détecteur sont une référence mondiale dans le domaine des détecteurs CdTe, d'où la pertinence de son application pour des caméras gamma.

Plus spécifiquement, mon travail de thèse s'articule en trois axes. Dans un premier temps, à travers une modélisation détaillée du détecteur, j'analyse les données spectrométriques issues de Caliste afin de comprendre comment l'information de la nature et de la quantité des radioéléments sont inscrites dans ces données. Cette modélisation et cette analyse me permettent de mettre en place des algorithmes de traitement des données spectrométriques nouveaux et originaux, basés sur les méthodes de *deep learning* et de réseaux de neurones convolutifs bayésiens, dans le but de répondre à la question de l'identification et de la quantification des radioéléments tout en apportant une estimation de l'incertitude sur les résultats. Dans un second temps, j'étudie l'utilisation des méthodes d'imagerie à masque codé avec le détecteur Caliste, et en particulier, les algorithmes de localisation des sources radioactives. Je montre l'application et les limites des algorithmes de reconstruction habituellement utilisés pour cette problématique puis je démontre le potentiel des algorithmes de *deep learning* pour surpasser ces limites. Enfin, je m'intéresse à la problématique de la localisation de points chauds radioactifs par imagerie Compton, à travers des algorithmes de traitement adaptés au détecteur Caliste, en comparant les performances des méthodes issues de l'état de l'art à de nouvelles méthodes que je mets en place.

Au long de ce travail de thèse, j'effectue les développements de mes algorithmes à partir de données issues de simulation de sorte à pouvoir transposer ces méthodes à d'autres systèmes de détection. Je prends ensuite le soin de les tester sur des données acquises avec le détecteur Caliste, afin d'obtenir une évaluation de leurs performances, fidèle aux conditions réelles et afin d'être confronté aux aléas des mesures, non pris en compte dans les simulations.

Mots-Clés : détecteurs CdTe ; spectrométrie gamma ; imagerie à masque codé ; imagerie Compton ; *deep learning* ; réseaux de neurones convolution

Abstract

The characterization of a radiological scene is a major issue in many fields, such as nuclear safety and security, Decommissioning & Dismantling or in the medical field. Such an analysis consists in answering two main questions. On the one hand, what radioelements are present in the scene and in what quantities? On the other hand, where are these radio-emitting sources located? The systems designed to answer these questions use the information carried by the X and gamma photons from these radioactive sources, hence the name "gamma camera" for these systems. Their performances naturally depend on those of the instrument, but also on the performances of the processing algorithms applied to the data acquired by the instrument.

In the framework of this thesis, I study the development of a gamma camera, based on the Caliste technology, a CdTe detector for high-energy photons, which results from more than fifteen years of development, initially for applications in astrophysical observation. The spectrometric and imaging performances of this detector are a world reference in the field of CdTe detectors, hence the relevance of its application for gamma cameras.

More specifically, my thesis work explores three axes. First, through a detailed modelling of the detector, I analyse the spectrometric data from Caliste in order to understand how they carry the information of the nature and quantity of radioelements. This modelling and this analysis allow me to develop new and original spectrometric data processing algorithms, based on deep learning methods and Bayesian convolutional neural networks, in order to answer the question of identification and quantification of radioelements while providing an estimation of the uncertainty of the results. In a second step, I study the use of coded mask imaging methods with Caliste, and in particular, the algorithms for locating radioactive sources. I show the application and the limits for the reconstruction provided by the algorithms usually used for this problem and then I demonstrate the potential of deep learning algorithms to overcome these limits. Finally, I analyse the problem of localization of radioactive hot spots by Compton imaging, through processing algorithms adapted to Caliste detectors, by comparing the performances of methods resulting from the state of the art to new methods that I develop.

Throughout this thesis work, I develop my algorithms from simulation data in order to transpose these methods to other detection systems. Then, I take care to test them on data acquired with Caliste detectors, in order to obtain an evaluation of their performances, faithful to real conditions and in order to confront these algorithms with the hazards of the measurements, not taken into account in the simulations.

Keywords: CdTe detectors; gamma spectrometry; coded mask imaging; Compton imaging; *deep learning*; Convolutional Neural Networks

Remerciements

Nous y sommes, à cette page qui apparaissait encore lointaine il y a quelques semaines et qui provoque maintenant un pincement au cœur, marquant la conclusion d'une aventure exceptionnelle. Soyons positif, c'est aussi un nouveau chapitre qui commence ! Il est temps avant cela de faire le bilan de ces dernières années, à travers les personnes qui m'auront laissé une marque inoubliable.

Je commence par le plus évident, mais aussi le plus difficile, Olivier. Le plus difficile, parce que je ne sais pas comment ne pas en écrire un chapitre entier, je vais essayer de faire court ! Tu le sais très bien, faire court n'est pas ma spécialité, je crois que c'est un point que nous avons en commun. J'ai toujours dit, une thèse, c'est certes un sujet, mais avant tout, un encadrement. Et de ce côté, dire que c'était un encadrement, ce ne serait même pas fidèle à ce que j'ai vécu. Tu disais qu'on fonctionnait comme un duo, c'est clairement ce que j'ai ressenti. Déjà, d'un point de vue scientifique, c'était un vrai plaisir de progresser ensemble, de débattre et surtout de « chou-fleuriser* ». L'ambition et l'enthousiasme contagieux dont tu faisais preuve ont été un moteur au cours de ces trois années. Mais au-delà de cela, c'était une aventure humaine. Je crois que le plus marquant est ton empathie et l'attention que tu portes aux autres, je pense que c'est l'engrenage du moteur qui a permis de créer la synergie dans laquelle on a évolué. Je ne te remercierai jamais assez, cette aventure a été exceptionnelle grâce à toi, et à la célèbre question « si c'était à refaire, que changeriez-vous ? », de ce point de vue, je répondrais « rien du tout » sans hésiter. Il est évident que je serais ravi d'écrire de nouveaux chapitres avec toi à l'avenir.

Je me tourne à présent vers Daniel, que je qualifierais d'expert du détail dans le détail et qui a joué un rôle officieux d'encadrant. Je te remercie pour les discussions poussées sur la physique et le goût que tu as et que tu sais transmettre sur la compréhension profonde des phénomènes. Les conversations de fins d'après-midi (pour ne pas dire débuts de soirée) ont toujours été stimulantes et ont été des moments de « chou-fleurisation » intense. Je te souhaite de t'épanouir entièrement dans ta nouvelle vie professionnelle, mais aussi personnelle (il y a déjà un beau projet en vue de ce côté, je te souhaite beaucoup de bonheur).

Si j'écris un paragraphe sur tout le monde, je n'ai pas fini, je vais être un peu plus bref. Déjà, je continue avec les collègues du DAp. Aline, merci d'avoir été présente et d'avoir apporté un regard un peu plus extérieur mais précieux à mes travaux, je te souhaite du bonheur avec ta petite famille qui s'est récemment agrandie. Claire, tu as marqué les premiers mois de ma thèse et tu m'as formé pour réaliser les manips qui ont été indispensables pour réaliser mes travaux, je te dois beaucoup et je te souhaite une très belle vie post-CEA. Alexandra, nos petites pauses discussions mathématiques (ou pas d'ailleurs) me manquent déjà. Diana, c'est toujours une joie d'échanger avec toi, n'hésite pas à m'envoyer encore des stagiaires, j'espère qu'on aura l'occasion de travailler ensemble dans le LabCom. Philippe L., la rivalité avec Centrale n'a pas empêché de faire des jolies études sur l'imagerie Compton. François D., il nous est arrivé plusieurs fois de refaire le monde, j'espère que nous aurons l'occasion de continuer, merci pour les conseils que tu as pu me donner dans ces trois années. Stéphane, je crois que je te dois le bon déroulement de la soutenance grâce à ton expertise en sono, merci beaucoup

* Chou-fleuriser : verbe intransitif du premier groupe, vient de chou-fleur, signifie qu'une discussion débouche sur 1000 idées à la seconde, processus difficile, voire impossible à arrêter une fois qu'il est lancé.

pour cela. Bertrand, merci pour les discussions enrichissantes qu'on a eues, je te souhaite bon courage pour la dernière ligne droite de SVOM, ainsi qu'à toute l'équipe. Philippe F., Jérôme et Arnaud, merci d'avoir suivi mes travaux avec attention, vos avis ont toujours été précieux. Merci à Luc, pour la gestion logistique des sources, sans cela, une partie de mon travail n'aurait pas pu se réaliser. À toute la team du midi, Aymeric (chanteur officiel du département et amateur d'anecdotes mathématiques), Nicolas R. (tu vas bientôt avoir des nuits complètes, ne t'inquiète pas), Kamshat (ta progression en français est impressionnante, j'espère que je ne t'ai pas trop embrouillée en t'expliquant les subtilités), Tatyana (merci d'avoir partagé tes petites gourmandises ouzbeks), Damien (c'est fascinant ! un peu trop d'ailleurs…), Alain, Rémi B., Rémi C., Élisabeth et Déborah (ancienne membre de la team), les repas à la cantine n'auraient pas eu la même saveur sans cette fine équipe.

Je remercie les personnes qui ont contribué à la fabrication de la caméra, François V. et Marin pour l'assemblage minutieux, Jérôme qui a conçu la mécanique, Dat pour les connecteurs, et Pierre-Anne pour la carte et le logiciel (je te souhaite un plein épanouissement dans le Sud). C'est simple, ma thèse n'aurait pas pu se réaliser sans vous.

Merci à tous les collègues du DAp pour leur accueil et les moments qu'on a pu partager, Éric, Frédéric, Benoît, Cyrille (pour le montage aussi du petit ventilateur qui a permis de faire de belles images), Vincent, Louis, Clément, Cyril, Andrea, Christian, Isabelle, Diego, Émeric, Thierry, Koryo, Pierre-Antoine, Marie, Hugues.

Je remercie aussi les collègues du DEDIP, Olivier G., je te dois beaucoup à travers Caliste et IDeF-X, ainsi que pour les moments qu'on a passés ensemble (surtout à Sydney !), Éric pour avoir pris le temps d'assister à une de mes répétitions de soutenance et pour vos conseils, ainsi que Sébastien, que j'ai pu voir soutenir avec brio il y a cinq ans.

J'en viens aux collègues du List, Frédérick et Vincent, grâce à qui le projet ORIGAMIX a vu le jour et je les remercie pour le regard qu'ils ont porté sur mes travaux, avec des conseils pertinents sur les applications.

Je remercie aussi les membres de InTheArt, nos échanges ont été de précieuses pistes de réflexion. En particulier, Dominique et Slava, j'ai hâte de travailler avec vous. Et bien évidemment, Valérie, son organisatrice (je ne t'ai pas oubliée). Merci pour m'avoir accueilli au sein du groupe et de m'avoir donné l'opportunité de faire cette petite formation. Au-delà de l'aspect professionnel, j'espère que nous nous retrouverons pour nos longues discussions, où je te tiendrai au courant de mes vacances pluvieuses (et je te ferai mes petites remarques taquines dont j'ai le secret, je suis sûr que cela te manquerait sinon).

La transition est parfaite pour remercier un autre membre de InTheArt, Jean-Marc. Je te remercie d'avoir pensé à moi pour le poste au LGLS, pour lequel ce premier mois me montre déjà que les projets, très variés, s'annoncent enthousiasmants, et je te remercie aussi pour tes conseils lors de la préparation de ma soutenance. Merci à Erwan pour m'accueillir au sein de ton laboratoire et pour t'être impliqué dans la (fastidieuse) constitution de mon dossier. Je remercie aussi tous les membres du LGLS pour leur accueil chaleureux, l'intégration est bien facile avec des collègues aussi sympathiques.

Je remercie les collègues de 3D PLUS, Marie-Cécile et Valérie, c'est un plaisir d'avoir travaillé avec vous sur ce beau projet et de pouvoir continuer aussi pendant le LabCom. Et un remerciement tout particulier à Paul, je t'ai vu au début de ta thèse, j'ai commencé la mienne quand tu as brillamment fini la tienne, c'est toujours une joie de travailler avec toi et ton enthousiasme communicatif.

Un remerciement tout particulier aussi à Christelle, relectrice vraiment très courageuse du manuscrit. Aucune erreur ne t'échappe et il en ressort encore bien amélioré après être passé entre tes mains.

Je passe maintenant à mes camarades de galère. Laura, merci d'avoir prêté une oreille attentive pour les résumés de Koh-Lanta, ton petit mug à son effigie m'est toujours très utile, j'ai été ravi que ton aventure se termine avec succès, tu as tenu bon ! Francesco, dit général Ceraudo du fond du couloir et conseiller scientifique du Centre Spatial de Toulouse-Matabiau, c'était une joie de partager ton bureau ces deux années, pour nos discussions plus ou moins scientifiques, j'ai beaucoup appris avec toi et ta thèse s'est finie avec toute la réussite que tu mérites. David, tu as pris la relève dans le bureau pour les deux mois avant ta soutenance que tu as menée avec excellence, c'était un plaisir de partager ces moments et tu peux me contacter quand tu veux si tu as des questions à poser sur des spectres. Nicolas, on a commencé la thèse ensemble, et on la finit presque ensemble, j'ai hâte d'assister à ta soutenance. Floriane et Thibault, je vous souhaite aussi beaucoup de courage dans cette dernière ligne droite, ainsi que pour l'organisation dans ces conditions particulières. Benjamin, tu nous as épatés en stage et je sais que tu vas encore nous montrer ton talent avec ta thèse. Ion, qui finit sa première année de thèse, je te souhaite une belle aventure, et je suis sûr que tu arriveras à le faire fonctionner ce détecteur ! Enfin, Hugo, tu prends la relève, je pense qu'il n'est pas nécessaire de t'avertir que le plus gros risque que tu prends, c'est de passer d'excellents moments, je te souhaite de t'épanouir dans ce projet.

J'ai croisé le chemin de beaucoup de stagiaires, qui m'ont aussi marqué par leur passage, Corentin et le fameux fer à souder, ainsi que Thierry, je te dois beaucoup avec ton travail sur DiagnostiX, je crois que sans cela, je n'aurais pas pu réaliser une grande partie des mesures qui ont conduit aux résultats de ma thèse, je t'en remercie ; les stagiaires de Bertrand, Marion (qui m'a initié à Pékin Express, parce que je ne regardais pas encore suffisamment ce genre d'émission), Katia, Sarah et Hugo. Je vous souhaite beaucoup de réussite pour la fin de vos études et votre vie professionnelle.

Parmi les stagiaires, je remercie tout particulièrement Yann, mon premier stagiaire. La période n'a pas été facile avec le confinement mais tu as su rester motivé, curieux et assidu dans ton travail, ce qui nous a permis d'obtenir des résultats au-delà de nos attentes. Je t'en félicite encore, et j'espère que tu pourras continuer de parcourir un petit morceau de chemin avec nous.

Je tiens à remercier également mon jury de thèse pour avoir participé en direct à la présentation de mon travail, malgré la situation délicate, ainsi que pour leurs questions et leurs remarques, toutes très riches et qui m'ont permis de prendre de la hauteur sur mon travail.

Je remercie Pascale Delbourgo, pour les comités de thèse et son suivi au cours de ces trois années pour s'assurer du bon déroulement de la thèse et son regard pertinent sur mes travaux, ainsi que Françoise Prêteux, pour ses précieux conseils lors des comités de thèse, je vous dois aussi beaucoup dans la réalisation de cette aventure pour l'aide que vous avez apportée dans la constitution du dossier de thèse. J'en profite pour remercier le Corps des Ponts, des Eaux et des Forêts pour m'avoir autorisé à réaliser ce beau projet scientifique.

Sur un plan plus personnel, je remercie l'ensemble de mes amis pour leur soutien lors de ces trois années. En particulier, Frobs et Tibo, j'ai été heureux d'avoir été votre témoin pour votre très beau mariage. Edo, merci d'être venu assister à ma soutenance, cela me touche beaucoup, et merci de ne pas avoir posé de question embarrassante « This is more a comment than a question… ». William, je te souhaite beaucoup de courage pour ta dernière ligne droite dans ta thèse. Merci à toute la petite équipe, Pauline, Mohammed, Alexis, Wei, ainsi qu'à la team Koh-Lanta, Rémi (maître des chiffres et des statistiques), Albert et Paul, les Skype Lanta ont été des moments importants pour bien vivre le confinement.

Je vais profiter de ces remerciements pour avoir une pensée pour les personnes qui ont forgé ma scolarité et m'ont donné le goût de poursuivre dans toutes ces années d'études. Je n'ai jamais eu l'occasion de les remercier : pour les professeurs qui ont marqué mon parcours et m'ont poussé à donner le maximum, Elisabeth Klein-Nave, Joëlle Pottier, Nadine Kühnapfel, Daniel Noircler, Martin Canals, Catherine Pederzoli, Gaëlle-Suzanne Mulard et Charles Vix.

Il est bientôt temps de conclure cette longue liste. Je remercie bien sûr ma famille qui m'a soutenu depuis toujours. Mes parents Marie-Laure et Laurent, vous avez toujours cru en moi (« mais si, ne t'inquiète pas, tu vas y arriver ») et je vous dois évidemment ce que je suis devenu. J'ai été heureux de voir votre bonheur après la soutenance et j'espère vous en apporter encore à l'avenir. Ma petite sœur Lisa, tu es la fille la plus forte que je connaisse, j'ai puisé beaucoup d'énergie dans ton beau sourire qui ne te quitte pas malgré tout ce que tu as vécu, je serai toujours à tes côtés comme tu l'as été pour moi. Tu m'as soutenu depuis toute petite et tu as continué même ces dernières années. Je vous aime tous les trois et je vous embrasse fort.

Ma Wendy, tu ne crois pas que je vais t'épargner. Tu es arrivée dans ma vie il y a près de deux ans et tu l'as immédiatement illuminée. Tu as été présente à chaque instant de ma thèse, en particulier pour les moments les plus intenses, tu as même réussi l'exploit d'absorber mon stress légendaire juste avant la soutenance. Tout le bonheur que tu m'apportes est indescriptible, je te le dis et le répète à longueur de journée, et j'espère t'en amener autant en retour, pour toujours. Je t'aime si fort mon petit cœur.

Geoffrey

Table des matières

Résumé 1

Abstract 2

Remerciements 3

Table des matières 7

Introduction 13

Chapitre 1 : Caliste – De l'espace aux applications nucléaires 17

1 L'instrumentation X et gamma en astrophysique 18

1.1 L'astronomie X et gamma 18

1.2 Détecter les photons X et gamma 20

1.3 Vers la conception de détecteurs CdTe miniatures pixellisés 23

2 Application des détecteurs Caliste aux caméras gamma 27

2.1 Caliste-HD et Caliste-O 27

2.2 Vers des applications nucléaires : le prototype de caméra gamma WIX 31

2.2.1 Des caméras gamma pour la sûreté nucléaire issues de l'astrophysique des hautes énergies 31

2.2.2 WIX : un prototype de caméra gamma basé sur Caliste HD 34

2.2.3 Performances de la caméra WIX-HD et comparaison avec les autres caméras gamma 38

3 Objectifs de thèse 39

3.1 Défis à relever dans cette thèse 39

3.1.1 La spectro-identification automatique des radioéléments 40

3.1.2 L'imagerie à masque codé 40

3.1.3 La reconstruction Compton 41

3.2 Structuration du manuscrit 41

4 Bibliographie 42

Chapitre 2 : La vie d'un photon, de la source aux données 45

1 Observation des données et spectrométrie avec Caliste 46

1.1 Les données brutes issues de Caliste 46

1.1.1 La liste d'événements 46

1.1.2 Carte des coups et spectre 48

Carte de coups 48

Spectre 49

1.2 Étalonnage en énergie 50

1.2.1 Principe 50

1.2.2 Principe et loi de calibration 50

1.2.3 Détermination des paramètres de calibration par la méthode ECC (Energy Calibration via Correlation) 51

1.2.4 Maximisation de la corrélation par recherche exhaustive et algorithme génétique 53

ECC-AMR 54

Algorithme génétique 55

1.2.5 Résultats de calibration 58

Spectre 58
Spectro-image 60
1.2.6 Calibration non affine 61
1.3 Conclusion et nécessité de construire un modèle 63

2 Modélisation des interactions photons/matière, de la source au détecteur 64
2.1 Simulation directe Monte-Carlo 64
2.2 Modélisation des diffusions Compton multiples 68
2.2.1 Diffusion Compton simple 68
2.2.2 Diffusions Compton multiples 70

3 Modélisation de la réponse du détecteur 74
3.1 Fluctuation du nombre de paires électrons-trous 74
3.2 Formation du signal 76
3.2.1 Induction du signal 76
3.2.2 Transport et perte de charge 77
3.2.3 Efficacité d'induction de charges 77
3.2.4 Résolution numérique de l'efficacité d'induction de charges (CIE) 78
Calcul du potentiel électrique et du potentiel de pondération 78
Calcul de l'efficacité d'induction de charges 82
3.2.5 Prise en compte de l'efficacité d'induction de charges dans la réponse du détecteur 85
3.3 Bruit électronique 85

4 Modélisation des spectres, simulations et réponse du détecteur 86
4.1 Utilisation de la simulation Geant4 et de la réponse du détecteur 86
4.2 Application à la simulation de spectre 89
4.2.1 Application et comparaison avec des mesures réelles 89
4.2.2 Utilisation des structures Compton multiples 90
4.3 Conclusion et utilité des simulations 92

5 Bibliographie 92

Chapitre 3 : Spectro-Identification automatique 95

1 Contexte de la spectro-identification 96
1.1 Problème à traiter 96
1.1.1 Description du problème 96
1.1.2 Données utilisées 96
1.1.3 Hypothèses et contexte 97
1.2 État de l'art 98
1.2.1 Algorithmes de recherche et d'ajustements de pics 98
1.2.2 Méthodes basées sur des modèles de spectres 99
1.2.3 Algorithmes de machine learning 100
1.3 Orientation vers des algorithmes de *deep learning* 103

2 Les réseaux de neurones 104
2.1 Calcul de la sortie du réseau : *forward pass* 104
2.1.1 Neurone fully-connected 104
2.1.2 Neurone de convolution 106
2.1.3 Réseau de neurones 107
2.2 Apprentissage des réseaux de neurones 109
2.2.1 Objectif et fonction de coût 109
2.2.2 Descente de gradient et backpropagation 110

2.2.3 Surveillance des performances et adaptation des hyperparamètres 111

3 Apprentissage de réseaux de neurones convolutifs pour la spectro-identification de radioéléments 112

3.1 Méthodologie mise en place 112

3.1.1 Indentification, quantification et bibliothèque de sources 112

3.1.2 Construction de la base de données 115

Application des simulations 115

Augmentation de données 116

3.1.3 Normalisation des spectres 118

3.2 Incertitudes et réseaux de neurones bayésiens 121

3.3 Construction des réseaux de neurones et apprentissage 123

4 Application des réseaux de neurones à la spectro-identification de radioéléments 127

4.1 Constitution de la base de données de test 127

4.2 Mise en œuvre des réseaux et quantification des performances 130

4.2.1 Représentation des résultats et cas d'application 130

4.2.2 Comment « réfléchit » mon réseau de neurones convolutif ? 134

4.3 Évaluation des performances des réseaux de neurones d'identification 135

4.3.1 Precision, Recall, Accuracy 135

4.3.2 Adaptation du seuil de décision 137

4.3.3 Sensibilité pour les sources individuelles 142

4.3.4 Performances sur la quantification 144

5 Application à la dosimétrie avec Caliste 148

5.1 Débit de dose : définition et calcul 148

5.2 Mesure du débit de dose avec Caliste 149

5.2.1 Principe de la mesure 149

5.2.2 Inversion Compton 150

5.2.3 Prise en compte de l'efficacité de détection 153

5.2.4 Application au calcul de la dose 154

5.3 « Démélange » de sources et dose par radioéléments individuels 155

6 Conclusion 161

7 Bibliographie 162

Chapitre 4 : Imagerie à masque codé – Présentation et méthodes classiques 167

1 L'imagerie à masque codé 168

1.1 Principe de ce système d'imagerie 168

1.2 Masque utilisé 169

1.3 Problème d'inversion 174

2 Algorithmes de localisation existants 176

2.1 Algorithme de déconvolution 176

2.2 *Maximum Likelihood Expectation Maximization* (MLEM) 178

3 Application des algorithmes de déconvolution et MLEM, aux données Caliste 181

3.1 Tests sur des sources ponctuelles 181

3.1.1 Quantification des résultats : CNR (Contrast-to-Noise Ratio) et GCR (Ghost Contrast Ratio) 182

3.1.2 Reconstruction d'images pour une source ponctuelle unique 183

3.1.3 Reconstruction d'images pour des sources ponctuelles multiples 187

Sources multiples avec différents radioéléments 187

Sources multiples d'un même radioélément 189
3.2 Sources étendues 192
3.2.1 Simulations de sources 192
3.2.2 Localisation de sources étendues issues de simulations 193
Quantification des résultats 193
Tests sur des sources étendues 194

4 Approfondissements sur l'algorithme MLEM 199
4.1 Détermination d'un critère d'arrêt objectif 199
4.2 Détermination de la distance à la source 200

5 Conclusion sur les algorithmes classiques et leurs limitations 203

6 Bibliographie 204

Chapitre 5 : Imagerie à masque codé – Méthodes de deep learning 207

1 Base de données et construction de réseaux de neurones 208
1.1 Génération continue d'une base de données 208
1.2 Architecture du réseau de neurones et apprentissage 211

2 Performances des réseaux de neurones sur la problématique de la localisation de sources par masque codé 214
2.1 Tests sur des données réelles avec des sources ponctuelles 214
2.1.1 Source ponctuelle unique dans l'axe 214
2.1.2 Source ponctuelle unique hors axe 216
Comparaison de sensibilité entre l'algorithme MLEM et le réseau de neurones 219
2.1.3 Sources ponctuelles multiples 220
Cas de deux sources ponctuelles hors axe, même radioélément 220
Cas de cinq sources ponctuelles, même radioélément 221
2.2 Tests sur des données simulées de sources étendues 223
2.2.1 Cas d'une source étendue 2°x2° 223
2.2.2 Cas d'une source étendue 4°x4° 224
2.2.3 Cas d'une source étendue 7°x7° 226
2.2.4 Cas d'une source étendue de forme et taille arbitraires 227
2.3 Conclusion sur l'application des réseaux de neurones à la reconstruction par masque codé 228

3 Superposition optique et sources étendues réelles 230
3.1 Superposition image gamma et image optique 230
3.2 Cas de sources étendues 234

4 Conclusion 240

5 Bibliographie 241

Chapitre 6 : Localisation de sources radioactives par imagerie Compton 243

1 Principe de la reconstruction Compton 244
1.1 Diffusion Compton et élargissement Doppler 244
1.2 Application de la diffusion Compton à la localisation de sources radioactives 247

2 Reconstruction Compton avec Caliste 249
2.1 Sélection des événements Compton 249
2.2 Ordre d'interaction et simulations 251
2.2.1 Problématique de l'ordre d'interaction 251

2.2.2 Simulations d'événements Compton 253
2.2.3 Validation de la méthode 256
2.3 Backprojection directe 257
2.3.1 Système de coordonnées 257
2.3.2 Calcul et projection des cônes sur la sphère 258
2.3.3 Application de la backprojection directe à des données réelles acquises avec Caliste 261

3 Algorithmes de traitement Compton 265
3.1 Algorithme LM-MLEM : *List-Mode Maximum Likelihood Expectation Maximization* 266
3.1.1 Principe de la méthode 266
3.1.2 Test LM-MLEM avec des données Caliste 268
3.2 SOE-RR : *Stochastic Origin Ensemble Resolution Recovery* 270
3.2.1 Principe de la méthode 270
3.2.2 Tests SOE-RR sur des données réelles acquises avec Caliste 271
3.2.3 Soustraction itérative de sources avec l'algorithme SOE-RR 276
3.3 Algorithme bayésien de reconstruction Compton 279
3.3.1 Principe de la méthode 279
3.3.2 Tests sur des données réelles acquises avec Caliste 280
3.4 Reconstruction par Inversion Compton 284
3.4.1 Principe de la méthode d'inversion Compton 284
3.4.2 Test de la reconstruction par inversion Compton avec des données réelles issues de Caliste 285
3.4.3 Test de l'inversion Compton sur une simulation de source étendue 288
3.5 Conclusion sur les algorithmes 289

4 Perspectives : vers des algorithmes de Deep Learning 291
4.1 Mise en place du problème à traiter et développement de réseaux de neurones 291
4.2 Test du réseau de neurones sur des données de simulation 293
4.3 Résultats sur des données réelles de sources ponctuelles uniques par deep learning 294

5 Conclusion 296

6 Bibliographie 297

Conclusion et perspectives 301

Annexe A : Interactions photon-matière 307

1 Les interactions photon-matière dans le domaine X et gamma 307
1.1 L'absorption photoélectrique 307
1.2 La diffusion Compton 308

2 Probabilité d'interaction 311

3 Bibliographie 313

Annexe B : Réseaux de neurones et backpropagation 315

1 Rappel du problème et notations 315

2 Calcul des dérivées partielles de la fonction de coût suivant les paramètres 316
2.1 Calcul sur la dernière couche du réseau de neurones 317
2.2 Calcul des dérivées partielles pour les autres couches 318
2.3 Conclusion sur l'algorithme de *backpropagation* 319

3 Bibliographie 319

Annexe C : Utilisation de deux sorties complémentaires softmax _____ *321*

Annexe D : L'algorithme MLEM _____ *325*

1 Problème et notations _____ **325**

2 Approche bayésienne _____ **325**

3 Algorithme MLEM et convergence _____ **327**

3.1 Calcul de la dérivée de la vraisemblance _____ 328

3.2 Concavité de la log-vraisemblance et existence d'un maximum _____ 330

3.3 Algorithme de recherche du maximum de vraisemblance _____ 331

3.4 Démonstration que le point de convergence maximise la vraisemblance _____ 332

4 Conclusion _____ **335**

5 Bibliographie _____ **335**

Annexe E : Reconstruction Compton sur des données IBIS/INTEGRAL _____ *337*

1 L'instrument IBIS _____ **337**

2 Application de la reconstruction Compton à un GRB observé par IBIS _____ **338**

2.1 Étude spectrométrique _____ 338

2.2 Reconstruction Compton _____ 340

3 Bibliographie _____ **342**

Introduction

La compréhension de l'Univers passe nécessairement par son observation, qui consiste en l'acquisition d'informations provenant d'objets ou de phénomènes lointains dans le cadre de l'astronomie. Ces informations sont en particulier portées par les ondes électromagnétiques, que l'on appelle « lumière » dans le langage commun. Le spectre de la lumière est subdivisé en différents domaines suivant sa longueur d'onde : des ondes radio pour les grandes longueurs d'onde, aux rayons X et gamma pour les petites longueurs d'onde, en passant par la lumière visible, bande infime du spectre électromagnétique à laquelle les humains ont accès.

L'observation de la lumière dans chacun de ces domaines apporte des informations complémentaires sur les objets et les phénomènes contenus dans notre Univers, de l'observation des amas de gaz interstellaire dans le domaine des rayons Infrarouges à celle des phénomènes les plus violents de l'Univers comme les sursauts gamma. Observer la lumière, cela signifie la détecter dans un premier temps, puis d'en acquérir des informations, comme mesurer sa longueur d'onde, c'est-à-dire effectuer de la spectrométrie, et en déterminer sa provenance, on parle d'imagerie. Or, la détection, la spectrométrie et l'imagerie de la lumière demandent à relever des défis propres à chaque domaine du spectre électromagnétique.

En particulier, les rayons X ou gamma, rayons à faible longueur d'onde ou aussi appelés rayons de haute énergie, présentent la particularité d'avoir une faible probabilité d'interaction avec la matière, ce qui rend complexe la détection même de ces photons de haute énergie. Il faut ajouter à cela que l'atmosphère terrestre, du fait de son épaisseur, est complètement opaque pour ces longueurs d'onde, ce qui contraint à l'utilisation de satellites observationnels pour l'astronomie X et gamma.

Dans ce cadre, le département d'astrophysique du CEA a développé une expertise en instrumentation pour la détection de photons de haute énergie, en participant à la conception de l'instrument SIGMA à bord du satellite soviétique Granat lancé en 1989 ou encore à celle de l'instrument IBIS du satellite INTEGRAL, lancé en 2002, dans une collaboration entre l'ESA, la NASA et l'agence spatiale russe. L'instrument IBIS embarque en particulier deux détecteurs de photons de haute énergie, un détecteur en semi-conducteur CdTe (Tellurure de Cadmium) et un détecteur à scintillation CsI (Iodure de Césium). Tous deux ont la capacité de détecter des photons de haute énergie, dans le domaine X et gamma, de mesurer l'énergie de ces photons pour la spectrométrie et de déterminer la position d'interaction des photons dans le détecteur pixellisé pour l'imagerie : on parle alors de spectro-imageurs. Ces plans de détection doivent être couplés à une optique afin d'être capable de déterminer l'origine des photons. Les lentilles utilisées pour la lumière visible ne focalisent pas les photons à ces énergies et d'autres systèmes d'imagerie doivent être utilisés. L'instrument IBIS embarque par exemple un masque codé qui est un système d'imagerie dite indirecte dans le sens où l'image formée sur le détecteur n'est pas une représentation directe du ciel observé, mais une projection qui doit être inversée pour retrouver l'image du ciel.

À la suite de ces expériences, les laboratoires d'instrumentation du département d'astrophysique du CEA ont poursuivi les développements sur les détecteurs semi-conducteurs, qui présentent l'avantage d'avoir une meilleure résolution spectrale que les détecteurs à scintillation, ainsi que la capacité de détecter des photons à plus basse énergie. Ces développements ont mené à la conception d'une

famille de détecteurs, nommée Caliste, qui se classe parmi les meilleurs détecteurs CdTe du point de vue des performances spectrométrique. La version Caliste-HD atteint une résolution de 670 eV à 60 keV pour un seuil bas inférieur à 2 keV. Ce sont aussi des détecteurs miniatures à haute densité de pixels, avec 16x16 pixels au pas de 625 µm, pour une épaisseur de cristal de 1 à 2 mm, à mettre en comparaison aux 16 384 pixels au pas de 4 mm du détecteur CdTe d'INTEGRAL. En outre, les détecteurs Caliste présentent un fonctionnement optimal à une température inférieure à -10°C, nécessitant donc un refroidissement pouvant être assuré par des systèmes embarqués comme des modules Peltier.

Ces performances et cette miniaturisation de détecteurs de photons X et gamma présentent un intérêt tout particulier pour des applications en astronomie. La mission Solar Orbiter, lancée en février 2020, embarque les premiers détecteurs Caliste pour une mission spatiale, servant de plans de détection pour l'instrument STIX muni d'un système d'imagerie à transformée de Fourier par motifs de Moiré. Cette mission a pour objectif d'observer des éruptions solaires, ou *flares* solaires, dans le domaine des rayons X, et en a détecté plus d'une soixantaine lors des quatre premiers mois suivant le lancement.

Les caractéristiques de Caliste permettent aussi d'envisager des applications qui sortent de l'astronomie, vers des domaines qui nécessitent la détection de photons X et gamma. Deux domaines d'application ressortent principalement dans ce cadre. D'une part, il existe des besoins en imagerie médicale nucléaire, à travers par exemple le suivi de traceurs radioactifs, la tomographie à émission de positons ou encore la mesure de dose dans les traitements par radiothérapie. Ainsi, Caliste a été à la base du projet SATBOT entre 2015 et 2018, qui visait à mesurer la dose reçue par des cellules cancéreuses du sein traitées par radiothérapie, en détectant les photons de fluorescence X émis par des nanoparticules d'or injectées préalablement dans ces cellules tumorales.

D'autre part, la sûreté nucléaire représente un secteur majeur dans lequel les systèmes de spectro-imagerie ont la capacité de fournir des informations précieuses à travers différentes applications, comme la caractérisation de scènes radiologiques accidentelles et post-accidentelles, de scènes d'assainissement-démantèlement ou de diagnostic d'environnements de travail en zone réglementée du point de vue de la radioprotection. En particulier, suite à l'accident de Fukushima en 2011, cet axe d'instrumentation a connu un nouvel essor à travers le développement d'imageurs X et gamma, connus sous le nom de caméra gamma.

Entre 2014 et 2018, le projet ORIGAMIX a visé au développement de caméras gamma de nouvelle génération pour la sûreté nucléaire, en visant l'utilisation des détecteurs Caliste-HD au sein d'un prototype de caméra gamma nommé WIX-HD. L'objectif de ce projet est d'utiliser la technologie de pointe en matière de détecteurs, développée pour l'astronomie astrophysique afin de la transférer vers des applications industrielles.

Ainsi, la caméra gamma hérite de cette technologie à travers le détecteur Caliste, et il est nécessaire aussi de développer les méthodes de traitements de données adaptées pour tirer au mieux profit des performances de ces détecteurs et répondre aux besoins que doivent couvrir la caméra gamma. En particulier, dans le cadre d'applications dans le domaine nucléaire, ce système de spectro-imagerie doit être capable de fournir deux informations essentielles : l'identification et la quantification des radioéléments présents dans la scène observée, à partir des informations spectrométriques, d'une part, et leur localisation dans la scène à partir des informations en imagerie d'autre part.

C'est dans le cadre du développement de traitements de ces données que s'inscrit mon travail de thèse présenté dans ce manuscrit. Dans un premier temps, à partir de la mise en œuvre du détecteur Caliste au sein de la caméra WIX, j'identifie les données qui doivent être traitées en les interprétant pour comprendre les phénomènes physiques qui sont à leur origine, notamment les processus d'interaction photon-matière et l'acquisition du signal au sein du détecteur CdTe. Cette étude passe par la simulation de la trajectoire des photons entre leur source et le détecteur ainsi que par la modélisation de la réponse du détecteur suite à un dépôt d'énergie par les photons dans le cristal CdTe. Cette étape est essentielle pour développer des méthodes d'identification et de quantification de radioéléments à travers l'utilisation des données spectrométriques.

Je m'intéresse à la spectro-identification des radioéléments en mettant en œuvre des algorithmes nouveaux en matière de traitements de données, issus de l'Intelligence Artificielle, et plus particulièrement du domaine du *Deep Learning* et des réseaux de neurones. Ces méthodes ont fait leurs preuves dans le cadre de l'analyse d'images ou encore de bandes sonores et présentent un potentiel très intéressant pour des applications pour le traitement de données scientifiques. Sur la base des simulations et de la modélisation du détecteur que j'ai mises en place, j'ai créé une base de données entièrement synthétique pour entraîner des réseaux de neurones bayésiens à couches de convolution.

Ces réseaux de neurones ont un objectif triple qui est d'identifier les radioéléments à l'origine des photons détectés, de quantifier les contributions respectives de chaque radioélément au signal enregistré et enfin de fournir une évaluation de l'incertitude sur les prédictions effectuées. Je porte une attention particulière à tester ces algorithmes avec des mesures réelles issues d'acquisitions effectuées avec la caméra gamma WIX, alors qu'ils ont été entraînés uniquement sur des données simulées. En particulier, j'évalue leurs performances dans des cas où le signal enregistré contient une faible statistique de comptage de photons, ainsi que de possibles dérives de l'étalonnage du détecteur, permettant d'effectuer une analyse automatique et en temps réel de la scène observée.

Ensuite, j'étudie une méthode d'imagerie basée sur l'utilisation de systèmes à masque codé, appliquée à la caméra WIX. Cette méthode d'imagerie indirecte nécessite des traitements particuliers de l'image enregistrée sur le détecteur, à travers la connaissance du masque utilisé et de sa projection sur le détecteur. Je m'intéresse aux algorithmes classiquement utilisés pour ce problème, à savoir un algorithme de déconvolution qui analyse la corrélation entre le masque et la projection effectivement enregistrée, et un algorithme nommé MLEM qui maximise la vraisemblance d'observer les données en fonction de la position des sources. En testant sur des données réelles issues du détecteur Caliste, je montre les performances de ces algorithmes dans mon cadre d'application, notamment leur sensibilité et leur précision pour des sources ponctuelles dans le champ de vue du masque. J'étudie aussi les défis à relever pour ces algorithmes, en particulier sur la question de l'imagerie de sources étendues, c'est-à-dire avec un rayon angulaire grand devant les éléments du masque et les pixels du détecteur, cas pour lesquels les algorithmes sont incapables de fournir une information de localisation.

Je développe une méthode basée sur des algorithmes de *Deep Learning* afin de dépasser ces limitations. J'effectue l'apprentissage de ces algorithmes à l'aide de données issues de simulations et je les applique ensuite à des données réelles issues du détecteur. Je montre leurs capacités de reconstructions dans les cas de sources ponctuelles et je les compare à celles des algorithmes

classiques. Puis, je montre le potentiel de ces algorithmes pour relever le défi de l'imagerie de sources étendues en les appliquant de nouveau à des données mesurées avec Caliste.

Enfin, j'étudie une autre méthode d'imagerie, basée sur le phénomène de diffusion Compton, qui permet de traiter des sources de photons de plus haute énergie pour lesquelles l'imagerie à masque codé devient difficile à appliquer, le masque devenant de plus en plus transparent lorsque l'énergie des photons augmente. La géométrie de détecteur que j'utilise n'est cependant pas favorable à l'imagerie Compton. En effet, j'utilise un seul détecteur plan dans lequel doivent s'effectuer les deux interactions, diffusion Compton et absorption photoélectrique, nécessaires pour la reconstruction Compton. Cela implique une sensibilité du détecteur très peu optimale pour la reconstruction Compton et demande une sélection minutieuse des événements à traiter pour réduire au maximum le bruit dans la reconstruction. Par ailleurs, du fait des incertitudes liées à cette géométrie, la reconstruction Compton génère un bruit intrinsèque qui empêche de localiser les sources avec précision et de séparer deux sources proches.

Dans ce cadre, j'ai développé des algorithmes de reconstruction Compton particuliers pour tenir compte de cette géométrie et améliorer la précision sur la reconstruction. J'évalue aussi ces algorithmes sur des données réelles issues du détecteur Caliste et je compare leurs performances à celles d'algorithmes issus de l'état de l'art sur ce sujet. Je propose pour finir une application des réseaux de neurones à cette problématique, en évaluant leur potentiel sur des données réelles et en envisageant la possibilité d'améliorer la sensibilité pour la reconstruction Compton.

Le développement de l'ensemble des méthodes que j'ai mises en œuvre, pour la spectro-identification, l'imagerie à masque codé et l'imagerie Compton, a été guidé par l'application sur les détecteurs Caliste. Cependant, j'ai pris le soin d'effectuer des développements qui soient généralisables à d'autres types de détecteurs, avec les adaptations propres au système de détection utilisé. Par ailleurs, les applications visées ici concernent principalement le domaine de la sûreté nucléaire, mais des retombées de ces travaux pourront être envisagées sur d'autres secteurs d'application, en particulier en imagerie médicale, ou encore avec un retour sur les données d'observation astrophysique.

Chapitre 1 : Caliste – De l'espace aux applications nucléaires

1 L'instrumentation X et gamma en astrophysique _____ 18

2 Application des détecteurs Caliste aux caméras gamma _____ 27

3 Objectifs de thèse _____ 39

4 Bibliographie _____ 42

Ce premier chapitre me permet d'introduire le contexte dans lequel s'inscrit mon travail de thèse. Les détecteurs spectrométriques pixélisés à base de CdTe Caliste que j'utilise sont issus de développements pour l'observation X et gamma en astrophysique spatiale, domaine que je présente dans une première partie pour comprendre la logique de développement qui a conduit aux détecteurs que j'utilise dans le cadre de mes travaux.

Dans un second temps, je décris plus précisément les caractéristiques et les performances de ces détecteurs. Puis, je montre les motivations, basées sur ces performances, pour un transfert de ces technologies dans le cadre de nouvelles applications, en particulier liées au domaine nucléaire, pour la conception de caméras gamma portables, bénéficiant tant des performances particulières des détecteurs Caliste que de leur miniaturisation.

Enfin, à travers la compréhension des limites des systèmes existants, j'identifie les verrous et questions scientifiques que j'étudie et que je tente de résoudre, en m'intéressant plus particulièrement à l'analyse de données d'acquisition par des méthodes innovantes. Je décris précisément les différentes problématiques que je vais aborder au long de ce manuscrit.

1 L'instrumentation X et gamma en astrophysique

1.1 L'astronomie X et gamma

L'astronomie et l'astrophysique trouvent leur origine à travers l'observation du ciel. L'observation, initialement faite à l'œil nu, a ensuite été complétée par des instruments comme la lentille astronomique de Galilée ou le télescope de Newton, marquant les premiers pas de l'instrumentation en astronomie. Ces instruments avaient pour objectif de former une image visible « grossie » de certaines régions du ciel, afin d'en tirer des informations, comme la découverte de nouveaux astres.

La compréhension des rayonnements électromagnétiques montre cependant que la lumière visible ne représente qu'une infime partie de l'ensemble du spectre électromagnétique, ce qui implique que l'observation dans le domaine visible ne permet d'avoir accès qu'à une information très partielle des régions du ciel observées et des phénomènes physiques à l'œuvre.

Ainsi, au cours du XXème siècle, de nouveaux instruments ont été développés avec l'objectif d'observer le ciel à différentes longueurs d'onde. La Figure 1-1 présente le spectre électromagnétique et l'observation à diverses longueurs d'onde de la nébuleuse du Crabe, vestige d'une supernova qui a eu lieu au XIème siècle. On parle d'astronomie multi-longueur d'ondes ou multi-spectrale.

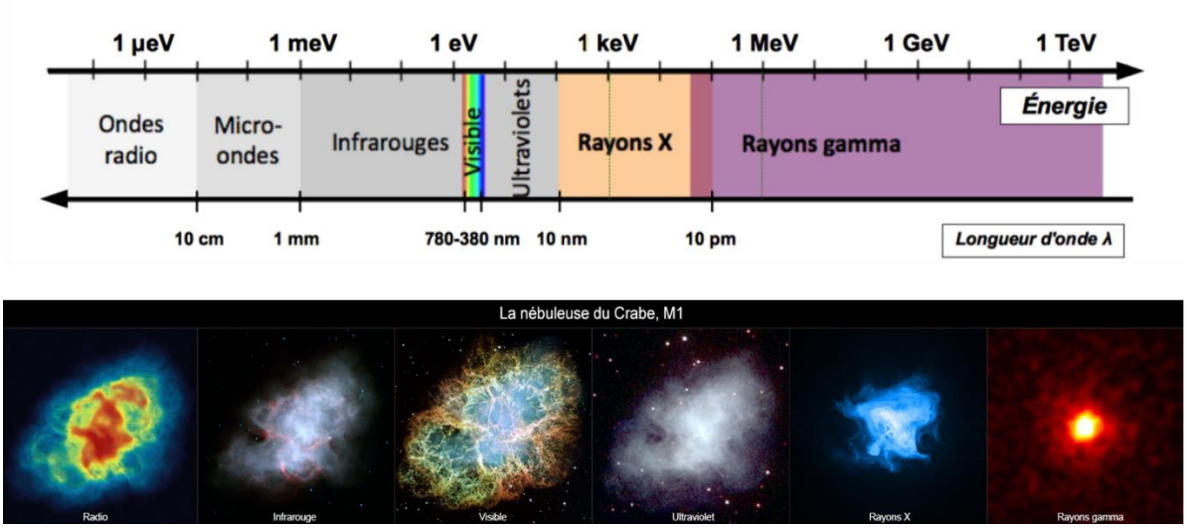


Figure 1-1 : En haut : Spectre électromagnétique (Dubos, 2015). En bas : Nébuleuse du Crabe observée dans différentes longueurs d'onde (Astrosurf, 2016).

Suivant le domaine du spectre électromagnétique observé, on peut accéder à des informations sur différents phénomènes physiques à l'œuvre dont l'interprétation conduit à la compréhension profonde de notre Univers. Par exemple, le domaine infrarouge permet d'étudier les zones froides de l'Univers comme les amas de gaz interstellaire, les zones dans lesquelles on observe les formations d'étoiles, ou encore les nébuleuses planétaires. Le domaine visible est adapté à l'observation des étoiles, à leur position, et l'étude plus fine du spectre visible renseigne sur la température des étoiles et leur masse, dont la classification a abouti au diagramme de Hertzsprung-Russell. Par ailleurs, l'étude précise des variations de luminosité des étoiles est aussi une méthode de détection d'exoplanètes.

Le domaine des rayons X et gamma représente le domaine des rayonnements de haute énergie. Ils sont une signature des phénomènes les plus violents de l'Univers. Ils sont émis par exemple par la collision d'objets compacts, comme deux étoiles à neutrons, vestiges d'étoiles massives. Ils peuvent aussi être émis par des supernovæ, explosions d'étoiles massives en fin de vie, en particulier à travers les chocs internes entre les différentes couches de l'explosion et à travers les chocs externes avec la matière environnante. Les rayonnements de haute énergie émis par ces phénomènes sont appelés sursauts gamma et font l'objet de nombreuses études à travers diverses missions spatiales comme SWIFT (NASA, 2000), mission lancée en 2004, ou encore la mission SVOM (SVOM, 2015) dont le lancement est prévu en 2021. Les missions d'observation de ces phénomènes cosmiques de haute énergie visent maintenant à combiner leurs informations avec celles issues des détecteurs d'ondes gravitationnelles, qui ont fait leur preuve récemment en détectant une telle onde émise par la fusion de deux étoiles à neutrons le 17 août 2017 (Abbott & al., 2017). La combinaison des informations issues du rayonnement électromagnétique et des ondes gravitationnelles, ainsi que de la détection de neutrinos, forme les prémices de ce qu'on appelle désormais l'astronomie multi-messagers. Cette nouvelle discipline constitue l'avenir proche de l'astrophysique moderne.

D'autres phénomènes astrophysiques peuvent aussi être à l'origine de l'émission du ciel en rayons X et gamma. Les disques d'accrétion, matière autour des objets compacts, peuvent atteindre des températures suffisantes pour émettre des rayons X. Par ailleurs, les particules chargées dans ces environnements extrêmes peuvent émettre un rayonnement X synchrotron lorsqu'ils sont soumis à un champ magnétique intense d'une étoile à neutrons en rotation rapide, appelée pulsar. Un tel exemple est donné par l'observation en rayons X de la nébuleuse du Crabe, présentée en Figure 1-1.

Dans le cadre de l'observation X et gamma, on adopte généralement la description corpusculaire des rayonnements électromagnétiques. On étudie alors les quanta de lumière, appelés photons, que l'on caractérise par une énergie E, reliée à la longueur d'onde λ du rayonnement, par la célérité de la lumière c et la constante de Planck h :

$$E = \frac{hc}{\lambda} \tag{1-1}$$

L'unité utilisée pour mesurer l'énergie E est l'électron-volt (eV)[*], énergie d'un électron dans un potentiel de 1 V, soit $1{,}6.10^{-19}$ J. Les photons de la lumière visible ont une énergie de l'ordre de quelques eV. Les photons de haute énergie, X ou gamma, ont des énergies allant de quelques keV à plusieurs MeV.

Les rayons X et gamma se distinguent par leur origine. Les rayons X ont une origine électronique : ils peuvent être émis par la désexcitation d'électrons au sein d'atomes ou lors de la modification de la trajectoire d'un électron dans un champ magnétique (rayonnement synchrotron) ou encore par freinage des électrons dans un milieu, appelé *Bremsstrahlung*. Les rayons gamma ont une origine nucléaire et sont émis lors du passage d'un noyau à un niveau d'énergie inférieur.

La distinction entre photons X et gamma ne se fait pas sur leurs énergies, même si typiquement, les photons X sont moins énergétiques que les photons gamma. Pour observer ce recoupement en énergie, on peut par exemple voir que la fluorescence X $\mathrm{K}\beta_1$ de l'Américium se fait à une énergie de

[*] L'unité dans le Système Internationale est le Joule (J), mais vu les ordres de grandeurs en jeu, j'utiliserai par la suite l'unité électronvolt comme il est d'usage dans la communauté scientifique de l'instrumentation gamma.

120,3 keV alors qu'un de ses isotopes radioactifs, le ^{241}Am, a une ligne d'émission gamma à 26,3 keV. De même, les photons X produits par les générateurs X peuvent atteindre des énergies au-delà de 200 keV.

Je finirai en soulignant l'intérêt de ces photons dans le cadre d'applications nucléaires. Avec les phénomènes de radioactivité, les noyaux radioactifs se désintègrent pour devenir des radio-isotopes plus stables, en transformant des protons en neutrons (radioactivité β^+) ou des neutrons en protons (radioactivité β^-) ou encore en émettant un noyau d'hélium, deux protons et deux neutrons (radioactivité α). Le nouveau noyau obtenu suite à ces désintégrations se trouve en général dans un état excité, et sa désexcitation s'accompagne alors d'une émission gamma, d'où l'intérêt des photons gamma pour détecter des sources radioactives. Par ailleurs, les atomes fils sont aussi soumis aux rayonnements qui peuvent provoquer une fluorescence X au niveau des électrons de ces atomes fils eux-mêmes, dont on peut se servir pour détecter des sources radioactives.

Ainsi, des instruments permettant de constituer des images à partir des photons X et gamma sont utiles pour localiser des sources radioactives, que ce soit des isotopes produits lors de supernovæ pour l'astronomie ou des isotopes naturels ou artificiels sur Terre. En outre, des instruments capables de mesurer l'énergie des photons détectés peuvent être utilisés pour identifier les éléments à l'origine de ces photons, les lignes d'émission gamma et de fluorescence X étant des signatures de chaque radioélément.

1.2 Détecter les photons X et gamma

L'observation de photons X et gamma représente en réalité un défi instrumental sur différents aspects.

Le premier aspect est spécifique à l'observation astronomique : l'atmosphère est opaque aux photons X et gamma. Par conséquent, il est nécessaire d'utiliser des télescopes spatiaux pour étudier les astres et phénomènes physiques à haute énergie. Cet aspect apporte des contraintes de dimensionnement et des contraintes opérationnelles importantes. La masse et la taille des instruments doivent être minimisées pour être embarqués à bord du satellite. La consommation en énergie doit être optimisée sur les instruments ainsi que pour les systèmes nécessaires à leur fonctionnement. Enfin, les instruments doivent être conçus avec une grande fiabilité et qualifiés pour des missions spatiales.

Le deuxième défi est celui de la détection des photons[*]. Pour être capable de détecter ces photons de haute énergie, il faut disposer d'un matériau dans lequel la probabilité d'interaction des photons est grande, on parle d'efficacité de détection. Ainsi, les matériaux avec un numéro atomique élevé ou avec une forte densité sont capables d'arrêter plus efficacement les photons X et gamma. Pour augmenter l'efficacité de détection, on peut aussi augmenter le volume de détection, augmentant dans le même temps la taille des instruments à utiliser.

Le troisième aspect, relié aussi à la détection des photons, est la capacité du système à enregistrer et transmettre un signal électrique mesurable lors de l'interaction des photons avec la matière. Les photons X et gamma sont des rayonnements dits « ionisants » et sont capables d'arracher des

[*] Les phénomènes physiques liés à l'interaction photon-matière sont détaillés en Annexe A.

électrons aux atomes du matériau et d'ioniser la matière. Divers instruments basés sur ce phénomène ont été développés :

- les scintillateurs, basés sur une ionisation indirecte, dans laquelle le photon X ou gamma excite les atomes du matériau, qui émettent un rayonnement UV ou visible en se désexcitant appelé scintillation, qui est ensuite elle-même détectée et convertie en électrons par un photodétecteur (photomultiplicateur, SiPM, …) ;
- les détecteurs gazeux, basés sur une ionisation directe, dans lesquels l'ionisation produit des charges, électrons et ions, qui sont collectées par application d'un champ électrique intense dans le gaz ;
- les détecteurs solides à semi-conducteurs, basés aussi sur une ionisation directe, dans lesquels l'interaction des photons avec le semi-conducteur produit des paires électrons-trous en fournissant une énergie suffisante aux électrons du matériau pour passer de la bande de valence à la bande de conduction du semi-conducteur.

Les scintillateurs sont en général utilisés en spectrométrie gamma pour des applications industrielles, avec une efficacité élevée au détriment d'une résolution spectrale plus faible que les autres types de détecteurs, typiquement de l'ordre de 7 % à 662 keV (ORTEC, 2019), pour un coût modéré et abordable d'où leurs applications industrielles.

Les détecteurs solides à semi-conducteurs offrent une alternative intéressante aux scintillateurs, avec une bonne efficacité puisque les matériaux utilisés sont généralement assez denses avec un numéro atomique élevé, en comparaison des détecteurs à gaz, ce qui permet de réduire le volume de détection. La conversion étant directe, leurs performances spectrales sont généralement meilleures que pour les scintillateurs. Parfois compatible avec des procédés de fabrication industriels des semi-conducteurs, leur utilisation rend également possible des géométries complexes avantageuses pour la spectrométrie et l'imagerie.

Parmi les détecteurs semi-conducteurs, les détecteurs Silicium ont un numéro atomique de 14 et une densité de 2,33, ce qui les rend peu efficaces pour détecter des photons au-delà de quelques dizaines de keV. Cependant, dans la bande 100 eV-10 keV, ils montrent de très bonnes performances spectrométriques, par exemple de l'ordre de 150 eV FWHM[*] (*Full Width at Half Maximum*) à 6 keV, à une température de -50°C pour le détecteur de l'instrument MXT (*Microchannel X-ray Telescope*) de la mission SVOM (Ceraudo, 2019). Pour des photons à des énergies plus élevées, les détecteurs Germanium présentent de bonnes capacités de détection avec un numéro atomique de 32 et une densité de 5,33, donnant une efficacité de détection de 77 % à 100 keV pour une épaisseur de 5 mm. Les performances spectrales sont les meilleures à haute énergie pour les détecteurs germanium haute-pureté, HPGe (*High Purity Germanium*), avec une résolution typiquement de 0,8 % à 59,5 keV et de 0,2 % à 662 keV (ORTEC, 2020). Cependant, les diodes Germanium ne peuvent pas être utilisées dans de bonnes conditions à température ambiante et doivent être refroidies à une température de 77 K. En plus de cela, leurs performances sont affectées par les neutrons dans l'environnement spatial, induisant un phénomène de déplacement atomique, qui nécessite un recuit du détecteur pour

[*] Largeur totale à mi-hauteur : sur un spectre en énergie, on mesure la largeur à mi-hauteur des pics photoélectriques à certaines énergies pour caractériser la résolution du détecteur.

recouvrer les caractéristiques initiales. Ce sont des contraintes fortes de masse et de consommation pour des applications spatiales.

Le CdTe (Tellurure de Cadmium) est un semi-conducteur très utilisé pour la détection de photons X et gamma. Ses caractéristiques sont données en Tableau 1-1. Son numéro atomique de 48 et 52 respectivement pour le Cadmium et le Tellure ainsi que sa forte densité de 5,85 le rendent de fait plus efficace que le Germanium pour arrêter les photons de haute énergie, avec 99 % d'efficacité à 100 keV pour 5 mm. Malgré des performances spectrales inférieures à celles du Germanium et du Silicium, comme par exemple une résolution de 0,6 % FWHM à 662 keV pour le détecteur Caliste-HD, ce matériau présente l'intérêt d'être utilisable à des températures proches de l'ambiante, avec des performances optimales en dessous de -10°C. Ces avantages font du CdTe un très bon candidat pour des missions spatiales, permettant de concevoir des instruments répondant aux contraintes de masse et de consommation liées au spatial.

Caractéristique	Valeur
Numéros atomiques	48 et 52
Densité (à 300 K)	5,85
Énergie de bande interdite (à 300 K)	1,51 eV
Permittivité relative	10,3
Énergie de création de paires (à 300 K)	4,42 eV/paire
Résistivité (à 300 K)	$10^9\ \Omega.\mathrm{cm}$
Mobilité des électrons **(Meuris, 2009)**	$950\ \mathrm{cm}^2.\mathrm{V}^{-1}.\mathrm{s}^{-1}$
Produit temps de vie-mobilité des électrons	$3{,}3.10^{-3}\ \mathrm{cm}^2/\mathrm{V}$
Mobilité des trous **(Meuris, 2009)**	$73\ \mathrm{cm}^2.\mathrm{V}^{-1}.\mathrm{s}^{-1}$
Produit temps de vie-mobilité des trous	$2.10^{-4}\ \mathrm{cm}^2/\mathrm{V}$
Facteur de Fano **(Meuris, 2009)**	0,15
Concentration en porteurs de charge (à 300 K)	$6{,}9.10^5\ \mathrm{cm}^{-3}$

Tableau 1-1 : Caractéristiques du CdTe. Sauf mention contraire, les valeurs sont données par (Limousin, 2016)

Un dernier défi lié à l'instrumentation X et gamma concerne l'imagerie de sources de photons de haute énergie. Les optiques classiques, utilisées dans le domaine visible, à base de lentilles ou de miroirs ne peuvent pas être utilisées à haute énergie, les photons traversant majoritairement la matière sans interagir. De ce fait, d'autres optiques doivent être conçues. Par exemple, pour les rayons X, des miroirs à incidence rasante ont été envisagés pour des missions spatiales comme Simbol-X (Meuris, 2009) afin d'imager des sources jusqu'à 80 keV. Ils ont été mis en œuvre dans les missions NuSTAR (Caltech, 2005), avec un détecteur CdZnTe, et ASTRO-H Hitomi (NASA, 2020) avec des diodes Schottky CdTe. Une des difficultés de ce type de système est sa taille, puisque la distance focale par exemple prévue pour Simbol-X est de l'ordre de 20 m et celle de NuSTAR est de 10 m, la mise en œuvre étant assurée par un vol en formation de deux satellites, contenant l'optique pour l'un et le plan de détection pour l'autre.

D'autres optiques sont basées sur les réflexions des photons X comme les optiques en œil de langouste, utilisées par l'instrument MXT de la mission SVOM. Leurs performances dépendent très fortement de la fabrication précise de l'optique et de l'alignement de ses éléments (Gosset, 2019). Ces optiques sont intéressantes pour des photons ayant une énergie inférieure à 15 keV.

Pour des photons de plus haute énergie, allant jusqu'à 150 keV, il est nécessaire d'utiliser d'autres types d'imagerie. Les optiques basées sur les masques codés répondent à ce besoin en utilisant une

méthode d'imagerie indirecte, mais nécessitant un traitement particulier des données mesurées par le plan de détection. J'utilise et détaille ce type d'optique dans les Chapitres 4 et 5. Pour des photons d'énergie supérieure à 150 keV, les masques codés ne sont plus suffisants pour arrêter les photons. D'autres méthodes d'imagerie doivent alors être utilisées, comme par exemple les télescopes Compton, basés sur l'effet du même nom, méthode d'imagerie que j'utilise et que je détaille en Chapitre 6.

Dans tous les cas, ces méthodes d'imagerie nécessitent que le plan de détection ait lui-même des capacités d'imagerie, c'est-à-dire, qu'il soit capable de fournir la position d'interaction des photons détectés. Cette information en position peut être donnée par des détecteurs à bandes double-face (*double-sided strip detector*) (Ishikawa, et al., 2010), constitués de deux bandes d'électrodes orthogonales qui, par analyse des événements en coïncidence temporelle, renvoient les positions X et Y sur le détecteur, avec des bandes qui peuvent être aussi fines que 60 µm pour le CdTe (Furukawa, et al., 2019). Cela se fait cependant au détriment des capacités spectrales (Meuris, 2009). L'autre possibilité est l'utilisation de détecteurs pixellisés, que je mets en œuvre dans le cadre de mes travaux, où chaque pixel rend compte de la position d'interaction des photons.

1.3 Vers la conception de détecteurs CdTe miniatures pixellisés

En 2002, la mission spatiale INTEGRAL (ESA, 2019) a été lancée dans le but d'observer l'Univers violent dans le domaine des rayons X et gamma, à laquelle le Département d'Astrophysique du CEA a beaucoup contribué. Le satellite embarque en particulier l'instrument IBIS, télescope à masque codé présenté en Figure 1-2, qui contient deux plans de détection : ISGRI, détecteur CdTe, et PICsIT, détecteur à scintillation CsI (Iodure de Césium).

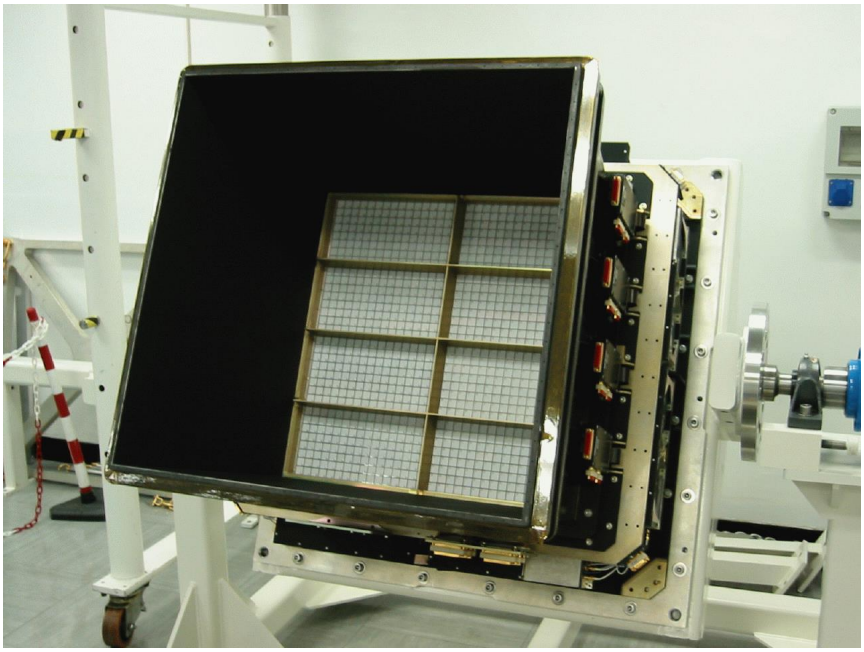


Figure 1-2 : L'instrument IBIS (ISDC, 2010). En blanc, le plan de détection ISGRI avec ses 128x128 pixels.

ISGRI est un détecteur semi-conducteur à base de CdTe, prévu pour mesurer des photons à une énergie allant jusqu'à 1 MeV, avec un seuil bas de 15 keV, caractéristique importante car l'efficacité est plus grande à basse énergie. Le plan de détection ISGRI est pixellisé : 128x128 pixels de 4x4 mm² pour une épaisseur de 2 mm, ce qui couvre une surface totale de détection de 2621 cm².

La mission INTEGRAL a permis de référencer près de 1000 sources, dont la plupart sont des candidats trous noirs ou étoiles à neutrons, et la mission acquiert encore des données à l'heure actuelle. À la suite de cette mission, au sein du Département d'Astrophysique du CEA, un programme de Recherche & Développement a été lancé en 2004 pour concevoir des détecteurs semi-conducteurs CdTe. Les développements se sont orientés vers la construction de plans de détections pour photons X-durs, de 5 à 100 keV, avec pour objectifs d'atteindre des performances spectrales et spatiales d'un ordre de grandeur meilleures que le détecteur ISGRI.

ISGRI est constitué de détecteurs ohmiques de grandes tailles, ce qui provoque un courant d'obscurité élevé et par conséquent un bruit électronique important. Cela réduit d'une part la résolution spectrale et augmente le seuil bas d'autre part. Les recherches se sont tournées alors vers des détecteurs de petite taille, pixellisés, à contact Schottky, réduisant drastiquement le courant d'obscurité et augmentant ainsi simultanément les capacités spatiales et spectrales du détecteur. Ce concept de détecteur a été nommé Caliste et décliné sur plusieurs versions ou générations sur ces quinze dernières années.

En 2007, la première version de Caliste, appelée Caliste-64 est née. Ce détecteur est composé d'un cristal CdTe pixellisé, de 64 pixels répartis sur une matrice carrée 8x8 au pas de 1 mm. Une des principales originalités qui démarque ce détecteur est l'intégration en 3D des électroniques de lecture à ultra bas bruit, les ASICs IDeF-X (Limousin, et al., 2005), capables de lire 16 canaux, placés perpendiculairement au cristal comme présenté en Figure 1-3. L'intégration électronique et le report des cristaux sur la matrice de pixels sont effectués par l'entreprise 3D PLUS.

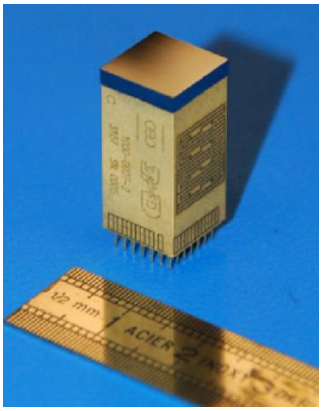


Figure 1-3 : Caliste-64. Sous le cristal CdTe, le corps du Caliste contient les ASICs disposés perpendiculairement au cristal (Limousin, 2016).

Entre 2008 et 2009, une nouvelle version de Caliste a été développée sous le nom de Caliste-256 avec une densité de pixels quatre fois plus importante que Caliste-64 : 16x16 pixels au pas de 580 µm, associés à une nouvelle version des ASICs IDeF-X capables de lire jusqu'à 32 canaux. Cette amélioration s'est faite au prix d'une augmentation de la consommation, passant de 200 mW pour Caliste-64 à 800 mW pour Caliste-256. La puissance consommée est un facteur important à maîtriser car cela dimensionne le système qui doit assurer le refroidissement du détecteur. Ce but a été atteint avec Caliste-HD, présenté en paragraphe 2.1, et l'électronique de lecture IDeF-X HD associée. C'est ce dernier détecteur que j'utiliserai principalement dans mes travaux.

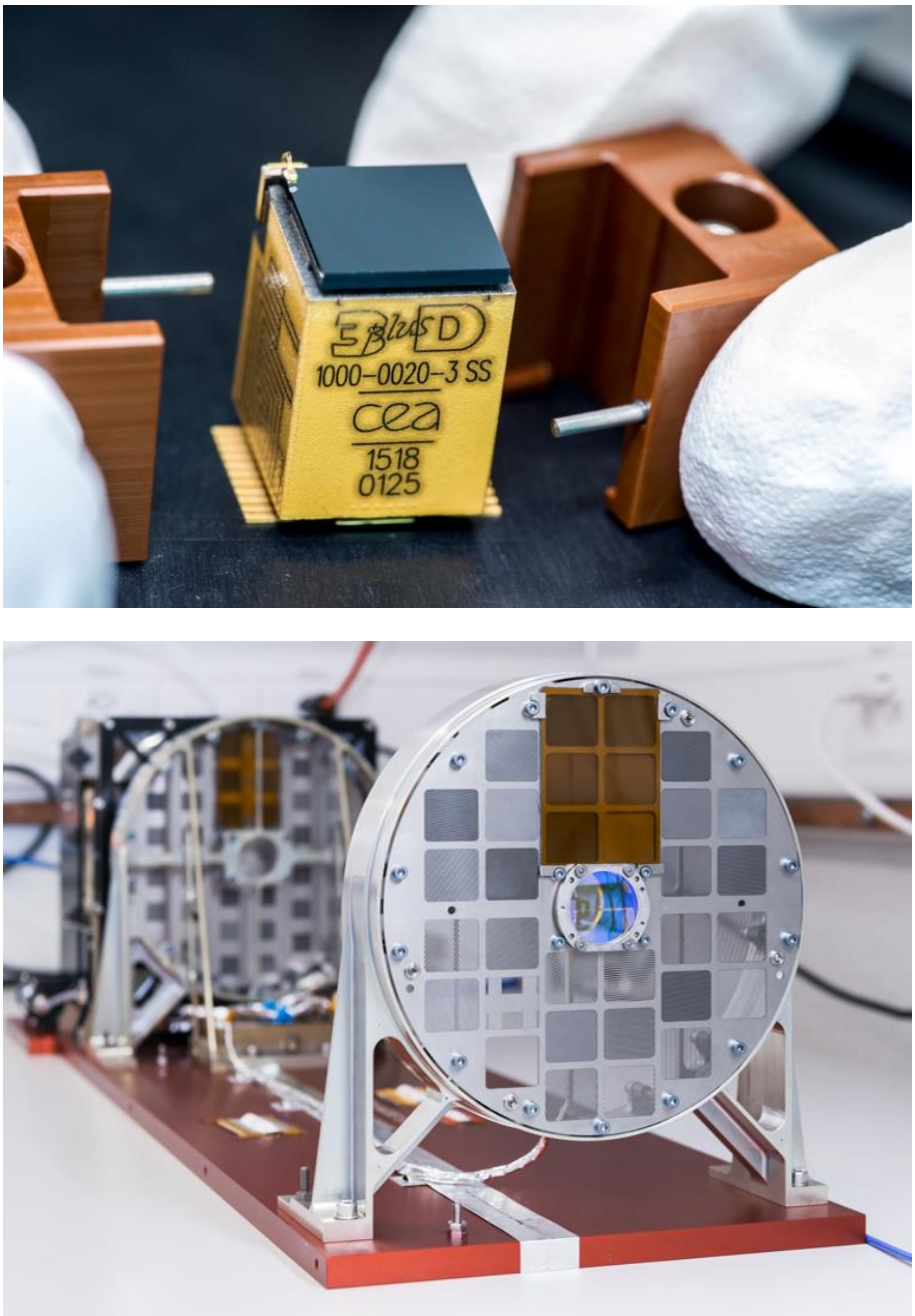


Figure 1-4 : En haut : Caliste-SO, détecteur utilisé par l'instrument STIX. En bas : Structure de l'instrument STIX, en face de chaque ouverture carrée est disposé un détecteur Caliste-SO.

Une version de Caliste a été dédiée à la mission Solar Orbiter (ESA, 2019), nommée Caliste-SO (Meuris, et al., 2014) et intégrée à l'instrument STIX pour l'observation et l'imagerie des éruptions solaires dans le domaine des rayons X. Caliste-SO et l'instrument STIX sont présentés en Figure 1-4. La matrice de pixels n'est pas régulière comme dans les précédentes versions, mais est composée de deux lignes de 4 grands pixels de 9,8 mm², entre lesquelles est située une ligne de 4 petits pixels de 0,9 mm², comme présenté en Figure 1-5. Cette forme est motivée par l'optique de STIX, optique à transformée de Fourier qui suit un *pattern* de Moiré. Les petits pixels sont utilisés dans le cas où le taux de comptage est trop important sur les grands pixels. Chaque Caliste-SO n'utilise qu'un seul ASIC IDeF-X HD. Les 12 pixels de la matrice ainsi que l'anneau de garde qui les entoure sont lus par l'électronique intégrée.

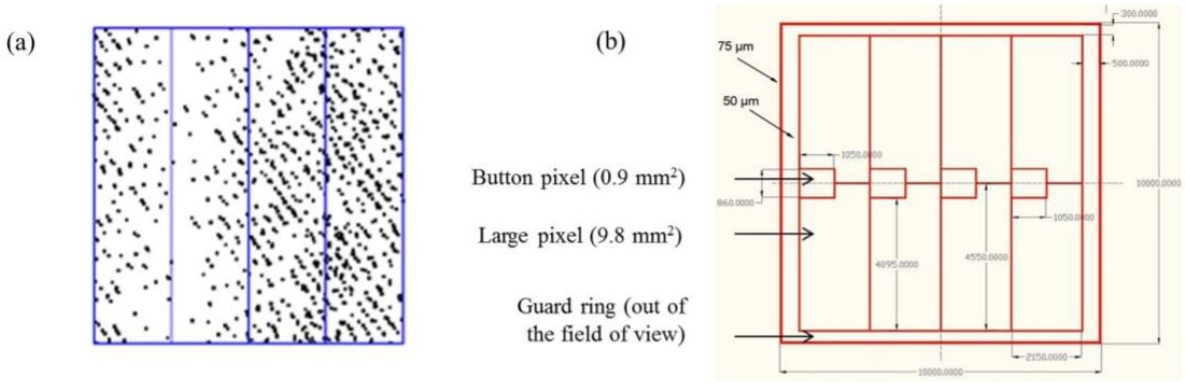


Figure 1-5 : (a) Simulation de photons à travers l'optique de l'instrument STIX (b) Structure des pixels de Caliste-SO (Meuris, et al., 2014).

La mission Solar Orbiter a été lancée avec succès le 9 février 2020, et les premières données des instruments à bord parviennent depuis fin mars 2020, dont on peut voir un spectre de calibration en vol en Figure 1-6.

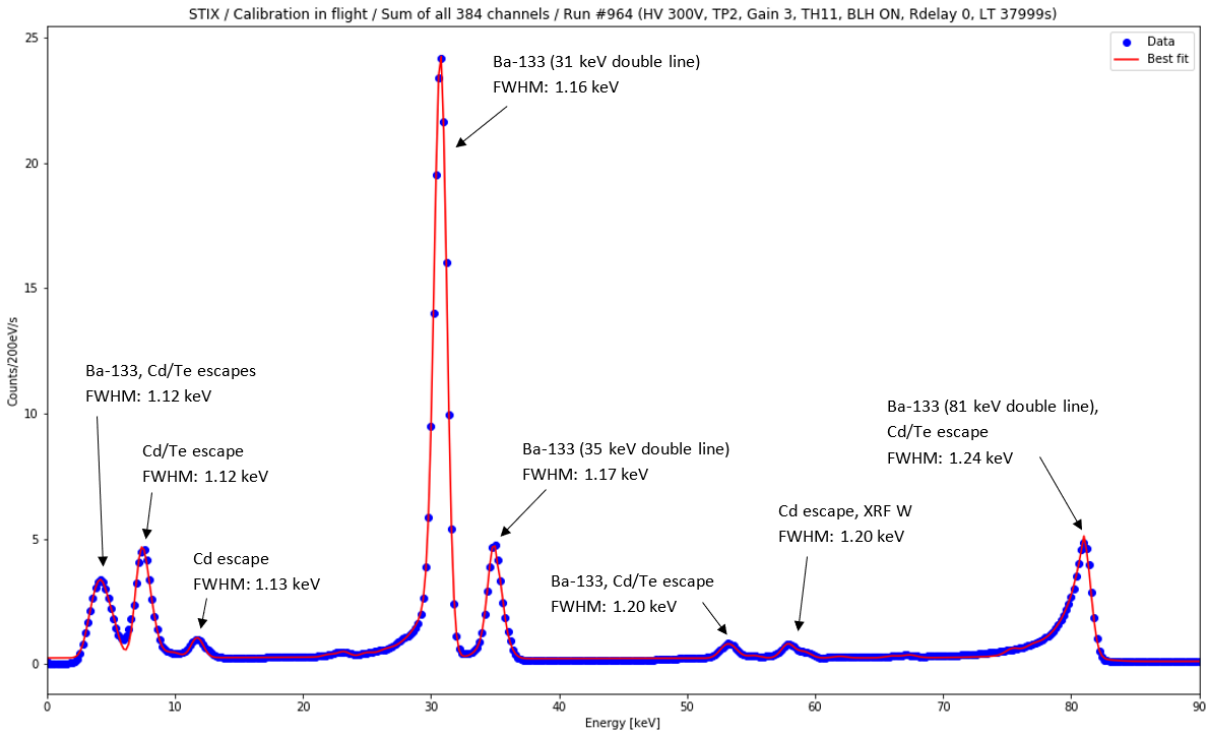


Figure 1-6 : Spectre de calibration en vol de l'ensemble des 32 Caliste-SO (384 canaux au total).

Le Tableau 1-2 résume les caractéristiques des trois versions de Caliste conçues pour des applications spatiales. Après avoir présenté l'historique des développements de Caliste avec les deux versions Caliste-64 et Caliste-256 ainsi que la version Caliste-SO, originale pour son développement dédié à l'instrument STIX, je présente en paragraphe 2.1 de manière plus détaillée la version que j'utilise dans le cadre de ma thèse Caliste-HD et sa version industrielle Caliste-O.

Paramètres	Caliste-64	Caliste-256	Caliste-SO
Années de développement	2005-2007	2008-2009	2011-2013
Matrice de pixels	8 x 8	16 x 16	4 x 3
Pas des pixels	900 µm	580 µm	2150/4550 µm
Taille de l'anneau de garde	900 µm	100 µm	500 µm
Électronique de lecture	IDeF-X V1.1 (16 canaux)	IDeF-X V2 (32 canaux)	IDeF-X HD (13 canaux utilisés)
Nombre d'ASICs	4	8	1
Interface	7 x 7 PGA	7 x 7 PGA	2 x 10 SOP
Consommation	200 mW	800 mW	20 mW
Dynamique en énergie	2 à 250 keV	1,5 à 250 keV	1,5 à MeV
Résolution (FWHM à 60 keV)	900 eV	860 eV	1000 eV
Dimensions	10x10x18,6 mm^3	10x10x18,6 mm^3	11x12x16,7 mm^3

Tableau 1-2 : Caractéristiques des Caliste-64, Caliste-256, Caliste-SO (Limousin, 2016).

2 Application des détecteurs Caliste aux caméras gamma

2.1 Caliste-HD et Caliste-O

Dans la lignée de Caliste-256, dont le principal inconvénient était sa consommation électrique de 800 mW, particulièrement contraignante pour le refroidissement efficace du détecteur CdTe, l'électronique de lecture IDeF-X a été déclinée dans une nouvelle version optimisée : IDeF-X HD (Gevin, et al., 2012). IDeF-X HD a une consommation réduite d'un facteur 4 par rapport à la version IDeF-X V2 utilisée dans Caliste-256.

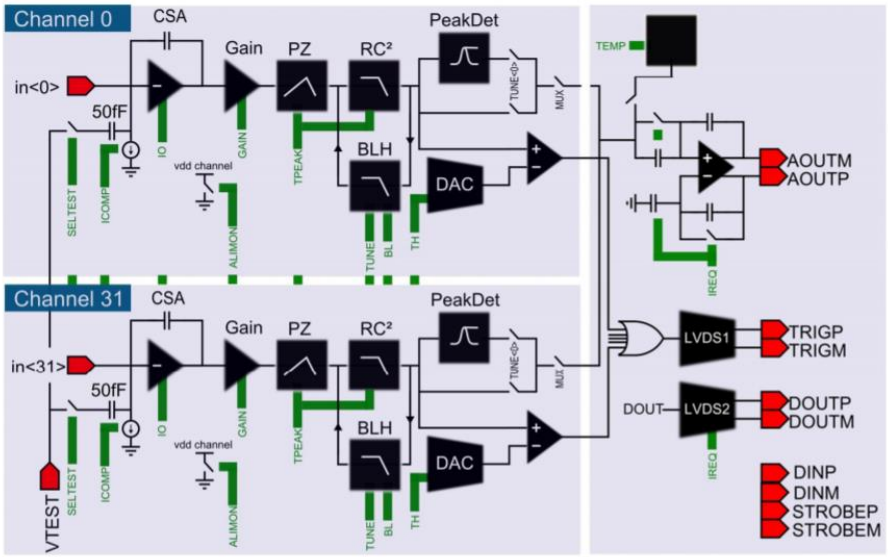


Figure 2-1 : Architecture de l'ASIC IDeF-X HD (Meuris, et al., 2014).

Par ailleurs, cette nouvelle version présente aussi l'avantage d'avoir un gain adaptable : 50 mV/fC, 100 mV/fC, 150 mV/fC ou 200 mV/fC. Cette fonctionnalité permet de changer la dynamique en énergie : de 2 keV à respectivement 1 MeV, 500 keV, 333 keV, 250 keV, environ sur les événements simples, c'est-à-dire les événements qui n'ont déclenché qu'un seul pixel. Ainsi, suivant l'application, on pourra préférer une plus petite dynamique en énergie, de façon à avoir une meilleure résolution, ou un gain plus élevé permettant d'avoir potentiellement une meilleure résolution en énergie. Dans

mon travail de thèse, j'ai par exemple principalement utilisé le gain de 50 mV/fC amenant à une dynamique de 1 MeV. Une dynamique importante est en effet plus avantageuse pour l'identification d'un plus grand nombre de radioéléments, en tirant profit de leurs signatures spectrales sur une plus grande gamme d'énergie.

La Figure 2-1 représente l'architecture de l'ASIC IDeF-X HD. Intégré dans le détecteur Caliste-HD que je présente ci-après, le bruit électronique associé est typiquement de l'ordre de 50 à 100 électrons rms (Gevin, et al., 2012).

De la même manière que pour la version précédente Caliste-256, Caliste-HD contient une matrice de 16x16 pixels, soit 256 pixels, lus par 8 ASICs IDeF-X HD disposés perpendiculairement au cristal. Chaque ASIC lit deux lignes de 16 pixels. Chaque pixel a un pas de 625 µm, soit 525 µm en taille de pixel et 100 µm de zone inter-pixel, pour une surface totale de détection de 1 cm². Dans sa version principale, le cristal utilisé a une épaisseur de 1 mm. Il est envisageable de reporter des cristaux ayant une épaisseur comprise entre 0.75 mm et 5 mm suivant les applications.

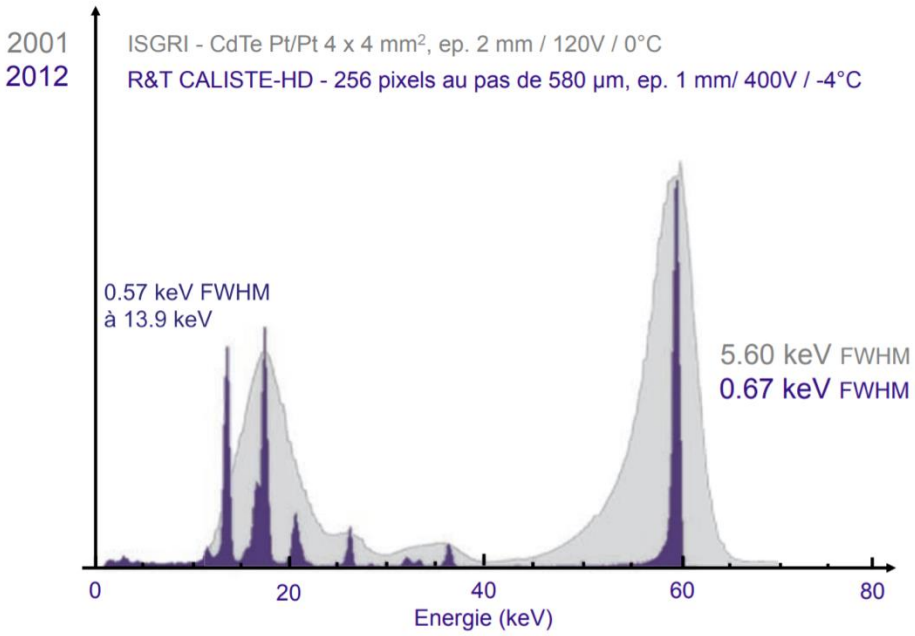


Figure 2-2 : Progrès en matière de spectrométrie X, d'INTEGRAL/ISGRI à Caliste-HD (Limousin, 2016).

Chaque interaction dans le cristal crée un nuage de charges dans le semi-conducteur, qui doivent être guidées vers les électrodes de façon à induire un signal à l'anode de lecture. Les charges sont guidées par une haute tension, amenée sur le cristal au niveau d'un plot en or installé sur le cristal, comme on peut le voir sur le dispositif MACSI présenté en Figure 2-3. La tension nominale utilisée est de 300 V/mm, des adaptations pouvant être faites suivant les conditions opérationnelles et les applications. Par exemple à très faible température, en-dessous de -30°C, une tension plus forte est applicable et permet d'obtenir une meilleure résolution spectrale. Les phénomènes à l'origine de l'induction du signal sur les électrodes sont étudiés plus en détail au Chapitre 2.

Avec le détecteur Caliste-HD, la résolution spectrale atteint 666 eV FWHM à 60 keV et 4,1 keV à 662 keV, ce qui place ce détecteur parmi les meilleurs détecteurs CdTe sur cet aspect, comme je le

montre en Tableau 2-1. Les progrès sur la résolution spectrale entre le détecteur ISGRI de l'instrument INTEGRAL et Caliste-HD sont présentés en Figure 2-2.

La compacité de ces détecteurs a permis de construire le spectro-imageur MACSI (*Modular Assembly of Caliste Spectro-Imagers*), en plaçant bout à bout deux lignes de quatre détecteurs Caliste-HD, pour aboutir à un total de 2048 pixels sur une surface de détection de 8 cm². Cela correspond à un huitième du détecteur ISGRI pour ce qui est du nombre de pixels, qui couvrent une surface de 328 cm² sur ce détecteur. Pour un même nombre de pixels, on aboutit ainsi à une réduction de surface d'un facteur 41, ce qui montre les performances atteignables en imagerie, en plus des performances spectrométriques.

La Figure 2-3 montre le système MACSI ainsi qu'une acquisition de carte de coups, avec une source de ^{241}Am éclairant un masque en tungstène de 1 mm d'épaisseur, en forme de spirale.

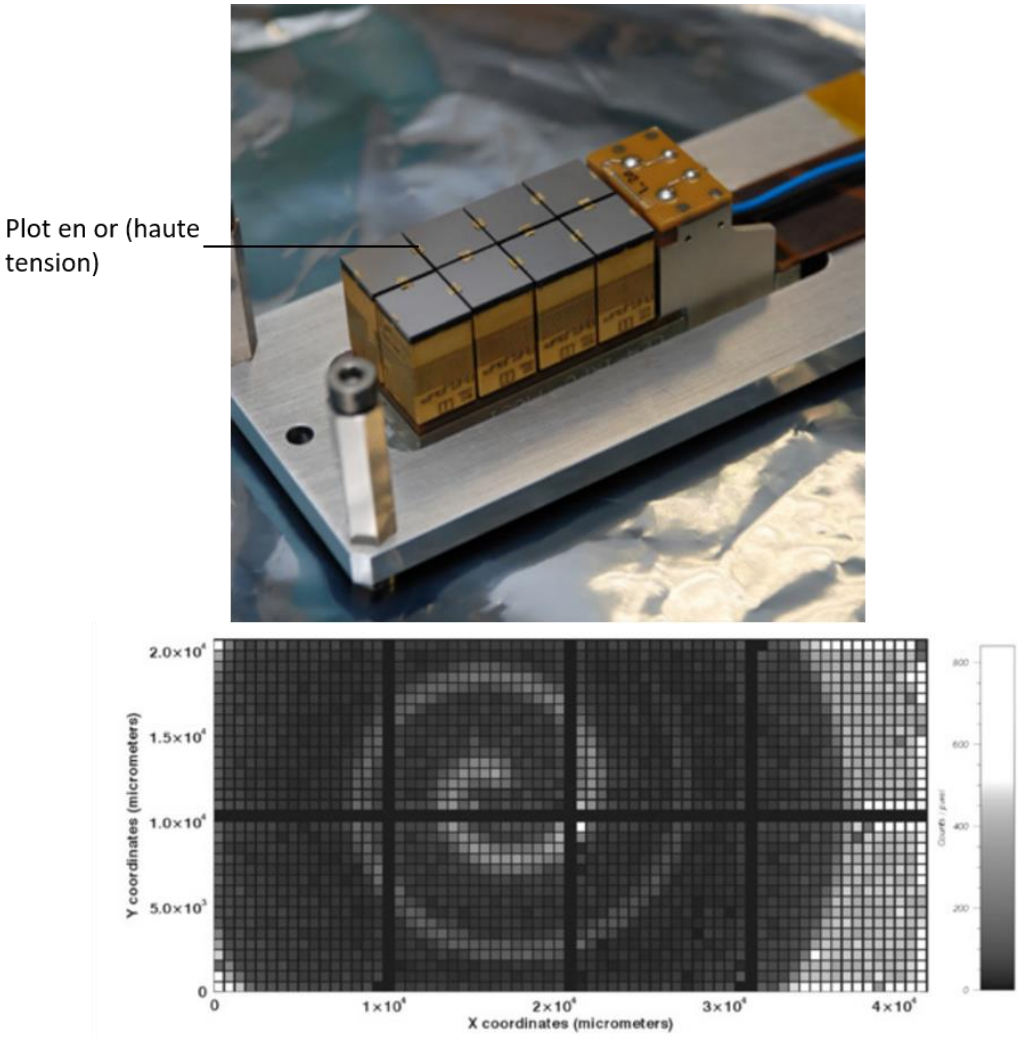


Figure 2-3 : En haut : Système MACSI. En bas : Carte de coups d'une acquisition de ^{241}Am éclairant un masque en tungstène de 1 mm d'épaisseur (Limousin, 2016).

Le détecteur Caliste-HD est à la base du prototype de caméra gamma WIX (nom donné au laboratoire sans signification particulière) que je décris en paragraphe 2.2.2. Ce prototype a été conçu dans le cadre du projet ANR (Agence Nationale de la Recherche) ORIGAMIX (Dubos, et al., 2015) qui avait entre

autres pour objectif le développement d'une caméra gamma hybride Compton/masque codé basé sur la technologie Caliste.

À l'issue de ce projet auquel j'ai contribué, une nouvelle version de détecteur a été développée à partir de 2017 compte tenu des besoins d'un détecteur plus efficace et donc plus volumineux. Le détecteur Caliste-O a été construit et évalué (Maier, et al., 2018)[*]. La matrice de pixels est toujours constituée de 16x16 pixels, associés aux électroniques de lecture IDeF-X HD. Cette fois, les pixels de 700 µm au pas de 800 µm couvrent une surface de de détection de 2 cm², soit le double de Caliste-HD. Les cristaux utilisés pour Caliste-O ont une épaisseur de 2 mm. Ces caractéristiques confèrent à Caliste-O un volume de détection quatre fois plus important que Caliste-HD.

Alors que Caliste-HD est produit de façon à répondre aux exigences de qualité pour des applications spatiales, Caliste-O est produit avec un standard de qualité de grade « industriel » et a vocation à être produit en plus grandes quantités. Le grade industriel ne signifie pas une baisse de performances mais il permet en revanche d'utiliser des procédés et technologies non qualifiés, souvent plus simples et tout aussi robustes.

Les performances spectrométriques de Caliste-O sont inférieures à celles de Caliste-HD en raison de la taille des pixels et de l'épaisseur du cristal principalement. La résolution est de 927 eV FWHM à 60 keV et de 6,7 keV à 662 keV, ce qui met Caliste-O au même niveau que les autres détecteurs CdTe existants du point de vue de la résolution, comme on peut le voir en Tableau 2-1 qui compare les détecteurs Caliste-HD et Caliste-O aux autres détecteurs CdTe existant pour des applications spatiales. Il reste meilleur sur la dynamique en énergie, allant même légèrement au-delà de 1 MeV, comme montré dans (Maier, et al., 2018), atteignant environ 1,2 MeV, la première raie à 1,17 MeV du ^{60}Co étant visible avec le détecteur en utilisant seulement les événements simples.

Détecteur	Caliste-HD	Caliste-O	Hexitec	JAXA group	Caltech	HXI/Astro-H
Pas des pixels/strips	625 µm	800 µm	**250 µm**	270 µm	498 µm	**250 µm (strips)**
Matrice de positions	16 x 16	16 x 16	80 x 80	12 x 12	24 x 44	**128 x 128**
Nombre de canaux	16 x 16	16 x 16	80 x 80	12 x 12	24 x 44	**128 x 2**
Épaisseur du cristal	1 mm	**2 mm**	1 mm	0,5 mm	0,5 mm	0,75 mm
Dynamique en énergie (événements simples)	2-1000 keV	**2-1200 keV**	4-200 keV	4-300 keV	5-100 keV	5-80 keV
FWHM à 60 keV	**666 eV**	927 eV	1 keV	900 eV	900 eV	1,5 keV
FWHM à 662 keV	**4,1 keV**	6,7 keV	N/A	N/A	N/A	N/A
Consommation par canal	800 µW	800 µW	-	**330 µW**	660 µW	500 µW
Consommation par unité de surface	2 mW/mm²	1 mW/mm²	-	1,9 mW/mm²	**0,16 mW/mm²**	0,42 mW/mm²
Éléments du module	CdTe, 8xASICs	CdTe, 8xASICs	CdTe, ASIC	CdTe, ASIC	CdTe, ASIC, ADC, processeur	CdTe, 8xASICs

Tableau 2-1 : Performances des détecteurs Caliste-HD (grade spatial), Caliste-O (grade industriel) et comparaison avec d'autres détecteurs CdTe pour le spatial.

[*] Publication à laquelle j'ai contribué pour le modèle dit de Comptonisation multiple détaillé en Chapitre 2.

2.2 Vers des applications nucléaires : le prototype de caméra gamma WIX

2.2.1 Des caméras gamma pour la sûreté nucléaire issues de l'astrophysique des hautes énergies

Le développement de systèmes d'imagerie de sources radioactives pour la sûreté nucléaire remonte principalement aux années 1990. L'objectif de ces systèmes est d'être capable de localiser des points chauds radioactifs dans une scène éventuellement inconnue à partir de la détection de photons X et gamma. Les premiers systèmes étaient principalement basés sur des scintillateurs, comme la caméra gamma CARTOGAM (Gal, et al., 2001) développée par le CEA LETI et EURISYS Mesures, qui repose sur des scintillateurs en Iodure de Césium dopé au Thallium, CsI:Tl, de 2 ou 4 mm d'épaisseur. L'imagerie est faite à l'aide d'un collimateur sténopé, ce qui nécessite un blindage particulier. De ce fait, l'ensemble de la tête de détection, bien que d'un diamètre assez réduit de 8 cm, a une masse de 15 kg.

À la fin des années 2000 puis plus spécifiquement après 2011, avec l'accident nucléaire de Fukushima, le développement de caméras gamma représente un intérêt particulier. Notamment, un des points clés étudiés a été la miniaturisation et la portabilité de ces systèmes, dans l'objectif d'intervenir rapidement dans des zones potentiellement difficiles d'accès. Une revue de ces systèmes a été effectuée par (Amoyal, 2019).

Le CEA List a développé la caméra gamma GAMPIX (Gmar, et al., 2011) à partir de 2007, système qui a été industrialisé et licencié par AREVA CANBERRA (actuellement Mirion Technologies) sous le nom d'iPIX (Mirion Technologies, 2019), montré en Figure 2-4. Cette caméra gamma utilise un détecteur CdTe Medipix (CERN, 2001), et plus particulièrement la puce de comptage Timepix. Le système d'imagerie est basé sur l'imagerie à masque codé, avec des masques de type MURA, d'épaisseur adaptable de 2 mm à 8 mm en fonction de l'énergie des photons utilisés pour la localisation. Des photons d'énergie élevée nécessitent une épaisseur importante pour être arrêtés. L'imagerie offre un champ de vue de 50° pour une résolution allant de 2,8° à 6,2° suivant le type de masque utilisé. En matière de sensibilité, ce système est capable de localiser une source de ^{241}Am à un débit de dose de 25 nSv/h au niveau de la tête de caméra en 30 secondes, et une source de ^{137}Cs à un débit de dose de 2 µSv/h en 30 secondes. Une déclinaison d'iPIX, nommée iPIX-NID, effectue des analyses spectrales et l'identification de radioéléments, basée sur des algorithmes de recherche de pics. Cette caméra gamma pèse 2,5 kg ce qui garantit sa portabilité.

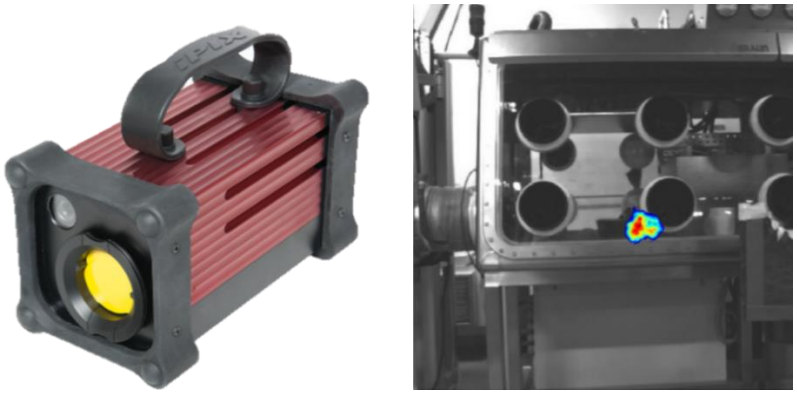


Figure 2-4 : Caméra iPIX (à gauche) et exemple de localisation de source (à droite) (Mirion Technologies, 2019).

En 2013, suite à l'accident nucléaire de Fukushima, une caméra gamma Compton (Takeda, et al., 2012) a été développée à partir des technologies utilisées pour le satellite ASTRO-H (ASTROnomie Haute énergie spatiale), et commercialisée sous le nom ASTROCAM 7000HS par Mitsubishi Heavy Industries. Cette caméra gamma utilise le principe de la diffusion Compton à travers plusieurs couches de détection. Les huit premières couches sont constituées de Silicium d'épaisseur 0,6 mm, matériau à faible numéro atomique Z=14, dans lequel la diffusion Compton est généralement plus probable que l'absorption photoélectrique. Les quatre couches suivantes sont constituées de CdTe d'épaisseur 0,75 mm, à plus haut numéro atomique (Z moyen de 50 environ) et utilisées pour arrêter les photons précédemment diffusés. Dans cette configuration, l'imagerie Compton offre un champ de vue de 180°, avec une résolution angulaire de 5,4° pour cette caméra gamma. Du point de vue de la sensibilité, cette caméra a besoin de 8 heures d'acquisition pour détecter une source de 1 MBq de ^{137}Cs à 3,2 m de la caméra, soit un débit de dose de 8 nSv/h. Avec cette configuration, la caméra gamma ASTROCAM 7000HS a une masse de 13 kg.

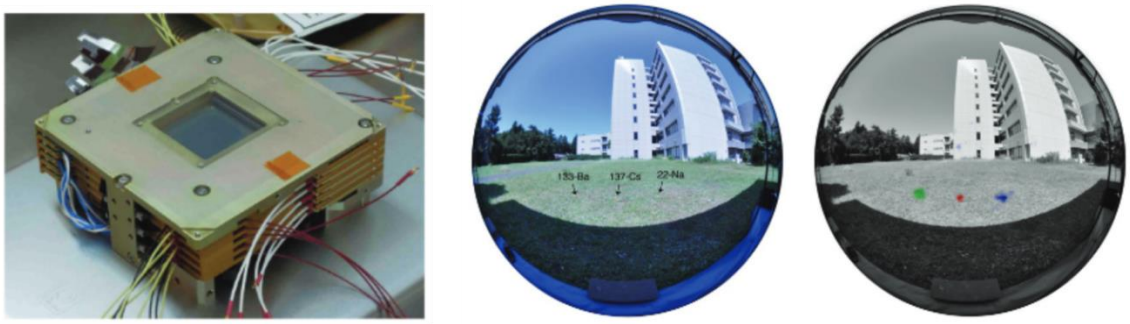


Figure 2-5 : À gauche : Détecteur à la base d'ASTROCAM 7000HS, on peut voir les différentes couches composant le détecteur. Au milieu et à droite : localisation de sources radioactives. (Takeda, et al., 2012).

Développée au milieu des années 2010 dans le cadre d'études menées par l'Université du Michigan et commercialisée par l'entreprise H3D, la caméra gamma Polaris (Wahl, et al., 2015) utilise un détecteur CZT. Son cristal épais possède un volume de plus de 19 cm^3 dans ses versions H400 et H420, avec une capacité de localisation de l'interaction en 3D dans le cristal, information intéressante pour utiliser l'imagerie Compton avec un unique détecteur*. L'utilisation d'un détecteur unique permet d'atteindre une masse de 4 kg. La résolution en énergie de ce détecteur est de 0,8 % FWHM à 662 keV. En ce qui concerne l'imagerie Compton, dans cette configuration et avec ces performances spectrométriques, le champ de vue qui peut être imagé est de 360° ou 4π stéradians, pour une résolution de 20° FWHM. Ce mode est utilisable lorsque l'énergie des photons incidents est suffisamment élevée, supérieure à 250 keV, pour produire des interactions Compton exploitables. Le volume de détection élevé de 19 cm^3 permet de localiser en mode Compton une source de ^{137}Cs produisant un débit de dose de 30 nSv/h au niveau de la caméra en 90 secondes. La version H420 de cette caméra est équipée d'un masque codé avec un champ de vue atteignant 86° pour une résolution de 5° FWHM. Du fait du pouvoir d'arrêt du masque codé, ce mode n'est utilisable que pour des photons d'énergie inférieure à 450 keV.

* ASTROCAM 7000HS utilise plusieurs couches de détection pour connaître la profondeur d'interaction, la couche d'interaction représentant la profondeur.

Figure 2-6 : À gauche : la caméra Polaris version H-420. À droite : Exemple de localisation avec la caméra Polaris. (H3D, 2020).

Plus récemment, la caméra gamma NuVISION (Montemont, et al., 2017) a été développée dans le cadre d'une collaboration entre le CEA Leti de Grenoble et l'entreprise NUVIA. Équipée d'un détecteur CZT de 4x4x0,6 cm^3 et des électroniques de lecture IDeF-X HD, la résolution spectrale de cette caméra atteint 1,5 % à 662 keV, pour une pixellisation de 16x16 pixels de 2,5 mm de côté. Cependant, un processus de sous-pixellisation est appliqué afin d'être capable de positionner plus précisément les interactions, et ultimement d'obtenir des images de 128x128 pixels. Cette caméra présente un fonctionnement hybride masque codé et imagerie Compton. Pour ce qui est du masque codé, le champ de vue est de 45° pour une résolution de 3,5° et pour l'imagerie Compton, le champ de vue est total, 360°, pour une résolution de 15°. Cette caméra est capable en 2 minutes de localiser une source de ^{137}Cs produisant un débit de dose de 500 nSv/h au niveau de la caméra. Le poids total de la caméra gamma est de 3 kg.

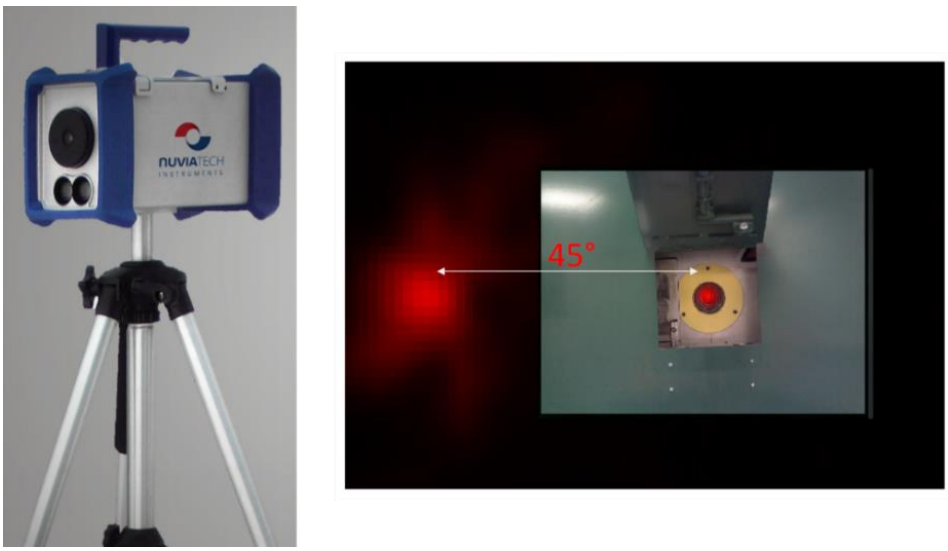


Figure 2-7 : À gauche : la caméra NuVISION (NUVIATECH Instruments, 2020). À droite : Exemple de reconstruction Compton de deux sources de ^{137}Cs produisant 10 µSv/h à gauche et de 18 µSv/h au niveau de la tête de détection (Montemont, et al., 2017).

2.2.2 WIX : un prototype de caméra gamma basé sur Caliste HD

Au vu de l'état de l'art des détecteurs utilisés pour les caméras gamma, les détecteurs Caliste ont des performances spectrométriques et en imagerie qui en font de bons candidats pour constituer les briques de base d'une caméra gamma.

J'ai étudié la faisabilité de l'utilisation des détecteurs Caliste dans le cadre de l'imagerie Compton, à travers des simulations et j'ai eu l'occasion de participer aux premières étapes de la conception des têtes de détection des caméras WIX, qui utilisent les détecteurs Caliste-HD.

L'intérieur de l'une de ces têtes de détection est montré en Figure 2-8. Le détecteur Caliste est maintenu sur un mât, enserré entre deux modules Peltier utilisés pour assurer le refroidissement du détecteur à une température de -10°C. Les sorties du détecteur Caliste sont reliées à un connecteur 21 points par un PCB flex, que j'ai conçu avec l'équipe du DEDIP (Département d'Électronique des Détecteurs et d'Informatique pour la Physique) du CEA. Les modules Peltier sont reliés à un connecteur 9 points pour leur alimentation, gérée en dehors de la tête de détection.

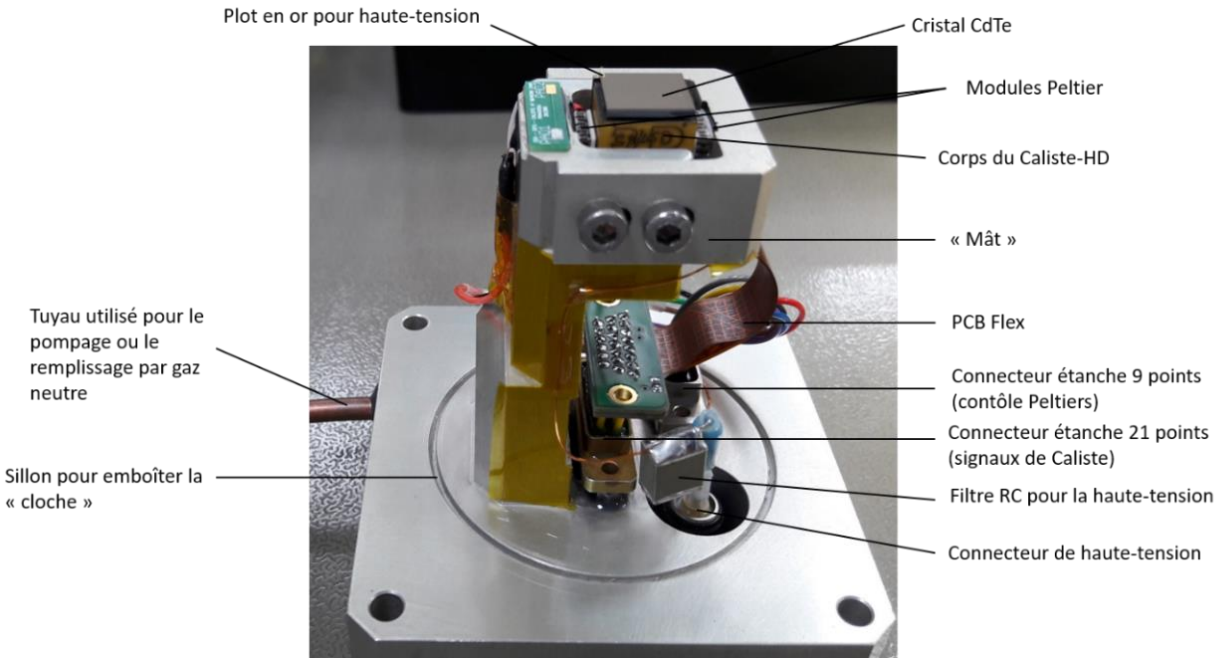


Figure 2-8 : Tête de détection WIX-HD.

La structure mécanique est composée d'Aluminium, qui présente l'intérêt d'avoir une densité assez basse (2,7) tout en étant assez résistant pour les applications prévues. La tête de détection est fermée par une « cloche » en Aluminium représentée en Figure 2-9, d'épaisseur moyenne de 1 mm environ. Une fenêtre plus fine est située en face du cristal Caliste, de 500 µm pour laisser passer les photons avec une énergie plus faible, jusqu'à 10 keV[*].

La tête de détection est fermée et l'intérieur est mis sous atmosphère neutre et sèche, avec un gaz inerte comme de l'Argon ou de l'Azote ou bien sous vide. Cela évite que le refroidissement des Peltier provoque de la condensation, qui serait alors dangereuse pour l'ensemble de l'électronique du Caliste. Ainsi, les éléments en interaction avec l'extérieur de la tête de détection, en particulier les

[*] Les photons de 10 keV traversent la fenêtre avec une probabilité de 10 %.

connecteurs, doivent assurer l'étanchéité. J'ai effectué le collage des connecteurs et leur *potting* avec de la résine EPOXY.

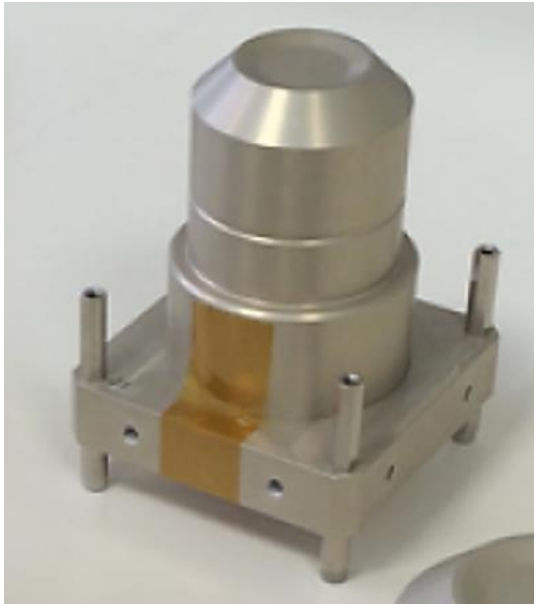


Figure 2-9 : Tête de détection fermée avec cloche.

Le connecteur 21 points est relié, en dehors de la tête de détection, à une carte d'interface « CIF ». Cette carte fait le lien entre le Caliste et une carte d'acquisition MicroZed, alimentée par un lien USB-C. La MicroZed assure le transfert des commandes de modification de configuration des ASICs IDeF-X HD d'une part et celui des signaux analogiques issus des acquisitions de Caliste vers l'ordinateur d'acquisition d'autre part. Cette carte intègre un *system-on-chip* Zynq pour le contrôle des ASICs, pour la configuration et la lecture par le FPGA (*Field Programmable Gate Arrays*), l'encodage des données par l'ADC (*Analog-to-Digital Converter*), la mise en forme et le transfert sur le lien Ethernet par le processeur ARM. Les données sont transmises par paquets à l'ordinateur d'acquisition, par liaison sans fil. Je présente plus précisément la nature des données transmises au début du Chapitre 2.

L'électronique d'acquisition est représentée en Figure 2-10. Par ailleurs, des modifications sont prévues au niveau de cette carte MicroZed, pour la remplacer par une carte CAMEHA, dans une collaboration entre le CEA et l'entreprise 3D PLUS. Cette carte aura pour objectif de contrôler en plus l'alimentation des modules Peltier, qui est actuellement assurée de manière externe.

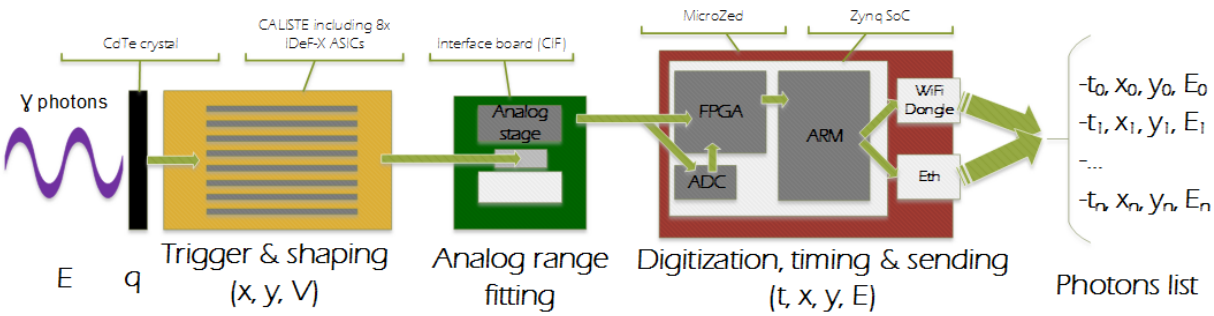


Figure 2-10 : Architecture électronique WIX-HD (Daniel, et al., 2018).

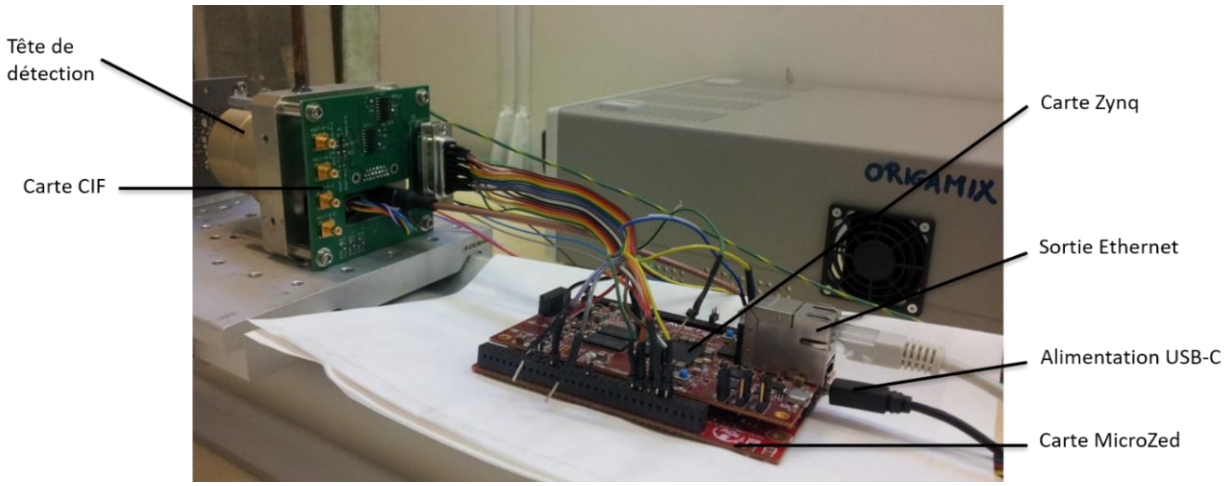


Figure 2-11 : Photo de l'électronique de la WIX.

Ces éléments sont intégrés dans un système compact qui constitue la caméra WIX. L'ensemble de ce système a des dimensions de 7,5x7,5x21 cm^3 pour une masse de 950 g environ, garantissant sa portabilité. La communication avec la caméra est effectuée sans fil, par un module Wi-Fi. L'alimentation des modules Peltier et de la carte MicroZed peut être opérée sur secteur ou sur batterie. La caméra est placée sur un pied de hauteur ajustable, d'une masse de près de 1 kg.

Sur la tête de détection, un support de masque codé peut être placé ou retiré suivant les applications, fabriqué par impression 3D. Dans ce support, il est possible de placer un masque codé, avec différentes épaisseurs suivant les énergies des photons à détecter, comme je le présente en Chapitre 4. L'ensemble du système, présenté en Figure 2-12, possède une masse totale de 2 kg.

Figure 2-12 : Prototype de caméra WIX sur son pied, équipée d'un masque codé aléatoire.

Le contrôle de la caméra et l'acquisition des données se fait à l'aide d'un ordinateur équipé du logiciel d'acquisition *DiagnostiX*, développé au Département d'Astrophysique du CEA spécifiquement pour cette application. Une première version de ce logiciel a été développée en 2016 avec des fonctionnalités de base. Elle permet de contrôler les différents registres des ASICs, comme les seuils bas de chaque pixel, le gain utilisé, le temps de *peaking* ainsi que le contrôle de la haute tension.

L'interface logiciel affiche en temps réel la carte des coups de l'acquisition. J'ai participé aux évolutions successives de *DiagnostiX*, notamment pour en définir les fonctions de représentation des données avancées :

- la tenue à des taux de comptage supérieurs à 10 000 coups par seconde ;
- la possibilité de charger et sauvegarder des configurations au niveau des ASICs et de la carte MicroZed, permettant de lancer rapidement des configurations préférées ;
- la possibilité de lancer des *batchs* de mesure à partir des configurations précédemment définies, ce qui permet notamment de tester plusieurs configurations (haute tension et temps de *peaking* différents) pour caractériser les détecteurs, ainsi que de faire des mesures longues à travers des cycles* ;
- l'affichage des spectres étalonnés en énergie avec une table de calibration que j'établis avec la méthode décrite en Chapitre 2 ;
- d'autres fonctionnalités pour une meilleure ergonomie ont été ajoutées.

Les algorithmes que j'ai développés et que je présente dans les chapitres suivants, trouveront leur place naturellement dans les versions futures du logiciel de *DiagnostiX*.

La Figure 2-13 montre une capture d'écran lors d'une acquisition de DiagnostiX.

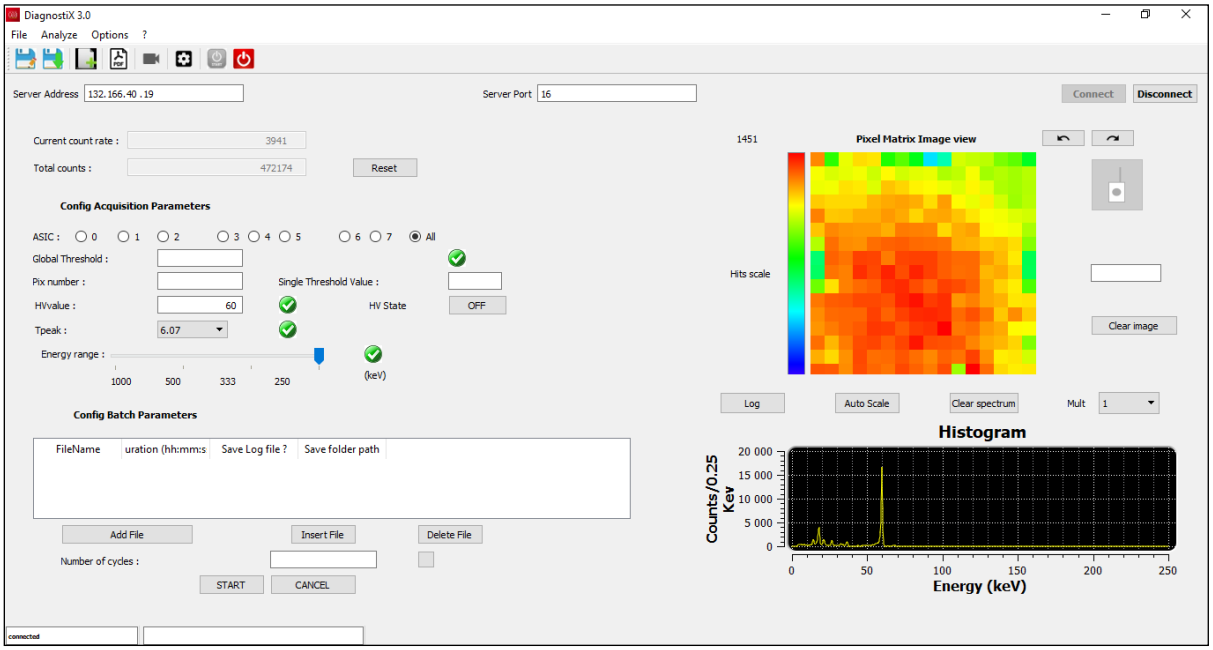


Figure 2-13 : Interface d'acquisition du logiciel DiagnostiX pour une source de 440 kBq de ^{241}Am placée devant la caméra, avec une haute-tension de 300 V, un temps de *peaking* de 6,07 µs et à gain de 200 mV/fC. La gauche de l'écran représente le panneau de contrôle des paramètres. En haut à droite se trouve la carte des coups en temps réel. En bas à droite se trouve le spectre calibré des événements simples, acquis en temps réel.

* Pour une mesure longue, il est en effet nécessaire de cycler régulièrement la haute tension, pour éviter l'effet de polarisation qui survient lorsque la haute tension est apportée au cristal pendant une longue durée. Cet effet a été décrit et caractérisé plus précisément par (Meuris, 2009)

Dans la suite, la WIX que j'utilise principalement est basée sur Caliste-HD, de ce fait nous précisons WIX-HD. Une adaptation de la structure a été développée pour utiliser des Caliste-O, on parle alors de WIX-O.

2.2.3 Performances de la caméra WIX-HD et comparaison avec les autres caméras gamma

La caméra WIX-HD reprend les performances spectrométriques de Caliste-HD : une résolution de 666 eV FWHM à 60 keV, soit 1 % et 4,7 keV FWHM à 662 keV, soit 0,7 %. La dynamique en énergie s'étend de 10 keV, du fait de la fenêtre en Aluminium, à 1 MeV avec les événements simples. En utilisant les événements multiples, c'est-à-dire des interactions dans le détecteur qui se sont étendues sur plusieurs pixels, la dynamique en énergie peut atteindre 2 MeV.

J'étudie en Chapitres 4 et 5 les capacités d'imagerie de la WIX-HD équipée du masque codé présenté en Figure 2-12. Dans ce chapitre, je montre que la précision atteignable est de 0,7° pour un champ de vue de 51.5° avec la configuration actuelle du masque, situé à 5,8 cm du détecteur. Une modification de cette configuration est envisageable.

Le masque utilisé est à épaisseur ajustable entre 0,3 mm et 3 mm. Je montre en Chapitre 4 l'impact de l'épaisseur sur l'efficacité du masque à arrêter les photons, mais inversement, je montre aussi qu'un masque plus épais apporte un effet de *vignetting* plus important qui diminue le champ de vue. Par exemple, un masque de 1 mm ne transmet que 50 % des photons à 250 keV sans diminuer le champ de vue de 51,5°, alors qu'un masque de 3 mm a une efficacité de 50% jusqu'à 450 keV mais en limitant le champ de vue à 37°. En matière de sensibilité, je montre que ce système est capable de localiser une source de ^{241}Am de 440 kBq située à 1 m de la caméra, soit un débit de dose de 9 nSv/h, en 50 secondes.

En Chapitre 6, j'étudie les capacités de reconstruction de la WIX-HD pour l'imagerie Compton. Les très bonnes capacités spectroscopiques de Caliste-HD sont intéressantes car la mesure de l'énergie est une étape clé pour exécuter une reconstruction en imagerie Compton. Cependant, je montre que la méconnaissance de la profondeur d'interaction est la principale source d'incertitude qui entache la reconstruction. Du fait de l'utilisation d'un seul détecteur plan, qui joue le rôle de plan de symétrie dans la reconstruction, le champ de vue atteignable est de 180°. Par ailleurs, en traitant les données avec les algorithmes adaptés que je présente en Chapitre 6, je montre que j'atteins une résolution de 8° FWHM, et que je suis capable de localiser indépendamment deux sources séparées de 10°[*]. La sensibilité reste un problème à surmonter au vu de la géométrie utilisée, consistant en un unique détecteur plan. Par exemple, il est nécessaire d'effectuer une acquisition de 7 h typiquement pour pouvoir localiser une source de ^{137}Cs générant un débit de dose de 1,4 µSv/h au niveau du détecteur.

Le Tableau 2-2 présente les caractéristiques et les performances des caméras gamma décrites en paragraphe 2.2.1 et les compare à celles de la caméra WIX-HD, décrite en paragraphes 2.2.2 et 2.2.3. J'ai établi les performances données pour la WIX-HD à travers les méthodes d'analyse des données que j'ai mises en œuvre et que je décris en Chapitres 4, 5 et 6.

[*] Le pouvoir séparateur est en effet très important en imagerie Compton. On peut avoir de très bonnes performances avec une seule source ponctuelle, mais, lorsque plusieurs sources sont présentes, chaque source représente un bruit pour les autres du fait de la méthode de reconstruction Compton. Par conséquent, la présence de plusieurs sources peut diminuer les performances de reconstruction.

Paramètre	iPIX	Astrocam 7000HS	Polaris-H	Nuvision	WIX-HD
Dimensions (cm^3)	19x9x9	45x34x24	24x10x18	10x12x23	**8x8x21**
Masse (kg)	2,5	13	4	3	**1**
Matériau détecteur	CdTe	Si et CdTe	CZT	CZT	CdTe
Dynamique énergie	10 keV à 1,5 MeV	10 keV à quelques MeV	50 keV à 3 MeV	20 keV à 1400 keV	10 keV à 1 MeV (2 MeV*)
Résolution en énergie (FWHM)	9,5 % à 60 keV 4,1 % à 662 keV†	2,2 % à 662 keV	0,8 % à 662 keV	2,5 % à 122 keV 1,5 % à 662 keV	1 % à 60 keV **0,7 % à 662 keV**
Mode d'imagerie	Masque codé	Compton	**Compton (> 250 keV) Masque codé (< 450 keV)**	**Masque codé Compton**	**Masque codé Compton**
Champ de vue Masque codé	50°	N/A	**86°**	45°	52°
Résolution angulaire Masque codé	2,8° à 6,2°	N/A	5°	3,5°	**0,7°**
Champ de vue Compton	N/A	180°	**360°**	**360°**	180°
Résolution angulaire Compton	N/A	**5,4°**	20°	15°	8° Pouvoir séparateur : 10°
Sensibilité	30 s pour ^{241}Am à 25 nSv/h 30 s pour ^{137}Cs à 2 µSv/h	8 h pour ^{137}Cs à 7,5 nSv/h	**90 s pour ^{137}Cs à 30 nSv/h (mode Compton)**	120 s pour ^{137}Cs à 500 nSv/h (mode non précisé)	**50 s pour ^{241}Am à 9 nSv/h (Masque codé)** 7 h pour ^{137}Cs à 1,4 µSv/h (Compton)

Tableau 2-2 : Caractéristiques et performances des caméras gamma existantes et comparaison avec la WIX-HD.

3 Objectifs de thèse

3.1 Défis à relever dans cette thèse

À travers la revue des systèmes existants, il existe toujours plusieurs défis scientifiques à relever concernant la caractérisation d'une scène radiologique, notamment à travers l'analyse des données acquises au niveau de ces scènes.

* Avec les événements multiples
† Données Timepix3 CdTe

3.1.1 La spectro-identification automatique des radioéléments

Dans un premier temps, les systèmes proposés sont principalement des systèmes d'imagerie et leurs capacités spectroscopiques sont exploitées principalement pour la localisation des sources en imagerie Compton. Certaines caméras, comme iPIX ou Polaris-HD, proposent un système d'identification des sources, basé sur la détection de pics. Cependant, les algorithmes de détection de pics nécessitent d'avoir détecté un nombre suffisamment élevé de photons. De ce fait, la spectro-identification automatique des sources radioactives et en temps réel, de surcroit avec un nombre réduit de photons dans le spectre, est un défi scientifique ouvert de grande importance.

Bien que les détecteurs aient de très bonnes performances spectrales du point de vue de la résolution, ils ne sont pas à l'abri de potentielles dérives d'étalonnage dues aux conditions opérationnelles telles qu'une moins bonne maîtrise de la température du détecteur ou de la haute tension, ce qui peut impacter les performances d'identification par le truchement de fluctuations des paramètres d'étalonnage. En outre, une information importante pour caractériser la scène radiologique est d'être capable, parmi les sources identifiées, de déterminer leurs proportions dans le signal, de façon à savoir si certaines sources sont largement dominantes dans la scène ou non.

Le premier axe de mon travail consiste à étudier et développer des méthodes de spectro-identification originales afin de répondre à ce défi scientifique, et de les appliquer au cas des détecteurs Caliste en les testant avec des données réelles issues de mesures.

À partir d'un spectre obtenu avec le détecteur Caliste, je réponds aux trois questions suivantes :

- *Quels radioéléments ont été à l'origine de ce spectre ?*
- *Dans quelles proportions chaque radioélément identifié a-t-il contribué au signal enregistré ?*
- *Comment évaluer les incertitudes liées aux résultats répondant aux deux questions précédentes ?*

3.1.2 L'imagerie à masque codé

Les concepts de systèmes d'imagerie à masque codé remontent aux années 1960 et les algorithmes de traitement associés sont très efficaces pour la localisation de sources ponctuelles, à l'infini, ce qui explique d'ailleurs leur utilisation pour des applications en astronomie. De ce fait, ces systèmes d'imagerie, associés à des détecteurs adaptés, ont des performances satisfaisantes pour la caractérisation de scènes radiologiques, à la fois en résolution spatiale et en sensibilité, comme montré en Tableau 2-2.

Néanmoins, lorsqu'il existe une multitude de sources ponctuelles dans le champ de vue ou dans le cas de la présence d'une source étendue, les algorithmes existants présentent des limitations, et peuvent alors être incapables de fournir des informations de localisation, comme je le présente en Chapitre 4. Pour autant, ces cas peuvent être rencontrés dans des scènes radiologiques.

Le deuxième axe de mon travail consiste à étudier et appliquer les méthodes usuelles de traitement de données d'imagerie à masque codé à la caméra gamma WIX-HD et à développer de nouvelles méthodes pour répondre au défi lié à la présence de plusieurs sources ponctuelles et de sources étendues.

Je formule la question scientifique à traiter ainsi :

À partir d'une carte des coups enregistrés par le détecteur Caliste devant lequel a été placé un masque codé, comment peut-on retrouver la position de la ou des sources à l'origine de cette acquisition ?

3.1.3 La reconstruction Compton

L'imagerie Compton avec un seul détecteur Caliste-HD est envisageable. Cependant, les incertitudes liées à la méconnaissance de la profondeur d'interaction des photons dans le cristal de CdTe, l'utilisation d'une unique couche de détection et le bruit intrinsèque lié aux méthodes classiques de reconstruction Compton apportent des défis à relever en matière de traitement des données. J'ai développé plusieurs algorithmes pour prendre en compte cette difficulté liée au bruit intrinsèque de la reconstruction Compton, qui peut dégrader la capacité du système à séparer deux sources proches, malgré une bonne résolution spatiale dans le cas de la reconstruction de l'image d'une source seule.

Par ailleurs, un autre point essentiel est la sensibilité de la caméra gamma, notre configuration étant peu favorable avec un cristal peu épais de 1 mm contrairement à la caméra Polaris-H qui possède un cristal épais ou la caméra ASTROCAM qui possède plusieurs couches de détection. Il est nécessaire de développer des algorithmes capables de localiser des sources avec un nombre réduit de détections Compton.

Le troisième axe de mes recherches consiste à mettre en œuvre des algorithmes avancés de traitement Compton appliqués à la caméra gamma WIX-HD, afin d'être capable d'une part de séparer des sources ponctuelles proches, en particulier de même énergie, ainsi que de traiter le cas de sources étendues, et d'autre part, de développer des algorithmes nouveaux pour aboutir à une meilleure sensibilité.

Je traite plus précisément la question scientifique suivante :

À partir d'une acquisition effectuée avec Caliste, comment extraire les événements pertinents pour la reconstruction Compton et retrouver la position de la ou des sources à l'origine de cette acquisition ?

3.2 Structuration du manuscrit

Dans ce manuscrit, je présente la méthodologie que j'ai mise en place pour répondre à ces trois axes de recherche. Je montre les résultats que j'ai obtenus en appliquant mes méthodes à des données préférentiellement issues de mesures réelles acquises avec la WIX-HD et le détecteur Caliste-HD, et par simulation dans certains cas où les mesures n'étaient pas réalisables.

Le Chapitre 2 présente précisément la nature des données issues des détecteurs Caliste. Je montre qu'un soin particulier doit être apporté à l'étalonnage en énergie et je propose un algorithme qui établit les paramètres de l'étalonnage. Dans un deuxième temps, je montre comment je peux expliquer les données qui sont observées avec Caliste, notamment en termes spectrométriques. Pour ce faire, je mets en place des simulations de la trajectoire des photons entre une source et le détecteur avec les éventuels dépôts d'énergie ayant eu lieu sur cette trajectoire, puis je modélise finement la réponse spectrométrique du détecteur. En plus de la compréhension des phénomènes à l'origine des spectres acquis avec Caliste, cette étape me permet de disposer d'un algorithme de simulation de spectres que j'utiliserai en Chapitre 3.

Le Chapitre 3 traite du premier axe de la thèse, sur la spectro-identification de sources radioactives. Pour relever les défis présentés en paragraphe 3.1.1, j'étudie l'application d'algorithmes de *deep learning* fondés sur l'utilisation de réseaux de neurones convolutifs. Je montre que je peux entraîner

ces réseaux de neurones à l'aide des simulations établies au Chapitre 2 et les appliquer à des données réelles. Par ailleurs, j'utilise cette méthode pour répondre à la fois à la question de l'identification des sources et à la question de la quantification de leur contribution au signal, tout en donnant une incertitude sur les deux résultats. À travers des tests issus de données réelles, j'évalue les performances de ces algorithmes. Enfin, je montre une application pour la détermination du débit de dose attribué à chaque source identifiée.

Les Chapitres 4 et 5 traitent du deuxième axe de recherche et concernent l'imagerie à masque codé pour la caméra gamma WIX. En Chapitre 4, je présente plus précisément le système de masque codé que j'utilise. Puis, je décris les algorithmes classiquement utilisés pour traiter les données d'imagerie à masque codé et je les applique au cas de la WIX-HD, dans le but d'établir les performances que j'ai données dans le Tableau 2-2, en montrant aussi les limites de ces algorithmes. Afin de dépasser ces limites, en particulier vis-à-vis de sources étendues, je mets en place au Chapitre 5 des algorithmes de *deep learning*, dont j'évalue les performances sur la problématique à la fois des sources ponctuelles et des sources étendues.

Le Chapitre 6 aborde le troisième axe de la thèse sur la reconstruction Compton avec Caliste-HD. Je présente d'abord le principe de l'imagerie Compton et la manière dont j'extrais les événements d'intérêt dans une acquisition. Puis, j'étudie quatre algorithmes, deux algorithmes existants et deux algorithmes que j'ai développés, qui peuvent être appliqués pour faire cette reconstruction Compton en montrant leurs avantages mais aussi leurs limitations. J'analyse notamment les performances du point de vue de la résolution spatiale, de la séparation de sources et de la sensibilité. Je montre enfin les possibilités d'application d'algorithmes de *deep learning* et leurs apports sur cette problématique.

4 Bibliographie

Abbott, B. & al., 2017. *GW170817: Observation of Gravitational Waves from a Binary Neutron Star Inspiral.* in: Phys. Rev. Lett. 119, 161101.

Amoyal, G., 2019. *Développement d'un imageur gamma hybride pour les applications de l'industrie.* in: Normandie Université.

Astrosurf, 2016. *Nébuleuses et amas stellaires : les rémanents de supernovae (III).* [En ligne]
Available at: http://www.astrosurf.com/luxorion/univers-nebuleuses-amas3.htm
[Accès le 2020].

Caltech, 2005. *NuSTAR Bringing the High Energy Universe into Focus.* [En ligne]
Available at: https://www.nustar.caltech.edu/

Ceraudo, F., 2019. *Characterization and performance optimization of the focal plane of the Micro-channel X-ray Telescope on-board the space astronomy mission SVOM.* in: Instrumentation and Methods for Astrophysic [astro-ph.IM], Université Paris-Saclay.

CERN, 2001. *The Medipix Collaboration.* [En ligne]
Available at: http://medipix.web.cern.ch/

Daniel, G. et al., 2018. *ORIGAMIX - Rapport final du Workpackage 3,* by: CEA - 3DPLUS.

Dubos, S., 2015. *Nouveau spectro-imageur CdTe à très haute résolution spatiale et spectrale pour l'astronomie X et gamma.* at: Université Paris 7 - Denis Diderot ; Commissariat à l'Energie Atomique et aux Energies Alternatives (CEA); Centre National d'Etudes Spatiales (CNES)..

Dubos, S. et al., 2015. *ORIGAMIX, a CdTe-based spectro-imager development for nuclear applications.* in: NIM-A, Volume 787, Pages 302-307.

ESA, 2019. *Integral.* [En ligne]
Available at: https://sci.esa.int/web/integral

ESA, 2019. *Solar Orbiter.* [En ligne]
Available at: https://www.esa.int/Science_Exploration/Space_Science/Solar_Orbiter

Furukawa, K. et al., 2019. *Development of 60 µm pitch CdTe double-sided strip detectors for the FOXSI-3 sounding rocket experiment.* in: NIM-A Volume 924, Pages 321-326.

Gal, O. et al., 2001. *CARTOGAM – a portable gamma camera for remote localisation of radioactive sources in nuclear facilities.* in: NIM-A, Volume 460, Issue 1, Pages 138-145.

Gevin, O. et al., 2012. *Imaging X-ray detector front-end with high dynamic range: IDeF-X HD.* in: NIM-A, Volume 695, Pages 415-419.

Gmar, M., Agelou, M., Carrel, F. & Schoepff, V., 2011. *GAMPIX: A new generation of gamma camera.* in: NIM-A, Volume 652, Issue 1, Pages 638-640.

Gosset, L., 2019. *Etude et optimisation des performances de l'instrument MXT, télescope X à micro-canaux, embarqué à bord de la mission spatiale d'astronomie SVOM.* at: Astrophysique [astro-ph]. Université Paris-Saclay.

H3D, 2020. *H400 - Gamma-Ray Imaging Spectrometer (datasheet).* [En ligne]
Available at: https://h3dgamma.com/H400Specs.pdf?

ISDC, 2010. *La mission INTEGRAL.* [En ligne]
Available at: https://www.isdc.unige.ch/integral/outreach/integral_fr

Ishikawa, S. et al., 2010. *Development of Double-Sided CdTe Strip Detectors for γ-Ray Imaging and Spectroscopy.* in: Japanese Journal of Applied Physics, Volume 49, Number 11R.

Limousin, O., 2016. *The story of Caliste: CdTe based pixel detectors for Hard X-Ray astronomy in space.* at: Habilitation à Diriger des Recherches, Université Paris-Diderot, CEA, http://irfu.cea.fr/Phocea/file.php?class=std&&file=Doc/Publications/Archives/irfu-16-38.pdf.

Limousin, O. et al., 2005. *IDeF-X ASIC for Cd(Zn)Te spectro-imaging systems.* in: IEEE Transactions on Nuclear Science, Volume: 52, Issue: 5.

Maier, D. et al., 2018. *Second generation of portable gamma camera based on Caliste CdTe hybrid technology.* in: Nuclear Instrumentation and Methods in Physics Research A, Volume 912, Pages 338-342.

Meuris, A., 2009. *Etude et optimisation du plan de détection de haute énergie en Cd(Zn)Te pour la mission spatiale d'observation astronomie X et gamma SIMBOL-X.* at: Université Paris-Diderot.

Meuris, A. et al., 2014. *Caliste-SO: the x-ray spectrometer unit of the STIX instrument onboard the Solar Orbiter space mission.* in: Proceedings Volume 9154, High Energy, Optical, and Infrared Detectors for Astronomy VI; 91540Y.

Mirion Technologies, 2019. *IPIX.* [En ligne]
Available at: https://www.mirion.com/products/ipix-ultra-portable-gamma-ray-imaging-system

Mirion Technologies, 2019. *iPIX - Ultra Portable Gamma-Ray Imaging System (datasheet).* [En ligne]
Available at: https://www.mirion.com/assets/ipix-ultra-portable-gamma-ray-imaging-system_gv3Kx6N.pdf

Montemont, G. et al., 2017. *NuVISION: a Portable Multimode Gamma Camera based on HiSPECT Imaging Module.* in: IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC).

NASA, 2000. *The Neil Gehrels Swift Observatory.* [En ligne]
Available at: https://swift.gsfc.nasa.gov/

NASA, 2020. *Hitomi.* [En ligne]
Available at: https://heasarc.gsfc.nasa.gov/docs/hitomi/

NUVIATECH Instruments, 2020. *NuVISION - Portable spectrometric gamma imaging system (specification sheet).* [En ligne]
Available at: https://www.nuviatech-instruments.com/wp-content/uploads/sites/3/2019/09/NVG-375022-FicheProduit-Sept2019-06-NuVISION-3.pdf

ORTEC, 2019. *905 Series - Na(Tl) Scintillation Detectors.* [En ligne]
Available at: https://www.htds.fr/wp-content/uploads/2019/04/instrumentation-nucleairespectrometre-alpha-et-gammadetecteurs_gammadetecteurs-a-scintillation-nai-et-labr3_905-2-1.pdf

ORTEC, 2020. *PROFILE "SP" Series - P-type Semi-planar HPGe Detectors.* [En ligne]
Available at: https://www.ortec-online.com/-/media/ametekortec/brochures/profile%20sp.pdf

SVOM, 2015. *SVOM.* [En ligne]
Available at: http://www.svom.fr/

Takeda, S. et al., 2012. *Applications and Imaging Techniques of a Si/CdTe Compton Gamma-Ray Camera.* in: Physics Procedia, Volume 37, Pages 859-866.

Wahl, C. G. et al., 2015. *The Polaris-H imaging spectrometer.* in: NIM-A, Volume 784, Pages 377-381.

Chapitre 2 : La vie d'un photon, de la source aux données

1 Observation des données et spectrométrie avec Caliste _____ 46

2 Modélisation des interactions photons/matière, de la source au détecteur_____ 64

3 Modélisation de la réponse du détecteur _____ 74

4 Modélisation des spectres, simulations et réponse du détecteur _____ 86

5 Bibliographie_____ 92

Afin de répondre aux problématiques de spectro-identification et de localisation de sources radioactives avec Caliste, nous devons d'abord nous familiariser avec les données qui sont issues du détecteur et les comprendre. À travers ce chapitre, je vais dans un premier temps montrer les données que je dois étalonner en énergie avec précision afin de les traiter en spectrométrie. Puis, je vais expliquer leur origine à partir des processus physiques d'émissions de photons et d'interaction rayonnement/matière en mettant en place des simulations Monte-Carlo et en modélisant la réponse spectrométrique du détecteur.

Cette étape est essentielle pour comprendre les caractéristiques des données qui sont discriminantes pour l'identification des radioéléments ainsi que pour être capable de simuler une large base de données de radioéléments qui ne se limite pas seulement aux sources que nous possédons en laboratoire. La base de données à traiter doit pouvoir contenir les radionucléides d'intérêt à traiter suivant les applications visées et, à travers ces simulations, elle pourra être adaptée selon les besoins de ces applications.

1 Observation des données et spectrométrie avec Caliste

1.1 Les données brutes issues de Caliste

1.1.1 La liste d'événements

Lorsqu'un événement a lieu dans le cristal en CdTe du détecteur Caliste, un signal est créé, acquis et traité par le détecteur. Un tel événement correspond généralement au dépôt d'énergie d'un photon dans le cristal, par absorption photoélectrique ou par diffusion Compton[*]. Plusieurs pixels peuvent être touchés en coïncidence temporelle par exemple si le dépôt d'énergie crée un nuage de charges s'étendant sur plusieurs pixels, ou encore si le dépôt initial crée un photon de fluorescence X dans le CdTe qui interagit ensuite dans un autre pixel. Dans tous les cas, si le signal enregistré est supérieur au seuil de détection défini par l'utilisateur, la chaîne d'acquisition renvoie les informations suivantes :

- le nombre de canaux qui ont enregistré un signal supérieur au seuil ;
- les numéros des ASIC qui ont enregistré l'événement ;
- les canaux des ASICs qui ont enregistré le ou les événements ;
- la date de l'événement ;
- l'amplitude maximale du signal PHA (*Pulse-Height Amplitude*) donnée en ADU (*Analog-to-Digital Unit*) après conversion du signal en sortie du détecteur de pic par l'ADC (*Analog-to-Digital Converter*), qui représente l'énergie déposée par l'interaction.

Le FPGA (*Field Programmable Gate Arrays*), en charge du contrôle de la séquence des ASICs, synchronise la conversion des données sous un format binaire dédié qui contient ces informations pour chaque événement. Ces données sont préparées sous forme de paquets de télémétrie et transmises via un lien Ethernet ou Wi-Fi vers l'ordinateur de contrôle et enregistrée par le logiciel d'acquisition. Notre équipe a développé un programme de conversion de ces fichiers binaires vers des fichiers sous format FITS[†] (*Flexible Image Transport System*), qui présentent l'intérêt d'être compatible avec tout système d'exploitation et importable avec différents langages de programmation (C, IDL, Python…) pour traitement.

Un exemple de liste d'événements visualisée avec le logiciel fitsview[‡] est présenté en Figure 1-1. Elle correspond à une mesure de source de ^{137}Cs avec Caliste. Au long de ce chapitre, je me baserai principalement sur l'exemple de ce radioélément pour expliquer les différentes étapes de d'étalonnage en énergie, que j'appelle aussi calibration, ainsi que ma modélisation des spectres de Caliste.

[*] Il est possible de détecter des signaux de bruit intempestif, dus par exemple aux fluctuations thermiques dans le détecteur.
[†] Le format FITS (https://fr.wikipedia.org/wiki/Flexible_Image_Transport_System) est un format de transfert habituellement utilisé pour les images astronomiques.
[‡] https://heasarc.gsfc.nasa.gov/ftools/fv/

Select	EVENT	MULT	MULT_i	TIME_s	PIX_DM	X	Y	PHA
All	1J	1B	1B	1D	1B	1B	1B	1I
				s				ADU
Invert	Modify	Modify	Modify	Modify	Modify	Modify	Modify	Modify
1972	1971	3	2	3.117970967293E+001	109	13	6	43937
1973	1972	1	0	3.119639801979E+001	68	4	4	10699
1974	1973	1	0	3.126165914536E+001	77	13	4	27669
1975	1974	1	0	3.127775478363E+001	24	8	1	10573
1976	1975	2	0	3.133321452141E+001	147	3	9	26596
1977	1976	2	1	3.133321452141E+001	148	4	9	13130
1978	1977	1	0	3.138047695160E+001	215	7	13	10628
1979	1978	1	0	3.139430046082E+001	235	11	14	10098
1980	1979	1	0	3.142004227638E+001	25	9	1	12082
1981	1980	1	0	3.145358181000E+001	231	7	14	9461
1982	1981	1	0	3.146714019775E+001	20	4	1	9996
1983	1982	2	0	3.151768660545E+001	248	8	15	19339
1984	1983	2	1	3.151768660545E+001	249	9	15	8659
1985	1984	1	0	3.153983998299E+001	179	3	11	10028
1986	1985	2	0	3.157222914696E+001	89	9	5	21400
1987	1986	2	1	3.157222914696E+001	90	10	5	10295
1988	1987	1	0	3.161406612396E+001	79	15	4	14402
1989	1988	1	0	3.165473365784E+001	181	5	11	13412
1990	1989	1	0	3.166722941399E+001	99	3	6	17866
1991	1990	1	0	3.171568346024E+001	189	13	11	10091
1992	1991	1	0	3.173738169670E+001	6	6	0	9974
1993	1992	1	0	3.177129340172E+001	66	2	4	26086
1994	1993	1	0	3.177526712418E+001	67	3	4	10500
1995	1994	1	0	3.182057952881E+001	177	1	11	9382
1996	1995	1	0	3.186107206345E+001	239	15	14	12696
1997	1996	1	0	3.187777471542E+001	0	0	0	9932
1998	1997	1	0	3.187963747978E+001	73	9	4	18506
1999	1998	1	0	3.190751886368E+001	98	2	6	31993
2000	1999	2	0	3.190970373154E+001	176	0	11	8680

Figure 1-1 : Exemple d'une liste d'événements enregistrée avec Caliste.

Le programme de conversion de fichier binaire à fichier FITS met en forme les données de sorte que leur exploitation soit simplifiée pour l'utilisateur, notamment pour les informations sur la position d'interaction. Ainsi, l'ensemble des informations temporelles, spatiales et spectrales des événements sont disponibles :

- EVENT contient le numéro de l'événement dans la liste ordonnée en temps ;
- MULT contient la multiplicité de l'événement, c'est-à-dire, le nombre de pixels touchés dans la même fenêtre temporelle, soit en coïncidence temporelle.
  On parle d'événement simple quand la multiplicité vaut 1, double pour 2, triple pour 3 et quadruple pour 4 et plus généralement, événement multiple si elle est supérieure ou égale à 2 ;
- MULT_i numérote de 0 à i-1 les événements d'une même multiplicité, où i est la multiplicité de l'événement suivant l'ordre de multiplexage ;
- TIME_s donne la date d'interaction en secondes, à partir du début de l'acquisition avec une précision de 10 ns, correspondant à la période de l'horloge d'acquisition ;
- PIX_DM donne le numéro du pixel d'interaction, de 0 à 255 ;
- X et Y donnent la ligne et la colonne du pixel touché, de 0 à 16, l'information est redondante avec PIX_DM mais simplifie le traitement en donnant directement les coordonnées du pixel ;
- PHA contient la valeur de l'amplitude PHA enregistrée en ADU, codée par l'ADC (*Analog/Digital Converter*) sur 16 bits, soit entre 0 et 65535.

1.1.2 Carte des coups et spectre

La représentation sous forme de listes d'événements nous empêche d'appréhender directement la mesure, d'autant qu'elles peuvent contenir de quelques éléments à plusieurs millions d'événements en fonction de la durée de la mesure et de l'intensité des rayonnements enregistrés par le détecteur. Il est nécessaire de représenter les données sous des formes interprétables physiquement, notamment la distribution spatiale des événements ainsi que leur distribution en énergie.

Carte de coups

La distribution spatiale des événements est représentée sous la forme d'un histogramme 2D, représentant la matrice du détecteur, et contenant le nombre d'événements enregistrés pour chacun des 256 pixels au cours d'une acquisition. Cet histogramme 2D est appelé carte de coups (*Hit map* en anglais) et un exemple est donné en Figure 1-2. Cette acquisition est faite pour une source de ^{137}Cs de 1,6 MBq, située directement devant la caméra, pendant 72 minutes, avec une haute tension de 300 V appliquée au cristal, à une température de détecteur d'environ -10°C.

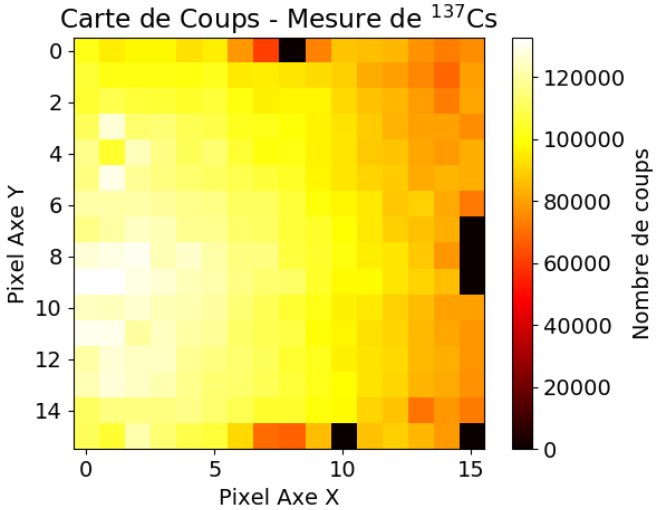


Figure 1-2 : Exemple de cartes de coups.
Acquisition d'une source de ^{137}Cs de 1,6 MBq placée à 1 cm du détecteur pendant 72 minutes.
Haute-tension appliquée au cristal : 300 V ; température du détecteur : -10°C ; temps de *peaking* : 6,07 µs.

Ce type de représentation porte déjà de l'information. La source a été placée légèrement sur la gauche du détecteur, ce que l'on observe sur la carte des coups par une zone plus brillante à gauche de l'image. L'alignement de la source est ici directement visible à travers le comptage total des pixels du fait que la source est positionnée très proche du détecteur, à près de 1 cm.

On peut voir des pixels enregistrant moins d'événements, apparaissant noirs. Ils correspondent à des pixels qui étaient bruyants dans ces conditions de mesure et dont le seuil en énergie a été augmenté pour éviter qu'ils ne génèrent un grand nombre de faux événements qui monopolisent le temps d'acquisition en créant par conséquent un temps mort important sur chaque événement.

Des pixels ont aussi un taux de comptage plus faible sur la première et sur la dernière ligne du détecteur, au milieu de chaque ligne. Ils sont situés sous des plots en or prévus pour l'interconnexion

de la haute tension, présents sur le cristal de ce modèle de détecteur. Ces plots absorbent une partie des photons de basse énergie, typiquement ici, à 32 keV et 37 keV pour le ^{137}Cs.

La distribution spatiale des événements est naturellement liée aux capacités d'imagerie du détecteur et elle est une information essentielle pour répondre à la problématique de la localisation des sources radioactives quelle que soit la méthode d'imagerie utilisée, comme je le présente en chapitres 4, 5 et 6.

Spectre

L'information en énergie est portée par la grandeur PHA et la distribution des valeurs de PHA représente la distribution des énergies déposées dans le cristal CdTe. La représentation de cette distribution se fait à partir d'un histogramme des valeurs des PHA, appelé spectre, dont un exemple est donné en Figure 1-3. La configuration d'acquisition est identique à celle donnée pour la Figure 1-2.

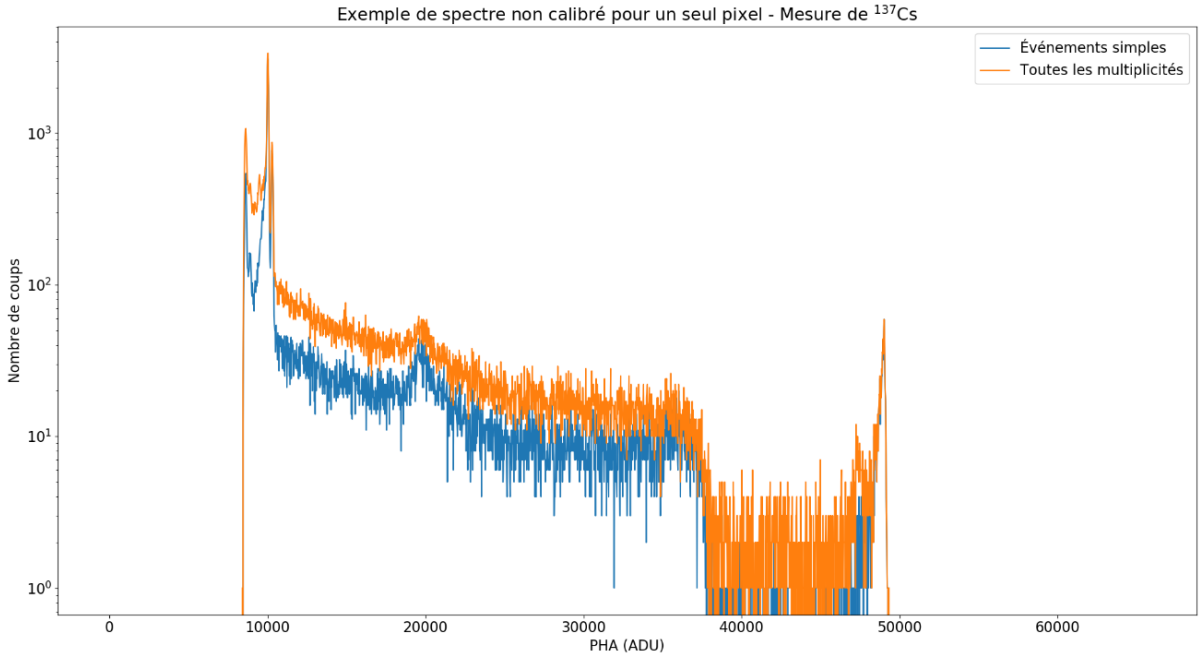


Figure 1-3 : Exemple de spectre non calibré sur un seul pixel.
Acquisition d'une source de ^{137}Cs de 1,6 MBq placée à 1 cm du détecteur pendant 72 minutes.
Haute-tension appliquée au cristal : 300 V ; température du détecteur : -10°C ; temps de *peaking* : 6,07 µs.

J'ai représenté le spectre pour un seul pixel avec d'une part uniquement les événements simples, qui n'ont déclenché que ce pixel, et d'autre part, avec toutes les multiplicités, c'est-à-dire les événements qui ont déclenché ce pixel mais qui ont aussi pu déclencher d'autres pixels en même temps.

L'intérêt des spectres en événements simples réside dans le fait qu'ils correspondent à des dépôts d'énergie uniquement dans un pixel et ne sont pas impactés par un éventuel partage de charge, c'est-à-dire, des dépôts d'énergie qui se partagent entre au moins deux pixels voisins, avec une énergie supérieure au seuil de détection dans chaque pixel.

L'analyse du spectre est essentielle. Le spectre est la signature de la source que l'on cherche à étudier. On peut notamment repérer des structures de pics qui correspondent à des interactions photoélectriques aux énergies correspondant aux lignes d'émission du radioélément, et des structures

continues au milieu du spectre qui correspondent aux dépôts d'énergie Compton (voir annexe A). L'objectif de ce chapitre est d'expliquer et comprendre finement l'ensemble de ces structures.

1.2 Étalonnage en énergie

1.2.1 Principe

Les spectres issus des données brutes sont donnés à partir des PHA en ADU, grandeur adimensionnée non physique. Représenter les spectres en énergie à travers une étape de calibration est essentiel et a plusieurs impacts. Cela permet d'effectuer une analyse spectrométrique plus fine basée sur des grandeurs physiques.

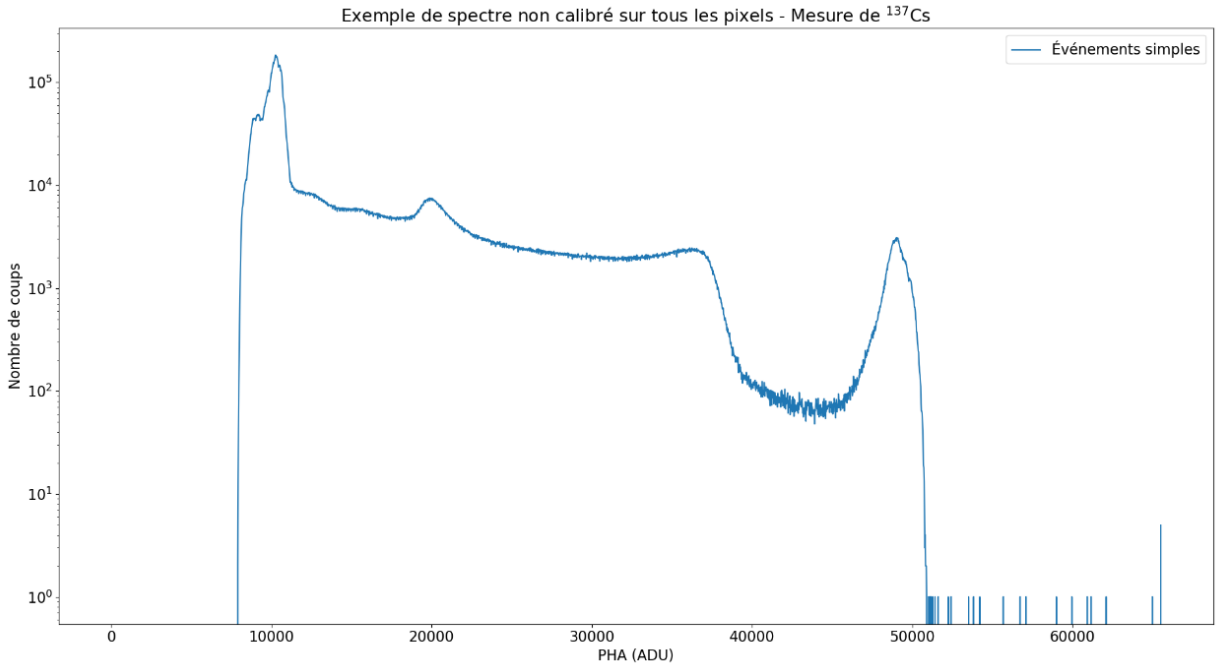


Figure 1-4 : Spectre sur l'ensemble des pixels en ADU, sans calibration.
Acquisition d'une source de ^{137}Cs de 1,6 MBq placée à 1 cm du détecteur pendant 72 minutes.
Haute-tension appliquée au cristal : 300 V ; température du détecteur : -10°C ; temps de *peaking* : 6,07 µs.

Un étalonnage en énergie permet de faire des cartes de coups en sélectionnant les événements suivant l'énergie déposée et construire des spectro-images dans différentes bandes d'énergie. Par ailleurs, pour pouvoir exploiter l'information issue de l'ensemble des pixels et faire un spectre somme, la calibration est nécessaire, comme on peut le voir sur la Figure 1-4. En effet, contrairement à la Figure 1-3, les pics photoélectriques sont beaucoup plus dispersés. Cela démontre notamment que la calibration à appliquer est différente pour chaque pixel individuellement.

1.2.2 Principe et loi de calibration

La calibration consiste à associer à chaque PHA une énergie E correspondante, c'est-à-dire trouver la fonction f telle que $E = f(\mathrm{PHA})$. Généralement, on utilise une fonction f_P qui dépend de paramètres P dont on recherche les valeurs par cette étape de calibration.

Lorsqu'un photon interagit dans le cristal du détecteur, un nombre de paires électrons/trous, proportionnel à l'énergie déposée par le photon, est créé, induisant donc un signal qui est lui-même

proportionnel à l'énergie déposée. Puis, l'électronique de lecture, l'ASIC IDeF-X HD a été conçu pour avoir une réponse affine en fonction de l'amplitude du signal mesuré (Michalowska, et al., 2010). Cependant, plusieurs phénomènes induisent des non linéarités. D'une part, la physique sur laquelle se base le fonctionnement des transistors n'est pas linéaire, mais exponentielle, on parle de non linéarité intégrale. D'autre part, des phénomènes de non linéarité différentielle ont lieu au niveau de l'ADC, du fait que les *bins* sur lesquels sont codées les données n'ont pas tous la même largeur.

Je modélise la loi de calibration de la manière suivante :

$$E = f_{P=(g,o,d)}(\mathrm{PHA}) = g \times (\mathrm{PHA} - o)^d \tag{1-1}$$

On appelle g le gain et o l'*offset* de la loi de calibration. d est un paramètre de distorsion qui intervient pour prendre en compte les phénomènes de non linéarité. g, o et d sont appelés paramètres de calibration.

Dans un premier temps, j'utilise une loi de calibration affine, avec $d = 1$, qui est la loi généralement utilisée du fait que la chaîne d'acquisition est conçue pour fournir une réponse la plus linéaire possible. Cela permet de ne prendre en compte que deux paramètres par pixel, ce qui est important pour certains domaines comme dans le cas des applications spatiales, où les performances de calcul et de stockage sont limitées. Dans un second temps, j'applique la loi de calibration complète en recherchant aussi le paramètre de distorsion d optimal en paragraphe 1.2.6.

1.2.3 Détermination des paramètres de calibration par la méthode ECC (Energy Calibration via Correlation)

Chaque pixel a son propre gain et son propre *offset* qui doivent être déterminés. Classiquement, la calibration se fait en présentant au détecteur une source d'étalonnage, dont les lignes d'émission sont connues et en effectuant la mesure et le spectre en ADU. Puis, les pics photoélectriques sont recherchés dans le spectre en ajustant une gaussienne pour chaque pic et en relevant leur position. Les paramètres de calibration sont déterminés pour associer au mieux les énergies des lignes d'émission à la position de chaque pic en ADU.

Cependant, cette méthode nécessite que les pics photoélectriques soient suffisamment intenses dans le spectre en ADU, soit avec un temps de pose suffisamment long, pour être ajustés par une gaussienne. Cela représente un inconvénient pour pouvoir corriger d'éventuelles dérives du détecteur en temps réel, de l'ordre de quelques secondes, avec des sources d'intensité faible ou pour lesquelles les pics photoélectriques sont peu intenses.

Une autre méthode, nommée ECC pour *Energy Calibration via Correlation*, a été développée (Maier & Limousin, 2016) et présente de meilleures performances que la méthode par recherche des pics, notamment lorsque le spectre ne contient qu'un faible nombre d'événements.

La méthode ECC consiste à créer un spectre synthétique de la source d'étalonnage puis à rechercher les paramètres de calibration qui doivent être appliqués pour obtenir la meilleure corrélation entre le spectre mesuré et ce spectre synthétique.

Cette méthode présente l'avantage d'exploiter des données complexes en ne se concentrant pas uniquement sur les pics photoélectriques. Par exemple, on peut auto-aligner les réponses des pixels

les uns par rapport aux autres avant de procéder à la calibration en énergie, en utilisant comme spectre de référence le spectre d'un des pixels. Par ailleurs, cette méthode est tolérante aux structures continues dans le spectre comme les structures Compton, qui portent aussi de l'information en énergie, ou le bruit de fond.

Formellement, notons $S_{\text{syn}}(E)$ l'intensité du spectre synthétique pour l'énergie E. Je désigne par $S_{\text{mes}} = \left(S_{\text{mes}_i}\right)_{i\in\text{bins}}$ le spectre mesuré pour une discrétisation des PHA donnée et par $\text{PHA} = (\text{PHA}_i)_{i\in\text{bins}}$, les canaux de la discrétisation.

Pour des paramètres de calibration (g, o) donnés, on calcule le spectre synthétique associé à ces paramètres : $\left(S_{\text{syn}}(E_i)\right)_{i\in\text{bins}}$ où $E_i = g \times (\text{PHA}_i - o)$, comme représenté en Figure 1-5. Notons qu'une autre loi de calibration f_P peut être appliquée avec un autre nombre de paramètres P, cette approche reste valable en prenant plus généralement $E_i = f_P(\text{PHA}_i)$.

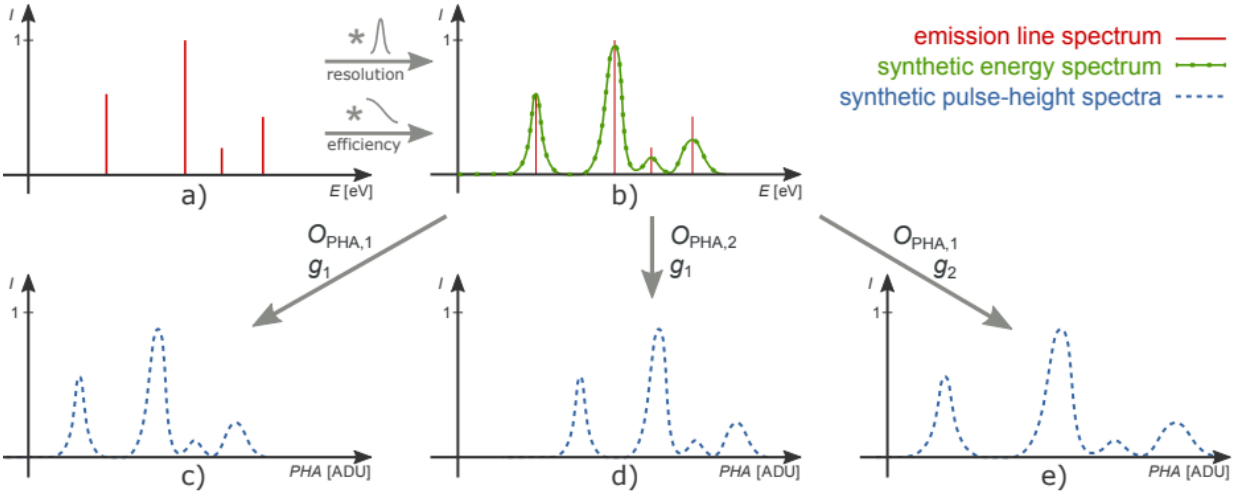


Figure 1-5 : Calcul du spectre synthétique (Maier & Limousin, 2016).

Le score de corrélation est donné par :

$$C(g, o) = \langle S_{\text{syn}}\big(g \times (\text{PHA} - o)\big) \mid S_{\text{mes}} \rangle = \sum_{i\in\text{bins}} S_{\text{syn}}(E_i) \times S_{\text{mes}_i} \tag{1-2}$$

où $\langle S_{\text{syn}}\big(g \times (\text{PHA} - o)\big) \mid S_{\text{mes}} \rangle$ est le produit scalaire entre le spectre synthétique et le spectre mesuré.

L'objectif est alors de trouver les paramètres (g, o) qui maximisent $C(g, o)$.

Il n'est pas nécessaire que le spectre synthétique soit entièrement fidèle au spectre mesuré. Par exemple, le spectre synthétique peut ne contenir que les pics photoélectriques, sans les continuums Compton. La méthode reste valable mais le score de corrélation maximum atteignable sera plus faible par rapport à un spectre synthétique plus fidèle.

En revanche, cela est aussi une force de la méthode puisqu'elle permet d'utiliser un spectre plus complet et fidèle aux mesures si l'utilisateur en dispose. Par exemple, il est possible d'utiliser un spectre calibré issu d'une ancienne mesure avec le détecteur comme spectre de référence.

Dans le cas où le radioélément présente certaines de ses lignes d'émission principales dans les hautes énergies, au-delà de 300 keV, du fait de la faible probabilité d'interagir par effet photoélectrique dans le détecteur (voir Annexe A), les pics correspondants apparaissent avec une faible intensité dans le spectre. De ce fait, ils n'ont pas beaucoup d'influence sur le score de corrélation et donc ne sont pas pris en compte pour la calibration. Afin de résoudre ce problème, une solution consiste à exprimer le spectre en échelle logarithmique, c'est-à-dire à calculer la corrélation suivante :

$$C(g,o) = \langle \log\big(S_{\mathrm{syn}}\big(g \times (\mathrm{PHA} - o)\big) + 1\big) \mid \log(\, S_{\mathrm{mes}} + 1)\rangle \tag{1-3}$$

On ajoute 1 à chaque canal avant d'appliquer la fonction logarithme pour éviter d'appliquer le logarithme à des canaux qui ne contiennent aucun coup.

J'utilise principalement cette transformation avec la fonction logarithmique sur le spectre de ^{137}Cs, de sorte que le pic à 662 keV soit important dans le score de corrélation.

Il est également possible d'augmenter volontairement et artificiellement l'intensité de certaines structures dans le spectre synthétique, ce qui revient à les pondérer. Cela permet de faire en sorte que le score de corrélation dépende plus fortement de ces structures, comme par exemple un pic photoélectrique sur lequel on veut mettre un poids plus important.

1.2.4 Maximisation de la corrélation par recherche exhaustive et algorithme génétique

La première approche directe pour trouver les paramètres qui maximisent la corrélation est de tester exhaustivement toutes les valeurs de gains et d'*offsets* sur une plage donnée, mais cette approche peut être relativement longue. Par exemple, pour une recherche en testant 1000 valeurs de gains et 1000 valeurs d'*offsets*, cette méthode demande 2 mn 30 de calibration par pixel, calculée sur un ordinateur de bureau classique, ce qui demande donc plus de 11 h de calibration pour l'ensemble des 256 pixels du détecteur.

Cette approche est envisageable pour une première calibration « d'usine » d'un détecteur, en laboratoire. En revanche, cela n'est pas compatible avec des besoins en temps courts, de l'ordre de quelques secondes à quelques minutes, pour des corrections de calibration dues à des dérives du détecteur. De même, si on souhaite caractériser les performances du détecteur suivant par exemple les différents temps de *peaking*, pour chaque paramètre testé, une nouvelle calibration est nécessaire, et avec 11 h de calibration à chaque essai, ce travail devient extrêmement laborieux.

Pour effectuer une calibration en temps réel, une autre méthode est nécessaire. Cependant, le problème d'optimisation consistant à trouver les paramètres (g,o) qui maximisent la corrélation $C(g,o)$ est un problème difficile car la fonction $C(g,o)$ n'est pas concave comme on peut le voir en Figure 1-6. Cette figure est la carte de corrélation correspondant au spectre de la Figure 1-3.

Sur cette carte de corrélation, il est intéressant de remarquer trois lignes verticales. La ligne du centre correspond à une forte corrélation liée au bon positionnement des pics à 32 keV et 37 keV qui sont les

pics les plus intenses du spectre. La corrélation est très forte grâce au bon positionnement de ces deux pics.

La ligne de gauche correspond à des cas où le pic à 32 keV est positionné en fait à 37 keV avec les paramètres de calibration induisant une forte corrélation à cette position. En revanche, le pic à 37 keV est positionné au-delà de 37 keV sur le spectre induisant une faible corrélation à cet endroit. Au total, la corrélation est moins bonne que lorsque les deux pics sont bien positionnés. De même, la ligne de droite correspond à des cas où le pic à 37 keV est positionné en fait à 32 keV.

La ligne diagonale en haut à gauche correspond à des paramètres de calibration qui positionnent le pic à 662 keV au bon endroit. Ainsi, la plus forte corrélation est obtenue à l'intersection des lignes les plus intenses, c'est-à-dire, lorsque les trois pics sont à la bonne position avec les paramètres de calibration donnés.

Il est important de noter sur cette carte de corrélation la présence de plusieurs maxima locaux. Cela empêche d'utiliser des algorithmes d'optimisation basés sur le calcul du gradient qui peuvent converger vers l'un ou l'autre de ces maxima locaux. En l'absence d'a priori, il est donc nécessaire d'utiliser d'autres algorithmes d'optimisation pour rechercher les paramètres de calibration.

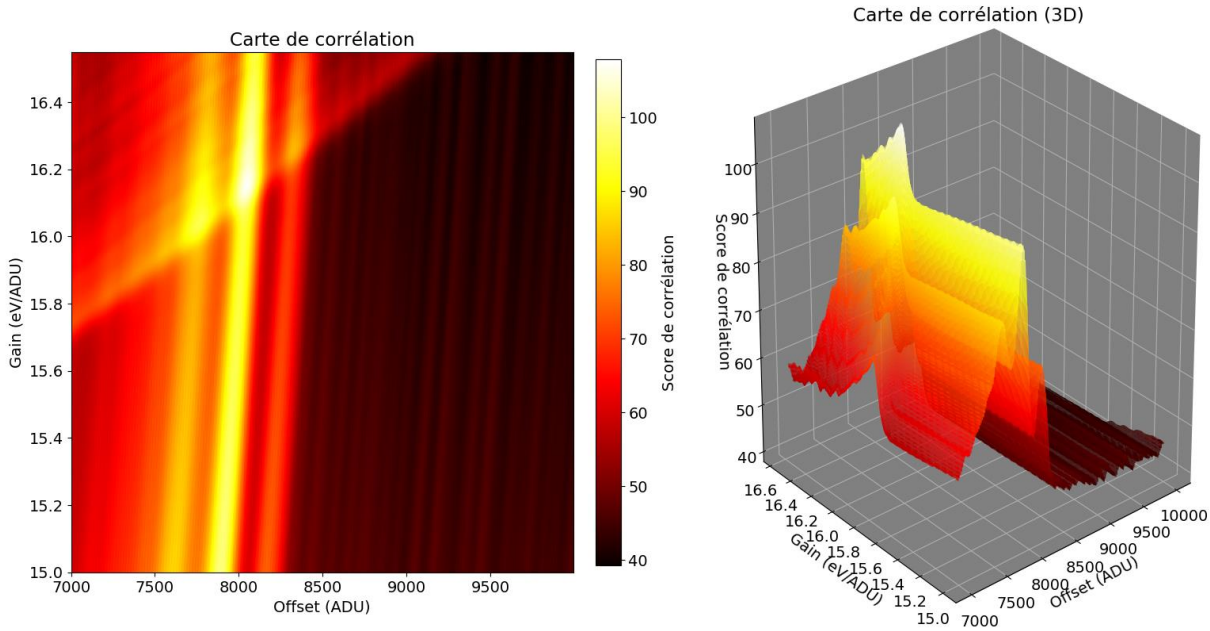


Figure 1-6 : Carte du score de corrélation pour une source de ^{137}Cs pour un pixel du détecteur. À gauche : carte de corrélation en 2D. À droite : représentation en 3D.

ECC-AMR

Une première méthode, nommée ECC-AMR (*Adaptive Mesh Refinement*)* (Maier, et al., 2020) consiste à évaluer le score de corrélation sur un petit nombre de paramètres, de l'ordre de la dizaine, répartis uniformément dans la zone de recherche des paramètres. Les paramètres donnant un score de corrélation supérieur à un certain seuil sont conservés et la recherche est affinée dans une zone autour

* Cette méthode a été présentée à la conférence ANIMMA 2019 et publiée en proceedings (Maier, et al., 2020)

de ces paramètres. Puis le processus est réitéré avec ces nouveaux paramètres jusqu'à atteindre un niveau de précision décidé par l'utilisateur. La Figure 1-7 schématise ce processus.

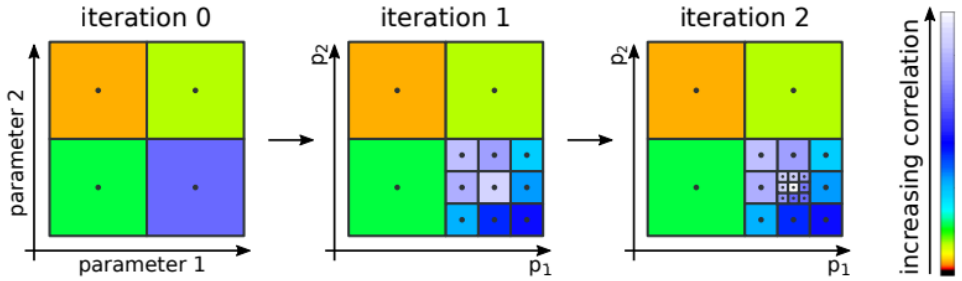


Figure 1-7 : Processus de recherche des paramètres ECC-AMR (Maier, et al., 2020).

Cette méthode peut cependant manquer les bons paramètres. En effet, sur la Figure 1-7, l'affinage lors de la première itération a lieu uniquement dans le carré en bas à droite. Cependant, la corrélation est évaluée pour les paramètres uniquement au centre de chaque zone. Il est donc possible que la corrélation maximale soit quand même dans une zone qui n'est pas affinée, par exemple, dans le carré en bas à gauche de la Figure 1-7. Pour résoudre ce problème, une étape d'expansion est ajoutée : si à l'itération i, une des zones du bord possède une corrélation supérieure au seuil, alors on fait aussi un affinage dans la zone voisine de l'itération i-1, comme cela est représenté en Figure 1-8.

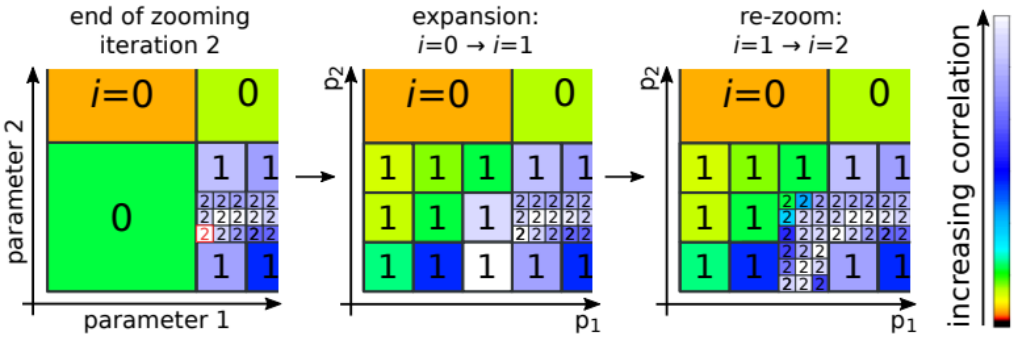


Figure 1-8 : Étape d'expansion ECC-AMR (Maier, et al., 2020).

Sur cette figure, grâce à cette étape d'expansion, une recherche est aussi effectuée dans le carré en bas à gauche de l'itération 0, car à l'itération 2, on a remarqué un score de corrélation élevé dans certaines zones voisines du carré à droite.

La méthode ECC-AMR accélère la recherche des paramètres de calibration optimaux en évaluant le score sur un nombre bien plus réduit de paramètres. Typiquement, le nombre de paramètres testés est diminué d'un facteur 450, tout en atteignant finalement la même précision, ce qui réduit d'un même facteur le temps de calcul, passant de 11 h à moins de 2 minutes pour calibrer l'ensemble du détecteur.

Algorithme génétique

J'ai développé une autre approche pour répondre au problème de la recherche des paramètres de calibration optimaux, basée sur un algorithme génétique. Les algorithmes génétiques constituent

une classe d'algorithme utilisée pour des optimisations de fonctions non convexes ou non concaves, à plusieurs paramètres, efficacement en temps d'exécution et en précision (Whitley, 2001).

Le principe de cet algorithme est le suivant :

- Initialisation : On choisit une population de paramètres (g, o) dans une plage de donnée, aléatoirement, composée de N individus ;
- Sélection : Pour tous les individus composant la population, on évalue leurs scores de corrélation $C(g, o)$ et on classe les individus selon leurs scores de corrélations. On sélectionne la fraction f_{graded} des meilleurs individus selon ce score. Parmi les individus non sélectionnés, on choisit une fraction $f_{\text{nongraded}}$ d'individus sur le reste de la population. Cette deuxième sélection évite que la population se bloque sur des maxima locaux.
- Hérédité : Les individus sélectionnés sont appelés les parents. On sélectionne aléatoirement des couples de parents (g_1, o_1) et (g_2, o_2) et on fabrique un nouvel individu à partir de ces couples. J'ai choisi de fabriquer un nouvel individu avec $g_{\text{enfant}} = \lambda_g g_1 + (1 - \lambda_g) g_2$ et $o_{\text{enfant}} = \lambda_o o_1 + (1 - \lambda_o) o_2$ où λ_g et λ_o sont tirés aléatoirement entre 0 et 1. On fabrique suffisamment d'individus pour que la population finale revienne à la population initiale N.
  La manière de concevoir un nouvel individu peut être adaptée et d'autres solutions peuvent être testées. On peut par exemple tirer aléatoirement λ_g et λ_o sur une gaussienne centrée en 0,5 et d'écart-type 0,25, pour donner plus de poids aux valeurs situées à égale distance des deux individus parents.
- Mutation : Parmi tous les nouveaux individus, on en choisit une certaine fraction f_{mut} qui va subir des mutations. Pour un individu $(g_{\text{init}}, o_{\text{init}})$, j'ai choisi de tirer un nouvel individu g_{mut} dans une loi $\mathcal{N}(g_{\text{init}}, \sigma_g)$ et o_{mut} dans une loi $\mathcal{N}(o_{\text{init}}, \sigma_o)$. Cette étape permet d'explorer au-delà des maxima locaux.
  Tout comme pour l'hérédité, la manière de fabriquer ces mutations peut être adaptée. On peut par exemple choisir de ne modifier que le gain ou l'*offset* aléatoirement, ou encore de tirer un gain ou un *offset* totalement nouveau.
- Les étapes sélection, hérédité et mutation sont réitérées sur plusieurs générations n_g définies initialement. L'individu finalement choisi est celui qui possède le meilleur score de corrélation dans la dernière génération, qui est le meilleur score de corrélation parmi tous les individus explorés sur toutes les générations, du fait que la phase de sélection garantit que les meilleurs individus sont reconduits à chaque génération.
  À chaque génération, il suffit d'évaluer le score de corrélation seulement pour les enfants, celui des parents ayant été évalué à la génération précédente.

J'ai déterminé les paramètres de l'algorithme en procédant à des essais sur des données réelles, de sorte à minimiser le nombre de mauvaises calibrations tout en maintenant un temps de calcul raisonnable.
Pour l'initialisation, j'ai choisi une population de $N = 200$ individus, sur $n_g = 20$ générations. Ce sont principalement ces deux paramètres qui influent à la fois sur les performances et le temps de calcul. Pour la sélection, j'ai choisi de retenir $f_{\text{graded}} = 20\ \%$ des individus sur la base de leurs scores, ainsi que $f_{\text{nongraded}} = 20\ \%$ des individus parmi le reste de la population. En effet, j'ai remarqué que si je choisis de retenir moins d'individus parmi le reste de la population, l'algorithme avait tendance à se bloquer dans des maxima locaux.

Pour la mutation, j'ai choisi une fraction de mutation $f_{\mathrm{mut}} = 20\,\%$ avec $\sigma_g = 1\ \mathrm{eV/ADU}$ et $\sigma_o = 500\ \mathrm{ADU}$, ces deux valeurs correspondent environ à 5 % des gains et *offsets* attendus sur mes mesures.

La plage de valeurs initiales que j'ai choisies pour le gain est comprise entre 15 eV/ADU et 16.5 eV/ADU et pour l'*offset* entre 7 000 et 10 000 ADU. Ces valeurs dépendent de l'ADC utilisé et éventuellement du détecteur. Certains pixels peuvent par exemple présenter un gain plus élevé ou plus faible du fait de la dispersion technologique de fabrication des circuits de lecture et capteurs CdTe. Cependant, un des intérêts de ce type d'algorithme est qu'il autorise les valeurs des individus à sortir des plages de valeurs initiales à travers l'étape de mutation.

La Figure 1-9 présente l'exécution de l'algorithme sur différentes générations.

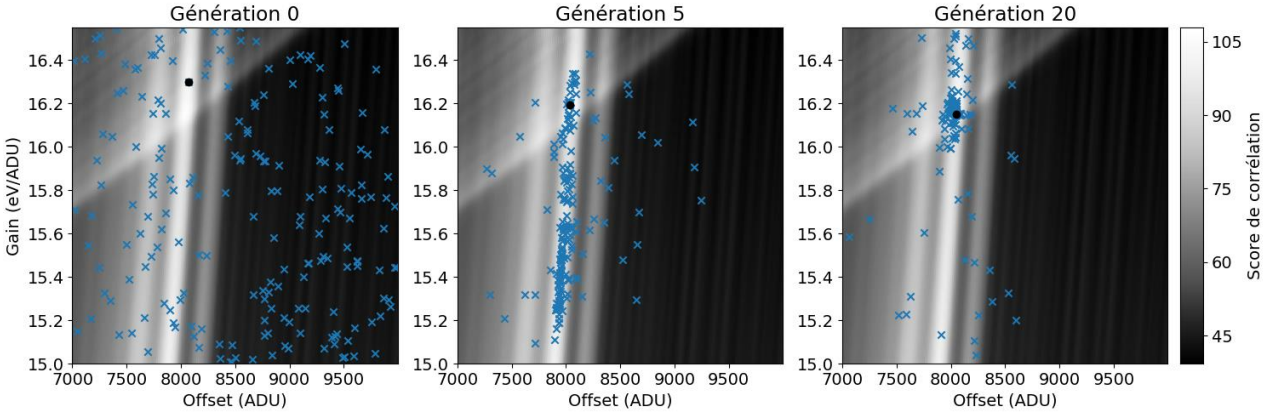


Figure 1-9 : Différentes générations de l'algorithme génétique. Les croix représentent les individus et le point noir représente l'individu avec le meilleur score de corrélation.

Sur cette figure, j'ai représenté la carte de corrélation. Les données correspondent toujours au spectre de la Figure 1-3. La carte de corrélation que j'affiche en gris est à titre d'illustration, c'est celle qui a été déterminée par la recherche exhaustive des paramètres, elle est inconnue pour l'algorithme génétique.

On remarque qu'à la génération 0, la population est répartie aléatoirement dans l'espace des paramètres. À la génération 5, les individus ont globalement convergé sur la ligne la plus intense, qui correspond au positionnement correct des pics à 32 keV et 37 keV. Enfin, à la génération 20, on voit que l'ensemble des individus ont globalement convergé autour du maximum de corrélation, les dernières générations ont affiné la recherche du maximum. Quelques points restent dispersés dans le reste de l'espace du fait du taux d'individu $f_{\mathrm{nongraded}}$ ainsi que des mutations.

Avec les paramètres choisis pour l'algorithme génétique, la calibration sur l'ensemble du détecteur donne entre 0 % et 2 % d'erreurs de calibration. Ces erreurs sont principalement dues à des pixels qui ont un fort pic de bruit aux basses énergies. Cela entraîne une forte corrélation avec le pic à 32 keV et l'algorithme tente alors de le positionner sur ce pic de bruit. Pour répondre à cet effet qui est présent en réalité quel que soit l'algorithme d'optimisation utilisé, il faut adapter la manière dont le score de corrélation est calculé. Cette adaptation est cependant complexe à réaliser de manière générale car la forme et la position des structures de bruit peuvent être très différentes suivant les pixels.

Par rapport à une recherche exhaustive, le temps de calcul avec ces paramètres pour l'algorithme génétique est réduit d'un facteur 580, ce qui représente donc un temps de calibration de 0,3 s par pixel. En augmentant le nombre de générations ou le nombre d'individus dans la population, cela allonge le temps de calcul mais permet aussi d'être plus conservatif en assurant la convergence de l'algorithme.

Une autre possibilité pour assurer la convergence de l'algorithme est d'utiliser une méthode de *boosting*, c'est-à-dire de réappliquer plusieurs fois l'algorithme, qui donne une réponse différente à chaque calcul du fait du caractère aléatoire dans l'algorithme, et donc de sélectionner à la fin la meilleure corrélation. Cela ralentira le calcul, autant de fois que l'algorithme sera réappliqué. Typiquement, avec les paramètres de l'algorithme que j'ai donnés, un *boosting* répété trois fois est suffisant pour obtenir les paramètres optimaux.

Cet algorithme calcule donc les paramètres optimaux de calibration en un temps raisonnable de 0,3 s par pixel, jusqu'à 0,9 s par pixel en utilisant une méthode de *boosting* répétée trois fois. Par ailleurs, un avantage de ce type d'algorithme est qu'il peut traiter des cas où le nombre de paramètres à optimiser est plus important, qui est une limitation de la méthode AMR. En effet, pour la méthode AMR, l'étape d'expansion est plus délicate à réaliser lorsque le nombre de paramètres est plus grand que 3, notamment pour définir la notion de zone « voisine » en dimension supérieure à 3.

1.2.5 Résultats de calibration

Spectre

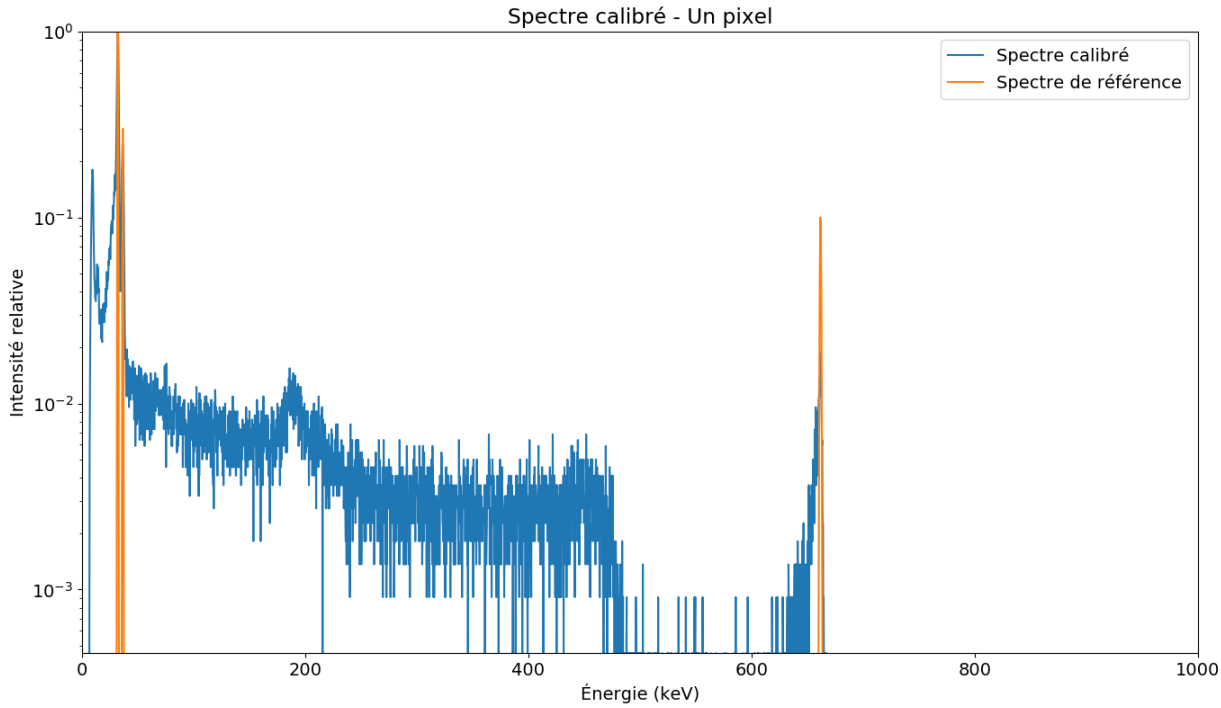


Figure 1-10 : Spectre calibré, un seul pixel, événements simples
Acquisition d'une source de ^{137}Cs de 1,6 MBq placée à 1 cm du détecteur pendant 72 minutes.
Haute-tension appliquée au cristal : 300 V ; température du détecteur : -10°C ; temps de *peaking* : 6,07 µs.

La Figure 1-10 présente le résultat de la calibration pour le spectre de la Figure 1-3, c'est-à-dire un spectre de ^{137}Cs d'événements simples sur un pixel unique. Les paramètres de calibration sont ceux

trouvés par l'algorithme génétique en Figure 1-9, en utilisant un spectre de référence composé des pics photoélectriques principaux du spectre de ^{137}Cs. Une transformation logarithmique est appliquée aux deux spectres et l'intensité relative du pic à 662 keV est volontairement accentuée dans le spectre de référence pour que la corrélation dépende fortement du bon positionnement de ce pic.

La Figure 1-11 présente la carte de gains et d'*offsets* issus de la calibration, calculée sur l'ensemble des pixels d'un module Caliste, soit 256 pixels. La carte des gains est assez uniforme sur l'ensemble du détecteur. Quatre pixels ont un gain manifestement un peu plus élevé. De même, la carte des *offsets* est uniforme, avec des valeurs un peu plus faibles pour trois pixels à droite, situés au niveau du plot d'interconnexion en or sur lequel est amenée la haute tension. On peut remarquer des ensembles de lignes de pixels qui ont des gains et *offsets* un peu plus proches. Elles correspondent aux différents ASICs présents dans le détecteur. Chaque ASIC lit en effet deux lignes de pixels.

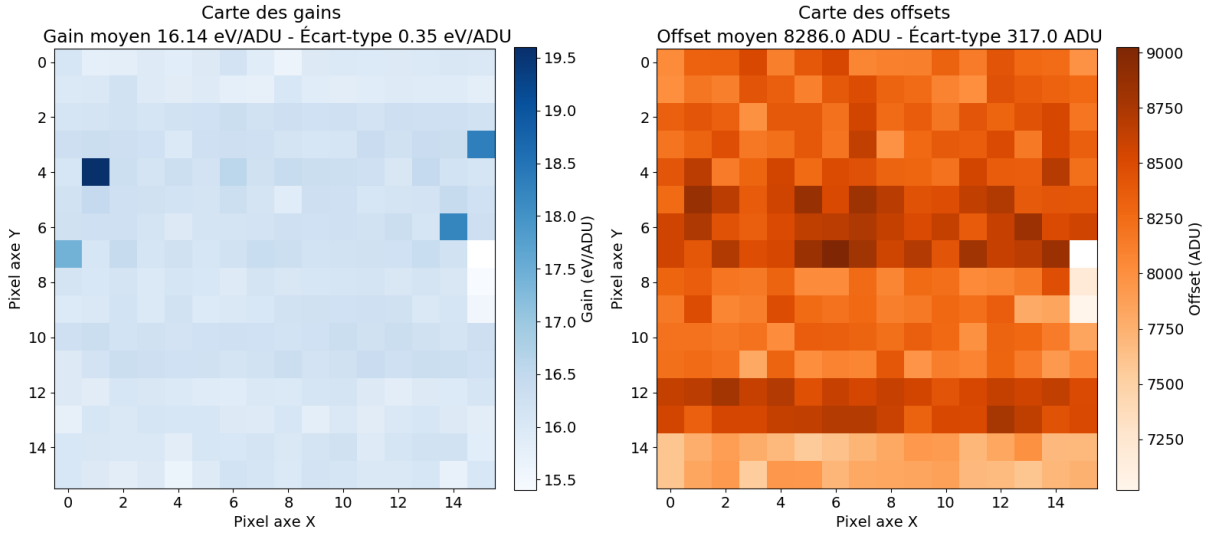


Figure 1-11 : Carte des gains et *offsets* du détecteur.

Les paramètres de calibration étant connus, je peux les appliquer sur l'ensemble des pixels et faire le spectre somme, qui correspond donc au spectre de la Figure 1-4, mais calibré cette fois.

Sur ce spectre, à l'aide des connaissances de lignes d'émission du ^{137}Cs et des processus physiques d'interaction photon-matière décrits en Annexe A, je peux désormais identifier avec précision les énergies associées à ces processus. Les différentes structures sont identifiées sur la Figure 1-12.

On peut distinguer plusieurs types de structures. Les pics photoélectriques correspondent à des photons ayant atteint le détecteur avec l'ensemble de leur énergie puis déposé cette énergie par absorption photoélectrique. Les pics d'échappement correspondent à des absorptions photoélectriques ayant créé un photon de fluorescence X de Cd (environ 23 et 26 keV) ou Te (environ 27 et 31 keV) qui s'est échappé du cristal. La structure continue Compton correspond à des photons ayant effectué une diffusion Compton dans l'environnement ou dans le détecteur.

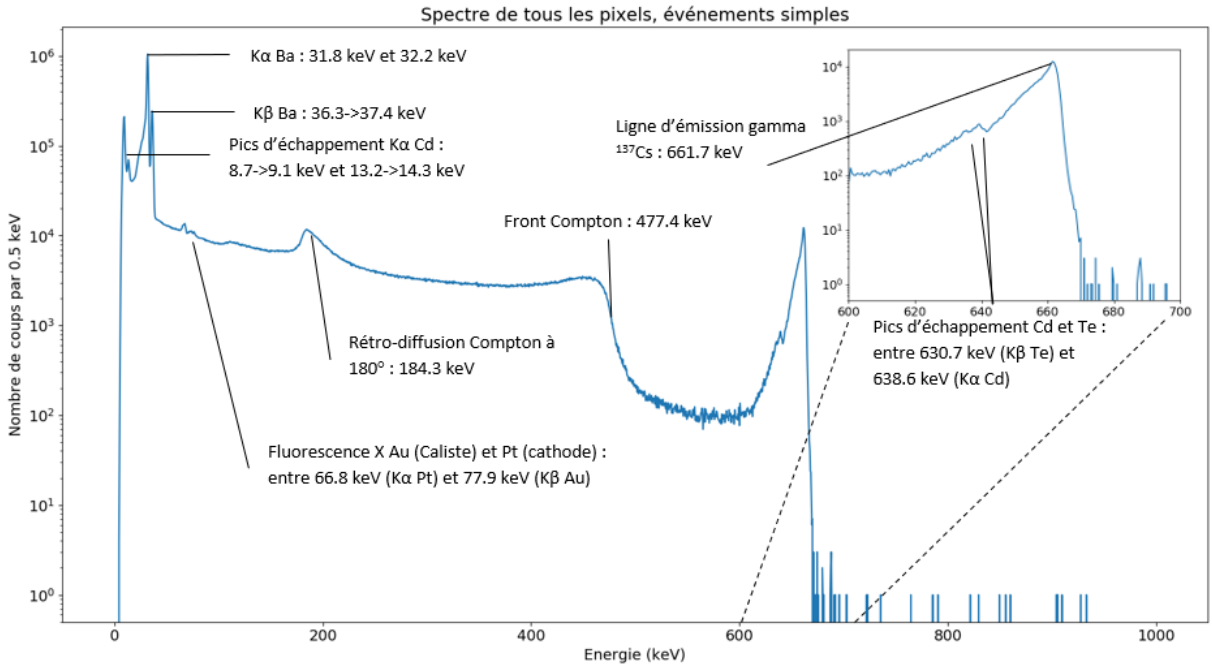


Figure 1-12 : Spectre de 137Cs calibré - Tous les pixels utilisés, événements simples seulement. Acquisition d'une source de ^{137}Cs de 1,6 MBq placée à 1 cm du détecteur pendant 72 minutes. Haute-tension appliquée au cristal : 300 V ; température du détecteur : -10°C ; temps de *peaking* : 6,07 µs.

Spectro-image

La calibration est nécessaire pour faire des spectro-images. Cela consiste à faire une carte des coups en sélectionnant les événements suivant leur énergie. Ces images en bande d'énergie sont aussi appelées images hyper spectrales dans certains domaines.

La Figure 1-13 présente une mesure de ^{137}Cs que j'ai acquise en plaçant un masque codé entre la source et le détecteur. J'ai effectué une sélection en énergie entre 30 keV et 40 keV, qui correspond à la bande comprenant les deux pics principaux du spectre de ^{137}Cs.

Les motifs du masque codé apparaissent déjà légèrement sans la sélection en énergie, mais on constate qu'ils apparaissent plus distinctement dans la bande 30-40 keV. Cela provient du fait qu'à 662 keV, les photons peuvent traverser le masque sans être absorbés ou en étant éventuellement diffusés, ce qui crée cette répartition plus uniforme sur le plan détecteur, visible sur l'image faite en dehors de la bande 30-40 keV.

L'utilisation de spectro-images présente deux avantages pour la localisation par masque codé :

- elles permettent d'obtenir potentiellement des images avec un meilleur contraste si on se focalise sur les pics photoélectriques à basse énergie, ce qui implique de meilleures performances de localisation ;
- on peut effectuer une sélection en énergie pour localiser différents radioéléments séparément et donc faire l'image d'un radioélément en particulier.

Je présenterai cela aux Chapitres 4 et 5, concernant l'utilisation du masque codé pour répondre à la problématique de localisation des sources radioactives. Par ailleurs, dans le cadre de l'imagerie

Compton en Chapitre 6, nous verrons que la connaissance des énergies déposées par les interactions est une donnée essentielle pour utiliser ce type d'imagerie.

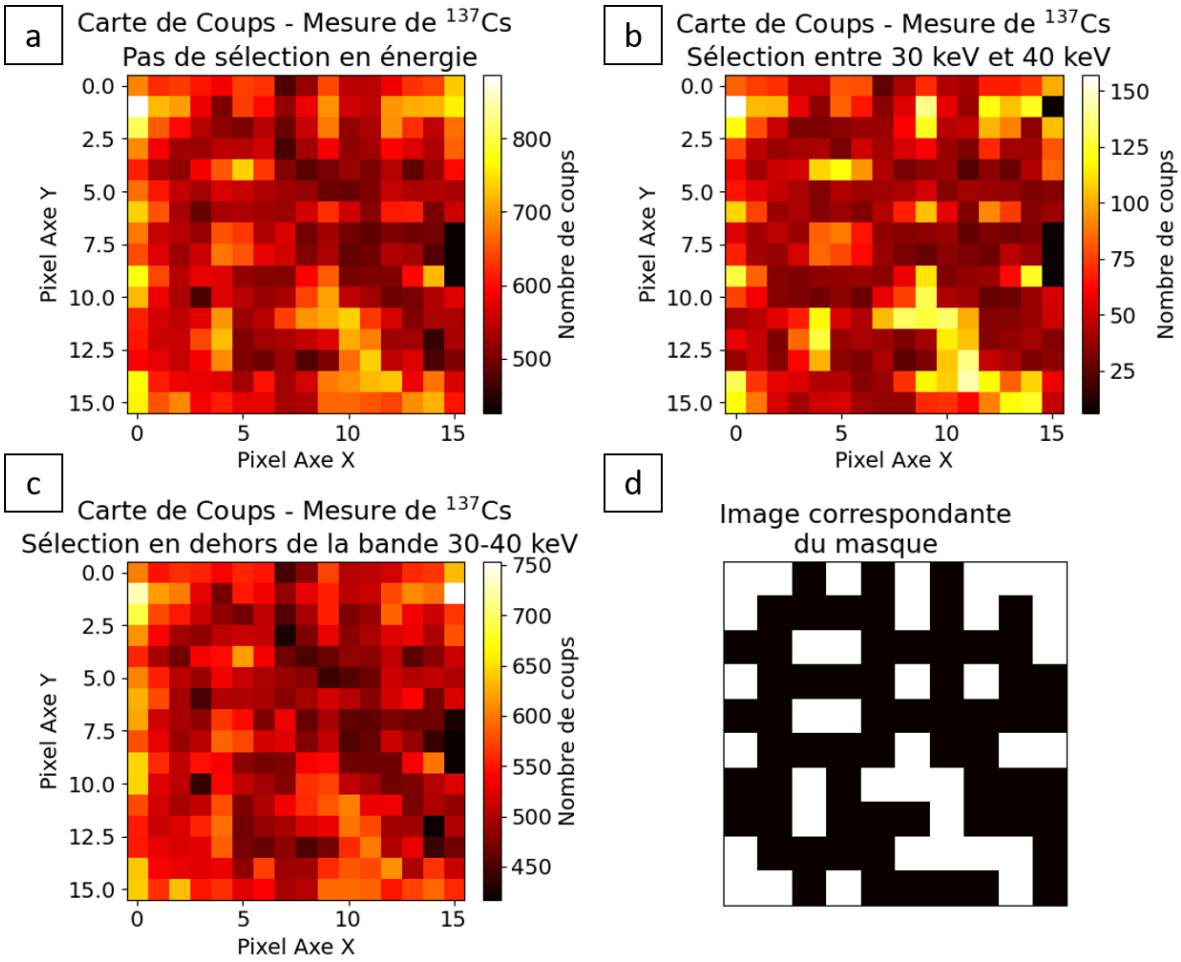


Figure 1-13 : a : Carte des coups d'une mesure d'une source de ^{137}Cs sans sélection en énergie. b : Sélection en énergie entre 30 et 40 keV. c : Sélection en dehors de la bande 30-40 keV. d : Motif du masque correspondant à la projection mesurée.
Acquisition d'une source de ^{137}Cs de 1,6 MBq placée à 1 m du détecteur pendant 18 h, masque codé placé devant le détecteur. Haute-tension appliquée au cristal : 300 V ; température du détecteur : - 10°C ; temps de *peaking* : 6,07 µs.

1.2.6 Calibration non affine

Les électroniques de lecture IDeF-X HD ont été conçues pour fournir une réponse la plus linéaire possible par rapport à l'amplitude du signal détecté. Cependant, des petites non-linéarités peuvent apparaître sur le gain, de l'ordre de 1 à 2 %, et produire des décalages dans les spectres.

De même, les processus de pertes de charges dans le cristal, que nous étudierons plus en détail en paragraphe 3.2, ne dépendent pas linéairement de l'énergie et sont aussi à l'origine de non-linéarités dans la réponse du détecteur.

Par conséquent, la non-linéarité qui découle de ces deux effets a une influence sur les spectres mesurés. En utilisant les lignes d'émission du ^{137}Cs à 32 keV, 37 keV et 662 keV, la calibration donne de bons résultats autour de ces pics mais produisent des décalages allant jusqu'à 3,5 keV entre ces pics.

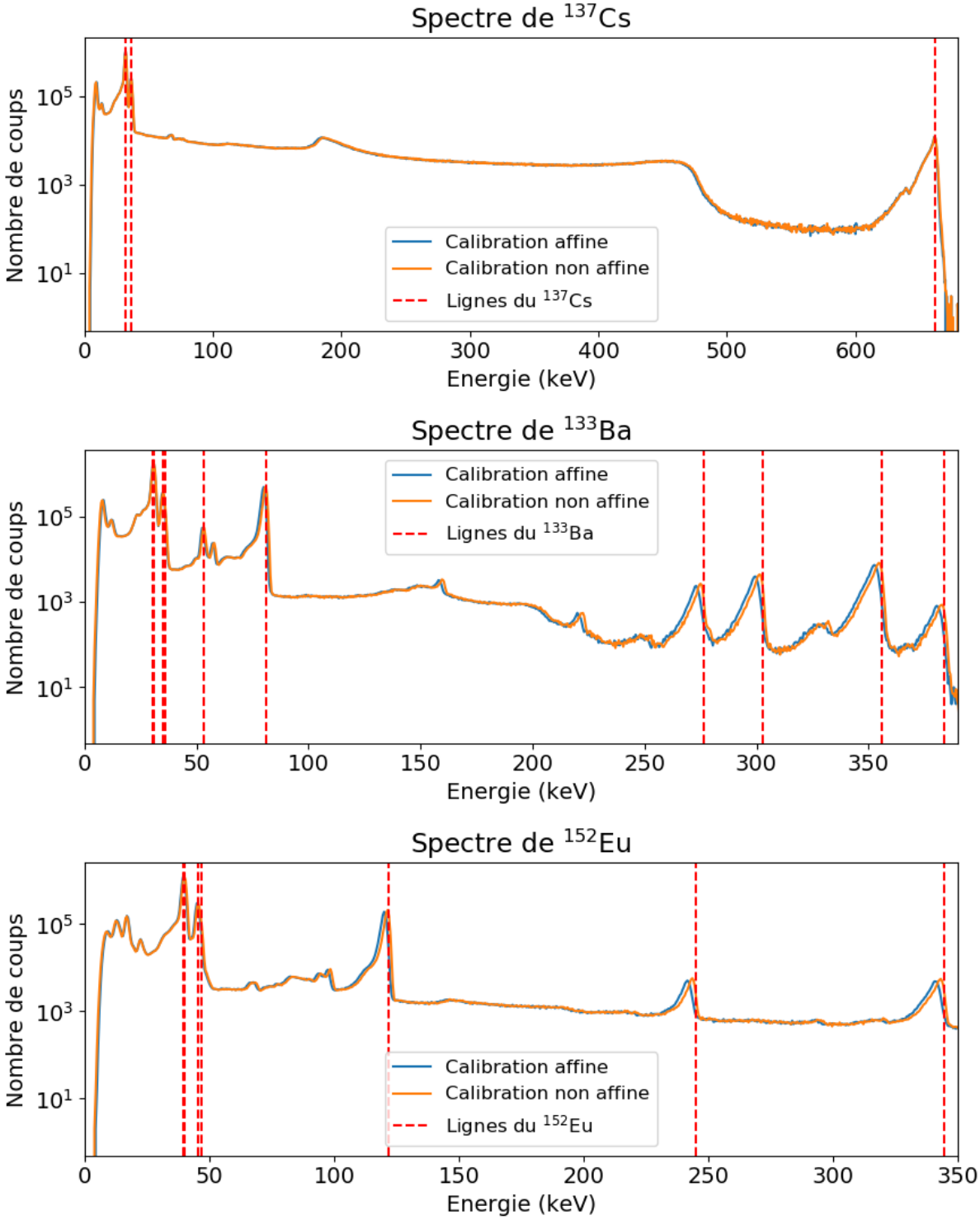


Figure 1-14 : Comparaisons calibration affine et calibration non affine. De haut en bas : acquisition d'une source de ^{137}Cs de 1,6 MBq placée à 1 cm du détecteur pendant 72 minutes, acquisition d'une source de ^{133}Ba de 130 kBq placée à 1 cm du détecteur pendant 90 minutes, acquisition d'une source de ^{152}Eu de 66 kBq placée à 1 cm du détecteur pendant 144 minutes. Haute-tension appliquée au cristal : 300 V ; température du détecteur : -10°C ; temps de *peaking* : 6,07 µs.

On peut illustrer ce phénomène en effectuant la mesure pour une source de ^{133}Ba et de ^{152}Eu qui présentent des lignes d'émission entre 100 keV et 400 keV. La Figure 1-14 présente un spectre calibré

utilisant les paramètres de calibration déterminés précédemment sur la loi de calibration affine. Le décalage est clairement visible sur les lignes identifiées du ^{133}Ba et ^{152}Eu pour les lignes entre 100 keV et 400 keV alors que les lignes d'émission du ^{137}Cs sont correctement positionnées.

Pour résoudre ce problème, j'ajoute la distorsion d de l'équation (1-1) aux deux autres paramètres de calibration g et o. Afin de trouver ces trois paramètres de calibration, il est alors nécessaire d'utiliser des radioéléments possédant des lignes dans la zone d'intérêt, et j'utilise ici les deux radioéléments ^{133}Ba et ^{152}Eu en plus du ^{137}Cs. J'utilise toujours le score de corrélation défini en (1-3) pour évaluer la qualité des paramètres de calibration et j'applique mon algorithme génétique, avec trois paramètres cette fois, pour trouver les paramètres optimaux. Les résultats de la cette calibration sont représentés en Figure 1-14.

On constate que les pics photoélectriques du spectre du ^{137}Cs étaient déjà bien positionnés avec la calibration affine, et restent toujours à la même position avec la calibration non affine. Pour les sources de ^{133}Ba et ^{152}Eu, les pics photoélectriques à basse énergie, inférieurs à 60 keV, sont aussi bien positionnés avec la calibration affine qu'avec la calibration non affine. En revanche, au-delà de 60 keV, la calibration non affine corrige en partie les erreurs sur le positionnement des pics. À titre d'exemple, le pic à 276,4 keV du ^{133}Ba est positionné à 273 keV avec la calibration linéaire et à 275 keV avec la calibration non affine.

On constate que la loi non affine améliore la calibration mais ne corrige pas entièrement la non-linéarité. Idéalement, il serait nécessaire d'introduire d'autres paramètres dans la loi de calibration, au risque cependant de complexifier encore l'optimisation de la fonction de corrélation. Cela est possible avec l'algorithme génétique et plus difficile avec l'algorithme ECC-AMR, du fait de la définition des zones voisines en dimension plus grande que 3.

1.3 Conclusion et nécessité de construire un modèle

J'ai introduit dans cette partie les données que j'utilise ensuite pour répondre à la problématique de spectro-identification de radioéléments et celle de leur localisation. Un premier traitement de calibration en énergie est nécessaire avant de construire des méthodes pour répondre à ces problèmes.

À travers les spectres calibrés en Figure 1-12 et Figure 1-14, nous pouvons expliquer la présence de certaines structures à l'aide des connaissances sur la physique des interactions photon/matière et les lignes d'émission des radioéléments. Cela illustre le fait que les signatures spectrales sont les informations fondamentales pour répondre au problème de spectro-identification des radioéléments.

Cependant, une compréhension plus détaillée de la forme de ces structures est essentielle pour pouvoir mettre en œuvre des méthodes avancées de spectro-identification. Cette compréhension passe par une modélisation fine de l'environnement et des interactions des photons dans le détecteur et par la modélisation de la réponse du détecteur que je détaille dans les deux prochaines parties.

2 Modélisation des interactions photons/matière, de la source au détecteur

Nos connaissances physiques nous permettent de comprendre les facteurs qui influencent les structures spectrales que nous enregistrons, à travers deux principaux types de structures. D'une part, les pics présents dans les spectres sont issus de photons atteignant le détecteur avec l'intégralité de leur énergie et interagissant dans le détecteur par effet photoélectrique en déposant l'ensemble de leur énergie, dans un ou plusieurs pixels. Ces photons n'ont donc pas subi d'interaction ayant conduit à un dépôt d'énergie dans l'environnement[*]. D'autre part, les fonds continus sont principalement dus à deux phénomènes : d'un côté, ils sont formés par des photons qui ont interagi dans l'environnement du détecteur en y déposant de l'énergie par diffusion Compton puis ont interagi dans le détecteur par effet photoélectrique. D'un autre côté, ils sont aussi dus à des photons qui ont interagi par diffusion Compton dans le détecteur en y déposant une partie de leur énergie, tout en ayant potentiellement interagi dans l'environnement au préalable. Des combinaisons plus complexes d'interactions existent.

Ces observations nous montrent qu'une simulation numérique est nécessaire pour modéliser le trajet des photons entre la source et le détecteur ainsi que les dépôts d'énergie des photons lors d'éventuelles interactions. Je présente ici les outils numériques que j'ai utilisés pour effectuer cette simulation. Dans un second temps, je porte une attention particulière à la modélisation des structures Compton, qui dépendent fortement des conditions de la mesure et de la présence de matériaux diffusants dans l'environnement, comme la structure mécanique de la caméra ou les matériaux présents dans la scène de mesure.

Dans cette partie, je distingue trois zones :

- le détecteur, constitué de la surface sensible du module Caliste, soit le cristal CdTe, dans lequel je relève les interactions des photons ;
- l'environnement proche ou direct du détecteur, constitué du corps du Caliste, de la structure de la tête de caméra et du socle de la tête de caméra. Cette structure est présente quelle que soit la scène de mesure, vu qu'elle contient le détecteur. Je la modélise en paragraphe 2.1 ;
- l'environnement éloigné, propre à la scène de détection, constitué de l'ensemble des éléments présents dans la scène, en dehors de la caméra gamma. Cela peut être les murs de la pièce, le blindage de la source… Cet environnement spécifique à la scène ne peut pas être modélisé de manière générale. Je le prends en compte en paragraphe 2.2 à travers le calcul analytique de profils de spectre suite à des interactions Compton dans l'environnement.

2.1 Simulation directe Monte-Carlo

La complexité de la géométrie et de la nature des matériaux constituant l'environnement direct du détecteur et de la source empêche de calculer des spectres uniquement de manière analytique. Une part de simulation numérique est nécessaire, notamment à travers des méthodes Monte-Carlo. Pour ce faire, j'ai utilisé le logiciel de simulation Geant4 (Agostinelli, et al., 2003), devenu une référence dans le domaine de la simulation Monte-Carlo en physique des particules.

[*] Cependant, ils peuvent interagir sans dépôt d'énergie par diffusion de Rayleigh.

Dans ce logiciel, la simulation s'effectue en générant des photons à partir d'une source, dans une direction tirée aléatoirement, soit uniformément sur tout l'espace soit suivant une loi de distribution définie, pour représenter une source avec un collimateur par exemple.

L'espace est discrétisé et la simulation suit la trajectoire du photon sur chaque pas d'espace. Si la trajectoire du photon traverse un matériau, le logiciel calcule à chaque pas la probabilité que le photon interagisse dans le matériau suivant son énergie et la nature du matériau, régie par les lois décrites en Annexe A. Il applique alors cette interaction aléatoirement suivant cette probabilité.

La liste de physique utilisée dans ma simulation est la liste Shielding-EMZ (Geant4 Collaboration (1), 2018), dans laquelle les phénomènes électromagnétiques et les interactions photons/matière sont modélisés le plus précisément. Les effets photoélectriques et diffusions cohérentes sont modélisés à travers le modèle de Livermore (Geant4 Collaboration (2), 2018), la diffusion Compton à travers le modèle de l'Université de Monash (Geant4 Collaboration (3), 2018), prenant en compte notamment les effets liés à l'élargissement Doppler, dont je détaille l'origine en paragraphe 2.2.

La géométrie et l'environnement modélisés correspondent à l'environnement direct du détecteur : le cristal en CdTe du détecteur, le corps du Caliste en résine époxy et la structure de la tête de caméra WIX en Aluminium. Le modèle reste simplifié, les matériaux constituant les ASICs, câbles, modules Peltier et circuits électroniques ne sont pas modélisés, du fait du ralentissement du calcul qu'ils induisent.

Les sources que je modélise sont des sources mono-énergétiques, allant de 0,5 keV à 2 MeV, par pas de 0,5 keV. Cela me permet d'obtenir la réponse impulsionnelle du système pour chacune de ces énergies. Par rapport au fait de simuler des sources issues de véritables radioéléments avec l'ensemble de leurs lignes d'émission, cette approche présente l'intérêt qu'il n'est pas nécessaire de procéder à de nouvelles simulations pour ajouter des radioéléments à la base de données ainsi constituée. Il suffit de considérer les lignes du nouveau radioélément et d'utiliser les simulations mono-énergétiques correspondantes, en prenant en compte les intensités relatives de chaque ligne d'émission.

Pour chaque source mono-énergétique, un million de photons sont simulés à partir de la source, dans un cône englobant la tête de la caméra WIX, pour éviter de simuler des photons qui n'interagiront pas dans le détecteur de façon certaine car ils ne sont pas orientés vers le détecteur. La Figure 2-1 montre la fraction des photons qui ont interagi dans le détecteur parmi le million de photons simulés.

On voit dans un premier temps que la fraction des photons détectés augmente avec l'énergie jusqu'à 60 keV environ. Ce phénomène provient du fait que la fenêtre de 500 µm d'épaisseur en Aluminium de la tête de détection absorbe de moins en moins les photons lorsque leur énergie augmente. À 26,7 keV et 31,8 keV, on voit deux structures sur la courbe. Ces énergies correspondent respectivement au *K-edge* de Cd et Te qui entraînent une diminution de la section efficace d'interaction avec le cristal. Dans un second temps, le nombre de photons détectés diminue avec l'énergie du fait de la diminution de l'efficacité de détection du CdTe avec l'énergie.

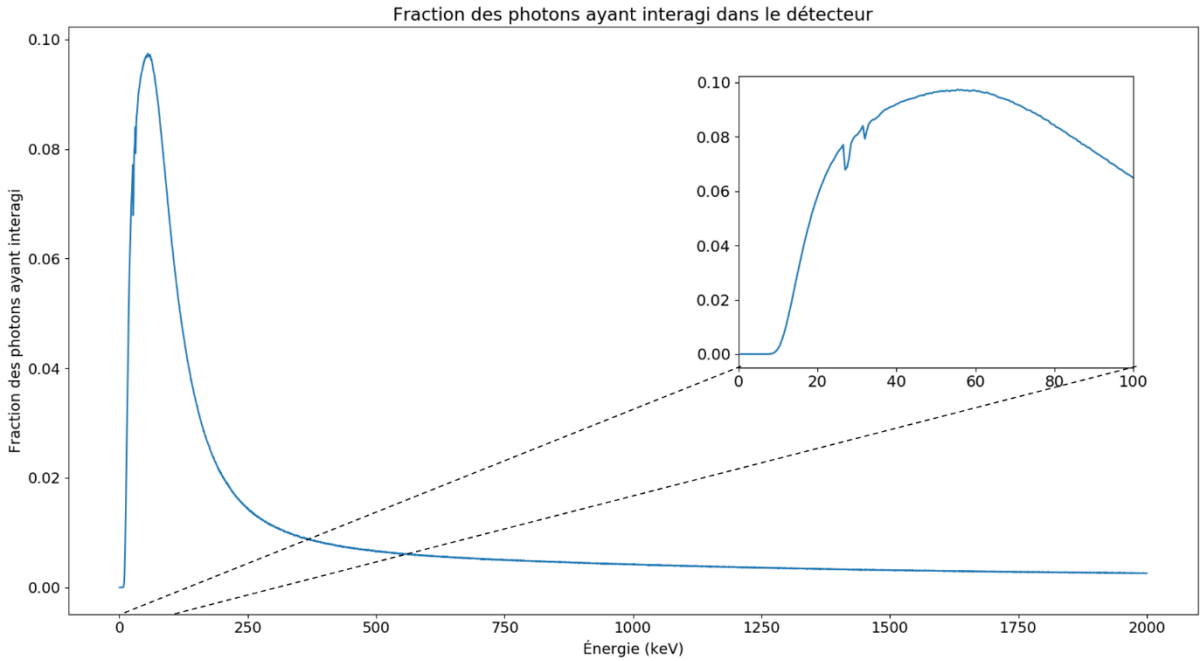


Figure 2-1 : Fraction des photons ayant interagi dans le détecteur en fonction de leur énergie d'émission.

La Figure 2-2 présente un exemple imagé de simulation. Sur cette figure, la structure en bleu est le support mécanique de la caméra, en Aluminium, la partie en rouge représente le cristal en CdTe et la partie grise transparente est le capot de la caméra, aussi appelé « cloche ». La longueur totale de la structure est de 10 cm. Les différentes lignes représentent les trajectoires des photons simulés.

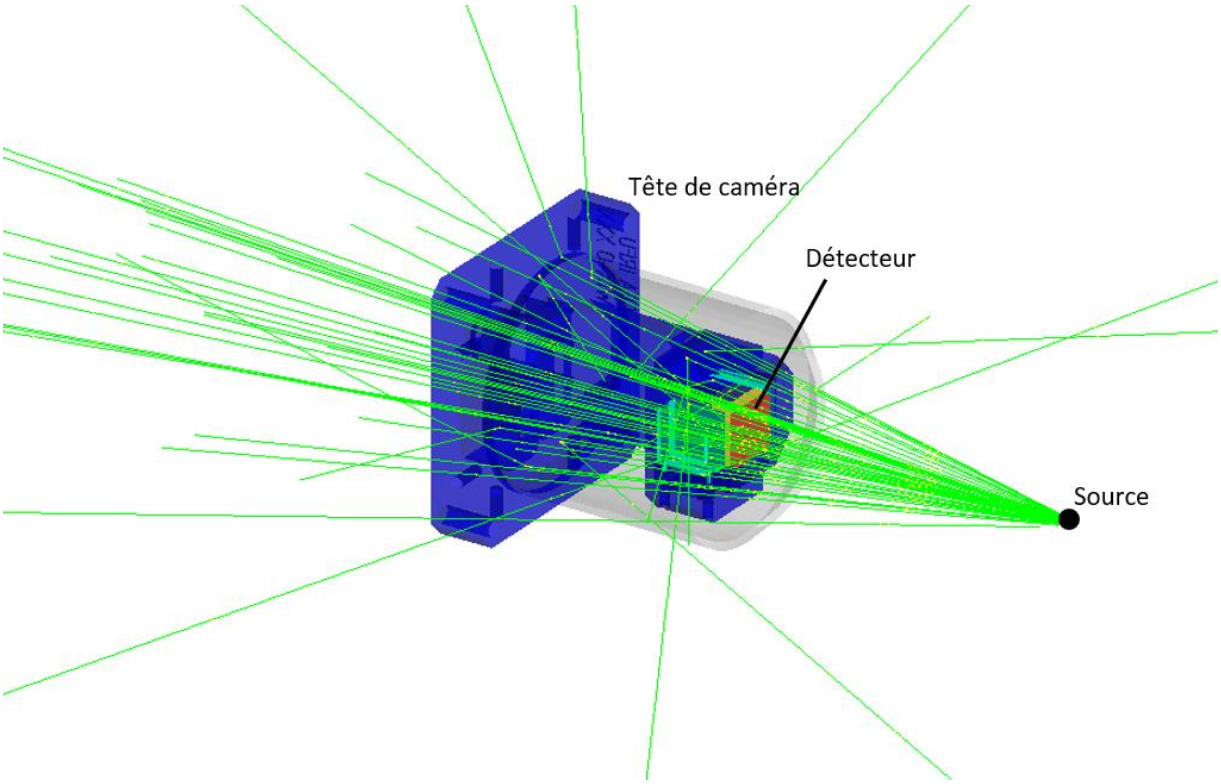


Figure 2-2 : Exemple de simulation Geant4. Les traits verts correspondent aux trajets des photons.

La simulation enregistre l'ensemble des photons qui interagissent dans le cristal du détecteur. Un des avantages majeurs de cette simulation est qu'elle prend en compte l'ensemble de la cascade d'interactions qui a lieu dans le détecteur. En effet, lorsqu'un photon interagit dans le détecteur, il produit un photoélectron qui effectue ensuite une trajectoire plus ou moins longue dans le cristal, dépendant de l'énergie que lui a transmise le photon. Le photoélectron cède son énergie à d'autres électrons en ionisant la matière le long de sa trajectoire. Au long de ce processus, les électrons qui ont reçu de l'énergie peuvent parfois émettre des photons de fluorescence. Ces photons peuvent ensuite éventuellement s'échapper du cristal ou interagir de nouveau dans le cristal.

La création de paires électrons-trous peut avoir lieu à diverses positions du cristal pour une même interaction, avec des écarts moyens attendus allant au-delà de 500 µm à 662 keV par exemple (Dirks, 2006). Dans la liste des événements enregistrés par Geant4, pour chaque photon, une liste des différents dépôts d'énergie est disponible, relevant l'énergie du dépôt et sa position. L'information en position se révélera essentielle pour appliquer en détail la réponse du détecteur que je présente en partie 3. La taille minimale de la cellule de calcul de production de charges ou photons secondaires a été fixée à 1 µm pour ces simulations, au regard de la précision spatiale de 10 µm que j'utilise dans ma modélisation de la réponse du détecteur en paragraphe 3.2.4. En effet, 10 µm correspond à la taille typique d'un nuage de charge créé par un photon de 10 keV, que l'on considère comme ponctuel au regard de la plus petite structure dans le cristal, qui est l'inter-pixel de 100 µm.

La Figure 2-3 donne un exemple de spectres obtenus par simulation à 32 keV et 662 keV. Ici, je fais seulement la somme des dépôts d'énergie, tels qu'ils sont enregistrés dans la simulation, ce qui correspond au dépôt total d'énergie de l'interaction, sans prendre en compte la résolution spectrale du détecteur.

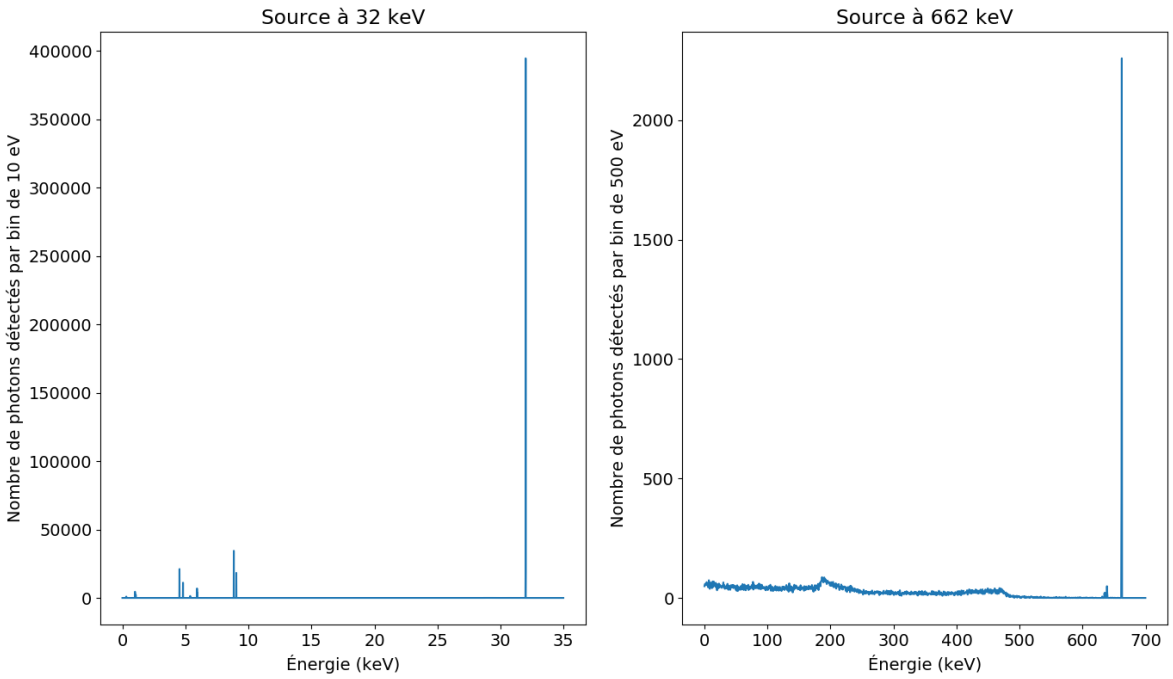


Figure 2-3 : Résultat des simulations de Geant4 à 32 keV et 662 keV.

La quasi-totalité des énergies déposées pour l'émission à 32 keV se retrouve à 32 keV, ce qui montre que les dépôts sont effectués par effet photoélectrique uniquement. À basse énergie, les pics

d'échappement de Cd et Te à (32 – 23) et (32 – 27) keV principalement, sont visibles et sont générés à partir des photons à 32 keV dont l'énergie est suffisante pour exciter la fluorescence X des couches K du Cd et du Te.

À 662 keV, on voit que le pic photoélectrique est très marqué lorsque la réponse du détecteur n'est pas appliquée. Le fond continu Compton est visible aussi. Il correspond d'une part à des diffusions dans l'environnement direct du détecteur, comme en témoigne le pic caractéristique de la rétrodiffusion (à 180°) à 184 keV. Ces diffusions ont lieu probablement dans le module Caliste lui-même. Le fond continu Compton correspond d'autre part à des diffusions dans le cristal CdTe, comme en témoigne le front Compton à 478 keV. On voit aussi distinctement les pics d'échappement à 662-23 et 662-27 keV.

L'étape suivante sur cette simulation est la prise en compte de la réponse du détecteur, dont je développe la réponse en partie 3. Cependant, avant cela, toujours dans le cadre des interactions entre les photons et l'environnement du détecteur, je m'intéresse à la possibilité de diffusions Compton multiples dans l'environnement éloigné du détecteur et leur modélisation.

2.2 Modélisation des diffusions Compton multiples

La simulation Geant4 prend en compte l'environnement direct du détecteur, à savoir la structure de la tête de détection et le corps du détecteur. Cet environnement direct a déjà un impact à travers le fond Compton visible en Figure 2-3.

Dans la réalité, l'ensemble de l'environnement est susceptible de provoquer des diffusions Compton supplémentaires, comme par exemple, dans les murs, le sol ou divers matériaux autour de la caméra et de la source. La simulation Geant4 que je réalise ne prend pas en compte toutes ces caractéristiques qui sont naturellement spécifiques à chaque scène à étudier.

Une modélisation de ces diffusions Compton multiples est cependant nécessaire pour que la simulation explique en détail les spectres observés. En utilisant les connaissances des processus physiques d'interaction Compton, j'ai calculé analytiquement la forme attendue des distributions spectrales des diffusions Compton multiples pour chaque énergie.

2.2.1 Diffusion Compton simple

Dans un premier temps, je calcule la distribution en énergie attendue pour un photon diffusé suite à une seule interaction Compton. Le calcul de la section efficace doublement différentielle en fonction de l'angle de diffusion et de l'énergie du photon diffusé doit prendre en compte le fait que l'électron sur lequel est effectué l'interaction est en mouvement et lié à l'atome.

Ce phénomène est appelé élargissement Doppler et la méthode pour le prendre en compte est décrite par (Ordonez, et al., 1997). On note E_0, l'énergie du photon incident avant la diffusion et $E_{\text{repos}}(\theta)$, l'énergie du photon diffusé avec un angle θ en considérant l'électron initialement au repos.

Comme expliqué en Annexe A, cette énergie vaut :

$$E_{\text{repos}}(\theta) = \frac{E_0}{1 + \frac{E_0}{m_e c^2}(1 - \cos(\theta))} \tag{2-1}$$

Or, dans le cas où l'électron n'est pas au repos, le lien entre l'impulsion initiale de l'électron et l'énergie du photon diffusé E_1 est donné par :

$$p_z = -m_e c \frac{\left(E_0 - E_1 - \frac{E_0 E_1 (1-\cos(\theta))}{m_e c^2}\right)}{\sqrt{E_0^2 + E_1^2 - 2E_0 E_1 \cos(\theta)}} \tag{2-2}$$

où p_z est la projection de l'impulsion initiale de l'électron sur l'impulsion de l'électron après la collision. Les équations (2-1) et (2-2) sont équivalentes dans le cas où $p_z = 0$.

Pour un électron de la n-ième sous-couche de l'atome auquel il appartient, la section efficace doublement différentielle selon l'énergie de diffusion E_1 et l'angle de diffusion θ est donnée par :

$$\left(\frac{d^2\sigma_{E_0}}{d\theta dE_1}\right)_n = \frac{m r_0^2}{2E_0}\left(\frac{E_{\text{repos}}(\theta)}{E_0} + \frac{E_0}{E_{\text{repos}}(\theta)} - \sin^2(\theta)\right)\frac{E_1}{\sqrt{E_0^2 + E_1^2 - 2E_0 E_1 \cos(\theta)}} J_n(p_z) 2\pi \sin(\theta) \tag{2-3}$$

Où r_0 est le rayon classique de l'électron et $J_n(p_z)$ est le profil Compton pour la n-ième sous-couche de l'atome de diffusion, donné par (Biggs, et al., 1975).

La formule (2-3) n'est valable que dans le cas où l'énergie déposée par le photon est suffisante pour arracher l'électron de l'atome. Cela signifie que $E_1 \leq E_0 - I_n$ où I_n est l'énergie de liaison de l'électron située à la n-ième sous-couche, sinon l'interaction Compton ne peut pas avoir lieu. La section efficace différentielle totale est donc :

$$\frac{d^2\sigma_{E_0}}{d\theta dE_1} = \sum_{n, E_1 \leq E_0 - I_n} \left(\frac{d^2\sigma_{E_0}}{d\theta dE_1}\right)_n \tag{2-4}$$

La section efficace doublement différentielle en fonction de l'énergie est finalement donnée par :

$$\frac{d\sigma_{E_0}}{dE_1} = \int_0^\pi \frac{d^2\sigma_{E_0}}{d\theta dE_1} d\theta \tag{2-5}$$

La méthode de calcul est la suivante. Pour une énergie E_1 donnée, on calcule l'intégrale (2-5) avec l'équation (2-4) dans laquelle $E_{\text{repos}}(\theta)$ et p_z sont données par les équations (2-1) et (2-2). La densité de probabilité, notée $f^Z_{E_0,\text{diff}}(E_1)$, d'obtenir un photon diffusé d'énergie E_1 est donnée par la normalisation suivante :

$$f^Z_{E_0,\text{diff}}(E_1) = \frac{\frac{d\sigma_{E_0}}{dE_1}(E_1)}{\int \frac{d\sigma_{E_0}}{dE_1}} \tag{2-6}$$

La Figure 2-4 montre cette densité de probabilité pour un photon incident d'énergie $E_0 = 662\ \text{keV}$ et pour une diffusion sur un atome d'étain* Z = 50, dont le numéro atomique est proche du numéro atomique moyen de CdTe.

* L'étain est un matériau modèle du CdTe du point de vue des interactions photons/matière, d'une part du fait de son numéro atomique (50 pour l'étain contre 49,1 comme numéro atomique moyen du CdTe) et d'autre part, du fait de sa densité (5,77 pour l'étain contre 5,85 pour le CdTe). Lorsque les données ne sont pas disponibles pour le CdTe ou les composés de manière générale, on utilise l'étain comme matériau modèle, comme ici dans les données de (Biggs, et al., 1975).

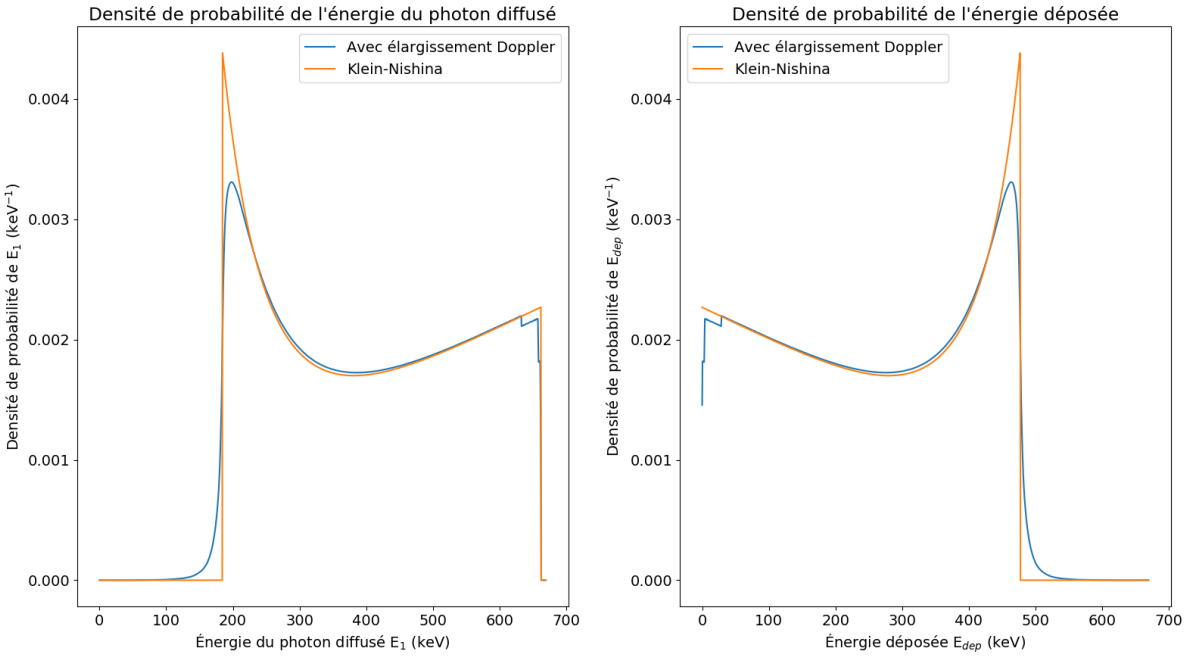


Figure 2-4 : Densités de probabilité de l'énergie du photon diffusé et de l'énergie déposée par interaction Compton, l'énergie initiale du photon est de 662 keV.

À droite, je représente la densité de probabilité $f_{E_0,\mathrm{dep}}^{Z=50}(E_{\mathrm{dep}})$ où $E_{\mathrm{dep}} = E_0 - E_1$ est l'énergie déposée lors de la diffusion Compton. À titre de comparaison, j'ai représenté aussi la densité de probabilité attendue en utilisant la section efficace différentielle de Klein-Nishina (voir Annexe A), sans prise en compte du phénomène d'élargissement Doppler.

Les deux distributions sont symétriques par rapport à $\frac{E_0}{2}$ du fait de la conservation de l'énergie totale lors de l'interaction. On peut noter ici que les énergies inférieures à 184 keV sont interdites pour le photon diffusé dans le cas utilisant la formule de Klein-Nishina. En prenant en compte l'élargissement Doppler, ces énergies sont autorisées, ce qui est une caractéristique du fait que l'électron cible possède une impulsion initiale et n'est pas au repos. On constate aussi que la formule (2-3) que j'utilise avec l'élargissement Doppler prend en compte les énergies de liaison dans les différentes sous-couches, ce qui se traduit par des discontinuités dans la densité de probabilité entre 600 keV et 662 keV.

2.2.2 Diffusions Compton multiples

À partir de la modélisation des diffusions Compton simples, je peux en déduire analytiquement une expression des densités de probabilité pour des diffusions Compton multiples dans l'environnement, moyennant certaines hypothèses simplificatrices.

Un photon d'énergie initiale E_0 peut enchaîner plusieurs diffusions Compton dans l'environnement. On notera E_i l'énergie du photon diffusé après la i-ème diffusion Compton. Notons aussi $p_i(E_i)$ la densité de probabilité sur cette énergie E_i.

En utilisant les lois sur les probabilités conditionnelles, je peux en déduire récursivement le lien entre les densités de probabilités $p_i(E_i)$ et $p_{i-1}(E_{i-1})$. Si on note $p(E_i|E_{i-1})$, la densité de probabilité d'obtenir un photon diffusé avec une énergie E_i à partir d'un photon initial d'énergie E_{i-1}, alors :

$$p_i(E_i) = \int_{\mathbb{R}^+} p(E_i|E_{i-1}) p_{i-1}(E_{i-1}) dE_{i-1} \tag{2-7}$$

Le terme $p(E_i|E_{i-1})$ doit prendre en compte trois effets :

- la densité de probabilité qu'un photon incident d'énergie E_{i-1} conduise à un photon diffusé d'énergie E_i. Cette densité de probabilité est en fait donnée par la diffusion Compton simple en (2-6) avec $f^Z_{E_{i-1},\mathrm{diff}}(E_i)$, où Z est le numéro atomique du matériau de diffusion ;
- la probabilité $P_{\mathrm{diff}}(E_{i-1})$ qu'un photon d'énergie E_{i-1} interagisse dans le matériau de longueur Z, de coefficient d'atténuation linéique total $\mu(Z, E_{i-1})$ et Compton $\mu_C(Z, E_{i-1})$ à l'énergie E_{i-1}, d'épaisseur L, donnée par :

$$P_{\mathrm{diff}}(E_{i-1}) = \frac{\mu_C(Z, E_{i-1})}{\mu(Z, E_{i-1})} \times \left(1 - e^{\mu(Z,E_{i-1})\times L}\right) \tag{2-8}$$

- la densité de probabilité qu'un photon d'énergie E_{i-1} atteigne le matériau de diffusion. Elle dépend en fait de la direction d'émission du photon source d'énergie E_0 et de la configuration de l'environnement. Il n'est donc pas possible de calculer cette densité de probabilité analytiquement, et afin de généraliser le problème, sans connaissance a priori de la scène, on doit faire l'hypothèse que ce terme est uniforme. Cette hypothèse est raisonnable du fait que la source est considérée comme émettant de manière isotrope et que les interactions Compton multiples dans l'environnement impliquent une diffusion des photons que l'on peut considérer comme isotrope[*]. Ce terme de densité de probabilité étant considéré comme uniforme et indépendant de l'énergie E_i, il n'a pas d'impact sur le calcul de (2-7) et donc je ne le représente pas dans $p(E_i|E_{i-1})$.

J'obtiens alors :

$$p(E_i|E_{i-1}) = f^Z_{E_{i-1},\mathrm{diff}}(E_i) \times \frac{\mu_C(Z, E_{i-1})}{\mu(Z, E_{i-1})} \times \left(1 - e^{\mu(Z,E_{i-1})\times L}\right) \tag{2-9}$$

On peut introduire (2-9) dans (2-8) pour obtenir le calcul récursif de $p_i(E_i)$. Pour $i = 0$, je prends $p_0(E_0) = \delta_{E_0}$, étant donné que le photon source est considéré d'énergie connue E_0. De cette manière, j'obtiens directement $p_1(E_1)$ par l'équation (2-10) : $p_1(E_1) = f^Z_{E_0,\mathrm{diff}}(E_1)$.

L'étape précédente donne la distribution en énergie pour les photons diffusés. Maintenant, pour connaître la distribution en énergie des photons effectivement détectés, il faut prendre en compte l'interaction dans le détecteur qui peut être soit une interaction photoélectrique, soit une interaction Compton.

Notons $E_{i,\mathrm{PEdep}}$ l'énergie finale déposée dans le détecteur par interaction photoélectrique par un photon suite à i diffusions Compton dans l'environnement. Dans ce cas, il faut que le photon incident ait une énergie $E_{i,\mathrm{PEdep}}$ et interagisse par effet photoélectrique avec une probabilité $P_{\mathrm{PE}}\left(E_{i,\mathrm{PEdep}}\right)$, ce qui donne donc la densité de probabilité $p\left(E_{i,\mathrm{PEdep}}\right)$:

[*] Notons que cette hypothèse n'est pas valable pour des géométries particulières pour lesquelles on aurait par exemple placé un collimateur devant la source, qui sélectionne les directions des photons.

$$p(E_{i,\mathrm{PEdep}}) = P_{\mathrm{PE}}(E_{i,\mathrm{PEdep}}) \times p_i(E_{i,\mathrm{PEdep}}) \tag{2-10}$$

Avec :

$$P_{\mathrm{PE}}(E_{i,\mathrm{PEdep}}) = \frac{\mu(\mathrm{CdTe}, E_{i,\mathrm{PEdep}})}{\mu_{\mathrm{PE}}(\mathrm{CdTe}, E_{i,\mathrm{PEdep}})} \times \left(1 - e^{\mu(\mathrm{CdTe}, E_{i,\mathrm{PEdep}}) \times d}\right) \tag{2-11}$$

où d est l'épaisseur de détecteur traversée par le photon.

De même, on note $E_{i,\mathrm{Cdep}}$ l'énergie finale déposée dans le détecteur par interaction Compton suite à i diffusions Compton dans l'environnement. Dans ce cas, en utilisant les lois de probabilités conditionnelles de manière analogue à (2-11) :

$$p(E_{i,\mathrm{Cdep}}) = \int_{\mathbb{R}^+} p_{\mathrm{dep}}(E_{i,\mathrm{Cdep}} | E_i) p_i(E_i) dE_i \tag{2-12}$$

$p_{\mathrm{dep}}(E_{i,\mathrm{Cdep}} | E_i)$ est la densité de probabilité de déposer l'énergie $E_{i,\mathrm{Cdep}}$ sur le détecteur à partir d'un photon incident d'énergie E_i. De même que pour obtenir l'équation (2-9), cette densité de probabilité se décompose selon :

- la probabilité que le photon incident d'énergie E_i effectue une diffusion Compton dans le détecteur ;
- le profil Compton $f_{E_i,\mathrm{dep}}^{Z=50}(E_{i,\mathrm{Cdep}})$ de l'énergie déposée dans le détecteur calculé au paragraphe 2.2.1 sur la diffusion Compton simple.

De manière analogue à (2-9), j'obtiens :

$$p(E_i | E_{i-1}) = f_{E_i,\mathrm{dep}}^{Z=50}(E_{i,\mathrm{Cdep}}) \times \frac{\mu_C(\mathrm{CdTe}, E_{i,\mathrm{Cdep}})}{\mu(\mathrm{CdTe}, E_{i,\mathrm{Cdep}})} \times \left(1 - e^{\mu(\mathrm{CdTe}, E_{i,\mathrm{Cdep}}) \times d}\right) \tag{2-13}$$

J'ai donc calculé les densités de probabilités (2-10) et (2-12) avec la méthode décrite précédemment.

J'ai choisi des diffusions dans l'environnement avec des matériaux de Z = 13, correspondant à de l'Aluminium. L'environnement est composé de matériaux différents, mais il n'est pas possible d'englober toutes les configurations possibles, d'où ce choix. De même, j'ai choisi une longueur moyenne $L = 1\ \mathrm{mm}$ de matériau, qui correspond environ à l'épaisseur moyenne d'aluminium autour du détecteur. J'ai fait ce choix sachant que cette longueur pour l'ensemble de l'environnement est dépendante de la scène de la mesure. Notons aussi que cette méthode ne traite pas des cas où le photon diffusé est réabsorbé dans le matériau s'il est trop épais, ce qui dépend de la longueur L. Enfin, j'ai choisi une épaisseur traversée $d = 1\ \mathrm{mm}$ de détecteur, qui correspond à l'épaisseur normale du cristal, mais cette épaisseur dépend aussi de la direction d'arrivée du photon incident. Par exemple, dans le cas extrême où le photon arrive sur la tranche du détecteur, l'épaisseur traversée est d'au moins 1 cm.

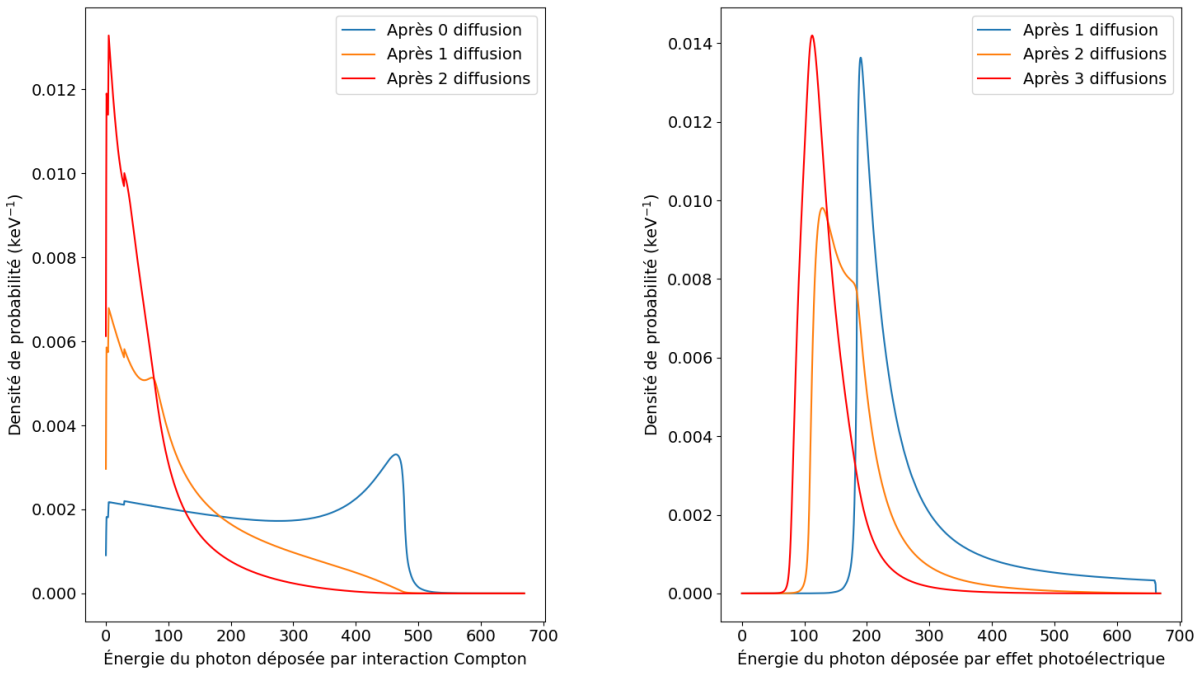


Figure 2-5 : Densités de probabilités des dépôts d'énergie dans le détecteur suite à des diffusions Compton multiples dans l'environnement. À gauche : énergie finalement déposée dans le détecteur par interaction Compton. À droite : énergie finalement déposée dans le détecteur par effet photoélectrique. Énergie du photon source de 662 keV.

Malgré ces hypothèses et approximations, les formes analytiques que je déduis des différents profils Compton sont suffisantes pour expliquer l'essentiel des spectres observés, comme je le montre en paragraphe 4.2.2 et en Figure 2-6. La configuration exacte de l'environnement a principalement une influence sur l'intensité de chaque composante dans le spectre. Par exemple, si l'environnement est clos avec des cloisons très proches du détecteur et de la source, comme dans un petit conteneur, les diffusions multiples seront favorisées par rapport à un environnement très ouvert.

La Figure 2-5 présente les résultats de ces profils suite à une, deux ou trois interactions Compton successives dans l'environnement, pour une énergie initiale $E_0 = 662\ \mathrm{keV}$. Je représente les densités de probabilité quand le dépôt d'énergie dans le détecteur s'effectue par une interaction photoélectrique et quand il s'effectue par une interaction Compton.

On constate bien l'effet de diminution des énergies au fur et à mesure que le nombre de diffusions Compton dans l'environnement augmente. J'utilise ces prévisions de profils Compton afin d'expliquer en détail certaines composantes du spectre en Figure 2-6.

On voit que les structures continues Compton sont expliquées par les différents profils de diffusions Compton multiples que j'ai calculées avec la méthode décrite dans ce paragraphe. Ce résultat a fait l'objet d'une publication dans la revue à comité de lecture NIM-A (Maier, et al., 2018).

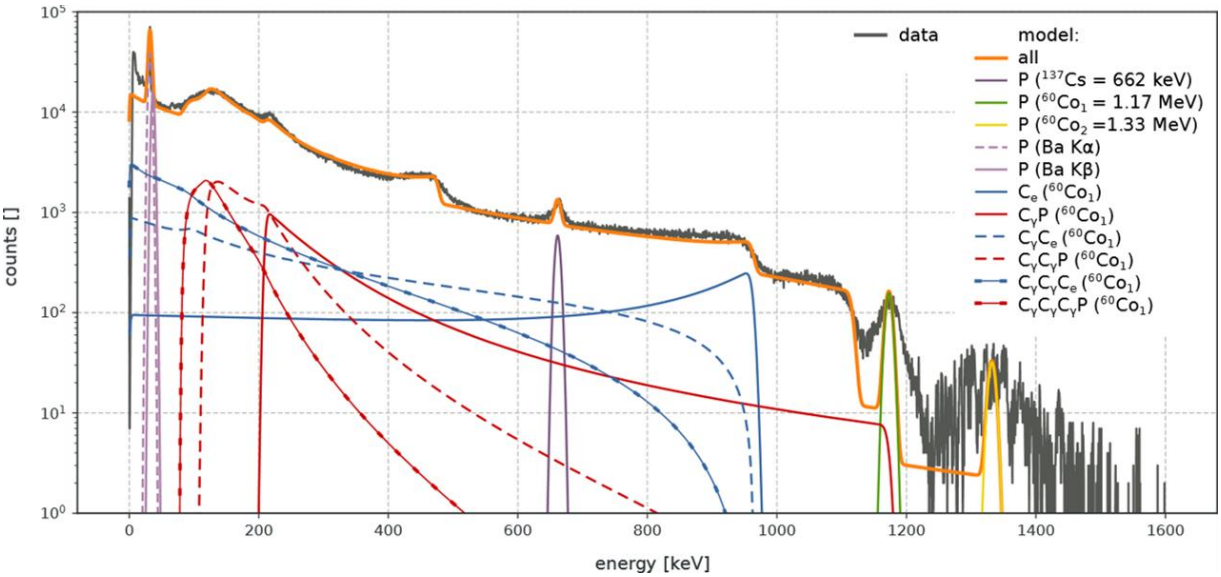


Figure 2-6 : Modèle analytique des diffusions Compton multiples en comparaison à un spectre de ^{137}Cs et ^{60}Co acquis avec Caliste-O. Les profils de diffusions Compton multiples sont représentés seulement pour la ligne à 1,17 MeV du ^{60}Co. En légende, P représente une absorption photoélectrique dans le détecteur, C_γ une diffusion Compton en dehors du détecteur et C_e une diffusion Compton dans le détecteur. Par exemple, C_γ C_e est le profil Compton des photons ayant fait une diffusion Compton dans l'environnement et un dépôt d'énergie par diffusion Compton dans le détecteur.

3 Modélisation de la réponse du détecteur

La connaissance des interactions des photons dans l'environnement et dans le détecteur est la première composante qui permet de comprendre les spectres et les éléments qui les structurent. Cependant, pour expliquer finement la forme des différentes composantes spectrales, il est nécessaire de modéliser précisément la réponse en énergie du détecteur. En effet, le spectre physique est convolué à la réponse instrumentale, définie comme la distribution des énergies effectivement mesurées E_m à partir d'une énergie déposée E_0.

Dans cette partie, je détaille les principales composantes d'un modèle de la réponse du détecteur : la fluctuation du nombre de paires électrons-trous créées, le transport et la perte de charge et le bruit électronique. On suppose tout au long de cette partie qu'un photon a déposé une énergie E_0 dans le détecteur, soit par effet photoélectrique, soit par interaction Compton. Je montre comment la mesure de l'énergie est impactée à chaque étape de la prise en compte de la réponse du détecteur.

3.1 Fluctuation du nombre de paires électrons-trous

En interagissant dans le détecteur, un photon cède suffisamment d'énergie à un électron d'un atome du cristal pour l'arracher de son atome. Au long de sa trajectoire, cet électron transmet une partie de son énergie pour arracher des électrons d'autres atomes. Certains de ces électrons vont éventuellement revenir dans leur état en émettant un photon de fluorescence X qui, lui-même, peut interagir de nouveau et recommencer ce processus.

Finalement, l'ensemble de l'interaction crée un nuage de charges, sous forme de paires électrons-trous. Le mouvement de ces charges, guidées par le champ électrique appliqué au détecteur, induit le

signal sur les anodes du cristal, où sont connectés les circuits de lecture. Le signal induit est proportionnel au nombre de charges créées.

Le nombre N_0 de paires électrons-trous créées dépend de l'énergie E_0 déposée et de l'énergie de création de paires w. Pour le CdTe, $w = 4{,}42$ eV (Meuris, 2009) par paire créée. Le lien entre E_0 et N_0 est donné par :

$$N_0 = \frac{E_0}{w} \tag{3-1}$$

Par exemple, pour un dépôt d'énergie $E_0 = 662\ \mathrm{keV}$, le nombre de paires électron-trous attendues est de 149 774.

N_0 est en réalité une valeur moyenne du nombre de paires créées, dont les fluctuations sont dues aux processus aléatoires dans la création de charges décrite au début de ce paragraphe.

Si ces processus sont purement statistiques, on s'attend à ce que le nombre de charges effectivement créées suive une loi de Poisson de moyenne N_0 et d'écart-type $\sqrt{N_0}$. Cependant, il a été démontré que les processus de création de charges ne sont pas purement indépendants (Fano, 1947). Cela est en partie dû au fait qu'il n'existe qu'un nombre limité de possibilités d'ionisation d'un atome, par la quantification des couches électroniques.

De ce fait, on peut toujours modéliser ces fluctuations par une loi de Poisson de moyenne N_0, mais on doit ajouter à la variance un facteur correctif adimensionné F, appelé facteur de Fano. L'écart-type de la loi de Poisson régissant les fluctuations de paires-électrons trous est alors $\sqrt{F.N_0}$. F dépend du matériau dans lequel a lieu l'interaction. Pour le CdTe, il vaut 0,15, ce qui signifie qu'il diminue in fine l'influence des fluctuations statistiques.

Le nombre de paires créées est grand aux énergies auxquelles je travaille. Le seuil bas de Caliste est au minimum de 2 keV, ce qui correspond à 453 paires de charge créées en moyenne. De ce fait, on peut faire l'approximation de la loi de Poisson par une loi normale de moyenne N_0 et d'écart-type $\sqrt{F.N_0}$.

À travers l'équation (3-1), on peut en déduire que l'énergie équivalente E_{Fano} au nombre de paires de charges créées va aussi subir des fluctuations, de moyenne $w.N_0 = E_0$ et d'écart-type $w.\sqrt{F.N_0} = \sqrt{F.w.E_0}$.

Ces fluctuations impliquent que la résolution que l'on peut atteindre avec ce matériau ne peut pas être meilleure que $2{,}355\sqrt{F.w.E_0}$ FWHM (*Full Width at Half Maximum*), indépendamment de toute l'optimisation qui pourra être effectuée sur le détecteur en matière de performances spectroscopiques. Cette optimisation concerne en particulier l'électronique de lecture et l'intégration de l'ensemble dans un module réel tel que Caliste. On parle alors de « résolution intrinsèque » du détecteur. Elle vaut par exemple 470 eV FWHM à 60 keV et 1,56 keV FWHM à 662 keV.

Pour prendre en compte ce phénomène dans la modélisation de la réponse du détecteur, à partir de E_0, l'énergie déposée par le photon, je considère que l'énergie directement vue par le détecteur est E_{Fano}, tirée suivant la loi suivante :

$$E_{\mathrm{Fano}} \sim \mathcal{N}\left(\mu = E_0, \sigma = \sqrt{F.w.E_0}\right) \tag{3-2}$$

3.2 Formation du signal

3.2.1 Induction du signal

Le signal vu par l'anode est dû au courant transitoire induit par le déplacement des porteurs photo-générés vers les électrodes. Sous l'effet d'un champ électrique appliqué au cristal semi-conducteur, le courant est la somme des composantes dues au mouvement des électrons s'approchant de l'anode et de celui des trous s'en éloignant*. Ce courant cesse lorsque les porteurs de charges ont atteint leurs électrodes respectives. La charge image qui s'établit à l'anode est une fonction du temps, de la géométrie des électrodes et du détecteur et par conséquent de la trajectoire des porteurs au voisinage des électrodes. Dans le cas de détecteurs pixélisés, la trajectoire peut être complexe et son calcul nécessite une résolution numérique.

Le théorème de Shockley-Ramo (Shockley, 1938) (Ramo, 1939) permet de calculer le courant électrique instantané induit par une charge se déplaçant au voisinage d'une électrode. Il est basé sur le concept que le courant induit dans l'électrode est dû au changement instantané des lignes de flux électrostatiques qui se terminent sur l'électrode, et non à la quantité de charge reçue par l'électrode par seconde.

L'expression calculant le courant induit sur l'anode par une charge Q située à la position $\vec{x}$ dans le détecteur, se déplaçant à une vitesse de dérive instantanée $\vec{v}$ est donnée par :

$$I_{\text{induit}} = Q.\overrightarrow{E_w}(\vec{x}).\vec{v}(\vec{x}) \tag{3-3}$$

$\overrightarrow{E_w}$ est le champ de pondération (*weighting field*) ; c'est un champ virtuel qui correspond à celui créé dans le cas où on fixe le potentiel de l'anode du pixel considéré à une valeur unitaire, et l'ensemble des autres plans de charge à 0. L'analyse de l'origine de ce champ de pondération est réalisée dans (He, 2001) et je détaille son calcul en paragraphe 3.2.4.

Si on considère un électron ou un trou de charge $q = \pm e$, e étant la charge élémentaire, de mobilité μ, évoluant dans un champ réel $\vec{E}$, l'expression de la vitesse de dérive s'exprime par :

$$\vec{v}(\vec{x}) = \frac{q}{e}\mu\vec{E}(\vec{x}) \tag{3-4}$$

Par ailleurs, la charge Q dans l'équation (3-3) est directement reliée au nombre de charges, à savoir $+e.N_h$ pour N_h trous et $-e.N_e$ pour N_e électrons. Ainsi, l'équation (3-3) se réécrit, en distinguant les trous (indice h) et les électrons (indice e) :

$$I_e = N_e e\mu_e\overrightarrow{E_w}(\overrightarrow{x_e}).\vec{E}(\overrightarrow{x_e}) \tag{3-5}$$

$$I_h = N_h e\mu_h\overrightarrow{E_w}(\overrightarrow{x_h}).\vec{E}(\overrightarrow{x_h}) \tag{3-6}$$

* De la même manière, un signal est induit sur la cathode, mais celle-ci n'est pas pixellisée et le signal n'est pas enregistré, d'où le fait que je ne considère pas son calcul ici. En revanche, l'enregistrement de ce signal à la cathode est une méthode qui peut permettre de déterminer la profondeur des interactions dans le cristal.

3.2.2 Transport et perte de charge

Les nombres de charges N_e et N_h évoluent en fonction du temps et dépendent de leur temps de parcours au sein du matériau. En raison des imperfections du cristal, il existe des états actifs dans la bande interdite du semi-conducteur susceptibles de capturer ces porteurs ou d'agir comme centre de recombinaison. Ces phénomènes affectent la densité de charges libres dans le volume dans le temps de transit des porteurs. En notant N_0 le nombre initial d'électrons créés et de trous créés, initialement égal pour les deux types de charges vu que ce sont des paires qui sont créées par les interactions, on modélise le nombre d'électrons et de trous à un instant t par une loi exponentielle :

$$N_e(t) = N_0 e^{-\frac{t}{\tau_e}} \tag{3-7}$$

$$N_h(t) = N_0 e^{-\frac{t}{\tau_h}} \tag{3-8}$$

où τ_e et τ_h sont les durées de vie respectives des électrons et des trous, dont les valeurs dépendent du matériau considéré, en l'occurrence le CdTe ici.

La charge finalement induite Q_{det} à l'anode s'obtient donc en intégrant la somme des courants (3-5) et (3-6) jusqu'aux temps d'arrivée des électrons sur l'anode t_{f_e} et des trous sur la cathode t_{f_h} :

$$Q_{\text{det}} = N_0 e \left(\int_0^{t_{f_e}} \mu_e e^{-\frac{t}{\tau_e}} \overrightarrow{E_w} \left(\overrightarrow{x_e(t)}\right) . \vec{E}\left(\overrightarrow{x_e(t)}\right) dt + \int_0^{t_{f_h}} \mu_h e^{-\frac{t}{\tau_h}} \overrightarrow{E_w} \left(\overrightarrow{x_h(t)}\right) . \vec{E}\left(\overrightarrow{x_h(t)}\right) dt \right) \tag{3-9}$$

3.2.3 Efficacité d'induction de charges

Sachant que la charge initiale créée est $Q_0 = N_0 . e$, on peut définir l'efficacité d'induction de charges CIE (*Charge Induction Efficiency*) qui correspond à la part de la charge initiale qui est effectivement induite sur le détecteur :

$$\mathbf{CIE}\left(\overrightarrow{X_0}\right) = \frac{\boldsymbol{Q_{\text{det}}}}{\boldsymbol{N_0 e}} = \int_0^{t_{f_e}} \boldsymbol{\mu_e e^{-\frac{t}{\tau_e}} \overrightarrow{E_w}} \left(\overrightarrow{\boldsymbol{x_e(t)}}\right) . \vec{\boldsymbol{E}}\left(\overrightarrow{\boldsymbol{x_e(t)}}\right) \boldsymbol{dt} + \int_0^{t_{f_h}} \boldsymbol{\mu_h e^{-\frac{t}{\tau_h}} \overrightarrow{E_w}} \left(\overrightarrow{\boldsymbol{x_h(t)}}\right) . \vec{\boldsymbol{E}}\left(\overrightarrow{\boldsymbol{x_h(t)}}\right) \boldsymbol{dt} \tag{3-10}$$

L'équation (3-10) est appelée équation de Hecht (Hecht, 1932), elle dépend de la position initiale d'interaction $\overrightarrow{X_0}$. Cette équation est souvent représentée sous une forme plus simple, en l'appliquant au cas où on considère les deux électrodes comme planaires et infinies. Dans ce cas, les champs de pondération et électrique sont uniformes et s'expriment par :

$$\vec{E} = \frac{V_0}{d} \overrightarrow{u_z} \tag{3-11}$$

$$\overrightarrow{E_w} = \frac{1}{d} \overrightarrow{u_z} \tag{3-12}$$

où $\overrightarrow{u_z}$ désigne le vecteur unitaire de la cathode vers l'anode, d est l'épaisseur du cristal et V_0 est la tension de polarisation appliquée entre les deux électrodes.

Sous l'hypothèse de deux électrodes planaires infinies, les déplacements des charges s'effectuent dans la direction suivant la profondeur du cristal. On peut exprimer les distances parcourues par les électrons z_e et les trous z_h par :

$$z_e = \frac{\mu_e V_0}{d} t \tag{3-13}$$

$$z_h = \frac{\mu_h V_0}{d} t \tag{3-14}$$

En faisant le changement de variable correspondant aux deux équations (3-13) et (3-14) dans l'équation (3-10), on obtient :

$$\mathrm{CIE}(z_0) = \frac{1}{d}\left(\int_0^{d-z_0} e^{-\frac{z_e d}{\mu_e \tau_e V_0}} dz_e + \int_0^{z_0} e^{-\frac{z_h d}{\mu_h \tau_h V_0}} dz_h\right) \tag{3-15}$$

en considérant que l'anode est à la profondeur $z = d$ et la cathode en $z = 0$ et que l'interaction s'est produite à la profondeur z_0. En intégrant l'équation (3-15), on obtient directement :

$$\mathrm{CIE}(z_0) = \frac{V_0}{d^2}\left(\mu_e \tau_e \left(1 - e^{-\frac{(d-z_0)d}{\mu_e \tau_e V_0}}\right) + \mu_h \tau_h \left(1 - e^{-\frac{d z_0}{\mu_h \tau_h V_0}}\right)\right) \tag{3-16}$$

L'équation de Hecht est le plus souvent représentée sous cette forme, faisant intervenir notamment le produit $\mu\tau$. Cependant, ce calcul est effectué dans l'hypothèse simplificatrice de deux électrodes planaires parallèles infinies. Dans ce cas, on voit avec l'équation (3-16) que la contribution des trous est forte lorsque l'interaction a lieu à proximité de l'anode et celle des électrons est forte lorsque l'interaction a lieu à proximité de la cathode.

Or, les trous ont un produit $\mu_h \tau_h$ bien plus faible que celui des électrons, principalement du fait de leur mobilité plus faible. À titre de comparaison, d'après les données de l'entreprise ACRORAD (ACRORAD, 2020), $\mu_h \tau_h = 2.10^{-4}\ \mathrm{cm}^2/\mathrm{V}$ alors que $\mu_e \tau_e = 3.3.10^{-3}\ \mathrm{cm}^2/\mathrm{V}$. Cela signifie que la perte de charge due aux trous est bien plus importante. Elle induit une CIE beaucoup plus faible proche de l'anode par rapport à la cathode, nécessitant de mettre une tension V_0 plus élevée pour réduire cet effet, au risque de créer du bruit supplémentaire.

Cependant, pour un détecteur pixellisé, le champ de pondération n'est pas uniforme, et il est en réalité plus fort à proximité de l'anode par rapport à la cathode. Cet effet, appelé « effet petit pixel » (Dirks, 2006), implique qu'une très grande partie du signal est induite à proximité de l'anode. Ainsi, si l'interaction a lieu à proximité de l'anode, les trous induisent une grande partie du signal avant de disparaître du fait des phénomènes pris en compte par l'équation (3-8).

Dans la suite de cette partie, je vais effectuer un calcul plus précis de la CIE, basé sur un calcul du champ de pondération dans le cas pixellisé, ainsi que le calcul numérique de la CIE à partir de l'équation (3-10). De même, je ferai aussi un calcul plus précis du champ électrique qui n'est pas complètement uniforme du fait de la pixellisation.

3.2.4 Résolution numérique de l'efficacité d'induction de charges (CIE)

Calcul du potentiel électrique et du potentiel de pondération

Le calcul des potentiels électriques et de pondération s'effectue à partir de l'équation locale de Maxwell-Gauss :

$$\vec{\nabla}.\vec{E} = \frac{\rho}{\varepsilon_r \varepsilon_0} \tag{3-17}$$

où ρ est la densité de charges dans le milieu considéré, ε_r est la permittivité relative du matériau et ε_0 est la permittivité diélectrique du vide. $\vec{\nabla}$ est l'opérateur nabla de sorte que $\vec{\nabla}.\vec{E}$ représente la divergence de $\vec{E}$.

Dans la suite, je ferai l'hypothèse que la densité de charges est uniformément nulle dans le cristal, c'est-à-dire, qu'il est un diélectrique parfait et qu'il n'existe donc pas de charge d'espace dans le matériau.

Ensuite, le lien entre champ électrique $\vec{E}$ et potentiel U étant $\vec{E} = -\vec{\nabla}(U)$, on en déduit de l'équation (3-17) à résoudre, sous l'hypothèse d'une densité de charge nulle :

$$\Delta U = \frac{\partial^2 U}{\partial x^2} + \frac{\partial^2 U}{\partial y^2} + \frac{\partial^2 U}{\partial z^2} = 0 \tag{3-18}$$

L'opérateur Δ est l'opérateur laplacien. Du fait de la symétrie des pixels en x et y, les pixels étant carrés, je résoudrai le problème en deux dimensions, soit en considérant $\frac{\partial^2 U}{\partial x^2} + \frac{\partial^2 U}{\partial z^2}$.

L'axe z représente l'axe suivant la profondeur du cristal et l'axe x est un des axes orthogonaux à z, suivant la longueur du détecteur.

L'équation (3-18) est appelée équation de Laplace. C'est une équation aux dérivées partielles et elle doit être résolue par des méthodes numériques dans le cas de géométries complexes, comme pour un détecteur pixellisé.

Pour résoudre ce type d'équation, deux méthodes principales existent, à savoir la méthode des différences finies et celle des éléments finis. La description de ces deux méthodes est effectuée au chapitre 5 de (Dubos, 2015).

Pour le calcul du potentiel de pondération, j'ai implémenté la méthode des différences finies, en utilisant la méthode de Gauss-Seidel* couplée à une méthode de relaxation SOR (*Successive Over Relaxation*). À partir d'une discrétisation de l'espace par pas de 10 µm, je note i l'indice en x de la discrétisation et j l'indice en z. Le calcul est effectué de manière itérative. À l'itération $k+1$, le calcul du potentiel de pondération se fait de la manière suivante :

$$U_{i,j}^{k+1} = (1-w)U_{i,j}^{k} + \frac{w}{4}\left(U_{i+1,j}^{k} + U_{i,j+1}^{k} + U_{i-1,j}^{k+1} + U_{i,j-1}^{k+1}\right) \tag{3-19}$$

w est le coefficient de relaxation, que j'ai pris égal à 1,8 de manière empirique. Si ce coefficient est égal à 1, il n'y a pas de relaxation et la convergence est plus lente. Inversement, si ce coefficient est trop grand, le calcul peut diverger.

* La méthode de Jacobi peut aussi être envisagée. L'équation (3-19) est dans ce cas remplacée par :

$$U_{i,j}^{k+1} = \frac{w}{4}\left(U_{i+1,j}^{k} + U_{i,j+1}^{k} + U_{i-1,j}^{k} + U_{i,j-1}^{k}\right)$$

La convergence nécessite en général plus d'itérations que la méthode de Gauss-Seidel, mais la méthode de Jacobi présente l'intérêt d'être plus facilement parallélisable, du fait qu'elle ne fait appel qu'au calcul de U à l'étape k et ne fait pas appel à certaines valeurs de U à l'étape $k+1$. Elle peut être in fine plus rapide à calculer grâce à la parallélisation malgré un nombre d'itérations plus élevé.

J'ai considéré un domaine composé de 5 pixels où j'ai calculé le potentiel de pondération induit sur le pixel central. Pour cela, j'ai fixé le potentiel du pixel central à 1*, et le potentiel des autres pixels et de la cathode à 0.

La méthode des différences finies exige que les conditions à tous les bords du domaine soient définies, du fait que l'équation (3-19) utilise les termes $i-1, j-1, i+1$ et $j+1$ de la discrétisation. De ce fait, sur l'axe x, j'ai défini des conditions de bord périodiques, de sorte que $U_{0,j} = U_{i_{\max},j}$. Sur l'axe y, j'ai étendu le domaine jusqu'à 6 mm, soit 1 mm de cristal et 5 mm vide au-delà des anodes†, de sorte à éviter des effets de bord, et j'ai mis le potentiel à 0 sur cette dernière frontière. Enfin, une condition initiale est nécessaire à l'itération $k = 0$, j'ai choisi de mettre le potentiel $U^0 = 0$ sur l'ensemble du domaine.

La Figure 3-1 représente le résultat du calcul du potentiel de pondération avec la méthode des différences finies. On voit que le potentiel de pondération n'est pas uniformément croissant au centre du pixel, la croissance est d'autant plus marquée avec la profondeur. Cela signifie que le signal est majoritairement induit en profondeur, à proximité de l'anode. Au centre de l'inter-pixel, on observe même une décroissance du potentiel de pondération, ce qui doit se traduire par une induction négative au niveau du pixel : une charge se déplaçant dans cette zone va retirer du signal au niveau de l'anode. Je montre cet effet en paragraphe 3.2.5 sur l'efficacité d'induction de charges.

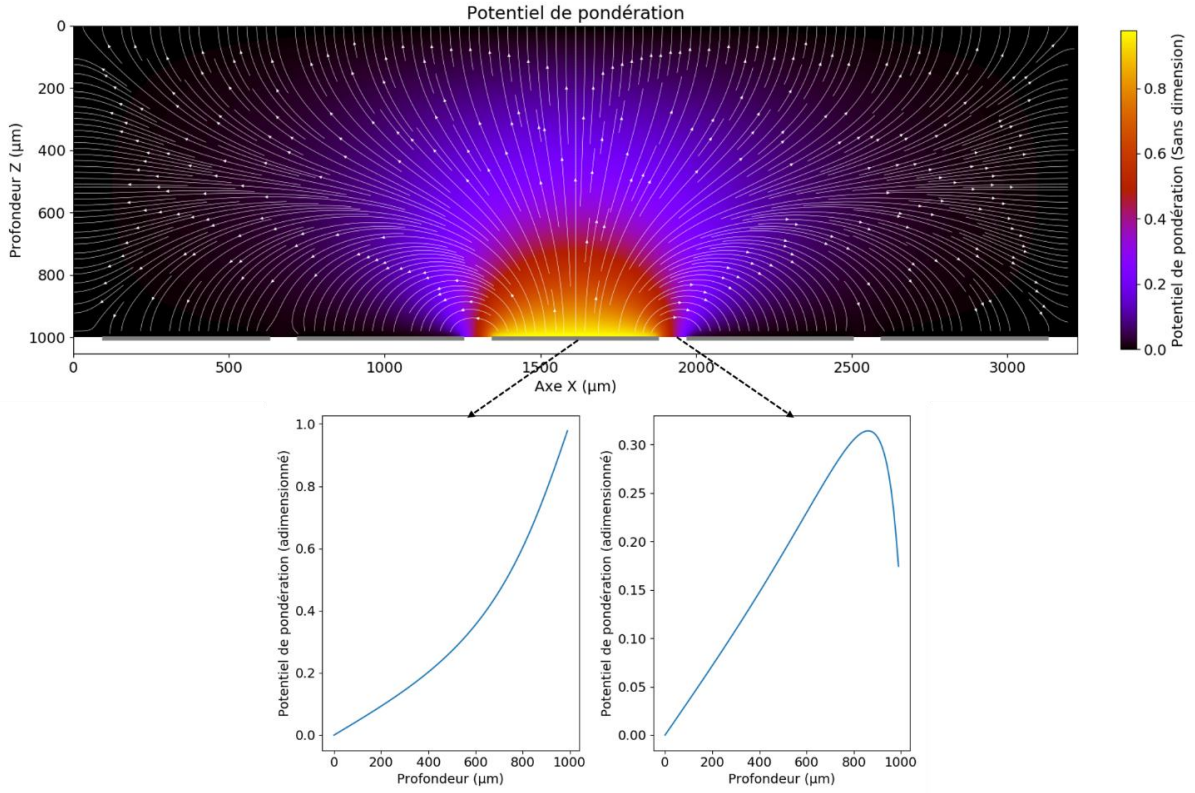


Figure 3-1 : En haut, potentiel de pondération de l'anode centrale. Les lignes de champ sont représentées en blanc, les pixels sont représentés en gris. En bas, coupe du potentiel de pondération au centre du pixel et au centre de l'inter-pixel.

* Le potentiel de pondération est en réalité un potentiel adimensionné

† Idéalement, cette distance est infinie. Une trop grande distance conduit à un nombre de mailles très important qui n'améliore pas significativement le résultat mais allonge considérablement le temps de calcul et la mémoire requise pour son exécution. 5 mm est ici un compromis.

À l'aide du logiciel de simulation et de calcul numérique COMSOL Multiphysics® (COMSOL AB), j'ai appliqué la méthode des éléments finis pour effectuer ce même calcul. La méthode des éléments finis, basée sur la recherche approchée d'une solution sur un maillage de l'espace, est en général plus adaptée pour des problèmes à géométrie complexe, et selon la précision du maillage, peut aussi donner une solution plus proche de la solution exacte. J'ai regardé la différence entre le résultat donné par mon algorithme de différences finies et celui donné en éléments finis, notamment afin de valider mon algorithme.

La Figure 3-2 représente la différence des deux calculs, en valeur absolue.

Les plus grandes différences sont situées aux bords des pixels voisins. Elles sont très localisées et relativement faibles, de l'ordre de 0,05, qui correspond à un écart relatif de 33 % à cet endroit, à mettre en rapport au potentiel du pixel central qui est de 1. Des différences un peu plus marquées sont aussi à noter aux bords du domaine sur l'axe X, qui viennent du fait que j'ai défini des conditions périodiques dans ma résolution en différences finies alors que ces conditions étaient libres sur ce bord du domaine dans le cas des éléments finis. Ces différences sont dans des zones très restreintes, seulement quatre *bins* de 10 µm ont des écarts supérieurs à 0,03, et ils concernent des zones à la frontière du domaine, au niveau des pixels voisins et non du pixel central d'intérêt.

Cela montre que je peux valider mon modèle au regard du calcul en éléments finis et utiliser son résultat. La méthode des différences finies présente l'intérêt d'être implémentable plus facilement dans un logiciel sans licence.

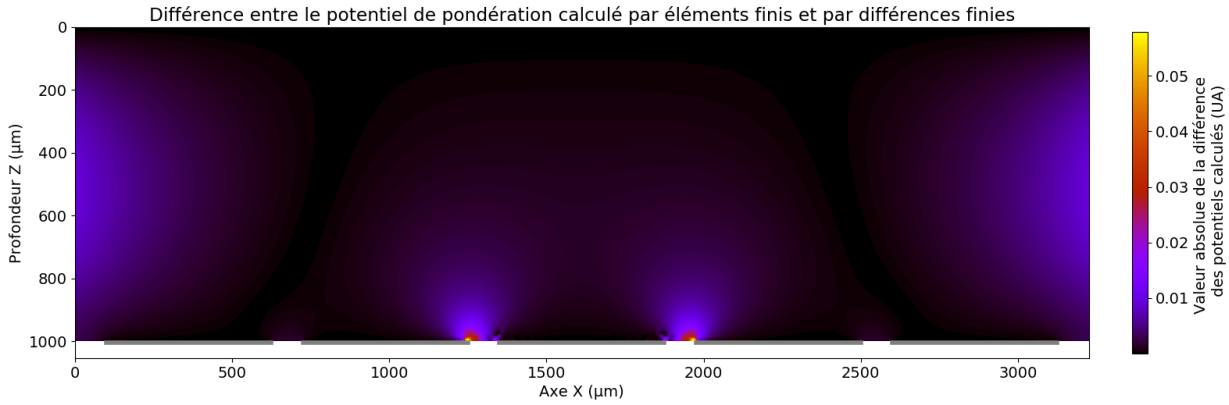


Figure 3-2 : Comparaison entre le calcul en éléments finis et celui en différences finies. Les différents pixels sont représentés en gris.

Dans un second temps, j'effectue un calcul plus précis du champ électrique, sans considérer qu'il est uniforme sur l'ensemble du domaine. En effet, le champ est créé par les pixels et il faut prendre en compte les zones entre les pixels pour lesquelles nous n'avons pas d'action sur le potentiel électrique. J'ai refait le calcul du potentiel électrique pour résoudre l'équation (3-18) en modifiant uniquement les conditions suivantes : j'impose un potentiel de 0 V à la cathode et de 300 V sur toutes les anodes.

J'ai choisi cette tension de $V_0 = 300$ V car elle correspond à la tension habituelle que j'applique sur le cristal (300 V/mm). Dans le cas où on appliquerait une tension différente V_{new}, il suffirait de multiplier le résultat de mon calcul par un facteur correctif $\frac{V_{\mathrm{new}}}{V_0}$.

La Figure 3-3 présente le résultat du calcul du potentiel électrique.

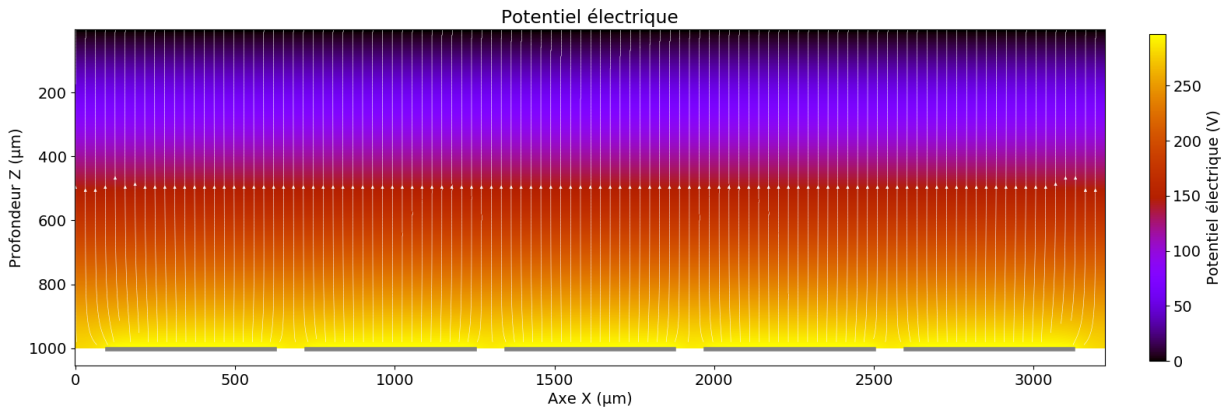


Figure 3-3 : Potentiel électrique. Les lignes blanches sont les lignes de champ électrique, les bandes grises sont les pixels

On constate, à travers les lignes de champ, que le champ électrique est presque uniforme sur l'ensemble du domaine. La non-uniformité intervient localement entre les pixels. Cependant, ce calcul reste important, notamment dans le cas où on aurait des pixels plus petits ou des écarts plus grands entre chaque pixel, comme montré par (Dubos, 2015).

Calcul de l'efficacité d'induction de charges

Les calculs des champs électriques et du champ de pondération étant effectués, je peux calculer numériquement l'intégrale de l'équation (3-10) le long de la trajectoire d'une charge afin d'obtenir la CIE.

Pour ce faire, je considère une charge élémentaire, électron ou trou, créée en un point $\overrightarrow{X_0} = (x_0, z_0)$ et je calcule sa trajectoire pas-à-pas, avec un pas de temps Δt. En utilisant le champ électrique calculé précédemment et la formule (3-4), je calcule à chaque pas de temps la vitesse instantanée $\vec{v}\left(\vec{X}(t)\right)$ de la charge en fonction de sa position puis je mets à jour numériquement sa nouvelle position : $\vec{X}(t + \Delta t) = \vec{X}(t) + \vec{v}\left(\vec{X}(t)\right)\Delta t$. Je répète ce processus jusqu'à ce que la charge atteigne le plan des pixels $z = d$ si c'est un électron ou le plan de la cathode $z = 0$ si c'est un trou.

De même, à chaque pas de temps, grâce au calcul du champ de pondération et à la formule (3-5) ou (3-6) en fonction du type de charge et connaissant la position de la charge, je détermine l'intensité instantanée du courant induit par la charge au pas de temps t, que je note $i(t)$.

J'utilise aussi le ratio de perte de charge grâce aux formules (3-7) et (3-8), qui est de $e^{-\frac{t}{\tau_{\text{charge}}}}$.

Je calcule ensuite l'efficacité d'induction de charges de la manière suivante :

$$CIE = \sum_{\text{pas de temps } t_e} \frac{I_e(t_e)}{e} e^{-\frac{t_e}{\tau_e}} \Delta t_e + \sum_{\text{pas de temps } t_h} \frac{I_h(t_h)}{e} e^{-\frac{t_h}{\tau_h}} \Delta t_h \qquad (3\text{-}20)$$

Le nombre de pas de temps et leur durée ne sont pas égaux pour les électrons et les trous, étant donné que les électrons ont une vitesse plus élevée que les trous dans le matériau. Pour la discrétisation du temps, j'applique un pas de temps Δt de 0,01 ns pour les électrons et de 0,1 ns pour les trous. Cela correspond typiquement à 2,5 % du temps nécessaire pour un électron pour parcourir 10 µm, soit

0,025 % du temps nécessaire à un électron produit à la cathode pour atteindre l'anode. De même, le pas de temps pour les trous correspond à 3 % du temps nécessaire pour parcourir 10 µm.

J'applique cette méthode pour des électrons et trous répartis initialement sur une grille au pas de 10 µm pour discrétiser l'espace et j'effectue le calcul précédent pour chaque charge.

Pour les données physiques, je me suis référé au produit $\mu\tau$ fourni par l'entreprise ACRORAD (ACRORAD, 2020). Cependant, j'ai besoin de distinguer la mobilité μ et le temps de vie τ des charges. J'ai choisi les valeurs de mobilité répertoriées par (Meuris, 2009) et j'ai calculé la valeur du temps de vie correspondant au produit $\mu\tau$ des données d'ACRORAD. Cela donne :

- Pour les électrons : $\mu_e = 950\ \mathrm{cm^2V^{-1}s^{-1}}$ et $\tau_e = 3{,}3\ \mu\mathrm{s}$
- Pour les trous : $\mu_h = 73\ \mathrm{cm^2V^{-1}s^{-1}}$ et $\tau_h = 2{,}7\ \mu\mathrm{s}$

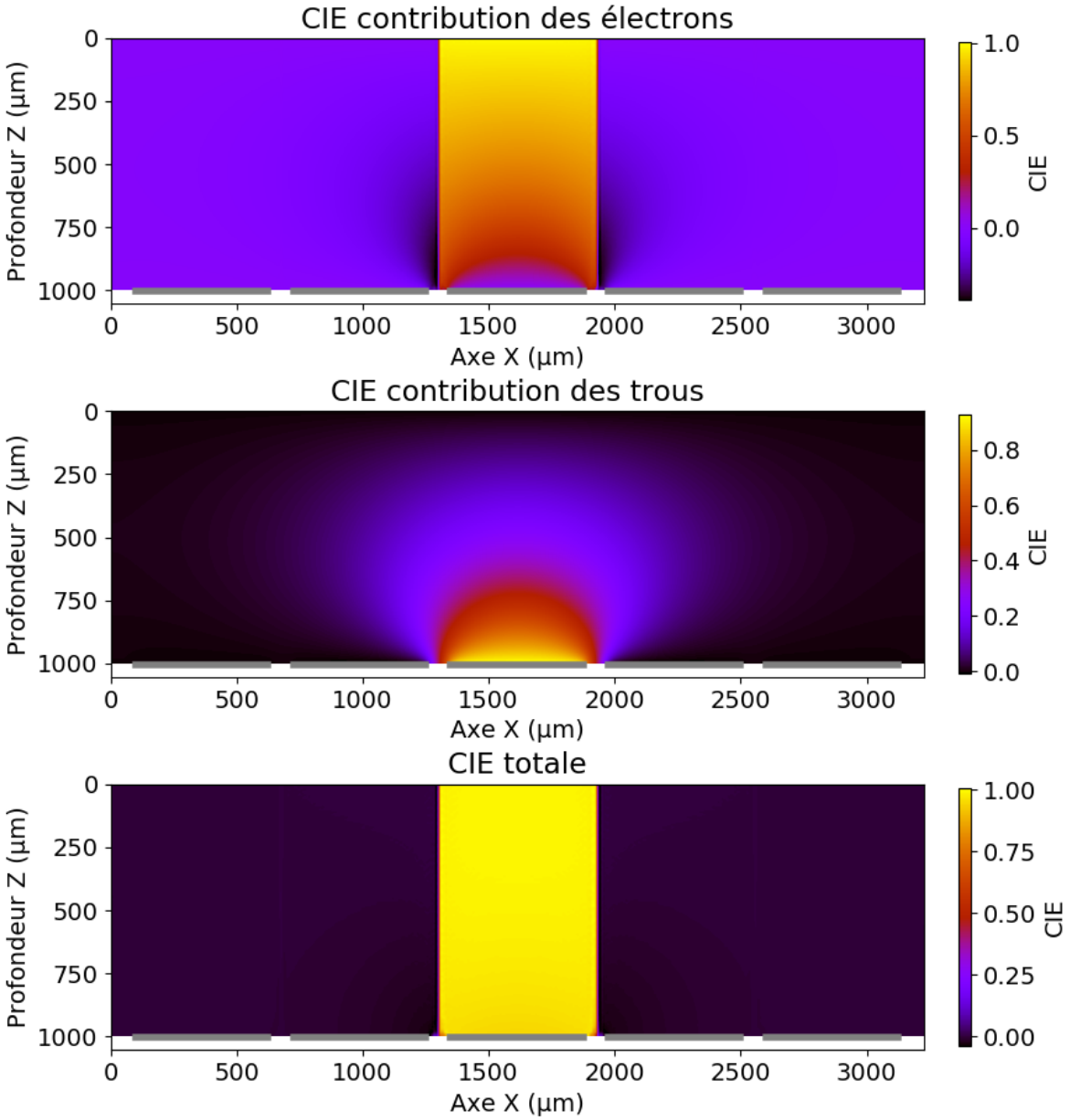


Figure 3-4 : Efficacité d'induction de charges. Dans l'ordre : contribution des électrons, contribution des trous, CIE totale.

La Figure 3-4 représente le calcul de l'efficacité d'induction de charges. Dans l'équation (3-20) et avec ma méthode de calcul, je peux différencier la contribution des trous et celle des électrons dans le signal induit dans le détecteur et calculer la CIE totale.

L'efficacité d'induction de charges est principalement portée par les électrons dans les trois premiers quarts du détecteur. Dans la zone la plus proche de l'anode, le signal est porté par les trous principalement, qui ont une mobilité plus faible. Cependant, comme le potentiel de pondération varie rapidement à proximité du pixel, comme on le voit en Figure 3-1, la plus grande partie du signal est induite dans cette zone. L'effet de perte de charge due aux trous est réduit dans cette zone.

Ainsi, en sommant les deux contributions, la CIE est très proche de 1 sur la zone directement au-dessus du pixel, jusqu'à la moitié de l'inter-pixel. Plus précisément, les calculs montrent qu'elle est de 95 % à proximité du pixel et monte à plus de 99 % à proximité de la cathode.

Un autre effet intéressant à noter est la présence de zones où la CIE est négative, pouvant aller jusqu'à -3 %. Cela se produit notamment lorsque l'interaction a lieu dans l'autre moitié de l'inter-pixel, comme je le montre sur les coupes en Figure 3-5. Cela s'explique par la forme du champ de pondération que l'on voit en Figure 3-1. En effet, dans l'autre moitié de l'inter-pixel, en profondeur, le champ de pondération est orienté vers le bas du cristal, au niveau des « boucles ». Cela implique que si une charge est créée dans cette zone, le courant induit par les charges est négatif et va se soustraire à la charge totale. On voit que cet effet peut retirer jusqu'à 3 % de la charge totale induite sur le pixel central, ce qui représente quelques keV pour des photons de haute énergie de plus de 100 keV. De ce fait, si une interaction a lieu dans la zone de l'inter-pixel, le nuage de charge créé peut s'étaler à la fois dans une zone où la CIE est proche de 1 et dans une zone où la CIE est négative ce qui retirerait du signal induit sur le pixel.

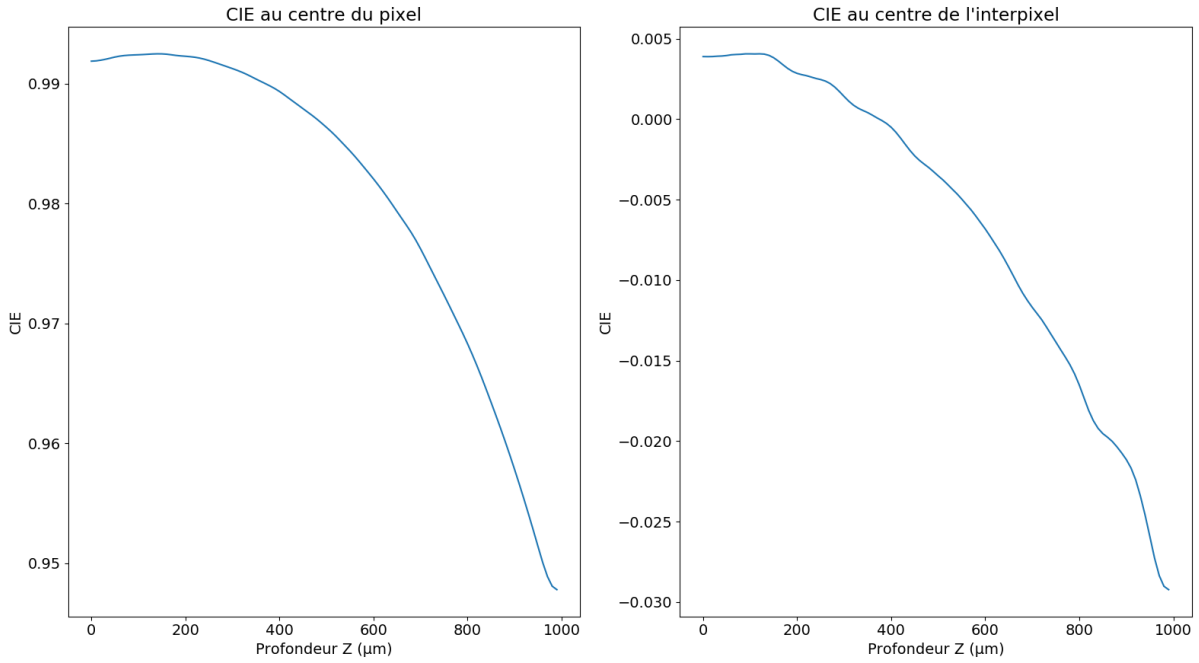


Figure 3-5 : Efficacité d'induction de charges au centre du pixel et au centre de l'inter-pixel. Les fluctuations visibles sur le graphique de droite sont dues aux erreurs numériques dans la méthode que j'ai mise en place, à la fois au niveau du calcul du potentiel de pondération et du calcul de l'induction de la charge.

3.2.5 Prise en compte de l'efficacité d'induction de charges dans la réponse du détecteur

En paragraphe 3.1, j'ai décrit l'influence des fluctuations du nombres de charges créées pour une énergie donnée E_0, qui conduit à une énergie effective que j'ai notée E_{Fano}, correspondant aux nombres de paires de charges effectivement créées.

Maintenant, en fonction de la position du dépôt d'énergie $\overrightarrow{X_0}$, j'utilise la CIE calculée précédemment afin de déterminer l'énergie E_{CIE} qui sera réellement enregistrée par le détecteur du fait de la perte de charge :

$$E_{\text{CIE}} = E_{\text{Fano}} \times CIE\left(\overrightarrow{X_0}\right) \tag{3-21}$$

3.3 Bruit électronique

Une fois que la charge est induite sur l'anode, la chaîne d'acquisition de l'électronique de lecture a pour rôle d'amplifier ce signal et de l'enregistrer. Dans notre cas, le détecteur Caliste utilise des électroniques de lectures spécifiques dédiées, les ASICs IDeF-X HD (Michalowska, et al., 2010). Ils sont principalement constitués de :

- un préamplificateur de charges, qui convertit la charge induite en un échelon de tension ;
- un ensemble de filtres qui met en forme le signal en tension ;
- un détecteur de pic qui mémorise le sommet de l'impulsion mise en forme ;
- un « trigger » qui déclenche la logique de lecture du circuit lorsqu'un événement a eu lieu avec une énergie au-delà d'un certain seuil.

Le bruit électronique est alors défini comme l'ensemble des sources de fluctuations autres que la fluctuation du nombre de paires de charges créées et que la perte de charge. Il a pour cause principale l'électronique de lecture (Michalowska, et al., 2010), dont l'optimisation consiste à minimiser ce bruit.

Il est d'autant plus élevé que le courant présenté en entrée de la chaîne d'acquisition est important. La raison principale est le courant d'obscurité, qui est le courant continu qui traverse le détecteur dès qu'il est mis sous tension, dû aux fluctuations thermiques dans le détecteur. Il doit être finement caractérisé pour optimiser les performances spectrales du détecteur.

Le bruit intrinsèque de l'électronique de lecture se mesure en charge équivalente de bruit, notée ENC (*Equivalent Noise Charge*). Il se mesure de la manière suivante. On note G le gain total de la chaîne d'acquisition qui convertit la charge Q induite sur l'anode en signal mis en forme de tension de sortie U_s, tel que $U_s = G.Q$. On désigne le bruit quadratique moyen, noté rms (*root mean square*) de cette tension de sortie, par σ_{U_s}. L'ENC est alors la charge, en nombre d'électrons, à apporter en entrée pour que la tension de sortie soit σ_{U_s} et donc pour que le rapport signal sur bruit soit égal à 1.

$$\text{ENC} = \frac{Q_{\text{bruit}}}{e} = \frac{\sigma_{U_s}}{e.G} \tag{3-22}$$

On exprime l'ENC en électrons rms. Le bruit électronique a une implication directe sur la mesure de l'énergie puisqu'il correspond aux fluctuations sur le nombre de charges que le signal enregistré représente. Si N charges sont créées par un dépôt d'énergie, le signal de sortie a des fluctuations qui représentent en moyenne N charges avec un écart-type de ENC charges.

Or, on a vu en paragraphe 3.1 qu'une charge représente $w = 4{,}42$ eV pour le CdTe, on en déduit que ce bruit électronique ajoute des fluctuations d'un écart-type $\sigma_{\mathrm{ENC}} = \mathrm{ENC} \times w$. L'ENC est modélisée ici comme étant indépendante de l'énergie déposée par l'interaction, et constitue une source de bruit intrinsèque au détecteur.

Pour prendre en compte l'ENC, à partir de l'énergie E_{CIE} calculée au paragraphe 3.2.5, je tire aléatoirement une énergie E_{final} dans une gaussienne de moyenne E_{CIE} et d'écart-type $\mathrm{ENC} \times w$.

$$E_{\mathrm{final}} \sim \mathcal{N}(\mu = E_{\mathrm{CIE}}, \sigma = \mathrm{ENC} \times w) \tag{3-23}$$

J'ai choisi $ENC = 80\ \mathrm{e}^{-}\mathrm{rms}$. Cette valeur est motivée par les mesures faites sur IDeF-X HD (Michalowska, et al., 2010) qui ont déterminé un bruit compris entre $68\ \mathrm{e}^{-}\mathrm{rms}$ et $100\ \mathrm{e}^{-}\mathrm{rms}$.

4 Modélisation des spectres, simulations et réponse du détecteur

4.1 Utilisation de la simulation Geant4 et de la réponse du détecteur

En partie 2.1, j'ai présenté comment je simule à partir de Geant4 une liste d'interactions à partir d'une source de photons mono-énergétique. J'utilise maintenant la modélisation de la réponse du détecteur que j'ai présentée en partie 3 afin de l'appliquer aux simulations pour obtenir des spectres mono-énergétiques. Ces spectres me permettront, en les combinant, de créer des spectres de sources.

Sur la liste d'interactions en sortie de la simulation Geant4, pour une interaction donnée j, il est possible que des dépôts de charge s'effectuent à différents endroits, sur la trajectoire des photoélectrons et par interaction des photons de fluorescence. Pour l'interaction j, je note $E_{(j,k)}$ l'énergie du dépôt k, et $\vec{X}_{j,k}$, la position de ce dépôt, qui est connue dans la simulation.

Sur cette énergie, je tire aléatoirement une énergie $E_{(j,k)_{\mathrm{Fano}}}$ comme décrit par l'équation (3-2) pour prendre en compte les fluctuations liées au nombre de paires électrons-trous créées.

Ensuite, j'applique l'efficacité d'induction de charges à travers l'équation (3-21). Comme je l'ai montré en paragraphe 3.2.5, une interaction qui a lieu dans un pixel voisin peut induire un signal, éventuellement négatif, sur le pixel central où a eu lieu l'interaction. Pour prendre cela en compte, je considère que le pixel central est celui au plus proche de la position du dépôt d'énergie, et je calcule une énergie $E_{(j,k)_{\mathrm{CIE}}}$ pour tous les pixels voisins du dépôt d'énergie. Cette étape est importante pour déterminer la multiplicité de l'événement. La Figure 4-1 représente ce phénomène.

Sur cet exemple, le dépôt d'énergie est de 200 keV dans le pixel principal mais a induit simultanément un signal négatif sur les pixels voisins. Ainsi, si par exemple l'interaction fait un autre dépôt d'énergie dans le pixel en haut à gauche, disons de 400 keV, l'énergie totale que verra ce pixel sera de $400 - 2 = 398\ \mathrm{keV}$. J'ai représenté ici des cas où l'induction dans les pixels voisins est négative ce qui correspond à des charges créées en profondeur dans la zone inter-pixel comme on l'a vu en Figure 3-5. En fonction de la position du dépôt d'énergie, cette induction peut aussi ajouter du signal sur les pixels voisins selon la CIE montrée en Figure 3-4.

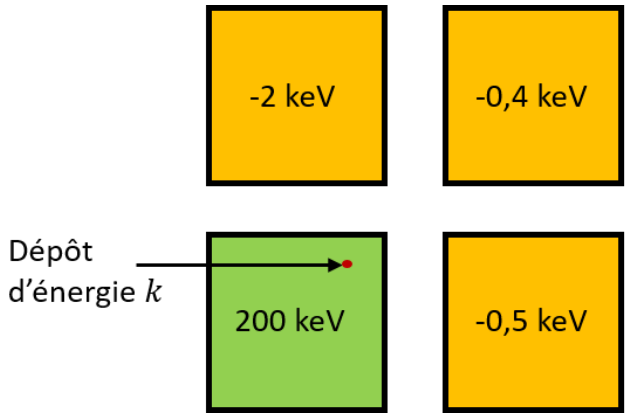


Figure 4-1 : Induction sur les pixels voisins, les valeurs écrites dans les pixels correspondent à l'énergie vue par ces pixels du fait de ce dépôt d'énergie.

Pour chaque pixel p dans lequel je calcule un dépôt d'énergie, éventuellement négatif, associé au dépôt k. Cette énergie $E_{(j,k)_{\mathrm{CIE}}}(p)$ est donnée par l'équation (3-21).

Puis, j'applique le bruit électronique en tirant aléatoirement pour chaque pixel p une énergie $E_{(j,k)_{\mathrm{ENC}}}(p)$ à partir de $E_{(j,k)_{\mathrm{CIE}}}(p)$ selon l'équation (3-23).

J'applique cette méthode sur l'ensemble des dépôts k associés à l'interaction j, et ceci pour l'ensemble des pixels du détecteur. Ainsi, pour un pixel p du détecteur, je calcule l'énergie finale que voit ce pixel en sommant toutes les contributions induites par l'ensemble des dépôts k produits lors de l'interaction j :

$$E_j(p) = \sum_k E_{(j,k)_{\mathrm{ENC}}}(p) \tag{4-1}$$

Pour les pixels non voisins de l'interaction, cette énergie sera de 0 car je ne calcule l'étape d'efficacité d'induction de charges que pour les pixels voisins de chaque dépôt.

Pour les pixels voisins de l'interaction en revanche, l'énergie calculée est non nulle et on peut y associer une énergie par pixel, comme schématisé en Figure 4-2.

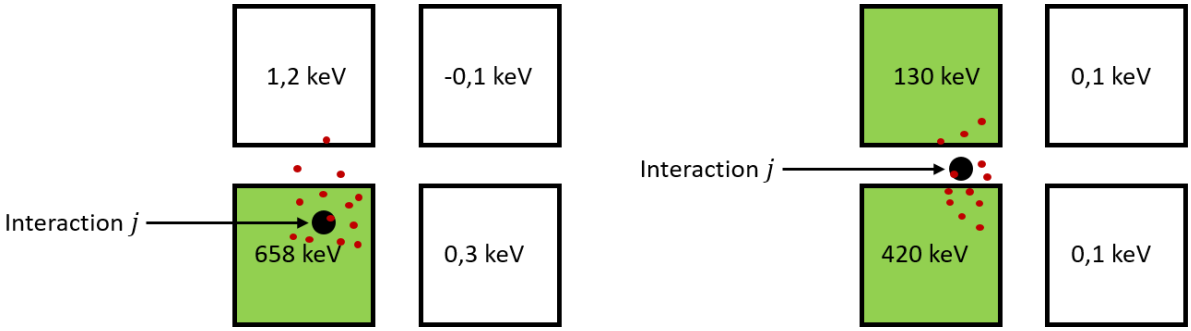


Figure 4-2 : Dépôt total d'énergie. Les points rouges sont les différents dépôts k d'énergie. À gauche, l'événement est simple, à droite, il est double.

Cela détermine donc la multiplicité des événements : si la somme des énergies enregistrées pour chaque dépôt dépasse un certain seuil dans un seul pixel, l'événement est simple, si c'est dans plusieurs pixels, il est multiple. Dans mes simulations, j'ai choisi un seuil de 5 keV. Cela correspond

globalement au seuil bas typique que j'utilise avec Caliste pour des mesures réelles. Par la suite, je ne traiterai que les événements simples, du fait de leur résolution que l'on attend meilleure que pour les événements multiples. Pour les événements multiples, la perte de charge attendue est plus importante entre les pixels et la résolution est aussi dégradée à cause des bruits intrinsèques de chaque canal électronique qui se combinent quadratiquement.

J'applique ensuite cette méthode sur l'ensemble des interactions *j* de ma simulation Geant4. J'obtiens en sortie une liste d'événements conforme aux données réelles issues du détecteur. L'ensemble de mon système de simulation et de modélisation de réponse du détecteur est résumé en Figure 4-3.

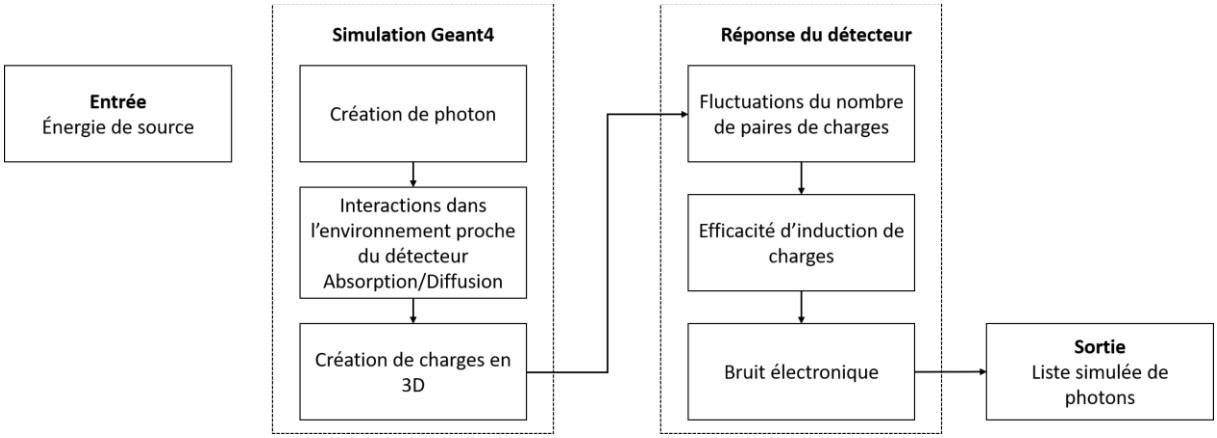


Figure 4-3 : Pipeline de simulation

À titre d'illustration, la Figure 4-4 montre le résultat de cette méthode sur la simulation que j'avais faite en Figure 2-3 à 32 keV et 662 keV.

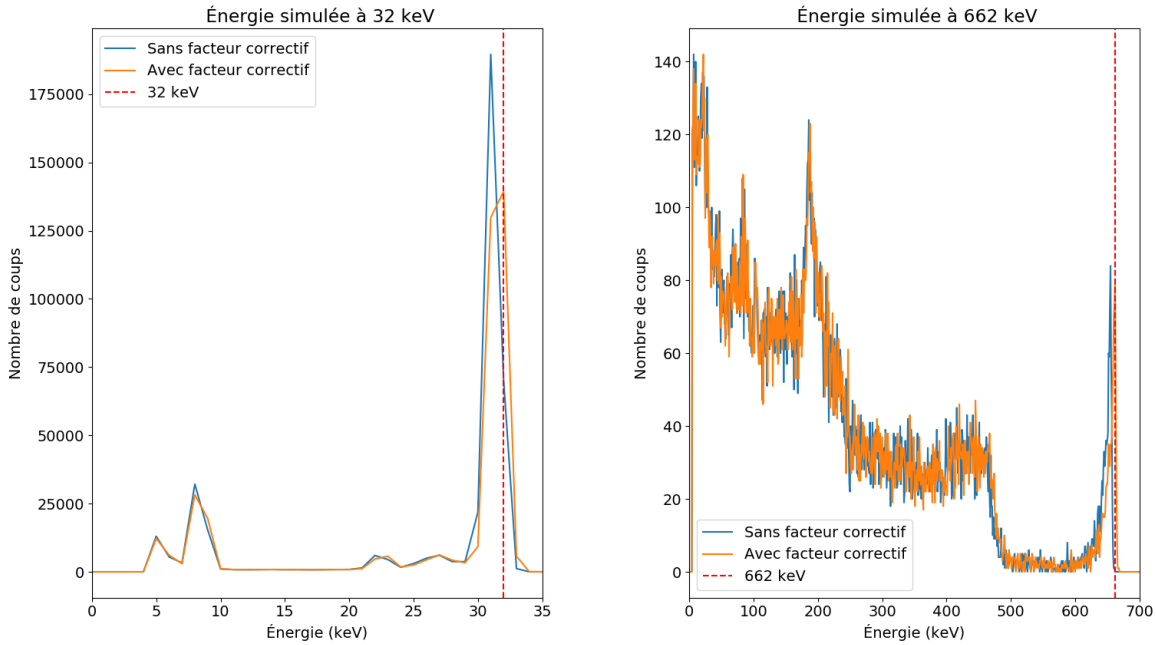


Figure 4-4 : Application de la réponse du détecteur aux simulations à 32 keV et 662 keV.

J'ai constaté que l'énergie que je retrouvais en appliquant directement la réponse du détecteur donnait un pic photoélectrique décalé vers les basses énergies d'environ 1 % par rapport à l'énergie

attendue. Cela provient en réalité de l'effet de perte de charge. L'efficacité d'induction de charges est au maximum de 100 %, mais elle est inférieure à 100 % sur une grande partie de la zone au-dessus du pixel d'interaction. En réalité, elle est plutôt majoritairement située entre 98 % et 100 %. De ce fait, je dois appliquer un facteur correctif sur mon résultat, car dans le cas de vraies données, la calibration compense justement cette perte de charge vu que les pics photoélectriques sont positionnés au niveau des lignes d'émission. Empiriquement, j'ai constaté qu'un facteur correctif de 1,01 recale les pics photoélectriques à la bonne position pour l'ensemble des énergies que j'ai simulées.

Je souligne aussi que dans ces simulations, j'ai considéré que tous les dépôts d'énergies sont ponctuels. En réalité, sous l'effet de processus de diffusion (Knoll, 2000), les nuages de charge créés s'élargissent au cours du temps, ce que l'on peut modéliser par une gaussienne en 3D d'écart-type $\sigma(t) = \sqrt{2Dt}$, où D est le coefficient de diffusion donné par la relation d'Einstein :

$$D = \frac{k_B T}{e} \mu_e \tag{4-2}$$

Cet effet interviendrait particulièrement dans le calcul de l'efficacité d'induction de charges, et cela alourdirait considérablement le calcul. Cependant, je montre dans le paragraphe suivant que l'approximation des charges ponctuelles est suffisante pour modéliser assez fidèlement la perte de charge dans les spectres.

4.2 Application à la simulation de spectre

4.2.1 Application et comparaison avec des mesures réelles

Pour prendre en compte une vraie source avec des lignes d'émission de différentes intensités, j'utilise mes simulations mono-énergétiques pour simuler des spectres de radioéléments selon leurs lignes d'émission.

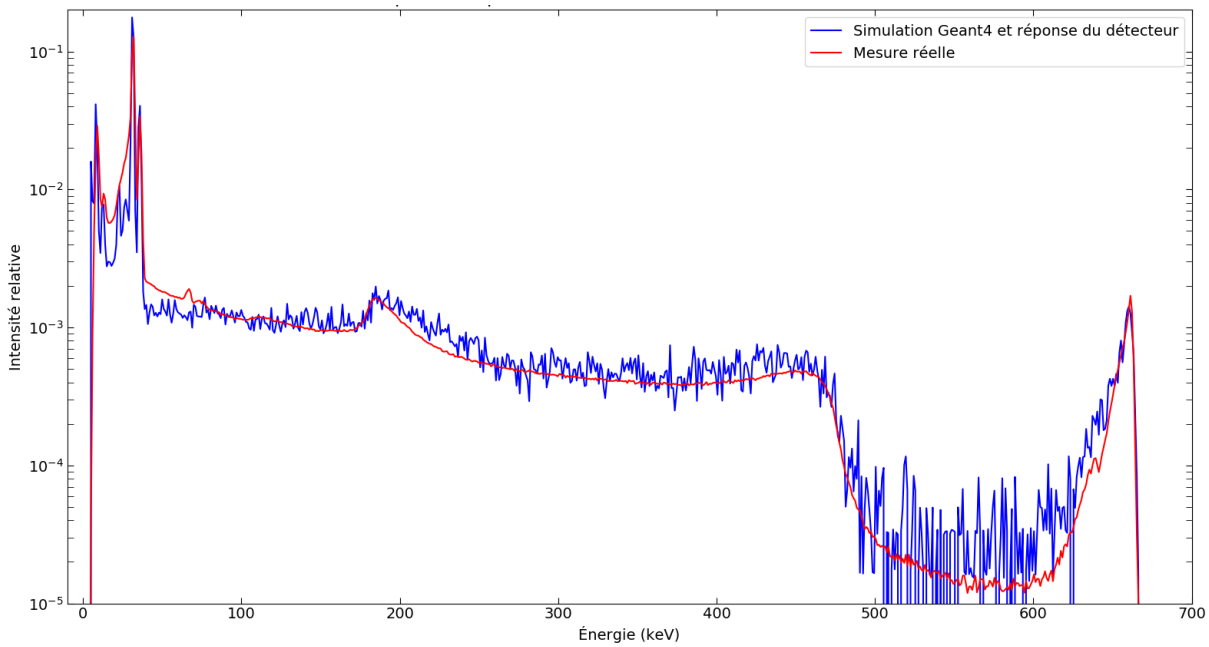


Figure 4-5 : Comparaison d'un spectre simulé de ^{137}Cs et d'un spectre réel. Acquisition d'une source de ^{137}Cs de 1,6 MBq placée à 1 cm du détecteur pendant 72 minutes. Module utilisé : Caliste-HD ; haute-tension appliquée au cristal : 300 V ; température du détecteur : -10°C ; temps de *peaking* : 6,07 µs.

La Figure 4-5 montre la comparaison de mes simulations avec le spectre en Figure 1-12. Les spectres sont normalisés par le nombre de coups total détectés ou simulés.

La simulation est plus bruitée que la mesure, du fait que la mesure a été effectuée avec un temps de pose long de 72 minutes, avec une source de 1,6 MBq à près de 1 cm du détecteur. Cela donne un nombre de photons important au total de plus de 14 millions d'événements simples. Il est possible d'augmenter le nombre de photons dans la simulation au prix d'un temps de calcul beaucoup plus important et de ressources de stockage des données en conséquence, vu que 4000 sources mono-énergétiques sont simulées.

On notera que l'ensemble des structures sont représentées dans la simulation. La structure principale Compton est présente dans la simulation avec des intensités très proches. De même, les pics photoélectriques sont correctement représentés. La perte de charge, dans la queue de distribution du pic à 662 keV, se retrouve dans la simulation. On notera que la perte de charge est un peu plus importante dans la simulation que sur la vraie mesure, ce qui peut être lié au choix des paramètres μ et τ des charges, que l'on peut chercher à adapter pour retrouver plus précisément le spectre. Pour le pic à 32 keV, la simulation semble mieux résolue que la mesure réelle à gauche du pic. Cet effet provient en réalité de la diffusion Compton des photons dans le *packaging* en plexiglas contenant la source de ^{137}Cs, qui n'est pas prise en compte dans la simulation. Ces diffusions peuvent diminuer l'énergie des photons de 32 keV jusqu'à 28,4 keV.

4.2.2 Utilisation des structures Compton multiples

L'exemple précédent est une acquisition effectuée avec la caméra WIX en laboratoire. L'environnement est peu fermé autour du banc d'acquisition dans ce cas. De ce fait, il n'y a pas vraiment d'artéfact visible sur la Figure 4-5 du fait des diffusions multiples Compton dans l'environnement.

En revanche, j'ai effectué une acquisition avec un Caliste-O, situé dans un banc de test clos, constitué d'une enceinte en acier, que l'on peut voir en Figure 4-6, d'une trentaine de centimètres de diamètre et de profondeur, que l'on utilise pour tester les détecteurs avant de les intégrer dans une caméra gamma ou un autre système. La source est posée sur une fenêtre de Béryllium au-dessus de l'enceinte.

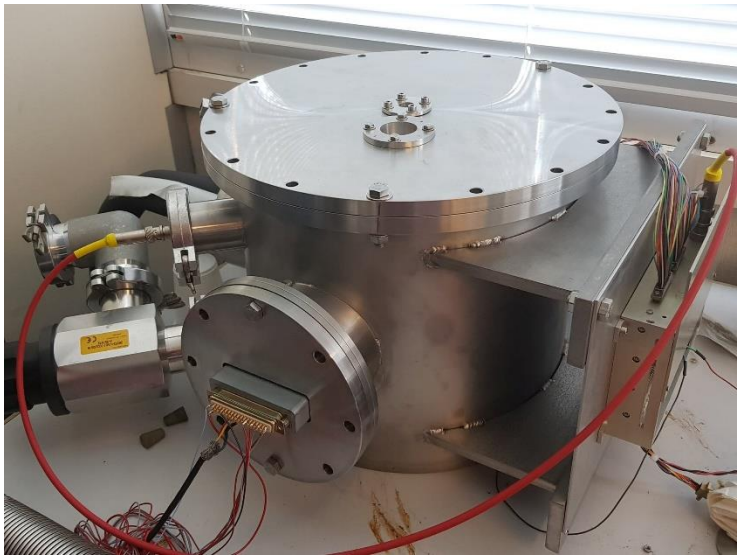


Figure 4-6 : Enceinte de test des détecteurs Caliste-O

L'environnement est donc clos et propice aux diffusions Compton multiples. La Figure 4-7 montre en rouge le résultat de la mesure. On voit notamment entre 100 keV et 200 keV une structure bombée qui n'était pas présente dans la mesure que j'ai présentée en Figure 4-5. Cette structure bombée est une signature typique des diffusions Compton multiples dans l'environnement.

J'ai représenté sur cette figure la simulation que j'ai faite précédemment, en introduisant un facteur correctif f_{eff} du fait que le Caliste utilisé est Caliste-O et non Caliste-HD, soit un cristal de 2 mm d'épaisseur au lieu de 1 mm. Ce facteur corrige de l'efficacité de détection, il doit être appliqué à chaque canal du spectre simulé, pour une énergie E donnée :

$$f_{\text{eff}}(E) = \frac{1 - e^{-\mu(\text{CdTe},E)\times(d=2\text{ mm})}}{1 - e^{-\mu(\text{CdTe},E)\times(d=1\text{ mm})}} \tag{4-3}$$

On constate que sur cette simulation, en vert sur la figure, je ne retrouve pas cette structure bombée entre 100 et 200 keV. J'ai donc cherché à expliquer cette structure en utilisant les formules analytiques des diffusions Compton multiples que j'ai calculées en partie 2.2. J'ai ainsi ajusté l'intensité de ces structures par rapport au spectre mesuré. J'ai utilisé seulement les cas où au moins deux diffusions Compton se sont produites dans l'environnement, constitué ici principalement de l'enceinte en acier contenant le détecteur. La simulation prend en compte déjà une partie de l'environnement proche du détecteur dans lequel il est probable qu'une seule diffusion ait lieu.

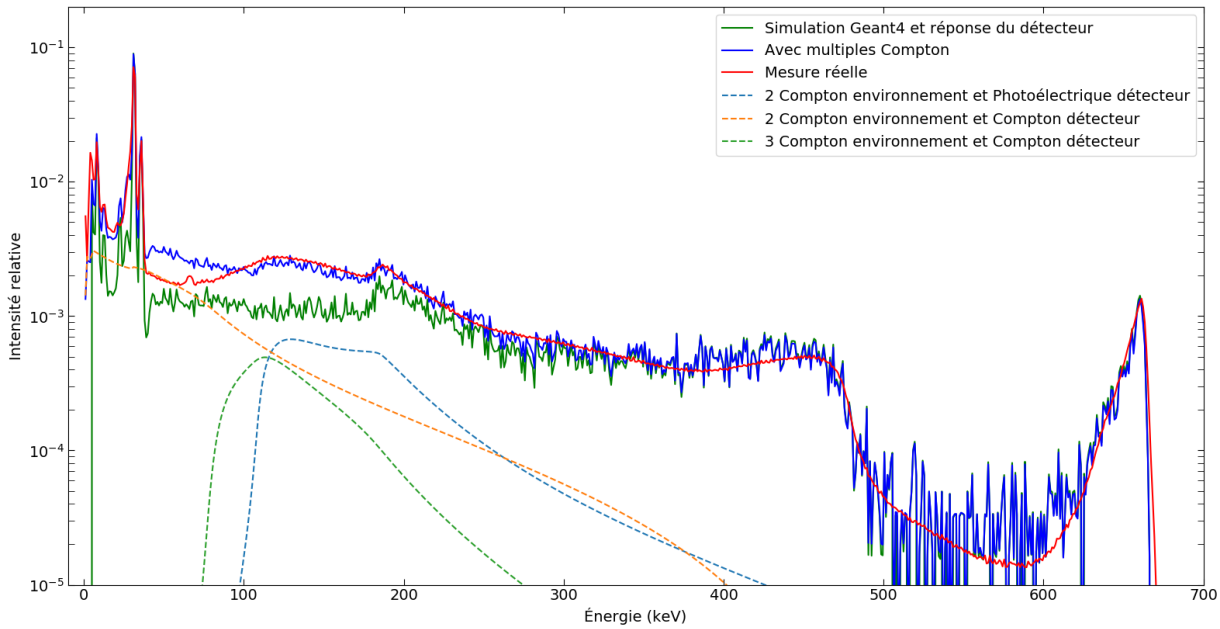


Figure 4-7 : Comparaison entre la simulation, l'ajustement avec les structures Compton multiples analytiques et les mesures réelles avec Caliste-O.
Acquisition d'une source de ^{137}Cs de 1,6 MBq placée à 8 cm du détecteur pendant 55 h. Module utilisé : Caliste-O ; haute-tension appliquée au cristal : 600 V ; température du détecteur : -30°C ; temps de *peaking* : 6,07 µs.

J'ai représenté sur la figure les structures Compton multiples utilisées ainsi que le spectre résultant en bleu. On voit que j'arrive à expliquer la zone entre 100 keV et 200 keV à partir de ces structures.

En revanche, pour la zone entre 50 keV et 100 keV, ma méthode surestime l'intensité du spectre. Cela est peut-être dû à l'hypothèse que j'avais faite lors du calcul des diffusions Compton multiples qui

considérait que les diffusions étaient isotropes et qu'il n'y avait donc pas de sélection en énergie du fait de ces diffusions. Cependant, cette hypothèse est fortement dépendante de l'environnement, et on voit ici que certaines énergies de diffusion sont certainement favorisées. Seule une simulation complète prenant en compte la scène exacte corrigerait ce point, mais cela compliquerait le calcul de manière considérable, pour un résultat qui n'apporte pas d'information déterminante au vu de l'utilisation que je souhaite faire des simulations, décrite en paragraphe 4.3 et en chapitre 3.

4.3 Conclusion et utilité des simulations

À travers ce chapitre, j'ai présenté la constitution des données que l'on enregistre avec le détecteur Caliste. J'ai montré l'enjeu de la calibration, ainsi que l'attention qui doit être portée à cette étape, et j'ai développé une méthode de calibration basée sur l'utilisation d'un algorithme génétique.

Puis, j'ai pu montrer des résultats de calibration que j'ai expliqués dans un premier temps à l'aide des processus physiques connus. Afin d'expliquer en détail la forme de ces spectres, j'ai procédé à des simulations Monte-Carlo Geant4 pour modéliser le trajet des photons de la source au détecteur, avec potentiellement des dépôts d'énergie dans l'environnement proche et immédiat du détecteur, ainsi que les interactions dans le détecteur lui-même. J'ai porté un soin particulier à modéliser analytiquement la possibilité d'avoir des environnements plus complexes dans lesquels plusieurs diffusions Compton peuvent avoir lieu.

Une fois que les photons interagissent dans le détecteur, j'ai expliqué comment le signal était produit et induit sur l'électronique de lecture ainsi que les phénomènes qui entrent en compte afin de modéliser finement la réponse du détecteur.

Enfin, j'ai pu appliquer la réponse du détecteur ainsi modélisée à la simulation Monte-Carlo précédente pour produire des spectres simulés que j'ai confrontés à des données réelles, acquises avec le détecteur Caliste-HD. J'ai aussi montré l'influence des structures Compton multiples pour expliquer des structures du spectre que l'on observe dans des environnements particuliers qui favorisent les diffusions multiples.

J'ai introduit le type de données que je vais utiliser par la suite pour répondre aux problématiques de spectro-identification de sources radioactives et de leur localisation, avec des méthodes de masque codé et d'imagerie Compton. La compréhension fine de l'origine des phénomènes est une étape fondamentale pour mettre en place des méthodes de traitement de ces données et répondre aux objectifs de ma thèse.

En particulier, l'algorithme de simulation me permettra de construire une base de données de spectres de radioéléments que je pourrai utiliser comme références pour les méthodes de spectro-identifications, et notamment comme données d'apprentissage pour les algorithmes de *deep learning* que je présente au Chapitre 3.

5 Bibliographie

ACRORAD, 2020. *Comparison of Typical Semiconductor Detector.* [Online]
Available at: https://www.acrorad.co.jp/index_en/products_en/technology/basic.html

Agostinelli, S. et al., 2003. *Geant4 - a simulation toolkit.* in: Nuclear Instruments and Methods in Physics Research A, Volume 506, Issue 3, Pages 250-303.

Biggs, F., Mendelsohn, L. & Mann, J., 1975. *Hartree-Fock Compton profiles for the elements.* in: Atomic Data and Nuclear Data Tables, Volume 16, Issue 3, Pages 201-309.

COMSOL AB, n.d. *COMSOL Multiphysics® v. 5.4..* Stockholm, Sweden.: COMSOL.

Dirks, B., 2006. *Study and modelling of the new generation Cd(Zn)Te X and gamma-ray detectors for space applications.* at: Université Paris 7 - Denis Diderot.

Dubos, S., 2015. *Nouveau spectro-imageur CdTe à très haute résolution spatiale et spectrale pour l'astronomie X et gamma.* at: Université Paris 7 - Denis Diderot ; Commissariat à l'Energie Atomique et aux Energies Alternatives (CEA); Centre National d'Etudes Spatiales (CNES).

Fano, U., 1947. *Ionization Yield of Radiations. II. The Fluctuations of the Number of Ions.* in: Phys. Rev., American Physical Society, Vol. 72, Issue 1, pp 26-29.

Geant4 Collaboration (1), 2018. [Online]
Available at: http://geant4-userdoc.web.cern.ch/geant4-userdoc/UsersGuides/PhysicsListGuide/html/electromagnetic/emphyslist.html

Geant4 Collaboration (2), 2018. [Online]
Available at: http://geant4-userdoc.web.cern.ch/geant4-userdoc/UsersGuides/PhysicsReferenceManual/html/electromagnetic/gamma_incident/photoelectric/livermore_photoelectric.html?highlight=livermore

Geant4 Collaboration (3), 2018. [Online]
Available at: http://geant4-userdoc.web.cern.ch/geant4-userdoc/UsersGuides/PhysicsReferenceManual/html/electromagnetic/gamma_incident/compton/monash_compton.html?highlight=monash

Hecht, K., 1932. *Zum Mechanismus des lichtelektrischen Primärstromes in isolierenden Kristallen.* in: Z. Physik 77, 235–245.

He, Z., 2001. *Review of the Shockley–Ramo theorem and its application in semiconductor gamma-ray detectors.* in: NIM-A, Volume 463, Issues 1–2, Pages 250-267.

Knoll, 2000. *Radiation Detection and Measurement.* in: John Wiley & Sons, 3rd edition.

Maier, D. et al., 2018. *Second generation of portable gamma camera based on Caliste CdTe hybrid technology.* in: Nuclear Instrumentation and Methods in Physics Research A, Volume 912, Pages 338-342.

Maier, D. & Limousin, O., 2016. *Energy calibration via correlation.* in: Nuclear Instruments and Methods in Physics Research A 812 43–49. arXiv:1512.04584, doi:10.1016/j.nima.2015.11.149..

Maier, D., Limousin, O. & Daniel, G., 2020. *Energy calibration via correlation using an adaptive mesh refinement.* in: EPJ Web Conf., 225 01003, ANIMMA 2019 Proceedings, DOI: https://doi.org/10.1051/epjconf/202022501003.

Meuris, A., 2009. *Etude et optimisation du plan de détection de haute énergie en Cd(Zn)Te pour la mission spatiale d'observation astronomie X et gamma SIMBOL-X.* at: Université Paris-Diderot.

Michalowska, A. et al., 2010. *IDeF-X HD : A low power multi-gain CMOS ASIC for the readout of Cd(Zn)Te detectors.* in: IEEE Nuclear Science Symposuim and Medical Imaging Conference, NSS/MIC p. 1556-1559 5874037.

Ordonez, E., Bolozdynya, A. & Chang, W., 1997. *Doppler broadening of energy spectra in Compton cameras.* in: 1997 IEEE Nuclear Science Symposium Conference Record, Albuquerque, NM, USA, pp. 1361-1365 vol.2.

Ramo, S., 1939. *Currents Induced by Electron Motion.* in: Proceedings of the IRE, vol. 27, no. 9, pp. 584-585.

Shockley, W., 1938. *Currents to Conductors Induced by a Moving Point Charge.* in: Journal of Applied Physics, Volume 9, Issue 10, p.635-636.

Whitley, L. D., 2001. *An overview of evolutionary algorithms: practical issues and common pitfalls.* in: Inf. Softw. Technol., Vol. 43, pp 817-831.

Chapitre 3 : Spectro-Identification automatique

1 Contexte de la spectro-identification _____ 96

2 Les réseaux de neurones _____ 104

3 Apprentissage de réseaux de neurones convolutifs pour la spectro-identification de radioéléments_____ 112

4 Application des réseaux de neurones à la spectro-identification de radioéléments_____ 127

5 Application à la dosimétrie avec Caliste _____ 148

6 Conclusion_____ 161

7 Bibliographie_____ 162

Dans ce chapitre, je développe la problématique de l'identification de radioéléments à partir de mesures spectroscopiques effectuées avec les détecteurs Caliste.

Dans un premier temps, je pose le problème de spectro-identification que je vais traiter et je présente les méthodes actuellement existantes pour ce problème et leurs limitations. Ensuite, je mets en place de nouvelles méthodes de traitement basées sur des algorithmes de *deep learning* par réseaux de neurones pour répondre à la problématique, dont je montre les motivations et les avantages attendus.

Je présente plus précisément comment je mets en place le problème dans le cadre du *deep learning*, à travers les données d'apprentissage utilisées, basées sur les simulations décrites au Chapitre 2, conditionnant le développement des architectures de réseaux de neurones et leur apprentissage. Enfin, je teste les performances de mes algorithmes sur des données réelles acquises avec Caliste et je montre un approfondissement de mes méthodes de traitement en les appliquant à des calculs de débit de dose par radioélément identifié.

1 Contexte de la spectro-identification

1.1 Problème à traiter

1.1.1 Description du problème

Supposons que nous avons effectué une mesure avec le détecteur Caliste, nous obtenons une liste d'événements comme présentée en Chapitre 2. À partir de cette liste d'événements, je vais répondre successivement à deux questions :

- Quels radioéléments ont émis les photons qui ont été détectés ?
- En quelles proportions chaque radioélément identifié a-t-il contribué au signal enregistré ?

Je porterai une attention particulière à la caractérisation de l'incertitude sur les réponses à ces questions qu'apporteront mes algorithmes.

La première question est une problématique dite de classification, l'objectif de sortie étant discret. Pour un radioélément donné, on répond à une question binaire : est-il présent ou non dans la scène où la mesure a été effectuée ?

La deuxième question est une problématique dite de régression, l'objectif de sortie est continu. Plus précisément, pour un radioélément donné, j'appellerai proportion le rapport du nombre de photons associés à ce radioélément sur le nombre total de photons détectés. Je cherche donc à donner un nombre continu pour chaque radioélément compris entre 0 et 1, qui vaut 0 si le radioélément n'est pas identifié par la première question.

Il est important de distinguer ces deux questions, car, nous le verrons dans l'état de l'art en paragraphe 1.2, une majorité des algorithmes existants répond à l'une ou l'autre des questions indépendamment. Or, répondre uniquement à la première question ne donne naturellement pas de quantification. Inversement, répondre uniquement à la deuxième question ne confirme pas la présence ou l'absence d'un radionucléide : si une faible proportion, par exemple de 1 %, est associée à un radioélément, cela ne permet pas de savoir si cette proportion est réellement présente ou si ces 1 % étaient nécessaires à l'algorithme pour expliquer une petite partie des données, sans que le radioélément ne soit présent.

1.1.2 Données utilisées

Pour répondre à ces questions à partir de la liste d'événements enregistrée, je représente les données sous forme de spectre en énergie, après une étape de calibration décrite en Chapitre 2. En effet, j'ai montré que les spectres sont des signatures des radioéléments et contiennent les informations essentielles liées aux lignes d'émission des sources et aux processus d'interaction photon-matière.

La représentation sous formes de spectre est la représentation classique dans le problème de la spectroscopie gamma. Nous le verrons d'ailleurs dans l'état de l'art, la très grande majorité de ces méthodes utilise cette représentation, même s'il existe aussi des exceptions avec des algorithmes qui utilisent les données photon par photon.

Pour anticiper sur les aspects de *deep learning*, cette représentation présente aussi deux avantages. D'une part, la signature spectrale présente des formes caractéristiques qui permettent d'identifier le

radioélément, ce qui est un type de tâche pour laquelle les réseaux de neurones ont été utilisés avec succès dans d'autres domaines, comme la reconnaissance d'image. D'autre part, les réseaux de neurones, qu'ils soient de type perceptron ou réseaux de neurones de convolution, comme décrits en partie 2, ont besoin que les données en entrée aient une taille fixe, ce que je peux garantir avec les spectres, en choisissant un nombre de canaux et une dynamique en énergie fixes.

La Figure 1-1 résume la problématique que je vais traiter dans ce chapitre.

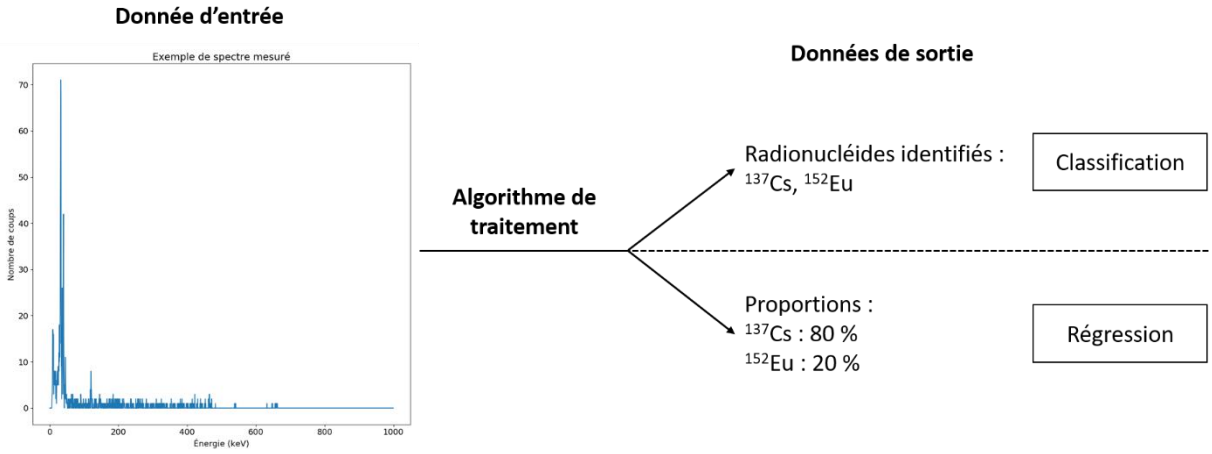


Figure 1-1 : Schéma du problème à résoudre en spectro-identification.

Par ailleurs, un des objectifs est d'être capable de quantifier l'incertitude sur les résultats renvoyés par mes algorithmes de traitement. Il faut pouvoir dire si les radioéléments sont détectés ou absents avec certitude ou non. De même, pour la quantification, les algorithmes doivent fournir une erreur associée au résultat qu'ils renvoient.

1.1.3 Hypothèses et contexte

Dans l'ensemble de cette partie, je ferai l'hypothèse que les données sont acquises avec un détecteur CdTe de type Caliste. Les méthodes que je développe dans ce chapitre sont transposables à d'autres types de détecteurs, à quelques adaptations près, notamment en ce qui concerne les données d'apprentissage.

Pour le contexte des mesures, je fais l'hypothèse que les sources mesurées se retrouvent parmi une bibliothèque de radioéléments connus et inclus dans la base de données d'apprentissage. Je reviendrai plus en détail sur cette hypothèse en paragraphe 3.1.

Les données mesurées peuvent être issues de mélanges de plusieurs radioéléments, dont le nombre peut varier de 0 si aucun radioélément n'est présent, à l'ensemble des radioéléments de la bibliothèque d'apprentissage. Cependant, selon les domaines d'application, comme la sûreté nucléaire, les portiques de contrôle ou encore l'intervention en situation accidentelle, les mélanges que l'on s'attend à retrouver peuvent contenir un nombre restreint de radioéléments d'intérêt pour un diagnostic. Si c'est le cas, les méthodes que je mets en place peuvent faire l'objet d'un apprentissage spécifique à chaque type de situation attendue par l'adaptation de la bibliothèque de radioéléments.

De même, je ne fais pas d'hypothèse particulière sur l'environnement du détecteur ou des sources. Des éléments absorbeurs ou diffuseurs peuvent être présents entre la source et le détecteur, induisant

d'éventuelles modifications dans les spectres mesurés. Cependant, il faut noter que l'apprentissage est effectué avec des données représentant un contexte particulier. Même si les méthodes ont une capacité de généralisation par rapport aux données d'apprentissage, elles seront plus particulièrement efficaces si on effectue un apprentissage spécifique par rapport à la situation attendue.

Enfin, je ne fais pas non plus d'hypothèse particulière sur le temps d'acquisition. Les temps de pose peuvent être très faibles avec seulement une dizaine de photons détectés au minimum, ou très longs avec plusieurs millions de photons détectés. J'étudie d'ailleurs l'influence du nombre de photons utilisés sur les résultats de mes méthodes en paragraphe 4.3.

1.2 État de l'art

Dans ce paragraphe, je fais un tour d'horizon des méthodes existantes pour répondre aux problématiques de spectro-identification. À travers cela, je pourrai identifier les limites de ces méthodes et donner les motivations qui m'ont entraîné à choisir des algorithmes de *deep learning* pour aborder ce problème.

1.2.1 Algorithmes de recherche et d'ajustements de pics

Les premières méthodes développées pour répondre à ce type de problématique consistent à rechercher les pics photoélectriques et à ajuster un modèle pour en trouver la forme. Cette classe d'algorithme a d'abord été développée pour des spectres à très haute résolution issus de détecteurs germanium HPGe (*High Purity Germanium*) (Phillips & Marlow, 1976). Leurs performances permettent d'obtenir des pics photoélectriques avec une résolution allant de 1 keV à un peu plus de 2 keV sur une dynamique en énergie de quelques keV à plus de 1 MeV (Szymańska, et al., 2008). Ces résolutions, combinées à l'efficacité de détection de ces détecteurs, donnent des pics photoélectriques intenses et précisément positionnés, d'où l'intérêt des algorithmes de recherche de pics et d'ajustements de pics.

Ces méthodes nécessitent que les pics photoélectriques apparaissent clairement, c'est-à-dire, avec un nombre de photons supérieur à $N_{\text{fond}} + \sqrt{N_{\text{fond}}}$ où N_{fond} est le nombre de photons dans le spectre de fond, qui provient du bruit de fond naturel et des continuums Compton. Cette condition est plus facilement assurée avec des spectres HPGe, pour lesquels la haute résolution spectroscopique permet d'accumuler plus efficacement les événements dans les canaux d'intérêt correspondant aux pics.

Ce type de méthode a été tout de même transposé à des spectres issus de détecteurs CdTe et CdZnTe (Gunnink & Arlt, 2001), dans lesquels il est nécessaire de porter une attention particulière à la forme des pics et notamment aux effets de perte de charge (cf. Chapitre 2). Les algorithmes d'ajustements de pics sont alors utilisés dans des cas très spécifiques où l'on cherche à effectuer des quantifications très précises, comme montré en Figure 1-2, pour déterminer l'enrichissement en Uranium à partir de ces spectres (Meleshenkovskii, et al., 2019). Les pics étant ajustés, le but est ensuite de calculer le ratio en intensité de chacun des pics afin de déterminer la composition isotopique de l'échantillon. Cependant, dans ce type de méthode, il faut un temps de pose suffisamment long pour que les pics apparaissent clairement, typiquement d'au moins 2 h pour (Meleshenkovskii, et al., 2019).

Certains algorithmes sont aussi conçus pour des applications très ciblées, dans lesquelles les radioéléments sont connus et dont le but est de faire de la quantification précise de fraction de masse isotopique (Simon, et al., 2009). Ces algorithmes se basent sur l'extraction de régions d'intérêts ROI (*Region Of Interest*), déterminées par le concepteur en fonction de l'application cible. Ils reposent

principalement sur le calcul de l'aire des pics, correspondant au nombre de coups dans le pic. Cela nécessite cependant un nombre de coups assez élevé pour chaque pic, au moins quelques dizaines de milliers, afin d'atteindre la précision voulue. Un algorithme d'inversion est ensuite mis en place pour déterminer les fractions de masse de chaque isotope en fonction de l'aire des pics calculée précédemment et de l'efficacité des détecteurs, supposée connue.

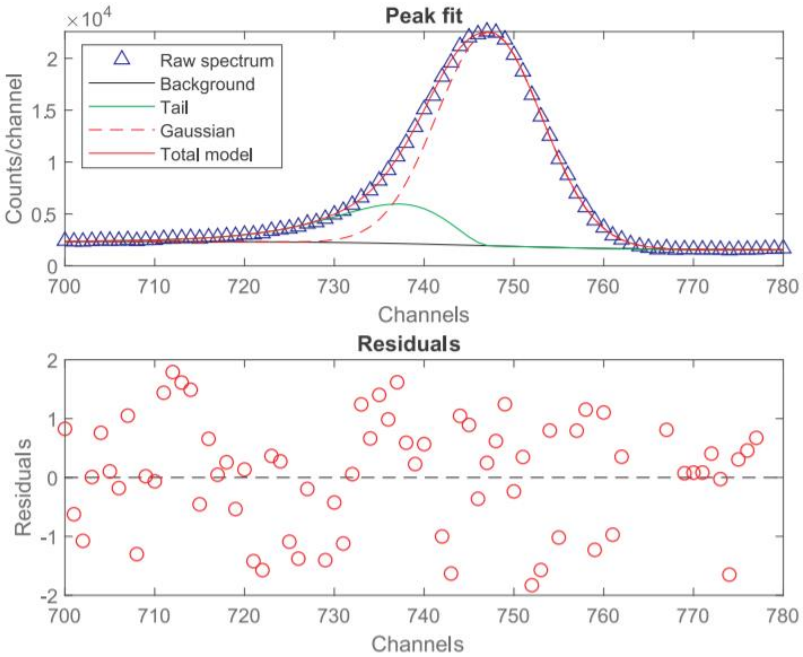


Figure 1-2 : Algorithme de *peak fitting* décrit par (Meleshenkovskii, et al., 2019) appliqué à la raie à 186 keV du ^{235}U, pour un spectre acquis avec un détecteur CZT de 500 mm³ pendant 40 h.

1.2.2 Méthodes basées sur des modèles de spectres

Dans les cas d'utilisation de détecteurs à plus basse résolution mais à grande efficacité de détection, comme les détecteurs NaI, les méthodes précédentes sont plus difficiles à appliquer du fait de la plus faible résolution, qui ne permet pas de localiser précisément les pics. Certaines méthodes proposent de passer par une transformation des spectres, notamment en utilisant des ondelettes ou *wavelets* en anglais (Sullivan & Lu, 2013), afin ensuite de mesurer précisément la position et l'aire de chaque pic. Ces méthodes peuvent être combinées avec un algorithme bayésien afin d'identifier les radioisotopes à partir d'une bibliothèque de radioisotopes (Sullivan & Stinnett, 2015), en maximisant la vraisemblance d'observer les données, c'est-à-dire, les pics avec leurs positions et leurs aires, en fonction d'un modèle de spectre des radioisotopes de la bibliothèque de sources utilisée.

Un autre algorithme itératif a été développé (Bobin, et al., 2016), basé sur une application de la formule de Bayes pour des applications avec des spectres avec un nombre limité de photons, à partir de 50 photons au total dans le spectre. Cet algorithme consiste à mettre à jour un *prior* $p(k)$ pour chaque radionucléide k recherché. Ce *prior* correspond à la proportion de ce radionucléide dans le spectre. Itérativement sur chaque photon j de la liste d'événement, le *prior* est mis à jour en calculant le *posterior* $p(k|j)$ à partir de la formule de Bayes. La formule de Bayes utilise le précédent *prior* $p(k)$ et la vraisemblance $p(j|k)$ qui correspond la probabilité d'observer l'énergie détectée du photon E_j avec le radionucléide k. Cette vraisemblance est en fait proportionnelle au spectre du radionucléide k. Il est donc nécessaire d'avoir un modèle précis du spectre attendu du radionucléide k avec le détecteur

utilisé, ce qui implique qu'il est plus difficile de prendre en compte d'éventuelles dérives de calibration ou de modifications du spectre liées à l'environnement par rapport au modèle. J'explique plus en détail une adaptation que j'ai faite de cet algorithme en paragraphe 5.2.2 pour le redressement des diffusions Compton dans les spectres ainsi que dans le cadre de l'imagerie Compton en Chapitre 6.

Il existe aussi des développements basés sur des algorithmes génétiques (Carlevaro, et al., 2008), dont le but est de chercher la combinaison de radionucléides qui explique au mieux la mesure. Pour ce faire, on dispose d'un modèle de spectre pour chaque radionucléide. L'objectif est de trouver la combinaison linéaire à appliquer sur ces modèles pour retrouver le spectre de la mesure, comme montré en Figure 1-3. On applique un algorithme génétique pour la recherche de ces coefficients, en évaluant leur score à partir de l'écart entre la combinaison linéaire obtenue et le spectre, avec la formule des moindres carrés.

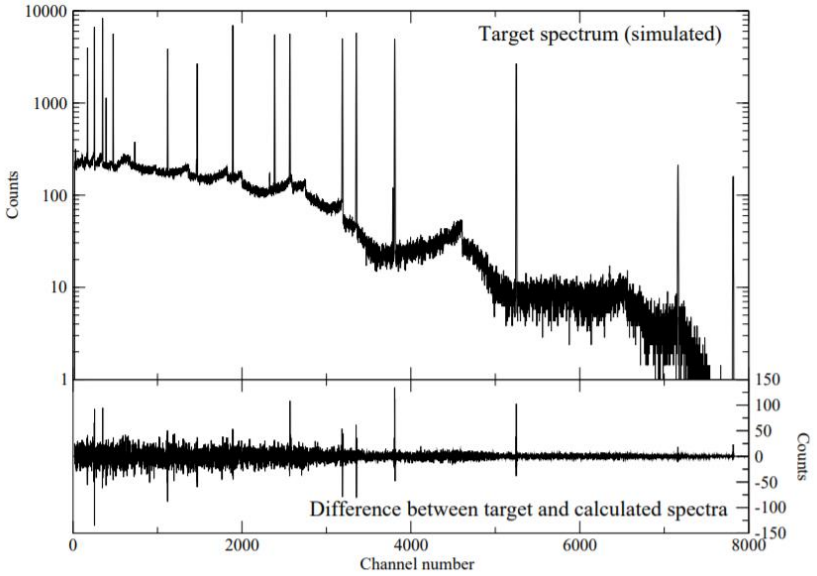


Figure 1-3 : Exemple de résultat d'application de l'algorithme génétique développé par (Carlevaro, et al., 2008) sur un spectre d'un mélange de 11 radioéléments issu d'une simulation d'un détecteur HPGe(^{241}Am, ^{109}Cd, ^{139}Ce, ^{57}Co, ^{60}Co, ^{137}Cs, ^{54}Mn, ^{113}Sn, ^{85}Sr, ^{88}Y, ^{65}Zn). En haut : résultat du spectre issu de l'algorithme génétique. En bas : différence entre le spectre de l'algorithme génétique et le spectre de la simulation Monte-Carlo.

1.2.3 Algorithmes de machine learning

Devenus incontournables à partir des années 1990, des algorithmes de *machine learning* ont été développés dans le domaine de la spectrométrie gamma. Le principe de ces algorithmes repose sur la recherche de caractéristiques pertinentes parmi des exemples qui leur sont présentés, afin de répondre à un objectif donné. Je détaille le principe de ces algorithmes en partie 2 dans le cadre des réseaux de neurones. Les applications en ingénierie nucléaire sont résumées par (Gomez-Fernandez, et al., 2020), dont un focus est fait sur la spectrométrie gamma.

On distingue deux types d'algorithmes de *machine learning*. D'une part, les algorithmes d'apprentissage supervisé sont une classe d'algorithme pour lesquels les exemples doivent comporter des données d'entrée et des données de sortie. Par exemple, dans le cadre de la spectro-identification, on utilisera souvent des données spectrales en donnée d'entrée et le nom du radioélément correspondant en donnée de sortie. D'autre part, il existe des algorithmes d'apprentissage non

supervisé pour lesquels les données de sortie ne sont pas connues. Cette classe d'algorithme est par exemple utilisée pour regrouper les données d'entrée en *clusters* en fonction de leurs caractéristiques communes, et un utilisateur expert peut ensuite annoter ces *clusters* en fonction de ses connaissances du problème. Dans le cadre de la spectroscopie gamma, un algorithme non supervisé par analyse en composantes principales ou PCA (*Principal Component Analysis*) (Boardman, et al., 2012) a été proposé. Il consiste à projeter les spectres suivant une matrice de vecteurs propres, préalablement calculée à partir d'une base de données, de sorte que le résultat de la projection contienne les informations discriminantes sur ses premières composantes. Les données sont ensuite regroupées en *clusters* suivant le radioélément qu'elles représentent. Lorsqu'une nouvelle donnée est testée, la distance à chaque cluster est calculée et le radioélément identifié est celui qui correspond au cluster situé à la plus petite distance. Il est à noter que, dans la très grande majorité des cas pour la spectroscopie gamma, les données de sortie sont connues puisque les exemples proviennent de simulations ou de données de laboratoire. En général, ce sont plutôt des algorithmes d'apprentissage supervisé qui sont utilisés, ce que je développe dans la suite de ce paragraphe.

Un algorithme de *machine learning* supervisé a par exemple été appliqué afin de rechercher automatiquement les régions d'intérêt dans le spectre et d'inférer des règles sur les intensités de ces régions pour identifier le ou les radionucléides (Abdel-Aal & Al-Haddad, 1997). Cet algorithme a principalement été développé pour des questions de rapidité de calcul. En effet, une fois l'apprentissage effectué, il n'est plus nécessaire d'appliquer des opérations de recherche de pics ou d'ajuster un modèle pour obtenir le résultat de l'identification. Cependant, il nécessite toujours un nombre de photons élevé, de l'ordre de plusieurs milliers à plusieurs dizaines de milliers.

Par ailleurs, à partir des années 1990, les méthodes de *deep learning* basées sur des réseaux de neurones de type perceptron (voir description en partie 2), ont émergé dans différents domaines et on peut voir leur utilisation dans le cadre de la spectro-identification. Le premier réseau de neurones, à une seule couche, a été développé pour l'analyse de spectres issus de détecteurs NaI et Germanium (Olmos, et al., 1991), mais nécessite des acquisitions de plus de 3 minutes, pour obtenir plusieurs dizaines de milliers de photons dans le spectre. De même, on peut citer des travaux plus spécifiques se concentrant sur des régions d'intérêt comme par exemple pour quantifier l'enrichissement en Uranium (Vigneron, et al., 1995), avec extraction préalable de pics (Yoshida, et al., 2002) (Medhat, 2012) ou encore après certaines transformations comme la réduction des spectres par décomposition en valeurs singulières ou SVD (*Singular Values Decomposition*) (Pilato, et al., 1999) ou par transformation K-L (Chen & Wei, 2009). Tous ces exemples ont en commun le fait que les réseaux de neurones utilisés n'ont aucune ou une seule couche cachée, limitation qui provenait à l'époque de la puissance de calcul des ordinateurs qui ne permettait pas le traitement de réseaux de neurones de plus grande taille. De ce fait, la réduction des données d'entrée par les techniques évoquées précédemment était inévitable.

Ces dernières années, le *deep learning* a connu une renaissance avec l'amélioration des performances de calcul des ordinateurs et l'explosion des données disponibles avec l'essor du *big data*. Les réseaux de neurones à plusieurs couches cachées ont été appliqués aux problématiques de spectro-identification de radioéléments. D'abord, avec deux couches (Kamuda, et al., 2017) ou trois couches cachées (Kim, et al., 2019), les réseaux de neurones permettent de résoudre la problématique de spectro-identification pour des spectres constitués de quelques centaines de photons acquis avec un détecteur NaI ou un scintillateur plastique à faible résolution, dont un exemple d'acquisition est illustré

en Figure 1-4. Cependant, du fait de l'utilisation de la fonction d'activation *softmax* sur la dernière couche du réseau, il n'est possible de ne détecter qu'un nombre limité de sources ou alors, il est nécessaire de prendre en compte toutes les combinaisons possibles d'un mélange de radioéléments, soit 2^N combinaisons pour N radioéléments dans la bibliothèque (Kim, et al., 2019). Cette limitation ne provient pas intrinsèquement de l'utilisation des réseaux de neurones, elle est due à la conception des réseaux de neurones propres à ces études. Il est intéressant de noter que dans ces différents exemples, aucun prétraitement, autre qu'une normalisation, n'est appliqué au spectre.

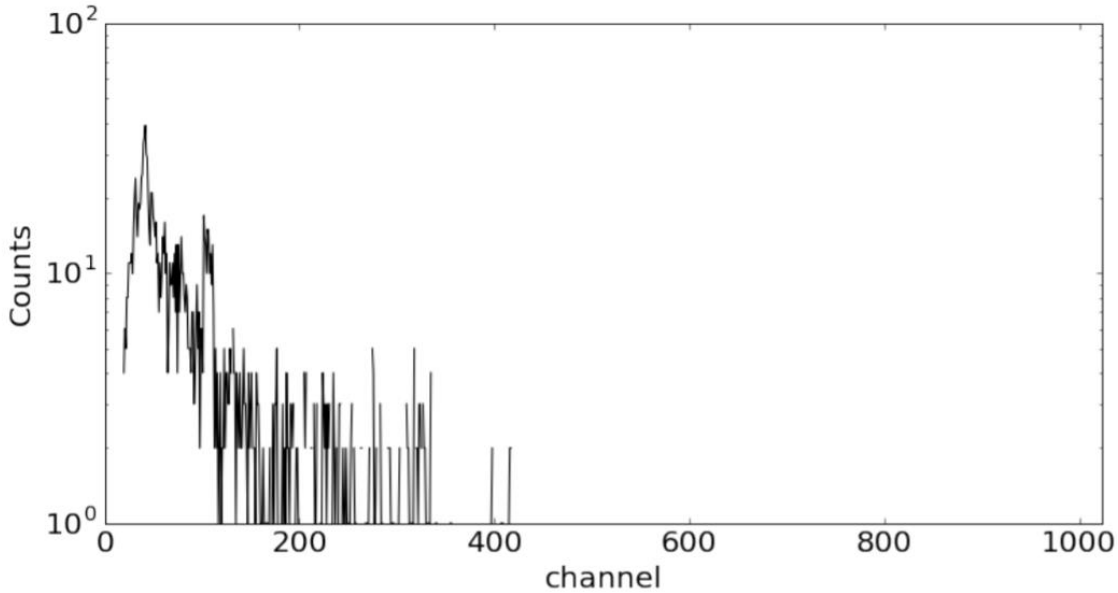


Figure 1-4 : Exemple de spectre utilisé par (Kendall & Gal, 2017) : source de ^{152}Eu de 8,85 µCi située à une distance de 74 cm d'un détecteur NaI, acquisition de 10 secondes.

Par ailleurs, les réseaux de neurones à couches de convolution sont intéressants car ils appliquent un filtrage du spectre sur la base d'un apprentissage, suivi d'un réseau de neurones de type perceptron pour procéder à l'identification. Cette opération peut être réalisée directement en dimension 1, sur le spectre lui-même (Kamuda, et al., 2020) ou après transformation en dimension 2, par transformée de Hilbert par exemple (Liang, et al., 2019) comme montré en Figure 1-5. Il est à noter que ces dernières références concernent des travaux contemporains à mes recherches.

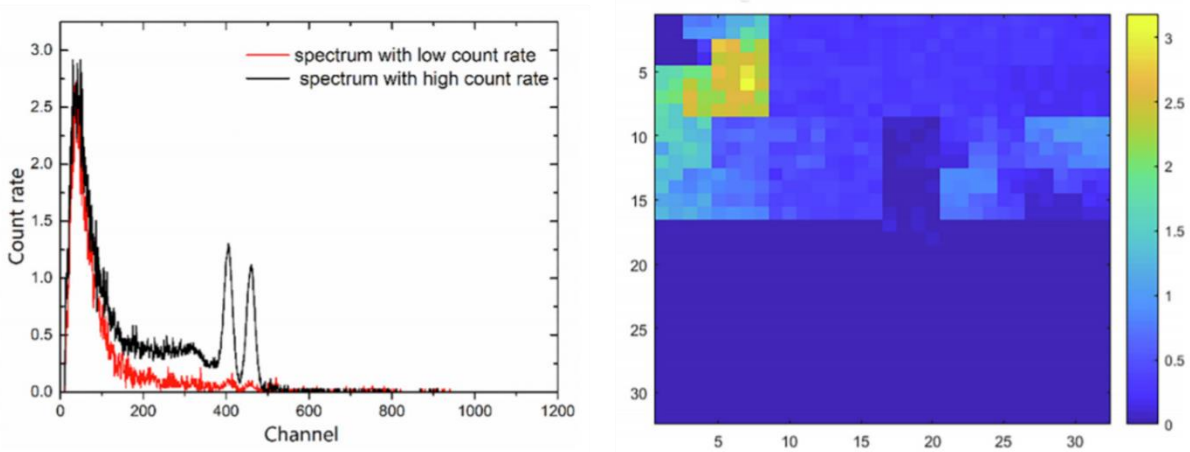


Figure 1-5 : À gauche : exemple d'une simulation de spectre de ^{60}Co par un code Monte-Carlo N-Particle transport (MCNP) pour un détecteur NaI(Tl). À droite : transformation d'un spectre de ^{60}Co en 2 dimensions par courbe de Hilbert sur laquelle est ensuite appliqué un réseau de neurones convolutif 2D (Liang, et al., 2019).

1.3 Orientation vers des algorithmes de *deep learning*

À travers cette revue bibliographique, on constate que trois grandes catégories d'algorithmes ont été développées afin de répondre aux problèmes de l'analyse de spectres gamma. Tout d'abord, les algorithmes basés sur la recherche et l'ajustement de pics nécessitent des détecteurs très bien résolus en énergie ainsi qu'un nombre de photons suffisant pour que les pics principaux des radioéléments apparaissent.

Ensuite, les algorithmes basés sur l'utilisation de modèles de spectres, avec des méthodes bayésiennes ou des méthodes d'ajustement de modèles, nécessitent l'utilisation de modèles fidèles aux données réelles attendues. Ils sont sensibles à des problématiques de modification de la calibration, que j'appelle par la suite « décalibration », mais aussi à la présence d'artéfacts dans le spectre que les algorithmes chercheront à compenser par exemple en utilisant une combinaison linéaire de radioéléments qui ne sont pourtant pas présents dans la scène.

Enfin, les méthodes les plus récemment utilisées sont presque exclusivement des algorithmes basés sur du *machine learning*, en particulier du *deep learning* avec des réseaux de neurones. Je vais utiliser et appliquer ces méthodes à nos données notamment dans l'objectif de relever plusieurs défis.

À l'aide de ce type d'algorithme combiné une base de données d'apprentissage adaptée, je vise à être capable de détecter des sources à partir d'un faible nombre de photons, au moins quelques dizaines de photons, et avec potentiellement des décalibrations appliquées aux spectres. Ces décalibrations peuvent être dues aux conditions opérationnelles, notamment liées à la température du détecteur ou encore à la haute-tension appliquée au cristal. Un autre exemple est l'application continue de la haute tension au cristal pendant une longue période, sur plusieurs heures, qui crée un effet de dérive de calibration (Meuris, 2009) appelé effet de polarisation.

Les capacités des algorithmes de *deep learning* sur les spectres avec peu de photons ainsi que le temps de calcul sont un avantage majeur attendu pour le traitement automatique en temps réel des événements. En effet, bien que l'apprentissage soit en général assez long, allant de plusieurs heures à plusieurs jours de calcul, une fois entraînés, les réseaux de neurones nécessitent des performances de calcul moindres et on peut en particulier envisager de les intégrer dans des systèmes embarqués.

Je vais aussi créer une base de données issue de spectres simulés, dont j'ai décrit la méthode de simulation au chapitre 2, et tester la réponse de mes algorithmes avec des spectres issus de données réelles prises avec Caliste. Le nombre de sources présentes dans les mélanges testés ne devra pas non plus être un facteur limitant pour les performances de l'algorithme.

J'utiliserai mes algorithmes à la fois pour faire de la classification, détecter la présence ou l'absence de sources, et de la quantification, c'est-à-dire, donner la proportion de chaque radioélément détecté, là où pour l'instant les algorithmes dans l'état de l'art ne font que l'une ou l'autre des tâches. Je donnerai aussi des moyens de quantifier l'erreur sur le résultat.

Enfin, la souplesse des algorithmes basés sur le *deep learning* permet de prendre en compte un grand nombre de situations, comme des cas où la source est située derrière un blindage absorbant ou derrière des matériaux diffusants. En effet, la base d'apprentissage peut être adaptée compte tenu de l'application que l'on souhaite effectuer, si on connaît les conditions qui sont susceptibles d'être rencontrées.

Dans mes recherches, je me baserai principalement sur les réseaux de neurones avec des couches de convolution, qui présentent l'intérêt d'effectuer le prétraitement des spectres par apprentissage, en extrayant les caractéristiques pertinentes dans les spectres pour l'identification des radioéléments. J'expliquerai en paragraphe 4.2.2 les avantages de ces réseaux par rapport aux perceptrons.

2 Les réseaux de neurones

Dans cette partie, je présente les réseaux de neurones de manière générale et notamment les calculs effectués dans ces algorithmes. Il est important de comprendre comment ils fonctionnent et comment ils apprennent avant de les appliquer à la problématique de la spectro-identification, et je présente notamment dans cette partie les notions que j'utiliserai par la suite. Un réseau de neurones d'une architecture donnée définit une famille de fonctions paramétriques, pour laquelle je décris d'abord la manière dont le calcul est effectué pour un jeu de paramètres donné puis je présente la méthode mise en œuvre pour déterminer les paramètres du réseau.

2.1 Calcul de la sortie du réseau : *forward pass*

Le calcul dit *forward pass* est l'opération effectuée par un réseau de neurones pour obtenir, sur un exemple donné, une donnée de sortie $\hat{Y} \in \mathbb{R}^{n_y}$, représentée sous forme d'un vecteur $\hat{Y} = (\hat{y}_i)_i$ de taille n_y, à partir d'une donnée d'entrée $X \in \mathbb{R}^{n_0}$ représentée sous forme d'un vecteur $X = (x_i)_i$ de taille n_0. On parle de calcul *forward pass* dans le sens où on part de l'entrée et on propage l'information « vers l'avant » pour obtenir la sortie.

2.1.1 Neurone fully-connected

Un neurone dit *fully-connected* est la brique élémentaire qui constitue un réseau de neurones de type perceptron. Considérons un neurone j d'une couche l du réseau de neurones. Je note $A^{[l-1]} \in \mathbb{R}^{n_{l-1}}$ le vecteur constitué des résultats de la couche précédente $l-1$. Ces résultats sont en fait les sorties des n_{l-1} neurones de la couche $l-1$, chaque neurone de cette couche contenant un nombre réel. Pour la première couche, le vecteur $A^{[0]}$ considéré est en fait le vecteur des données d'entrée : $A^{[0]} = X$.

Le neurone j de la couche l dont je détaille le calcul possède les paramètres suivants :

- un vecteur de poids $W_j^{[l]}$ de la même dimension que le vecteur $A^{[l-1]}$;
- un nombre réel $b_j^{[l]}$ appelé biais.

Dans mes notations, je note entre crochets le numéro de la couche considérée, et en indice, le numéro du neurone considéré.

Ces paramètres sont les valeurs que le réseau va chercher à adapter lors de la phase d'apprentissage pour répondre à l'objectif qui lui est donné.

Le calcul effectué par le neurone j, soit la composante $a_j^{[l]}$ du neurone $A^{[l]}$, est le suivant :

$$a_j^{[l]} = f\left(z_j^{[l]} = {W_j^{[l]}}^T . A^{[l-1]} + b\right) \tag{2-1}$$

${W_j^{[l]}}^T . A^{[l-1]} = \sum_{j'} w_{j,j'}^{[l]} . a_{j'}^{[l-1]}$ est le produit scalaire entre $W_j^{[l]}$ et $A^{[l-1]}$. $a_j^{[l]}$ est la sortie du neurone, c'est par construction ici un nombre réel. $z_j^{[l]}$ est le résultat du neurone avant l'application de la fonction f.

f est appelée la fonction d'activation. C'est une fonction qui est, en général, choisie non linéaire, de sorte que le réseau de neurones n'effectue pas uniquement une suite de calculs linéaires ou affines. Je donne ici deux exemples de fonction d'activation que j'utiliserai par la suite :

- la fonction ReLU (*Rectified Linear Unit*), définie par $\mathrm{ReLU}(x) = \max(0, x)$, c'est une fonction qui vaut 0 sur les nombres négatifs et $y = x$ sur les nombres positifs. Cette fonction est principalement utilisée sur les couches intermédiaires des réseaux de neurones, du fait de la facilité à calculer sa dérivée (0 sur les nombres négatifs et 1 sur les nombres positifs). Un autre intérêt est que la valeur de sa dérivée est constante, donc elle ne peut pas prendre de trop petites valeurs ou de trop grandes valeurs. Cela prévient des effets d'explosion ou d'évanescence du gradient que j'explique en paragraphe 2.2.2 ;
- la fonction *softmax*, définie sur les sorties $z_j^{[l]}$ par $\mathrm{softmax}\left(z_j^{[l]}\right) = \frac{e^{z_j{}^{[l]}}}{\sum_{j'} e^{z_{j'}{}^{[l]}}}$ est une généralisation de la fonction logistique. Par définition, les sorties des neurones après application de la fonction *softmax* sont positives et ont une somme qui vaut 1, on peut alors les assimiler à une distribution de probabilités.

En fonction de la couche concernée et du résultat souhaité, il faudra choisir parmi différentes fonctions d'activation. Lorsque ce sera le cas, je présenterai les motivations qui m'ont conduit à utiliser l'une ou l'autre de ces fonctions.

La Figure 2-1 résume le calcul effectué par un neurone de type *fully-connected*.

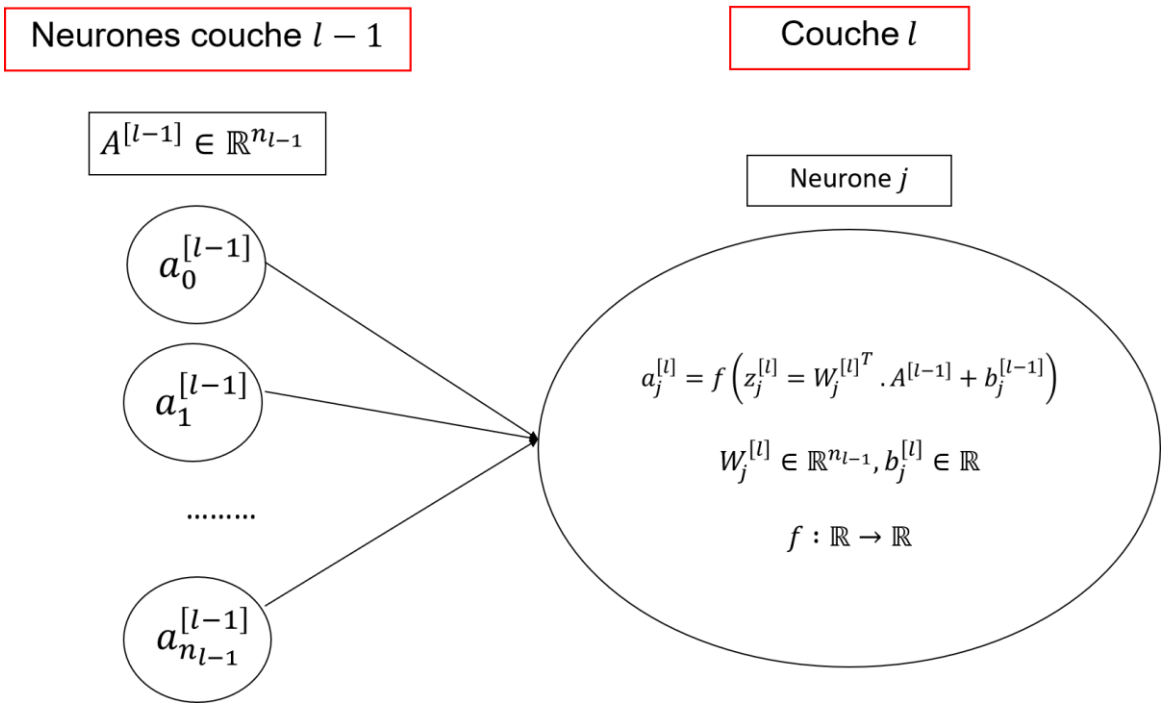


Figure 2-1 : Schéma du calcul d'un neurone *fully-connected*. Le calcul intermédiaire $z_j^{[l]}$ est utilisé en Annexe B pour décrire l'algorithme de *backpropagation* pour l'apprentissage.

2.1.2 Neurone de convolution

Depuis le début des années 1990, les réseaux de neurones de convolution (Aloysius & Geetha, 2017) ont permis d'effectuer améliorations considérables dans le cadre du *deep learning*, et en particulier pour l'analyse d'images.

Le principe de ces réseaux de neurones est l'utilisation d'opérations de convolutions à la place des opérations de produits scalaires d'un neurone *fully-connected*.

On note A un tableau $n \times m \times c$, K un filtre $n_K \times m_K \times c$ avec $n_K \leq n$, $m_K \leq m$. La convolution de A par K pour un pixel de coordonnées (i, j), $i \in [\![0, n-1]\!]$, $j \in [\![0, m-1]\!]$ est définie par :

$$(A \star K)[i,j] = \sum_{c'=1}^{c} \sum_{k=-\frac{n_K}{2}}^{\frac{n_K}{2}} \sum_{l=-\frac{m_K}{2}}^{\frac{m_K}{2}} A[i+k, j+l, c'] \times K\left[k + \frac{n_K}{2}, l + \frac{m_K}{2}, c'\right] \tag{2-2}$$

Ici, je considère un tableau A à 3 dimensions. On peut se représenter les deux premières dimensions comme une image si $m \geq 2$ ou un vecteur à une dimension si $m = 1$. Dans ce chapitre, on considère par exemple un spectre donc un vecteur à une dimension.

La troisième dimension c est le nombre de canaux. Par exemple, si l'image d'entrée est une image RGB, $c = 3$. Dans le cas des couches intermédiaires, c représente le nombre de neurones de la couche précédente.

Pour considérer les effets de bord, si $i + k$ ou $j + l$ sort de l'image A (c'est-à-dire, $i + k \geq n$ par exemple), alors on peut prendre $A[i + k, j + l] = 0$, on parle de « *zero-padding* », de sorte qu'à la fin, le résultat de la convolution ait la même dimension que A. D'autres solutions de *padding* sont possibles, ou on peut aussi choisir de ne pas en appliquer, mais dans ce cas, la sortie du réseau aura une dimension réduite par rapport à A. J'ai choisi de mon côté que toutes mes opérations de convolution se fassent avec « *zero-padding* » afin de conserver les dimensions du spectre.

Le résultat à la suite de l'opération de convolution est donc un tableau de taille $n \times m$ ou éventuellement de taille réduite dans le cas sans *padding*. On applique ensuite à ce résultat une fonction d'activation, comme celles définies au paragraphe 2.1.1, élément par élément sur $A \star K$.

Tout comme les paramètres W et b d'un neurone *fully-connected*, les valeurs contenues dans le filtre K sont des paramètres qui seront ajustés lors de la phase d'apprentissage. Un intérêt majeur de ce type d'opération est que le nombre de paramètres à entraîner est bien plus faible par rapport à une couche *fully-connected* : les filtres K comportent typiquement de l'ordre de quelques dizaines de paramètres, en comparaison des poids W qui contiennent typiquement plusieurs centaines de paramètres. Ainsi, ces opérations permettent d'accélérer l'apprentissage et d'éviter des phénomènes d'*overfitting* du fait de l'apprentissage d'un trop grand nombre de paramètres. En outre, ces opérations présentent l'intérêt de rechercher des structures locales, invariantes par translation dans le signal.

La Figure 2-2 schématise le calcul d'un neurone de convolution, en reprenant les indices de notation introduits au paragraphe précédent. Notons qu'un biais b peut aussi être ajouté à la suite de l'opération de convolution, de manière similaire au cas du neurone *fully-connected*.

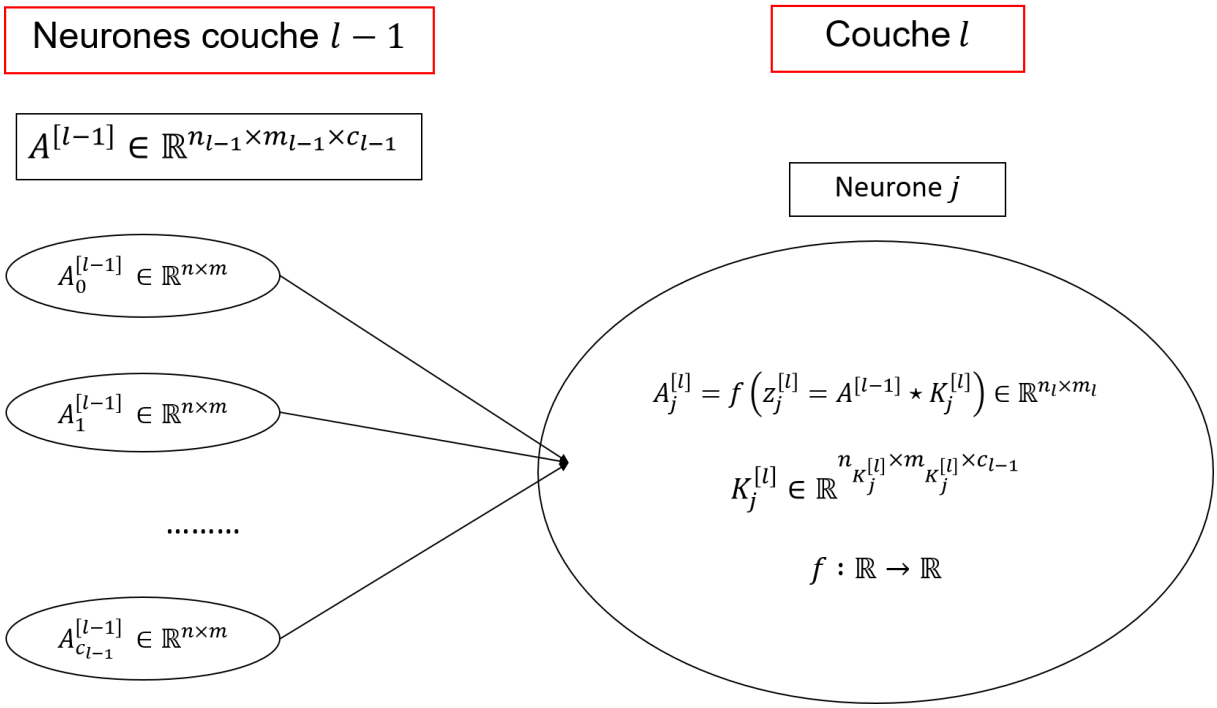


Figure 2-2 : Calcul d'un neurone de convolution.

2.1.3 Réseau de neurones

Le principe d'un réseau de neurones est de combiner ces briques élémentaires. Une couche peut contenir plusieurs neurones, des neurones *fully-connected* ou des neurones de convolution. Il existe d'autres types de neurones, comme les neurones récurrents, que je n'ai pas introduits ici car je ne les applique à ma problématique. Classiquement, une couche ne contient qu'un seul type de neurones. Chaque neurone a ses propres paramètres, poids W et biais b ou filtre K.

Ensuite, les couches sont enchaînées, toujours en effectuant les calculs décrits précédemment, les dimensions des paramètres de chaque couche étant adaptées aux dimensions de la couche précédente. Dans le cas d'un réseau de neurones de type perceptron multicouches, toutes les couches sont *fully-connected* et dans le cadre d'un réseau de neurones à couches de convolution, les premières couches sont des couches de convolution et les dernières couches sont en général des couches *fully-connected*. Certains réseaux de neurones sont purement convolutifs pour des applications particulières dans le domaine de l'imagerie, comme le débruitage ou la segmentation d'images.

Les Figure 2-3 et Figure 2-4 schématisent les structures que l'on retrouve généralement pour les réseaux de neurones de type perceptron et de convolution respectivement.

On parle de réseaux de neurones profonds lorsqu'ils contiennent beaucoup de couches intermédiaires, aussi appelées couches cachées, en général avec plus de deux couches cachées. La dernière couche est appelée couche de sortie et contient la prédiction $\hat{Y}$ du réseau de neurones.

L'aspect « boîte noire » des réseaux de neurones ne provient pas des calculs qui sont effectués par le réseau, car mathématiquement et de manière élémentaire, on est capable de comprendre le calcul étape par étape avec les formules que j'ai décrites dans ce paragraphe. En réalité, cela provient du fait que la structure devient très complexe lorsque le réseau contient beaucoup de couches intermédiaires, au point que le comportement du réseau devient difficilement interprétable.

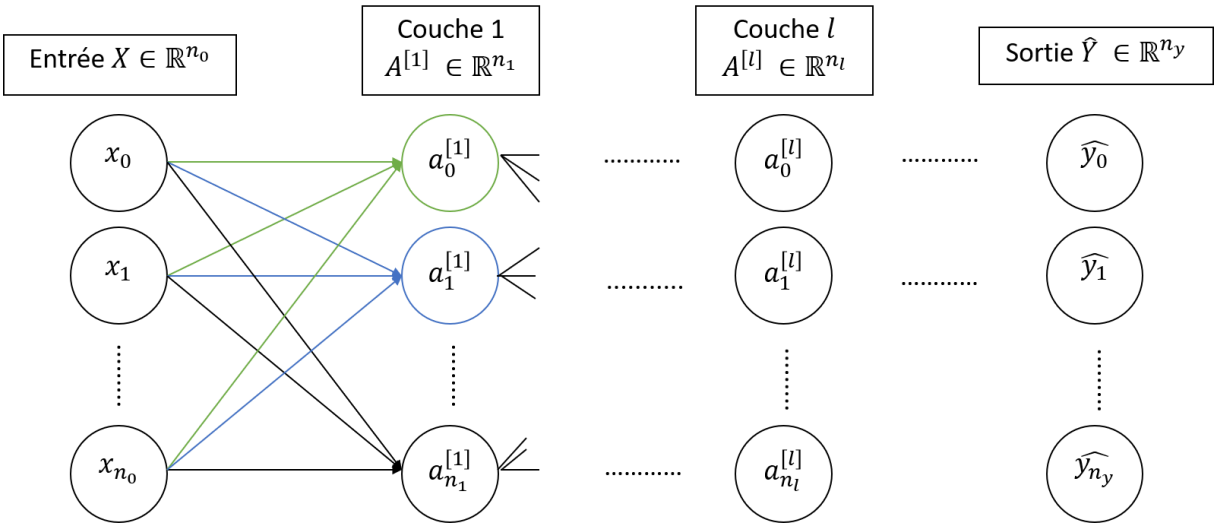


Figure 2-3 : Structure typique d'un réseau de neurones de type perceptron.

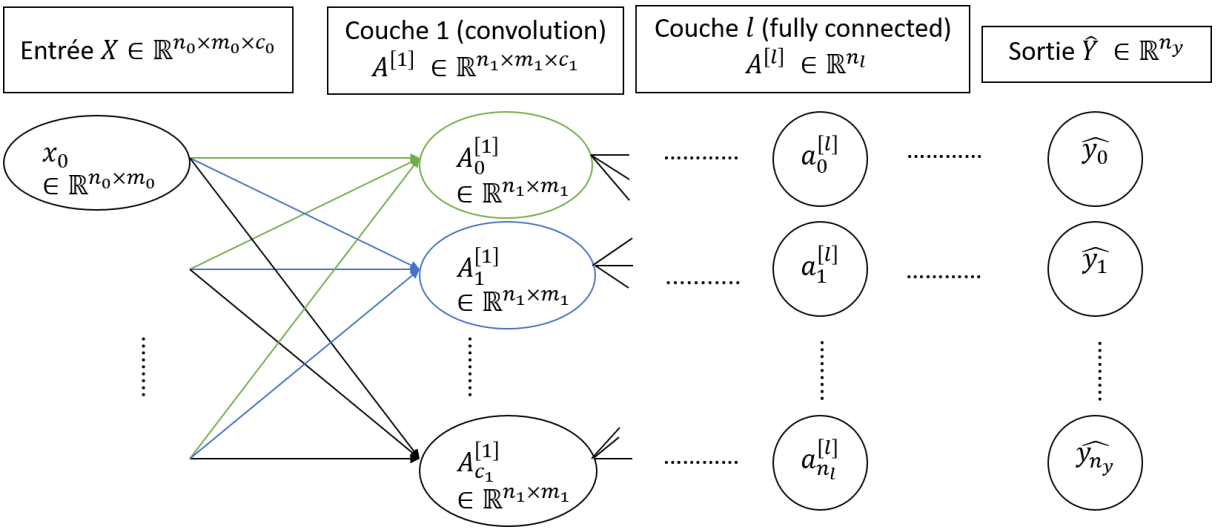


Figure 2-4 : Structure typique d'un réseau de neurones à couches de convolution.

Je souligne aussi que des couches spéciales peuvent être utilisées. En particulier, dans le cas des réseaux de convolution, on peut utiliser des couches dites de « *pooling* », qui réduisent la dimension des données. Par exemple, un *max-pooling* de taille 2 sur un vecteur A consiste à prendre toutes les deux valeurs successives du vecteur A et à ne conserver que la plus grande des deux. Cela réduit la dimension de A d'un facteur 2, ce qui donne un premier intérêt de ce type d'opération en manipulant des objets avec des dimensions réduites et ainsi nécessitant des performances de calcul moindres. Par ailleurs, ces opérations appliquées au fur et à mesure des couches permettent au réseau de neurones d'analyser les détails fins de la donnée d'entrée dans les premières couches, puis au fur et à mesure que les opérations de *pooling* sont appliquées, d'utiliser des structures plus « grossières » de la donnée. Dans le cadre de la spectrométrie, les détails fins peuvent s'apparenter aux pics photoélectriques, qui sont bien résolus dans le cadre des détecteurs Caliste. Les structures Compton sont des structures plus grossières, qui peuvent s'étaler sur plusieurs centaines de canaux du spectre.

2.2 Apprentissage des réseaux de neurones

L'apprentissage des réseaux de neurones consiste à trouver les paramètres optimaux, c'est-à-dire les poids W, biais b ou filtres K de chaque neurone, afin d'atteindre un objectif donné.

2.2.1 Objectif et fonction de coût

Les réseaux de neurones sont des algorithmes adaptés pour l'apprentissage supervisé, ce qui signifie que la base de données d'apprentissage doit être constituée de données d'entrées X et de données de sortie Y. Cette base de données est constituée d'un certain nombre d'exemples N, que je note X_i et $Y_i = \left(y_{ij}\right)_j$. Je note avec un deuxième indice j le neurone de sortie correspondant : y_{ij} est le résultat du neurone de sortie j pour l'exemple i.

En paragraphe 2.1, j'ai montré comment un réseau de neurones effectue une prédiction $\hat{Y}$ à partir d'une donnée X et de paramètres θ constitués des filtres, poids et biais des neurones de toutes les couches. Je note $\hat{Y}(X,\theta)$ la prédiction effectuée à partir des données X et des paramètres θ.

On définit une fonction de coût qui évalue la pertinence d'une prédiction $\hat{Y}$ par rapport aux données attendues Y. Je note $L(\theta)$ cette fonction de coût.

On peut choisir différentes fonctions de coût, comme par exemple, la distance euclidienne moyenne définie par :

$$L(\theta) = \frac{1}{N}\left\|Y - \hat{Y}(X,\theta)\right\|_2^2 = \frac{1}{N}\sum_{i,j}\left(y_{ij} - \hat{y}_{ij}(X,\theta)\right)^2 \tag{2-3}$$

Cette distance est une des métriques utilisées par des algorithmes classiques. On utilise en général cette métrique dans des problèmes de régression, c'est-à-dire quand la sortie y_{ij} attendue est un nombre continu.

J'utiliserai aussi, dans des cas de classification, l'entropie croisée binaire (*binary cross-entropy*) :

$$L(\theta) = -\frac{1}{N}\sum_{i,j} y_{ij}\log\left(\hat{y}_{ij}(X,\theta)\right) + \left(1 - y_{ij}\right)\log\left(1 - \hat{y}_{ij}(X,\theta)\right) \tag{2-4}$$

Cette fonction de coût est intéressante quand la sortie y_{ij} est binaire, c'est-à-dire, si elle vaut 0 ou 1, et quand la prédiction $\hat{y}_{ij}$ est une valeur comprise entre 0 et 1. Cela correspond à des problèmes dans lesquels on souhaite prédire des classes, ce qui est naturellement adapté à la problématique de l'identification des radioéléments. On peut aussi utiliser cette fonction de coût pour prédire des valeurs continues entre 0 et 1, en remarquant qu'elle atteint son minimum quand $Y = \hat{Y}$ mais ce minimum ne vaut pas 0 si $y_{ij} \neq 0$ ou $y_{ij} \neq 1$.

Quelle que soit la fonction de coût utilisée, l'objectif sera toujours de chercher les paramètres optimaux θ_{opt} qui minimisent la fonction de coût, de sorte à prédire au mieux les vraies valeurs de sortie y_{ij} :

$$\theta_{\text{opt}} = \underset{\theta}{\operatorname{argmin}}\, L(\theta) \tag{2-5}$$

2.2.2 Descente de gradient et backpropagation

Afin de résoudre ce problème d'optimisation, l'algorithme classiquement utilisé est un algorithme de descente de gradient. Cela consiste à calculer le gradient de la fonction de coût ∇L et d'optimiser itérativement les paramètres en suivant ce gradient. En réalité, cet algorithme est adapté à des problèmes convexes, or ce n'est généralement pas le cas du problème défini par l'équation (2-5). L'optimisation permet ici de ne trouver qu'un minimum local, dont la pertinence est quantifiée en testant les performances du réseau sur une base de données de test.

À une étape n du processus d'optimisation aussi appelée époque, on met à jour les paramètres de la manière suivante :

$$\theta^{n+1} \coloneqq \theta^{n} - \lambda \nabla L(\theta) \tag{2-6}$$

λ désigne ici le taux d'apprentissage, sa valeur est fixée par l'utilisateur ou peut être optimisée par certains algorithmes spécifiques d'apprentissage.

L'intérêt des réseaux de neurones est que le calcul de ce gradient s'effectue par une succession d'opérations élémentaires, appelée *backpropagation* de l'erreur. Je décris en Annexe B la manière dont ce calcul est effectué en détail dans le cas d'un réseau de neurones de type perceptron.

L'idée générale est de calculer la dépendance de L par rapport aux paramètres de la dernière couche n_y, c'est-à-dire : $\frac{\partial L}{\partial W^{[n_y]}}, \frac{\partial L}{\partial b^{[n_y]}}$. Ce calcul est direct à partir de $\frac{\partial L}{\partial \hat{Y}}$ et des formules *forward pass* décrites en paragraphe 2.1 en utilisant les formules sur les dérivées des fonctions composées.

Ensuite, pour une couche l, on peut calculer $\frac{\partial L}{\partial W^{[l]}}$ et $\frac{\partial L}{\partial b^{[l]}}$ à partir de $\frac{\partial L}{\partial a^{[l]}}$ et $\frac{\partial L}{\partial z^{[l]}}$, qui s'obtiennent de manière récursive avec les valeurs de $\frac{\partial L}{\partial a^{[l+1]}}$ et $\frac{\partial L}{\partial z^{[l+1]}}$, des formules *forward pass* et des opérations sur les dérivées de fonctions composées.

Pour les réseaux de neurones, la plupart de ces calculs sont en fait des sommes et des produits, ce qui les rend simples et parallélisables sur ordinateur. Les fonctions d'activation sont justement choisies pour que leur dérivée soit simple à calculer. Le choix de la fonction ReLU sur les couches intermédiaires, dont les dérivées ont des valeurs fixes, 0 ou 1, évite deux phénomènes dans l'algorithme de *backpropagation*. D'une part, cela empêche le phénomène d'évanescence du gradient, c'est-à-dire l'apparition d'une dérivée trop petite dans une couche qui risque de se propager aux autres couches récursivement et donc de ne plus mettre à jour les paramètres correctement. D'autre part, inversement, cela évite les phénomènes d'explosion du gradient, où une dérivée très grande se propage et fait diverger le calcul du gradient dans les couches suivantes.

L'application directe de la formule (2-6) et de cet algorithme est appelée *Batch Gradient Descent*. D'autres algorithmes basés sur la même approche introduisent des moments dans l'équation (2-6) qui accélèrent le calcul dans certains cas (Kingma & Ba, 2015) ou encore qui adaptent le taux d'apprentissage λ. Le choix de l'algorithme se faisant empiriquement, il n'existe pas de règle prédéfinie qui permet de juger de l'efficacité de l'algorithme a priori du fait de la complexité des structures à traiter.

Par ailleurs, le calcul en (2-6) peut être effectué non pas sur l'ensemble de la base de données, mais sur un certain nombre d'exemples défini par l'utilisateur, pris aléatoirement à chaque itération sur la

base de données, qu'on appelle *batch* ou *mini-batch*. En fonction des performances de calcul de l'ordinateur utilisé, cela peut garantir la parallélisation des opérations sur l'ensemble du *mini-batch* et donc accélérer l'apprentissage. Lorsque le calcul (2-6) est effectué élément par élément de la base de données, on parle de *Stochastic Gradient Descent* (SGD).

Le nombre d'époques sur lesquelles on applique la mise à jour des paramètres dépend du problème à traiter et des performances d'apprentissage. On peut par exemple ne faire qu'une seule époque d'apprentissage si la base de données contient un nombre élevé d'éléments, typiquement plusieurs dizaines de millions d'éléments, ou alors utiliser plusieurs milliers d'époques sur des bases d'apprentissage plus modestes ne contenant que quelques centaines d'exemples. Une manière d'arrêter l'apprentissage est d'utiliser une méthode d'*early stopping* qui consiste à surveiller l'évolution de la fonction de coût sur un ensemble d'exemples, appelé jeu de validation, qui n'est pas utilisé pour l'apprentissage des paramètres. La méthode d'*early stopping* consiste à arrêter l'apprentissage lorsque la fonction de coût ne diminue plus sur le jeu de validation. J'utilise cette méthode au paragraphe 3.3.

2.2.3 Surveillance des performances et adaptation des hyperparamètres

Au cours de l'apprentissage, on peut surveiller l'évolution de la fonction de coût pour s'assurer que sa valeur diminue bien au fur et à mesure des itérations. Si la fonction de coût ne diminue pas ou si les performances ne sont pas satisfaisantes une fois qu'elle a convergé, l'utilisateur peut modifier différentes options dans son réseau de neurones, que l'on appelle hyperparamètres. On distingue les « hyperparamètres », qui sont les valeurs que le concepteur peut modifier, des « paramètres », qui sont les valeurs (filtres, poids, biais) que le réseau cherche à optimiser par l'apprentissage.

Dans les hyperparamètres, on retrouve le taux d'apprentissage λ de l'équation (2-6). Si ce taux est trop petit, l'apprentissage va converger de manière sûre mais lente, et s'il est trop grand, il risque de diverger et on peut voir la fonction de coût augmenter au lieu de diminuer. C'est donc un hyperparamètre à choisir avec soin. La valeur typique par défaut du taux d'apprentissage est de 0,01 pour la plupart des bibliothèques d'apprentissage, mais cette valeur doit être adaptée suivant le problème à traiter, comme je le montre dans le cas de la problématique de la spectro-identification en paragraphe 3.3. De même, l'algorithme d'optimisation et les hyperparamètres qui lui sont associé doivent être choisis judicieusement. En général, cela se fait de manière empirique, en surveillant l'évolution des performances sur un jeu de validation, constitué de données qui ne sont pas utilisées pour entraîner le réseau de neurones.

De même, l'architecture du réseau est en soi un hyperparamètre majeur. Le nombre de couches, le nombre de neurones par couches, le type de couches ou encore la taille des filtres utilisés pour les couches de convolution sont autant de bras de levier possibles. Ils peuvent aussi être un frein à l'utilisation du *deep learning* quand les néophytes ne savent pas vraiment par où commencer. Je montre en paragraphe 3.3 une méthode que j'ai mise en place pour choisir l'architecture de mon réseau.

Il faut noter qu'un réseau profond arrivera mieux à ajuster ses paramètres pour faire les bonnes prédictions $\hat{Y}$ sur le jeu d'apprentissage au risque d'aboutir à un surapprentissage, aussi appelé *overfitting*, de la base de données. Un réseau trop peu profond risque en revanche de ne pas avoir assez de paramètres pour apprendre. La complexité du réseau est à adapter en fonction de ses performances.

Il existe des techniques pour éviter le phénomène d'*overfitting* tout en conservant une architecture profonde, de sorte que le réseau de neurones n'apprenne pas « par cœur » sur les données mais soit capable de généraliser. Ces techniques sont appelées techniques de régularisation. J'en montre trois ici qui sont couramment utilisées.

D'abord, la régularisation des paramètres consiste à pénaliser l'apprentissage de paramètres avec des valeurs trop grandes. L'idée sous-jacente est que les données doivent être expliquées avec un minimum de caractéristiques. Dans ce cadre, on ajoute un terme de pénalisation dans l'équation (2-3) ou (2-4), de sorte que la fonction de coût totale devienne :

$$L_{tot}(\theta) = L(\theta) + \alpha\|W\| + \beta\|b\| + \gamma\|K\| \tag{2-6}$$

Les normes utilisées peuvent être la norme 1 ou la norme euclidienne au carré, pénalisant les paramètres avec des valeurs trop élevées.

Les paramètres de régularisation α, β et γ sont des hyperparamètres fixés par le concepteur et peuvent être différents sur chaque couche du réseau. Si ces paramètres sont trop grands, la régularisation va dominer et le réseau risque de ne plus apprendre, et s'ils sont trop petits, aucune régularisation n'est effectuée.

Une méthode pour accélérer l'apprentissage est l'ajout de fonctions de *batch-normalization*, qui consistent à normaliser en moyenne et en variance le contenu des couches. Cela permet notamment de s'assurer que certaines couches ne prennent pas de trop grandes ou de trop petites valeurs et donc que le calcul de *back-propagation* ne soit pas ralenti par des gradients trop petits ou divergent par des gradients trop élevés. Cette méthode sert aussi à la régularisation.

Une dernière méthode de régularisation que j'utilise est le *dropout*. Il consiste à éteindre, c'est-à-dire à mettre à 0, aléatoirement les neurones de certaines couches, selon un taux défini par le concepteur. Cela évite que certains neurones ne se spécialisent pas pour reconnaître des exemples particuliers et donc évite le sur-apprentissage. Le taux d'extinction constitue un nouvel hyperparamètre. En général, le *dropout* n'est utilisé que dans la phase d'apprentissage et il est désactivé lors de l'utilisation du réseau. Cependant, je vais montrer en paragraphe 3.2 une utilisation particulière du *dropout* dans la phase d'utilisation du réseau de neurones pour évaluer l'incertitude sur le résultat du réseau. Sur la bibliothèque Keras (Chollet, 2015) que j'utilise pour concevoir et entraîner les réseaux de neurones, je précise que j'utilise la fonction *SpatialDropout1D* sur les couches de convolution. Elle éteint un neurone de convolution entier, contrairement à la fonction *Dropout* qui n'éteint que des éléments aléatoirement dans ces neurones.

3 Apprentissage de réseaux de neurones convolutifs pour la spectro-identification de radioéléments

3.1 Méthodologie mise en place

3.1.1 Indentification, quantification et bibliothèque de sources

Comme je l'ai décrit en partie 1, je souhaite être capable de répondre à deux questions. D'une part, je veux pouvoir dire quelles sources sont présentes et lesquelles sont absentes dans une acquisition. Dans un second temps, parmi celles qui sont présentes, je souhaite être en mesure de

quantifier leurs proportions, si plusieurs radioéléments sont effectivement présents. Enfin, je souhaite établir un niveau de confiance sur les résultats obtenus pour ces deux questions.

Pour ce faire, je distingue le problème d'identification et le problème de quantification en mettant en place un réseau de neurone distinct pour chaque problème. Dans le premier cas, le réseau de neurones doit renvoyer une réponse binaire, présence ou absence de source, ce qui est un problème de classification. Dans le deuxième cas, le réseau de neurones doit renvoyer un nombre continu entre 0 et 1, en fonction de la proportion de source détectée, ce qui est un problème de régression.

J'ai choisi d'utiliser une bibliothèque de sources constituée de 14 radioéléments, en me basant notamment sur (U.S. Department of Homeland Security, 2004), qui contient les spécifications auxquelles les détecteurs de sources radioactives gamma doivent répondre. Je prends aussi en compte les sources que nous possédons en laboratoire ou pour lesquelles j'ai pu obtenir des données réelles avec Caliste. J'ai aussi fait en sorte de choisir prioritairement des radioéléments dont les lignes d'émission principales sont dans la dynamique de Caliste, 2 keV à 1 MeV[*], avec une forte probabilité d'émission.

Les radioéléments que j'ai choisis sont les suivants :

- Sources d'intérêt d'après (U.S. Department of Homeland Security, 2004) et que nous possédons en laboratoire : ^{241}Am, ^{133}Ba, ^{57}Co, ^{137}Cs ;
- Autre source que nous possédons en laboratoire : ^{152}Eu ;
- Sources d'intérêt pour lesquelles j'ai pu avoir des données : ^{22}Na, ^{226}Ra-D[†] ;
- Autres sources d'intérêt : ^{67}Ga, ^{131}I, ^{192}Ir, ^{235}U-D, ^{99m}Tc, ^{201}Tl, ^{60}Co.

Je souligne aussi que, suivant les applications, certains radioéléments ou certains mélanges de radioéléments seront plus ou moins pertinents. Par exemple, en Assainissement & Démantèlement, un intérêt particulier est porté sur la détection du ^{241}Am alors que dans le cadre d'un accident d'exploitation nucléaire, c'est le ^{137}Cs qui est principalement étudié, notamment comme traceur des rejets radioactifs. Je présente ici principalement la méthode en prenant un nombre relativement large de radioéléments dans la bibliothèque pour montrer qu'elle est généralisable. Cependant, l'utilisation d'une base de données plus restreinte et plus spécifique présentera l'avantage d'avoir de meilleures performances si elle est dédiée à l'application visée.

Le dernier point important de ma méthode est que je vais créer deux réseaux de neurones propres[‡] à chaque radioélément, contrairement à la plupart des références où un seul réseau de neurones global est créé avec une sortie par radioélément d'intérêt. Cela est motivé par le fait que les réseaux de neurones à couches de convolution ont pour particularité d'extraire les caractéristiques pertinentes des données par les neurones de convolution dans le but de répondre à l'objectif donné. En spécialisant les réseaux de neurones pour chaque radioélément, les caractéristiques extraites seront spécifiques à chacun d'entre eux. Le réseau de neurones sera ainsi plus performant a priori et moins dépendant du nombre de radioéléments dans la scène à analyser. Un autre avantage est que le réseau

[*] J'utilise aussi le ^{60}Co, dont les lignes sont au-delà de 1 MeV, afin de vérifier la réponse de ma méthode pour ce type de source.

[†] ^{226}Ra-D signifie ^{226}Ra et ses descendants dans la chaîne radioactive, que l'on retrouve généralement à l'équilibre avec ^{226}Ra. De même pour ^{235}U-D.

[‡] Un réseau pour le problème d'identification et un pour le problème de quantification.

nécessaire pour chaque source pourra être réduit en nombre de couches et de neurones, par rapport à un réseau de neurones général pour toutes les sources. L'objectif pourra être d'embarquer à termes les réseaux pré-entraînés puis procéder à leurs calculs sur un FPGA par exemple.

En résumé, j'utilise 14 radioéléments. Pour chaque radioélément, je vais construire deux réseaux de neurones, l'un permettant de confirmer la présence ou l'absence de la source, le deuxième, qui, en cas de présence de la source, donne sa proportion dans le spectre enregistré. Ces réseaux de neurones sont des réseaux de neurones de convolution, dont la première partie, composée de neurones de convolution, extrait les caractéristiques pertinentes pour identifier les radioéléments. La seconde partie, constituée de neurones *fully-connected*, identifie ou quantifie les radioéléments à partir de ces caractéristiques extraites.

Cette stratégie est résumée en Figure 3-1 et Figure 3-2. Chaque ligne de chaque stratégie constitue un réseau de neurones propre. Pour N sources dans la bibliothèque, on construit $2N$ réseaux de neurones. On pourrait penser à utiliser les mêmes couches de convolution pour tous les radioéléments puis de n'entraîner que les couches *fully-connected* indépendamment pour chaque radioélément. Néanmoins, l'intérêt d'avoir des couches de convolution propres à chaque radioélément permet d'extraire à travers ces couches les caractéristiques propres au radioélément d'intérêt. En revanche, on peut utiliser cette méthode dans le cas d'un agrandissement de la bibliothèque de radioéléments : si on ajoute un radioélément dans la bibliothèque de radioéléments, on peut envisager de ne relancer l'apprentissage que sur les couches *fully-connected* pour les réseaux de neurones qui ont précédemment été entraînés, ce qui accélère la phase d'apprentissage en figeant les couches de convolution. Pour les radioéléments qui ont été ajoutés à la bibliothèque, il faut en revanche faire l'apprentissage de l'ensemble des couches.

Le nombre de réseaux peut être assez élevé suivant le nombre de radioéléments à identifier. Cela ne posera pas de problème du point de vue des applications car en réalité, pour des applications spécifiques, le nombre de radioéléments d'intérêt est en général relativement réduit, inférieur à 5 typiquement.

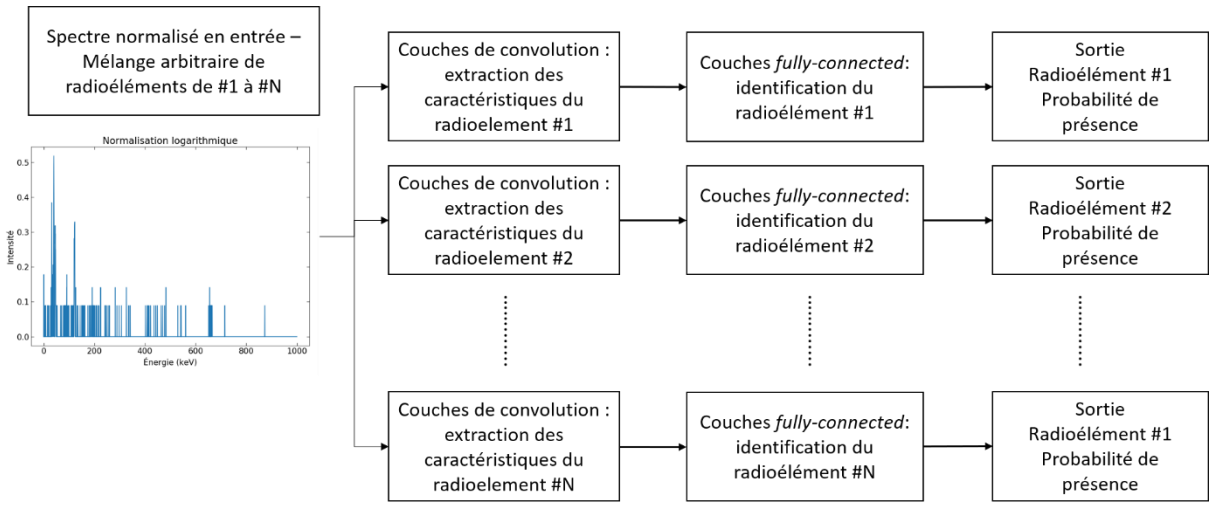


Figure 3-1 : Stratégie sur les réseaux d'identification.

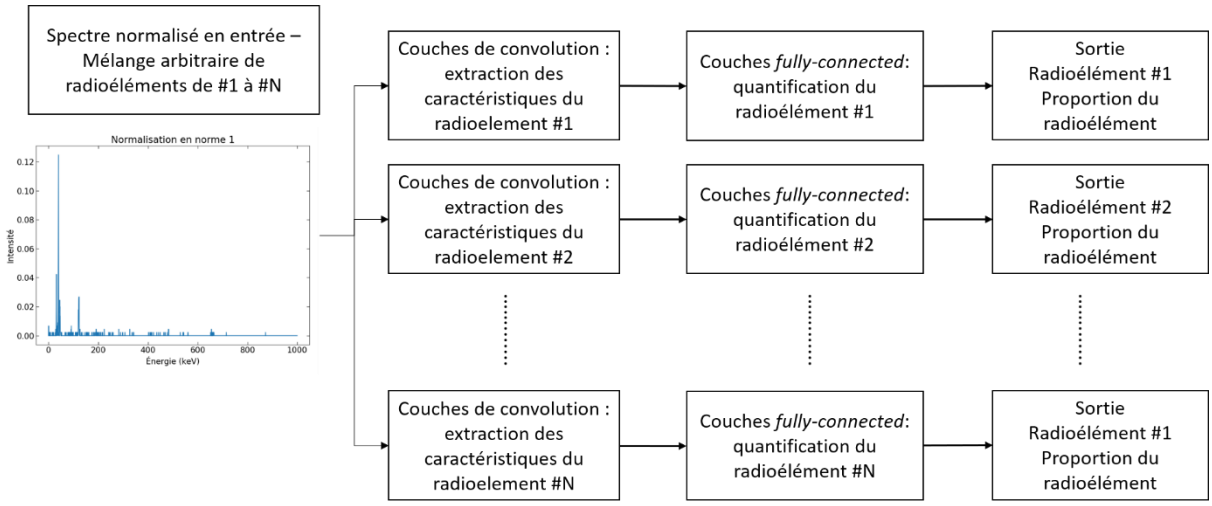


Figure 3-2 : Stratégie sur les réseaux de quantification.

3.1.2 Construction de la base de données

Pour construire la base de données d'apprentissage, les réseaux de neurones étant des algorithmes d'apprentissage supervisé, il est nécessaire d'avoir des données d'entrée X et des données de sortie Y. Je montre dans cette partie comment je mets en place ces données d'entrée X à partir de mes algorithmes de simulation et d'un processus d'augmentation de données que j'ai imaginé et développé.

Ce problème se distingue des applications usuelles d'Intelligence Artificielle, comme par exemple sur l'analyse de bases de données d'images ou des bases de données d'utilisateurs de sites *webs*. Ce sont en général des bases de données très grandes, contenant des milliers d'images ou des millions d'utilisateurs, d'où l'appellation de *big data*. Dans mon cas, je dispose par le Chapitre 2 d'une simulation par radioélément à identifier, simulations à partir desquelles je dois construire une base de données. Un processus d'augmentation de données est alors nécessaire pour obtenir une base de données constituée de plusieurs milliers d'exemples par radioélément à identifier.

Application des simulations

Les données d'entrée que j'utilise sont des spectres gamma de sources radioactives, information à partir de laquelle les algorithmes doivent être capables de répondre aux problèmes de spectro-identification et de quantification. J'utilise mon algorithme de simulation décrit en Chapitre 2 pour simuler des listes de photons pour chaque radioélément de ma bibliothèque. Ma base de données est uniquement synthétique et repose sur une modélisation physique de l'interaction des photons avec la matière et la formation du signal dans le détecteur.

L'intérêt d'utiliser des simulations est que je peux constituer une base de données synthétique qui ne se limite pas aux sources radioactives étalons dont nous disposons en laboratoire. J'utilise tout de même des données simulées pour les sources que nous possédons, car cela me permettra de valider avec des données réelles que l'apprentissage fonctionne avec ces données simulées. L'idée est notamment de faire l'apprentissage avec des « portraits-robots » des sources et de tester ensuite avec des données réelles.

Les simulations présentent aussi d'autres intérêts, car elles permettent par la suite, suivant les applications, de prendre en compte des matériaux absorbants, diffuseurs ou éventuellement d'autres spécificités liées à l'application. Les complexifications envisagées du problème résideront principalement dans la fidélité et la représentativité des modèles physiques employés, et non dans l'optimisation des réseaux de neurones eux-mêmes.

Par conséquent, à l'aide de mon algorithme de simulation, je génère pour chaque radioélément une liste de 300 000 photons* que j'utilise pour construire ma base de données. Je ne prends en compte que les événements simples dans ces simulations, ce qui implique qu'il sera nécessaire de faire de même dans les données de tests. Les événements simples présentent l'intérêt d'avoir une meilleure résolution spectrale, mais l'utilisation de ces événements seuls se fait au prix d'une perte de sensibilité qui n'est pas négligeable. En particulier à haute énergie, au-delà de 500 keV, typiquement la moitié au moins des événements enregistrés sont des événements multiples.

En plus des simulations des radioéléments, j'ai aussi ajouté une mesure de bruit de fond naturel, qui provient d'une véritable acquisition avec Caliste, dans notre laboratoire, sans source étalon. Ce bruit de fond ne peut pas être simulé facilement et dépend de chaque scène, mais cela montre au réseau de neurones des cas avec bruit de fond réaliste dans un environnement contrôlé. Dans la situation d'un bruit de fond différent, celui-ci se comporterait comme un autre type de source à identifier et pourrait aussi constituer un élément de la base d'apprentissage que le réseau de neurones apprendrait avec un entraînement dédié.

Augmentation de données

Pour l'instant, j'ai une liste de photons pour chaque radioélément de ma bibliothèque. J'utilise maintenant un processus d'augmentation de données ou *data augmentation* qui consiste ici à créer des mélanges de radioéléments avec des statistiques de comptage variées.

Pour créer une nouvelle donnée d'entrée, je tire aléatoirement le nombre de sources N_{sources} entre 1 et 15 qui seront dans le mélange, et je tire ensuite aléatoirement ces N_{sources} parmi les 14 radioéléments de la bibliothèque ainsi que le bruit de fond.

Puis, je tire aléatoirement un nombre de photons N_{photons} de la manière suivante :

$$N_{\text{photons}} = 10^{7x+1}, x \sim \text{Unif}([0,1]) \tag{3-1}$$

Cela permet de tirer un nombre de photons entre 10 et 10 millions, en faisant en sorte de ne pas pénaliser les échantillons avec une faible statistique de comptage.

Les proportions p_i de chaque source i choisie sont ensuite tirées aléatoirement, et je calcule le nombre de photons par source i : $N_i = p_i . N_{\text{photons}}$. Dans la liste de la source i, je tire N_i photons aléatoirement, avec la possibilité de tirer plusieurs fois le même photon si besoin. Les listes sont en

* 300 000 photons permettent d'obtenir des spectres des radioéléments avec une statistique de comptage suffisante pour faire apparaître l'ensemble des caractéristiques discriminantes de chaque radioélément. Par ailleurs, utiliser un nombre plus élevé de photons augmente considérablement la durée de création de la base de données décrite dans le paragraphe suivant. Du fait du tirage aléatoire dans une liste contenant 300 000 photons, la durée est déjà de près de 4 heures pour cette liste de 300 000 photons et de plusieurs jours pour 10 millions de photons.

effet constituées de 300 000 photons et N_i peut aller jusque 10 millions. Le fait de tirer plusieurs fois le même photon peut créer un biais dans les données. Cependant, l'influence de ce biais est réduite par le fait que le nombre de photons initial dans les listes est suffisamment grand d'une part et par l'application d'une décalibration aléatoire décrite ci-après d'autre part.

À la fin de cette première étape d'augmentation de données, j'obtiens une liste de taille N_{photons}, provenant des différents radioéléments sélectionnés aléatoirement.

À ce moment, je définis les données de sortie associées à l'exemple généré ici. Pour les réseaux de neurones d'identification, je code en binaire les sorties, connaissant les sources que j'ai tirées : je mets $y_i^{\text{ID}} = 1$ si la source i est présente et $y_i^{\text{ID}} = 0$ si elle est absente. De même, pour le réseau de neurones de quantification des proportions, j'affecte les proportions de chaque source que j'ai tirées aléatoirement précédemment : $y_i^{\text{prop}} = p_i$. Si la source est absente, je prends $y_i^{\text{prop}} = 0$.

La deuxième étape de mon processus d'augmentation de données consiste à appliquer une modification de la calibration sur la liste d'événements qui simule une éventuelle dérive du détecteur ou la sensibilité aux conditions opérationnelles comme la température ou la haute-tension. Ainsi, l'identification par les réseaux de neurones est rendue plus robuste. Pour ce faire, je tire un facteur de gain aléatoire g dans une loi normale centrée en 1 et d'écart-type 0,005, et un *offset* aléatoire o dans une loi normale centrée en 0 keV et d'écart-type de 0,5 keV. À partir de l'énergie E_j de chaque photon j de la liste d'événements, je calcule une énergie décalibrée $E_{j_{\text{new}}} = g(E_j - o)$. En fonction de la décalibration attendue, on peut changer les paramètres d'écart-type si on s'attend à un phénomène plus ou moins fort. La décalibration appliquée est propre à chaque mélange créé. Il est à noter qu'en utilisant une décalibration plus forte, je peux rendre les réseaux de neurones plus robustes, mais potentiellement au prix d'un niveau de confiance plus faible des réseaux de neurones sur leur prédiction, que j'évalue avec la méthode décrite en paragraphe 3.2.

Pour la dernière étape de mon processus d'augmentation de données, je tire aléatoirement un seuil bas en énergie, entre 0 et 20 keV. Le seuil bas est un paramètre que l'utilisateur du détecteur peut choisir et qui peut varier selon les mesures. En outre, cela permet aussi de simuler des cas où un absorbeur est placé devant la source, et faire ainsi en sorte que le réseau de neurones n'utilise pas seulement les lignes à basse énergie pour procéder à la spectro-identification.

Suivant ce protocole, j'ai créé 200 000 listes* de photons au total, correspondant à des mélanges de radioéléments, avec potentiellement un bruit de fond naturel, comprenant entre 10 et 10 millions de photons. Chaque radioélément est dans des proportions aléatoires, avec une décalibration aléatoire appliquée aux énergies des photons.

* Il est possible de stocker 200 000 spectres de 2000 canaux dans 1,5 Go de mémoire, correspondant aux capacités de mémoire vive d'un ordinateur classique. En fonction des performances de calcul, il est possible d'augmenter le nombre d'exemples utilisés, pour garantir une meilleure robustesse de l'apprentissage, ou inversement, de diminuer le nombre d'exemples si le matériel de calcul n'est pas adapté, au risque de diminuer les performances d'apprentissage. Utiliser plusieurs centaines de milliers d'exemples est en général suffisant pour avoir un apprentissage de bonne qualité. Avec des essais sur une base de données contenant 100 000 exemples, je n'ai pas observé de différence significative avec les performances des réseaux entraînés sur 200 000 exemples.

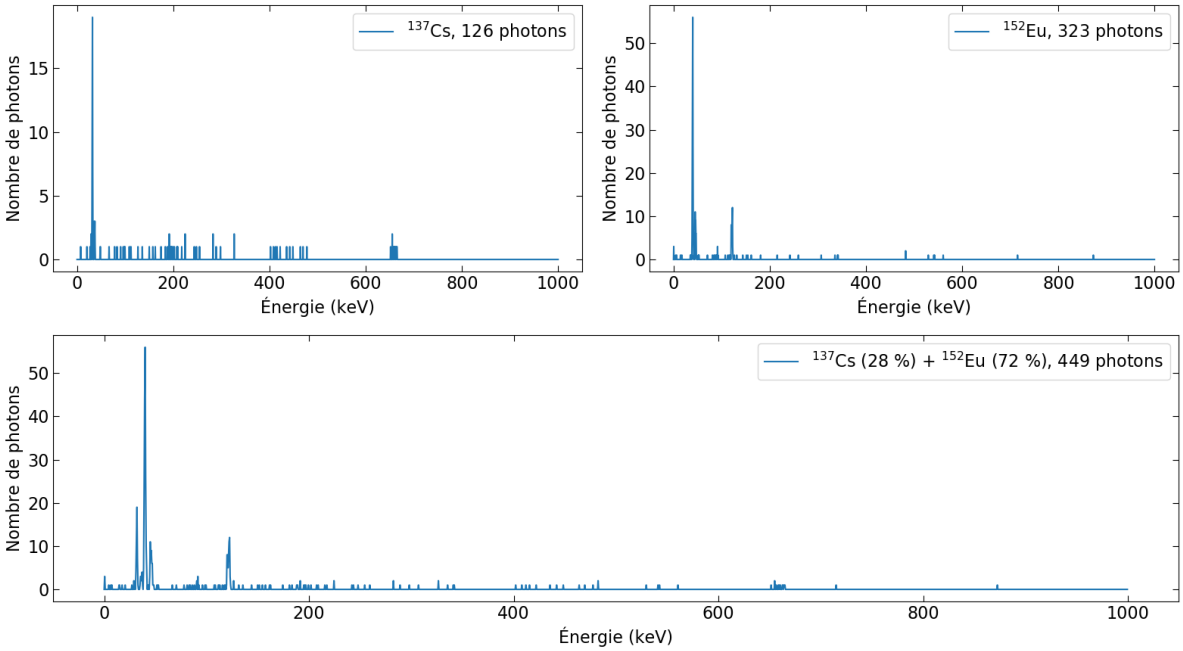


Figure 3-3 : Exemple d'un spectre mélange de deux radioéléments.

Pour chacune des listes, je crée le spectre correspondant, en faisant l'histogramme des énergies de chaque liste de photons avec 2000 canaux sur une dynamique de 0 à 1 MeV. Ce format fixe aussi celui des spectres que je construirai avec les données de test issues de Caliste. D'autres discrétisations du spectre sont possibles selon les propriétés du détecteur et de la précision de sa chaîne d'acquisition.

La Figure 3-3 représente un exemple de mélange contenant deux sources, une source de ^{137}Cs et une source de ^{152}Eu avec des proportions aléatoires associées.

3.1.3 Normalisation des spectres

Le nombre de photons dans le spectre variant de 10 à 10 millions, la diversité de statistiques de photons peut a priori poser problème dans l'apprentissage des réseaux de neurones. En effet, ces algorithmes sont en général utilisés avec des données d'entrée dont les valeurs sont comprises dans des plages relativement uniformes. Par exemple, les réseaux de neurones dédiés à l'analyse d'images ont des données d'entrée dont les intensités des pixels ont été normalisées de manière à être situées entre 0 et 1.

J'applique en conséquence un processus de normalisation aux spectres créés à l'étape précédente. Pour ce faire, j'utilise deux types de normalisations suivant la question à traiter, du fait des connaissances que l'on a, a priori, sur les spectres.

Pour la question de l'identification pure, on sait que les pics à haute-énergie peuvent porter des informations essentielles qui sont des signatures des radioéléments, comme le pic à 662 keV du ^{137}Cs par exemple. Or, du fait de la diminution de l'efficacité de détection en fonction de l'énergie, je sais aussi que ces pics apparaissent avec une intensité assez faible dans les spectres. On utilise généralement l'échelle logarithmique pour faire apparaître ces pics sur les représentations des spectres.

De ce fait, pour la question de l'identification, je vais utiliser une normalisation logarithmique définie de la manière suivante. Considérons un spectre $S = (S_i)_{i \in canaux}$, contenant les nombres de photons détectés dans chaque canal i d'énergie. J'applique une première transformation intermédiaire :

$$S'_i = \log\left(\frac{S_i + 1}{\sum_j (S_j + 1)}\right) \tag{3-2}$$

Dans l'équation (3-2), l'ajout de 1 à chaque canal évite d'avoir 0 en argument du logarithme. La division par $\sum_j (S_j + 1)$ rend le spectre indépendant du nombre total de photons dans le spectre. À la suite de ce calcul, les valeurs de S' sont alors comprises entre $-\log\left(\sum_j (S_j + 1)\right)$ et 0. J'applique la transformation (3-3) afin de faire en sorte que les valeurs du spectre normalisé soient comprises entre 0 et 1. S^{ID} est le spectre normalisé que j'utilise in fine pour les réseaux de neurones d'identification :

$$S_i^{\mathrm{ID}} = \frac{S'_i}{-\min(S')} + 1 \tag{3-3}$$

Pour la question des proportions en revanche, je n'utilise pas cette normalisation car je sais que la proportion dépend linéairement du nombre de photons émis par chaque source. La normalisation logarithmique « casserait » cette linéarité. Pour normaliser et faire en sorte de rendre le spectre le plus indépendant possible du nombre de photons, je divise directement par le nombre de photons dans le spectre, ce qui correspond à une normalisation en norme 1. S^{prop} est le spectre normalisé que j'utilise pour les réseaux de neurones de quantification des proportions de chaque radioélément dans le spectre :

$$S_i^{\mathrm{prop}} = \frac{S_i}{\sum_j S_j} \tag{3-4}$$

La Figure 3-4 représente l'effet de ces deux normalisations sur le spectre mélange de la Figure 3-3.

Lors de la phase de test, il faudra naturellement appliquer la même normalisation aux spectres à tester avant d'appliquer les réseaux de neurones.

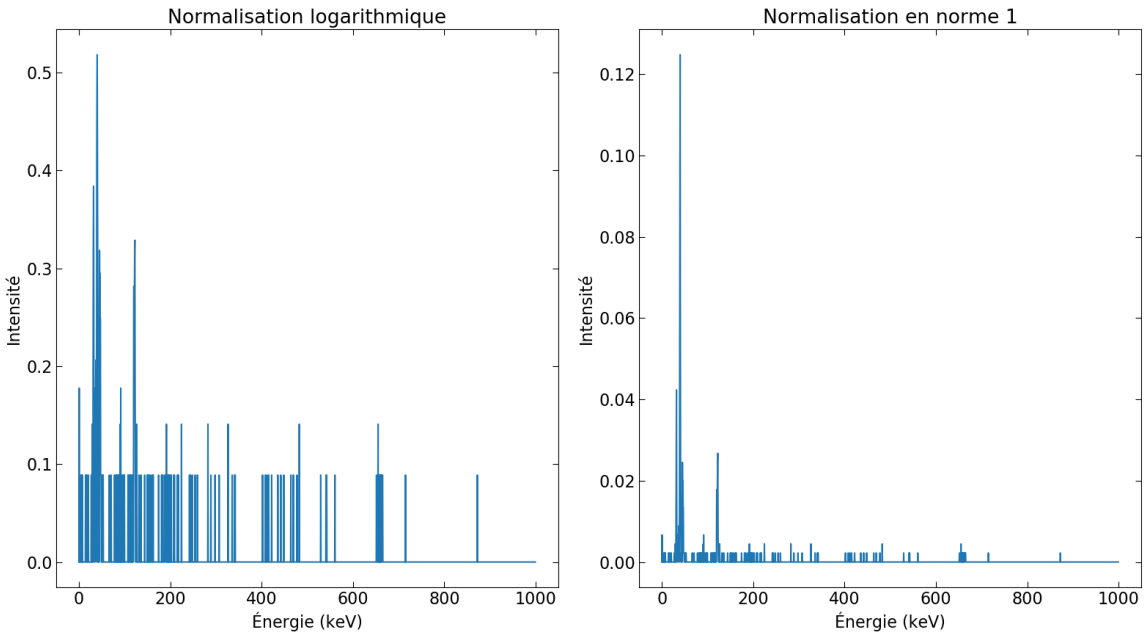


Figure 3-4 : Normalisations appliquées aux spectres, logarithmique pour l'identification (à gauche) et en norme 1 pour la quantification des proportions (à droite).

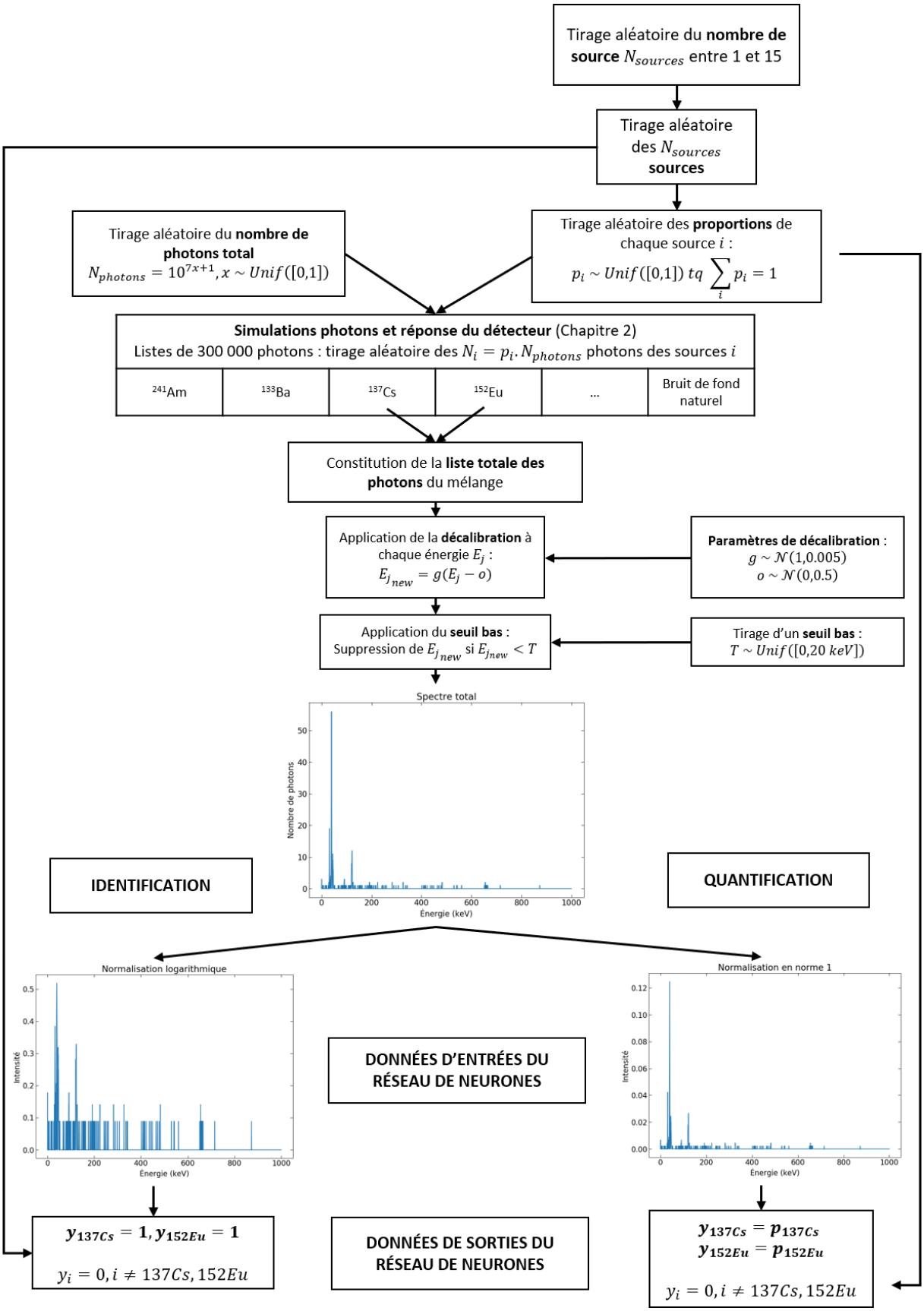


Figure 3-5 : Processus de création de la base de données, par *data augmentation*.

La Figure 3-5 schématise l'ensemble du processus de préparation des données décrit dans ce paragraphe, processus que j'itère 200 000 fois pour créer l'ensemble des exemples de la base de données.

Un point auquel il faut faire attention est que la base de données ne contient que les radioéléments qui sont dans la bibliothèque choisie à la base. Les réseaux de neurones ne seront pas entraînés pour identifier une source qui ne fait pas partie de la bibliothèque de sources. Si on présente un radioélément qui n'est pas connu des réseaux de neurones, chacun d'entre eux pourra seulement indiquer qu'il ne s'agit pas du radioélément pour lequel il a été entraîné, contrairement par exemple à un algorithme de « *fitting* » qui pourra chercher à expliquer un spectre à travers une combinaison ad-hoc des spectres des différents radioéléments.

3.2 Incertitudes et réseaux de neurones bayésiens

L'évaluation des incertitudes sur les résultats est un point essentiel, notamment dans le cadre de mesures physiques. Il est nécessaire de fournir l'imprécision attendue sur les sorties des réseaux de neurones, afin de pouvoir en attribuer un niveau de confiance.

Généralement, la fiabilité des réseaux de neurones est évaluée à l'aide d'une base de données de test. Cette base de données doit comporter des éléments différents de ceux de la base de données d'apprentissage et être au plus proche des données attendues pour l'application du réseau de neurones. Pour des réseaux de neurones de classification, on évalue alors la précision du réseau par le taux d'erreur par rapport aux labels attendus, de faux-positifs et de faux-négatifs, ce qui donne une évaluation des performances du réseau de neurones. Pour les réseaux de neurones de régression, on évalue l'écart entre la prédiction et la valeur attendue. Je procède à l'évaluation des performances du réseau de ce point de vue en paragraphe 4.3.4.

Cependant, cette méthode ne permet pas, pour un test sur un exemple donné, d'évaluer la fiabilité sur le résultat renvoyé par le réseau pour cet exemple en particulier. D'autres facteurs doivent être pris en compte. Par exemple, un plus grand bruit sur l'exemple, un plus faible nombre de photons sont des éléments qui peuvent augmenter l'incertitude sur le résultat du réseau. A contrario, certaines données très peu bruitées avec un grand nombre de photons permettront au réseau de neurones de donner un résultat avec un grand taux de confiance.

Pour prendre en compte cela, il faudrait que les réseaux de neurones soient en mesure de fournir non pas une sortie déterministe, comme on l'attend avec les réseaux de neurones tels qu'ils sont habituellement conçus, mais plutôt une distribution de résultats, permettant ainsi de faire une évaluation statistique du résultat.

Les réseaux de neurones bayésiens (Gal, 2016) ont été justement conçus dans le but d'apprendre non pas des poids W, biais b et filtres K fixes et déterministes mais des distributions de ces paramètres. Ainsi, dans ce type de réseau, pour chaque neurone, ces paramètres suivent une distribution et ce sont les paramètres de ces distributions qui sont appris lors de l'entraînement des réseaux de neurones. Généralement, la loi utilisée est la loi normale caractérisée par sa moyenne μ et son écart-type σ. Ainsi, à chaque application du réseau de neurones, les paramètres W, b et K sont tirés aléatoirement pour chaque neurone dans la loi qui a été apprise, et donc le résultat à chaque application est différent. En appliquant un certain nombre de fois le réseau de neurones bayésien sur le même exemple, on obtient alors une distribution de résultats, qui caractérise une incertitude « autoévaluée » par le réseau.

En revanche, toute la difficulté des réseaux de neurones bayésiens réside dans leur apprentissage, puisque ce sont des distributions qui sont apprises et non plus des paramètres fixes. Généralement, l'apprentissage est effectué en utilisant l'inférence variationnelle, en minimisant la divergence de Kullback-Leibler (*KL-divergence*) (Gal, 2016), complexe à mettre en place.

Une solution possible est alors de faire une approximation des réseaux bayésiens par *dropout* (Gal, 2016) (Gal & Ghahramani, 2016). Comme je l'ai présenté au paragraphe 2.2.3, le *dropout* est une méthode de régularisation permettant d'éviter le surapprentissage, en procédant à l'extinction aléatoire des neurones selon un taux défini par le concepteur du réseau. Le *dropout* n'est habituellement utilisé que dans la phase d'apprentissage et désactivé dans la phase de test, où l'ensemble du réseau est pris en compte. Or, pour l'approximation des réseaux de neurones bayésiens par *dropout*, l'idée est justement de le maintenir aussi dans la phase de test et d'application. Du fait de l'extinction aléatoire, à chaque test sur un même exemple, le réseau de neurones fournit un résultat différent. Il suffit alors d'appliquer plusieurs fois le même réseau de neurones pour obtenir une distribution des résultats.

Je présente en Figure 3-6 le type de résultat auquel on peut s'attendre avec un réseau de neurones bayésien appliqué à une tâche de classification. On considère le résultat d'un neurone dont les résultats sont compris entre 0 et 1, 1 signifiant que le neurone est activé. En faisant le parallèle avec le problème de spectro-identification, on considère avec la valeur 1 que la source est présente. En appliquant le réseau de neurones bayésien approximé par *dropout* 100 fois sur le même exemple, on obtient à chaque calcul un résultat différent, dont on peut faire l'histogramme.

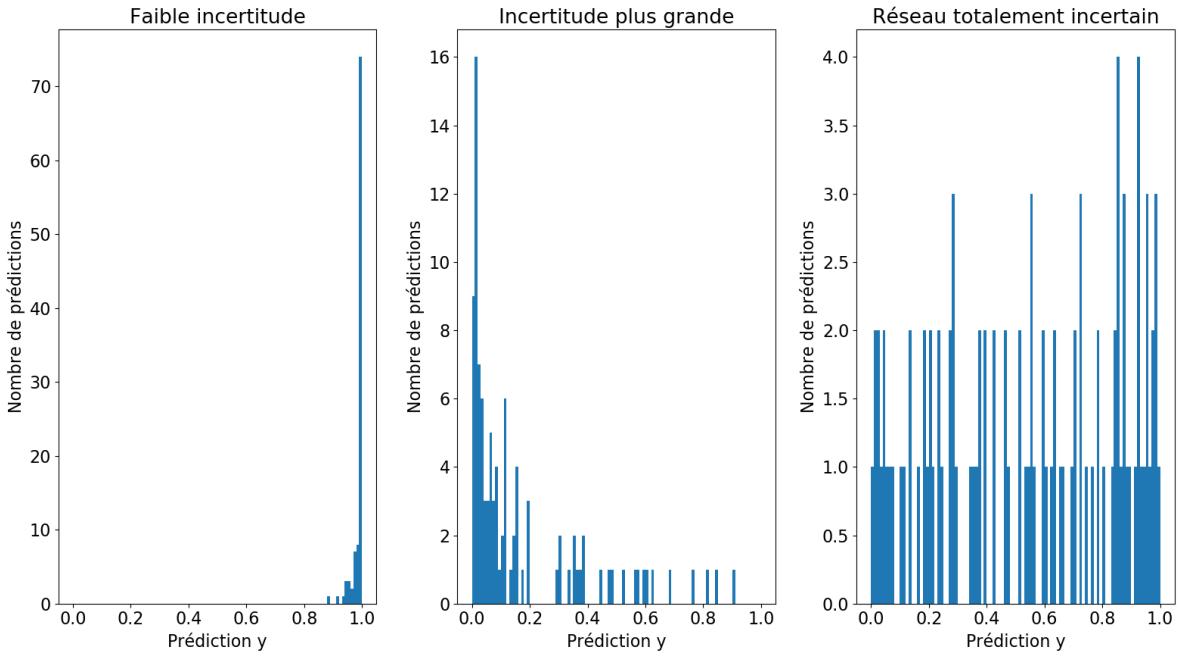


Figure 3-6 : Exemples d'incertitudes « autoévaluées » par un réseau bayésien.

Le premier exemple de la Figure 3-6 présente un cas où la dispersion des résultats renvoyés est faible, ils sont tous très proches de 1. Dans le cas de la spectro-identification, je considérerai avec ce type d'exemple que le radioélément est présent avec un fort taux de confiance. Dans le deuxième exemple, les résultats sont principalement concentrés relativement proches de 0, le radioélément est considéré comme étant absent par le réseau, avec un taux de confiance cependant plus faible que pour le

premier exemple. Enfin, dans le dernier cas, on voit que le réseau de neurones est très incertain, ses résultats étant très dispersés entre 0 et 1.

L'idée de l'utilisation de ce type de réseaux de neurones est de simuler en réalité un comportement assez similaire aux *Random Forests*, appliquées aux arbres de décisions (*Decision Trees*). Cela revient à artificiellement interroger plusieurs réseaux de neurones, chacun représenté par un *dropout* différent, qui ne vont pas tous s'intéresser aux mêmes caractéristiques dans les données. Ainsi, si dans les données, un artéfact risque d'être surinterprété avec un des *dropout*, qui augmente fortement la probabilité de présence de la source, il pourra être ignoré par d'autres *dropout*.

Pour la caractérisation des incertitudes, deux types d'erreurs peuvent être distingués (Kendall & Gal, 2017). D'une part, il existe une incertitude sur les données acquises, du fait par exemple dans le cas de la spectro-identification, d'une faible statistique de comptage (bruit statistique) ou encore de dérives de la calibration. Cette incertitude est appelée erreur aléatorique, elle correspond au bruit intrinsèque dans les observations. D'autre part, l'algorithme ou le modèle utilisé pour traiter les données n'est pas parfait et possède lui-même une incertitude intrinsèque, appelée erreur épistémique. Les réseaux de neurones bayésiens renvoient une incertitude globale qui prend en compte à la fois l'erreur aléatorique et l'erreur épistémique. In fine, il serait intéressant d'être capable, sur chaque test individuel du réseau de neurones, d'estimer la contribution de l'erreur épistémique et celle de l'erreur aléatorique indépendamment. Cette information serait importante pour comprendre les origines des incertitudes sur le résultat des réseaux.

L'erreur épistémique dépend fortement de deux facteurs. Le premier est le nombre d'exemples dans la base d'apprentissage : un nombre d'exemples élevé permet aux réseaux de neurones de mieux généraliser et de diminuer l'erreur épistémique. Le second facteur est la fidélité de la base d'apprentissage aux données d'application. En effet, si le réseau de neurones est entraîné sur des données d'apprentissage suivant une distribution, et que la distribution des données de l'application visée est différente, même avec une bonne généralisation, le réseau de neurones pourra commettre des erreurs à cause de cette différence entre les données d'entraînement et les données de l'application. Dans le cadre de la spectro-identification, on peut penser à des cas où on entraîne le réseau de neurones uniquement avec des spectres avec une grande statistique de comptage, mais où on les applique à des spectres avec de faibles statistiques de comptage : du fait de la différence entre les données d'entraînement et les données d'application, l'erreur épistémique du réseau de neurones sera plus élevée.

Dans la suite de ce manuscrit, je me concentrerai principalement sur l'utilisation de réseaux de neurones bayésiens, avec évaluation de l'incertitude. Je souligne que cette incertitude est autoévaluée par le réseau et qu'il faut prendre la précaution de ne pas la considérer comme une incertitude absolue. En particulier, sa validité statistique doit être vérifiée. J'ai aussi effectué dans un premier temps le travail de conception de réseaux de neurones non bayésiens et d'évaluation de leurs performances, avec les six sources que nous possédions en laboratoire, dont la description et les résultats ont été publiés dans un article de la revue IEEE Transactions On Nuclear Sciences (Daniel, et al., 2020).

3.3 Construction des réseaux de neurones et apprentissage

Afin de construire et mettre en œuvre l'apprentissage des réseaux de neurones, j'ai utilisé la bibliothèque Python *Keras* (Chollet, 2015), reposant sur l'outil *tensorflow*, qui présente l'avantage de pouvoir concevoir directement des réseaux de neurones et pour laquelle les algorithmes

d'optimisation décrits en paragraphe 2.2.2 sont déjà implémentés. Par ailleurs, cette bibliothèque permet l'utilisation des GPU (*Graphics Processing Units*) qui peuvent paralléliser les opérations élémentaires de la *backpropagation* et accélérer grandement le calcul, jusqu'à un facteur 30 dans mon application par rapport à un CPU (*Central Processing Unit*).

J'ai effectué ces calculs avec une carte graphique RTX 2080 Ti. À titre de comparaison, la vitesse d'exécution d'une itération du calcul sur l'ensemble de la base de données est de 30 secondes avec GPU alors qu'il est de 10 minutes sur CPU classique.

Comme je l'ai décrit en paragraphe 3.1.1, je construis deux réseaux de neurones par radioélément de la bibliothèque : un pour l'identification du radioélément et l'autre pour la quantification de ses proportions dans le signal. Pour les deux réseaux, j'applique la méthode du dropout dans la phase d'application pour estimer les incertitudes sur les résultats de ces réseaux.

La base de données d'apprentissage est construite sur le principe décrit en paragraphe 3.1.2 avec la normalisation des spectres présentée en paragraphe 3.1.3. J'ai divisé cette base de données en deux parts. 90 % de cette base de données est effectivement utilisée comme données d'entraînement. Les 10 % restantes sont utilisées comme jeu de validation. Cela signifie qu'à chaque itération de l'apprentissage, la fonction de coût est aussi évaluée sur ce jeu de validation, ses éléments ne sont pas utilisés comme exemples d'apprentissage.

Ce jeu de validation permet notamment d'évaluer la bonne convergence du réseau de neurones et de s'assurer qu'il n'y a pas de surapprentissage. En effet, si les performances sont bien meilleures sur le jeu d'apprentissage que sur le jeu de validation, cela signifie qu'il y a surapprentissage sur les données d'entraînement. Par ailleurs, le jeu de validation permet aussi de mettre en place un *early stopping*, qui arrête l'apprentissage lorsque la fonction de coût sur le jeu de validation ne diminue plus après un certain nombre d'époques, que j'ai fixé à 10 dans mon cas.

En ce qui concerne les fonctions de coût, j'ai utilisé la *binary cross-entropy* pour les réseaux de neurones d'identification, vu qu'ils répondent à un problème de classification, et la distance euclidienne pour les réseaux de quantification, vu qu'ils répondent à un problème de régression. La seule exception concerne le réseau de neurones d'identification du ^{137}Cs, pour lequel j'ai utilisé la distance euclidienne comme fonction de coût, pour laquelle j'ai constaté en évaluant les performances comme décrit en 4.3.1 que l'apprentissage aboutissait à de meilleurs résultats.

J'ai utilisé l'algorithme d'optimisation *batch*-SGD pour les réseaux d'identification, avec un taux d'apprentissage de 0,01, un moment de 0,9 et le gradient accéléré de Nesterov (Sutskever, et al., 2013). Pour les réseaux de neurones de quantification, j'ai utilisé l'algorithme Adam (Kingma & Ba, 2015) avec des moments $\beta_1 = 0{,}9$ et $\beta_2 = 0{,}999$, et un taux d'apprentissage de 10^{-6}, et une diminution de 10^{-7} du taux d'apprentissage à chaque itération. J'ai utilisé des *batchs* de taille 1000, à la fois pour l'identification et la quantification des proportions.

J'ai déterminé l'ensemble de ces paramètres empiriquement : dans un premier temps, en surveillant l'apprentissage et la fonction de coût sur quelques radioéléments, puis dans un deuxième temps, en testant sur quelques exemples le réseau de neurones pour vérifier que le résultat était bien cohérent. L'arrêt de l'apprentissage s'effectue soit avec l'*early stopping* au bout de 10 itérations successives pour lesquelles la fonction de coût ne diminue plus sur le jeu de validation, soit au bout de 100 itérations totales. Il est possible d'aller au-delà de 100 itérations, au risque que le réseau de neurones entre dans une phase de sur-apprentissage. Par ailleurs, j'ai choisi ces 100 itérations en raison du temps de calcul

total, qui était de l'ordre de 24 h pour faire l'apprentissage de l'ensemble des 14 réseaux de neurones avec 100 itérations. Ce délai me permet un développement et une optimisation rapides. Réduire au maximum la durée de l'apprentissage est important pour le développement d'algorithmes de *deep learning* car le principe de développement est itératif. Il faut tester différents hyperparamètres (architectures de réseaux, taux d'apprentissage, régularisation…), effectuer l'apprentissage puis évaluer les performances, et ceci à chaque test d'hyperparamètres, d'où l'importance que la phase d'apprentissage soit la plus rapide possible.

Maintenant, j'en viens à la conception de l'architecture du réseau de neurones. Il est assez courant pour ce type de problème de s'inspirer de la littérature. Cependant, au moment où j'ai commencé à étudier les réseaux de neurones de convolution, il n'existait pas dans la littérature de travaux qui s'étaient intéressés à l'utilisation de ce type de réseaux de neurones pour l'analyse en spectrométrie gamma. Je suis donc parti d'une « feuille blanche » et j'ai construit méthodiquement mon réseau de neurones. Il faut noter que depuis 2017, diverses équipes ont publié leurs travaux sur cette problématique (Kamuda, et al., 2017) (Liang, et al., 2019) (Kamuda, et al., 2020).

Pour ce faire, j'ai commencé par un réseau de neurones *fully-connected* avec une seule couche cachée et une dizaine de neurones. Puis, j'ai augmenté le nombre de neurones dans cette couche et à chaque fois, j'ai entraîné mon réseau et étudié ses performances sur le jeu de validation. Lorsque les performances ne s'amélioraient plus significativement, j'ai ajouté une couche de convolution avec quatre neurones, suivie d'une opération de *batch-normalization* et de *pooling*, décrites en paragraphe 2.2.3.

Sur les neurones de convolution, j'ai d'abord adapté la taille des filtres en l'augmentant progressivement, de même jusqu'à ce que je ne constate plus d'augmentation significative des performances du réseau.

J'ai réitéré ce processus en ajoutant d'abord des neurones, puis de nouvelles couches de convolution pour obtenir mon réseau de neurones final. J'ai utilisé deux radionucléides, le ^{241}Am et le ^{133}Ba pour vérifier que la même architecture obtenait de bonnes performances sur ces deux radionucléides. J'ai appliqué ensuite la même architecture à l'ensemble des radioéléments et pour les deux problèmes, identification et quantification. Il est possible d'adapter l'architecture à chaque radioélément, cependant, il faut le faire pour chacun des 28 réseaux de neurones. Cela est cependant envisageable si on pense à des applications plus spécifiques où le nombre de radioéléments d'intérêt est plus restreint, de l'ordre de deux ou trois. Dans certains cas, la recherche de l'architecture du réseau et des hyperparamètres se fait de manière algorithmique. Par exemple, l'architecture peut être optimisée avec une recherche par algorithme génétique (Whitley, et al., 1990).

Sur chaque couche intermédiaire, j'ai utilisé la fonction d'activation ReLU, décrite au paragraphe 2.1.1. Pour la couche de sortie, comme chaque réseau de neurones est spécifique à un radioélément, on peut penser ne mettre qu'un seul neurone de sortie, avec une fonction d'activation sigmoïde[*]. En effet, pour l'identification, on attend un nombre qui vaut 0 ou 1 et pour la quantification, des valeurs

[*] La fonction sigmoïde est un cas particulier de la fonction logistique. Elle est définie par :

$$\sigma(x) = \frac{1}{1 + e^{-x}}$$

Elle est strictement croissante, vaut 0 en $-\infty$ et 1 en $+\infty$. Cette fonction est souvent utilisée pour des problèmes de classification, dans des cas multi-classes non exclusives, c'est-à-dire où plusieurs classes peuvent valoir 1 simultanément pour un même exemple.

comprises entre 0 et 1. En pratique, les réseaux de neurones sont plus performants lorsque j'utilise deux sorties complémentaires, avec une fonction d'activation *softmax*, allant jusqu'à une réduction absolue de 5 % du taux d'erreur sur certains radioéléments comme le ^{133}Ba par rapport à une seule sortie sigmoïde, comme je le montre en Annexe C.

Pour le réseau d'identification, le premier neurone vaut 1 si la source est présente et 0 si elle est absente, et inversement, le second neurone vaut 0 si elle est présente et 1 si elle est présente. De même, pour le réseau de quantification, le premier neurone contient la proportion du signal qui provient du radioélément d'intérêt et le deuxième neurone la proportion du signal qui ne provient pas de ce radioélément.

L'idée derrière l'utilisation de deux sorties est de compenser la tendance des valeurs uniques en sortie de la fonction sigmoïde à être supérieures à 0,5 dans le cas des réseaux de neurones d'identification. En utilisant deux sorties complémentaires, cela montre au réseau de neurones que s'il attribue une valeur supérieure à 0,5 pour la présence de la source sur le premier neurone, il attribue par conséquent une valeur inférieure à 0,5 pour l'absence de la source sur le second neurone. Cette méthode agit comme une double pénalité si le réseau de neurones fait une erreur dans la classification[*]. Expérimentalement, j'ai vérifié cela en testant les performances avec le réseau de neurones de convolution non bayésien, sur nos six sources de laboratoire, le résultat est présenté en Annexe C.

Quant à l'aspect bayésien, j'ai appliqué des opérations de *dropout* sur l'ensemble des couches intermédiaires, avec un taux de 50 %. L'ajustement du taux de *dropout* peut faire l'objet d'une optimisation, un taux trop faible ne permettant pas d'estimer correctement l'incertitude alors qu'un taux trop élevé peut empêcher l'apprentissage de s'effectuer correctement. Pour ma part, ce taux de 50 % correspond à celui utilisé dans (Gal & Ghahramani, 2016), appliqué aux données de la base de données MNIST pour la reconnaissance de chiffres manuscrits. Je rappelle que ces opérations de *dropout* restent actives dans la phase de test pour l'évaluation de l'incertitude.

Finalement, la Figure 3-7 résume l'architecture de réseaux de neurones à laquelle je suis parvenu avec la méthode décrite précédemment. L'ensemble de ces méthodes a fait l'objet d'une demande de dépôt de brevet en mai 2019 sous la référence 72124 FR, demande actuellement étudiée par l'INPI (Institut National de la Propriété Industrielle).

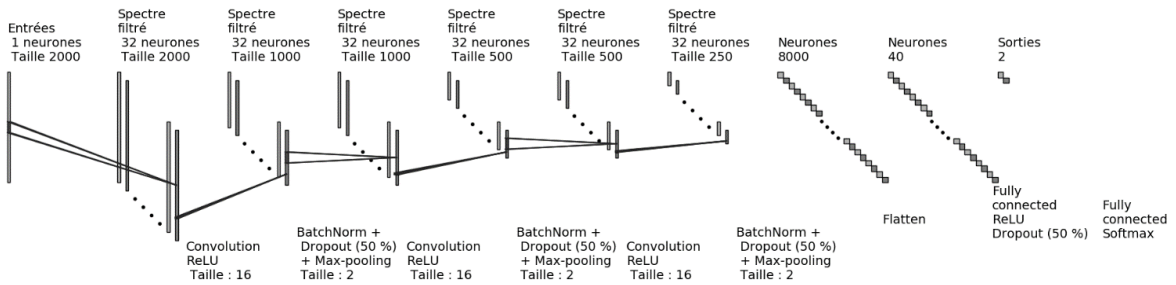


Figure 3-7 : Architecture des réseaux de neurones.

[*] Je souligne qu'ici, j'ai utilisé dans les deux cas la fonction d'entropie croisée binaire définie par l'équation (2-4) et non l'entropie croisée « classique » définie par $L(Y,\hat{Y}) = -\sum_i y_i \log(\hat{y}_i)$. Utiliser deux sorties avec une fonction d'activation *softmax* et l'entropie croisée « classique » serait équivalent à utiliser une seule sortie avec une fonction d'activation sigmoïde et l'entropie croisée binaire.

J'utilise trois couches de convolution, chacune constituée de 32 neurones, avec des filtres à une dimension de taille 16. Chaque couche de convolution est suivie d'une opération de *batch-normalization*, d'un *dropout* avec un taux de 50 % et d'une opération de *max-pooling* de taille 2. À la suite de la partie convolutive, les 32 spectres filtrés sont concaténés par l'opération *flatten*. Puis j'ai ajouté une couche *fully-connected* de 40 neurones avec une opération de *dropout* et enfin la couche de sortie constituée des deux neurones de sortie associés à une fonction d'activation *softmax*.

La Figure 3-8 montre l'évolution de la fonction de coût au fur et à mesure de l'apprentissage pour les jeux d'apprentissage et de validation avec l'exemple du radioélément ^{241}Am.

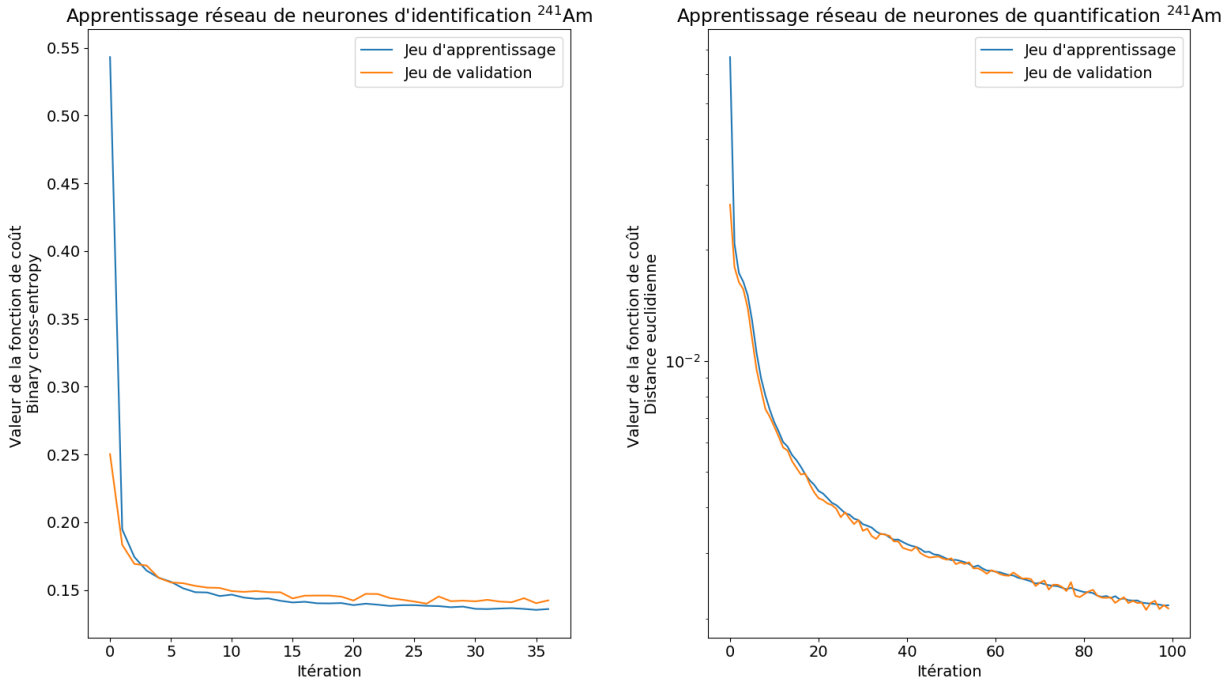


Figure 3-8 : Évolution de la fonction de coût au cours de l'apprentissage pour les réseaux de neurones dédiés au radioélément ^{241}Am. À gauche : sur le réseau de neurones d'identification. À droite : sur le réseau de neurones de quantification.

Dans cet exemple, l'apprentissage du réseau de neurones d'identification a été arrêté par l'*early stopping* alors que pour le réseau de neurones de quantification, l'apprentissage s'est arrêté au bout de 100 itérations.

Une fois les réseaux de neurone entraînés, je peux les évaluer sur des données de test, qui sont des données réelles issues de Caliste, que je présente dans la partie suivante.

4 Application des réseaux de neurones à la spectro-identification de radioéléments

4.1 Constitution de la base de données de test

Afin de constituer une base de données de test, j'ai utilisé des données issues de mesures que j'ai prises dans notre laboratoire. En utilisant les sources que j'avais à ma disposition, j'ai utilisé le

détecteur Caliste-HD dans la caméra gamma WIX (Maier, et al., 2018) pour faire des acquisitions longues, la source étant placée directement devant le détecteur, de façon à avoir un très grand nombre de photons.

Pour la source de ^{22}Na, j'ai utilisé le détecteur Caliste-O dans son banc de test décrit en Chapitre 2. Même si le détecteur Caliste-O n'a pas exactement les mêmes caractéristiques que Caliste-HD au niveau de l'épaisseur et de la taille des pixels, j'estime de cette manière la robustesse de mon apprentissage à des variations de configuration du système de détection. Pour la source de ^{226}Ra-D, j'ai utilisé un ancien réveil comportant une peinture en Radium, dont l'activité est inconnue mais reste très faible.

Pour chacune de ces mesures, je n'utilise que les événements simples enregistrés avec le détecteur, pour leur meilleure résolution spectrale et de façon à être en accord avec les données d'apprentissage.

Le Tableau 4-1 résume les caractéristiques des sources utilisées et des mesures effectuées.

Radioélément	Activité (kBq)	Distance au détecteur (cm)	Durée de la mesure	Détecteur utilisé	Nombre d'événements simples
^{241}Am	436	~ 1	30 min	Caliste-HD	9 220 768
^{133}Ba	130	~ 1	1 h 30 min	Caliste-HD	15 852 843
^{57}Co	52	~ 1	2 h 52 min	Caliste-HD	3 394 653
^{137}Cs	$1.6.10^3$	~ 1	1 h 12 min	Caliste-HD	14 217 618
^{152}Eu	66	~ 1	2 h 24 min	Caliste-HD	12 487 896
^{22}Na	3.5	~ 8	133 h	Caliste-O	6 853 970
^{226}Ra-D	Inconnue – Peinture de réveil (très faible)	~ 1	17 h	Caliste-HD	658 532

Tableau 4-1 : Caractéristiques des sources et mesures.

La Figure 4-1 montre l'ensemble des spectres réalisés à partir de chacune de ces mesures. J'ai utilisé une calibration linéaire, faite à partir de données acquises pour le ^{137}Cs. Idéalement, on pourrait prendre en compte la non-linéarité décrite dans le Chapitre 2, mais l'intérêt ici est de pouvoir aussi vérifier le résultat à partir d'une calibration linéaire. Cette configuration sera privilégiée pour la caméra, et cela me permet d'évaluer la réponse des réseaux de neurones avec ce type de calibration.

Les mesures que j'ai faites sont des acquisitions sur des radioéléments individuels, un à un. Pour chaque mesure, je dispose de la liste de chaque événement individuel enregistré (position, temps d'arrivée, énergie et multiplicité). À partir des listes d'événements de chaque mesure individuelle, je

peux former artificiellement des mélanges en utilisant un processus d'augmentation de données réelles.

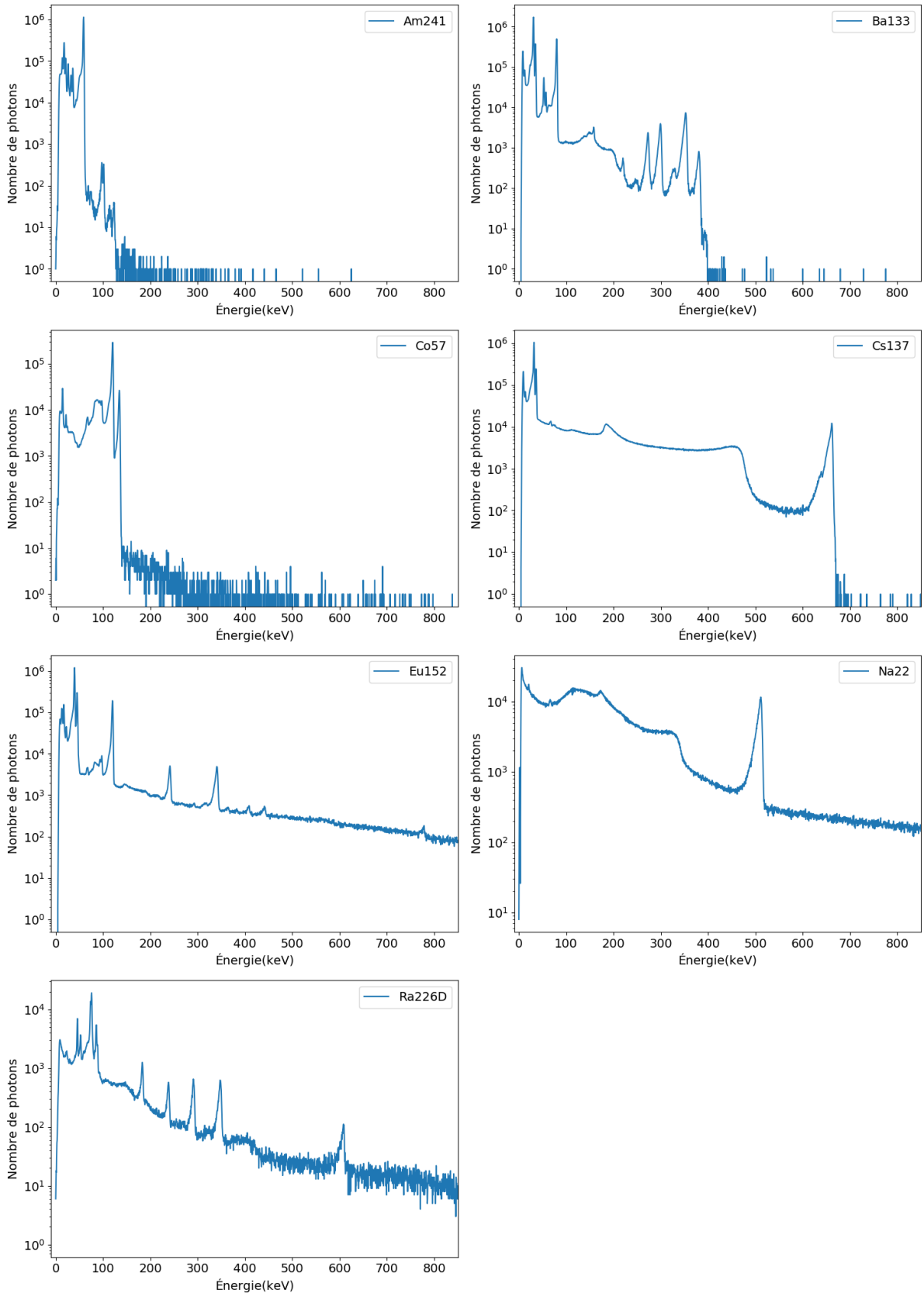


Figure 4-1 : Spectres issus des mesures avec Caliste.

Ainsi, j'applique le même processus d'augmentation de données que décrit en 3.1.2, à la différence près que j'utilise non pas des données de simulation, mais les mesures décrites précédemment. J'applique de même une décalibration en gains et *offsets* sur une partie de la base de test, de façon à vérifier la robustesse de l'algorithme à des décalibrations, ce qui d'ailleurs compense l'effet de la non-linéarité de la calibration évoquée précédemment.

Par ailleurs, lors de la création des mélanges, je ne tire pas aléatoirement le nombre de photons utilisés au total, j'utilise des valeurs fixes (10, 20, 30…, 100, 200 etc.), de façon à évaluer les performances des réseaux de neurones en fonction du nombre de photons dans le spectre.

Un autre intérêt de mon processus de *data augmentation* est que je peux contrôler le nombre de photons provenant de chaque mesure individuelle. En effet, pour des mélanges réels, on ne connaît que l'activité de la source. Pour connaître le nombre de photons appartenant à chaque radioélément détecté, il faudrait idéalement connaître la distance de la source au détecteur, ses lignes d'émission, l'efficacité du détecteur pour chaque photon émis, mais aussi prendre en compte l'ensemble des diffusions Compton dans l'environnement. Tout cela rendrait très complexe la connaissance de la proportion de chaque radioélément dans le signal. C'est une inconnue que je peux contrôler par mon processus d'augmentation de données, créant des mélanges artificiels à partir de mesures réelles.

4.2 Mise en œuvre des réseaux et quantification des performances

4.2.1 Représentation des résultats et cas d'application

Avant d'évaluer les performances sur un jeu de données complet, je montre ici quelques exemples de résultats en détaillant ma méthode de test à l'aide des réseaux de neurones convolutifs bayésiens que j'ai conçus.

Je montre dans ce paragraphe l'application des réseaux de neurones sur trois exemples :

- un premier exemple avec une acquisition à grande statistique de comptage, 30 millions de photons, d'un mélange de ^{137}Cs, ^{57}Co et ^{152}Eu qui illustre le fonctionnement des réseaux sur un mélange ;
- un deuxième exemple avec 150 photons de ^{241}Am et de ^{152}Eu, qui illustre les performances des réseaux de neurones dans un cas à faible statistique de comptage ;
- un troisième exemple avec une acquisition comportant deux radioéléments, 10 000 photons de ^{241}Am et 200 photons de ^{57}Co, qui illustre les performances des réseaux dans le cas où une des sources est en faible proportion dans le signal.

Pour chaque spectre à tester, j'applique les normalisations que j'ai décrites en paragraphe 3.1.3. Dans un premier temps, j'applique les réseaux de neurones d'identification. Du fait de l'approximation bayésienne par *dropout* que j'ai présentée en paragraphe 3.2, je dois appliquer plusieurs fois chaque réseau de neurones au même exemple afin d'obtenir une distribution de résultats, qui caractérisera l'incertitude. J'ai choisi d'appliquer 100 fois le réseau de neurones, donnant ainsi une distribution avec 100 résultats. Le nombre d'applications peut être plus grand si on souhaite caractériser plus finement l'incertitude ou plus petit si on se soucie de la vitesse d'exécution en priorité, en fonction des performances du calculateur utilisé.

Par exemple, sur un ordinateur de bureau équipé d'un processeur Intel® Core™ i5-7300U CPU (2.6 GHz, 2.7 GHz), la vitesse d'exécution pour un seul réseau de neurones sur 100 exemples est de 120 ms, soit 1,7 s pour 14 sources. Si on souhaite être plus proche d'un calcul en temps réel, de l'ordre du dixième de seconde, on peut réduire à 10 le nombre d'applications du réseau de neurones, l'incertitude correspondante sera cependant moins bien estimée.

Sur la distribution ainsi obtenue, je conserve la médiane des valeurs comme résultat du réseau de neurones et je caractérise l'incertitude par le premier quartile et le troisième quartile, que je représente en Figure 4-2. Il est bien sûr possible de choisir d'autres valeurs ou d'autres représentations, à partir des résultats obtenus, on peut par exemple prendre les valeurs des quantiles à 5 % et 95 % si on souhaite être plus conservatif.

En sortie des réseaux de neurones d'identification, on obtient 100 valeurs comprises entre 0 et 1, qui se représentent usuellement par une probabilité de présence de la classe. Il faut ensuite appliquer un seuil de décision, c'est-à-dire, décider à partir de quelle probabilité on considère que la source est effectivement présente ou non. Dans ces exemples, j'ai pris un seuil de décision à 50 %, en tenant compte des incertitudes de la manière suivante :

- si le troisième quartile de la distribution des résultats est inférieur à ce seuil, je considère que la source est absente, ce que je représente en bleu sur la figure ;
- si le premier quartile est supérieur à ce seuil, je considère que la source est présente, ce que je représente en rouge sur la figure ;
- si le premier quartile et le troisième quartile sont de part et d'autre de ce seuil, je considère que le réseau de neurones est incertain, ce que je représente en orange sur la figure.

La valeur de ce seuil peut être adaptée à l'aide d'un jeu de test comme je le montre en paragraphe 4.3.2. Je fais une adaptation du seuil qui dépend du radioélément considéré, car certains sont plus facilement détectés que d'autres, comme le ^{241}Am qui a des pics caractéristiques à des énergies où le détecteur Caliste est très efficace, à 60 keV. Il est aussi possible de faire un choix en fonction du nombre de photons reçus, en utilisant un seuil plus ou moins élevé, c'est-à-dire en étant plus ou moins conservatif.

La représentation que j'ai choisie peut bien sûr être sujette à discussion en fonction de l'application visée. Dans certains cas, on préfère s'assurer de l'absence de sources et il est préférable d'éviter les faux négatifs, comme dans des applications de démantèlement, alors que dans d'autres cas, on veut être en mesure d'éviter les fausses alarmes et n'intervenir que dans le cas où un radioélément en quantité significative a été détecté.

Une fois le ou les radioéléments détectés, j'applique les réseaux de neurones de quantification des proportions, à nouveau en effectuant 100 fois le même calcul sur le spectre. J'utilise les caractérisations avec la médiane, le premier et le troisième quartile, ainsi que les valeurs à 5 % et 95 % pour caractériser ici plus finement l'incertitude dans ce cas de régression. Le résultat est présenté en deuxième partie de la Figure 4-2.

Sur cet exemple, j'ai formé un mélange de ^{137}Cs, ^{57}Co et ^{152}Eu en utilisant la méthode décrite en paragraphe 4.1. Avec cette méthode, je maîtrise les proportions de chaque radioélément : le mélange

contient 11,3 % de ^{57}Co, 47,2 % de ^{137}Cs et 41,5 % de ^{152}Eu, sur un total de 30 millions de photons. Je montre ici un cas avec un grand nombre de photons, où le spectre est plus facilement reconnaissable.

Les trois sources utilisées sont détectées avec certitude, les barres d'incertitudes étant au-dessus du seuil de détection que j'ai fixé à 50 %. Les autres sources ne sont pas détectées, les valeurs associées et leur incertitude étant en-dessous du seuil de détection. Certaines sources ont une incertitude un peu plus grande comme le ^{235}U-D. Cela provient vraisemblablement du pic de rétrodiffusion Compton à 184 keV dû à la source de ^{137}Cs qui peut être assimilé à la ligne d'émission à 185 keV du ^{235}U-D. Cependant, cette ligne, bien qu'augmentant un peu l'incertitude sur une présence potentielle du radioélément, n'est pas suffisante pour affirmer à tort la détection du ^{235}U-D. Au niveau des proportions, on voit que les réseaux de neurones les estiment correctement au vu de l'incertitude sur la quantification.

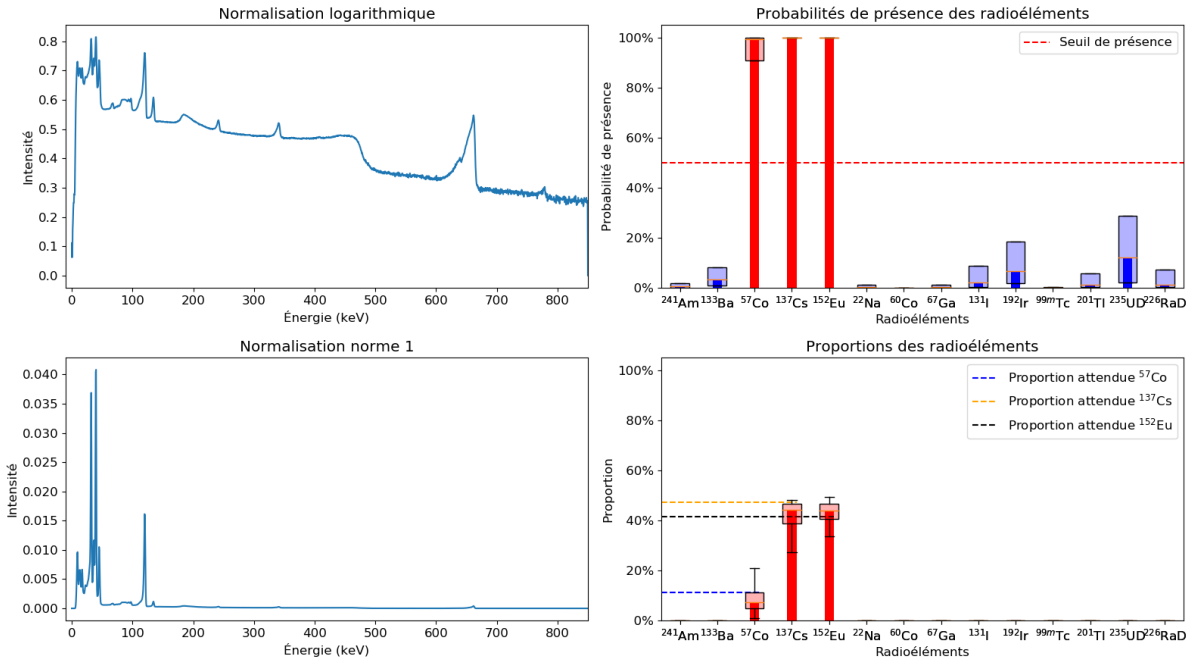


Figure 4-2 : Exemple de test, sources utilisées : ^{137}Cs, ^{57}Co et ^{152}Eu.

Je montre en Figure 4-3 un autre exemple avec un nombre plus limité de photon, où un expert pourrait plus difficilement affirmer ou rejeter la présence des sources. La mesure est composée de 100 photons de ^{241}Am et 50 photons de ^{152}Eu.

Les sources sont de nouveaux détectées correctement avec une faible incertitude, inférieure à 1 %, et aucune autre source n'est détectée. De même, les proportions sont estimées correctement, la barre d'erreur est assez large, près de 40 % sur les quantiles à 5 % et 95 %, du fait du faible nombre de photons dans le spectre utilisé.

Dans le spectre, le nombre de photons est suffisant pour identifier visuellement le ^{241}Am, dont le spectre est concentré dans les basses énergies avec des pics photoélectriques caractéristiques, notamment à 60 keV. En revanche, la source de ^{152}Eu est beaucoup moins évidente à reconnaître avec si peu de photons, son spectre étant beaucoup plus étalé.

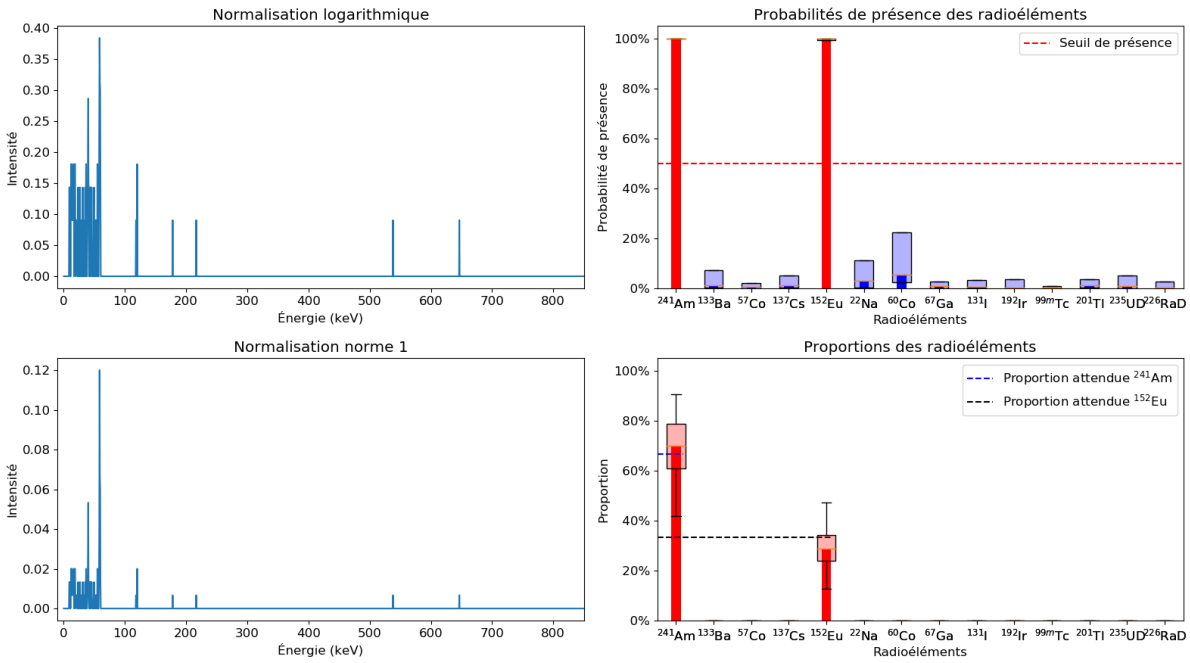


Figure 4-3 : Exemple de test, faible nombre de photons, sources utilisées : ^{241}Am et ^{152}Eu.

Enfin la Figure 4-4 montre un exemple avec un spectre contenant 10 000 photons de ^{241}Am et de seulement 200 photons de ^{57}Co.

On voit notamment l'intérêt de l'utilisation des deux réseaux de neurones séparés, où la proportion de ^{57}Co est faible. Les réseaux d'identification la détectent malgré sa faible intensité, certes avec une plus grande barre d'incertitude mais qui reste toujours supérieure au seuil de détection ce qui permet d'affirmer que la source est présente. Avec le résultat du réseau de neurones de quantification seul, il n'est pas possible de confirmer la présence de cette source.

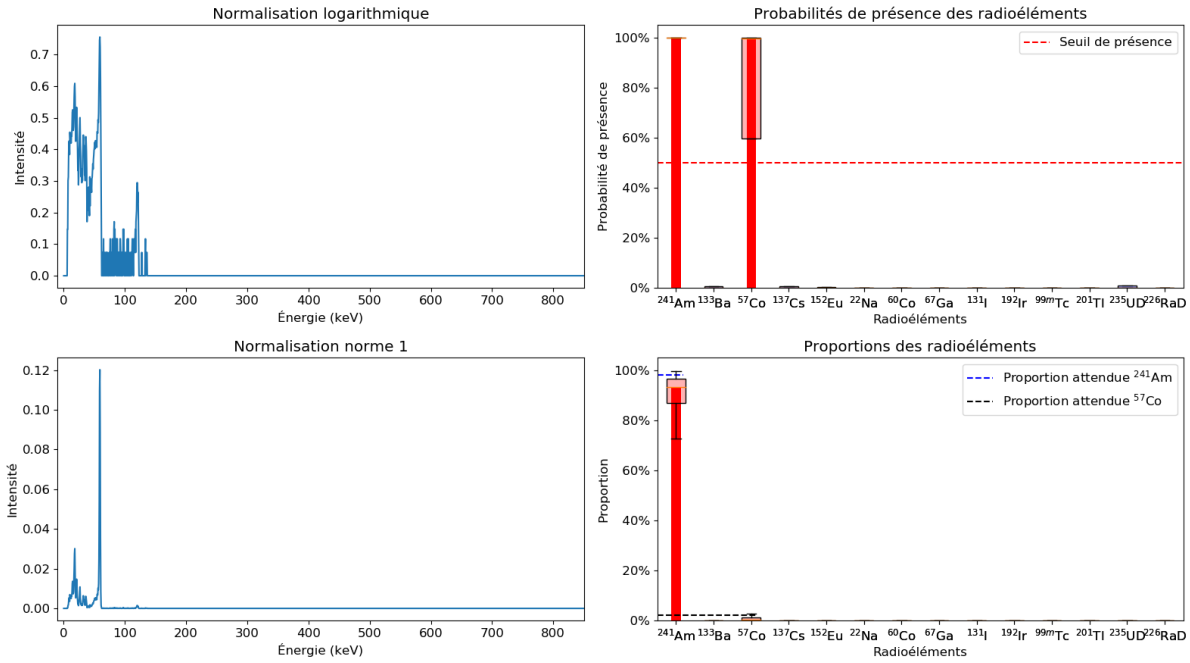


Figure 4-4 : Exemple de test, source à faible proportion, sources utilisées : ^{241}Am et ^{57}Co.

4.2.2 Comment « réfléchit » mon réseau de neurones convolutif ?

Un point très intéressant à étudier est la compréhension des structures extraites par les réseaux de neurones à couches de convolution. Ces couches agissent comme un prétraitement des données avant la couche *fully-connected* en filtrant les caractéristiques pertinentes du point de vue du réseau pour identifier les radioéléments ou les quantifier. Certaines de ces caractéristiques sont interprétables physiquement.

La Figure 4-5 montre quelques neurones à l'intérieur des couches de convolution pour le réseau de neurones d'identification du ^{137}Cs, appliqué à un spectre de mélange de ^{137}Cs et ^{57}Co.

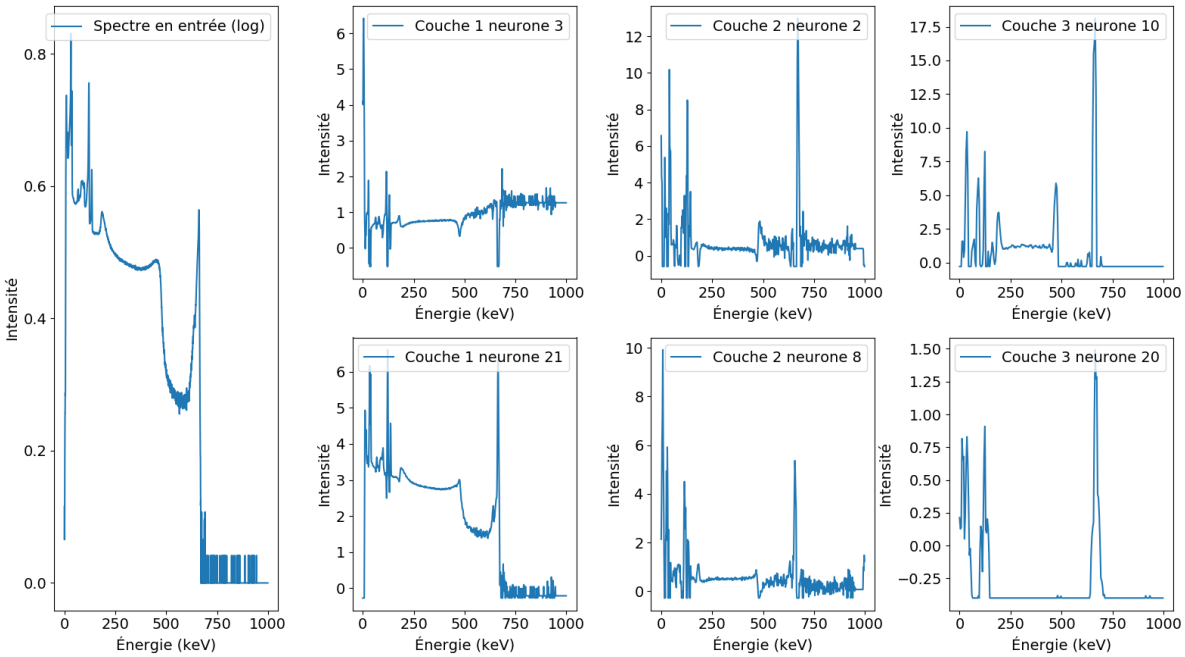


Figure 4-5 : À l'intérieur des couches de convolution du réseau de neurones d'identification du ^{137}Cs.

Pour la première couche de convolution, on peut voir que le rôle du neurone 21 (cf. Figure 4-5) est principalement de débruiter les données et d'intensifier les zones à fort gradient. Le neurone 3 extrait de son côté les positions des pics en cherchant vraisemblablement les zones où le gradient est négatif.

La deuxième couche met en valeur les structures physiques d'intérêt du spectre, notamment les pics photoélectriques, où le pic à 662 keV devient très intense pour le neurone 2. Pour le neurone 8, ce sont les pics à basse énergie qui sont principalement importants.

Le résultat de la troisième couche est le plus important, car c'est ce résultat qui est présenté aux neurones *fully-connected* utilisés pour l'identification. Si on commence par le neurone 20, on voit que ce sont tous les pics photoélectriques qui sont extraits : les pics à basse énergie, notamment 14 keV, 121 keV et 136 keV pour le ^{57}Co et 32 keV, 36 keV, qui apparaissent confondus, et enfin principalement 662 keV pour le ^{137}Cs. On voit notamment que le pic à 662 keV apparaît très distinctement, ce qui montre que, sur cet exemple, la forme de ce pic est discriminante. Les pics du ^{57}Co sont aussi extraits et c'est ensuite à la partie *fully-connected* de trier les éléments qui caractérisent le ^{137}Cs.

Le neurone 10 de la couche 3 met aussi en avant les pics photoélectriques, mais il est aussi particulièrement intéressant car il donne beaucoup d'importances aux structures Compton présentes

dans le spectre. En effet, on voit que le pic de rétrodiffusion à 180°, situé à 184 keV, et le front Compton à 478 keV apparaissent comme des pics dans ce neurone, ce qui montre le caractère discriminant de ces structures dans le cas de cet exemple. Elles renforcent l'identification du ^{137}Cs au-delà de la détection de ses raies principales. Par ailleurs, la structure continue Compton, entre 184 keV et 478 keV est aussi bien visible.

De ce fait, le réseau de neurones semble utiliser les caractéristiques que nous connaissons du ^{137}Cs, à savoir les pics photoélectriques et l'information dans la structure Compton qui est particulièrement discriminante.

À travers cet exemple, on comprend l'importance de la partie convolution dans les réseaux de neurones, notamment dans le filtrage des caractéristiques pertinentes pour l'identification des sources. J'ai pu montrer en particulier que nous pouvons interpréter physiquement la manière dont un réseau de neurones analyse les spectres, qui est ici assez proche de la manière dont nous pouvons les analyser à l'aide de nos connaissances physiques, auxquelles le réseau de neurones n'a pas accès de son côté.

4.3 Évaluation des performances des réseaux de neurones d'identification

4.3.1 Precision, Recall, Accuracy

L'évaluation des réseaux de neurones d'identification se fait en utilisant les méthodes d'évaluation liées aux algorithmes de classification. Pour ce type d'algorithme, on évalue habituellement le taux de classifications correctes ainsi que l'influence des faux-positifs et des faux-négatifs, que l'on dénomme par erreurs de type I et erreurs de type II respectivement.

Dans le cadre de la classification, on parle de :

- Vrais positifs (*True Positive*) lorsque les réseaux de neurones ont détecté une source effectivement présente, je note TP le nombre de vrais positifs ;
- Vrais négatifs (*True Negative*) lorsqu'ils n'ont pas détecté une source effectivement absente, je note TN le nombre de vrais négatifs ;
- Faux positifs (*False Positive*) lorsque les réseaux de neurones ont détecté une source qui est en réalité absente, je note FP le nombre de faux positifs ;
- Faux négatifs (*False Negative*) lorsqu'ils n'ont pas détecté une source qui est en réalité présente, je note FN le nombre de faux négatifs.

Je rappelle qu'on dit qu'une source est considérée comme détectée lorsque la médiane des sorties du réseau de neurones est supérieure à un certain seuil, que je fixe à 50 % dans ce paragraphe. L'influence du seuil est étudiée en paragraphe 4.3.2.

À partir de ces notations, je définis trois quantités qui permettent d'évaluer les performances des réseaux de neurones. Ces quantités sont classiques dans l'étude des performances des réseaux de neurones en général.

L'*Accuracy* sert à mesurer le taux de classifications correctes par le réseau de neurones, elle est définie par :

$$\mathrm{Accuracy} = \frac{\mathrm{TP} + \mathrm{TN}}{\mathrm{TP} + \mathrm{TN} + \mathrm{FP} + \mathrm{FN}} \tag{4-1}$$

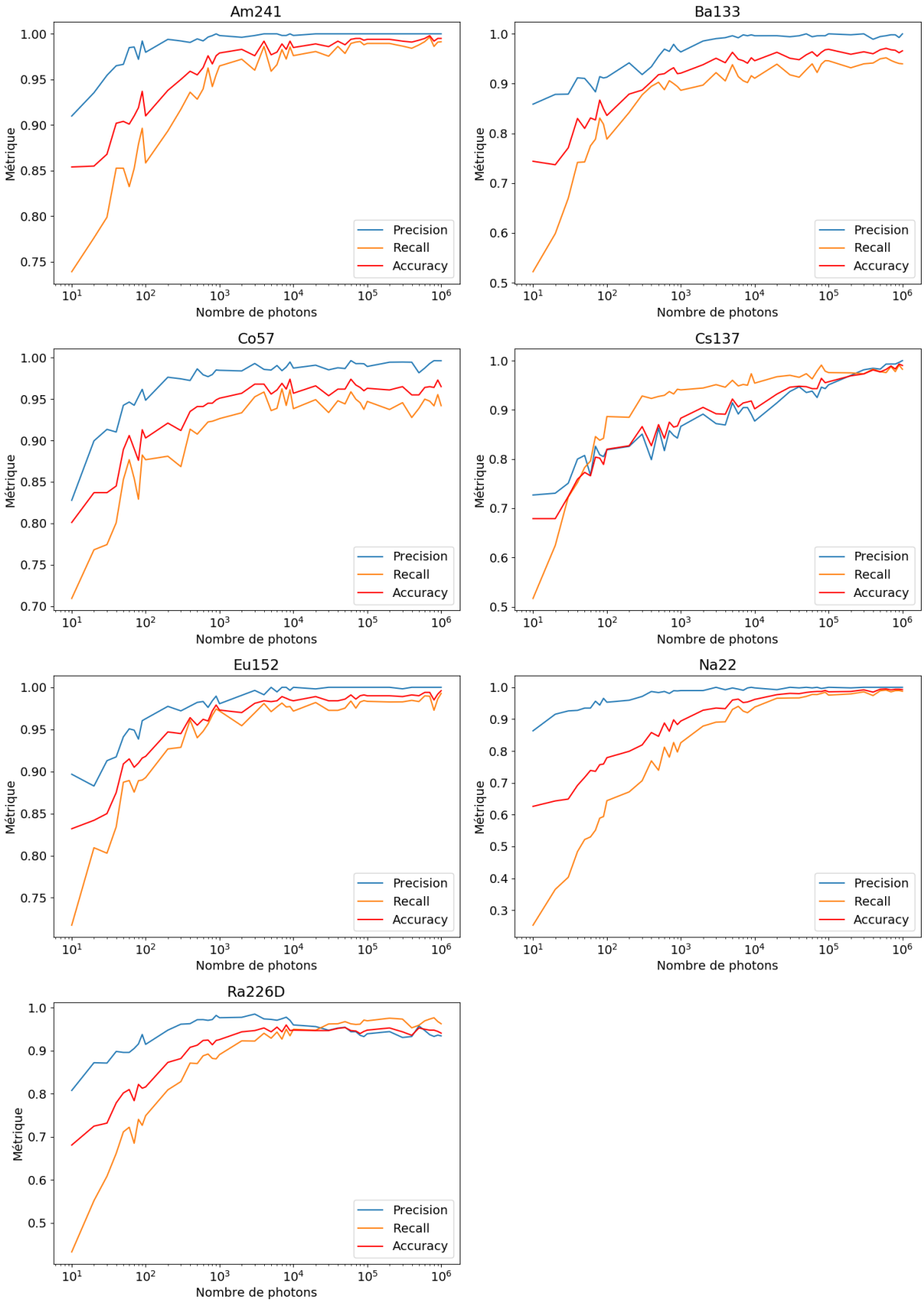


Figure 4-6 : Performances des réseaux de neurones, pour un seuil à 50 %. En utilisant 1000 exemples pour chaque statistique de comptage, la marge d'erreur des calculs effectués est de 3 %. Je n'ai pas représenté ces marges d'erreur pour des raisons de lisibilité du graphique.

La *Precision* sert à quantifier l'influence des faux-positifs. Elle est définie comme le taux de vrais positifs parmi les exemples que le réseau a classé comme positifs :

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{4-2}$$

Le *Recall* quantifie l'influence des faux-négatifs. Il est défini comme le taux de vrais positifs parmi les exemples qui contiennent en réalité le radioélément d'intérêt :

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{4-3}$$

À partir de mon jeu de test, j'ai calculé ces trois quantités pour les réseaux de neurones d'identification des radioéléments décrits en paragraphe 4.1, en fonction du nombre de photons détectés au total dans les mélanges que j'ai faits. Le résultat est présenté en Figure 4-6.

On constate que l'*accuracy* est de plus de 80 % avec une centaine de photons pour toutes les sources. Elle dépasse les 90 % avec un peu moins de 1000 photons, sauf pour le ^{137}Cs et le ^{22}Na qui ont besoin de quelques milliers de photons pour atteindre ces performances. En effet, pour ces deux sources, les deux pics discriminants sont à 662 keV et 511 keV respectivement, soit à haute énergie, ce qui explique pourquoi plus de photons sont nécessaires pour affirmer leur présence avec certitude compte tenu de l'efficacité du détecteur Caliste dans cette bande d'énergie (efficacité à 511 keV de 9,6 % et à 662 keV de 4,1 % pour 1 mm d'épaisseur de CdTe).

On notera aussi que la *precision* est plus faible pour le ^{137}Cs, ce qui s'explique notamment par un taux de faux positifs plus élevé en raison des interférences des pics de fluorescence X à 32 keV et 36 keV avec ceux du ^{133}Ba. On comprend à nouveau ici que la sensibilité du réseau dépend du choix du cocktail de sources qu'il a appris puisque les interférences entre les pics en dépendent.

Par ailleurs, j'ai montré dans (Daniel, et al., 2020) que les moins bonnes performances sur le *recall* proviennent de cas où la source est en très faible proportion dans le signal, moins de 5 % du signal, ce qui les rend plus difficiles à détecter par nature, quelle que soit la méthode.

4.3.2 Adaptation du seuil de décision

Les performances de détection dépendent fortement du seuil de décision fixé. Dans ce paragraphe, je reprends les résultats des réseaux de neurones et je regarde les performances moyennes en fonction du seuil, représentées en Figure 4-7.

Comme on peut s'y attendre, la *precision* augmente avec le seuil. Cela est normal puisqu'un seuil plus élevé est plus conservatif et donc diminue le nombre de faux positifs. Inversement, le *recall* diminue avec le seuil, du fait qu'un seuil plus élevé augmente les chances de ne pas détecter une source et donc d'avoir des faux négatifs.

En se référant à l'*accuracy*, les performances optimales sont en fait obtenues pour un seuil inférieur à 50 %, entre 25 % et 45 %, excepté pour le ^{137}Cs, pour lequel le seuil optimal est de 54 %. Ces paramètres sont réglables et ce type d'analyse permet de choisir les seuils de façon rigoureuse.

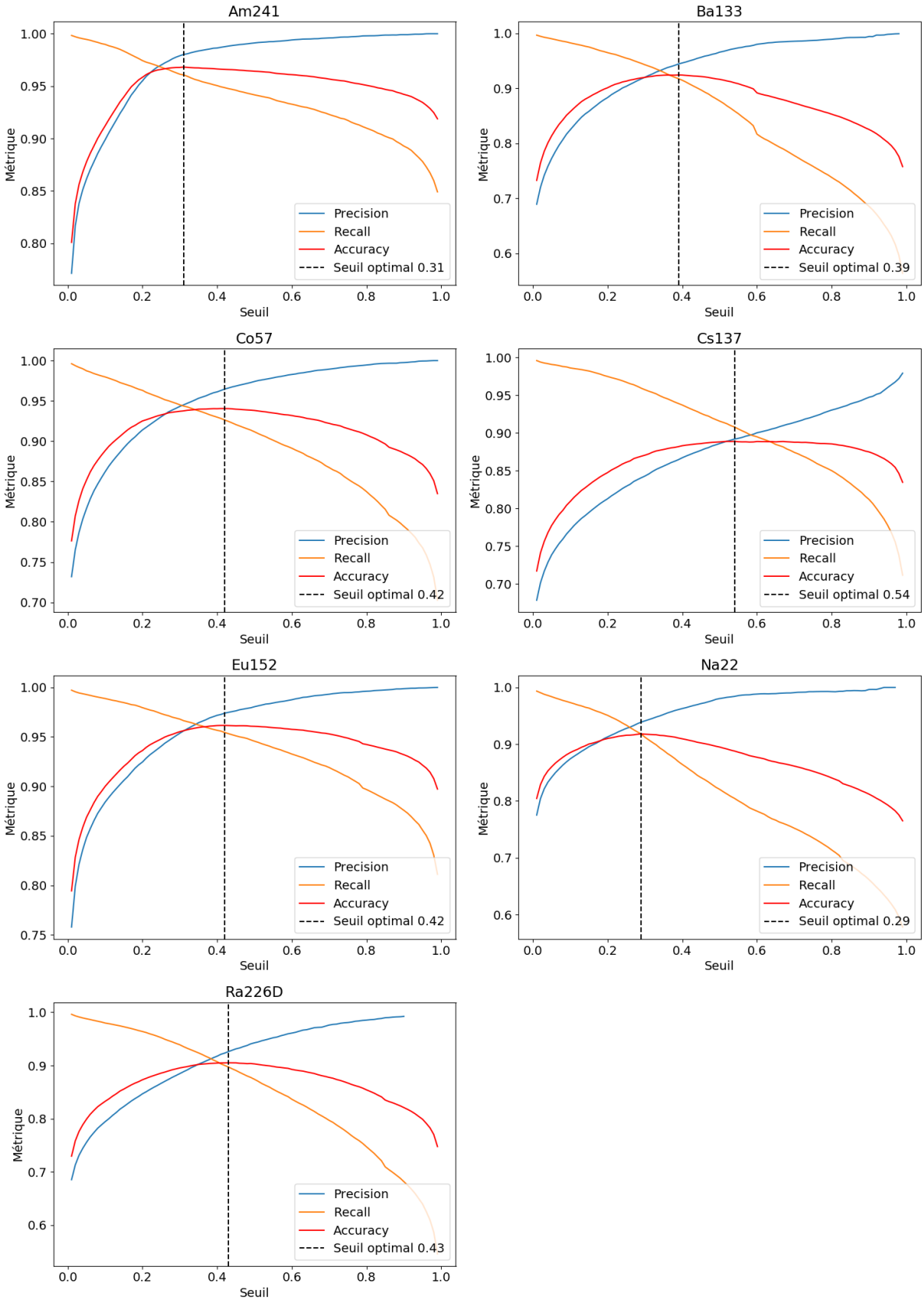


Figure 4-7 : Performances selon le seuil utilisé.

La Figure 4-8 montre les performances en fonction du nombre de photons en utilisant le seuil optimal pour chaque radioélément déterminé précédemment.

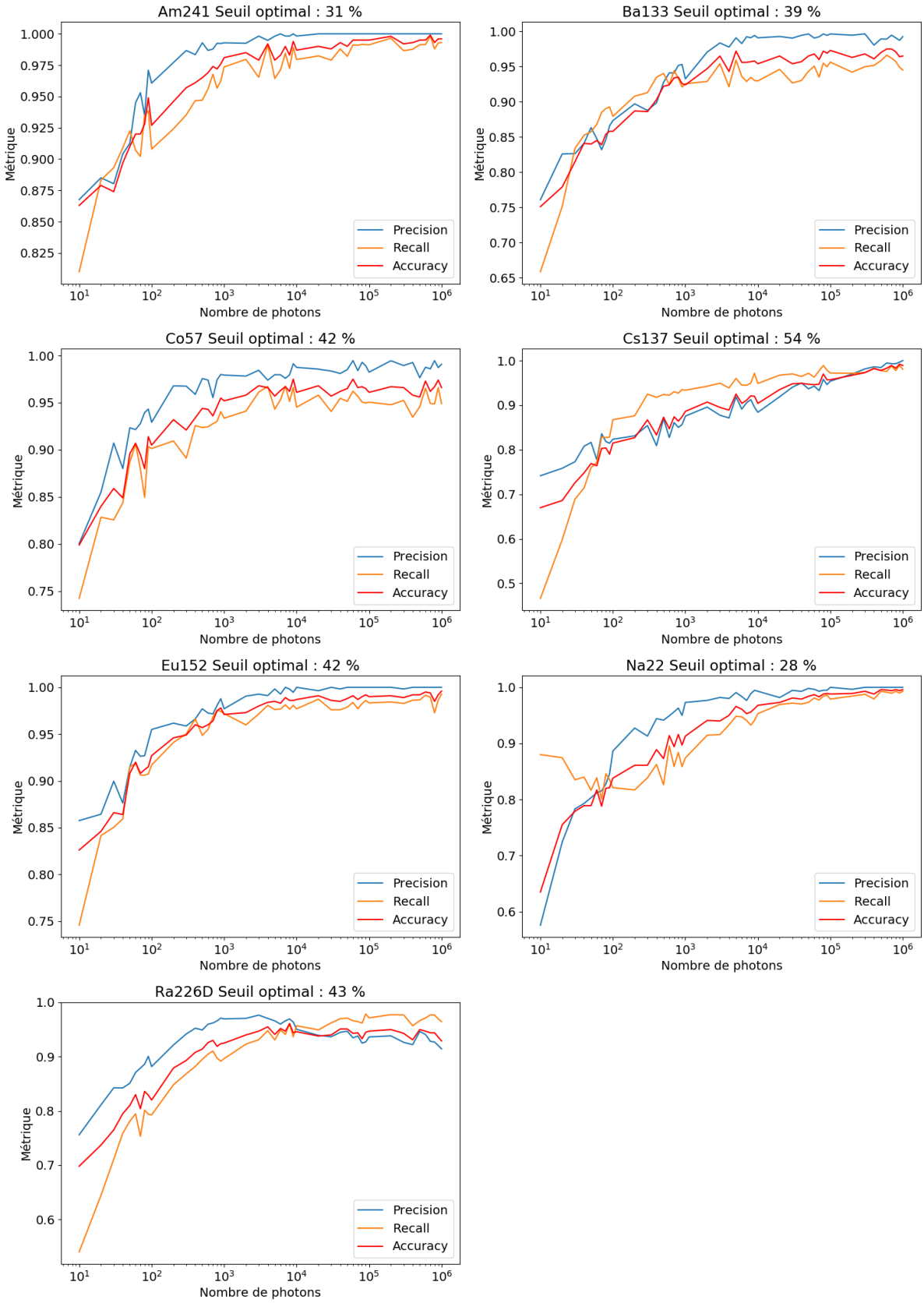


Figure 4-8 : Performances en utilisant le seuil optimal. En utilisant 1000 exemples pour chaque statistique de comptage, la marge d'erreur des calculs effectués est de 3 %. Je n'ai pas représenté ces marges d'erreur pour des raisons de lisibilité du graphique.

On voit sur l'ensemble des sources, excepté pour le ^{137}Cs, une amélioration de l'*accuracy* pour les faibles statistiques de photons, de l'ordre de 5 %, à moins de 100 photons. L'amélioration pour le ^{137}Cs est moins perceptible, le seuil optimal étant déjà proche de 50 %.

J'évalue aussi le seuil optimal pour le premier et le troisième quartile des distributions des réseaux de neurones bayésiens. Le premier quartile sert à s'assurer de la présence d'une source et le troisième quartile sert à confirmer l'absence d'une source, c'est la raison pour laquelle j'évalue le seuil optimal à partir de la *precision* pour le premier quartile et à partir du *recall* pour le troisième quartile. Je trouve le seuil optimal comme celui pour lequel la *precision* évaluée sur le premier quartile dépasse 95 % et celui pour lequel le *recall* passe sous les 95 % pour le troisième quartile. Cela correspond à un seuil de confiance de 95 % sur la présence et l'absence des sources, qui peut être ajusté suivant les spécifications de l'application visée.

Les résultats sont présentés en Figure 4-9. Cela donne de nouvelles méthodes de décision en prenant en compte les incertitudes :

- une source est considérée comme non détectée si le premier quartile et le troisième quartile sont inférieurs à leurs seuils optimaux respectifs ;
- une source est considérée comme détectée si le premier quartile et le troisième quartile sont supérieurs à leurs seuils optimaux respectifs ;
- sinon, le réseau de neurones ne peut confirmer ni l'absence ni la présence de la source.

Le Tableau 4-2 résume les seuils de décision ainsi calculés :

Source	^{241}Am	^{133}Ba	^{57}Co	^{137}Cs	^{152}Eu	^{22}Na	^{226}Ra-D
Seuil premier quartile	13 %	29 %	18 %	76 %	17 %	21 %	34 %
Seuil médiane	31 %	39 %	42 %	54 %	42 %	29 %	43 %
Seuil troisième quartile	55 %	40 %	42 %	43 %	59 %	35 %	42 %

Tableau 4-2 : Seuils de décision optimaux.

Comme cela a été remarqué précédemment, le comportement des réseaux de neurones pour le ^{137}Cs est différent de celui des autres sources. Il faut en effet un seuil beaucoup plus élevé, à 76 %, pour le premier quartile, pour exclure de manière plus certaine les faux positifs.

Cette méthode de décision nécessite de disposer d'un jeu de test avec des données réelles afin d'évaluer le seuil optimal. On peut transposer cette étude avec un jeu de test constitué de données de simulation, mais cela rend les estimations des seuils moins robustes. Pour les sources pour lesquelles on ne dispose pas de données réelles, j'utilise par défaut le seuil de 50 %.

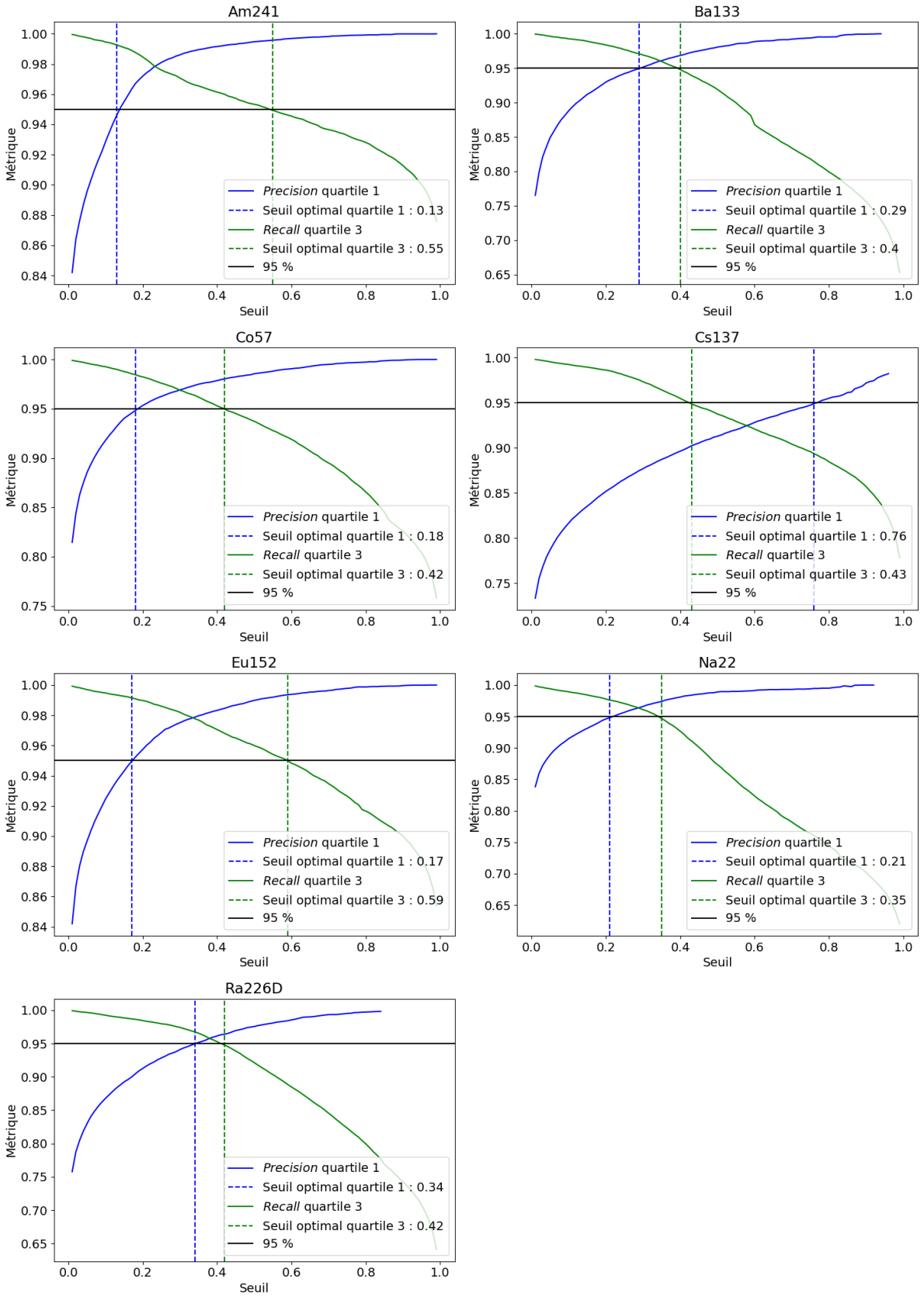


Figure 4-9 : Détermination des seuils optimaux pour les premiers et troisièmes quartiles.

4.3.3 Sensibilité pour les sources individuelles

Les tests précédents ont été faits sur des mélanges aléatoires de radioéléments, à différentes proportions et avec modification de la calibration. Dans ce paragraphe, j'évalue la sensibilité des réseaux de neurones pour des sources individuelles.

Pour ce faire, je constitue des spectres de radioéléments individuels, avec différentes statistiques de comptage de photons, afin d'évaluer le nombre minimum de photons nécessaires dans le spectre pour détecter la source.

Pour le seuil de détection, j'utilise les seuils optimaux sur le premier quartile, évalués au paragraphe 4.3.2, mais je compare aussi avec un seuil plus restrictif, à 90 % sur le premier quartile, qui confirme avec une très grande certitude la présence de la source. Les résultats sont représentés en Figure 4-10**.**

On voit que pour certaines sources comme le ^{241}Am, le ^{57}Co ou le ^{152}Eu, une dizaine de photons est nécessaire pour pouvoir détecter la source avec le seuil optimal. Avec le seuil à 90 %, quelques dizaines de photons sont nécessaires pour les identifier avec une grande confiance. Cela provient notamment du fait que les pics photoélectriques principaux de ces radioéléments sont situés dans le domaine d'énergie où Caliste est très efficace, inférieur à 140 keV.

Pour les autres radioéléments, une centaine à plusieurs centaines de photons sont nécessaires pour les identifier. Pour le ^{133}Ba, le ^{57}Co, le ^{22}Na et le ^{226}Ra-D, le seuil à 90 % sur le premier quartile est trop conservatif. Par conséquent, il n'est pas possible d'atteindre 100 % d'identifications correctes sur ces exemples.

Je résume dans le Tableau 4-1 le nombre de photons nécessaires pour identifier les sources dans 100 % des cas que j'ai testés. Je donne la sensibilité correspondante en débit de dose détectable en 5 secondes, sauf pour le ^{226}Ra-D, pour laquelle je ne connais pas l'activité de la source.

Radioélément	Nombre d'événements simples pour 100 % d'identifications, seuil optimal	Sensibilité correspondante : débit de dose en nSv/h détectable en 5 s	Nombre d'événements simples pour 100 % d'identifications, seuil à 90 %	Sensibilité correspondante : débit de dose en nSv/h détectable en 5 s
^{241}Am	10	11	30	35
^{133}Ba	100	130	/	/
^{57}Co	10	8	/	/
^{137}Cs	60	1000	80	1400
^{152}Eu	20	85	40	170
^{22}Na*	600	6300	/	/
^{226}Ra-D	60	/	/	/

Tableau 4-3 : Sensibilités des réseaux d'identification pour des sources individuelles.

* Mesure et sensibilité données avec Caliste-O

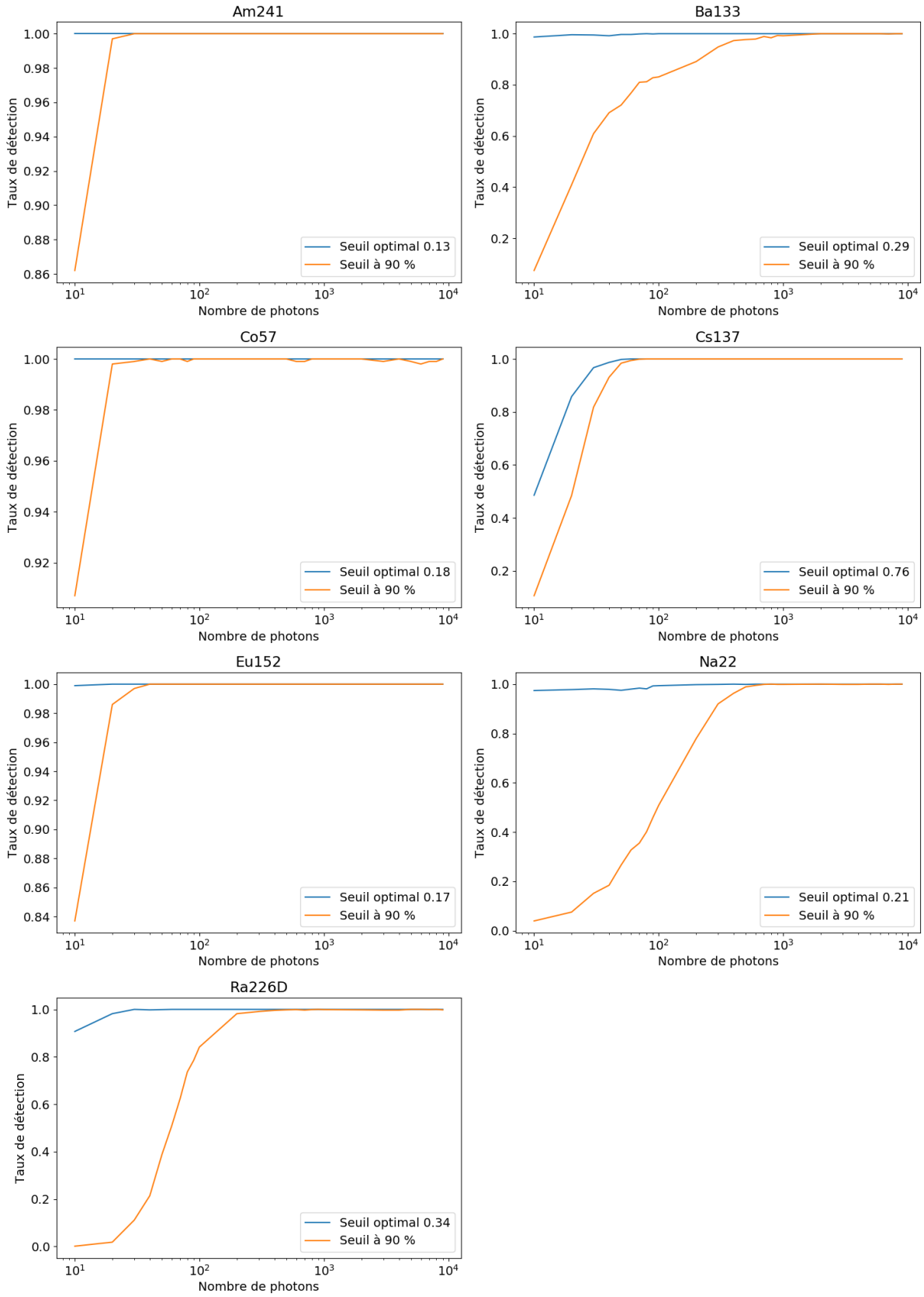


Figure 4-10 : Nombre de photons nécessaire pour détecter chaque source.

La sensibilité des réseaux de neurones permet d'identifier les sources individuelles avec très peu de photons. Cependant, cette étude est faite avec des spectres très purs, avec peu de bruit de fond étant

donné que les sources étaient très proches du détecteur dans les mesures, à moins de 10 cm. Le bruit de fond ajoute une composante qui peut masquer les sources. Cependant, si la source n'est plus détectable dans le bruit de fond ambiant, cela signifie qu'elle est trop faible et qu'on peut généralement la considérer comme absente dans le signal.

4.3.4 Performances sur la quantification

Afin d'évaluer les performances sur la quantification, dans la base de données que j'ai constituée, je connais les proportions de chaque radioélément dans les mélanges artificiels que j'ai créés à partir de données réelles à travers le processus d'augmentation de données décrit en paragraphe 4.1.

Pour chacun de ces mélanges, j'évalue l'écart entre la prédiction par les réseaux de neurones de quantification et la proportion réelle dans le mélange. Pour ce faire, pour chaque radioélément, je ne considère que les mélanges contenant ce radioélément et au minimum une autre source. En effet, si le radioélément n'est pas présent, le réseau de neurones de quantification n'est pas utilisé et de même, si ce radioélément est seul, la proportion de ce radioélément dans le signal est de 100 % par comparaison aux autres radioéléments.

Dans un premier temps, en Figure 4-11, j'évalue l'écart entre la médiane des prédictions et la valeur réelle. L'erreur est algébrique, je regarde la différence sans valeur absolue.

On voit qu'en moyenne, la médiane sous-estime la proportion réelle de -3,5 % pour le ^{241}Am et le ^{57}Co. Ce sont les radioéléments pour lesquels les lignes principales se situent dans le domaine où le détecteur est le plus efficace, et donc pour lesquels l'évaluation des proportions se fait principalement sur la base des pics photoélectriques. Inversement, pour le ^{22}Na et le ^{226}Ra-D, l'écart est plus élevé, en moyenne de -11,4 %, du fait que pour ces radioéléments, le spectre est beaucoup plus étalé. D'autre part, les sources de ^{22}Na et ^{226}Ra-D sont beaucoup moins intenses que les autres, comme décrit en Tableau 4-1, et une part non négligeable du signal est due au bruit de fond. De ce fait, une partie du signal devrait en fait être affectée au bruit de fond et non à la source, ce que je n'ai pas fait dans la création de mes mélanges artificiels. La sous-estimation des proportions apparaît dans tous les cas comme un effet systématique qui pourrait être redressé a posteriori.

J'évalue maintenant les performances en prenant en compte les incertitudes autoévaluées par le réseau. Pour ce faire, je regarde les incertitudes en prenant en compte d'une part le premier et le troisième quartile en Figure 4-12 et d'autre part, les quantiles à 5 % et 95 %, pour une incertitude plus large, en Figure 4-13. Dans ces figures, je regarde la différence entre le quantile le plus proche de la valeur et la valeur réelle. Si la valeur réelle se situe entre les deux quantiles, je mets une différence de 0, ce qui signifie que la valeur réelle se situe dans la barre d'incertitude.

Si l'incertitude autoévaluée par le réseau est assimilable à une incertitude statistique, cela signifierait que, dans 50 % des cas de test, la valeur réelle des proportions se situerait entre le premier et le troisième quartile des prédictions renvoyées par le réseau. De même, 90 % des valeurs réels des tests devraient se situer entre les quantiles à 5 % et à 95 % des prédictions du réseau.

En Figure 4-12, on voit que l'incertitude utilisant le premier et le troisième quartile permet d'estimer correctement la valeur réelle dans plus de 50 % des exemples de test pour le ^{241}Am, le ^{57}Co et le ^{152}Eu. Cependant, on constate pour les autres sources que les proportions sous toujours sous-estimées. On ne retrouve pas les 50 % de prédictions correctes avec les incertitudes : l'incertitude renvoyée par le réseau est sous-estimée.

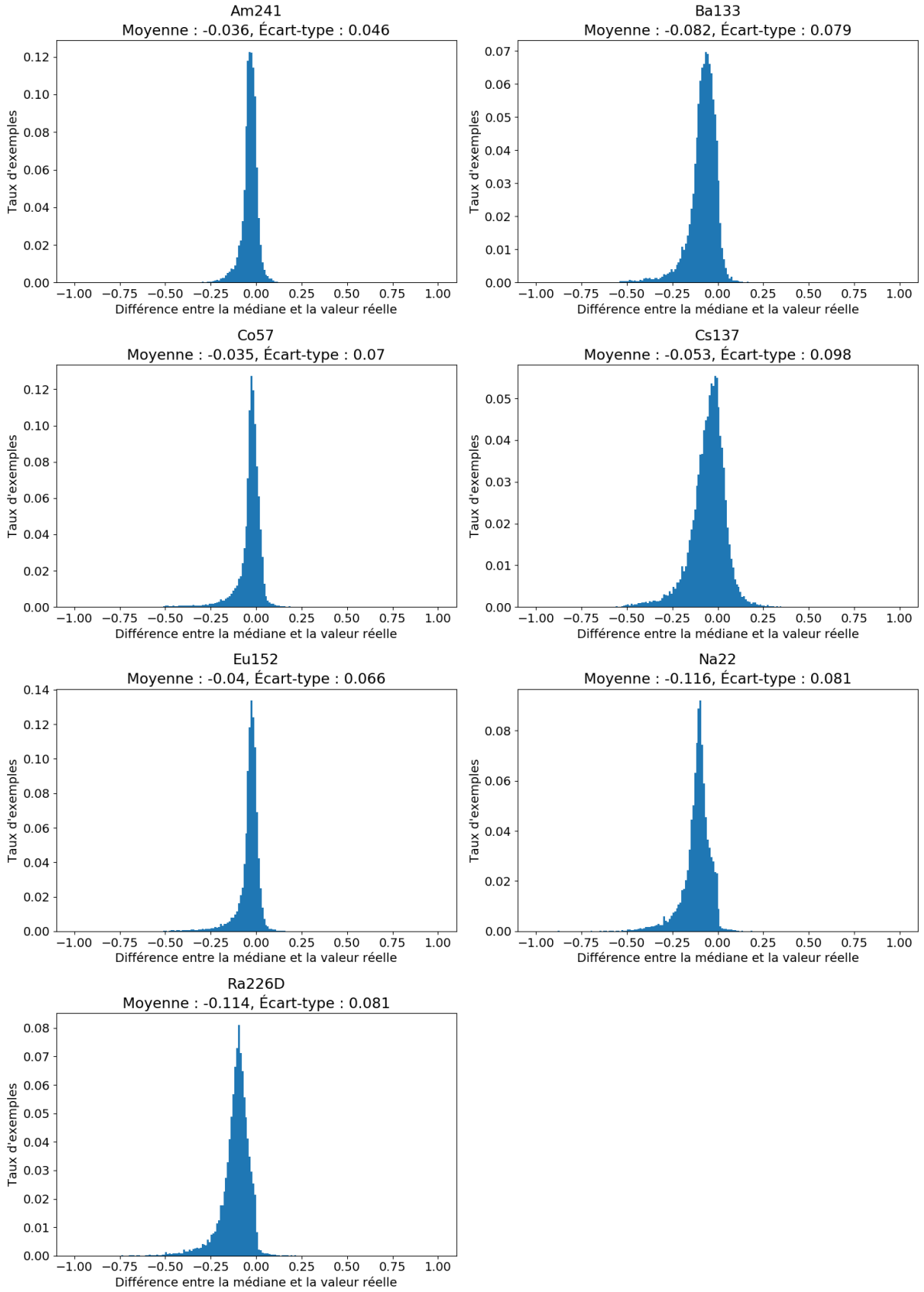


Figure 4-11 : Écart de prédiction entre la médiane et la valeur réelle.

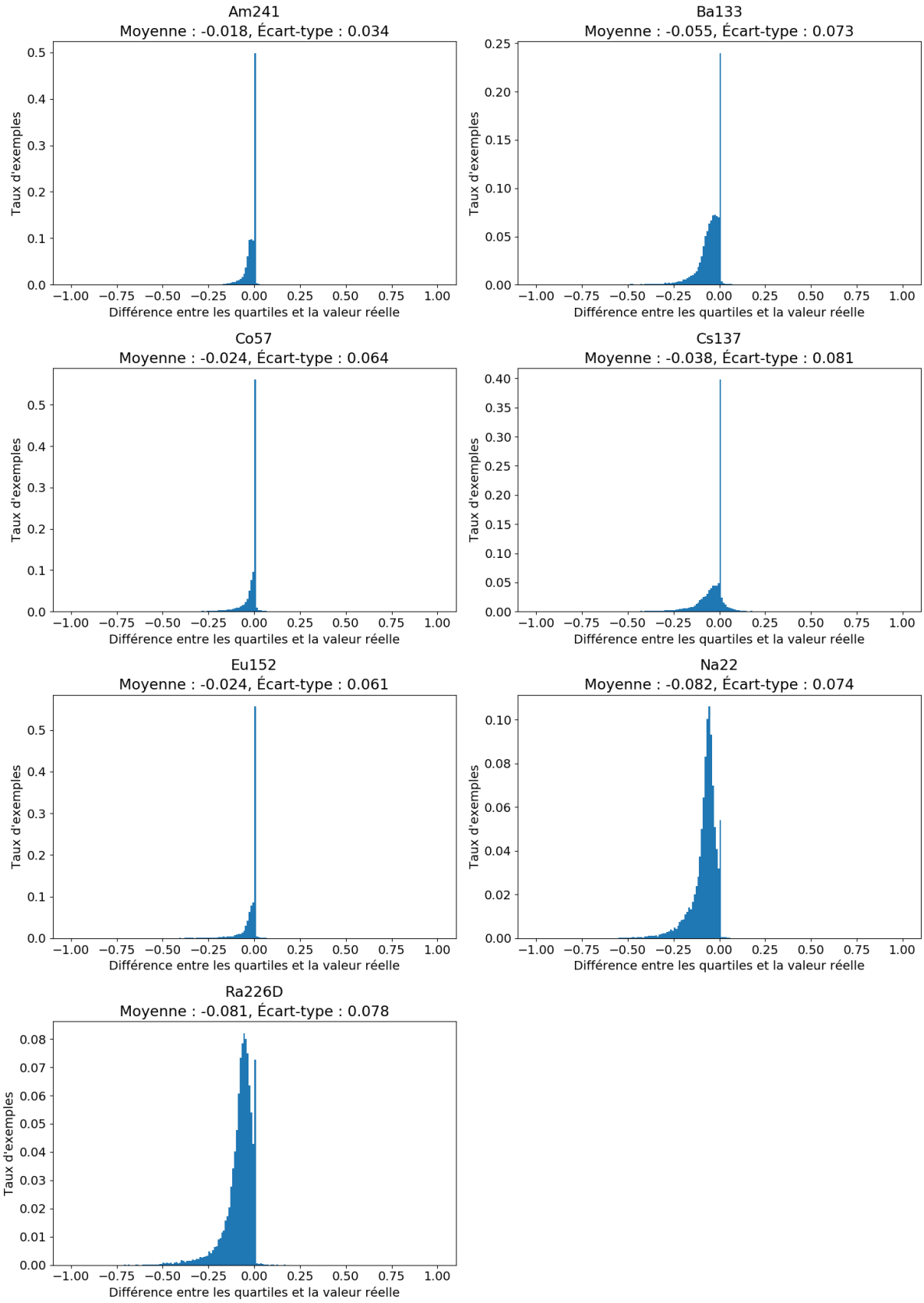


Figure 4-12 : Écart de prédiction entre les quartiles (25 % et 75 %) et la valeur réelle.

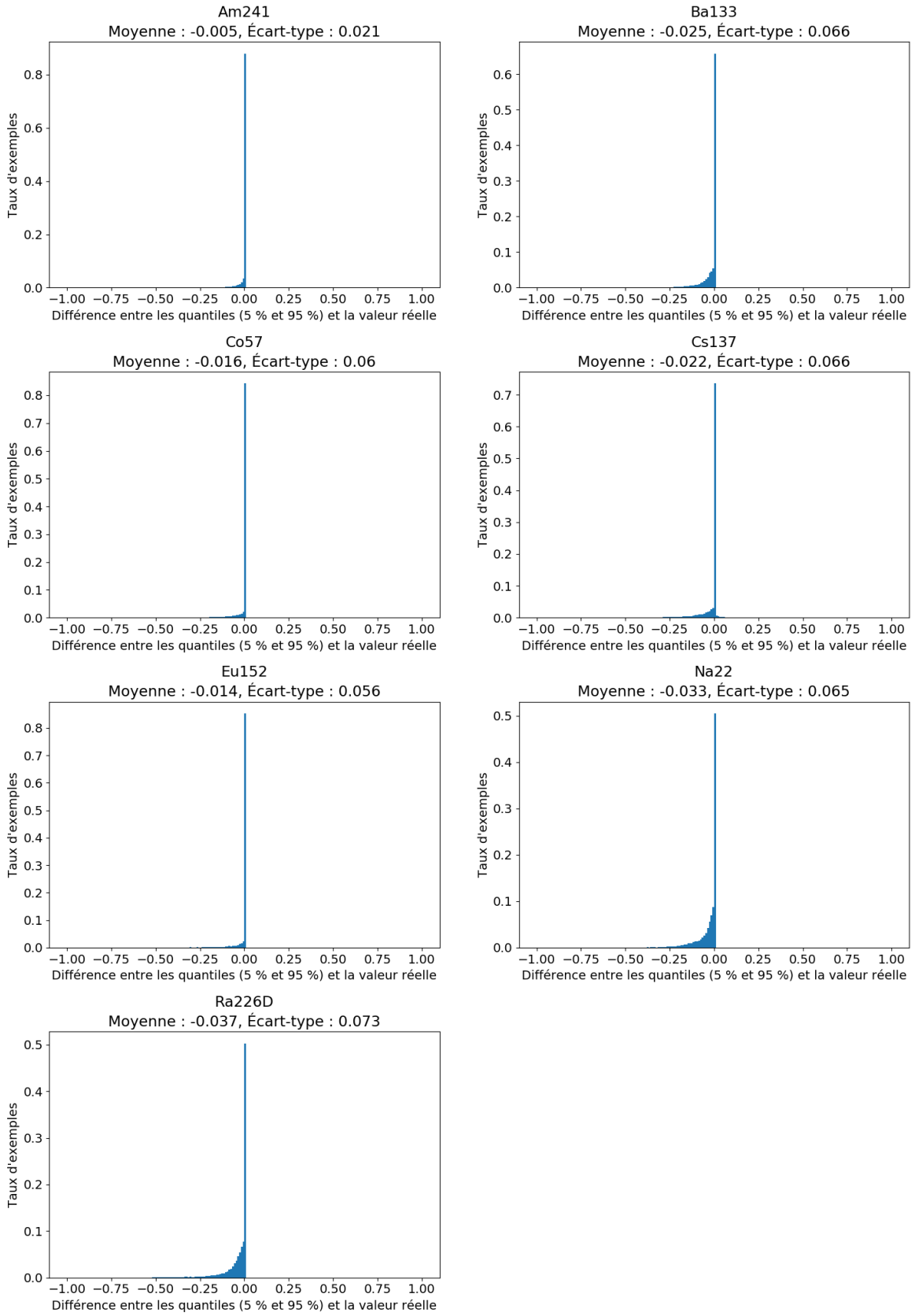


Figure 4-13 : Écart de prédiction entre les quantiles (5 % et 95 %) et la valeur réelle.

L'utilisation des quantiles à 5 % et 95 % permet de retrouver la valeur réelle des proportions dans un plus grand nombre de cas, comme on le voit en Figure 4-13. Avec ces quantiles, pour l'ensemble des radioéléments, j'encadre la valeur réelle dans plus de 50 % des cas pour l'ensemble des radioéléments.

De nouveau, en utilisant ces deux quantiles, l'incertitude telle que je l'évalue à travers les réseaux de neurones bayésiens est sous-estimée par rapport à l'incertitude à laquelle on s'attend dans les tests. Une étude plus avancée des incertitudes autoévaluées du réseau serait nécessaire pour caractériser plus en détail les performances du réseau. Il serait intéressant par exemple de comparer le nombre d'exemples pour lesquels la valeur attendue des proportions se situe effectivement entre deux quantiles des prédictions du réseau, en fonction de l'écart interquantile utilisé. Un redressement a posteriori pourra alors être envisagé afin de faire correspondre l'incertitude autoévaluée par le réseau à une incertitude statistique correcte.

5 Application à la dosimétrie avec Caliste

5.1 Débit de dose : définition et calcul

Dans le domaine de la radioprotection, une grandeur indispensable à mesurer est le débit de dose d'un environnement radiatif, qui sert à évaluer la dangerosité des radiations dans l'environnement pour un organisme biologique, en particulier pour les humains.

On définit la dose absorbée, qui s'exprime en Gray (Gy), équivalent au J/kg, qui correspond à l'énergie déposée par des rayonnements ionisants par unité de masse. Soit un élément de masse dm et dans lequel est déposée une énergie dE, la dose D déposée dans cet élément de masse est donnée par :

$$D = \frac{dE}{dm}$$

Ensuite, on définit la dose équivalente, qui s'exprime en Sievert (Sv), équivalent au Gray, qui correspond à la dose absorbée corrigée d'un facteur de pondération lié au type du rayonnement et à sa dangerosité pour des tissus biologiques. Ce facteur de pondération est défini par la Commission Internationale de Protection Radiologique (Nénot, et al., 2007). Il vaut par exemple 1 pour les photons X ou gamma, ce qui signifie que la dose équivalente et la dose absorbée sont égales, mais il vaut par exemple 20 pour les particules alpha du fait de leur dangerosité.

Enfin, on définit la dose efficace, exprimée aussi en Sievert, qui correspond à la dose équivalente corrigée d'un facteur de pondération selon le type de tissu ou organe touché, l'impact biologique d'un rayonnement étant plus ou moins dangereux selon l'organe considéré. Il est par exemple de 0,01 pour la peau et de 0,12 pour les poumons.

En radioprotection, c'est la dose équivalente en Sv qui est le principal indicateur utilisé pour caractériser un environnement, et plus précisément le débit de dose exprimé en Sv/h. Cela permet notamment de définir des types de zones suivant leur débit de dose. Par exemple, une zone bleue est une zone surveillée car soumise au rayonnement mais dont le débit de dose ne dépasse pas 7,5 µSv/h. Cela correspond à une dose de 1 mSv par an, soit la limite réglementaire d'exposition non naturelle pour le personnel civil non habilité à travailler en environnement radiatif. À l'inverse, une zone rouge

est une zone interdite, dépassant 100 mSv/h, débit de dose à partir duquel la vitesse des dégâts au niveau de l'ADN sont supérieurs à son taux de réparation, et donc très dangereux pour les organismes biologiques.

La dose équivalente découle de la dose absorbée, qui dépend donc du matériau utilisé, du fait que l'énergie déposée par un rayonnement dépend de la probabilité d'interaction de ce rayonnement dans le matériau. Il dépend aussi du type d'interaction, absorption photoélectrique ou diffusion Compton, qui conditionne l'énergie effectivement déposée dans le matériau lors de l'interaction. Le matériau utilisé généralement pour caractériser la dose est l'air sec à 20°C (Chabot, 2016).

On peut donner la relation entre l'activité d'une source, ses lignes d'émission, la distance au point de mesure et le débit de dose à ce point de mesure. On considère une source d'activité A, émettant des photons à différentes énergies E_i, où i est la ligne d'émission, d'intensité I_i, située à une distance d du point de mesure.

Le flux en énergie reçu (en keV/s/m²) correspondant à la ligne i, à cette distance d vaut :

$$\varphi_i = \frac{A \times I_i \times E_i}{4\pi d^2} \tag{5-1}$$

Connaissant le flux, on peut calculer le débit de dose $\dot{D}_i$ en utilisant le coefficient d'absorption masse-énergie $\left(\frac{\mu_{\text{en}}}{\rho}\right)_{\text{air}}$ (*mass-energy absorption coefficient*). Ce coefficient prend en compte les probabilités d'interaction d'un photon d'énergie E_i ainsi que l'énergie qui sera finalement déposée par le photon, soit par absorption photoélectrique, soit en retirant l'énergie des photons secondaires issus de la diffusion Compton, ou éventuellement lors du processus de créations de paires (Hubbell & Seltzer, 2004) :

$$\dot{D}_i = \left(\frac{\mu_{\text{en}}(E_i)}{\rho}\right)_{\text{air}} \times \varphi_i \tag{5-2}$$

Pour l'ensemble des lignes d'émission de la source, le débit de dose total associé vaut alors :

$$\dot{D} = \sum_i \left(\frac{\mu_{\text{en}}(E_i)}{\rho}\right)_{\text{air}} \times \frac{A \times I_i \times E_i}{4\pi d^2} \tag{5-3}$$

5.2 Mesure du débit de dose avec Caliste

5.2.1 Principe de la mesure

Maintenant, je souhaite utiliser les mesures réalisées avec Caliste pour déterminer le débit de dose. Le module de détection de Caliste est constitué d'un cristal de CdTe et non d'air. On ne peut donc pas utiliser directement la mesure faite avec le détecteur pour déterminer le débit de dose tel que défini précédemment.

La mesure de Caliste détermine le flux $\varphi_{\text{Caliste}}(E)$, que je définis comme le flux en keV/s/cm²* d'énergie déposée dans le cristal lors de la mesure. L'étape essentielle est d'être capable de déterminer le flux $\varphi_i(E)$ incident, c'est-à-dire avant interaction dans le cristal.

Retrouver le flux incident à partir de la mesure est un problème qui nécessite la prise en compte de plusieurs facteurs :

- l'efficacité du détecteur, c'est-à-dire, la probabilité qu'un photon à une énergie donnée interagisse dans le détecteur ;
- le dépôt partiel d'énergie dans le cas des interactions Compton dans le détecteur, qui nécessite de retrouver l'énergie du photon d'origine. Cette énergie peut ne pas être une des lignes d'un radioélément si le photon avait initialement été diffusé dans l'environnement ;
- les échappements dus aux atomes Cd et Te ;
- la résolution du détecteur et en particulier les phénomènes liés à l'efficacité d'induction de charges ;
- les énergies déposées par processus de création de paires.

Dans cette partie, je m'intéresse principalement à l'efficacité du détecteur et aux énergies partielles déposées par interaction Compton. La résolution du détecteur modifie assez peu l'énergie, compte tenu des performances du détecteur Caliste décrites en Chapitre 1†. Les échappements de Cd et Te ne concernent qu'une faible proportion des événements, typiquement de moins de 5 %. Par ailleurs, aux énergies que j'étudie, inférieures à 1 MeV, il n'y a pas de phénomène de création de paires.

5.2.2 Inversion Compton

Soit un spectre donné en flux par unité d'énergie et de surface (en $s^{-1}.cm^{-2}.keV^{-1}$) mesuré avec Caliste. L'objectif ici est de retrouver le spectre en énergie avant les diffusions Compton.

Une méthode d'inversion directe consisterait à calculer, pour chaque canal d'énergie E du spectre, la matrice contenant la distribution de Klein-Nishina, corrigée de l'élargissement Doppler comme je l'ai présenté dans le Chapitre 2 ainsi que le pic photoélectrique, dans les proportions correspondant aux probabilités d'interaction Compton et d'absorption photoélectrique à l'énergie E donnée. Puis, on inverse cette matrice et on l'applique au spectre mesuré. Le problème de cette méthode est sa grande sensibilité aux petites fluctuations dans la mesure, dues à la statistique de Poisson sur le comptage des photons.

Je propose une autre méthode, basée sur une utilisation de la formule de Bayes et d'une variante de l'algorithme présenté par (Bobin, et al., 2016). Cet algorithme a été utilisé initialement pour répondre à un autre problème qui est celui de la quantification des proportions de radioéléments dans un mélange, problème pour lequel j'ai proposé une autre méthode à travers les réseaux de neurones dans ce chapitre. Cependant, je montre ici que l'algorithme proposé par (Bobin, et al., 2016) peut être adapté au problème du redressement des interactions Compton dans le spectre.

* On utilise le flux par en keV/s/cm² et non en keV/s/m² (unité internationale) du fait que la surface de détection de Caliste-HD est de 1 cm².

† Typiquement, la largeur totale à mi-hauteur des pics est inférieure à 5 % au-delà de 20 keV.

Je note $p^0(E_{\text{inc}})$, la probabilité *prior* qu'un photon détecté avait une énergie E_{inc} avant interaction. J'utilise ici le spectre mesuré pour représenter ce *prior*, mais on peut utiliser un autre *prior*, par exemple uniforme sur toute la bande en énergie. L'avantage d'utiliser le spectre mesuré comme *prior* est que les canaux n'ayant pas détecté d'énergie resteront vides, ce qui évite une surinterprétation où l'on remplirait des canaux d'énergie qui devraient rester vides. Cela présuppose qu'on a détecté au moins un photon ayant déposé toute son énergie par absorption photoélectrique.

L'idée de l'algorithme est de mettre à jour le *prior* en fonction des énergies effectivement déposées E_{dep}, et la probabilité a posteriori qu'un photon d'énergie E_{inc} soit à l'origine de E_{dep}.

Pour ce faire, on attribue un compteur $k(E_{\text{inc}})$ à chaque énergie E_{inc} du spectre incident que l'on cherche à estimer.

À l'étape n de l'algorithme, on tire dans le spectre mesuré un photon E_{dep}. Pour chaque énergie E_{inc}, on calcule la probabilité $p^{n+1}(E_{\text{inc}}|E_{\text{dep}})$, c'est-à-dire la probabilité que le photon ait pour énergie initiale E_{inc} sachant qu'on a mesuré E_{dep}. Le théorème de Bayes donne :

$$p^{n+1}\left(E_{\text{inc}}\middle|E_{\text{dep}}\right) = \frac{p\left(E_{\text{dep}}\middle|E_{\text{inc}}\right) \times p^n(E_{\text{inc}})}{p\left(E_{\text{dep}}\right)} \tag{5-4}$$

Dans cette équation, $p^n(E_{\text{inc}})$ est le *prior*, qui vaut initialement $p^0(E_{\text{inc}})$ et qui sera mis à jour au fur et à mesure de l'algorithme. Il correspond à la probabilité que le photon incident ait une énergie E_{inc}.

$p\left(E_{\text{dep}}\middle|E_{\text{inc}}\right)$ correspond à la probabilité de mesurer l'énergie E_{dep} pour un photon ayant interagi avec une énergie incidente E_{inc}. Il existe deux possibilités :

- si $E_{\text{dep}} = E_{\text{inc}}$, l'interaction est une absorption photoélectrique, et la probabilité que l'interaction de E_{inc} conduise à une mesure de E_{dep} est directement la probabilité que l'interaction soit une absorption photoélectrique, ce qui est donné par $\frac{\mu_{\text{PE}}(E_{\text{inc}})}{\mu(E_{\text{inc}})}$, où $\mu_{\text{PE}}(E_{\text{inc}})$ est le coefficient d'absorption photoélectrique pour le CdTe et $\mu(E_{\text{inc}})$ est le coefficient d'absorption totale ;
- sinon, E_{dep} peut provenir d'une interaction Compton issue de E_{inc}, et dans ce cas, il faut prendre en compte la probabilité $\frac{\mu_{\text{Compton}}(E_{\text{inc}})}{\mu(E_{\text{inc}})}$ que l'interaction soit une diffusion Compton d'une part, et que la probabilité $p_{E_{\text{inc}}}^{\text{Compton}}\left(E_{\text{dep}}\right)$ l'interaction conduise à un dépôt d'énergie E_{dep}, cette probabilité a été calculée au Chapitre 2.

Je résume ainsi ces deux possibilités dans la formule suivante :

$$p\left(E_{\text{dep}}\middle|E_{\text{inc}}\right) = \frac{1}{\mu(E_{\text{inc}})}\left(\mu_{\text{PE}}(E_{\text{inc}})\delta_{E_{\text{inc}}}\left(E_{\text{dep}}\right) + \mu_{\text{Compton}}(E_{\text{inc}}) \times p_{E_{\text{inc}}}^{\text{Compton}}\left(E_{\text{dep}}\right)\right) \tag{5-5}$$

Dans l'équation (5-4), $p\left(E_{\text{dep}}\right)$ est un facteur de normalisation, il n'est pas nécessaire de le calculer. Il suffit de calculer les quantités $p\left(E_{\text{dep}}\middle|E_{\text{inc}}\right) \times p^n(E_{\text{inc}})$ pour chaque énergie E_{inc}, puis d'appliquer le facteur $\frac{1}{\sum_{E_{\text{inc}}} p\left(E_{\text{dep}}\middle|E_{\text{inc}}\right) \times p^n(E_{\text{inc}})}$ à chacune de ces quantités pour faire cette normalisation.

On met à jour le compteur avec la probabilité calculée par l'équation (5-4) :

$$k(E_{\mathrm{inc}}) \coloneqq k(E_{\mathrm{inc}}) + p^{n+1}\left(E_{\mathrm{inc}} \middle| E_{\mathrm{dep}}\right) \tag{5-6}$$

Et enfin, on met à jour le *prior* à l'aide du compteur :

$$p^{n+1}(E_{\mathrm{inc}}) = \frac{k(E_{\mathrm{inc}})}{\sum_E k(E)} \tag{5-7}$$

Le processus est ensuite réitéré à l'étape suivante.

Il est conseillé de ne faire la mise à jour (5-7) du *prior* qu'après un certain nombre d'étapes, de sorte à éviter que les premiers événements utilisés aient un poids trop important dans le calcul. J'ai choisi de mettre à jour le *prior* au bout de 50 000 itérations.

En répétant ce processus sur un grand nombre d'étapes, j'aboutis à l'estimation du spectre incident des photons ayant interagi dans le détecteur. J'applique cette procédure sur 1 million d'itérations. Avec un nombre d'étapes plus élevé, le calcul peut devenir très long, de l'ordre de plusieurs dizaines de minutes ce qui n'est pas acceptable pour obtenir une information dosimétrique en temps réel, et inversement, avec un nombre d'étapes plus faible, l'estimation sera moins fiable. Avec 1 million d'itérations, le calcul est de l'ordre d'une minute sur un ordinateur de bureau classique.

Une fois la distribution $p(E_{\mathrm{inc}})$ retrouvée, il suffit de la multiplier par le flux total du spectre mesuré pour retrouver le spectre incident $S_{\mathrm{inc}}(E)$ en flux par unité d'énergie et de surface, donné en $s^{-1}cm^{-2}keV^{-1}$.

Je montre en Figure 5-1 un exemple de cette procédure appliquée à un spectre de ^{137}Cs, dont les composantes sont facilement identifiables, avec deux pics principaux à basse énergie et seulement un pic à haute énergie à 662 keV. Contrairement aux spectres que j'ai présentés précédemment, j'utilise ici le spectre avec tous les événements, y compris les événements multiples, de sorte à ne pas avoir de perte au niveau de l'estimation de la dose dans le cas où je ne sélectionnerais que les événements simples. De ce fait, la résolution de ce spectre est moins bonne que sur les spectres en événements simples.

On voit dans cette procédure d'inversion que le pic photoélectrique à 662 keV est plus intense, d'un facteur 4 environ, contrairement à la zone entre 200 keV et 478 keV qui perd en intensité, d'un facteur 1,5 au maximum. Cela montre que des photons de ce domaine d'énergie, qui correspond aux dépôts Compton dans le détecteur, sont réattribués au pic photoélectrique, ce qui est justement le but de cette procédure d'inversion Compton.

En revanche, dans la zone des basses énergies, inférieures à 200 keV, l'intensité du spectre n'est pas impactée par la procédure, du fait que dans ce domaine d'énergie, le principal mode d'interaction est l'absorption photoélectrique.

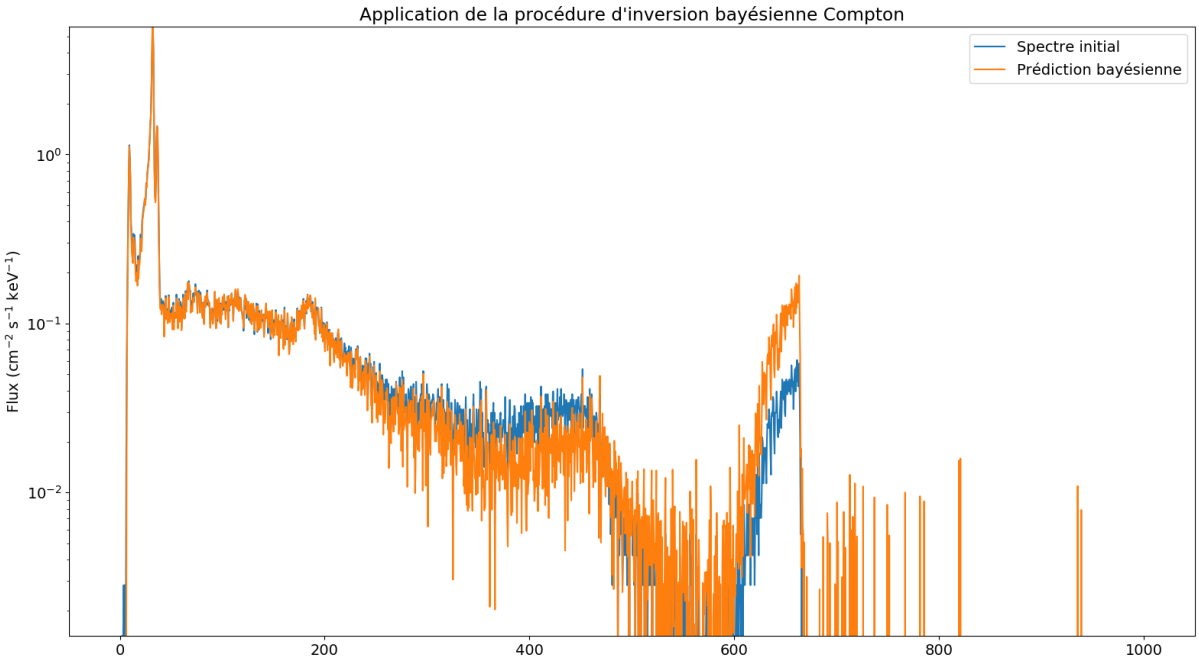


Figure 5-1 : Spectre initial et spectre après inversion bayésienne Compton.

5.2.3 Prise en compte de l'efficacité de détection

À la suite de l'étape précédente, je retrouve le spectre des photons qui ont interagi, à la résolution du détecteur près. Maintenant, je vais retrouver le spectre incident au niveau de la fenêtre d'entrée du détecteur, c'est-à-dire, en prenant en compte le fait que les photons qui ont traversé le détecteur n'ont pas tous interagi dans le cristal.

Pour cela, il suffit de redresser de l'efficacité de détection. En notant toujours $\mu(E)$ le coefficient d'absorption linéique du CdTe pour l'énergie E et L l'épaisseur du détecteur, l'intensité du spectre en entrée du détecteur est donnée par :

$$S_{\text{entree}}(E) = \frac{S_{\text{inc}}(E)}{1 - e^{-\mu(E)\times L}} \tag{5-8}$$

On peut objecter que les photons peuvent aussi provenir des côtés du détecteur et pas seulement de la fenêtre d'entrée, ce qui modifierait la longueur L traversée. Cependant, il n'est pas possible de prendre en compte l'ensemble des géométries existantes et des provenances des photons autrement que par des simulations numériques lourdes. Cette méthode est ici assez générale et on considère en moyenne que les photons arrivent directement sur la face avant du détecteur.

En Figure 5-2, j'applique la formule (5-8) au spectre calculé dans la Figure 5-1.

La correction en efficacité commence à avoir un impact significatif à partir de 60 keV, où l'efficacité de détection passe sous les 97 %. Elle est ensuite notable pour les plus hautes énergies. En particulier, le pic à 662 keV a une intensité bien plus importante avec la correction, qui est d'un facteur 24 environ, la probabilité d'interaction à cette énergie étant de 4,3 %.

Le spectre apparaît plus bruité à haute énergie qu'à basse énergie. Cela provient du fait que le redressement en intensité que j'effectue propage le bruit de Poisson dû à la statistique de comptage.

La statistique de comptage étant plus faible à haute énergie, elle contient plus de bruit, ce qui se répercute après correction de l'intensité.

Le spectre ainsi obtenu correspond au spectre estimé en entrée du détecteur, avant toute interaction dans le détecteur.

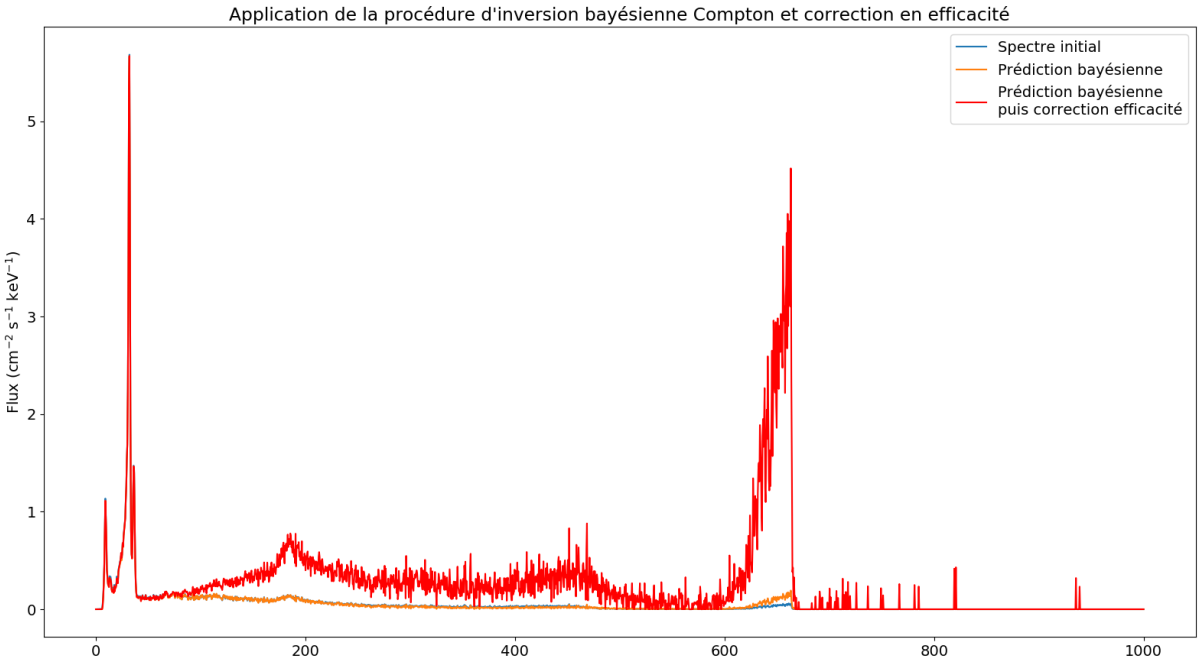


Figure 5-2 : Application de l'inversion bayésienne Compton et de la correction en efficacité.

5.2.4 Application au calcul de la dose

À la fin de l'étape précédente, j'obtiens donc une estimation du spectre des photons en entrée du détecteur, ce qui correspond donc au flux φ en formule (5-2). J'applique alors l'équation (5-2) à l'ensemble des énergies du spectre pour obtenir une mesure de la dose.

J'ai réalisé différentes mesures avec des sources individuelles, placées à 20 cm en face du détecteur. Le Tableau 5-1 suivant résume les caractéristiques des sources et les valeurs trouvées. La mesure utilisée pour les figures précédentes sont celles de la source de ^{137}Cs. J'effectue une estimation de l'incertitude à partir des fluctuations attendues sur le nombre de photons dans le spectre. Pour ce faire, si N_i représente le nombre de photons dans le *bin* i du spectre, j'effectue l'estimation basse du débit de dose avec le spectre contenant $N_i - \sqrt{N_i}$ photons pour chaque *bin* i et l'estimation haute avec le spectre contenant $N_i + \sqrt{N_i}$ photons.

J'ai aussi effectué le calcul théorique en utilisant la méthode décrite au paragraphe 5.1, étant donné que je connais les sources, leurs activités et la distance au détecteur. Quant aux incertitudes, l'erreur que je commets sur la mesure des distances est de l'ordre du cm, causée principalement sur l'incertitude sur la position du détecteur dans la tête de détection, et la position de la source dans son *packaging*. L'incertitude sur l'activité est estimée à 5 % de l'activité, qui est l'incertitude typique donnée au moment de la fabrication de la source.

Afin de confirmer le débit de dose, j'ai effectué une mesure avec un radiamètre Radiagem-2000 (MIRION Technologies, 2017), l'incertitude présentée correspond aux fluctuations du débit de dose observées lors de la mesure.

Conditions des mesures			Débit de dose théorique (µSv/h)	Mesure Radiamètre (µSv/h)	Estimation du débit de dose à partir des mesures (µSv/h)
Source	Activité (kBq)	Distance (cm)			
^{241}Am	436 ± 22	20 ± 1	$(2{,}3 \pm 0{,}4).10^{-1}$	$(6{,}5 \pm 3{,}5).10^{-2*}$	$(2{,}1 \pm 0{,}2).10^{-1}$
^{133}Ba	134 ± 7	20 ± 1	$(2{,}4 \pm 0{,}4).10^{-1}$	$(2{,}5 \pm 0{,}3).10^{-1}$	$(3{,}0 \pm 0{,}5).10^{-1}$
^{57}Co	53 ± 3	20 ± 1	$(2{,}4 \pm 0{,}4).10^{-2}$	$(2{,}1 \pm 0{,}8).10^{-2}$	$(5{,}4 \pm 2{,}1).10^{-2}$
^{137}Cs	1660 ± 83	20 ± 1	$3{,}2 \pm 0{,}5$	$3{,}2 \pm 0{,}6$	$1{,}9 \pm 0{,}4$
^{152}Eu	68 ± 4	20 ± 1	$(2{,}6 \pm 0{,}4).10^{-1}$	$(3{,}5 \pm 1{,}0).10^{-1}$	$(2{,}7 \pm 1{,}8).10^{-1}$

Tableau 5-1 : Comparaison entre les débits de dose calculés et estimés avec la méthode décrite précédemment à partir des mesures.

Dans tous les cas, l'ordre de grandeur de la dose est effectivement retrouvé. Pour le ^{241}Am, le ^{152}Eu et le ^{133}Ba, l'estimation du débit de dose est compatible avec le débit de dose calculé théoriquement et avec celui mesuré au radiamètre (sauf pour le ^{241}Am*). Cela provient du fait que les lignes principales de ces radioéléments se situent à basse énergie, donc dans des domaines où moins de corrections sont effectuées du fait que la probabilité d'interaction est proche de 100 % et que la très grande majorité des interactions sont des interactions photoélectriques.

Par ailleurs, on voit que ma méthode sous-estime le débit de dose pour le ^{137}Cs. On peut expliquer en partie cette sous-estimation par le fait qu'en utilisant les événements multiples, il existe des phénomènes de perte de charge entre les pixels du détecteur, qui impliquent que les énergies mesurées sont plus petites que l'énergie effectivement déposée (Dubos, 2015). Pour répondre à ce problème, on peut envisager d'appliquer cette méthode avec les événements simples uniquement, que l'on redresse in fine de la fraction d'événements simples attendue dans le cristal en fonction de l'énergie des photons incidents.

Inversement, ma méthode surestime le débit de dose d'un facteur 2 pour le ^{57}Co. Les lignes principales du ^{57}Co se situent à 121 keV et 136 keV, et les structures Compton se situent à basse énergie, entre 0 et 50 keV. Ces énergies sont celles qui contribuent majoritairement à la dose, du fait de leur forte probabilité d'interaction, proche de 100 %. Ainsi, si l'inversion Compton sous-estime la réattribution de ces énergies aux pics photoélectriques, on peut observer ce phénomène.

L'algorithme que je propose présente des limites. Il donne un ordre de grandeur du débit de dose mais il ne permet pas d'évaluer suffisamment précisément le débit de dose.

5.3 « Démélange » de sources et dose par radioéléments individuels

Dans le paragraphe 5.2, j'ai montré comment évaluer la dose totale au niveau du détecteur à partir d'un spectre. Une donnée importante du point de vue de la radioprotection est aussi d'être capable

* La gamme d'énergie couverte par le radiamètre va de 30 keV à 2 MeV. Or les lignes d'émission principales du ^{241}Am se concentrent à moins de 27 keV (hormis la ligne la plus intense à 60 keV), et de ce fait le radiamètre n'est pas capable de détecter les photons pour ces lignes d'émission, d'où la sous-estimation observée par rapport au calcul théorique.

d'estimer la dose de chaque radioélément individuel, dans une scène comportant plusieurs radioéléments différents.

Je procède en deux étapes : dans un premier temps, à partir du spectre total, j'extrais les spectres de chaque source individuelle avec la méthode que je décris ci-dessous et dans un second temps, sur chaque spectre individuel, j'applique la méthode d'évaluation de la dose pour des sources individuelles, décrite au paragraphe 5.2.

Considérons un spectre d'un mélange de radioéléments, comme présenté en Figure 5-3 que j'utiliserai pour illustrer la méthode que je décris ici. Ce spectre provient d'un mélange que j'ai artificiellement formé à partir des mesures réelles individuelles de ^{133}Ba, ^{137}Cs et ^{152}Eu que j'ai utilisées en paragraphe 5.2.4. L'intérêt d'utiliser un mélange artificiel est que je connais les spectres de chaque source individuelle et donc je peux vérifier la validité de la méthode que je vais présenter.

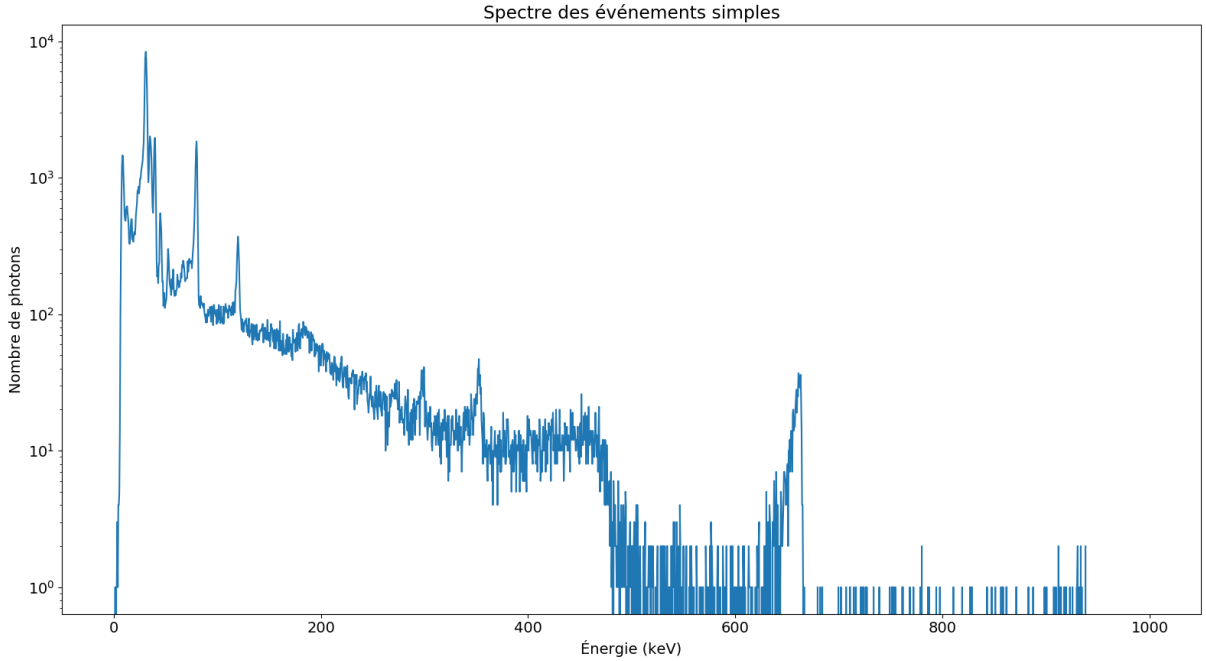


Figure 5-3 : Spectre du mélange utilisé.

La première étape consiste à identifier les radioéléments à l'origine de ce spectre et les quantifier, ce que je fais en utilisant mon approche par réseaux de neurones que j'ai présentée en partie 4. La Figure 5-4 présente le résultat des réseaux neurones sur le spectre précédent.

Les trois sources sont détectées de manière certaine. Les réseaux sont incertains sur la présence du ^{226}Ra-D, je considère la source comme absente ici si le réseau de neurones n'est pas certain de sa présence.

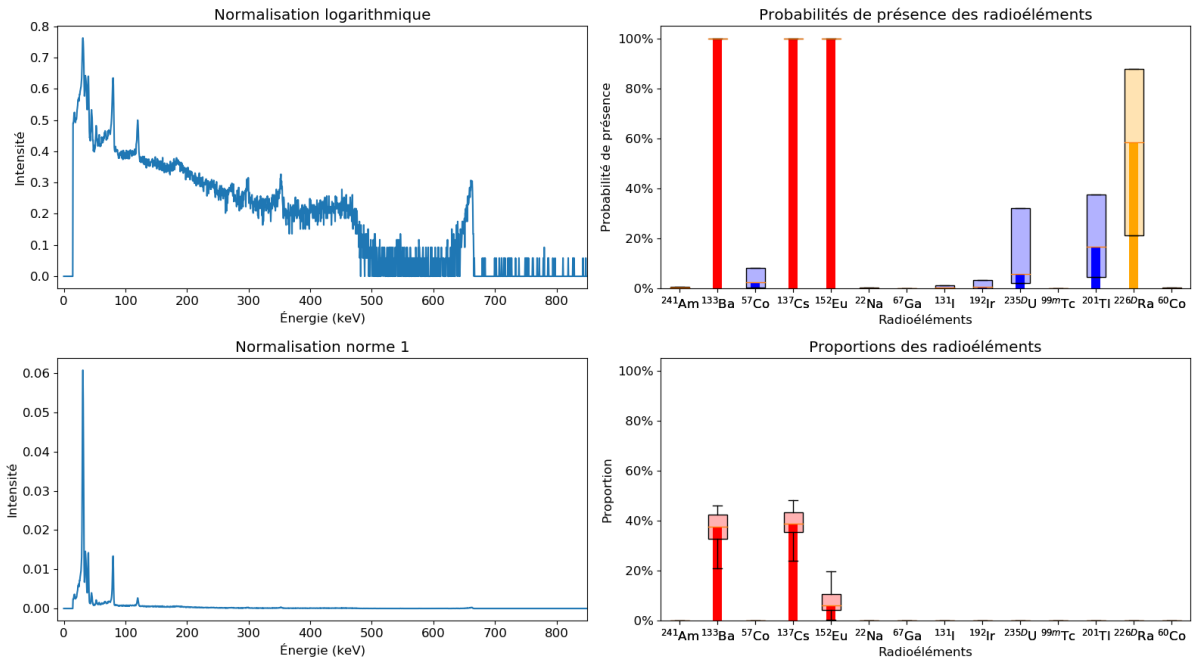


Figure 5-4 : Application des réseaux de neurones au spectre précédent.

Pour chaque source i détectée, je note $p(i)$ la proportion médiane déterminée par les réseaux de neurones. Je peux maintenant extraire les spectres individuels de la manière suivante.

Pour chaque canal j du spectre, associé à l'énergie E_j, ayant enregistré un nombre de photons N_j, je calcule la probabilité qu'un photon de cette énergie appartienne à la source i, que je note $p(i|j)$. Pour la calculer, j'applique la formule de Bayes :

$$p(i|j) = \frac{p(i) \times p(j|i)}{p(j)} \tag{5-9}$$

Le terme $p(i)$ est la probabilité qu'un photon d'énergie quelconque, détecté, provienne de la source i, c'est le terme de proportion calculé par le réseau de neurones.

Le terme $p(j|i)$ est la probabilité que la source i émette un photon d'énergie E_j : c'est donc le spectre de la source i. Pour évaluer ce terme, j'utilise le spectre issu des simulations de la source i.

Le terme $p(j)$ est un facteur de normalisation, que l'on peut calculer après avoir obtenu les termes du numérateur pour toutes les sources i.

$$p(j) = \sum_{i'} p(i') \times p(j|i') \tag{5-10}$$

Le nombre de photons associé à la source i pour le canal j du spectre est donné par N_{ij} :

$$N_{ij} = p(i|j) \times N_j \tag{5-11}$$

En appliquant cette méthode pour chaque source identifiée dans la Figure 5-4, j'obtiens les différents spectres individuels associés à chaque source en Figure 5-5. Je superpose le spectre attendu du fait que je connais chacun des spectres individuels par construction du mélange artificiel.

On voit que les spectres estimés par cette méthode sont proches de ceux attendus pour le ^{133}Ba et le ^{137}Cs. Les principaux écarts se voient sur le spectre du ^{137}Cs dans la zone entre 80 keV et 130 keV, où se situent les pics principaux du ^{152}Eu et les structures Compton associées.

Les différences sont plus visibles sur le spectre de ^{152}Eu. Bien qu'il reste relativement proche du spectre attendu, l'intensité du spectre estimé est globalement plus faible dans les structures continues Compton. Cela provient notamment du fait que le nombre de photons associés pour ce radioélément est beaucoup plus faible que pour les deux autres sources. Dans les structures Compton, le nombre de photons attendu pour le ^{152}Eu est de l'ordre de la dizaine par canal voire moins, alors que pour le ^{137}Cs, il est de l'ordre de la centaine à plusieurs dizaines.

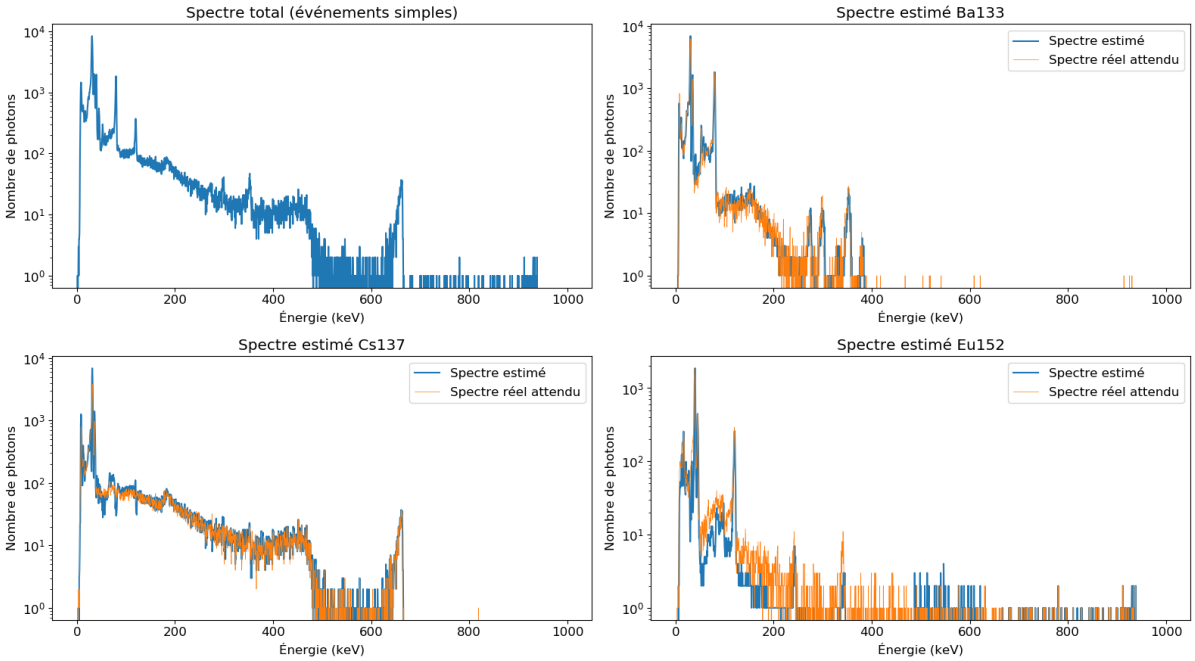


Figure 5-5 : Résultats du démélange.

Sur chaque spectre démélangé, j'applique ensuite chacun de mes réseaux de neurones de nouveau pour vérifier notamment que dans le spectre, on ne retrouve que le seul radioélément que je souhaitais obtenir. Les résultats correspondants sont présentés en Figure 5-6, Figure 5-7 et Figure 5-8.

Ces tests confirment la présence seule du radioélément correspondant au spectre extrait par le démélange. D'ailleurs, on note qu'on peut confirmer l'absence du ^{226}Ra-D pour lequel les réseaux de neurones ne savaient déterminer la présence ou l'absence avec certitude en Figure 5-4.

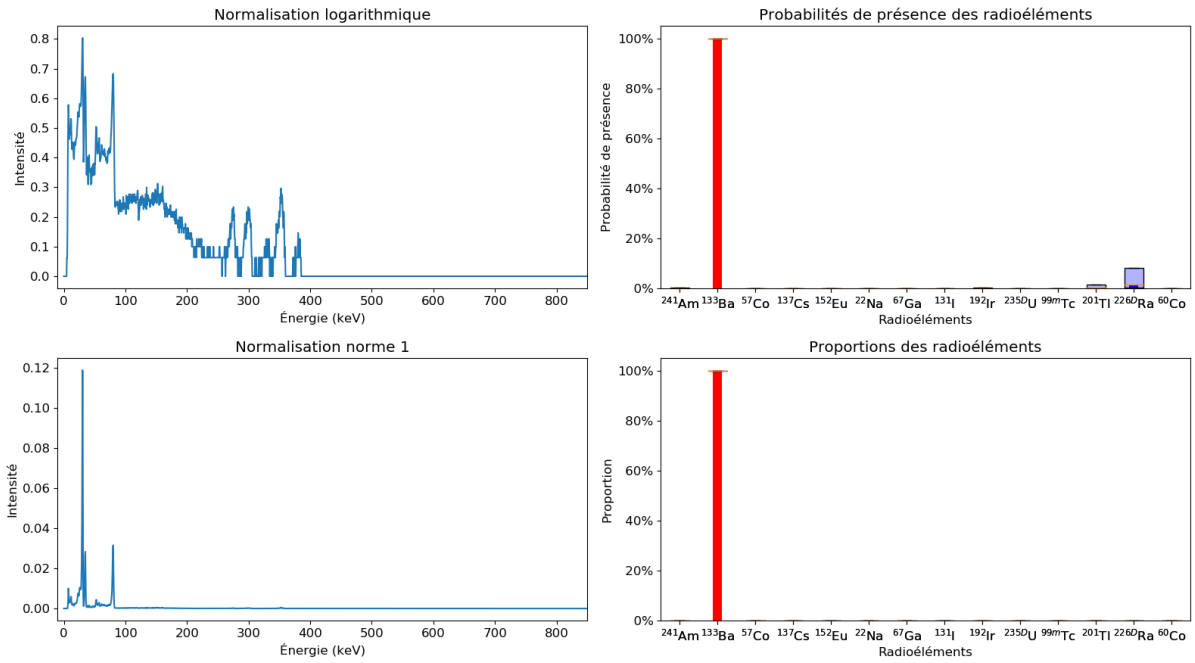


Figure 5-6 : Réseaux de neurones appliqués au spectre démélangé de ^{133}Ba.

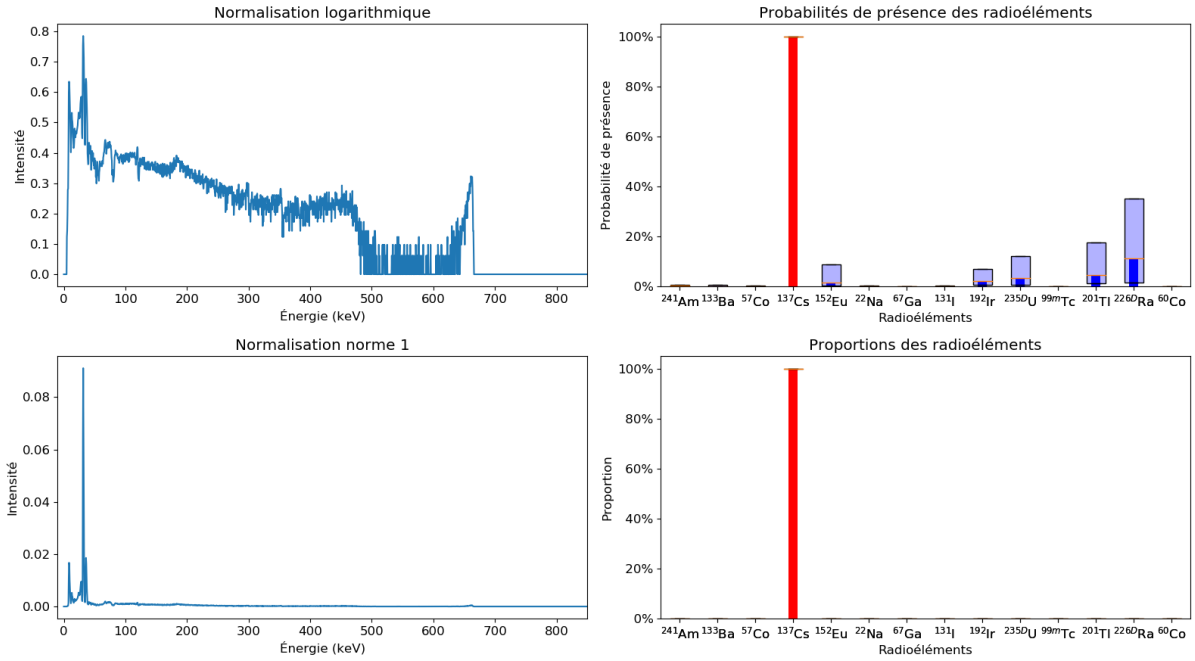


Figure 5-7 : Réseaux de neurones appliqués au spectre démélangé de ^{137}Cs.

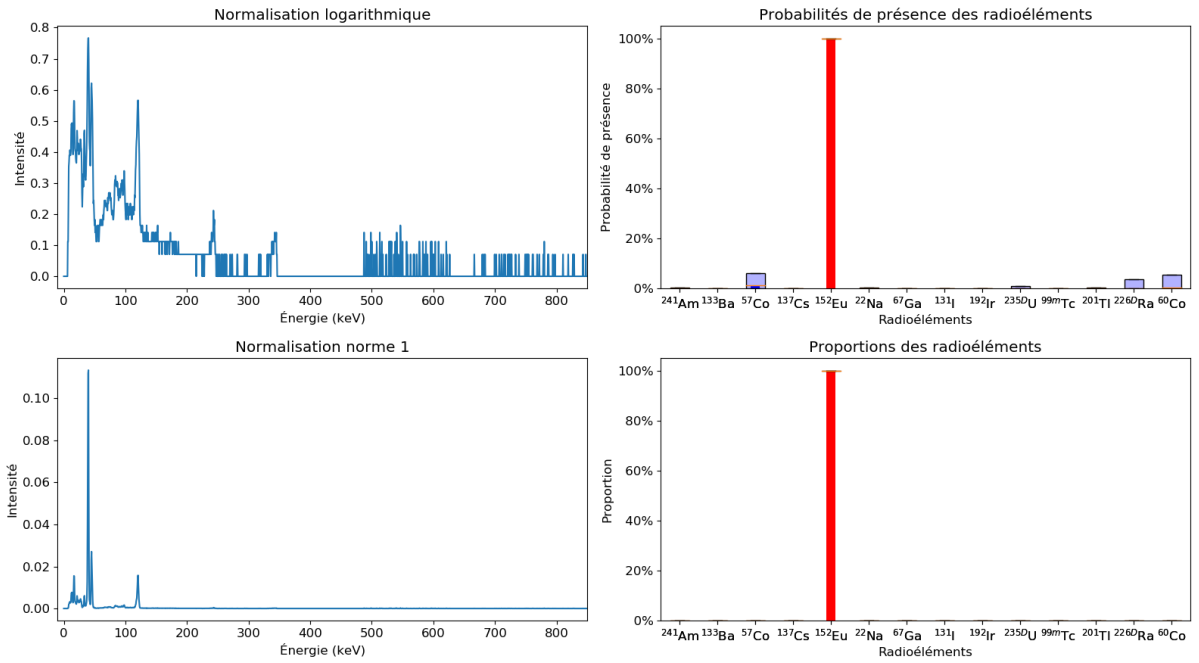


Figure 5-8 : Réseaux de neurones appliqués au spectre démélangé de ^{152}Eu.

Enfin, la dernière étape consiste à évaluer la dose de chaque source en utilisant la méthode décrite au paragraphe 5.2 sur chacun des spectres démélangés. Les résultats sont présentés dans le Tableau 5-2 et comparés aux valeurs théoriques et aux valeurs estimées précédemment.

Source	Débit de dose théorique (µSv/h)	Mesure Radiamètre (µSv/h)	Estimation du débit de dose à partir des mesures individuelles (µSv/h) (cf. paragraphe 5.2.4)	Estimation du débit de dose à partir des mesures dans le mélange (µSv/h)
^{133}Ba	$(2{,}4 \pm 0{,}4).10^{-1}$	$(2{,}5 \pm 0{,}3).10^{-1}$	$(3{,}0 \pm 0{,}5).10^{-1}$	$(2{,}1 \pm 0{,}2).10^{-1}$
^{137}Cs	$3{,}2 \pm 0{,}5$	$3{,}2 \pm 0{,}6$	$1{,}9 \pm 0{,}4$	$1{,}5 \pm 0{,}3$
^{152}Eu	$(2{,}6 \pm 0{,}4).10^{-1}$	$(3{,}5 \pm 1{,}0).10^{-1}$	$(2{,}7 \pm 1{,}8).10^{-1}$	$(0{,}9 \pm 0{,}3).10^{-1}$

Tableau 5-2 : Débits de dose estimés pour chaque radioélément individuel dans le mélange.

On voit sur cet exemple qu'on retrouve les mêmes ordres de grandeur pour les débits de dose que ceux estimés en utilisant les mesures individuelles. Pour le ^{133}Ba, le débit de dose que je retrouve est compatible avec le calcul théorique et celui mesuré avec le radiamètre. Pour le ^{137}Cs, on retrouve un débit de dose sous-estimé par rapport à la valeur théorique et à celle obtenue par le radiamètre, mais qui est compatible avec l'estimation faite en paragraphe 5.2.4. Enfin, le débit de dose estimé pour le ^{152}Eu dans le mélange sous-estime le débit de dose théorique d'un facteur presque 3. Cela provient probablement du fait que, dans le démélange, l'intensité du spectre estimé pour ce radioélément est globalement plus faible que l'intensité attendue. Notamment le domaine entre 60 keV et 120 keV est important dans le calcul du débit de dose du fait de la forte probabilité de détection à ces énergies.

Je souligne que, dans cette méthode, je n'ai pas pris en compte les incertitudes liées au calcul des proportions par les réseaux de neurones. Pour pouvoir faire cela, il faut notamment prendre en compte cette incertitude et la propager à la fois dans l'application de la formule de Bayes en équation (5-9) mais aussi à travers l'inversion bayésienne Compton, d'où la difficulté principale d'utiliser l'incertitude calculée dans ce type d'approche. Il serait par ailleurs intéressant d'évaluer les performances de cette méthode sur d'autres exemples.

6 Conclusion

À travers ce chapitre, j'ai montré l'intérêt et l'utilisation de réseaux de neurones que j'ai développés pour répondre au problème de spectro-identification et de quantification de la proportion de chaque radioélément identifié dans le signal.

J'ai mis en œuvre une méthode basée sur un apprentissage à partir de données de simulation, sur une base de données entièrement synthétique. Cette base de données est fondée sur un modèle physique ainsi que sur la réponse du détecteur que j'ai également développée au Chapitre 2. J'ai porté un soin particulier à la manière dont je normalise les données. Ma méthode distingue le problème d'identification et de quantification à travers deux réseaux de neurones spécifiques à chaque tâche pour chaque radioélément.

Les réseaux de neurones que j'ai utilisés sont des réseaux de neurones à couches de convolution, dont j'ai montré l'intérêt dans la manière dont ils extraient les caractéristiques discriminantes du signal pour répondre au problème d'identification des sources. Par ailleurs, ce sont des réseaux de neurones bayésiens qui permettent de fournir une incertitude « autoévaluée » sur le résultat qu'ils renvoient, information essentielle en instrumentation, mais qui est en général difficile à évaluer avec les approches par *deep learning*. La caractérisation de cette incertitude « autoévaluée » doit aussi être étudiée pour pouvoir l'assimiler à une incertitude réelle.

J'ai ensuite évalué les performances de mes réseaux de neurones sur des données réelles, sur lesquelles j'ai appliqué un changement volontaire de calibration pour montrer la capacité des réseaux de neurones à pouvoir répondre au problème malgré une modification de la calibration. J'ai montré que le taux de bonnes identifications atteint plus de 95 % lorsque les spectres sont constitués de quelques milliers de photons. Par ailleurs, j'ai montré aussi qu'avec un faible nombre de photons, de l'ordre de la centaine, le taux d'identifications correctes est déjà supérieur à 80 % pour l'ensemble des radioéléments testés. De même, j'ai montré que, grâce à l'évaluation des incertitudes, les réseaux de neurones peuvent répondre au problème de quantification, même si un soin doit encore être apporté au fait que les proportions des radioéléments dans le signal sont en général sous-estimées.

Actuellement, les réseaux de neurones utilisent les spectres construits à partir des événements simples des acquisitions, qui garantissent une meilleure résolution spectrale par rapport à l'utilisation de l'ensemble des événements, au détriment de la sensibilité du système, du fait que je n'utilise pas tous les événements. Il serait intéressant d'adapter cette méthode à l'utilisation de l'ensemble des événements pour évaluer le gain potentiel en sensibilité dans l'identification des radioéléments et la quantification de leurs proportions.

La Figure 6-1 présente un schéma de l'ensemble de la stratégie que j'ai mise en place pour répondre au problème de spectro-identification à partir des réseaux de neurones. J'ai mis en œuvre cette stratégie dans un code informatique qui traite en temps réel les acquisitions effectuées à partir des caméras WIX et qui fournit les résultats des réseaux de neurones au cours de l'acquisition.

Enfin, j'ai montré une application de cette approche à la dosimétrie, et plus particulièrement au calcul du débit de dose à partir de mesures issues du détecteur Caliste. J'ai présenté un traitement qui remonte au spectre incident à partir du spectre mesuré. Les méthodes que j'ai mises en place permettent d'estimer l'ordre de grandeur des débits de dose de chaque source individuelle dans un

mélange. Cependant, cette méthode doit encore être complétée notamment afin d'être plus performante du point de vue de la précision.

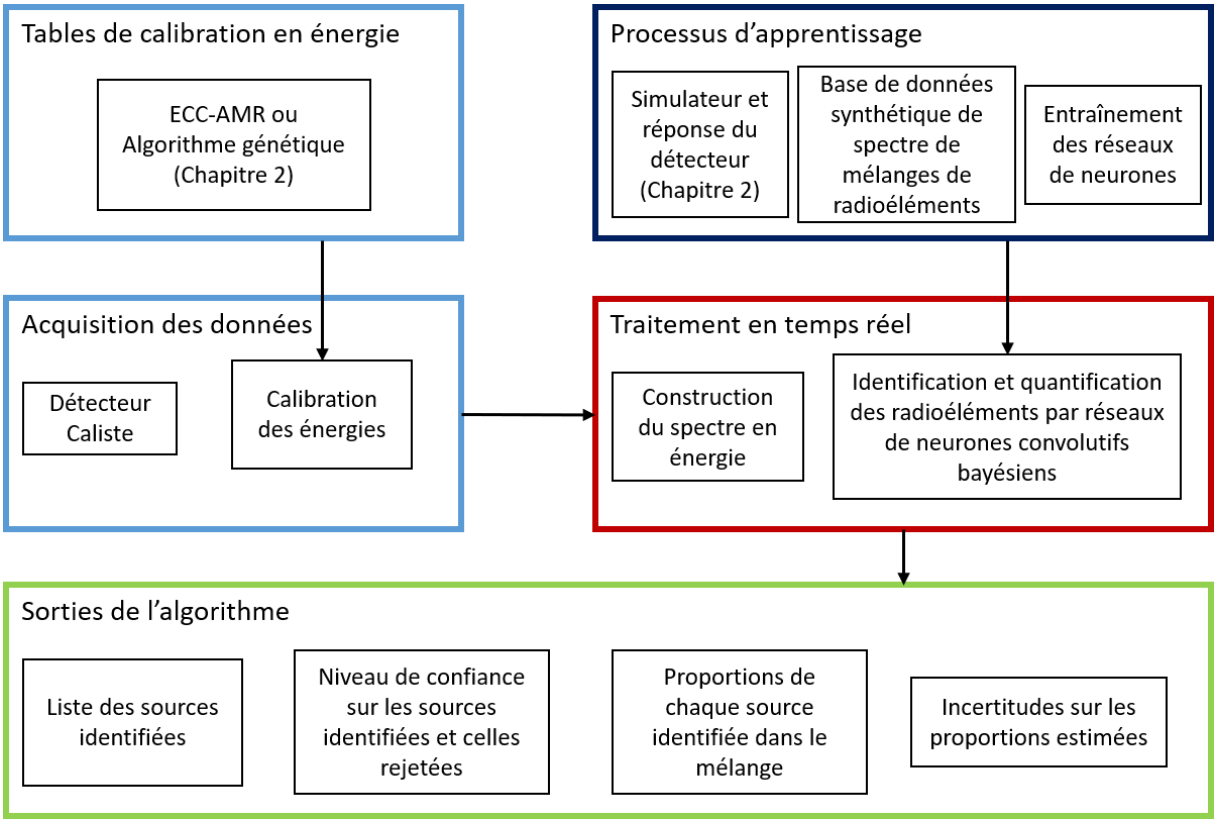


Figure 6-1 : Schéma de la méthode basée sur les réseaux de neurones que j'ai mise en place pour répondre au problème de spectro-identification.

7 Bibliographie

Abdel-Aal, R. & Al-Haddad, M., 1997. *Determination of radioisotopes in gamma-ray spectroscopy using abductive machine learning.* in: NIM-A, Volume 391, Issue 2, Pages 275-288.

Aloysius, N. & Geetha, M., 2017. *A review on deep convolutional neural networks.* in: 2017 International Conference on Communication and Signal Processing (ICCSP).

Boardman, D., Reinhard, M. & Flynn, A., 2012. *Principal Component Analysis of Gamma-Ray Spectra for Radiation Portal Monitors.* in: IEEE Transactions on Nuclear Science, vol. 59, no. 1, pp. 154-160.

Bobin, C. et al., 2016. *Real-time radionuclide identification in γ-emitter mixtures based on spiking neural network.* in: Applied Radiation and Isotopes, Volume 109, Pages 405-409.

Carlevaro, C., Wilkinson, M. & Barrios, L., 2008. *A genetic algorithm approach to routine gamma spectra analysis.* in: Journal of Instrumentation, Volume 3.

Chabot, G., 2016. *Relationship Between Radionuclide Gamma Emission and Exposure Rate.* in: Health Physics Society.

Chen, L. & Wei, Y., 2009. *Nuclide identification algorithm based on K–L transform and neural networks.* in: NIM-A, Volume 598, Issue 2, Pages 450-453.

Chollet, F., 2015. *Keras.* [Online]
Available at: https://github.com/fchollet/keras

Daniel, G. et al., 2020. *Automatic and Real-time Identification of Radionuclides in Gamma-ray Spectra: A new method based on Convolutional Neural Network trained with synthetic data set,.* in: IEEE Transactions on Nuclear Science, Proceedings of ANIMMA 2019.

Dubos, S., 2015. *Nouveau spectro-imageur CdTe à très haute résolution spatiale et spectrale pour l'astronomie X et gamma.* at: Université Paris 7 - Denis Diderot ; Commissariat à l'Energie Atomique et aux Energies Alternatives (CEA); Centre National d'Etudes Spatiales (CNES)..

Gal, Y., 2016. *Uncertainty in Deep Learning.* at: University of Cambridge, Thesis.

Gal, Y. & Ghahramani, Z., 2016. *Dropout as a Bayesian Approximation: Representing Model Uncertainty in Deep Learning.* in: Proceedings of the 33 rd International Conference on Machine Learning, New York, NY, USA, JMLR: W&CP volume 48.

Gomez-Fernandez, M. et al., 2020. *Status of research and development of learning-based approaches in nuclear science and engineering: A review.* in: Nuclear Engineering and Design, Volume 359, 110479.

Gunnink, R. & Arlt, R., 2001. *Methods for evaluating and analyzing CdTe and CdZnTe spectra.* in: NIM-A, Volume 458, Issues 1–2, Pages 196-205.

Hubbell, J. H. & Seltzer, S. M., 2004. *Tables of X-Ray Mass Attenuation Coefficients and Mass Energy-Absorption Coefficients from 1 keV to 20 MeV for Elements Z = 1 to 92 and 48 Additional Substances of Dosimetric Interest.* in: Radiation Physics Division, PML, NIST Standard Reference Database 126.

Kamuda, M., Stinett, J. & Sullivan, C., 2017. *Automated Isotope Identification Algorithm Using Artificial Neural Networks.* in: IEEE Transactions on Nuclear Science, Volume 64, Issue 7, Pages 1858-1864.

Kamuda, M., Zhao, J. & Huff, K., 2020. *A comparison of machine learning methods for automated gamma-ray spectroscopy.* in: NIM-A, Volume 954, 161385.

Kendall, A. & Gal, Y., 2017. *What uncertainties do we need in Bayesian deep learning for computer vision?.* in: NIPS'17: Proceedings of the 31st International Conference on Neural Information Processing Systems, Pages 5580–5590.

Kim, J., Park, K. & Cho, G., 2019. *Multi-radioisotope identification algorithm using an artificial neural network for plastic gamma spectra.* in: Applied Radiation and Isotopes, Volume 147, Pages 83-90.

Kingma, D. & Ba, J., 2015. *Adam: A Method for Stochastic Optimization.* in: Conference paper at the 3rd International Conference for Learning Representations, San Diego.

Liang, D. et al., 2019. *Rapid nuclide identification algorithm based on convolutional neural network.* in: Annals of Nuclear Energy, Volume 133, Pages 483-490.

Maier, D. et al., 2018. *Second generation of portable gamma camera based on Caliste CdTe hybrid technology.* in: Nuclear Instrumentation and Methods in Physics Research A, Volume 912, Pages 338-342.

Medhat, M., 2012. *Artificial intelligence methods applied for quantitative analysis of natural radioactive sources.* in: Annals of Nuclear Energy, Volume 45, Pages 73-79.

Meleshenkovskii, I., Pauly, N. & Labeau, P.-E., 2019. *Determination of the uranium enrichment without calibration standards using a 500 mm3 CdZnTe room temperature detector with a hybrid methodology based on peak ratios method and Monte Carlo counting efficiency simulations.* in: Applied Radiation and Isotopes, Volume 148, Pages 277-289.

Meuris, A., 2009. *Etude et optimisation du plan de détection de haute énergie en Cd(Zn)Te pour la mission spatiale d'observation astronomie X et gamma SIMBOL-X.* at: Université Paris-Diderot.

MIRION Technologies, 2017. *Radiagem 2000 : Radiamètre individuel portable.* [Online] Available at: http://www.canberra.com/fr/produits/hp_radioprotection/pdf/Radiagem-2000-SS-C38559_FR.pdf

Nénot, J.-C.et al., 2007. *Recommandations 2007 de la Commission internationale de protection radiologique.* in: PUBLICATION 103 DE LA CIPR, IRSN, Éditions Tec&Doc, Lavoisier.

Olmos, P. et al., 1991. *A new approach to automatic radiation spectrum analysis.* in: IEEE Transactions on Nuclear Science, vol. 38, no. 4, pp. 971-975.

Phillips, G. W. & Marlow, K. W., 1976. *Automatic analysis of gamma-ray spectra from germanium detectors.* in: Nuclear Instruments and Methods, Volume 137, Issue 3, Pages 525-536.

Pilato, V., Tola, F., Martinez, J. & Huver, M., 1999. *Application of neural networks to quantitative spectrometry analysis.* in: Nuclear Instruments and Methods in Physics Research A, Volume 422, Issues 1–3, Pages 423-427.

Simon, A. et al., 2009. *Determination of actinide isotopic composition: Performances of the IGA code on plutonium spectra according to the experimental setup.* in: 2009 1st International Conference on Advancements in Nuclear Instrumentation, Measurement Methods and their Applications, Marseille, pp. 1-8..

Sullivan, C. J. & Lu, J., 2013. *Automated photopeak detection and analysis in low resolution gamma-ray spectra for isotope identification.* in: 2013 IEEE Nuclear Science Symposium and Medical Imaging Conference (2013 NSS/MIC), Seoul, pp. 1-6.

Sullivan, C. J. & Stinnett, J., 2015. *Validation of a Bayesian-based isotope identification algorithm.* in: NIM-A, Volume 784, Pages 298-305.

Sutskever, I., Martens, J., Dahl, G. & Hinton, G., 2013. *On the importance of initialization and momentum in deep learning.* Atlanta, Georgia: Proceedings of the 30th International Conference on Machine Learning, Volume 28, Pages 1139-1147.

Szymańska, K. et al., 2008. *Resolution, efficiency and stability of HPGe detector operating in a magnetic field at various gamma-ray energies.* in: NIM-A, Volume 592, Issue 3, Pages 486-492.

U.S. Department of Homeland Security, 2004. *American National Standard Performance Criteria for Hand-held Instruments for the Detection and Identification of Radionuclides.* in: IEEE Standards.

Vigneron, V., Morel, J., Lepy, M. & Martinez, J., 1995. *Statistical modelling by neural networks in gamma-spectrometry.* in: Conference and International Symposium on Radionuclide Metrology, Paris, France.

Whitley, D., Starkweather, T. & Bogart, C., 1990. *Genetic algorithms and neural networks: optimizing connections and connectivity.* in: Parallel Computing, Volume 14, Issue 3, Pages 347-361.

Yoshida, E., Shizuma, K., Endo, S. & Oka, T., 2002. *Application of neural networks for the analysis of gamma-ray spectra measured with a Ge spectrometer.* in: NIM-A, Volume 484, Issues 1–3, Pages 557-563.

Chapitre 4 : Imagerie à masque codé – Présentation et méthodes classiques

1 L'imagerie à masque codé _____ 168

2 Algorithmes de localisation existants _____ 176

3 Application des algorithmes de déconvolution et MLEM, aux données Caliste _____ 181

4 Approfondissements sur l'algorithme MLEM _____ 199

5 Conclusion sur les algorithmes classiques et leurs limitations _____ 203

6 Bibliographie _____ 204

La localisation de sources radioactives émettrices de photons X et gamma représente une problématique majeure dans différents domaines liés aux applications nucléaires, comme la sûreté nucléaire, le démantèlement ou la médecine nucléaire. Les méthodes d'imagerie conventionnelle utilisées en optique classique ne sont pas adaptées à l'imagerie X et gamma et d'autres méthodes doivent être mises en œuvre.

Dans ce chapitre, je m'intéresse à l'imagerie X et gamma par masque codé, une des méthodes principalement utilisées pour répondre à cette problématique, nécessitant l'utilisation de détecteurs X et gamma pixellisés.

Je présente le principe de l'imagerie à masque codé et je pose le problème à traiter en termes d'analyse de données, à savoir comment retrouver la position de la ou des sources à partir de la mesure du détecteur.

Puis, je présente les algorithmes classiquement utilisés (déconvolution et MLEM) pour répondre à ce problème. J'applique ces méthodes à des données réelles acquises avec la caméra gamma WIX basée sur la technologie Caliste dotée d'un masque codé aléatoire. J'en étudie précisément les performances, notamment sur la base de critères quantitatifs et objectifs que je propose, non usuels dans le domaine de l'imagerie gamma. Enfin, j'analyse les limitations de ces méthodes et explore des voies d'approfondissement, notamment s'agissant de critères d'arrêt MLEM ou encore de localisation 3D.

C'est au chapitre suivant que je proposerai une méthode radicalement nouvelle basée sur le *deep learning* et les réseaux de neurones pour résoudre le même problème d'imagerie à masque codé en surmontant les limitations identifiées ici.

1 L'imagerie à masque codé

1.1 Principe de ce système d'imagerie

Envisagée à partir des années 1960 pour des applications dans le domaine de l'astronomie X et gamma (Dicke, 1968) et mise en œuvre à partir des années 1980 (Skinner, 1984), l'imagerie à masque codé peut être vue comme une généralisation du sténopé.

Dans le cas d'un sténopé, une surface opaque* percée d'un unique trou est placée devant le plan de détection. Suivant la position de la source à imager, les photons traversent le trou suivant la direction de la source et interagissent sur le plan de détection à la position alignée avec le trou du sténopé et la source. Le principal inconvénient du sténopé est le faible flux de photons atteignant le détecteur ainsi que la grande sensibilité au bruit.

L'idée du masque codé est de placer une surface percée non pas d'un seul trou devant le plan de détection, mais d'une multitude de trous répartis suivant un *pattern* défini. Si une source se situe dans le champ de vue du masque, « l'anti-ombre »† des trous est projetée sur le plan de détection, comme montré en Figure 1-1.

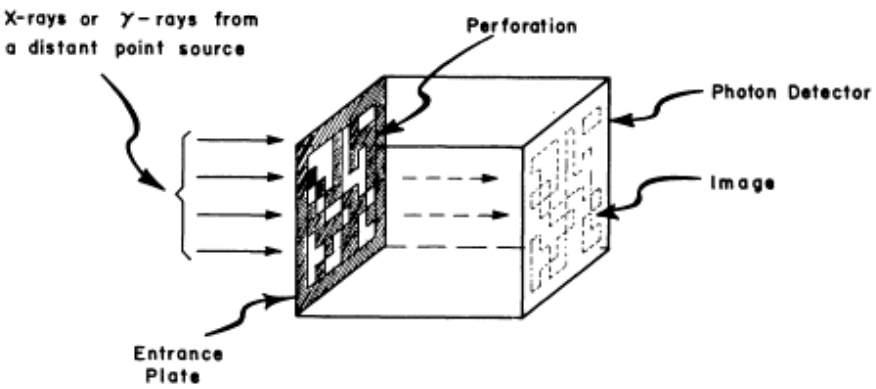


Figure 1-1 : Schéma du principe de l'imagerie à masque codé (Dicke, 1968).

L'image obtenue sur le détecteur est appelée *shadowgram*. Dans le cadre de l'imagerie X et gamma, le détecteur doit être un détecteur de photons X et gamma pixellisé, comme le détecteur Caliste. Le *shadowgram* représente alors le nombre de photons détectés par pixel du détecteur, ce qui correspond à une carte de coups.

Il existe différentes géométries et motifs de masque. Les masques peuvent être constitués de trous avec des motifs régulièrement répétés, on parle de masque URA (*Uniformly Redundant Array*), comme celui utilisé par le télescope SIGMA dans le cadre de la mission Granat (Bouchet, et al., 2001), ou de masque MURA (*Modified Uniformly Redundant Array*) comme celui de l'instrument IBIS à bord de la mission Integral (Goldwurm, et al., 2003). L'avantage de ce type de masque est le fait que l'algorithme de déconvolution que je présente en paragraphe 2.1 présente une matrice de corrélation inverse

* Opaque pour les photons aux longueurs d'onde (ou d'énergie) d'intérêt, X et gamma ici

† En général, on se représente plutôt l'ombre d'un objet qui est projetée sur un plan. Cependant, dans le cadre de l'imagerie X et gamma, il arrive souvent que l'on travaille avec un faible nombre de photons, et dans ce cas, ce n'est pas par l'ombre du masque que l'on retrouve les sources mais par l'apparition de zones « lumineuses » sur le détecteur correspondant aux trous du masque, d'où l'appellation « anti-ombre ».

simple à calculer à partir du motif du masque. En revanche, le principal inconvénient est que la redondance des motifs peut faire apparaître des « fantômes », des fausses sources, sur les images de reconstruction.

Ce type de masque a aussi trouvé son application dans le domaine de la sûreté nucléaire, utilisé par la caméra gamma Gampix (Gmar, et al., 2011). Le masque MURA utilisé pour cette caméra présente la particularité suivante : par une rotation de 90°, le motif du masque est parfaitement inversé, dans le sens où les trous du masque deviennent des éléments opaques et les éléments opaques deviennent des trous. Une procédure appelée « masque/anti-masque » consiste à effectuer une mesure avec le masque en position normale et une mesure avec l'anti-masque avec une rotation de 90°. Ce faisant, les deux mesures sont alors différentes mais la composante de bruit de fond, qui n'est pas due à la source, est identique pour les deux mesures. Elle devient identifiable par soustraction des deux mesures sur les deux *shadowgrams* et peut être retranchée avant le traitement pour localiser les sources. Cette procédure, applicable dans des mesures au sol comme dans le cas de la sûreté nucléaire, traite directement la composante due au bruit (bruit de fond et bruit instrumental). Par ailleurs, cette méthode permet aussi d'identifier les « fantômes » qui apparaissent du fait de la redondance des motifs du masque et de les supprimer en conséquence.

D'autres géométries reposent sur des masques aléatoires comme l'instrument BAT à bord du satellite SWIFT (Barthelmy, et al., 2005) ou encore celui du télescope ECLAIRs à bord de la mission SVOM dont le lancement est prévu en 2021 (Schanne, et al., 2014) pour détecter et observer des sursauts gamma. Du fait de la non redondance des motifs, la probabilité de faire apparaître des sources fantômes à cause de la géométrie du masque est faible. Elle est éventuellement possible si, malgré le caractère aléatoire du masque, certaines zones du masque présentent des corrélations entre elles.

Les géométries des masques codés, à savoir leur type (aléatoire ou redondant), les motifs des trous, la taille des trous, la distance entre le masque et le détecteur, sont à optimiser suivant le plan de détection utilisé et l'application visée. En particulier, la pixellisation du détecteur, directement reliée à ses capacités d'imagerie, est un facteur essentiel à prendre en compte pour la conception du masque. L'optimisation de la géométrie du masque codé est un objet d'étude en soi, mais il n'est pas le sujet de mon travail. Mes recherches se concentrent principalement sur le développement et l'application d'algorithmes qui, à partir d'une mesure effectuée sur le détecteur et connaissant le masque utilisé, retrouvent la position des sources radioactives à l'origine des photons détectés.

1.2 Masque utilisé

La géométrie du masque que j'utilise se base sur un modèle réduit, adapté au détecteur Caliste, de la première version du masque d'ECLAIRs. Le masque finalement utilisé pour la mission SVOM a subi des modifications, notamment pour renforcer sa structure afin de satisfaire à des contraintes mécaniques liées à la phase de décollage de la fusée embarquant le satellite. Ce masque, présenté en Figure 1-2, est un masque aléatoire, dans le sens où les motifs des trous ont été générés aléatoirement, en respectant la contrainte liée au fait qu'un élément de masque ne peut naturellement pas être entouré d'éléments vides[*].

[*] On peut cependant imaginer des cas où les trous ne seraient pas constitués d'éléments vides mais d'éléments transparents aux énergies d'intérêt. Un tel masque pourrait être construit à l'aide de techniques de fabrication additive.

Une structure élémentaire de ce masque est un carré de 1 mm de côté. La matrice du masque est composée de 46x46 éléments. Le masque est un carré de 46 mm de côté, placé à 58 mm du détecteur Caliste-HD, dont les dimensions sont pour rappel de 1 cm x 1 cm.

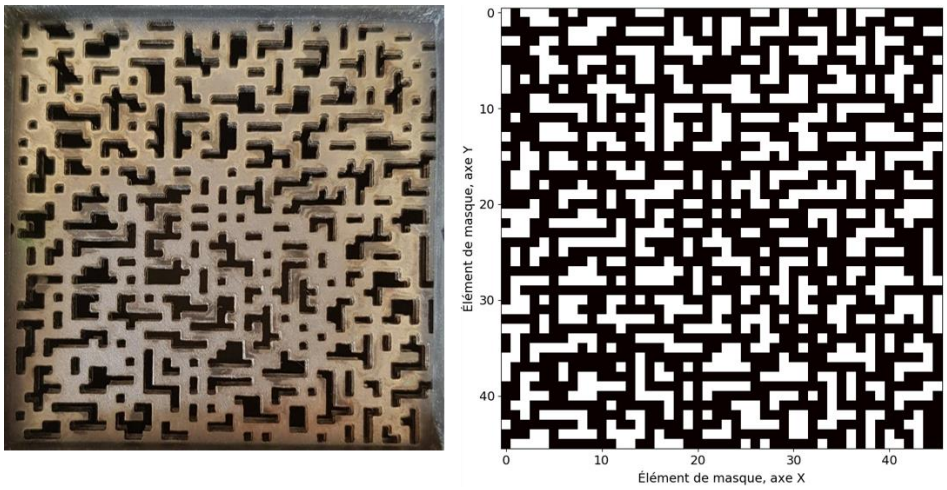


Figure 1-2 : À gauche : photo du masque codé utilisé. À droite : modèle numérique du masque (en noir : les éléments opaques, en blanc : les trous). Chacun des 46x46 éléments du masque a une taille de 1x1 mm², pour une taille totale de masque de 46x46 mm². L'épaisseur du masque peut varier entre 0,3 mm et 3 mm, par couche de 0,3 mm.

Avec ces éléments, on calcule le champ de vue entièrement codé et le champ de vue total. Une source se situe dans le champ de vue entièrement codé si la projection correspondante du masque sur le détecteur est complète, elle n'est pas partiellement sur le bord du masque, comme représenté en Figure 1-3. Le champ de vue total correspond à l'ensemble du champ de vue couvert par le masque par rapport au détecteur, comprenant les positions entièrement codées et partiellement codées, c'est-à-dire, où la projection est faite au niveau du bord du masque.

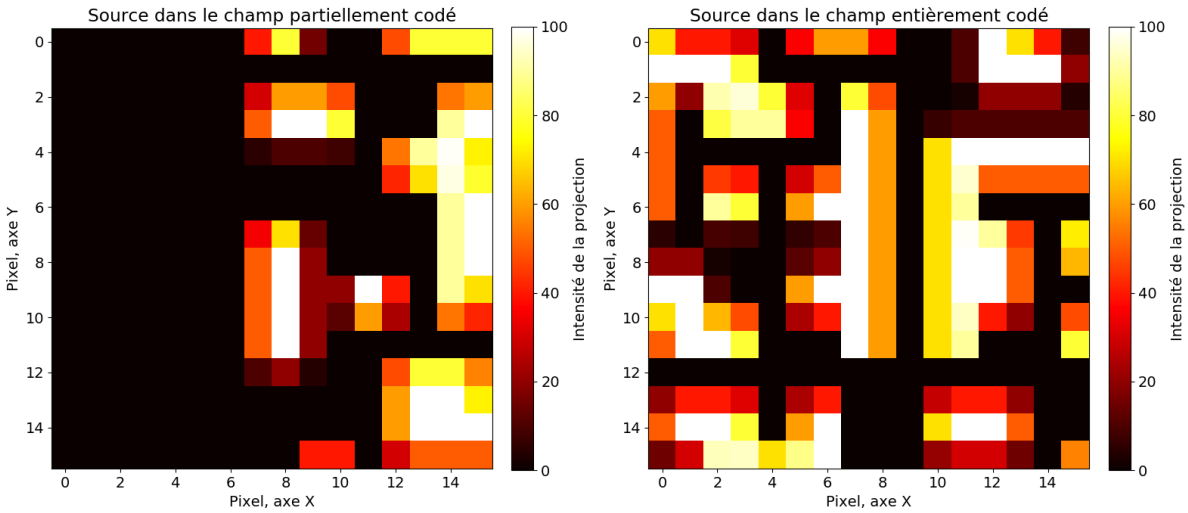


Figure 1-3 : Projection d'une source dans le champ de vue partiellement codé (à gauche) et entièrement codé (à droite). Certains pixels du détecteur n'ont pas une intensité maximale, ce qui correspond au fait qu'une partie du pixel est située sous un élément opaque du masque et l'autre partie est située sous un élément vide, autrement dit, l'ombre du masque ne recouvre qu'une partie du pixel.

La Figure 1-4 représente la taille du champ de vue entièrement codé θ_{ec} et celle du champ de vue total θ_{total}, qui se calculent avec les équations (1-1) et (1-2). Dans ces équations, L_d représente la longueur du détecteur, ici 10 mm, L_m la longueur du masque, soit 46 mm, et D, la distance entre le masque et le détecteur, soit 58 mm.

$$\theta_{ec} = 2\arctan\left(\frac{L_m - L_d}{2D}\right) \tag{1-1}$$

$$\theta_{total} = 2\arctan\left(\frac{L_m + L_d}{2D}\right) \tag{1-2}$$

Dans la configuration que j'utilise, $\theta_{ec} = 34{,}5°$ et $\theta_{total} = 51{,}5°$.

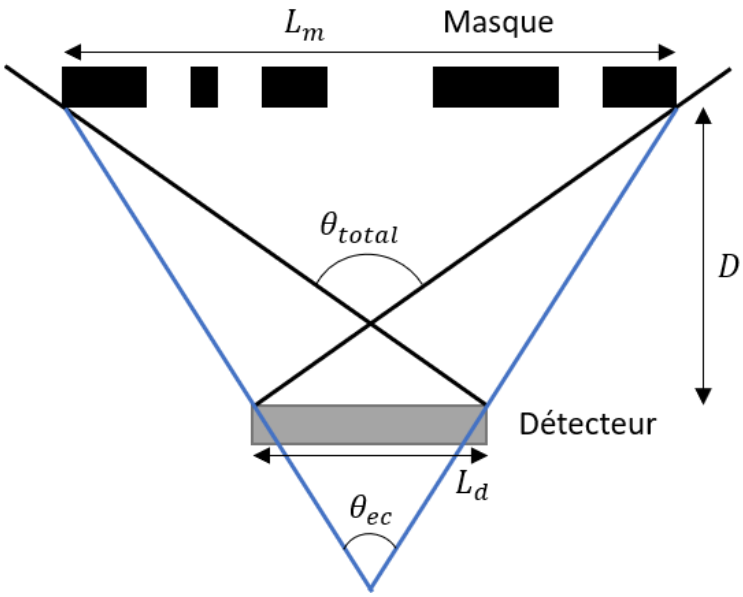


Figure 1-4 : Champ de vue entièrement codé (en traits bleus) et champ de vue total (en traits noirs).

Deux derniers paramètres importants pour définir le masque sont le matériau qui le compose et son épaisseur. Ces deux paramètres sont nécessaires pour connaître l'opacité du masque en fonction de l'énergie des photons qui vont le traverser.

Le modèle utilisé pour la caméra WIX est un masque en Tantale, matériau qui présente le double intérêt d'avoir un numéro atomique élevé, Z = 73, et une densité élevée, de 16,7. L'intérêt du système que j'utilise est la modularité de l'épaisseur du masque. En effet, le masque est composé de plaques de 0,3 mm d'épaisseur que je peux ajouter ou enlever suivant les applications. Je dispose de 10 plaques, ce qui me permet de faire varier l'épaisseur de 0,3 mm à 3 mm.

En Figure 1-5, je représente la probabilité d'interaction dans le masque en fonction de l'énergie du photon incident, pour chaque épaisseur de masque possible.

Je souligne notamment le fait que cela correspond à la probabilité d'interaction totale, donc l'interaction effectuée peut être soit une interaction photoélectrique, qui absorbe entièrement le photon, soit une diffusion cohérente (Rayleigh) ou incohérente (Compton). De ce fait, si l'interaction est une diffusion, le photon diffusé peut ressortir du masque et interagir dans le détecteur, constituant ainsi une source de bruit sur la mesure. Cependant, si le masque est épais, le photon diffusé peut aussi

interagir à son tour dans le masque. La cascade d'interactions ne peut cependant pas être modélisée simplement analytiquement et une simulation numérique Monte Carlo est nécessaire pour évaluer rigoureusement les probabilités d'absorption. Par ailleurs, si une diffusion Compton est effectuée, le photon perd une partie de son énergie, et grâce à une sélection en énergie sur les événements enregistrés par le détecteur, il est possible de ne pas tenir compte de ces photons.

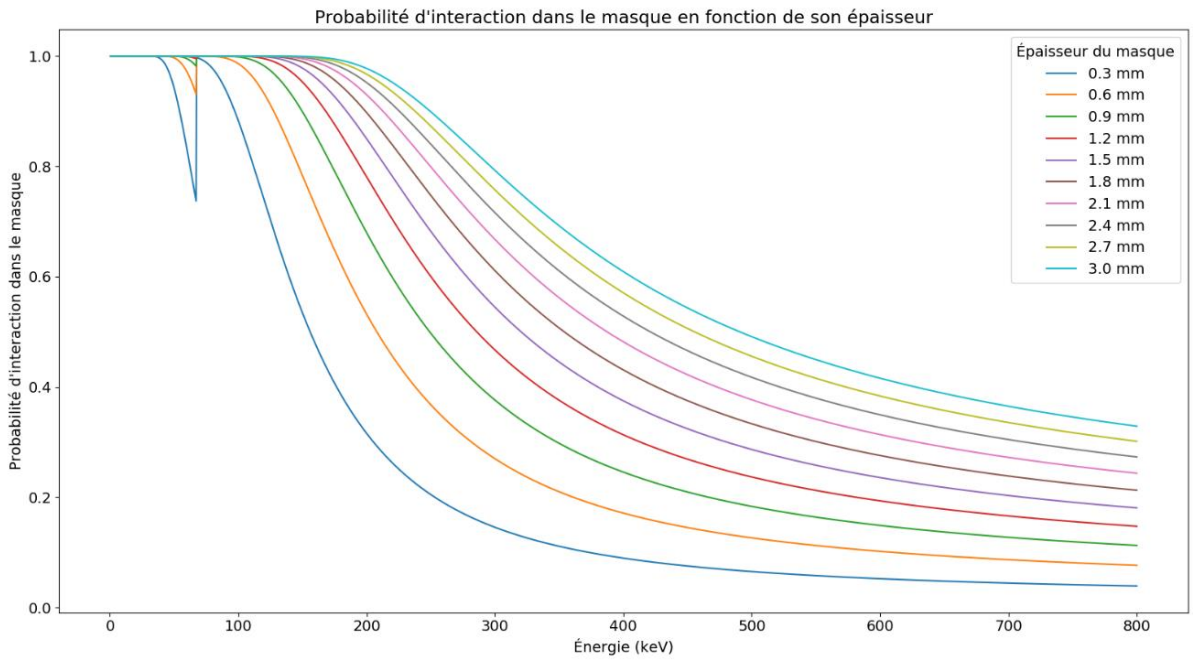


Figure 1-5 : Probabilité d'interaction dans le masque, en fonction de l'énergie du photon incident et de l'épaisseur du masque.

À travers la Figure 1-5, on notera que la probabilité d'interagir dans le masque diminue fortement autour de 67,4 keV pour les masques peu épais. Cela correspond à l'énergie du *K-edge* du Tantale. La probabilité d'interaction dans le masque passe sous les 50 % à partir de 150 keV environ pour le masque de 0,3 mm et à partir de 500 keV pour le masque de 3 mm. Cette méthode d'imagerie est ainsi plus difficilement utilisable pour les radioéléments émettant des photons d'énergie élevée, comme la ligne à 662 keV du ^{137}Cs.

Par conséquent, il semblerait intéressant de prendre le masque le plus épais possible, de façon à ce qu'il soit le plus opaque possible sur un grand domaine d'énergie. Cependant, cet avantage est contrebalancé par l'effet de *vignetting*, qui correspond au fait que l'épaisseur du masque réduit la fraction d'ouverture réelle des photons, suivant leur angle d'inclinaison, comme illustré en Figure 1-6.

La fraction d'ouverture f_{ouv} effectivement perçue en fonction de la direction θ de la source, pondérée par l'inclinaison géométrique de la direction de la source par rapport au détecteur $\cos(\theta)$, se calcule de la manière suivante, en notant e l'épaisseur du masque et L, la taille élémentaire du masque :

$$f_{\text{ouv}} = \max\left(0{,}1 - \frac{e}{L}\tan(\theta)\right) \times \cos(\theta) \tag{1-3}$$

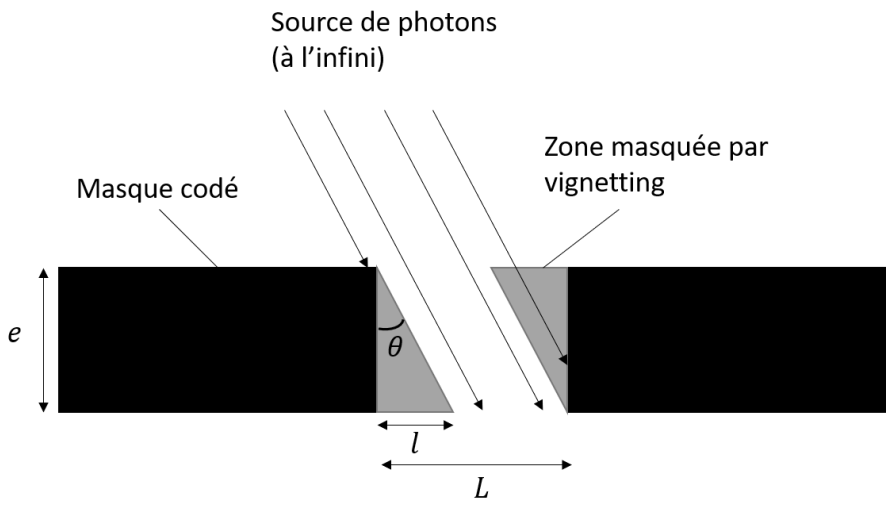


Figure 1-6 : Effet de *vignetting*.

La Figure 1-7 représente l'influence de l'épaisseur du masque à travers cet effet de *vignetting*. On voit notamment que pour un masque d'épaisseur de 0,3 mm, cet effet est assez faible, la fraction d'ouverture étant réduite au maximum de 22 %. Mais pour un masque de 3 mm d'épaisseur, cet effet est beaucoup plus marqué, la fraction d'ouverture est réduite à seulement 7 % au niveau de la limite du champ de vue entièrement codé, et tombe à 0 à partir de 18,5° d'inclinaison.

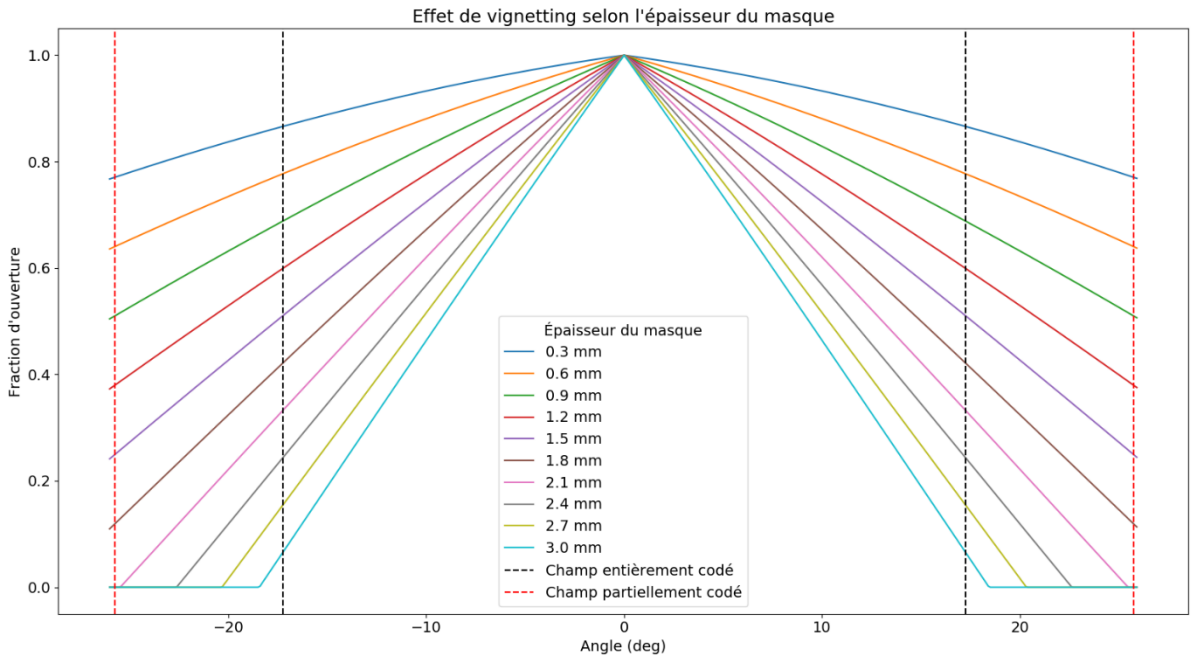


Figure 1-7 : Effet de *vignetting* selon l'épaisseur du masque en fonction de l'angle d'inclinaison des photons incidents.

Ce calcul est effectué en considérant que le masque est entièrement opaque. En fonction de l'énergie des photons incidents, une partie des photons peuvent tout de même traverser le masque, si l'épaisseur apparente est suffisamment faible pour que la probabilité d'interaction dans le masque ne soit pas trop élevée, comme calculé en Figure 1-5.

Dans la suite, j'effectue des mesures avec 3 couches de masque, soit une épaisseur de 0,9 mm. Ce choix garantit une fraction d'ouverture de 50 % au minimum sur tout le champ de vue, en étant opaque avec une probabilité de 50 % pour les photons d'énergie inférieure à 250 keV.

1.3 Problème d'inversion

L'objectif de l'imagerie à masque codé est de localiser les sources radioactives à partir d'une image en carte de coups enregistrés sur le détecteur, comme par exemple en Figure 1-8.

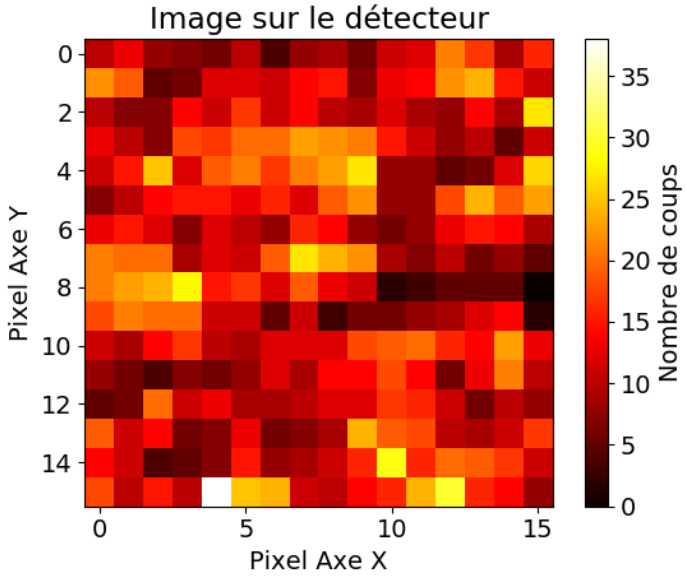


Figure 1-8 : Exemple de carte de coups. Acquisition avec Caliste-HD équipé du masque codé, mesure de deux sources de ^{241}Am situées à 1 m du détecteur, pendant 24 minutes.

Le problème peut alors se formuler de la manière suivante (Fenimore & Cannon, 1978). Considérons une mesure Y effectuée sur le détecteur. Y est un vecteur de taille N_{pix} où N_{pix} est le nombre total de pixels sur le détecteur.

On note X, un vecteur de taille N_{sky}, qui correspond à la discrétisation de l'espace à reconstruire, souvent appelé « ciel », notamment dans le cadre de l'astronomie. Je conserve cette dénomination par la suite.

Je note M la matrice de projections du masque, de taille $N_{\mathrm{pix}} \times N_{\mathrm{sky}}$. M_{ij} correspond à la projection sur le pixel i du détecteur d'une source à la position j sur le ciel. Une colonne $M_{.j}$ de M, de taille N_{pix} peut être représentée en 2D, comme une image de détecteur, à l'instar des images présentées en Figure 1-3.

Dans la suite de ce chapitre, je note les positions du ciel par un unique indice j. Pour un espace du ciel en deux dimensions, cet indice j représente classiquement deux coordonnées. De même, je noterai les pixels du détecteur par un unique indice i, qui représente le numéro du pixel.

J'introduis un vecteur ε de taille N_{pix} qui correspond au bruit attendu dans chaque pixel. Ce bruit est un terme général qui prend en compte d'une part les fluctuations statistiques au niveau du taux de comptage dans chaque pixel, obéissant à la statistique de Poisson. D'autre part, ce terme tient compte de l'ensemble des photons enregistrés qui ne proviennent pas de la source ou éventuellement ceux

qui ont pu traverser le masque sans interagir ou après diffusion, comme je l'ai mentionné au paragraphe 1.2.

L'équation (1-4) représente le problème général à résoudre :

$$Y = M.X + \varepsilon \tag{1-4}$$

Ce problème est en apparence relativement classique. Cependant, toute la difficulté réside dans le fait que le terme de bruit ε, inconnu, rend la résolution de ce problème particulièrement complexe, en particulier lorsque le nombre de photons détectés est faible. Par ailleurs, la matrice M n'est a priori pas inversible.

J'étudie les algorithmes classiquement utilisés pour traiter ce problème pour les masques codés en partie 2, un algorithme dit de déconvolution et un algorithme MLEM (*Maximum Likelihood Expectation Maximisation*). Ensuite, j'aborde ce problème avec des méthodes de *deep learning* en Chapitre 5, méthodes qui sont nouvelles dans le domaine de l'imagerie à masque codé.

Pour finir de poser le problème, il faut déterminer la manière dont on représente la pixellisation du ciel pour le vecteur X, ce qui est fondamental pour déterminer la matrice de projections M.

La projection du masque attendue sur le détecteur ne dépend pas seulement de la direction de la source mais aussi de la distance à la source. Dans le cas où la source observée se situe « à l'infini », comme dans le cas des applications en astronomie, la taille de l'ombre projetée sur le détecteur est égale à la taille du masque. Cependant, dans le cas où la source observée est plus proche du masque, un effet de grossissement est visible, l'ombre projetée a une taille apparente plus grande que la taille réelle du masque. Le facteur de grossissement γ est donné en équation (1-6), en fonction de la distance à la source D_{source} et la distance entre le masque et le détecteur $D = 58\ \text{mm}$:

$$\gamma = 1 + \frac{D}{D_{\text{source}}} \tag{1-5}$$

La Figure 1-9 représente l'évolution de ce facteur en fonction de la distance entre le masque et la source.

La taille apparente des pixels p_{ciel} du ciel correspond alors à la taille d'un élément de masque $p_{\text{mask}} = 10\ \text{mm}$ multiplié par le facteur de grossissement γ :

$$p_{\text{ciel}} = \gamma \times p_{\text{mask}} \tag{1-6}$$

Le nombre de pixels par ligne du ciel $N_{\text{sky}_{\text{row}}}$ situés dans le champ total du masque correspond au nombre d'éléments sur une ligne du masque $N_{\text{mask}_{\text{row}}} = 46$ multiplié par le rapport entre la taille apparente des pixels du ciel et la taille des pixels du détecteur $p_{\text{dét}} = 625\ \mu\text{m}$:

$$N_{\text{sky}_{\text{row}}} = N_{\text{mask}_{\text{row}}} \times \frac{p_{\text{ciel}}}{p_{\text{dét}}} \tag{1-7}$$

La discrétisation du ciel à reconstruire dépend donc de la distance à laquelle se situent les sources recherchées, distance qui n'est pas connue a priori. Dans la suite, on considérera que les sources sont suffisamment lointaines pour que le grossissement soit négligeable, sauf mention contraire. Je calcule

la matrice de projection M pour une distance de 3 m, distance à laquelle, d'après la Figure 1-9, le grossissement est de moins de 2 %, ce que je considère comme étant à l'infini.

À cette distance, d'après l'équation (1-7), le ciel est discrétisé en 75x75 éléments couvrant l'ensemble du champ de vue. Chaque élément a une taille angulaire de 0,7° en considérant le champ de vue total calculé en paragraphe 1.2. Pour chaque position du ciel, je calcule la projection, comme celles données en Figure 1-3.

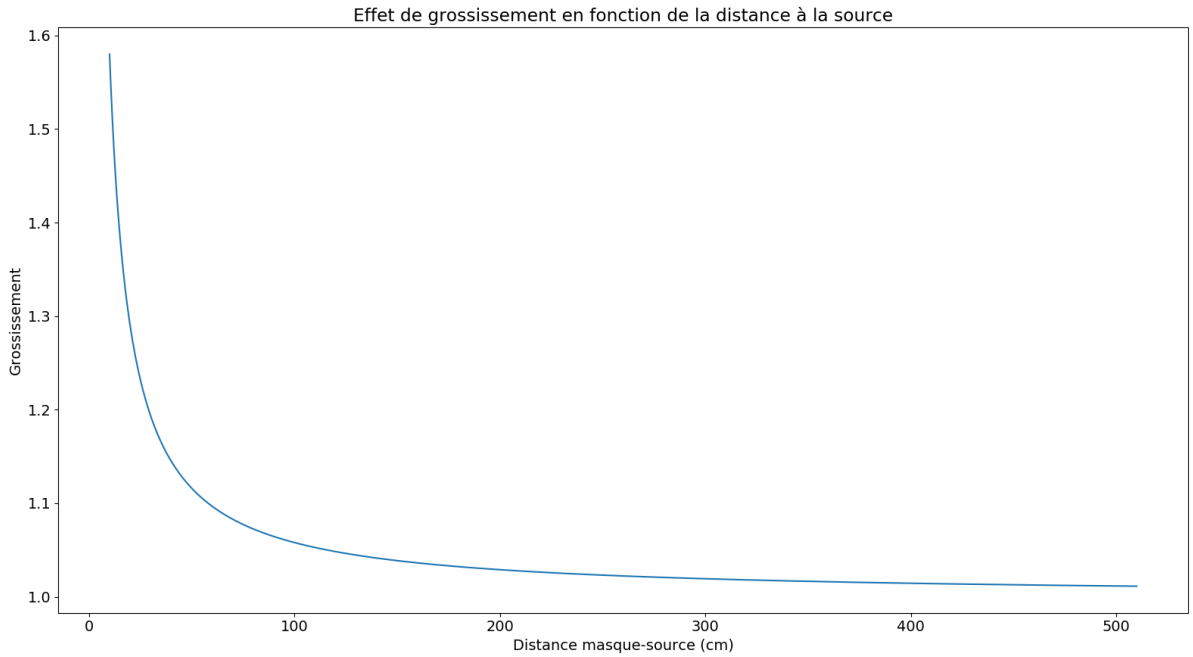


Figure 1-9 : Facteur de grossissement en fonction de la distance entre le masque et la source.

2 Algorithmes de localisation existants

Afin de résoudre le problème établi en paragraphe 1.3, il existe principalement deux algorithmes dans la littérature. Le premier est un algorithme de déconvolution, basé sur la corrélation entre le masque et les coups observés sur le détecteur. Le second est un algorithme itératif, appelé MLEM, qui cherche à maximiser la vraisemblance de l'observation par rapport aux paramètres du ciel qui permettent d'obtenir cette observation.

2.1 Algorithme de déconvolution

Dans le cadre de l'instrument ECLAIRs à bord de la mission SVOM, un algorithme de déconvolution est développé pour la reconstruction des sources dans le ciel (Antier-Farfar, 2016) (Goldwurm, et al., 2003).

Pour ce faire, une matrice dite de déconvolution G est construite à partir de la matrice M du masque de la manière suivante :

$$G = 2M - 1 \tag{2-1}$$

Dans cette équation, la soustraction par 1 est effectuée élément par élément.

À partir de cette matrice de déconvolution, on construit deux autres matrices G^+ et G^-, représentant respectivement les valeurs positives et négatives de G :

$$G^+ = G.(G > 0) \tag{2-2}$$

$$G^- = -G.(G < 0) \tag{2-3}$$

Par définition de G dans l'équation (2-1), G^+ représente la projection des éléments du masque ouverts à plus de 50 %, c'est-à-dire, les valeurs de M supérieures à 0,5, et inversement, G^- représente les projections des éléments du masque ouverts à moins de 50 %.

Considérons une mesure Y sur le détecteur. Y est un vecteur de taille 256, Y_i représente le nombre de coups observés dans le pixel i.

Le signal $\hat{X}_j$ reconstruit à la position j du ciel s'exprime de la manière suivante (Goldwurm, et al., 2003) :

$$\hat{X}_j = \frac{\sum_i G_{ij}^+ W_i Y_i}{\sum_i G_{ij}^+ W_i} - \frac{\sum_i G_{ij}^- W_i Y_i}{\sum_i G_{ij}^- W_i} \tag{2-4}$$

Où W est un vecteur de pondération appliqué aux pixels du détecteur, qui prend en compte les caractéristiques du détecteur. Par exemple, W_i peut être pris égal à 0 si le pixel i est éteint, ou il peut prendre une valeur comprise entre 0 et 1 si le seuil du pixel i est plus élevé que les autres pixels. Par défaut, si le pixel i est en fonctionnement normal, W_i est égal à 1.

La variance sur l'intensité reconstruite dans la position j du ciel est donnée par (Goldwurm, et al., 2003) :

$$V_j = \frac{\sum_i \left(G_{ij}^+ W_i\right)^2 Y_i}{\sum_i \left(G_{ij}^+ W_i\right)^2} + \frac{\sum_i \left(G_{ij}^- W_i\right)^2 Y_i}{\sum_i \left(G_{ij}^- W_i\right)^2} \tag{2-5}$$

À partir des équations (2-4) et (2-5), je peux calculer le rapport signal sur bruit noté SNR (*Signal-to-Noise Ratio*) associé à la position j du ciel :

$$SNR_j = \frac{\hat{X}_j}{\sqrt{V_j}} \tag{2-6}$$

Cette notion est particulièrement importante pour des instruments comme ECLAIRs, qui joue le rôle de détecteur de sursauts gamma à bord de la mission SVOM. De ce fait, un sursaut est considéré comme détecté si le rapport signal sur bruit dépasse une certaine valeur définie suivant les besoins en sensibilité de la mission. Une valeur élevée de ce seuil assure la fiabilité des détections, quitte à manquer des sursauts de plus faible SNR, alors qu'une valeur faible permet de s'assurer de manquer moins de sursauts, au risque d'avoir de fausses détections.

Les valeurs typiques de seuil en SNR utilisées dans le cadre d'ECLAIRs sont de l'ordre de 6 à 7 (Dagoneau, et al., 2018), du fait qu'il est important de rejeter au maximum les fausses détections. J'étudierai avec des données réelles les valeurs de SNR que je peux obtenir avec la caméra gamma en paragraphe 3. Notons aussi qu'avec l'équation (2-4), le signal $\hat{X}_j$ reconstruit peut être négatif, pouvant amener à un SNR négatif. Dans les figures que je présenterai, pour une meilleure lisibilité, je seuillerai les valeurs de façon à ne montrer que les SNR positifs.

2.2 *Maximum Likelihood Expectation Maximization* (MLEM)

Une autre façon d'aborder le problème décrit au paragraphe 1.3 est de le considérer d'un point de vue probabiliste. En conservant les notations $Y = (Y_i)_i$ pour le vecteur du nombre de coups observés dans chaque pixel et $X = \left(X_j\right)_j$ le nombre d'événements associés à la position j du ciel, on note $p(Y|X)$, la probabilité d'observer le vecteur Y sachant X.

Cette quantité est appelée la vraisemblance de X. Les méthodes basées sur le *Maximum Likelihood* consistent à maximiser cette vraisemblance, c'est-à-dire à trouver les paramètres X qui maximisent $p(Y|X)$. Dans un premier temps développé dans le cadre de la reconstruction tomographique (Shepp & Vardi, 1982), l'algorithme MLEM est un algorithme itératif qui atteint cet objectif dans le cas spécifique où $p(Y|X)$ se décompose suivant des lois de Poisson indépendantes.

Considérons un pixel i du détecteur et une position du ciel j. La probabilité qu'un photon soit détecté dans le pixel i sachant qu'il a été émis à la position j est donné par :

$$p(i|j) = \frac{M_{ij}}{\sum_{i'} M_{i'j}} \tag{2-7}$$

où M représente la matrice du masque définie au paragraphe 1.3.

De ce fait, si X_j photons sont émis par la position j et détectés par le capteur, le nombre moyen de photons attendus dans le pixel i est alors :

$$\lambda(i,j) = p(i|j)X_j \tag{2-8}$$

La probabilité de détecter, dans le pixel i, $N_{i,j}$ photons issus de la position j du ciel sachant que X_j photons ont été émis à cette position et détectés par le capteur est alors donnée par la loi de Poisson de paramètre $\lambda(i,j)$:

$$p\left(N_{i,j}\middle|X_j\right) = \frac{e^{-\lambda(i,j)}\lambda(i,j)^{N_{i,j}}}{\left(N_{i,j}\right)!} = \frac{e^{-p(i|j)X_j}\left(p(i|j)X_j\right)^{N_{i,j}}}{\left(N_{i,j}\right)!} \tag{2-9}$$

Avec ces notations, on a :

$$Y_i = \sum_j N_{i,j} \tag{2-10}$$

$$X_j = \sum_i N_{i,j} \tag{2-11}$$

Les nombres de coups obtenus dans chaque pixel i étant indépendants, et de même, les nombres de photons émis à chaque position du ciel j étant indépendants, la probabilité totale $p(Y|X)$ s'écrit comme le produit des probabilités élémentaires $p\left(N_{i,j}\middle|X_j\right)$:

$$L(X) = p(Y|X) = \prod_{i,j} p\left(N_{i,j}\middle|X_j\right) = \prod_{i,j} \frac{e^{-p(i|j)X_j}\left(p(i|j)X_j\right)^{N_{i,j}}}{\left(N_{i,j}\right)!} \tag{2-12}$$

en notant $L(X)$, la vraisemblance de X, donnée par $p(Y|X)$.

Dans ce cadre, un algorithme d'espérance-maximisation (EM) est utilisable pour aborder ce problème de maximisation.

Pour appliquer cet algorithme, on suppose que l'on possède déjà une estimation $\hat{X}^{\mathrm{old}}$ des paramètres optimaux X^{opt} qui maximisent la vraisemblance. On cherche avec l'algorithme EM à mettre à jour itérativement $\hat{X}^{\mathrm{old}}$ par $\hat{X}^{\mathrm{new}}$ avec les étapes d'espérance et de maximisation.

L'étape de calcul d'espérance consiste à calculer l'espérance conditionnelle $\mathbb{E}\left(X_j \middle| \hat{X}^{\mathrm{old}}, Y\right)$.

Du fait de l'équation (2-11) et de la linéarité de l'espérance :

$$\mathbb{E}\left(X_j \middle| \hat{X}^{\mathrm{old}}, Y\right) = \sum_i \mathbb{E}\left(N_{i,j} \middle| \hat{X}^{\mathrm{old}}, Y\right) \tag{2-13}$$

Calculons le terme élémentaire pour le pixel i : $\mathbb{E}\left(N_{i,j} \middle| \hat{X}^{\mathrm{old}}, Y\right)$. Comme on ne s'intéresse qu'au pixel i avec l'équation (2-10), on a :

$$\mathbb{E}\left(N_{i,j} \middle| \hat{X}^{\mathrm{old}}, Y\right) = \mathbb{E}\left(N_{i,j} \middle| \hat{X}^{\mathrm{old}}, Y_i\right) = \mathbb{E}\left(N_{i,j} \middle| \hat{X}^{\mathrm{old}}, \textstyle\sum_{j'} N_{i,j'}\right) \tag{2-14}$$

Or, si A et B sont deux variables qui suivent une loi de Poisson, on a l'identité suivante :

$$\mathbb{E}(A|A+B) = \frac{(A+B)\mathbb{E}(A)}{\mathbb{E}(A)+\mathbb{E}(B)} \tag{2-15}$$

Avec $A = N_{i,j}$ et $B = \sum_{j' \neq j} N_{i,j'}$, on obtient, en se rappelant que $p\left(N_{i,j} \middle| X_j\right)$ suit une loi de Poisson de paramètre $\lambda(i,j)$ et avec les équations (2-10) et (2-8) :

$$A + B = \sum_{j'} N_{i,j'} = Y_i \tag{2-16}$$

$$\mathbb{E}(A) = \mathbb{E}\left(N_{i,j} \middle| \hat{X}^{\mathrm{old}}\right) = \mathbb{E}\left(N_{i,j} \middle| \hat{X}_j^{\mathrm{old}}\right) = \lambda(i,j) = p(i|j)\hat{X}_j^{\mathrm{old}} \tag{2-17}$$

$$\mathbb{E}(B) = \mathbb{E}\left(\textstyle\sum_{j \neq j'} N_{i,j'} \middle| \hat{X}^{\mathrm{old}}\right) = \sum_{j \neq j'} \mathbb{E}\left(N_{i,j'} \middle| \hat{X}_j^{\mathrm{old}}\right) = \sum_{j \neq j'} \lambda(i,j') = \sum_{j \neq j'} p(i|j')\hat{X}_{j'}^{\mathrm{old}} \tag{2-18}$$

Cela donne donc en appliquant l'identité (2-15) avec les équations (2-16), (2-17) et (2-18) :

$$\mathbb{E}\left(N_{i,j} \middle| \hat{X}^{\mathrm{old}}, Y\right) = \frac{Y_i p(i|j)\hat{X}_j^{\mathrm{old}}}{\sum_{j'} p(i|j')\hat{X}_{j'}^{\mathrm{old}}} \tag{2-19}$$

Enfin, en reportant dans l'équation (2-13), on obtient :

$$\mathbb{E}\left(X_j \middle| \hat{X}^{\mathrm{old}}, Y\right) = \hat{X}_j^{\mathrm{old}} \sum_i \frac{Y_i p(i|j)}{\sum_{j'} p(i|j')\hat{X}_{j'}^{\mathrm{old}}} \tag{2-20}$$

L'étape « Espérance » de l'algorithme (EM) est terminée avec l'équation (2-20).

L'étape de maximisation est directe : l'équation (2-20) donne le nombre moyen estimé d'événements émis à la position j. Je peux identifier directement $\hat{X}_j^{\mathrm{new}}$ à $\mathbb{E}\left(X_j \middle| \hat{X}^{\mathrm{old}}, Y\right)$ qui maximise $p\left(X_j \middle| \hat{X}^{\mathrm{old}}, Y\right)$:

$$\hat{X}_j^{\text{new}} = \hat{X}_j^{\text{old}} \sum_i \frac{Y_i p(i|j)}{\sum_{j'} p(i|j') \hat{X}_{j'}^{\text{old}}} \tag{2-21}$$

L'équation (2-21) donne alors la procédure itérative qui met à jour l'estimateur des paramètres X^{opt} qui maximisent la vraisemblance, pour chaque position j du ciel.

Pour l'initialisation de l'algorithme, on définit $\hat{X}^0$ de sorte que l'ensemble des termes $\hat{X}_j^0$ soient strictement positifs et on calcule $\hat{X}^n$ avec l'équation (2-21) à chaque étape n. Les termes initiaux doivent être strictement positifs car, si un terme $\hat{X}_j^0$ est nul, par l'équation (2-21), l'ensemble des termes $\hat{X}_j^n$ resteront nuls à toutes les étapes.

Les détails de la convergence de la suite $\hat{X}^n$ vers les paramètres X^{opt} avec cet algorithme sont donnés dans (Shepp & Vardi, 1982), à travers la démonstration de la concavité de la log-vraisemblance $\log\big(L(X)\big)$ et du fait que cette log-vraisemblance obéit aux conditions de Karush-Kuhn-Tucker (KKT), garantissant respectivement la présence d'un maximum global et la convergence de la suite $\log\left(L\left(\hat{X}^n\right)\right)$ vers ce maximum. Je reprends cette démonstration en Annexe D et je propose aussi une autre justification de l'algorithme décrit par l'équation (2-21) avec un point de vue bayésien.

Dans l'équation (2-21), en fixant $\hat{X}_j^0$ à la première étape, tous les termes sont connus : Y_i provient directement de la mesure, c'est le nombre de coups détectés dans le pixel i et $p(i|j)$ provient de la connaissance du masque, comme donné en équation (2-7).

Le nombre d'itérations sur lequel il faut appliquer cet algorithme pour atteindre la convergence reste une question ouverte pour MLEM. Empiriquement, j'ai vu que 100 itérations étaient typiquement suffisantes pour atteindre une convergence avec les données issues du détecteur Caliste, ce qui représente 3 secondes de calcul sur un ordinateur classique. Il faut cependant noter qu'un nombre élevé d'itérations peut introduire du bruit supplémentaire dans la reconstruction par un phénomène de surinterprétation des données.

Pour conclure sur cette méthode, je souligne qu'on cherche ici à maximiser la vraisemblance $p(Y|X)$, c'est-à-dire que je cherche les paramètres X qui maximisent la probabilité d'observer les données Y. Cependant, ce ne sont pas forcément les paramètres X les plus probables au vu des données Y.

Pour trouver les paramètres X les plus probables, il faudrait idéalement calculer $p(X|Y)$, ce qui peut se faire à partir de la formule de Bayes :

$$p(X|Y) = \frac{p(Y|X)p(X)}{p(Y)} \tag{2-22}$$

Dans l'équation (2-22), la vraisemblance $p(Y|X)$ est celle que j'ai utilisée au long de ce paragraphe. Cependant, pour maximiser le terme en équation (2-22) sur X, il faut aussi prendre en compte le *prior* $p(X)$, ce qui présente deux difficultés. D'une part, il faut déterminer la forme de ce *prior*, ce qui n'est pas évident sans connaissance de la scène à analyser. D'autre part, il faut construire un nouvel algorithme pour faire cette maximisation, celui déterminé par l'équation (2-21) n'étant valable que pour la vraisemblance.

Un tel algorithme s'appelle MAP-EM (*Maximum A Posteriori Probability Expectation Maximization*). (Levitan & Herman, 1987) a démontré que l'utilisation de ce type d'algorithme peut se décliner sous un schéma itératif similaire à celui de MLEM dans le cas où le *prior* consiste en l'estimation a priori du nombre de photons émis à chaque position j du ciel, pondéré par la confiance que l'on a dans cette estimation. L'algorithme MAP-EM consiste alors à trouver un compromis entre l'algorithme MLEM et l'estimation a priori, qui joue le rôle de terme de régularisation. Cette méthode peut réduire le bruit dû à la surinterprétation des données dans le cas où on applique cet algorithme avec un nombre trop élevé d'itérations.

Cependant, lorsqu'on souhaite introduire des *priors* plus complexes, il devient beaucoup plus difficile de déterminer un schéma d'algorithme itératif. Un *prior* plus complexe peut consister à donner une plus forte probabilité sur la localisation de sources étendues peu nombreuses par rapport à une multitude de sources ponctuelles, ou en fonction de la scène, on peut aussi estimer plus probable a priori la présence de certaines formes, comme des tuyaux ou des bras de tuyaux, qui contiendraient plus vraisemblablement des sources radioactives.

3 Application des algorithmes de déconvolution et MLEM, aux données Caliste

3.1 Tests sur des sources ponctuelles

J'applique ici les deux algorithmes que j'ai décrits en partie 2 à des données acquises avec la caméra gamma WIX basée sur la technologie Caliste-HD. Un masque codé est placé devant le détecteur dans la configuration décrite au paragraphe 1.2.

Je montre des tests sur des sources ponctuelles uniques puis multiples. Dans un deuxième temps, je montre des résultats sur des sources étendues, issues de simulations cette fois, du fait que nous ne possédons pas de sources radioactives étendues en laboratoire. En Chapitre 5, je montre une manière de « mimer » des sources radioactives avec nos sources ponctuelles de laboratoire pour effectuer des tests sur des données réelles issues de mesures.

Pour les acquisitions réelles, je peux faire les images avec les données brutes, non calibrées en énergie, mais je peux aussi appliquer une calibration comme décrit au Chapitre 2, afin de faire des images dans des bandes d'énergie d'intérêt. Je préciserai dans les exemples si les données sont faites dans des bandes d'énergie ou si j'utilise la carte des coups totale sur l'ensemble des photons détectés. Par ailleurs, je n'utilise que les événements simples.

Les acquisitions étant souvent assez longues pour détecter un nombre de photons très grand pour des sources relativement éloignées du détecteur, de l'ordre d'un mètre de distance, il est possible que des mesures comportent des pixels bruyants suite au phénomène de polarisation du détecteur. Une difficulté est l'identification de ces pixels bruyants, qui nécessiterait une étude entière. En effet, du fait du masque codé, on s'attend à voir des pixels du détecteur plus lumineux que d'autres. Il n'est donc pas possible de simplement identifier les pixels bruyants en comparant simplement leur intensité à la moyenne du nombre de coups détectés.

J'ai cherché le quantile à 95 % du nombre de photons détectés par l'ensemble des pixels et j'ai retiré les pixels qui ont compté deux fois plus que ce quantile. Cela retire les pixels qui ont un comptage vraiment aberrant par rapport au reste des pixels, mais cela suppose qu'on n'a au maximum que 5 % de pixels bruyants, ce qui est une hypothèse raisonnable vu les mesures que j'ai pu effectuer. Par ailleurs, il apparaît que dans la plupart des mesures, aucun pixel bruyant n'apparaît, et cette règle conserve alors l'ensemble des pixels, en particulier pour des mesures courtes, de moins d'une heure.

Les résultats des algorithmes sont représentés sur une carte appelée image gamma. J'appellerai « fantômes » la localisation de sources au mauvais endroit, on parle aussi de faux-positifs.

3.1.1 Quantification des résultats : CNR (Contrast-to-Noise Ratio) et GCR (Ghost Contrast Ratio)

Dans la suite, je quantifierai la qualité des résultats pour les sources ponctuelles en analysant deux grandeurs, le *Contrast-to-Noise Ratio* (CNR) et *Ghost Contrast Ratio (GCR)*.

Considérons une zone d'intérêt dans laquelle se situe réellement la source ponctuelle. Le CNR dans cette zone, utilisé pour quantifier les résultats par exemple par (Zhang, et al., 2019), est défini tel que :

$$\mathrm{CNR} = \frac{(V_{\mathrm{int}} - \mu_{\mathrm{noise}})}{\sigma_{\mathrm{noise}}} \tag{3-1}$$

où V_{int} est la plus grande valeur de l'intensité des pixels dans la zone d'intérêt, dans l'image reconstruite du ciel, μ_{noise} et σ_{noise} sont respectivement la moyenne et la déviation standard des valeurs des pixels en dehors de cette zone.

Dans le cas des mesures réelles, je prends comme zone d'intérêt la position de la source avec l'incertitude de son positionnement soit 2 cm à 1 m, de l'ordre de grandeur de la taille du *packaging* de la source. Cette incertitude correspond à 1,14° à 1 m, c'est-à-dire une couronne de 1 pixel élémentaire du ciel autour de la position supposée. Cette zone est représentée en traits discontinus gris sur les images suivantes.

Dans le cas où plusieurs sources sont présentes dans le champ de vue, je calcule une valeur de CNR propre à chaque source, mais j'exclus les pixels du ciel correspondant aux autres sources pour le calcul de μ_{noise} et σ_{noise}. En effet, les autres sources ne doivent pas être considérées comme du bruit de reconstruction si elles sont effectivement présentes.

Le CNR quantifie la qualité de la reconstruction avec une valeur représentant le contraste entre la zone d'intérêt dans la reconstruction et le bruit de reconstruction en dehors de cette zone d'intérêt. Plus la valeur du CNR est élevée, plus le contraste donné par la reconstruction à la position de la source par rapport au bruit de fond est grand. Un CNR supérieur à 20 montre un contraste qui se détache du bruit de fond et que l'on considère comme signature de la présence d'une source. Le CNR peut aussi a contrario avoir des valeurs négatives, dans le cas où l'intensité de la reconstruction au niveau de la source est noyée dans le bruit de fond.

Le CNR est une grandeur qui quantifie la détection de la source dans la zone dans laquelle on s'attend à la trouver, mais il ne permet pas de juger de la présence de sources fantômes données par la reconstruction. Pour ce faire, je propose une seconde grandeur mesurable à chaque reconstruction, le *Ghost Contrast Ratio (GCR),* défini comme suit :

$$\mathrm{GCR} = \frac{V_{\mathrm{int}}}{V_{\mathrm{noise}}} \tag{3-2}$$

Où V_{int} est défini comme pour l'équation (3-1) et V_{noise} est la valeur maximale donnée par la reconstruction en dehors de la zone d'intérêt et des autres sources.

Un GCR inférieur à 1 signifie que l'algorithme a reconstruit une source fantôme plus intensément que la source attendue. Plus le GCR est grand, plus la reconstruction de la source se détache des potentiels fantômes. Par exemple, avec un GCR de 10, on peut considérer qu'il n'y a pas de fantôme significatif, la source est reconstruite avec une intensité 10 fois supérieure aux potentiels fantômes.

CNR	Interprétation	GCR	Interprétation
CNR < 0	Reconstruction moins intense que le bruit de fond	**GCR < 1**	Source reconstruite moins intensément qu'un fantôme
1 < CNR < 10	Peu de contraste par rapport au bruit de fond	**1 < GCR < 2**	Source ne se détache pas significativement des fantômes
10 < CNR < 20	Présence de source probable	**2 < GCR < 10**	Source au moins deux fois plus intense que les fantômes
CNR > 20	Signature de la présence de la source	**GCR > 10**	Source significativement plus intense que les fantômes
CNR >> 20	Source confirmée	**GCR >> 10**	Fantômes négligeables

Tableau 3-1 : Interprétations des quantités CNR et GCR.

3.1.2 Reconstruction d'images pour une source ponctuelle unique

Pour les acquisitions, j'utilise une source de ^{241}Am de 440 kBq, placée à 1 m de la caméra gamma. Le masque est utilisé avec 3 couches de 0,3 mm, soit un total de 0,9 mm. D'après la Figure 1-5, cette épaisseur est suffisante pour qu'au moins 99,5 % des photons interagissent, vu qu'ils ont une énergie très majoritairement inférieure à 60 keV pour ce radioélément.

Pour la première acquisition, j'ai placé la source dans l'axe du détecteur. L'acquisition est faite sur une longue durée de presque 3 jours, de façon à avoir un nombre de photons suffisant pour faire apparaître de manière visible la projection du masque sur le détecteur. J'utilise l'ensemble de la mesure, sans sélection en énergie en Figure 3-1 et avec sélection en énergie entre 50 keV et 70 keV, autour du pic principal du ^{241}Am en Figure 3-2.

La sélection en énergie permet de conserver les photons pertinents pour la localisation du ^{241}Am et de retirer les événements provenant potentiellement du bruit de fond ou du bruit à basse énergie de l'instrument lors de la mesure.

La comparaison entre la mesure et la projection attendue à la position de la source est donnée en Figure 3-3.

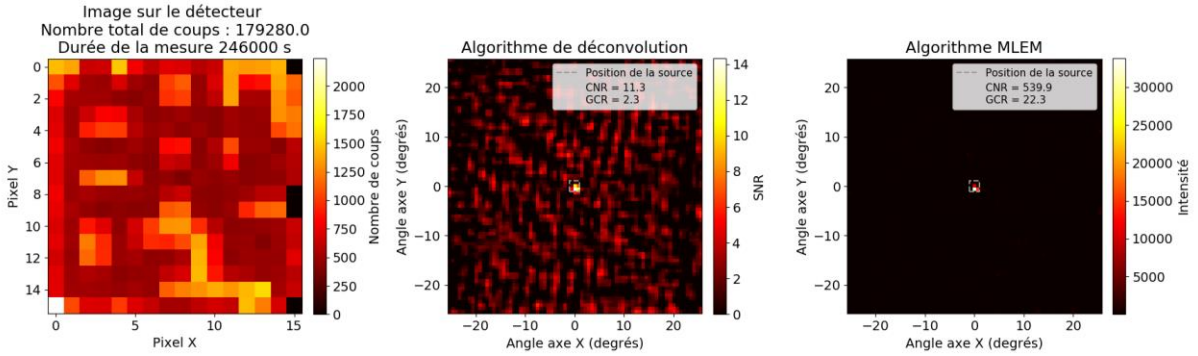


Figure 3-1 : Acquisition d'une source de ^{241}Am de 440 kBq placée à 1 m de la caméra gamma pendant 3 jours. Aucune sélection en énergie. Paramètres d'acquisition : temps de *peaking* 6,07 µs, gain 200 mV/fC, température du cristal -10 °C. Dans l'ordre de gauche à droite : carte de coups, résultat de l'algorithme de déconvolution et résultat de l'algorithme MLEM.

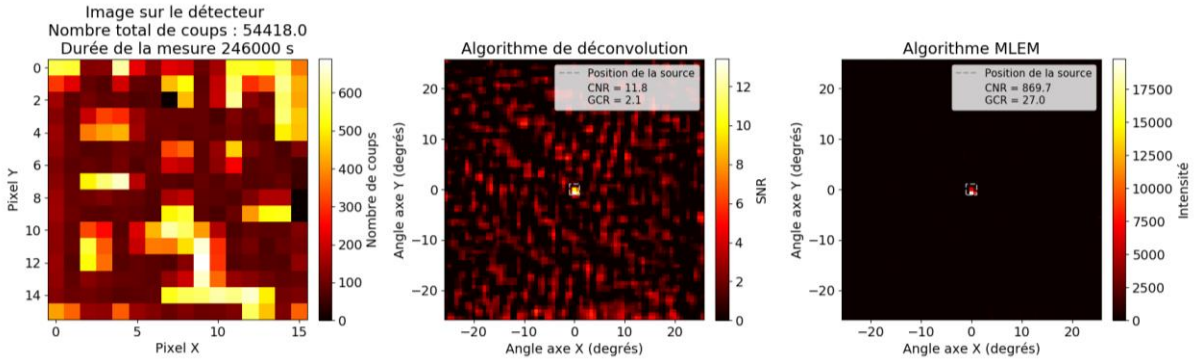


Figure 3-2 : Acquisition d'une source de ^{241}Am de 440 kBq placée à 1 m de la caméra gamma pendant 3 jours. Sélection en énergie entre 50 keV et 70 keV. Paramètres d'acquisition : temps de *peaking* 6,07 µs, gain 200 mV/fC, température du cristal -10 °C. Dans l'ordre de gauche à droite : carte de coups, résultat de l'algorithme de déconvolution et résultat de l'algorithme MLEM.

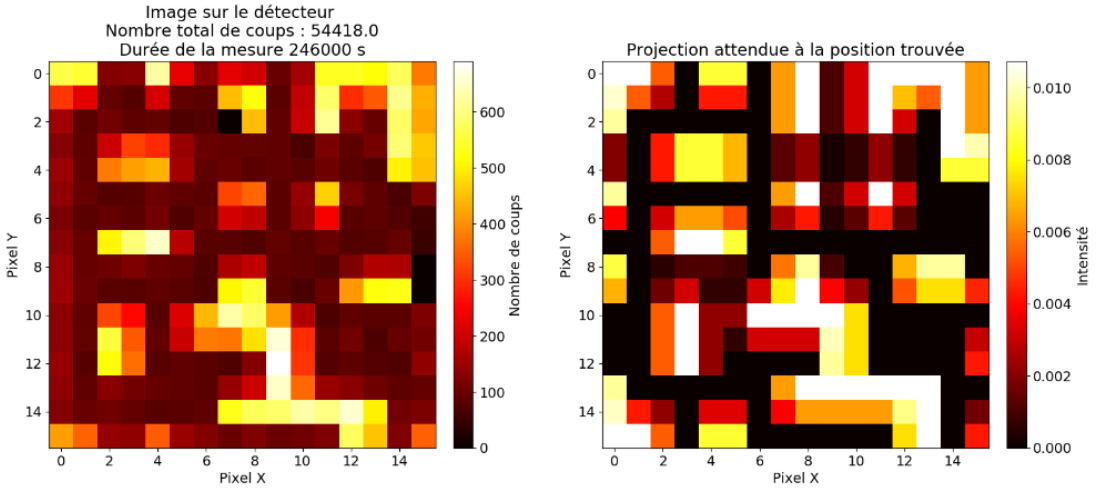


Figure 3-3 : Comparaison entre la projection mesurée et la projection attendue à la position trouvée par les algorithmes.

Les localisations réalisées par les deux algorithmes sont très similaires, avec et sans sélection en énergie, ce qui montre une certaine robustesse de ces algorithmes. Le CNR varie de 11,3 à 11,8 pour l'algorithme de déconvolution et de 540 à 870 pour MLEM. La source est localisée au centre du champ de vue, avec un SNR de 14 environ pour l'algorithme de déconvolution, et de plus de 10 sur deux

positions adjacentes du ciel, à comparer à des SNR inférieurs à 5 sur le reste de la carte du ciel. Le GCR est supérieur à 2 pour cet algorithme.

L'algorithme MLEM retrouve aussi la position de la source sans ambiguïté au centre du champ de vue. L'intensité donnée par l'algorithme est de plus de 30 000 sans sélection en énergie et de près de 20 000 avec sélection en énergie au niveau de la position centrale. Dans le reste du champ de vue, l'intensité est de moins de 1000, ce qui conduit à un GCR supérieur à 20.

Pour l'instant, j'ai utilisé la mesure longue de 3 jours de façon à avoir un très grand nombre de photons, notamment pour pouvoir faire la comparaison en Figure 3-3. Maintenant, je regarde en Figure 3-4 le résultat des algorithmes pour des temps d'acquisition beaucoup plus courts en sélectionnant les événements des premières dizaines de secondes de la même mesure.

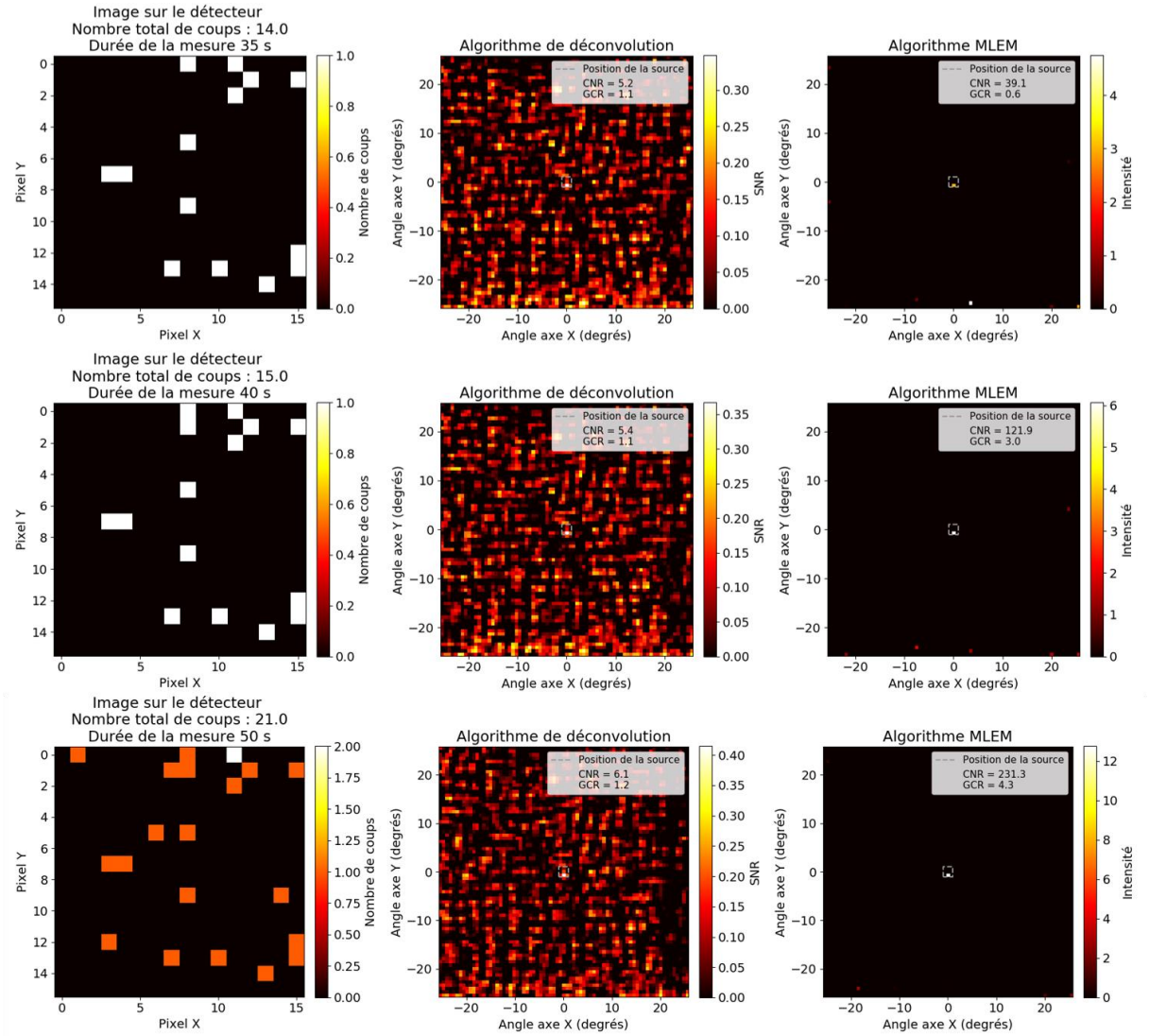


Figure 3-4 : Acquisition d'une source de ^{241}Am de 440 kBq placée à 1 m de la caméra gamma. Dans l'ordre de gauche à droite : carte de coups, résultat de l'algorithme de déconvolution et résultat de l'algorithme MLEM. De haut en bas : acquisition de 35 s, de 40 s et 50 s.

La Figure 3-4 illustre que 40 secondes suffisent pour retrouver la source à la bonne position avec l'algorithme MLEM avec un GCR de 3. En 50 secondes l'intensité attribuée à la bonne position est nettement supérieure et le GCR vaut 4,3. En revanche, avec seulement 35 secondes d'acquisition, la

position reconstruite la plus intense se situe en bas du champ de vue, soit au mauvais endroit - le GCR est inférieur à 1. Notons qu'entre la mesure à 35 secondes et celle à 40 secondes, il y a seulement un photon d'écart sur la carte des coups, et le résultat MLEM est pourtant très différent. Cela provient du phénomène de surinterprétation des données évoqué en fin de paragraphe 2.2.

Pour l'algorithme de déconvolution, le maximum de SNR est trouvé à la bonne position de la source dans les trois cas, mais ce maximum ne se distingue pas clairement du reste des positions, avec un GCR inférieur à 1,2. Par ailleurs, on voit que le SNR trouvé est de l'ordre de 0,4.

Pour atteindre un SNR de 6, qui correspond aux valeurs attendues pour l'instrument ECLAIRs dans le cadre de la mission SVOM, il faut une acquisition de près de 14 heures comme je le montre en Figure 3-5, alors que pour une telle acquisition, l'algorithme MLEM donne la bonne localisation en moins 40 secondes. Cela illustre le fait que l'interprétation de l'intensité en SNR telle que faite pour l'instrument ECLAIRs n'est pas transposable pour mon application.

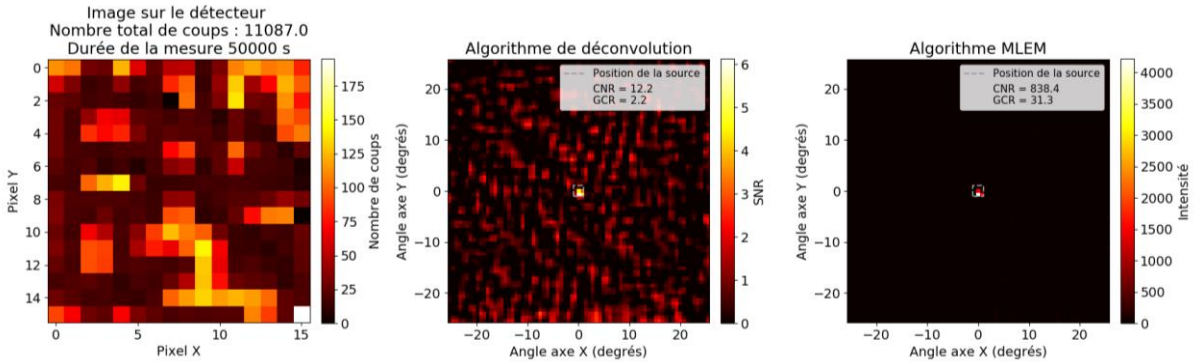


Figure 3-5 : Acquisition d'une source de ^{241}Am de 440 kBq placée à 1 m de la caméra gamma pendant 14 heures. Dans l'ordre de gauche à droite : carte de coups, résultat de l'algorithme de déconvolution et résultat de l'algorithme MLEM.

En Figure 3-6, je représente les résultats pour une acquisition de ^{241}Am hors axe du détecteur, à environ 10° à droite dans le champ de vue. De nouveau, en utilisant l'acquisition totale de près de 4 h 45 mn, la source est bien localisée par les deux algorithmes, en particulier par MLEM avec un GCR supérieur à 8.

En revanche, par rapport à la source positionnée au centre, environ 400 secondes sont nécessaires pour obtenir une localisation correcte avec MLEM pour ne plus avoir de fantômes apparaissant du fait de la surinterprétation des données par l'algorithme. Le GCR est en effet de 1,4 et de 7 à 300 et 400 secondes respectivement. 21 photons sont suffisants pour bien localiser la source dans l'axe alors qu'il faut 128 photons pour bien la localiser à 10° sans fantôme.

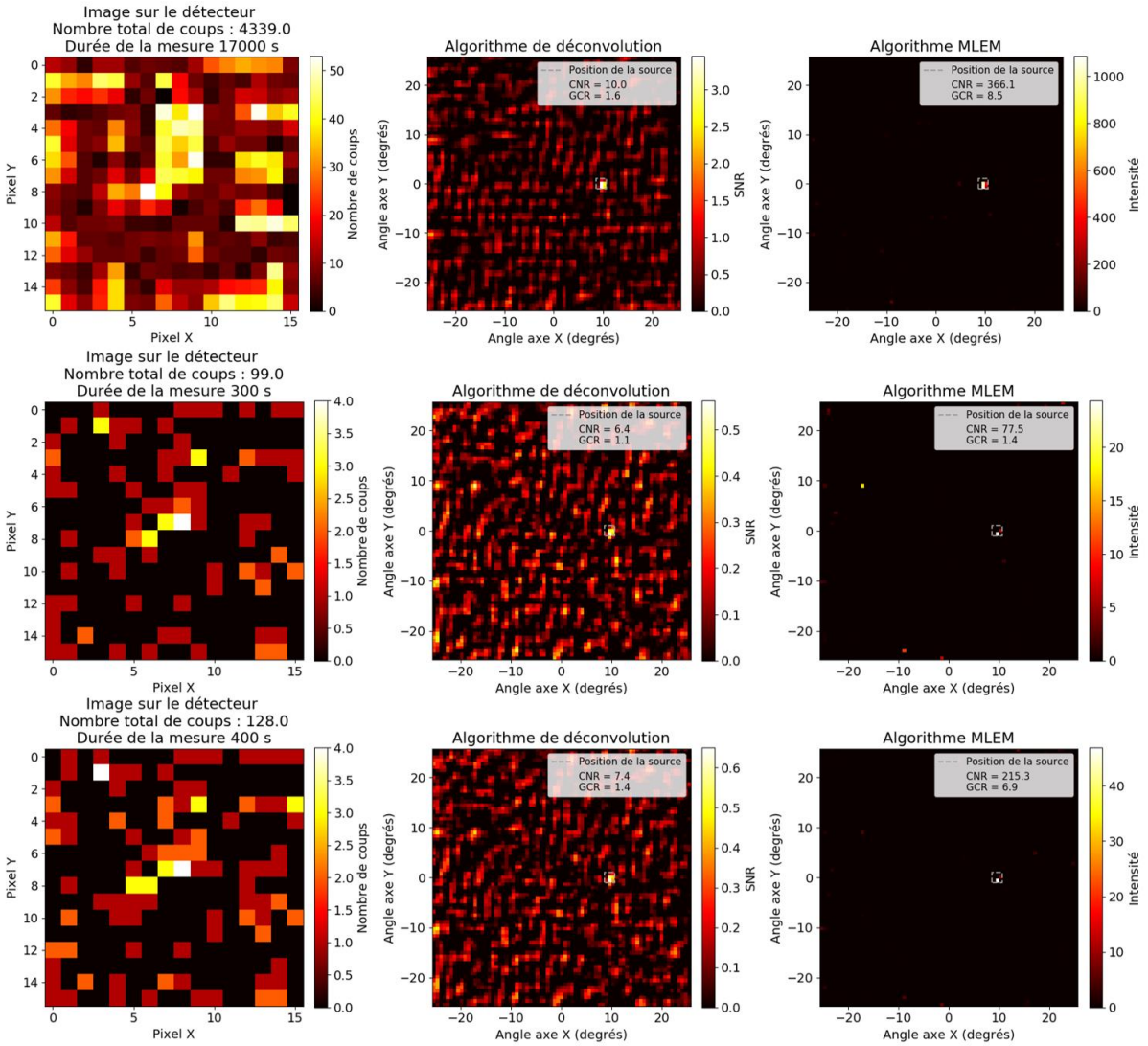


Figure 3-6 : Acquisition d'une source de ^{241}Am de 440 kBq placée à 1 m de la caméra gamma. Sélection en énergie entre 50 keV et 70 keV. Paramètres d'acquisition : temps de *peaking* 6,07 µs, gain 200 mV/fC, température du cristal -10 °C. Dans l'ordre de gauche à droite : carte de coups, résultat de l'algorithme de déconvolution et résultat de l'algorithme MLEM. De haut en bas : acquisition de 17 000 s, de 300 s et 400 s.

3.1.3 Reconstruction d'images pour des sources ponctuelles multiples

Sources multiples avec différents radioéléments

Je montre l'application des algorithmes sur une mesure prise avec deux radioéléments différents, une source de ^{137}Cs de 1,6 MBq et une source de ^{141}Am de 440 kBq, en sélectionnant les énergies respectivement entre 30 et 38 keV et entre 58,3 et 60,6 keV. La sélection est beaucoup plus fine que pour les tests précédents sur le ^{241}Am pour deux raisons. D'une part, la source de ^{137}Cs émet des photons à 662 keV qui ont la capacité de provoquer la fluorescence X du Tantale à 57,5 keV, phénomène impossible avec les photons du ^{241}Am d'énergie inférieure au *K-edge* du Tantale à 67,4 keV. D'autre part, les photons à 662 keV du ^{137}Cs créent un continuum Compton qui ajoute un bruit supplémentaire sur les données.

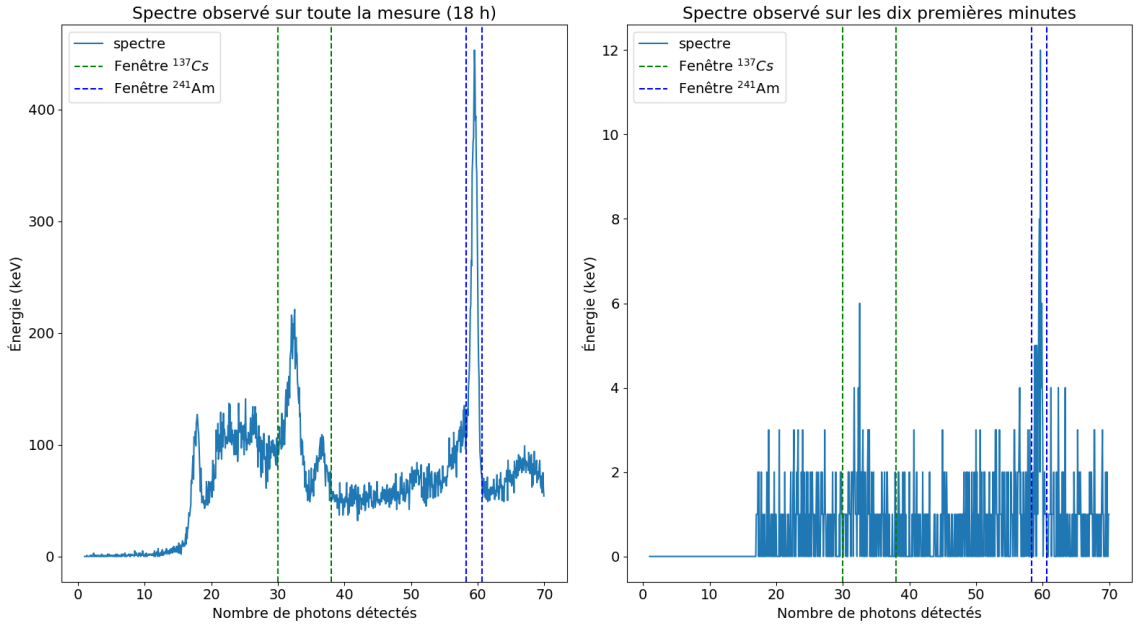


Figure 3-7 : Acquisition de deux sources de ^{241}Am de 440 kBq et de ^{137}Cs de 440 kBq 394 kBq placées à 1 m de la caméra gamma. Sélection en énergie entre 50 keV et 70 keV. Paramètres d'acquisition : temps de *peaking* 6,07 µs, gain 200 mV/fC, température du cristal -10 °C. À gauche : acquisition totale de 18 h. À droite : 10 premières minutes de l'acquisition.

La Figure 3-7 présente le spectre obtenu sur toute l'acquisition de 18 heures avec les sélections en énergie décrites ci-dessus, ainsi que le même spectre sur les 10 premières minutes d'acquisition, que je vais utiliser pour faire appliquer les algorithmes de localisation. J'effectue les reconstructions en Figure 3-8. Je ne montre que les figures avec les durées d'acquisition nécessaires pour identifier les sources et rejeter les fantômes, sans sélection en énergie dans un premier temps et avec les deux sélections en énergie ensuite.

Sans sélection en énergie, les deux sources sont localisées en 2000 secondes. On note la présence de fantômes qui restent visibles, avec une intensité supérieure à celle de la source de ^{241}Am, le GCR étant de 0,8 pour cette source.

Grâce à la sélection en énergie, il est aisé de localiser indépendamment la source de ^{137}Cs et ^{241}Am à partir des spectro-images correspondantes, en 10 minutes, avec des GCR respectifs de près de 3,9 et de 11,7. Cela montre l'intérêt de ces spectro-images : la réduction du bruit par la sélection en énergie permet de localiser plus rapidement les sources sans fantôme d'une part, et d'autre part, elle permet de les localiser indépendamment. En particulier, MLEM atteint un CNR de 0 sur les sources n'appartenant pas à la sélection en énergie.

Ainsi, la présence de plusieurs sources avec des radioéléments de natures différentes ne constitue pas un obstacle supplémentaire grâce à la sélection en énergie, possible après étalonnage du détecteur Caliste. La principale difficulté vient lorsque l'une des sources possède des photons de haute énergie créant des structures Compton d'intensité trop importante, se superposant avec les pics photoélectriques d'intérêt des autres sources.

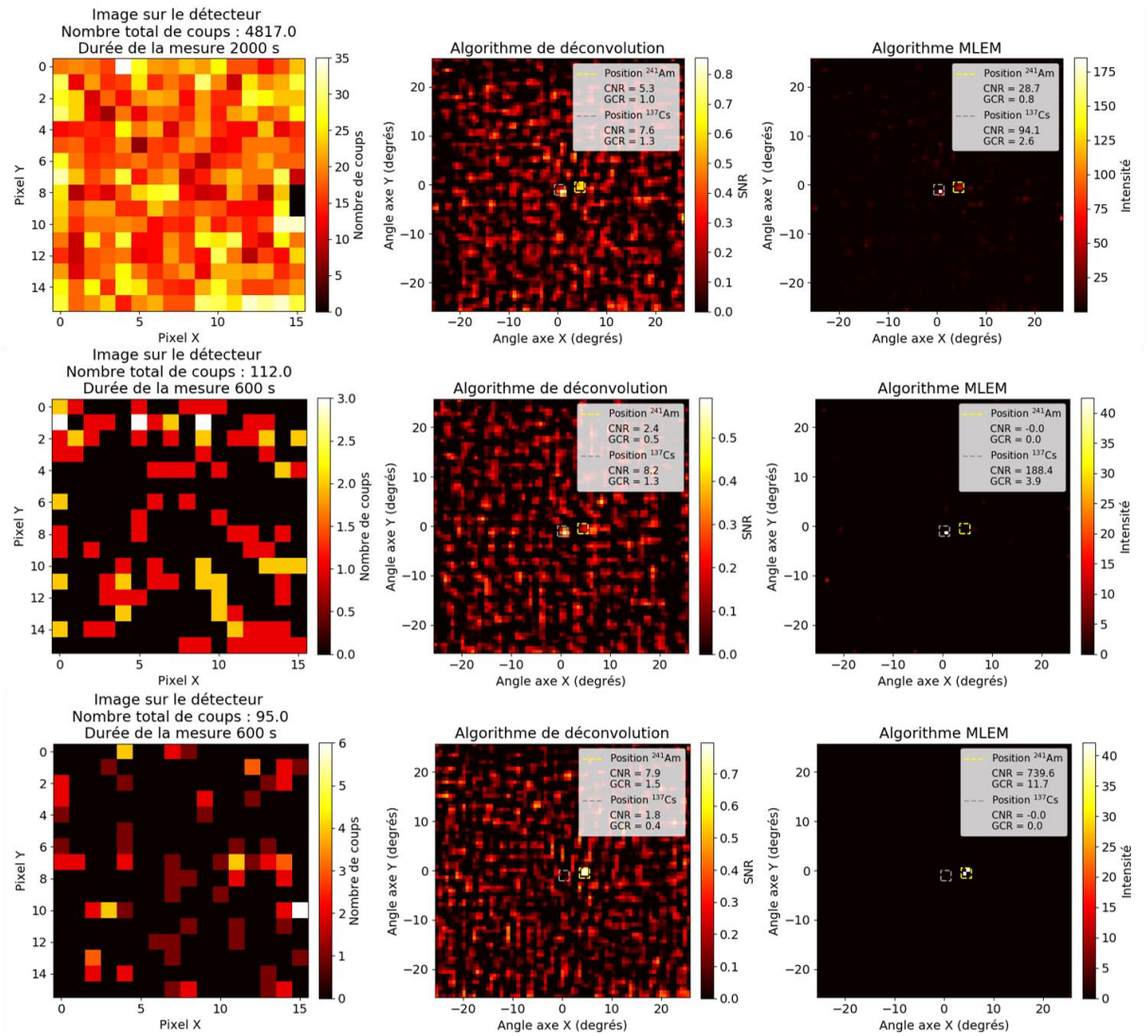


Figure 3-8 : Acquisition de deux sources de ^{241}Am de 440 kBq et de ^{137}Cs de 440 kBq 394 kBq placées à 1 m de la caméra gamma. Paramètres d'acquisition : temps de *peaking* 6,07 µs, gain 200 mV/fC, température du cristal -10 °C. Dans l'ordre de gauche à droite : carte de coups, résultat de l'algorithme de déconvolution et résultat de l'algorithme MLEM. De haut en bas : acquisition de 2000 s sans sélection en énergie, acquisition de 600 s avec sélection entre 30 et 38 keV et acquisition de 600 s avec sélection entre 58,3 et 60,6 keV.

Sources multiples d'un même radioélément

Je teste maintenant l'application des algorithmes de localisation en présence de plusieurs sources d'un même radioélément. J'utilise deux sources de ^{241}Am qui émettent donc dans la même bande d'énergie, une de 440 kBq, placée à gauche dans le champ de vue, et une de 394 kBq, placée à droite. Les résultats pour cette mesure sont montrés en Figure 3-9.

Les sources sont reconstruites aux bonnes positions avec les deux algorithmes. Un temps de pose plus long est nécessaire pour être capable de supprimer les fantômes. En 500 s d'acquisition, un fantôme apparaît toujours en bas à droite pour l'algorithme MLEM, ce qui donne un GCR pour la source de droite de 2,4. Avec l'algorithme de déconvolution, le GCR est de moins de 1,3 pour les deux sources.

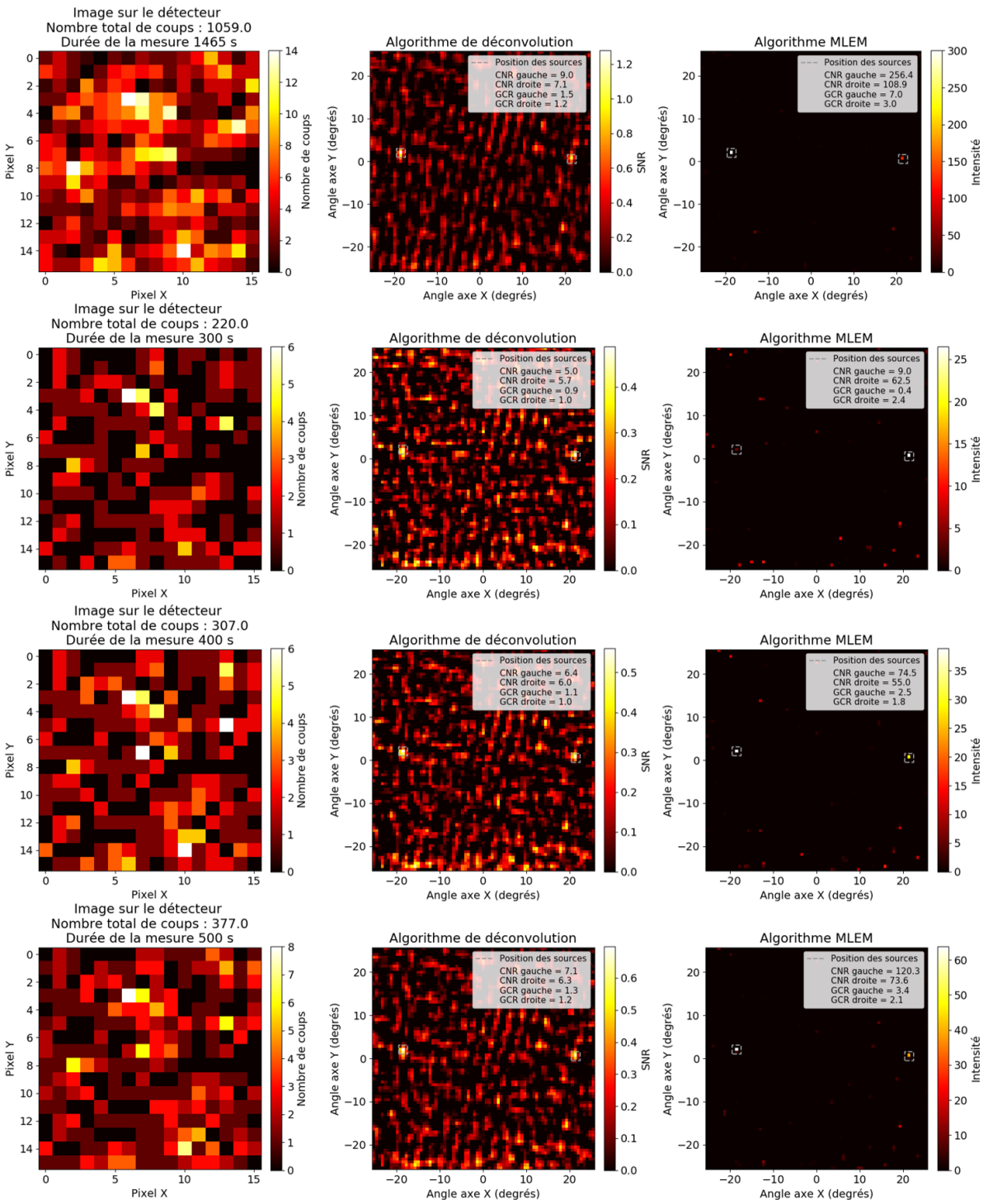


Figure 3-9 : Acquisition de deux sources de ^{241}Am de 440 kBq et 394 kBq placées à 1 m de la caméra gamma. Sélection en énergie entre 50 keV et 70 keV. Paramètres d'acquisition : temps de *peaking* 6,07 µs, gain 200 mV/fC, température du cristal -10 °C. Dans l'ordre de gauche à droite : carte de coups, résultat de l'algorithme de déconvolution et résultat de l'algorithme MLEM. De haut en bas : acquisition de 25 mn, de 300 s, 400 s et de 500 s.

Notons aussi que les intensités reconstruites par l'algorithme MLEM ne correspondent pas aux intensités attendues dont l'écart ne devrait être que de 9 % vu l'activité des sources. Sur la totalité de la mesure, la source de droite apparaît avec une intensité deux fois plus faible que celle de gauche. Les deux sources se situent dans le champ partiellement codé, au-delà de la limite de $\pm 17{,}3°$ du champ

entièrement codé, à $-18°$ à gauche et à $20°$ à droite, mais cela n'explique qu'en partie la différence d'intensité observée. En réalité, cette différence d'intensité est principalement due au fait que les photons sont redistribués par l'algorithme MLEM sur l'ensemble de la carte du ciel. De ce fait, une plus grande partie des photons de la source de droite est redistribuée sur les fantômes, par rapport à la source de gauche.

Si on fixe un seuil de détection pour chaque source à 20 % de l'intensité maximale détectée, on peut filtrer les sources et relancer l'algorithme MLEM en ne considérant que les projections associées aux deux positions identifiées. Ainsi, on empêche l'algorithme de redistribuer les photons sur l'ensemble de la carte. Puis, je peux redresser l'intensité retrouvée par la fraction d'ouverture, du fait que la source de droite se situe plus profondément dans le champ partiellement codé. Le résultat de cette méthode appliquée à la mesure complète de 25 minutes est illustré en Figure 3-10. Cette redistribution rapproche l'intensité relative de chaque source. L'intensité relative de droite par rapport à celle de gauche passe de 42 % (intensité de 300 à gauche et 127 à droite) avant redistribution à 71 % (intensité de 1553 à gauche et 1081 à droite) après ce processus, à comparer avec l'intensité relative attendue de 91 % vu les activités des sources. Par ailleurs, comme je n'utilise l'algorithme MLEM que sur les positions que j'ai déjà précédemment localisées, le reste des positions est mis à 0 d'où les CNR et GCR infinis avec cette méthode. Ce processus améliore donc l'intensité relative mais ne la retrouve pas précisément.

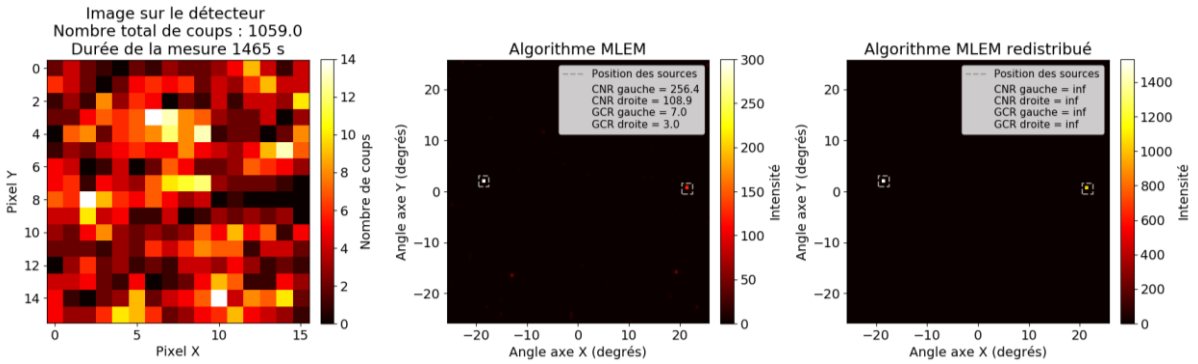


Figure 3-10 : Acquisition de deux sources de ^{241}Am de 440 kBq et 394 kBq placées à 1 m de la caméra gamma pendant près de 25 mn. Sélection en énergie entre 50 keV et 70 keV. Paramètres d'acquisition : temps de *peaking* 6,07 µs, gain 200 mV/fC, température du cristal -10 °C. Dans l'ordre de gauche à droite : carte de coups, résultat direct de l'algorithme MLEM et résultat de l'algorithme MLEM après redistribution sur les sources localisées et redressement par la fraction d'ouverture.

Je montre en Figure 3-11 un exemple avec 5 sources du même radioélément ^{241}Am. Comme, je n'ai pas accès à autant de sources en laboratoire, j'utilise une procédure d'augmentation de données, qui consiste à prendre 5 mesures individuelles de ^{241}Am que je fusionne ensuite. La source ayant une activité de 440 kBq à 1 m du détecteur, avec cette méthode, je n'introduis pas de biais qui pourrait être lié à un éventuel pile-up du fait d'un flux trop important de photons, le taux de comptage est inférieur à 10 coups/s. Je montre les résultats pour le temps total d'acquisition de près de 4 h 45 mn.

On note que les deux algorithmes ont beaucoup plus de difficulté à retrouver les sources, notamment la source au centre du champ de vue et la source la plus à droite avec des GCR inférieurs à 1, alors que toutes les sources ont la même activité et sont placées à la même distance du détecteur. L'algorithme de déconvolution fait apparaître un ensemble de fantômes en bas à droite du champ de vue, et l'algorithme MLEM fait apparaître un fantôme en-dessous des sources. On voit ici que le cas avec une

multitude de sources devient plus complexe à résoudre, ce qui fait apparaître les premières limitations de ces algorithmes associés au système d'imagerie de la caméra gamma.

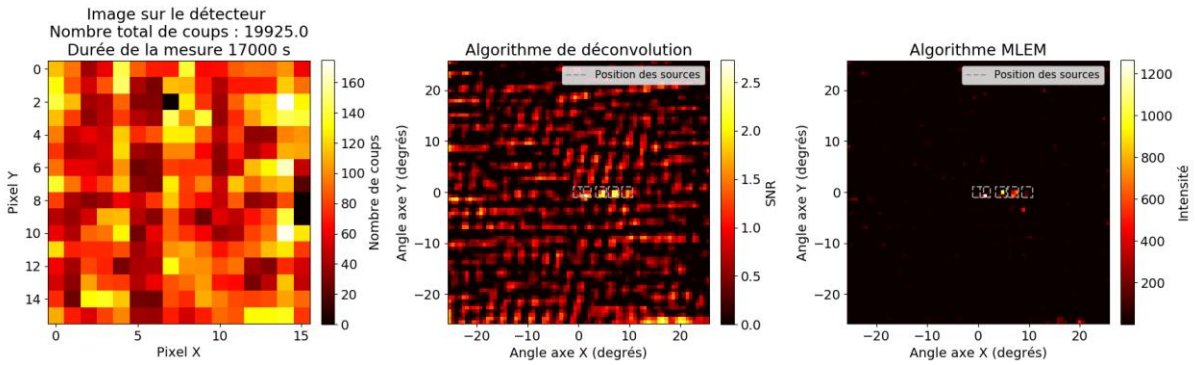


Figure 3-11 : Acquisition de cinq sources de ^{241}Am de 440 kBq placées à 1 m de la caméra gamma pendant 4 h 45 mn. Sélection en énergie entre 50 keV et 70 keV. Paramètres d'acquisition : temps de *peaking* 6,07 µs, gain 200 mV/fC, température du cristal -10 °C. Dans l'ordre de gauche à droite : carte de coups, résultat de l'algorithme de déconvolution et résultat de l'algorithme MLEM.

Position		(0°,0°)	(1.4°,0°)	(4.5°,0°)	(6.9°,0°)	(9.6°,0°)
Algorithme	Valeur					
Déconvolution	CNR	4,4	7	6,4	5,9	4,1
	GCR	0,8	1,1	1,1	1	0,7
MLEM	CNR	**13,4**	**90,2**	**72,6**	**47,6**	**10,5**
	GCR	0,4	**2,6**	**2,1**	1,4	0,3

Tableau 3-2 : Valeurs de CNR et GCR correspondant à la Figure 3-11. Les valeurs en gras identifient les CNR supérieurs à 10 et GCR supérieurs à 2 (voir Tableau 3-1).

3.2 Sources étendues

3.2.1 Simulations de sources

Connaissant les projections $M = (M_{ij})_{ij}$ du masque codé pour une source à la position j sur le pixel i, je peux simuler des images de détecteur Y à partir de sources dans le ciel X. Pour ce faire, j'utilise l'équation (1-4) de la manière suivante.

Je note X le vecteur qui représente les positions des sources dans le ciel, avec leurs intensités. X peut être fixé ou issu d'un tirage aléatoire. Je note n une constante qui représente un niveau de bruit uniforme à appliquer sur le détecteur, n peut de même être fixé volontairement ou tiré aléatoirement. Avec ces notations, je calcule Y_{th} de la manière suivante :

$$Y_{\text{th}} = M.X + \frac{n}{N_{\text{pixels}}} \tag{3-3}$$

Y_{th} représente la projection théorique des sources représentées par le vecteur X sur le détecteur, associée à un bruit uniforme n réparti sur les $N_{\text{pixels}} = 256$ pixels du détecteur. Je normalise ensuite Y_{th} en norme 1 pour obtenir Y_{thn} :

$$Y_{\text{thn}} = \frac{Y_{\text{th}}}{\|Y_{\text{th}}\|_1} = \frac{Y_{\text{th}}}{\sum_i Y_{\text{th}_i}} \tag{3-4}$$

Y_{thn_i} représente la probabilité qu'un photon interagissant dans le détecteur soit détecté dans le pixel i.

Maintenant, pour avoir une simulation qui prend en compte les fluctuations statistiques, je fixe un nombre N de photons détectés. Le nombre moyen de photons attendus dans le pixel i du détecteur est alors $Y_{\mathrm{thn}_i} \times N$. Le nombre de photons Y_i effectivement détectés dans ce pixel i suit une loi de Poisson dont le paramètre est donné par le nombre moyen de photons attendus :

$$Y_i \sim Poisson\left(Y_{\mathrm{thn}_i} \times N\right) \tag{3-5}$$

Enfin, je simule les pixels potentiellement bruyants, en éteignant un certain nombre N_{bruyants} de pixels, que je tire ensuite aléatoirement sur le détecteur.

Cette méthode permet d'obtenir rapidement des simulations de source puisqu'elle ne se base que sur des tirages directs. Une méthode de simulation Monte-Carlo de type Geant4 comme décrite en Chapitre 2 serait plus fidèle à la réalité mais son temps d'exécution dépendrait du nombre de photons à simuler, alors que le temps d'exécution de la méthode que j'utilise est constant : il ne dépend ni du nombre de sources, ni du nombre de photons à simuler.

Pour simuler des sources étendues et tester les capacités des algorithmes de déconvolution et MLEM, cette simulation sera suffisante, comme je le montre par la suite. Par ailleurs, cette technique de simulation rapide sera aussi importante en Chapitre 5 pour simuler de manière continue une base de données d'exemples pour l'apprentissage des réseaux de neurones.

3.2.2 Localisation de sources étendues issues de simulations

Dans ce paragraphe, j'étudie notamment l'influence de la taille des sources étendues sur les capacités de localisation. Les sources que je simule sont carrées et uniformes en intensité, le but étant de montrer les limites liées à la taille des sources. Je simulerai des sources aux formes plus complexes en Chapitre 5 lorsque j'appliquerai des méthodes de *deep learning* pour la localisation. Je tire un nombre aléatoire de pixels bruyants N_{bruyants} dans une loi de Poisson de paramètre 0,5.

Quantification des résultats

Pour évaluer les performances des algorithmes, je vais m'intéresser à leur capacité à imager correctement l'ensemble de la source étendue, sans introduire de bruit en dehors de la source attendue.

Je définis pour ce faire deux valeurs que j'appelle inROI (*Region of Interest*) et outROI de la manière suivante :

$$\mathrm{inROI} = \frac{V_{\mathrm{inROI}}}{S_{\mathrm{ROI}}} \tag{3-6}$$

$$\mathrm{outROI} = \frac{V_{\mathrm{outROI}}}{S_{\mathrm{outROI}}} \tag{3-7}$$

où V_{inROI} est le nombre de pixels du ciel donné par la reconstruction qui se situent à l'intérieur de la zone correspondant à la source étendue, pour lesquels la valeur renvoyée par l'algorithme est supérieure à un certain seuil. J'ai pris ce seuil comme étant égal à 10 % de la valeur maximale des pixels

de la reconstruction. L'optimisation de ce seuil peut faire l'objet d'une étude à part entière. S_{ROI} est le nombre de pixels du ciel effectivement couverts par la source étendue.

De manière analogue, V_{outROI} est le nombre de pixels du ciel au-dessus du seuil défini précédemment se situant en-dehors de la zone correspondant à la source étendue et S_{outROI} est le nombre de pixels du ciel de la zone non couverte par la source étendue.

inROI représente la part de la zone couverte par la zone étendue qui est effectivement reconstruite par l'algorithme. Si inROI vaut 1, cela signifie que la source est entièrement reconstruite. Inversement, outROI représente la part de la zone en-dehors de la source qui est reconstruite à tort par l'algorithme. Si outROI vaut 0, cela signifie qu'aucune zone en dehors de la source étendue n'est reconstruite. La zone en dehors de la source étendue est en général très grande : les sources étendues que je simule ne couvrent que moins de 10 % du champ de vue. Il faut mettre outROI en perspective avec la zone effectivement couverte par la source en le comparant à ROI/total que je définis de la manière suivante :

$$\mathrm{ROI/total} = \frac{S_{\mathrm{ROI}}}{S_{\mathrm{ROI}} + S_{\mathrm{outROI}}} \tag{3-8}$$

Cette grandeur représente la part du champ de vue couverte par la source.

Tests sur des sources étendues

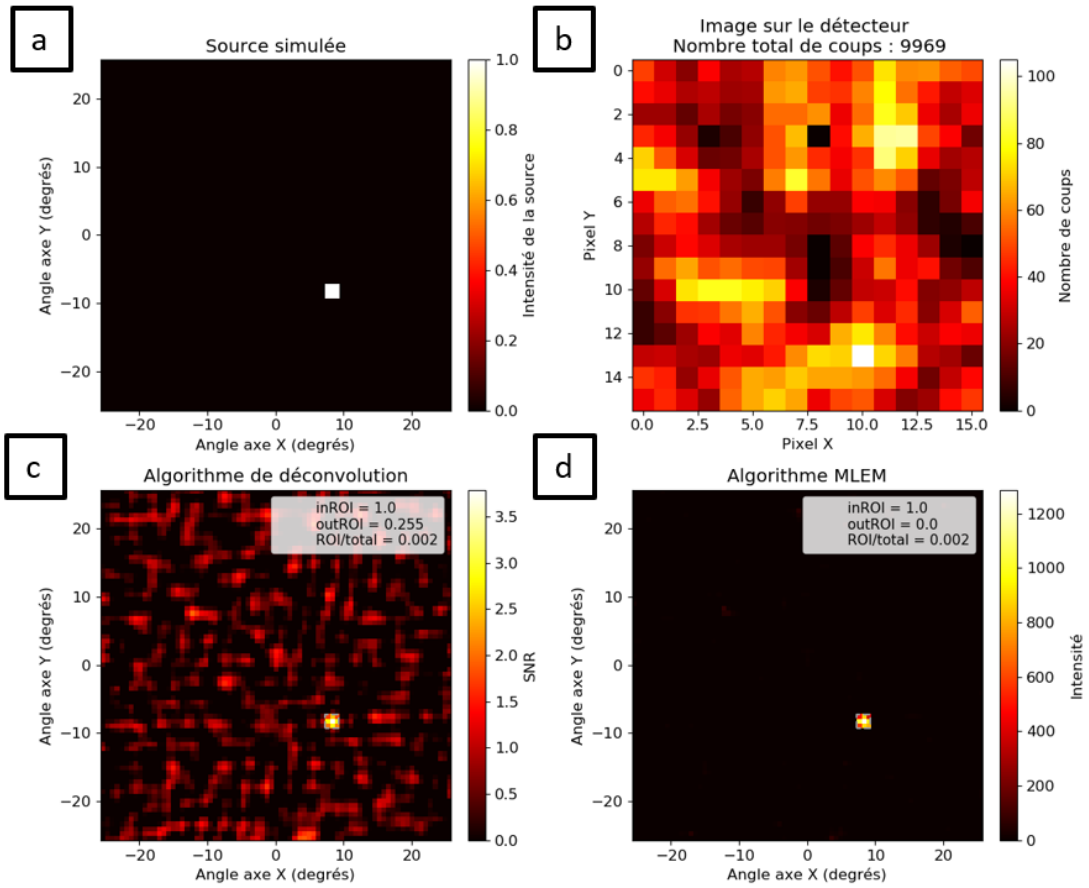


Figure 3-12 : Résultat de simulation d'une source carrée 2°x2° uniforme avec 10 000 photons, sans bruit. a : Source simulée. b : Résultat de la simulation sur le détecteur. c : Résultat de l'algorithme de déconvolution. d : Résultat de MLEM.

La première simulation est une source de 2°x2°, simulée avec 10 000* photons sans bruit. Le résultat est présenté en Figure 3-12. Les deux algorithmes localisent la source au bon endroit et la taille et la forme sont relativement fidèlement retrouvées. L'intensité n'apparaît pas uniforme dans les deux reconstructions, le centre de la source apparaît plus brillant que les coins. Avec MLEM, la source est reconstruite dans son ensemble avec un inROI de 1, sans aucune zone à l'extérieur avec un outROI de 0.

En Figure 3-13, je représente le résultat avec 10 000 photons et un bruit uniforme d'intensité $n = 3 \times \sum_j X_j$, c'est-à-dire un bruit trois fois plus intense que la source. Les algorithmes localisent de nouveau les sources, avec un inROI de 1. Cependant, la présence du bruit fait apparaître plus de fantômes. MLEM a un outROI de 2,4 % soit dix fois plus important que la part de la zone couverte par la source qui est de 0,2 %. Ces exemples montrent tout de même la capacité des algorithmes à traiter le cas de sources légèrement étendues.

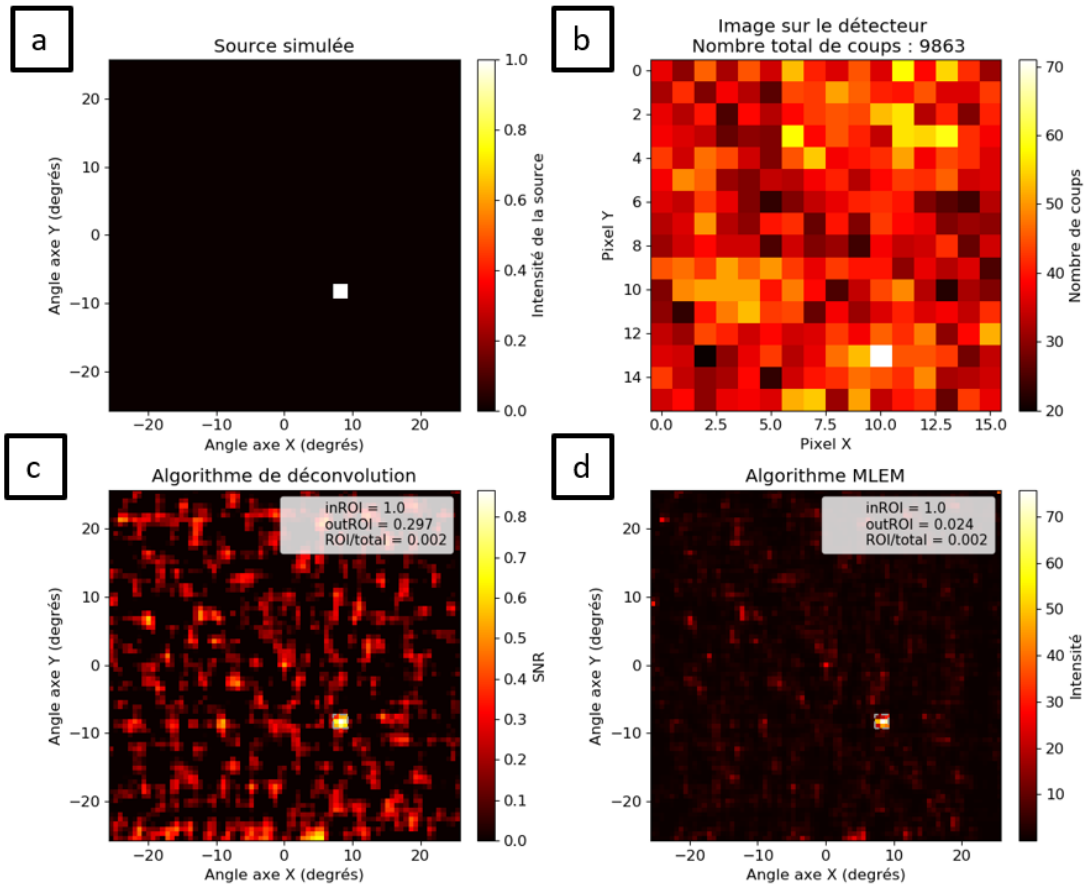


Figure 3-13 : Résultat de simulation d'une source carrée 2°x2° uniforme avec 10 000 photons, avec un bruit trois fois plus intense que la source. a : Source simulée. b : Résultat de la simulation sur le détecteur. c : Résultat de l'algorithme de déconvolution. d : Résultat de MLEM.

Je simule maintenant une source de taille 4°x4° avec 100 millions de photons et sans bruit. Pour illustrer, une taille angulaire de 4° représente environ 7 cm à 1 m. Le résultat est présenté en Figure 3-14. On constate que même avec ce nombre très élevé de photons et sans aucun bruit, les deux

* Le nombre de photons finalement présents sur le détecteur n'est pas de 10 000 du fait que je tire le nombre de photons de manière indépendante sur chaque pixel, donc la somme de tous les photons tirés subit aussi des fluctuations statistiques.

algorithmes ont plus de difficulté à retrouver les sources. L'algorithme de déconvolution reconstruit 94 % de la source mais près du tiers du reste de l'image est reconstruite à tort. Inversement, MLEM reconstruit la moitié de la source avec un inROI de 0,5, et fait apparaître des fantômes avec un outROI de 0,7 %, comparable à la part du champ de vue occupé par la source qui est de 0,6 %. Bien que le maximum d'intensité soit retrouvé au bon endroit, on constate la présence de beaucoup plus de fantômes et la source étendue n'est pas reconstruite dans son ensemble.

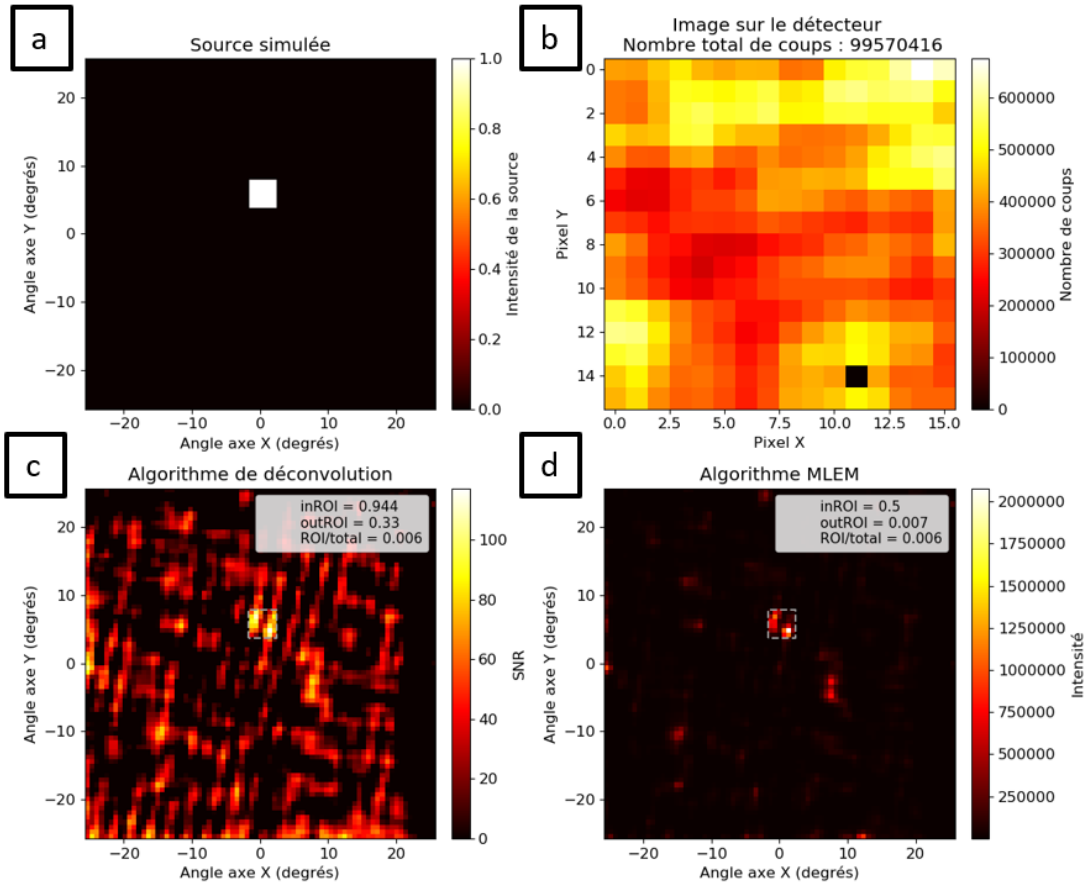


Figure 3-14 : Résultat de simulation d'une source carrée 4°x4° uniforme avec 100 millions de photons, sans bruit. a : Source simulée. b : Résultat de la simulation sur le détecteur. c : Résultat de l'algorithme de déconvolution. d : Résultat de MLEM.

En augmentant la taille de la source, les capacités de localisation se dégradent encore comme montré en Figure 3-15 avec une source de taille 7°x7°. Toujours simulée avec un nombre très élevé de 100 millions de photons et sans bruit, la source n'est plus du tout localisée par aucun des deux algorithmes. L'algorithme de déconvolution reconstruit à tort plus de 30 % du champ de vue et l'algorithme MLEM reconstruit à tort une zone sept fois plus grande que la vraie source, tout en ne reconstruisant correctement que 32 % de la source.

Sur ce résultat, on voit clairement les limites des algorithmes de reconstruction pour des sources étendues et leur incapacité à donner une réponse, même partielle, lorsque la taille de la source est trop importante.

Cette incapacité à reconstruire des sources trop étendues est liée à deux facteurs. D'une part, dans le problème à résoudre, j'essaye de reconstruire une carte du ciel X de taille 75x75 positions à partir

d'une carte des coups Y de 16x16 pixels, ce qui implique une non injectivité du problème. Il peut exister plusieurs cartes du ciel X pour générer une même projection Y sur le détecteur. D'autre part, le deuxième facteur est directement lié au premier par le fait que les algorithmes que j'ai mis en place n'ont pas de méthode de régularisation qui privilégie les sources étendues par rapport à une multitude de sources ponctuelles dans le champ de vue du détecteur, ce qui produit alors des images comme en Figure 3-15.

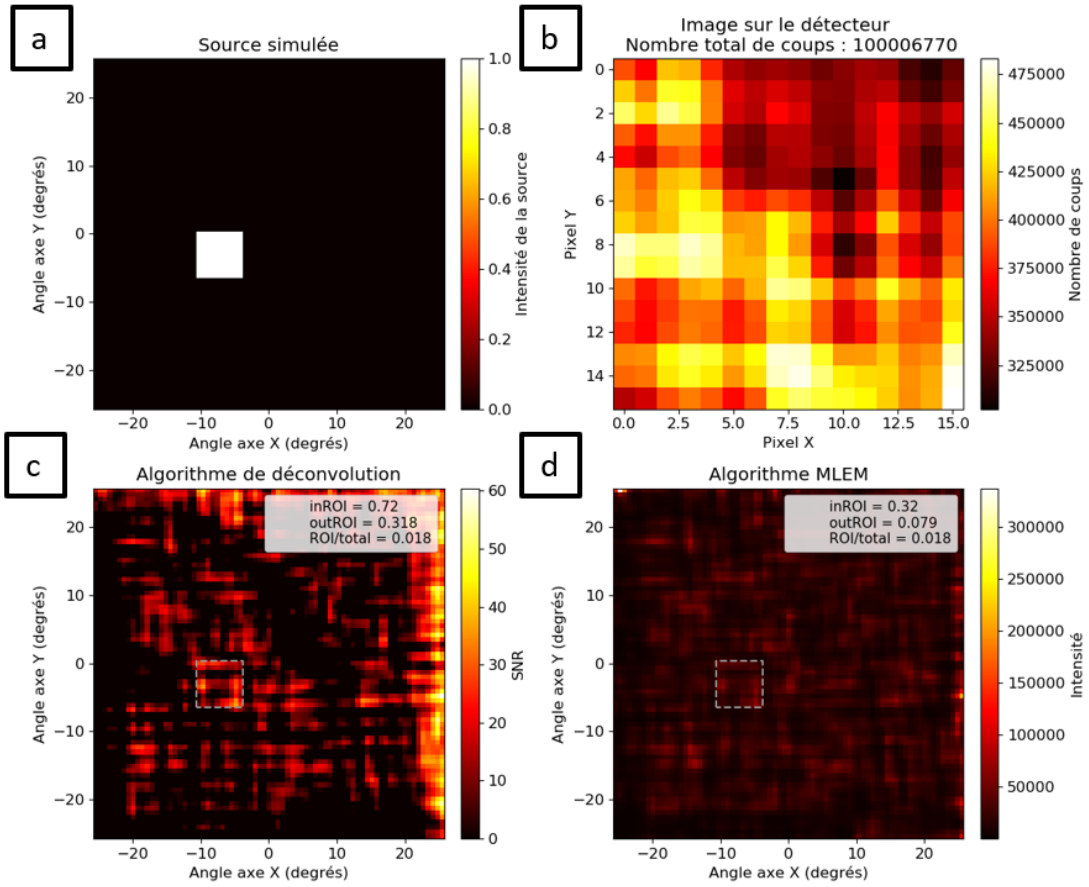


Figure 3-15 : Résultat de simulation d'une source carrée 7°x7° uniforme avec 100 millions de photons, sans bruit. a : Source simulée. b : Résultat de la simulation sur le détecteur. c : Résultat de l'algorithme de déconvolution. d : Résultat de MLEM.

Afin d'illustrer la non injectivité mentionnée plus haut, je reprends l'exemple de la simulation de la Figure 3-15 (7°x7°). Avec l'équation (3-6), je calcule la projection théorique $Y_{\text{th}} = M.X$ associée à la source simulée. Ensuite, à partir de Y_{th}, j'essaye de retrouver X en minimisant la fonction suivante sous la contrainte que tous les éléments de X doivent être positifs :

$$L(X) = \|Y_{\text{th}} - M.X\|_2^2 \tag{3-9}$$

J'effectue cette minimisation par un algorithme classique de descente de gradient, à partir d'une carte du ciel initiale X_0 uniforme, dont tous les éléments valent $\frac{1}{N_{\text{sky}}}$ où N_{sky} est le nombre de positions sur le ciel, soit 75x75 positions.

La Figure 3-16 montre en haut à gauche la source simulée, qui correspond à celle de la Figure 3-15. En haut à droite, je montre la projection directe sur le détecteur Y_{th}. En bas à gauche, je montre le résultat

de la minimisation X_{opt}. Enfin, en bas à droite, je montre la projection qui correspond à X_{opt}, c'est-à-dire, $Y_{\text{opt}} = M.Y_{\text{opt}}$. Je calcule aussi la différence relative Diff entre la projection trouvée Y_{opt} et la projection réelle Y_{th} :

$$\text{Diff} = \frac{\left\|Y_{\text{th}} - Y_{\text{opt}}\right\|_2}{\|Y_{\text{th}}\|_2} \tag{3-10}$$

On constate notamment que la différence relative entre la projection attendue et celle trouvée avec la carte du ciel X_{opt} est extrêmement faible, de l'ordre de 10^{-15}. En revanche, la carte du ciel trouvée X_{opt} est clairement différente de la carte simulée X. Cela montre que l'on peut obtenir une projection extrêmement proche de Y_{th} avec une carte du ciel complètement différente de celle qui l'a générée.

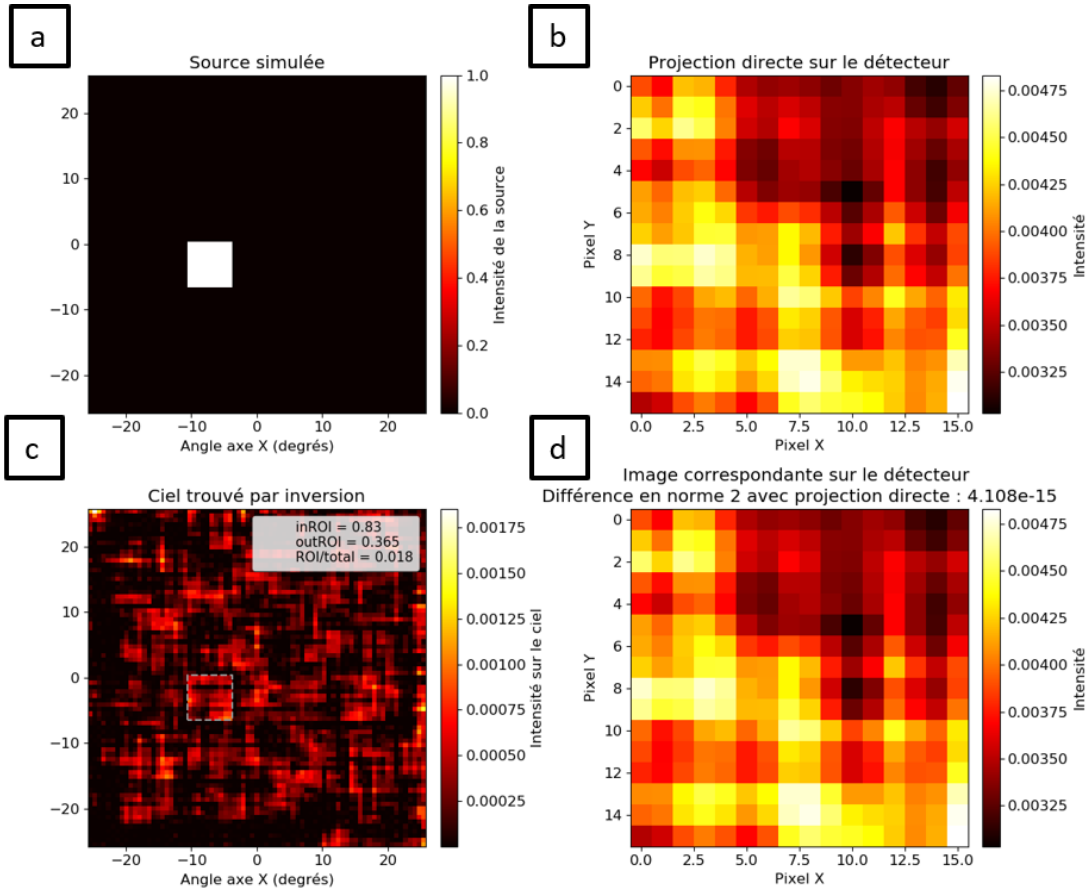


Figure 3-16 : Inversion de la projection directe. a : Ciel simulé. b : Projection directe correspondante sur le détecteur. c : Carte du ciel issue de l'inversion directe de la projection. d : Projection associée à la carte issue de l'inversion.

La conséquence de cette observation constitue un facteur d'échec des algorithmes de déconvolution et MLEM. Pour résoudre ce problème, on peut envisager d'utiliser un algorithme MAP-EM qui donne une probabilité plus grande a priori aux cartes du ciel constituées d'un minimum de sources étendues, par rapport à une multitude de sources ponctuelles. Cependant, il est nécessaire de formaliser ce type de *prior* et d'adapter l'algorithme de recherche du maximum de la probabilité a posteriori.

En Chapitre 5, je montre une méthode basée sur des algorithmes de *deep learning* pour effectuer ces inversions et je montre qu'il est possible, à partir de la construction de la base de données, de rechercher ce type de sources étendues.

4 Approfondissements sur l'algorithme MLEM

4.1 Détermination d'un critère d'arrêt objectif

L'un des principaux inconvénients de cet algorithme concerne l'absence de détermination d'un critère d'arrêt des itérations, autre que par l'analyse d'un expert.

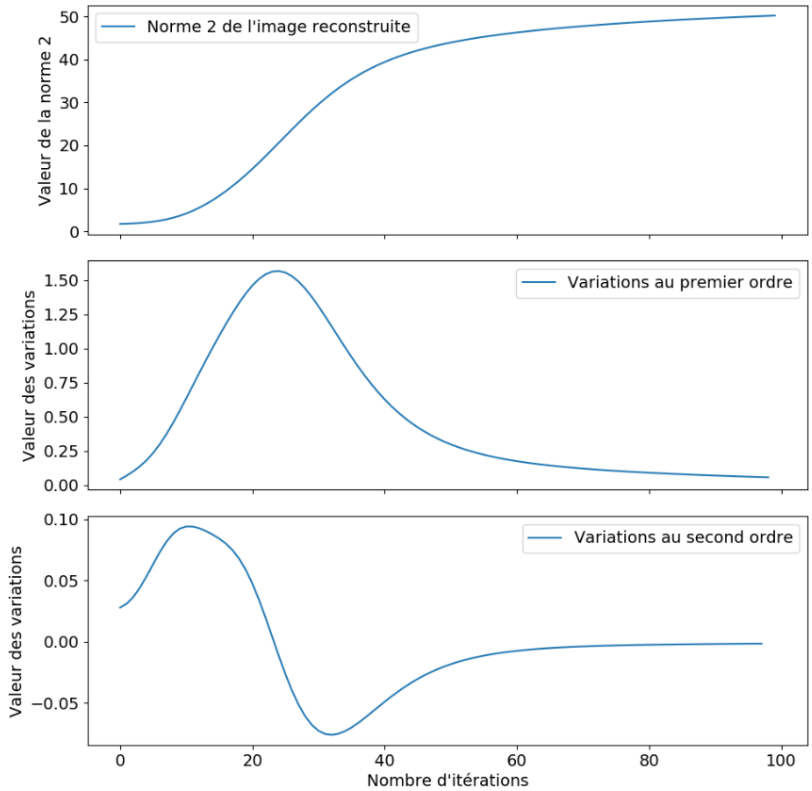


Figure 4-1 : Norme 2 de la reconstruction $\widehat{X}^{(n)}$ et ses variations pour chaque itération n de l'algorithme MLEM. L'exemple utilisé est celui présenté en Figure 3-6 avec 400 s d'acquisition.

Dans ce cadre, j'ai voulu étudier la possibilité d'introduire un critère d'arrêt basé sur la méthode suivante. Au cours des itérations de MLEM, j'ai étudié l'évolution de la norme 2* de l'image $\hat{X}^{(n)}$ obtenue par cet algorithme : $\left\|\hat{X}^{(n)}\right\|_2$. La motivation d'étudier le comportement de cette norme provient du fait qu'elle n'est a priori pas constante au fur et à mesure des itérations de l'algorithme. D'autre part, plus l'intensité des points de $\hat{X}^{(n)}$ est concentrée dans un nombre réduit de positions du ciel, plus cette norme $\left\|\hat{X}^{(n)}\right\|_2$ doit être élevée.

* On peut penser utiliser la norme 1, sauf qu'en réalité l'évolution de cette norme n'apporte pas d'information supplémentaire : par construction de l'algorithme MLEM, $\left\|\hat{X}^{(n)}\right\|_1$ est en fait égal au nombre total de coups enregistrés sur le détecteur dès la première itération de l'algorithme.

La Figure 4-1 montre des exemples des itérations de MLEM, avec l'évolution de la norme $\left\|\hat{X}^{(n)}\right\|_2$ ainsi que ses variations au cours des différentes itérations successives au premier ordre : $\Delta\left\|\hat{X}^{(n)}\right\|_2 = \left\|\hat{X}^{(n+1)}\right\|_2 - \left\|\hat{X}^{(n)}\right\|_2$, et au deuxième ordre : $\Delta^2\left\|\hat{X}^{(n)}\right\|_2 = \Delta\left\|\hat{X}^{(n+1)}\right\|_2 - \Delta\left\|\hat{X}^{(n)}\right\|_2$.

Sur cette figure, les formes particulières des variations de la norme 2 de $\hat{X}^{(n)}$ au premier et au second ordre montrent que l'on peut envisager d'établir un critère d'arrêt à partir de ces variations. On peut prendre par exemple l'itération pour laquelle la variation au premier ordre est maximale ou encore lorsque la variation au second ordre est minimale. Cette piste est intéressante à étudier plus avant pour établir un critère d'arrêt objectif sur l'algorithme MLEM, dont l'absence reste un inconvénient majeur de la méthode.

4.2 Détermination de la distance à la source

Les reconstructions que j'ai faites dans la partie précédente étaient des reconstructions angulaires de la direction de la source, en deux dimensions. Cependant, dans le cadre d'une caractérisation complète d'une scène radiologique, l'information de la distance entre la caméra et la source est très intéressante.

L'imagerie à masque codé peut retrouver cette information dans certaines limites, avec l'application de l'algorithme MLEM. En effet, j'ai montré en partie 1 l'influence de la distance à la source sur le grossissement induit sur la projection du masque. J'utilise ici cet effet de grossissement qui contient alors une information sur la distance à la source.

Pour ce faire, dans le problème inverse que j'ai défini par l'équation (1-4), la matrice M des projections du masque que j'avais définie était effectuée à une distance unique. Maintenant, je fais en sorte que cette matrice contienne des projections à différentes distances.

Pour se représenter cette matrice M, je la note en trois dimensions : M_{dij}, où i représente la position du ciel qui est projetée, j le pixel du détecteur sur lequel est effectuée la projection et d la distance à la source depuis laquelle est faite la projection. Il faut discrétiser l'espace sur la distance d et j'ai choisi d'effectuer une discrétisation de 10 cm à 1 m par pas de 10 cm. En effet, d'après la Figure 1-9, dans cette zone, le grossissement dépend très fortement de la distance.

Ensuite, j'applique l'algorithme MLEM de la manière que j'ai décrite dans la partie 2, à la différence près que, en sortie, j'utilise les différentes « couches » correspondant aux différentes distances de sources utilisées pour la projection. Je regarde l'intensité maximale renvoyée par l'algorithme MLEM pour chaque distance de projection.

J'applique cette méthode sur des exemples de données réelles avec une source de 440 kBq placée à différentes distances de la caméra.

En Figure 4-2, j'ai placé la source à 30 cm du détecteur. L'intensité maximale MLEM est effectivement donnée dans le cas où j'utilise la projection à 30 cm, avec une valeur en intensité de près de 600 coups attribués à cette position pour la projection à 30 cm. En comparaison, la projection à 20 cm a une intensité à peine supérieure à 500 coups. Cela montre que la distance est reconstruite avec cette méthode.

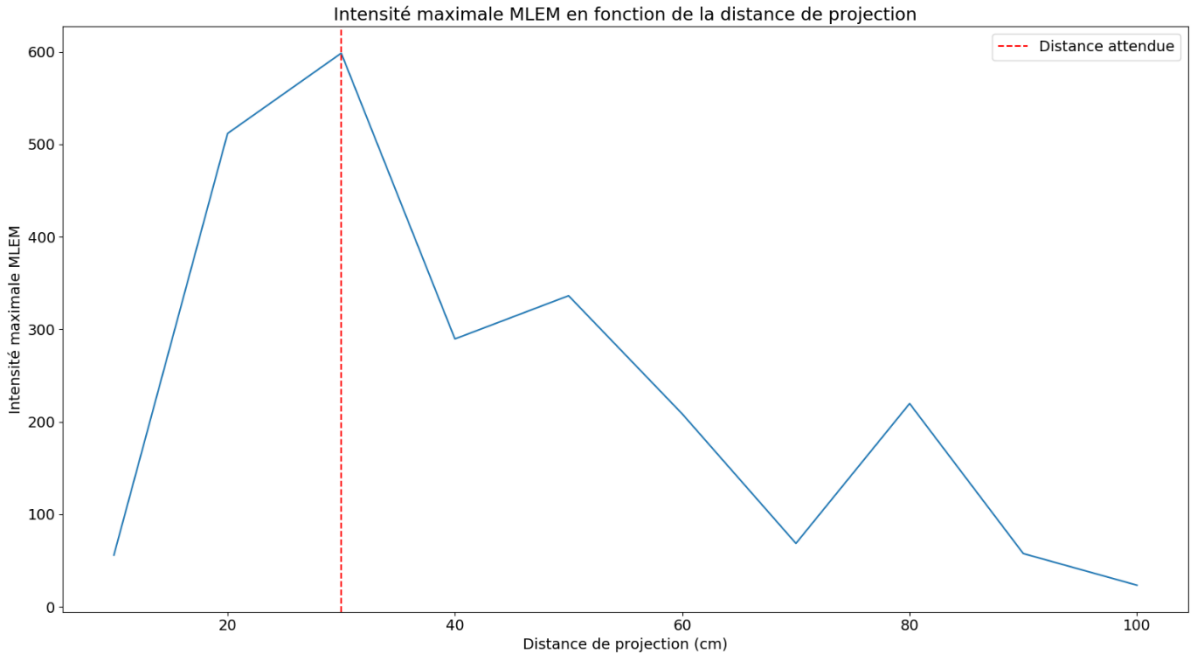


Figure 4-2 : Intensité maximale renvoyée par MLEM en fonction de la distance de projection. Acquisition d'une source de ^{241}Am de 440 kBq placée à 30 cm de la caméra gamma pendant 10 mn, nombre de photons au total : 4476 photons. Sélection en énergie entre 50 keV et 70 keV. Paramètres d'acquisition : temps de *peaking* 6,07 µs, gain 200 mV/fC, température du cristal -10 °C.

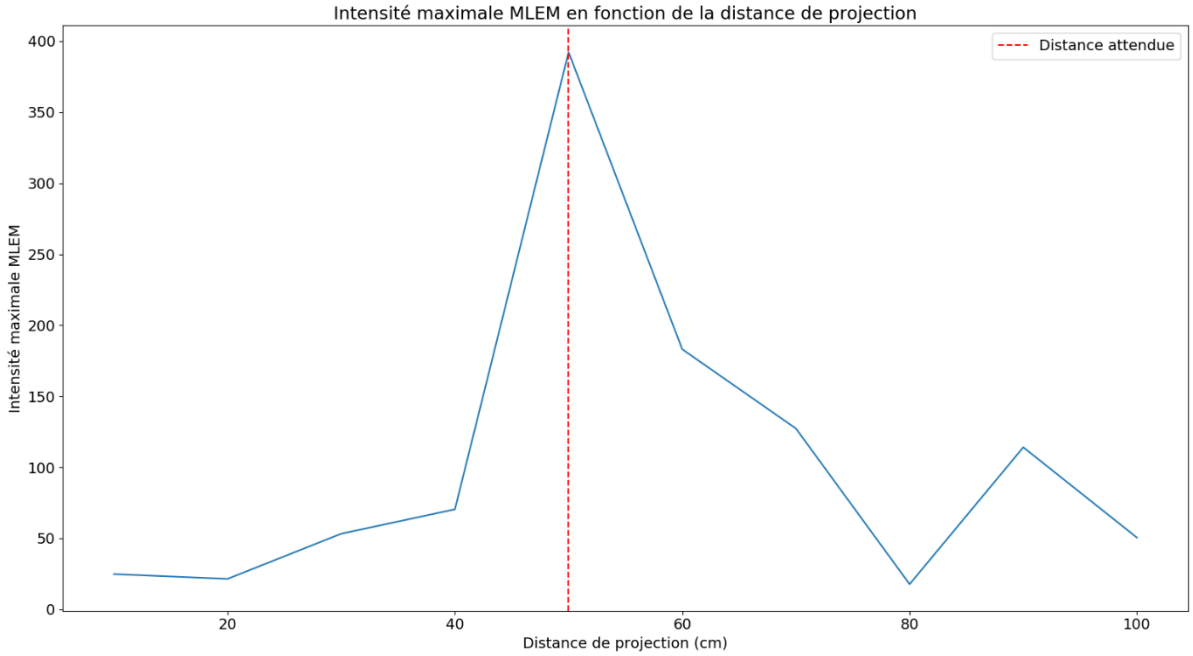


Figure 4-3 : Intensité maximale renvoyée par MLEM en fonction de la distance de projection. Acquisition d'une source de ^{241}Am de 440 kBq placée à 50 cm de la caméra gamma pendant 10 mn, nombre de photons au total : 1716 photons. Sélection en énergie entre 50 keV et 70 keV. Paramètres d'acquisition : temps de *peaking* 6,07 µs, gain 200 mV/fC, température du cristal -10 °C.

En Figure 4-3, j'ai réalisé la même expérience en plaçant la source à 50 cm, et de nouveau l'intensité maximale reconstruite par MLEM est donnée lorsque j'utilise la projection du masque à 50 cm. Elle est

près de deux fois plus grande que pour la projection à 60 cm, qui arrive en deuxième position. La distance est bien estimée par cette méthode.

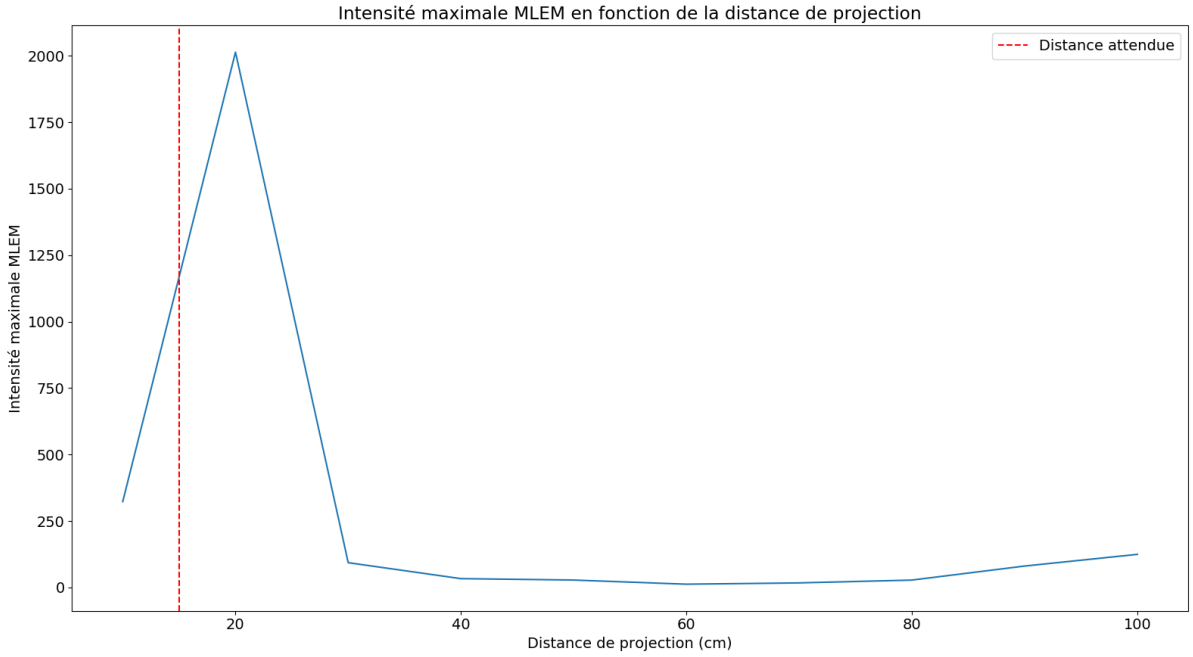


Figure 4-4 : Intensité maximale renvoyée par MLEM en fonction de la distance de projection. Acquisition d'une source de ^{241}Am de 440 kBq placée à 15 cm de la caméra gamma pendant 510 s, nombre de photons au total : 9398 photons. Sélection en énergie entre 50 keV et 70 keV. Paramètres d'acquisition : temps de *peaking* 6,07 µs, gain 200 mV/fC, température du cristal -10 °C.

Enfin, en Figure 4-4, j'ai placé la source à 15 cm de la caméra gamma, distance qui ne fait pas partie des projections que j'utilise. L'intensité maximale est alors reconstruite à 20 cm, suivie par la projection à 10 cm. Cette méthode est potentiellement utilisable même si la source se situe à des distances qui ne sont pas prises en compte dans les projections. Ici, en discrétisant tous les 10 cm, on a une incertitude de 5 cm sur la position de la source.

On pourrait envisager de faire une pondération de la distance en fonction des intensités retrouvées. Cependant, comme on le voit en Figure 4-2, l'intensité est très partagée entre 20 cm et 30 cm et une pondération donnerait une distance d'un peu plus de 25 cm alors que j'ai positionné la source à 30 cm. Inversement, en Figure 4-4, l'intensité est clairement dominante à 20 cm, alors que la source est positionnée à 15 cm, soit entre 10 cm et 20 cm.

En fonction des applications, la discrétisation peut être adaptée, mais une discrétisation plus fine nécessite un nombre de photons plus élevé de façon à avoir suffisamment d'information pour discriminer les différentes projections.

Par ailleurs, cette méthode est limitée par le système de détection et de l'effet de grossissement attendu. D'après la Figure 1-9, l'effet de grossissement est de moins en moins sensible avec la distance, et notamment, il sera plus complexe d'utiliser cet effet au-delà de 1 m, distance à partir de laquelle l'effet de grossissement est inférieur à 7 %. Cette méthode apporte donc une information de distance principalement lorsque celle-ci reste suffisamment faible pour que l'effet de grossissement soit significatif. En revanche, une modification du système d'imagerie à masque codé pourrait exploiter

cette méthode, par exemple, en augmentant la distance entre le masque et le détecteur, conformément à l'équation (1-5), au détriment du champ de vue cependant. Une autre possibilité pour pouvoir évaluer des distances plus importantes serait d'utiliser l'acquisition de deux caméras à deux positions différentes et d'effectuer une reconstruction gamma stéréoscopique, telle que décrite par (Paradiso, et al., 2018).

Dans ce paragraphe, j'ai montré la possibilité d'accéder à la distance entre la source et la caméra en utilisant l'effet de grossissement, en testant cette méthode à l'aide de données réelles pour une unique source ponctuelle. Il sera intéressant par la suite de conduire des tests avec plusieurs sources ponctuelles, situées à des distances différentes de la caméra pour pouvoir caractériser l'intérêt de cette méthode dans ce type de situation.

5 Conclusion sur les algorithmes classiques et leurs limitations

À travers ce chapitre, j'ai montré que les algorithmes classiquement utilisés pour traiter les données des systèmes d'imagerie à masque codé trouvent leur application dans la caméra gamma que j'utilise, basée sur la technologie Caliste et le masque codé issu de la première version du masque d'ECLAIRs de la mission SVOM.

Il faut noter que l'adaptation de ces algorithmes et leur fonctionnement n'était pas forcément immédiat car les plans de détection habituellement utilisés pour les missions spatiales comportent beaucoup plus de pixels, de 4 096 pour l'instrument PICsIT de la mission Integral (Goldwurm, et al., 2003) à 32 768 pour l'instrument BAT de la mission Swift (Barthelmy, et al., 2005). Par comparaison aussi dans la sûreté nucléaire, la caméra gamma GAMPIX contient 65 536 pixels (Gmar, et al., 2011). Les détecteurs Caliste possèdent 256 pixels, ce qui diminue potentiellement les capacités d'imagerie par rapport aux instruments que je viens de citer, avec en contrepartie, de bonnes performances spectrométriques.

Malgré ces différences, les algorithmes de déconvolution et MLEM appliqués à ce système d'imagerie permettent de localiser des sources ponctuelles individuelles, avec un nombre de photons assez réduit, allant d'une vingtaine de photons à un peu plus d'une centaine de photons détectés, sur des données réelles issues de mesures avec le détecteur. Dans certains cas, comme montrés en Figure 3-4, les résultats peuvent être très spectaculaires avec des localisations précises avec une information très faible de moins de 20 photons.

Les résultats obtenus avec l'algorithme de déconvolution en rapport signal sur bruit permettent d'identifier avec plus de fiabilité la présence des sources, mais les images fournies sont moins facilement interprétables par un utilisateur final de l'instrument, non expert. A contrario, les résultats obtenus par l'algorithme MLEM sont beaucoup plus nets, mais cet algorithme a une tendance à la surinterprétation des données qui laisse souvent apparaître des fantômes.

Les algorithmes de déconvolution et MLEM donnent aussi des résultats satisfaisants en étant capables de localiser plusieurs sources ponctuelles, qu'elles soient constituées du même radioélément ou de radioéléments différents. La sélection en énergie construisant des spectro-images avec Caliste permet même de localiser indépendamment des sources composées de radioéléments différents.

Un inconvénient est que l'intensité relative des sources n'est correctement reconstruite par aucun des deux algorithmes. J'ai présenté une nouvelle méthode de redressement qui prend en compte à la fois la fraction d'ouverture suivant la position de la source et la redistribution de l'intensité sur les fantômes par l'algorithme MLEM. Ce redressement rapproche les intensités relatives reconstruites de celles attendues mais il ne les retrouve pas fidèlement.

Lorsque le nombre de sources ponctuelles devient plus important, les algorithmes ont plus de difficulté à localiser l'ensemble des sources, comme je l'ai montré avec cinq sources ponctuelles. Par ailleurs, lorsque j'applique ces algorithmes à des sources trop étendues, au-delà d'une taille angulaire de 4°x4°, les algorithmes échouent à localiser correctement ces sources, quand bien même les images sont issues de simulations avec un nombre très élevé de photons et sans aucune source de bruit extérieure. Cette difficulté provient vraisemblablement du fait que, pour ce système d'imagerie, le problème inverse à résoudre possède en réalité plusieurs solutions, et les algorithmes classiques ne permettent pas d'identifier les différentes solutions possibles et d'en choisir la plus probable.

Par ailleurs, j'ai recherché la possibilité d'établir un critère d'arrêt objectif pour l'algorithme MLEM, dont l'absence est un des principaux inconvénients. J'ai pu voir qu'une piste intéressante consisterait à utiliser les variations de la norme 2 des images reconstruites pour établir ce critère. Cette méthode doit être approfondie notamment pour vérifier si elle évite la surinterprétation des données par l'algorithme MLEM et ainsi prévient l'apparition de fantômes lors de la reconstruction.

Enfin, j'ai établi une méthode qui évalue la distance entre la source et la caméra gamma. Cette méthode est basée sur l'association de l'algorithme MLEM avec l'effet de grossissement dû à la distance entre la source et la caméra. J'ai montré qu'elle évalue les distances jusqu'à 1 mètre, avec un pas de 10 cm. De meilleures capacités peuvent être attendues suivant le système de détection utilisé, en particulier suivant la distance entre le masque et le détecteur.

6 Bibliographie

Antier-Farfar, S., 2016. *La détection des sursauts gamma par le télescope ECLAIRs pour la mission.* at: Instrumentation et méthodes pour l'astrophysique [astro-ph.IM]. Université Paris-Saclay.

Barthelmy, S. et al., 2005. *The Burst Alert Telescope (BAT) on the Swift MIDEX Mission.* in: Space Science Reviews, Volume 120, Issue 3-4, Pages 143-164.

Bouchet, L. et al., 2001. *The SIGMA/Granat Telescope: Calibration and Data Reduction.* in: The Astrophysical Journal, Volume 548, Number 2.

Dagoneau, N., Schanne, S., Gros, A. & Cordier, B., 2018. *Detection capability of Ultra-Long Gamma-Ray Bursts with the ECLAIRs telescope aboard the SVOM mission.* in: Week of French Astrophysics.

Dicke, R. H., 1968. *Scatter-Hole Cameras for X-Rays and Gamma Rays.* in: Astrophysical Journal, vol. 153, p.L101.

Fenimore, E. & Cannon, T., 1978. *Coded aperture imaging with uniformly redundant arrays.* in: Applied Optics, Vol. 17, Issue 3, pp. 337-347.

Gmar, M., Agelou, M., Carrel, F. & Schoepff, V., 2011. *GAMPIX: A new generation of gamma camera.* in: NIM-A, Volume 652, Issue 1, Pages 638-640.

Goldwurm, A. et al., 2003. *The INTEGRAL/IBIS scientific data analysis.* in: A&A, Volume 411, Number 1, Pages L223 - L229.

Levitan, E. & Herman, G., 1987. *A Maximum a Posteriori Probability Expectation Maximization Algorithm for Image Reconstruction in Emission Tomography.* in: IEEE Transactions on Medical Imaging, Vol. 6, Issue 3.

Paradiso, V. et al., 2018. *3-D localization of radioactive hotspots via portable gamma cameras.* in: NIM-A, Volume 910, Pages 194-203.

Schanne, S. et al., 2014. *The ECLAIRs GRB-trigger telescope on-board the future mission SVOM.* in: PoS, proceedings of the conference Swift: 10 Years of Discovery.

Shepp, L. & Vardi, Y., 1982. *Maximum Likelihood Reconstruction for Emission Tomography.* in: IEEE Transactions on Medical Imaging, Vol. 1, Issue 2.

Skinner, G., 1984. *Imaging with coded-aperture masks.* in: NIM, Volume 221, Issue 1, Pages 33-40.

Zhang, R. et al., 2019. *Reconstruction method for gamma-ray coded-aperture imaging based on convolutional neural network.* in: Nuclear Inst. and Methods in Physics Research, A, Volume 934, p. 41-51..

Chapitre 5 : Imagerie à masque codé – Méthodes de *deep learning*

1 Base de données et construction de réseaux de neurones 208

2 Performances des réseaux de neurones sur la problématique de la localisation de sources par masque codé 214

3 Superposition optique et sources étendues réelles 230

4 Conclusion 240

5 Bibliographie 241

Dans ce chapitre, j'aborde une approche entièrement nouvelle de traitement des données du système d'imagerie à masque codé à travers des algorithmes basés sur le *deep learning* et en particulier sur les réseaux de neurones de convolution.

Je présente dans un premier temps la méthode que j'ai mise en place pour construire et entraîner les réseaux de neurones. J'utilise des données d'entraînement issues de simulations développées spécifiquement.

Puis, j'évalue les performances de ces réseaux sur des données réelles acquises avec des sources ponctuelles de laboratoire. Sur des données simulées, je démontre la faisabilité d'une reconstruction d'image de sources étendues. Des résultats sur des données réelles sont aussi présentés en fin de chapitre.

Les performances des algorithmes d'apprentissage profonds sont comparées à celles des algorithmes classiques présentés au chapitre précédent, avec une attention particulière portée à MLEM dont je démontre que certaines limitations peuvent être surpassées avec ces méthodes d'intelligence artificielle.

Finalement, je propose une fusion des données gamma avec des acquisitions vidéo. Des premiers résultats de reconstruction sur des données réelles avec des sources ponctuelles et des sources étendues concluent ce chapitre.

1 Base de données et construction de réseaux de neurones

Au niveau de l'état de l'art, l'approche par *deep learning* est relativement nouvelle pour cette application puisqu'elle n'a été utilisée qu'une seule fois par le passé sur l'imagerie à masque codé (Ballesteros, et al., 2001). Dans cette étude, des réseaux de neurones sont entraînés sur des images après traitement par un algorithme MLEM. L'objectif de cette étude n'est pas d'obtenir de meilleures performances que l'algorithme MLEM en matière de localisation, mais en réalité d'entraîner le réseau de neurones à exécuter le même calcul que MLEM sans l'aspect itératif, afin d'en accélérer l'exécution.

De façon contemporaine à mes travaux, deux nouvelles publications d'une même équipe sont parues successivement sur l'application des réseaux de neurones convolutifs (Zhang, et al., 2019) et de réseaux de neurones de type perceptron (Zhang, et al., 2020) à cette problématique. Il est intéressant de constater que dans ces études, le détecteur utilisé est muni d'un cristal CZT pixellisé, de 17x17 pixels de 2 mm d'épaisseur, configuration de détecteur relativement similaire à celle que j'utilise. Le masque utilisé est de type MURA.

Il faut principalement noter que ces études se basent sur des simulations, à la fois pour entraîner les réseaux de neurones et pour les tester. D'autre part, les images reconstruites ont la même pixellisation que le détecteur, soit 17x17, et l'étude montre des limites à ces méthodes en présence de plusieurs sources ponctuelles, avec des difficultés dans les reconstructions de scènes contenant plus de trois sources dans le champ de vue.

Par conséquent, l'absence de littérature antérieure à mes recherches sur ce sujet m'a conduit à mettre en œuvre une méthodologie nouvelle concernant la création de la base de données et la conception du réseau de neurones.

1.1 Génération continue d'une base de données

La première étape pour appliquer des méthodes de *deep learning* est la construction d'une base de données, avec des exemples d'apprentissage que les réseaux de neurones vont utiliser pour optimiser leurs paramètres afin de diminuer la fonction de coût qui leur est associée, comme je l'ai décrit en Chapitre 3.

Les exemples doivent contenir des données d'entrée et des données de sortie. Dans le cadre du problème de la localisation par masque codé, la donnée d'entrée à utiliser qui apparaît naturelle est la carte des coups enregistrée sur le détecteur. Quant aux données de sortie, le résultat que le réseau de neurones doit être en mesure de fournir est l'intensité des sources à chaque position du ciel.

Pour construire un exemple de la base de données, je procède de la manière suivante. Je tire aléatoirement des positions sur le ciel, parmi les 75x75 positions du champ de vue. Je reviens à la suite de ce paragraphe sur la manière dont j'effectue ce tirage aléatoire. À chaque position tirée, je tire aléatoirement une intensité relative entre 0 et 1, de sorte que la somme de toutes les intensités relatives de chaque source soit égale à 1. Par cette méthode, j'obtiens une carte des positions des sources sur le ciel avec leur intensité relative. Cette donnée sera utilisée comme donnée de sortie du réseau de neurones. En réalité, je présenterai un vecteur de taille 5625 = 75x75, chacun des 5625 neurones de sortie représentant une classe correspondant à l'intensité à une position du ciel.

À partir de cette carte des sources, j'utilise ma méthode de simulation décrite en Chapitre 4 afin de générer une carte de coups correspondante sur le détecteur. Le nombre de coups N que j'utilise est tiré aléatoirement suivant la loi suivante :

$$N = 10^{7x+1}, x \sim \mathrm{Unif}([0,1]) \tag{1-1}$$

De cette manière, N est compris entre 10 et 100 millions, en variant de manière logarithmique pour ne pas défavoriser les cas à faible statistique de comptage.

Pour le bruit de fond uniforme, je tire un bruit de fond n d'intensité comprise entre 0 et 1, 0 signifiant qu'il n'y a aucun bruit de fond et 1 signifiant que le bruit de fond est d'intensité égale à l'ensemble des sources. Pour les pixels bruyants, je tire un nombre aléatoire de pixels bruyants N_{bruyants} dans une loi de Poisson de paramètre 0,5.

Ma méthode simule une carte de coups que je normalise ensuite par sa valeur maximale, les valeurs prises par la carte normalisée variant alors entre 0 et 1. J'utilise cette carte normalisée comme donnée d'entrée du réseau de neurones.

Pour tirer aléatoirement les sources sur le ciel, je distingue deux cas conduisant à la génération de deux bases de données d'apprentissage indépendantes :

- le cas de plusieurs sources ponctuelles : pour ce faire, je tire aléatoirement entre 1 et 10 positions sur le ciel, ce qui limite le nombre de sources détectables à 10. Des études supplémentaires seront nécessaires pour déterminer plus précisément le nombre de sources détectables par cette méthode basée sur des réseaux de neurones ;
- le cas d'une source étendue. La génération de formes physiques « cohérentes » reste encore un problème ouvert. Je me suis pour l'instant concentré sur la génération aléatoire de formes élémentaires à partir de la bibliothèque Python *scikit-image* (Van der Walt, et al., 2014) et la fonction *random-shapes* qui génère aléatoirement des carrés, des disques et des triangles, à des positions aléatoires de la carte du ciel, avec des tailles et rotations aléatoires. Comme je l'ai montré en Chapitre 4, les algorithmes existants ne sont pas capables de traiter ce type de sources avec des formes élémentaires. Dans la base de données d'apprentissage, pour chaque exemple, je ne génère qu'une seule source étendue, mais l'utilisation d'autres formes de sources ainsi que de plusieurs sources étendues pourra être étudiée dans des analyses ultérieures.

La Figure 1-1 présente des exemples de simulations dans les deux cas que je viens de décrire.

J'entraînerai deux réseaux de neurones indépendants pour chacun de ces deux cas, multi-sources ponctuelles et une source étendue. Mon objectif est de montrer le potentiel des réseaux de neurones dans des cas pour lesquels les algorithmes de déconvolution et MLEM présentés en Chapitre 4 montrent des limites et de vérifier si, avec un apprentissage adapté, les réseaux de neurones sont capables ou non de surmonter ces limites. Pour des utilisations futures, il sera intéressant d'étudier le potentiel d'un réseau de neurones unique, entraîné sur une base d'apprentissage adaptée, pour pouvoir traiter des cas plus généraux, constitués de mélanges de plusieurs sources ponctuelles et étendues. La base d'apprentissage peut aussi être adaptée suivant l'application visée. Par exemple, dans des applications médicales, on peut penser à l'intérêt de suivre des traceurs radioactifs dans des

zones très localisées du corps, ce qui rend plus intéressant la localisation de sources ponctuelles. À l'inverse, dans des cas d'assainissement et démantèlement, on peut s'attendre plutôt à retrouver des sources étendues.

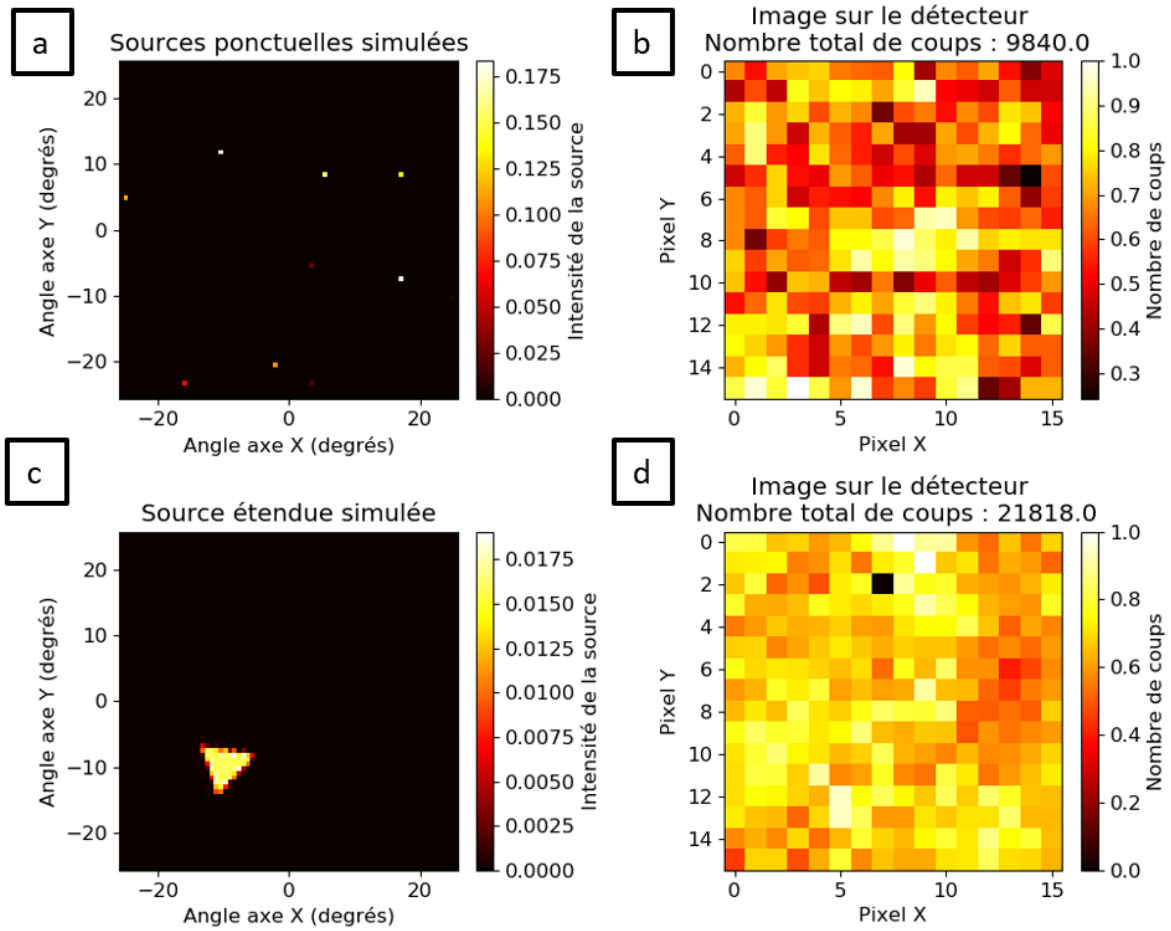


Figure 1-1 : Exemples de sources simulées. a : Plusieurs sources ponctuelles. b : Image correspondante sur le détecteur. c : Source étendue. d : Image correspondante sur le détecteur.

Le dernier point important pour une base de données est le nombre d'exemples qu'elle doit comporter. Sur ce point, j'apporte une méthode spécifique à mon problème. Les données de sortie contiennent un nombre important de 5625 classes et plusieurs classes peuvent être « actives » en même temps, dans le cas de plusieurs sources ponctuelles ou d'une source étendue. Le nombre de scénarios possibles est alors infini en prenant en compte aussi le fait que l'on peut générer des exemples avec des statistiques de comptage différentes et des sources d'intensité variable. En considérant un ordinateur avec une mémoire vive utilisable de 2 Go, en comptant le stockage à la fois des 5625 classes de sortie de chaque exemple et des images de détecteur 16x16 pour l'apprentissage, cela permet d'utiliser une base de données contenant au maximum 85 000 exemples, stockés sous un format *float32*. Cela permet de ne produire qu'en moyenne 15 exemples par classe de sortie, ce qui apparaît trop limité. En effet, rien que si on considère deux sources ponctuelles, il existe 5625 – 1 = 5624 possibilités de combinaisons pour chaque classe (toutes les autres positions du ciel), et cela sans compter les intensités relatives des deux sources ainsi que les statistiques de comptage qui peuvent être différentes.

Un moyen de surmonter cet obstacle est de ne pas utiliser une base de données fixe, mais de la générer au fur et à mesure de l'entraînement du réseau de neurones. J'ai développé l'algorithme de simulation décrit en Chapitre 4 de sorte qu'il soit très rapide : sur un ordinateur de bureau classique, il génère 1000 exemples en 210 ms en moyenne. De ce fait, lors de l'apprentissage, je génère la base de données de manière continue : je génère 1000 exemples sur lesquels j'entraîne le réseau de neurones pendant une époque seulement, puis je génère 1000 nouveaux exemples pour un nouvel apprentissage d'une époque, processus que je répète pendant plusieurs centaines de milliers d'itérations. Le nombre d'exemples à chaque étape peut être optimisé, j'en ai choisi 1000 car cela permet de faire une étape d'apprentissage en 500 ms.

Cette méthode me permet de présenter au réseau de neurones plusieurs centaines de millions d'exemples, au lieu de n'en présenter que quelques dizaines de milliers. Elle présente aussi l'avantage d'éviter naturellement le surapprentissage sur un nombre limité d'exemples, vu que les réseaux de neurones utilisent en permanence des exemples différents lors de l'apprentissage.

1.2 Architecture du réseau de neurones et apprentissage

Les réseaux de neurones sur lesquels je me suis concentré sont les réseaux de neurones à couches de convolution bayésiens. J'ai décrit leur fonctionnement dans le Chapitre 3, dans un cas à une dimension, pour les spectres. Ici, pour les images des détecteurs, je les utilise avec des couches de convolution contenant des filtres à deux dimensions, le fonctionnement étant ensuite similaire à celui que j'ai décrit au Chapitre 3.

Les motivations pour utiliser des réseaux de neurones à couches de convolution proviennent du fait que, sur les images des détecteurs, les données sont localement corrélées à travers la forme des projections des trous du masque codé. Les couches de convolution tirent profit de ces corrélations locales. Par ailleurs, l'utilisation de réseaux bayésiens établit une incertitude sur les résultats comme indiqué au Chapitre 3, et cette incertitude est très importante pour caractériser pleinement les résultats des réseaux de neurones.

Aux débuts de ma thèse, il n'existait aucun exemple d'architecture de réseaux de neurones dans la littérature, appliqué à la problématique de la localisation de sources par masque codé. L'absence d'état de l'art antérieur m'a poussé à partir d'une « page blanche » pour concevoir une architecture répondant au problème. En procédant méthodiquement par ajout de couches et de neurones tel que décrit au Chapitre 3 et en effectuant des tests sur les réseaux que j'ai construits, j'ai abouti à l'architecture présentée en Figure 1-2. J'ai utilisé une architecture identique pour le réseau dédié aux sources ponctuelles et celui spécifique aux sources étendues.

La principale difficulté que j'ai eue lors de la construction du réseau de neurones est le fait que, lorsque l'architecture n'est pas adaptée, le réseau de neurones se « bloque » dans un minimum local, dans lequel la réponse qu'il fournit en sortie est une carte du ciel de type « bruit blanc », où une intensité aléatoire est renvoyée pour chaque position du ciel. Ce phénomène s'explique principalement par le fait que, parmi les 5625 classes présentes sur les exemples que j'utilise, seules quelques classes sont actives pour chaque exemple. La très grande majorité des classes, plus de 99 % d'entre elles, valent 0, d'où le fait que le réseau est proche d'un optimum lorsqu'il tente de mettre toutes les classes à des valeurs proches de 0, les classes ayant une valeur non nulle ont ultimement peu de poids dans la

fonction de coût. L'apprentissage converge lorsque le réseau de neurones arrive à dépasser cet optimum.

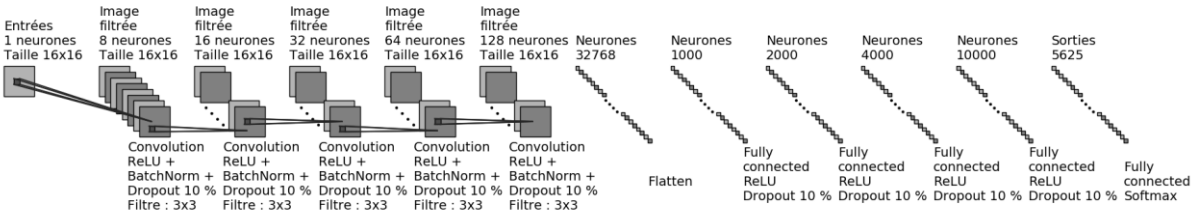


Figure 1-2 : Architecture du réseau de neurones que j'ai développée pour la localisation de sources radioactives par masque codé. Cette architecture est utilisée à la fois pour l'entraînement sur les sources ponctuelles et sur les sources étendues.

Je n'effectue pas d'opération de *pooling* entre chaque couche de convolution, opération qui réduit la taille des données d'un facteur k pour un pooling de taille k, k étant un nombre entier au moins égal à 2. En effet, la taille des données d'entrée est déjà assez modeste, 16x16, répéter par exemple deux fois une opération de *pooling* de taille 2 conduirait à des données finales de taille 4x4, sur lesquelles il n'est plus concevable d'effectuer des opérations de convolution.

J'ai développé un réseau assez profond, avec notamment cinq couches *fully-connected*, ce qui est sensiblement différent des réseaux de neurones développés au Chapitre 3 qui ne contenaient que deux couches de ce type. Lors des tests, j'ai constaté que l'ajout de couches *fully-connected* est essentiel pour garantir une bonne convergence du réseau de neurones. En utilisant seulement trois couches de ce type par exemple, l'apprentissage n'était pas capable de converger sur les essais d'entraînement que j'ai effectués. En revanche, je n'ai pas pu faire de test avec des couches *fully-connected* supplémentaires, les capacités GPU que j'utilisais ne permettant pas de faire des apprentissages avec un réseau aussi profond.

Pour l'évaluation des incertitudes dans le cas du réseau bayésien, j'ai utilisé un taux de *dropout* de 10 %, et non de 50 % comme en Chapitre 3. Avec un taux plus élevé que 10 %, l'apprentissage n'est pas capable de converger d'après les tests que j'ai effectués.

Les couches intermédiaires ont toutes une fonction d'activation ReLU, de façon à avoir un meilleur contrôle du gradient du fait des propriétés de cette fonction. La couche de sortie est une fonction d'activation *softmax*, puisque la somme des sorties doit être égale à 1, par la construction que j'ai faite des données de sortie décrites au paragraphe 1.1.

Les données de sorties ont des valeurs comprises entre 0 et 1, correspondant à l'intensité relative de chaque source. Le problème à traiter est donc un problème de régression et non un problème de classification, une fonction de coût naturelle à utiliser serait la distance euclidienne. Cependant, d'après les tests que j'ai effectués, l'utilisation de la distance euclidienne ne fait pas converger l'apprentissage, qui reste bloqué dans l'optimum consistant à mettre des valeurs aléatoires de type « bruit blanc », proches de 0, à l'ensemble des classes.

À la place de la distance euclidienne, l'utilisation de l'entropie binaire croisée, *binary cross-entropy*, a permis de surpasser ce problème. Le problème de cette fonction de coût est qu'elle ne vaut 0 en son minimum que si les classes réelles utilisées valent 0 ou 1, correspondant à un problème de

classification. Pour les problèmes de régression où les valeurs attendues sont comprises entre 0 et 1, cette fonction de coût reste utilisable dans le sens où elle atteint son minimum lorsque les données prédites sont égales aux données attendues pour les exemples. Il faut seulement souligner le fait que ce minimum ne vaut pas 0.

Le fait que l'apprentissage fonctionne avec l'entropie binaire croisée et non avec la distance euclidienne doit encore être étudié plus en détail. Ce phénomène n'est cependant pas forcément surprenant, car la fonction de coût à optimiser conditionne le gradient utilisé pour mettre à jour les paramètres : la modification de la fonction de coût implique une modification du gradient. Certaines fonctions de coût peuvent être plus avantageuses selon les situations, mais il n'est pas simple d'expliquer clairement ce phénomène du fait de la complexité des réseaux de neurones. Cela fait partie des aspects « boîte noire » de ce type de méthode.

J'utilise l'algorithme d'optimisation ADAM (Kingma & Ba, 2015) avec les moments $\beta_1 = 0{,}9$ et $\beta_2 = 0{,}999$ et un taux d'apprentissage de 10^{-4}, qui garantit la convergence de l'apprentissage d'après les essais que j'ai effectués. Je génère la base de données de manière continue comme je l'ai décrit au paragraphe 1.1, en créant des petites bases de données de 1000 exemples sur lesquels j'effectue un apprentissage d'une époque, avec des *mini-batchs* de taille 100. Cette base de données est utilisée sur 900 exemples comme jeu d'apprentissage et sur 100 exemples comme jeu de validation. Je répète ce processus plus de 200 000 fois.

Par ailleurs, j'effectue un entraînement avec des exemples représentant plusieurs sources ponctuelles et un entraînement avec des exemples représentant une seule source étendue pour chacun d'entre eux. J'obtiens, in fine, deux réseaux de neurones, avec chacun la même architecture, mais entraînés sur des bases de données différentes.

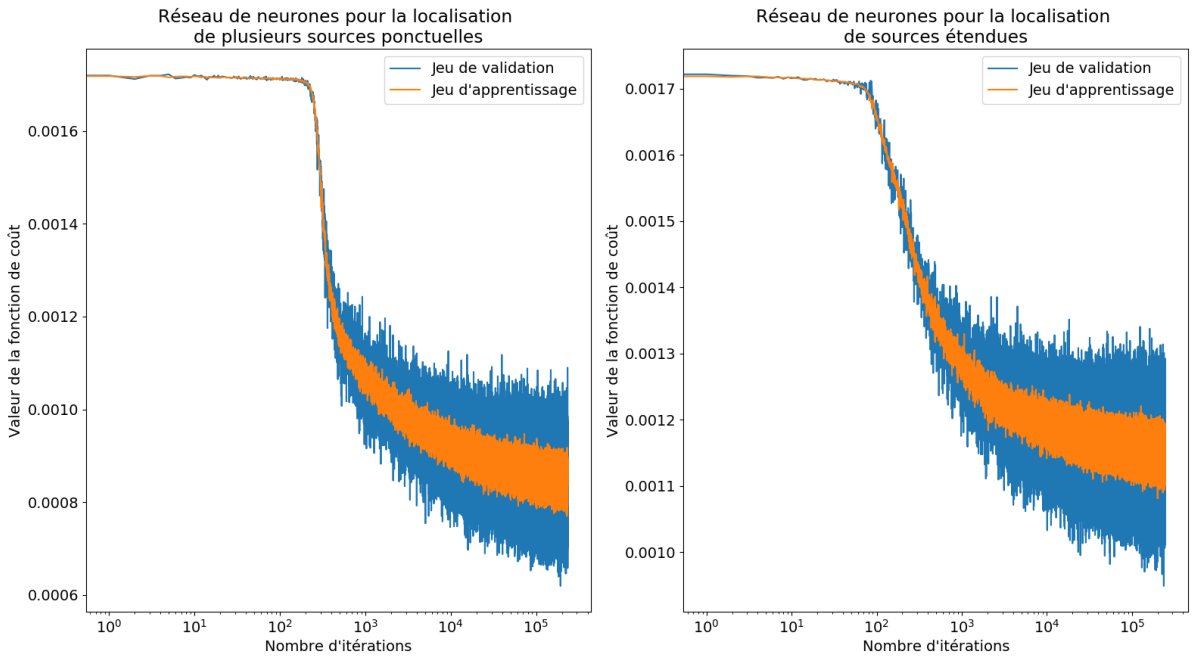


Figure 1-3 : Évolution de la fonction de coût lors de l'apprentissage des réseaux de neurones.

L'évolution de la fonction de coût lors de l'apprentissage est représentée en Figure 1-3. On constate en particulier que la fonction de coût stagne au niveau des premières itérations à une valeur de 0,0017.

Cela correspond au moment où le réseau de neurones renvoie en sortie un « bruit blanc », lorsqu'il est dans l'optimum local évoqué précédemment. Après quelques centaines d'itérations, la fonction de coût finit par diminuer, ce qui montre que le réseau arrive à être entraîné. Elle finit par atteindre des valeurs d'environ 8.10^{-4} pour le réseau de neurones dédié aux sources ponctuelles et 10^{-3} pour celui spécifique aux sources étendues.

À ce niveau, la fonction de coût oscille autour de ces valeurs. Elle est très bruitée notamment du fait que j'utilise une petite base de données à chaque itération, contenant seulement 1000 exemples, dont seulement 100 pour le jeu de validation. À ce moment, j'ai laissé l'apprentissage tourner sur un grand nombre d'itérations, mais il est envisageable de s'arrêter lorsque la fonction de coût atteint la zone d'oscillations autour de 8.10^{-4} et 10^{-3}. Cependant, vu que les oscillations sont assez importantes, il est difficile de déterminer un critère d'arrêt très précis pour l'apprentissage.

2 Performances des réseaux de neurones sur la problématique de la localisation de sources par masque codé

2.1 Tests sur des données réelles avec des sources ponctuelles

Une fois les réseaux de neurones entraînés, j'effectue des tests sur leurs performances. Dans un premier temps, j'utilise les données réelles sur des sources ponctuelles, testées en Chapitre 4, qui servent d'éléments de comparaison des performances par rapport aux algorithmes classiques.

Pour chaque figure suivante, je compare la reconstruction effectuée par les réseaux de neurones avec l'algorithme MLEM. Par rapport à l'algorithme de déconvolution, j'ai montré que l'algorithme MLEM donnait des résultats plus simplement interprétables. Même si les sources sont ici ponctuelles, j'utilise à la fois le réseau de neurones entraîné sur des simulations de sources ponctuelles et celui entraîné sur des simulations de sources étendues, afin d'analyser les capacités de ce second réseau de neurones dans le cas où des sources ponctuelles lui sont présentées. Pour chaque exemple, les réseaux de neurones sont exécutés 100 fois avec un *dropout* aléatoire pour obtenir une incertitude autoévaluée par le réseau. Les résultats que je présente utilisent la médiane des 100 essais. Enfin, sur les différentes acquisitions, j'effectue une sélection dans la bande d'énergie d'intérêt, ici celle du ^{241}Am, entre 50 keV et 70 keV.

2.1.1 Source ponctuelle unique dans l'axe

La Figure 2-1 représente les résultats dans le cas d'une source ponctuelle au centre du champ de vue, à différents temps d'acquisition.

Avec 40 s d'acquisition, l'algorithme MLEM est bien capable de reconstruire la source, avec un GCR* de 3 et un CNR de 122. Le réseau de neurones localise la position correcte de la source, avec un CNR de presque 20. En tenant compte de l'incertitude† autoévaluée par le réseau bayésien, l'intensité à la

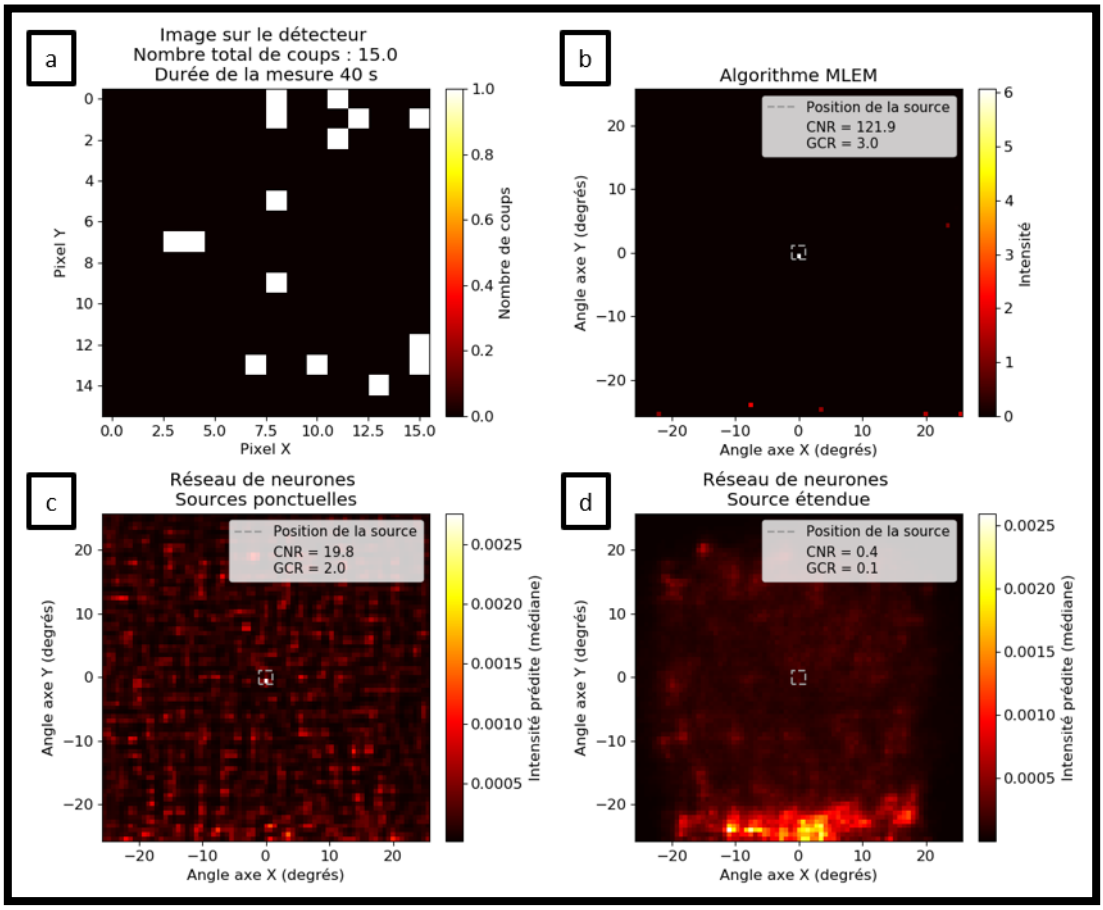
* Les notions de CNR (*Contrast-to-Noise Ratio*) et GCR (*Ghost Contrast Ratio*) sont définies au Chapitre 4.
† La représentation des incertitudes de manière synthétique en une seule image est difficile. J'avais pu le faire en Chapitre 3, car les données attendues étaient en une dimension. Ici, avec des données en deux dimensions, il faudrait utiliser une représentation des incertitudes en trois dimensions.

bonne position de la source varie de $1{,}3.10^{-3}$ (quantile à 5 %) à 8.10^{-3} (quantile à 95 %), soit d'un facteur 6. Cette incertitude tient notamment au fait que le nombre de photons détectés est très faible, seulement 15 photons.

Avec 1000 s d'acquisition, le réseau de neurones localise la source avec un CNR très élevé de plus de 16 000 et un GCR de plus de 500, permettant d'affirmer l'absence de fantôme. En parallèle, MLEM présente un CNR de 274 et un GCR de 7.

Avec une statistique de 244 photons au total, compte tenu de l'incertitude autoévaluée par le réseau, l'intensité à la position de la source varie de 0,75 (quantile à 5 %) à près de 0,9 (quantile à 95 %). Dans ces conditions, le réseau localise la source avec certitude.

Le réseau de neurones entraîné sur des sources étendues n'est pas capable de reconstruire les sources avec 40 secondes d'acquisition, en donnant un CNR inférieur à 1. En revanche, avec 1000 secondes d'acquisition, il reconstruit la source à la bonne position, avec un GCR de plus de 100. Il est intéressant de noter que ce réseau donne une « largeur » à la source, qui s'étale sur deux pixels du ciel, ce réseau étant entraîné justement à reconnaître des sources étendues, alors que les deux autres algorithmes renvoient une source effectivement ponctuelle sur un seul pixel du ciel.



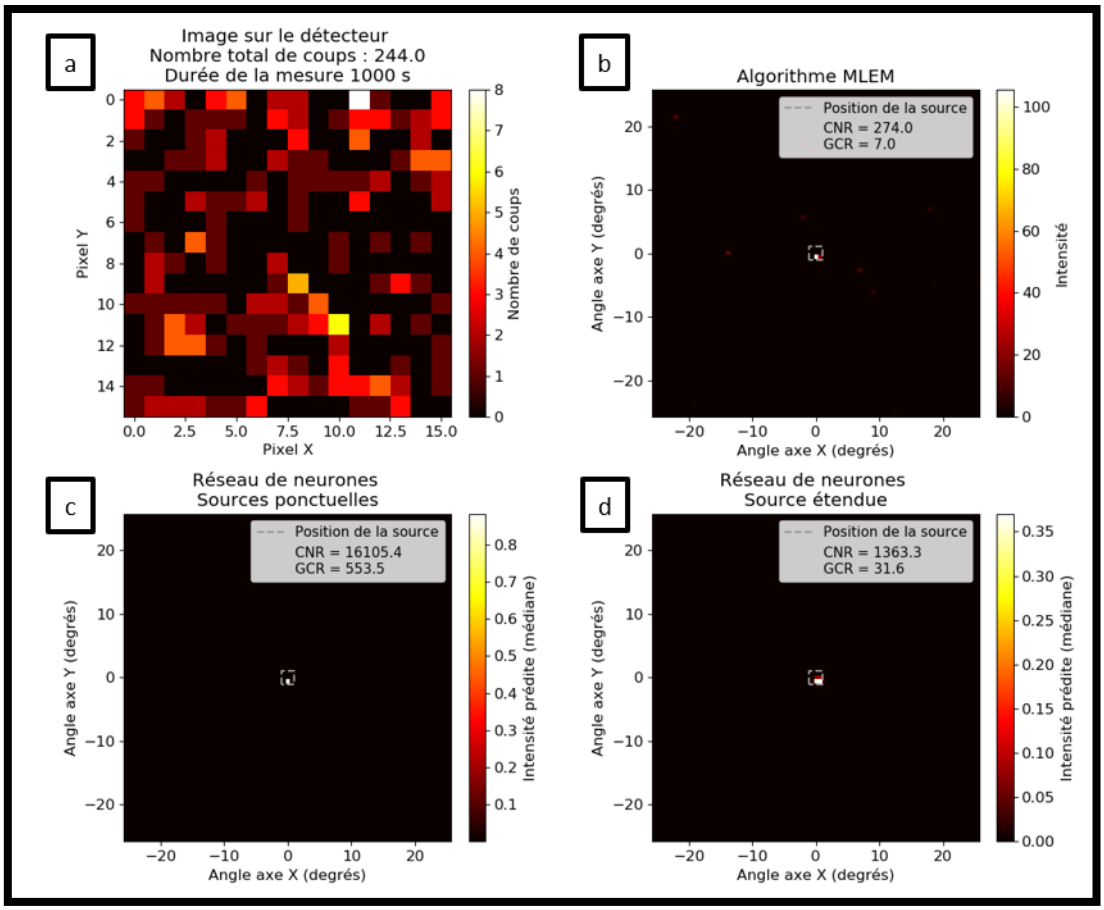
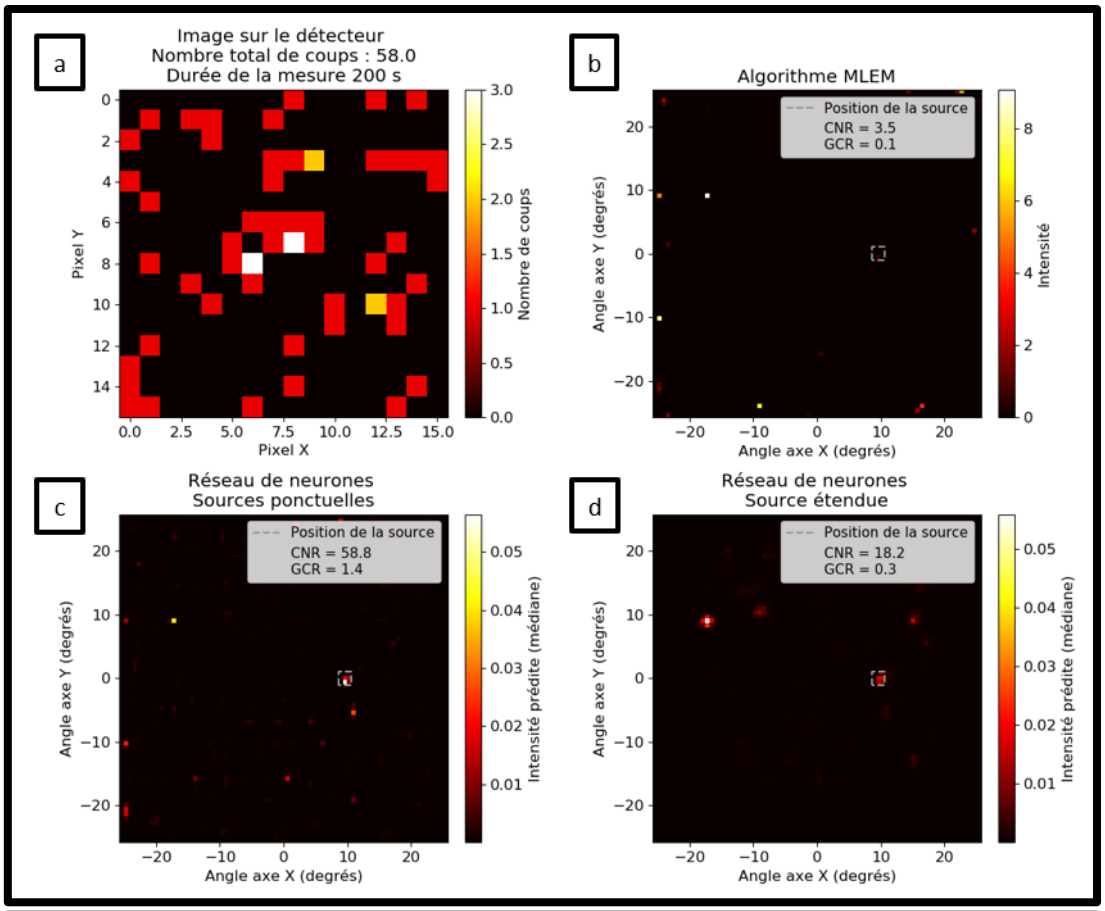
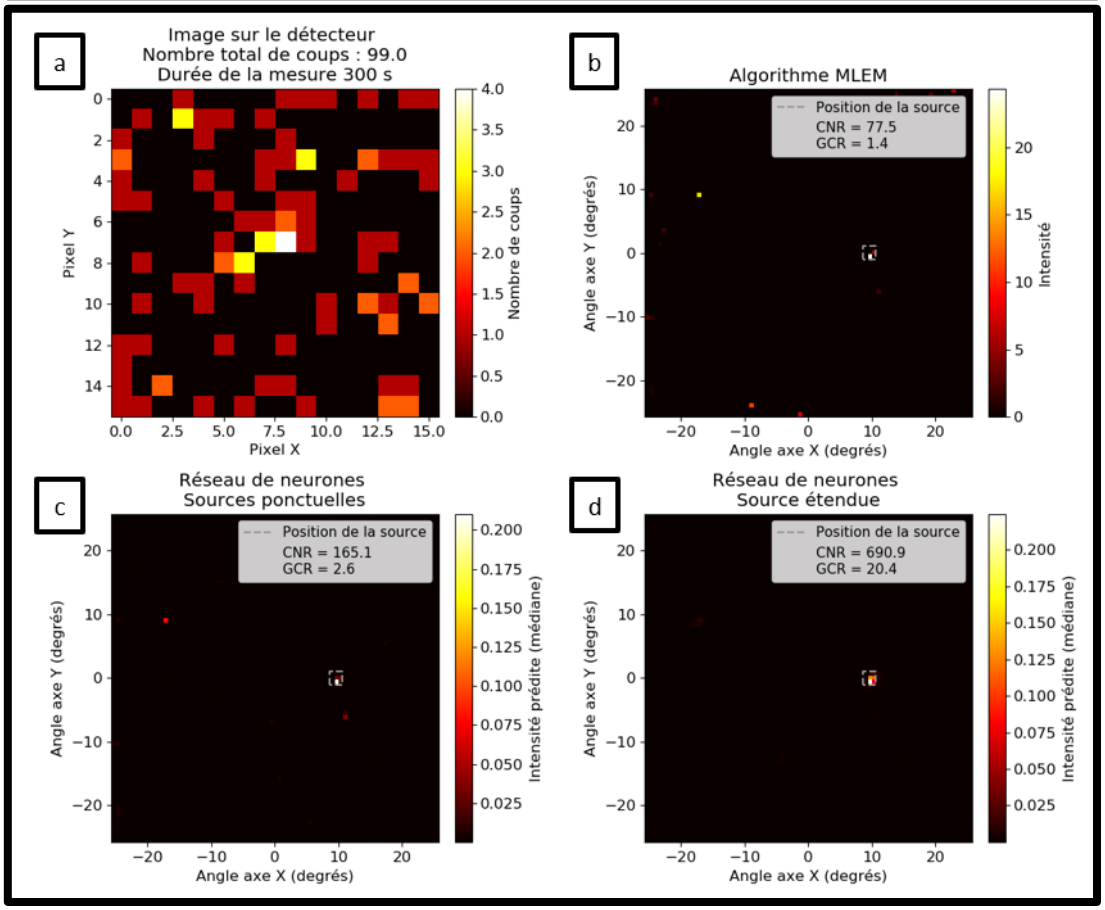


Figure 2-1 : Acquisition d'une source de ^{241}Am de 440 kBq placée à 1 m de la caméra gamma. Sélection en énergie entre 50 keV et 70 keV. Paramètres d'acquisition : temps de *peaking* 6,07 µs, gain 200 mV/fC, température du cristal -10 °C.
Pour chaque panel de 4 images : 40 s et 1000 s d'acquisition. a : Image sur le détecteur. b : Reconstruction MLEM. c : Reconstruction par le réseau de neurones entraîné sur des sources ponctuelles. d : Reconstruction par le réseau de neurones entraîné sur des sources étendues.

2.1.2 Source ponctuelle unique hors axe

La Figure 2-2 montre les résultats des algorithmes dans le cas d'une source ponctuelle excentrée, avec différents temps d'acquisition. Avec 200 secondes d'acquisition, l'algorithme MLEM ne localise pas la source et ne fait apparaître que des fantômes (GCR = 0,1 et CNR = 3,5). Le réseau de neurones entraîné sur des sources ponctuelles est capable de localiser la source à la bonne position avec CNR de 59, qui nous assure de sa présence. Cela illustre une meilleure spécificité des réseaux de neurones, c'est-à-dire, une meilleure robustesse aux faux-positifs, par rapport à l'algorithme MLEM.


a
Image sur le détecteur
Nombre total de coups : 58.0
Durée de la mesure 200 s
Pixel Y
Pixel X
Nombre de coups
b
Algorithme MLEM
Position de la source
CNR = 3.5
GCR = 0.1
Angle axe Y (degrés)
Angle axe X (degrés)
Intensité
c
Réseau de neurones
Sources ponctuelles
Position de la source
CNR = 58.8
GCR = 1.4
Intensité prédite (médiane)
d
Réseau de neurones
Source étendue
Position de la source
CNR = 18.2
GCR = 0.3


a
Image sur le détecteur
Nombre total de coups : 99.0
Durée de la mesure 300 s
Pixel Y
Pixel X
Nombre de coups
b
Algorithme MLEM
Position de la source
CNR = 77.5
GCR = 1.4
Angle axe Y (degrés)
Angle axe X (degrés)
Intensité
c
Réseau de neurones
Sources ponctuelles
Position de la source
CNR = 165.1
GCR = 2.6
Intensité prédite (médiane)
d
Réseau de neurones
Source étendue
Position de la source
CNR = 690.9
GCR = 20.4

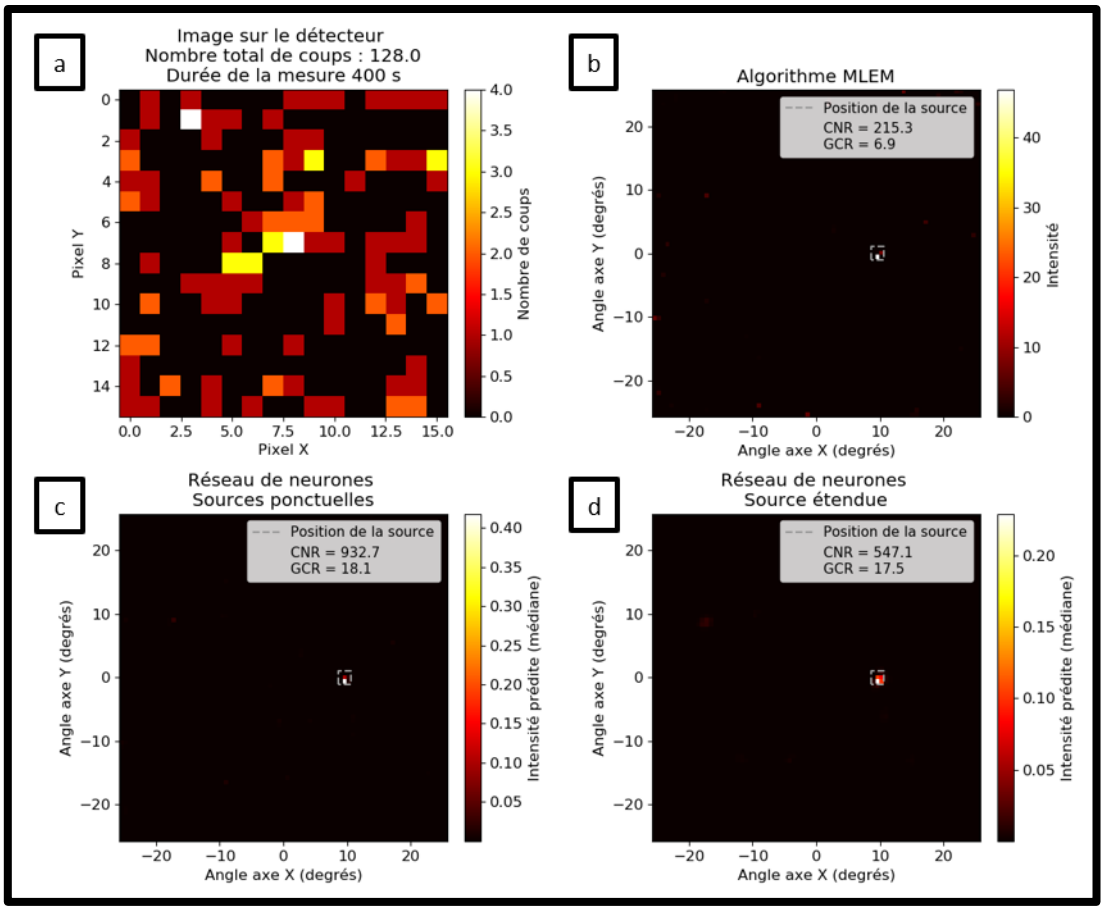


Figure 2-2 : Acquisition d'une source de ^{241}Am de 440 kBq placée à 1 m de la caméra gamma. Sélection en énergie entre 50 keV et 70 keV. Paramètres d'acquisition : temps de *peaking* 6,07 µs, gain 200 mV/fC, température du cristal -10 °C.
Pour chaque panel de 4 images : respectivement 200 s, 300 s et 400 s d'acquisition. a : Image sur le détecteur. b : Reconstruction MLEM. c : Reconstruction par le réseau de neurones entraîné sur des sources ponctuelles. d : Reconstruction par le réseau de neurones entraîné sur des sources étendues.

On peut d'ailleurs noter que certains fantômes sont communs aux deux algorithmes, comme à la position (-19°,9°). L'algorithme entraîné sur des sources étendues fait apparaître la source à la bonne position, même s'il attribue l'intensité maximale sur le fantôme à gauche de l'image avec un GCR de 0,3.

Avec 300 secondes d'acquisition, la plupart des fantômes disparaissent, même s'il reste encore celui en (-19°,9°) pour l'algorithme MLEM, avec un GCR de 1,4. Le réseau de neurones spécifique aux sources ponctuelles a un GCR de 2,6. Le réseau de neurones entraîné sur des sources étendues ne fait apparaître que la bonne source, avec un GCR supérieur à 20, mais cela doit provenir de la manière dont je l'ai entraîné : comme je ne lui ai montré qu'une seule source étendue dans chaque exemple, il privilégie des résultats dans lesquels il ne fait apparaître qu'une seule source.

Enfin, avec 400 secondes d'acquisition, l'ensemble des fantômes a disparu pour les algorithmes, avec des GCR supérieurs à 6 pour MLEM et à 17 pour les réseaux de neurones.

Comparaison de sensibilité entre l'algorithme MLEM et le réseau de neurones

Afin de comparer la sensibilité entre l'algorithme MLEM et le réseau de neurones pour une source ponctuelle, j'ai utilisé cinq acquisitions, dans lesquelles j'ai tiré aléatoirement des photons pour obtenir une carte de coups, à partir de laquelle je reconstruis la position des sources.

J'ai appliqué l'algorithme MLEM et le réseau de neurones spécifique aux sources ponctuelles et j'ai dénombré le nombre de tests pour lesquels chaque algorithme localise correctement la source, c'est-à-dire, pour lesquels il attribue l'intensité maximale à la position attendue de la source. Comme en Chapitre 3, j'appelle *accuracy* le taux de localisations correctes.

J'ai effectué cette étude en fonction du nombre de photons que j'ai tirés dans la carte de coups, avec 1000 exemples pour chaque statistique de comptage que j'ai étudiée. Les résultats sont présentés en Figure 2-3.

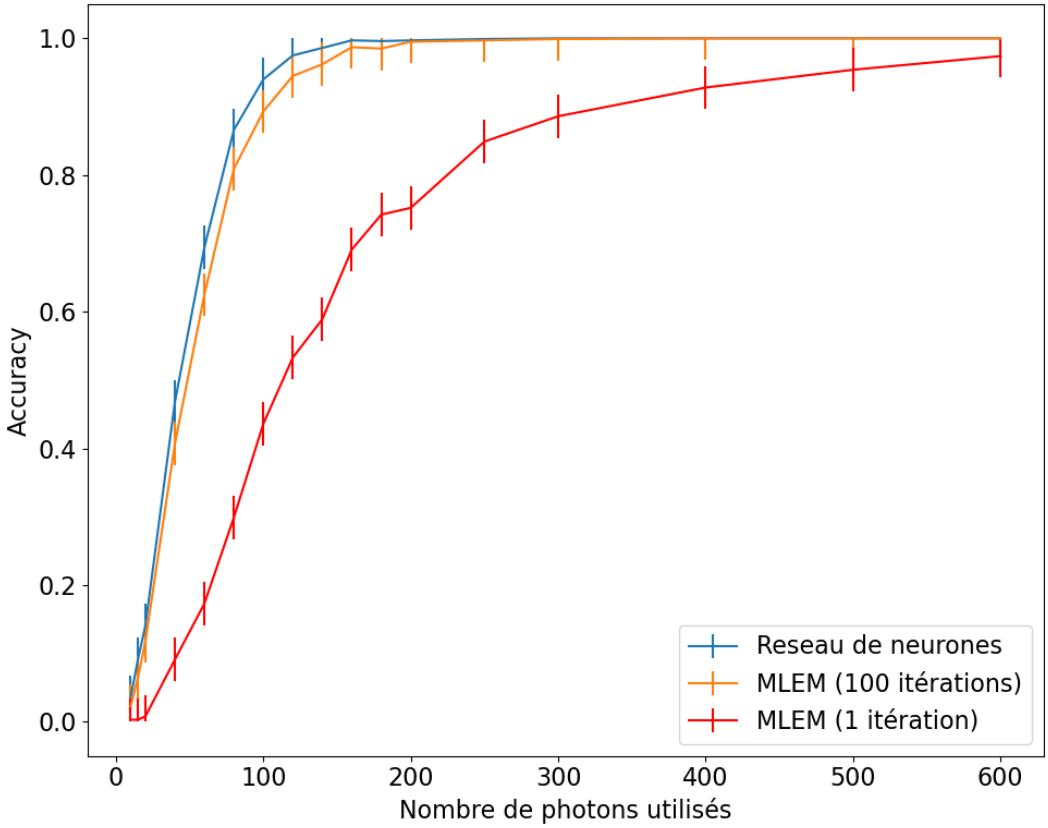


Figure 2-3 : *Accuracy* des algorithmes en fonction du nombre de photons utilisés pour former la carte de coups.

On voit que les performances au niveau de l'*accuracy* sont similaires pour le réseau de neurones et l'algorithme MLEM avec 100 itérations, quelle que soit la statistique de photons. Cela dénote une sensibilité équivalente entre les deux algorithmes. Il faut notamment de l'ordre de 200 photons pour obtenir une *accuracy* supérieure à 99 %.

Par ailleurs, il faut noter que le temps de calcul du réseau de neurones est de l'ordre de 0,035 s, alors qu'il faut 3,5 s pour 100 itérations de l'algorithme MLEM. Le réseau de neurones a donc un temps de calcul équivalent à une itération de MLEM et comme on le voit en Figure 2-3, il a de meilleures performances que MLEM avec une seule itération.

2.1.3 Sources ponctuelles multiples

Cas de deux sources ponctuelles hors axe, même radioélément

La Figure 2-4 montre un exemple avec deux sources ponctuelles et 400 secondes d'acquisition, permettant de détecter un peu plus de 300 photons. L'algorithme MLEM fait apparaître les deux sources aux bonnes positions, mais avec la présence de plusieurs fantômes dans le champ de vue, le GCR d'une des deux sources étant inférieur à 2. Le réseau de neurones entraîné sur des sources étendues localise uniquement la source la plus intense à gauche, ce qui provient de nouveau du fait que ce réseau est entraîné à rechercher une seule source étendue.

Le réseau de neurones spécifique aux sources ponctuelles fait apparaître principalement la source la plus intense à gauche. La source à droite apparaît avec un CNR de plus de 60, mais une intensité relative plus faible qu'attendue par rapport à celle de gauche. J'avais déjà montré en Chapitre 4 que cet effet peut s'expliquer par la fraction d'ouverture du masque et par l'effet de *vignetting* ainsi que par le fait qu'une partie de l'intensité de la source de droite est attribuée à d'autres positions fantômes sur le ciel. Ici, on constate que le réseau de neurones a une sensibilité équivalente à l'algorithme MLEM mais une meilleure spécificité du fait qu'il fait apparaître moins de fantômes sur le ciel.

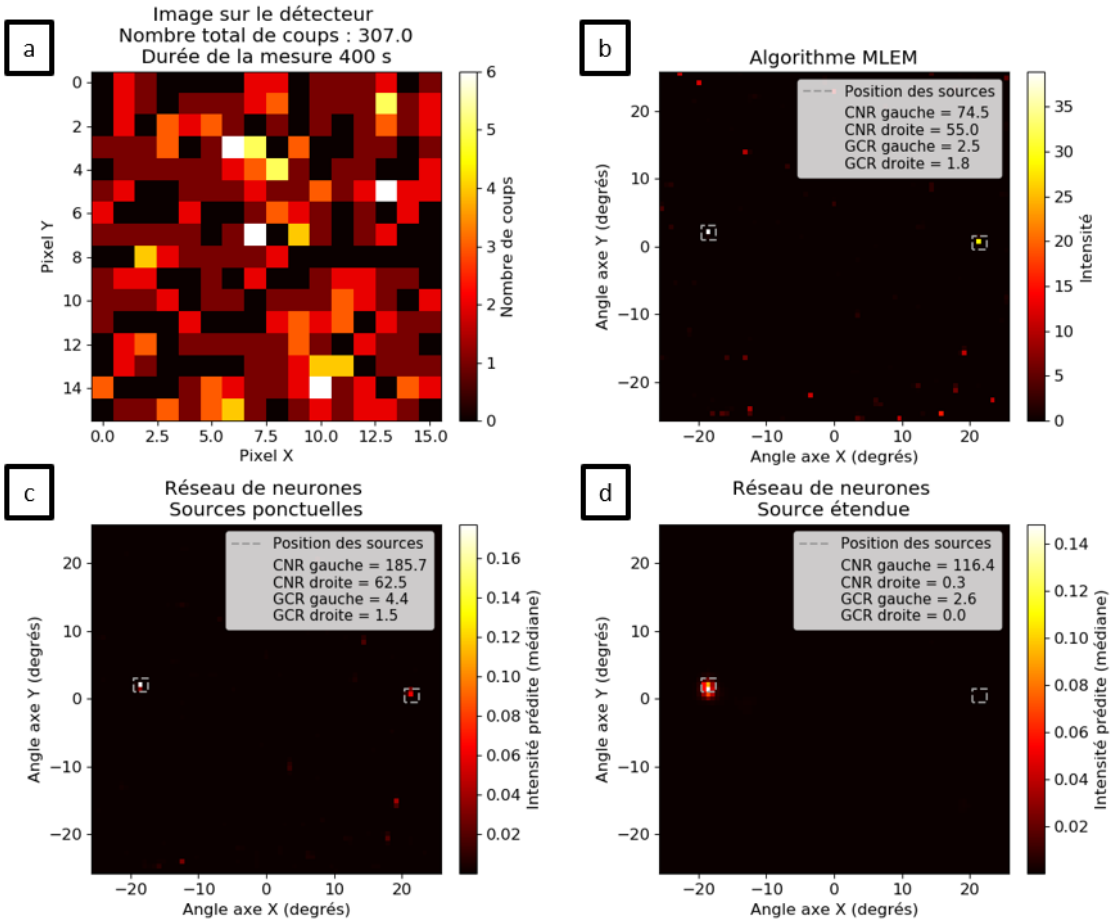


Figure 2-4 : Acquisition de deux sources de ^{241}Am de 440 kBq et 394 kBq placées à 1 m de la caméra gamma. Sélection en énergie entre 50 keV et 70 keV. Paramètres d'acquisition : temps de *peaking* 6,07 µs, gain 200 mV/fC, température du cristal -10 °C. a : Image sur le détecteur. b : Reconstruction MLEM. c : Reconstruction par le réseau de neurones entraîné sur des sources ponctuelles. d : Reconstruction par le réseau de neurones entraîné sur des sources étendues.

Cas de cinq sources ponctuelles, même radioélément

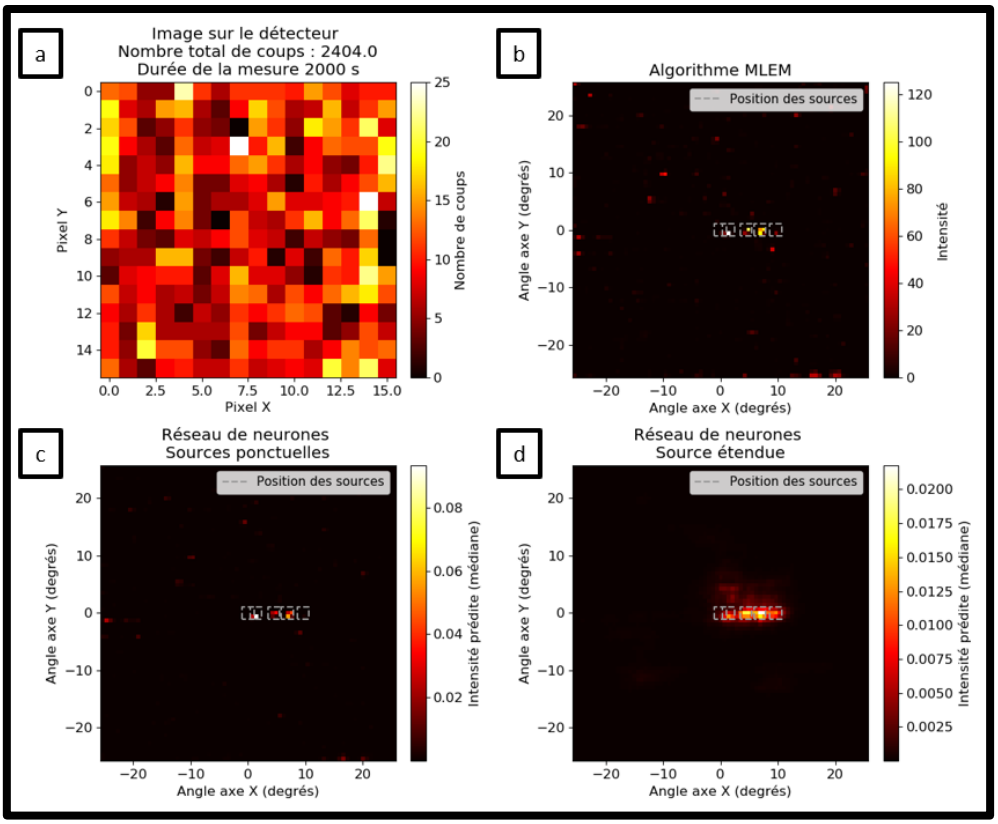
Je présente en Figure 2-5 le cas avec cinq sources ponctuelles à deux temps d'acquisition différents et le Tableau 2-1 résume les performances en CNR et GCR pour les deux temps d'acquisition.

L'algorithme MLEM a des difficultés à reconstruire l'ensemble des sources et on remarque que le réseau de neurones entraîné sur des sources ponctuelles présente les mêmes difficultés de reconstruction. La zone de présence des sources est bien identifiée, mais les algorithmes ne localisent pas l'ensemble des cinq sources. La source la plus à droite n'est pas localisée : elle a un CNR est inférieur à 5 et un GCR proche de 0.

Par ailleurs, il est intéressant de noter que le résultat du réseau de neurones spécifique aux sources étendues est correct et localise bien la zone de présence des sources. Cela provient principalement du fait que les sources ont des positions proches et elles peuvent être considérées comme une seule source étendue dans ce cas. Le réseau de neurones est alors en mesure de renvoyer une réponse cohérente avec la position de la zone concernée.

Position des sources			(0°,0°)	(1.4°,0°)	(4.5°,0°)	(6.9°,0°)	(9.6°,0°)
Acquisition	Algorithme	Valeur					
2000 s	MLEM	CNR	10	**66,3**	**54,9**	**48,8**	**11,7**
		GCR	0,4	**2,6**	**2,2**	1,9	0,5
	CNN sources ponctuelles	CNR	**38,7**	**154,4**	**54,8**	**93,9**	4,1
		GCR	1	**4,1**	1,5	**2,5**	0,1
	CNN sources étendues	CNR	15,4	**28,3**	**39,6**	**48,5**	27,1
		GCR	0,6	1,1	1,5	1,8	1
17000 s	MLEM	CNR	**13,4**	**90,2**	**72,6**	**47,6**	10,5
		GCR	0,4	**2,6**	**2,1**	1,4	0,3
	CNN sources ponctuelles	CNR	**66,3**	**203,5**	**58,7**	**84,2**	2,9
		GCR	**2,2**	**6,9**	**2**	**2,8**	0,1
	CNN sources étendues	CNR	**42,4**	**107,4**	**50,6**	**78,3**	**16,9**
		GCR	1,6	**4**	1,9	**2,9**	0,6

Tableau 2-1 : Valeurs de CNR et GCR correspondant à la Figure 2-5. Les valeurs en gras représentent celles qui sont satisfaisantes du point de vue des critères établis au Chapitre 4 (CNR > 10 et GCR > 2).



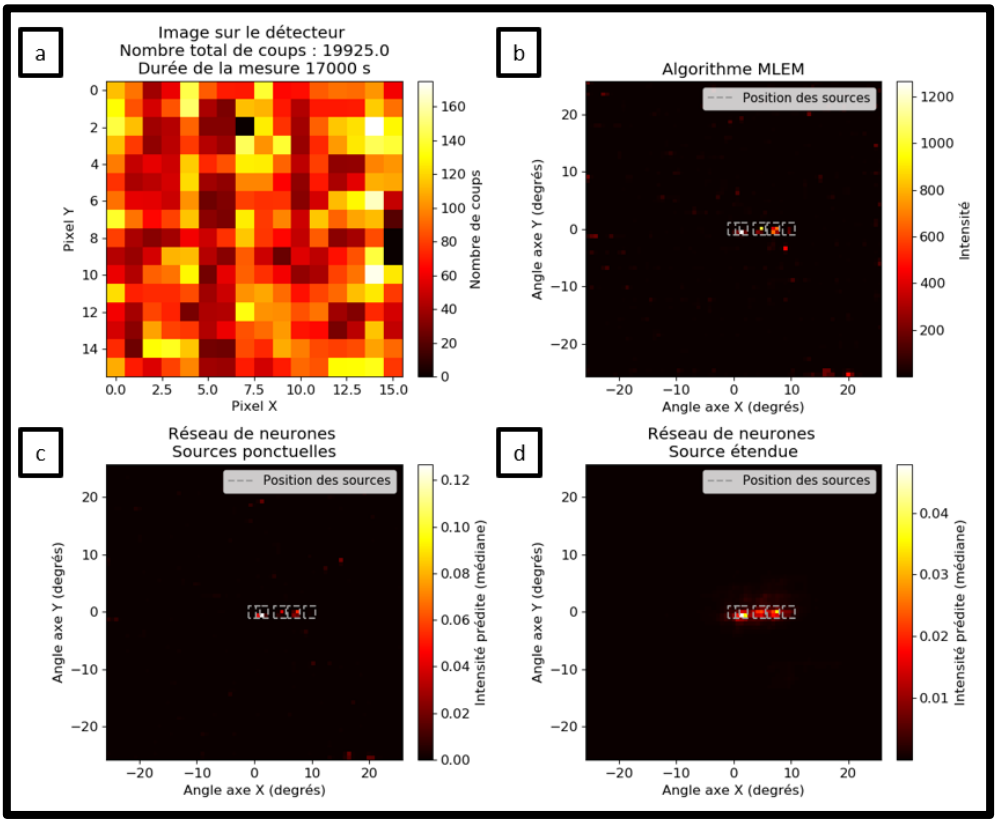


Figure 2-5 : Acquisition de cinq sources de ^{241}Am de 440 kBq placées à 1 m de la caméra gamma pendant 4 h 45 mn. Sélection en énergie entre 50 keV et 70 keV. Pour chaque panel de 4 images : respectivement 33 minutes et 4 h 45 mn d'acquisition. a : Image sur le détecteur. b : Reconstruction MLEM. c : Reconstruction par le réseau de neurones entraîné sur des sources ponctuelles. d : Reconstruction par le réseau de neurones entraîné sur des sources étendues.

2.2 Tests sur des données simulées de sources étendues

Cette fois, je m'intéresse aux capacités des algorithmes à reconstruire des sources étendues. Je compare les performances du réseau de neurones dédié aux sources étendues à celles du réseau de neurones entraîné sur des sources ponctuelles ainsi qu'aux algorithmes classiques. Dans cette section, la faisabilité est démontrée sur des sources étendues de complexité croissante, d'un simple carré à des formes arbitraires, par nature indisponibles dans un laboratoire.

Pour les tests, j'utilise des sources qui n'ont pas été incluses dans la base de données d'apprentissage. L'intérêt sur les simulations est que je maîtrise entièrement la taille et la forme de la source, facteurs importants pour pouvoir caractériser en détail les performances du réseau.

Toutefois, la démonstration n'est complète qu'après que j'effectue des tests sur des données réelles en paragraphe 3.2, en utilisant un système qui « mime » une acquisition d'une source étendue circulaire avec des sources ponctuelles de laboratoire mises en rotation rapide devant le temps de pose.

2.2.1 Cas d'une source étendue 2°x2°

En Figure 2-6, je commence avec le cas d'une simulation d'une source légèrement étendue, de 2°x2°, soit 3 pixels du ciel avec la discrétisation que j'utilise. Ce cas est simulé sans bruit de fond ajouté. La simulation utilise 224 photons.

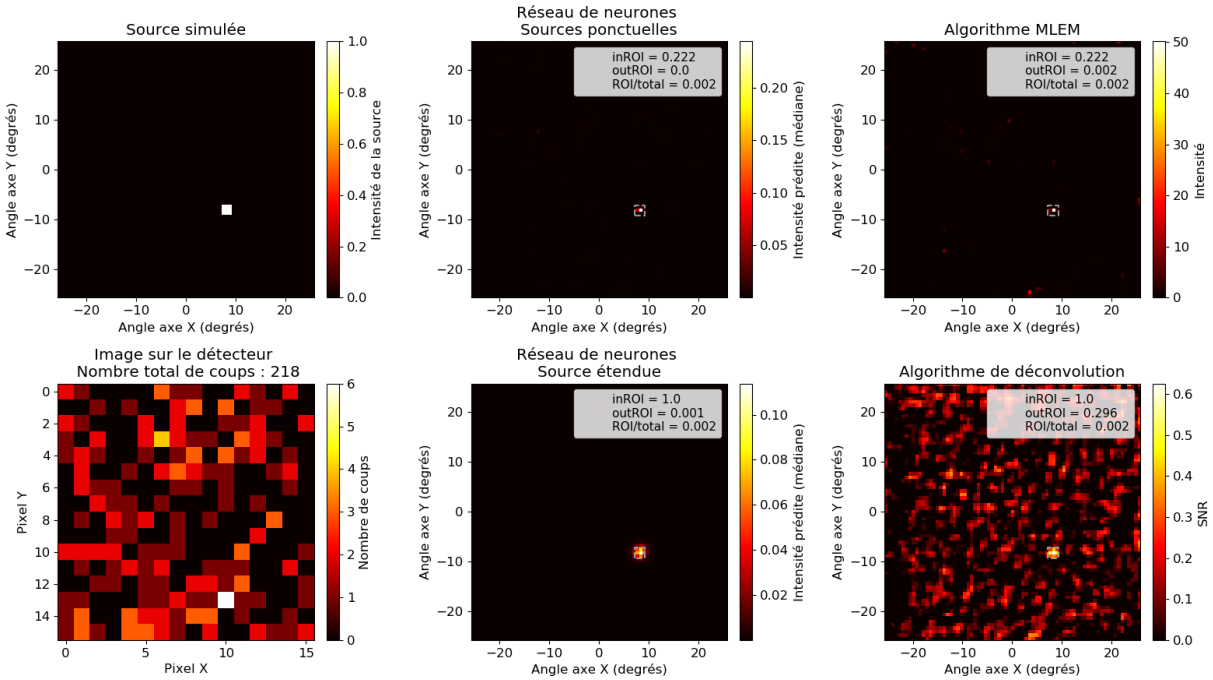


Figure 2-6 : Résultat de simulation d'une source carrée 2°x2° uniforme avec 224 photons, sans bruit. En haut à gauche : source simulée. En bas à gauche : image de la simulation sur le détecteur. Au milieu et à droite : reconstructions par les différents algorithmes.

Les algorithmes classiques positionnent la source avec la présence de fantômes du fait de la statistique assez faible. En effet, pour MLEM par exemple, la zone la plus intense se situe dans la zone

correspondant à la source simulée, mais ne la couvre pas entièrement avec un inROI* de 22 %, ce qui correspond à un pixel du ciel. De même, le réseau de neurones spécifique aux sources ponctuelles retrouve la position de la source, en ne reconstruisant aucun point en dehors de la source, avec un outROI de 0. Il ne reconstruit que 22 % de la source, à l'instar de MLEM. Le réseau de neurones entraîné pour les sources étendues retrouve la source avec une taille cohérente à celle attendue : la source est entièrement reconstruite avec un inROI de 1. Le outROI est de 0,1 %, ce qui montre que quelques pixels en dehors de la source sont reconstruits, mais ils se situent tous à la frontière de la zone attendue.

La Figure 2-7 montre la simulation de la même source avec l'ajout d'un bruit uniforme de même intensité que la source.

Cette fois, 9000 photons simulés sont nécessaires pour que tous les algorithmes retrouvent la position de la source. Le réseau de neurones entraîné sur des sources étendues rend compte du fait que la source n'est pas ponctuelle, le bruit uniforme rend moins nettes la forme et la taille de la source : le inROI est toujours de 1 mais le outROI passe à 0,7 %, les positions supplémentaires reconstruites s'étalent sur deux pixels du ciel par rapport à la position de la vraie source, donnant un diamètre apparent de la source 70 % plus grand que ce que j'avais simulé.

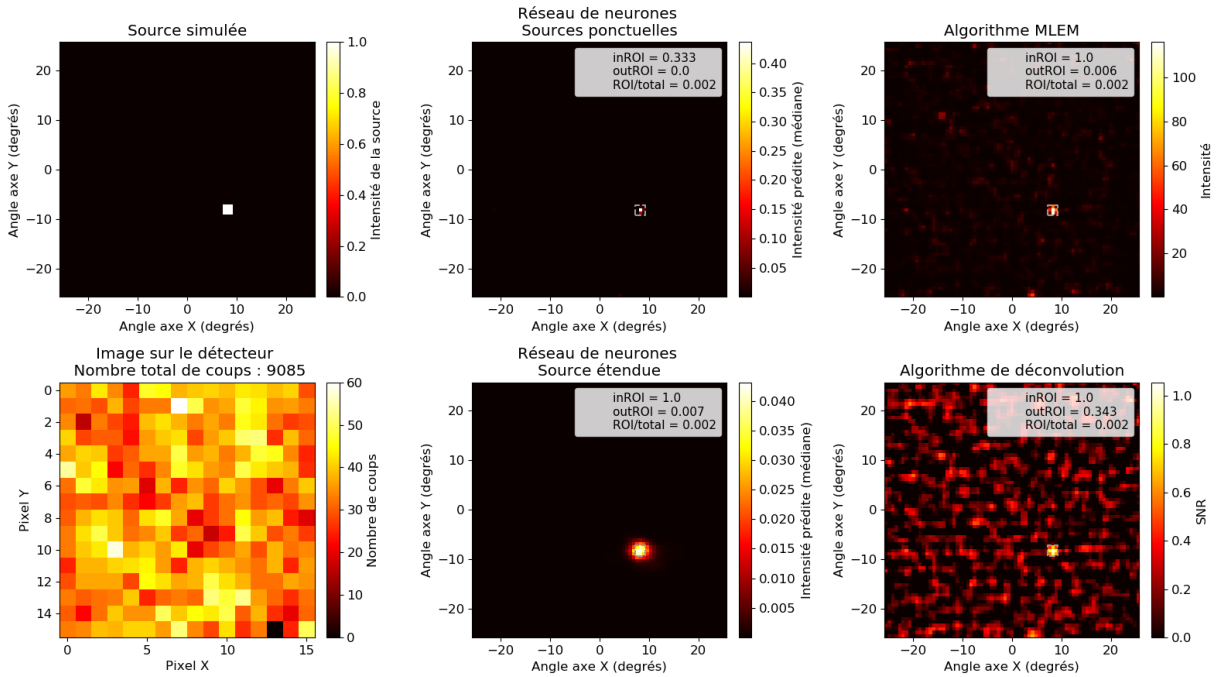


Figure 2-7 : Résultat de simulation d'une source carrée 2°x2° uniforme avec 9000 photons, avec bruit uniforme. En haut à gauche : source simulée. En bas à gauche : image de la simulation sur le détecteur. Au milieu et à droite : reconstructions par les différents algorithmes.

2.2.2 Cas d'une source étendue 4°x4°

En Figure 2-8, je passe à une source plus grande, de taille 4°x4°. J'ai montré en Chapitre 4 que les algorithmes de reconstruction classiques sont incapables de traiter ce cas, même avec un nombre très élevé de photons de l'ordre de la centaine de millions. Ici, j'effectue une simulation avec seulement 1000 photons. De nouveau, les algorithmes classiques sont incapables de traiter ce type de

* Les notions de inROI, outROI et ROI/total sont définies au Chapitre 4.

donnée. Seulement 14 % de la source est reconstruite par MLEM. L'algorithme reconstruit 0,4 % du reste du ciel alors que la source ne représente que 0,6 % de l'ensemble du ciel. Le réseau de neurones spécifique aux sources ponctuelles renvoie aussi un résultat dans lequel la source n'apparaît pas, en reconstruisant seulement 20 % de la source.

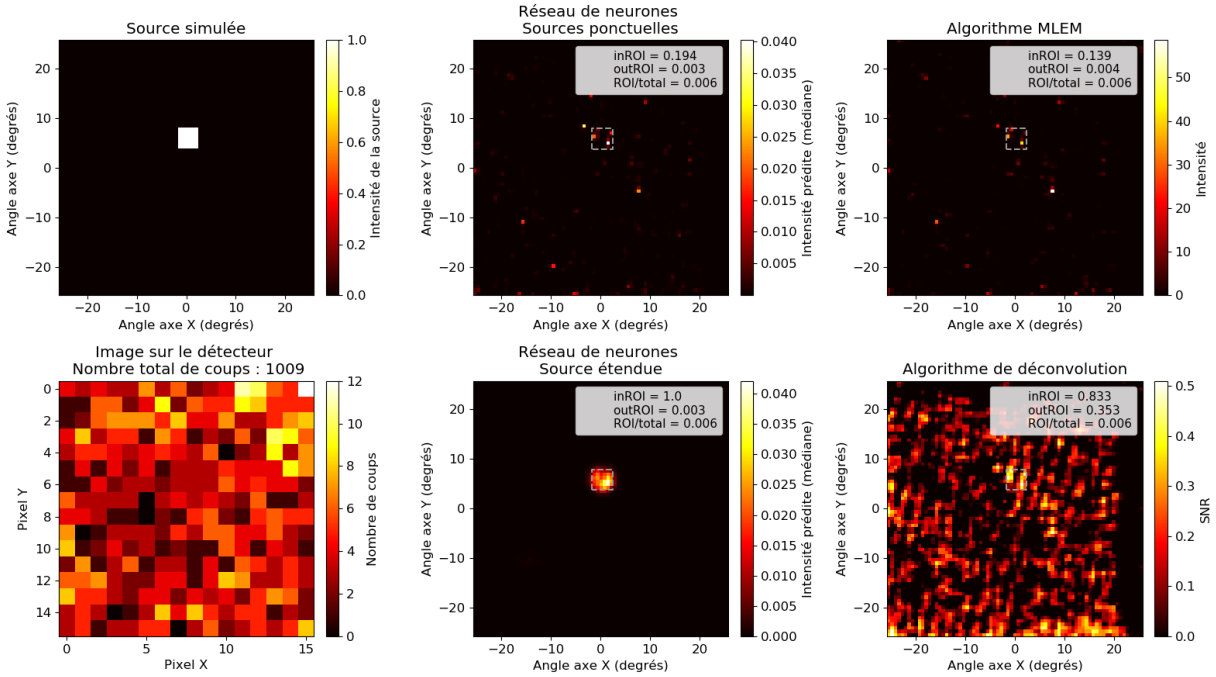


Figure 2-8 : Résultat de simulation d'une source carrée 4°x4° uniforme avec 1000 photons, sans bruit. En haut à gauche : source simulée. En bas à gauche : image de la simulation sur le détecteur. Au milieu et à droite : reconstructions par les différents algorithmes.

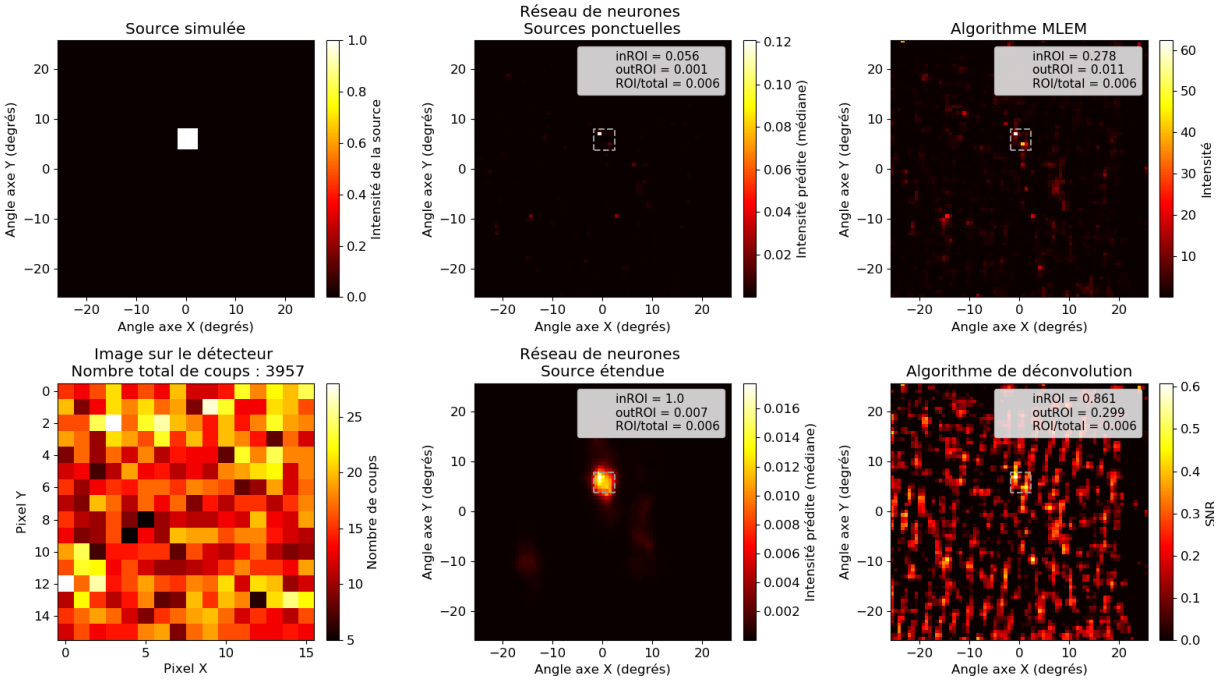


Figure 2-9 : Résultat de simulation d'une source carrée 4°x4° uniforme avec 4000 photons, avec bruit uniforme d'intensité égale à celle de la source. En haut à gauche : source simulée. En bas à gauche : image de la simulation sur le détecteur. Au milieu et à droite : reconstructions par les différents algorithmes.

En revanche, le réseau de neurones entraîné pour les sources étendues est capable de retrouver la position de la source, et renvoie aussi une image dont la taille et la forme sont cohérentes avec la source simulée, avec un inROI de 1 et un outROI de 0,3 %, sachant que les positions reconstruites en dehors de la zone correspondant à la source sont situées à la frontière de cette zone.

Il est particulièrement intéressant de noter que seuls 1000 photons sont nécessaires au réseau de neurones pour effectuer cette reconstruction dans le cas où aucun bruit uniforme n'est appliqué. En appliquant un bruit de fond uniforme d'intensité équivalente à celle de la source, un nombre plus élevé de photons est nécessaire, environ 4000 ici, comme je le montre en Figure 2-9. Le réseau de neurones retrouve la source avec un inROI de 1, mais un outROI plus grand que précédemment, de 0,7 %, l'étalement sur l'axe Y augmente de 50 % la taille de la source dans cet axe.

2.2.3 Cas d'une source étendue 7°x7°

La Figure 2-10 et la Figure 2-11 montrent une simulation avec une source encore plus grande, de 7°x7°, sans bruit uniforme, avec 6000 et 20 000 photons respectivement. Les algorithmes classiques et le réseau de neurones spécifique aux sources ponctuelles sont incapables de fournir une information, même partielle, ni sur la position ni sur la forme de la source.

Le réseau de neurones spécifique aux sources ponctuelles a un inROI de 1 % seulement et MLEM a un inROI de 9 %. En revanche, le réseau de neurones entraîné pour les sources étendues est capable de localiser la source avec 6000 photons, en renvoyant une forme et une taille correspondant à celle de la source : le inROI est de 1 et le outROI est de 1,3 %, inférieur aux 1,8 % occupés par la source dans le ciel. D'ailleurs, de nouveau, les positions reconstruites à tort sont situées au maximum à une distance de 3 pixels du ciel par rapport à la source. En augmentant la statistique de comptage, la forme et la taille de la source reconstruite sont plus précises et correspondent mieux à la source effectivement simulée, en diminuant le outROI à 0,9 %.

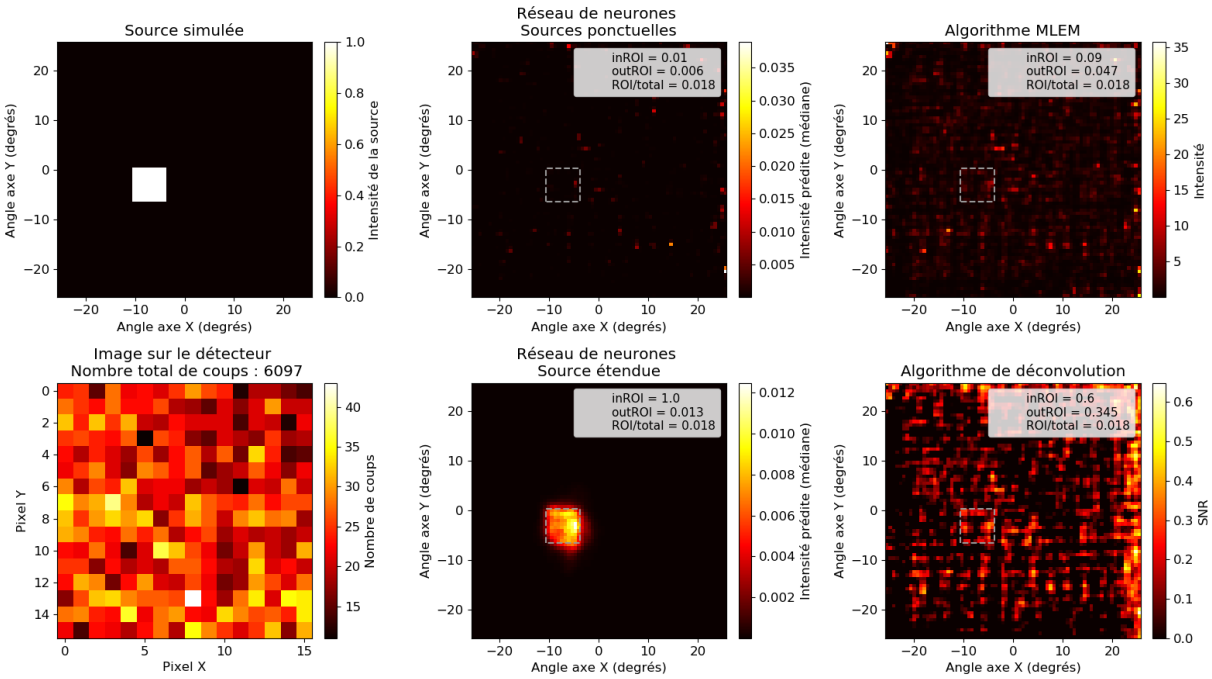


Figure 2-10 : Résultat de simulation d'une source carrée 7°x7° uniforme avec 6000 photons, sans bruit. En haut à gauche : source simulée. En bas à gauche : image de la simulation sur le détecteur. Au milieu et à droite : reconstructions par les différents algorithmes.

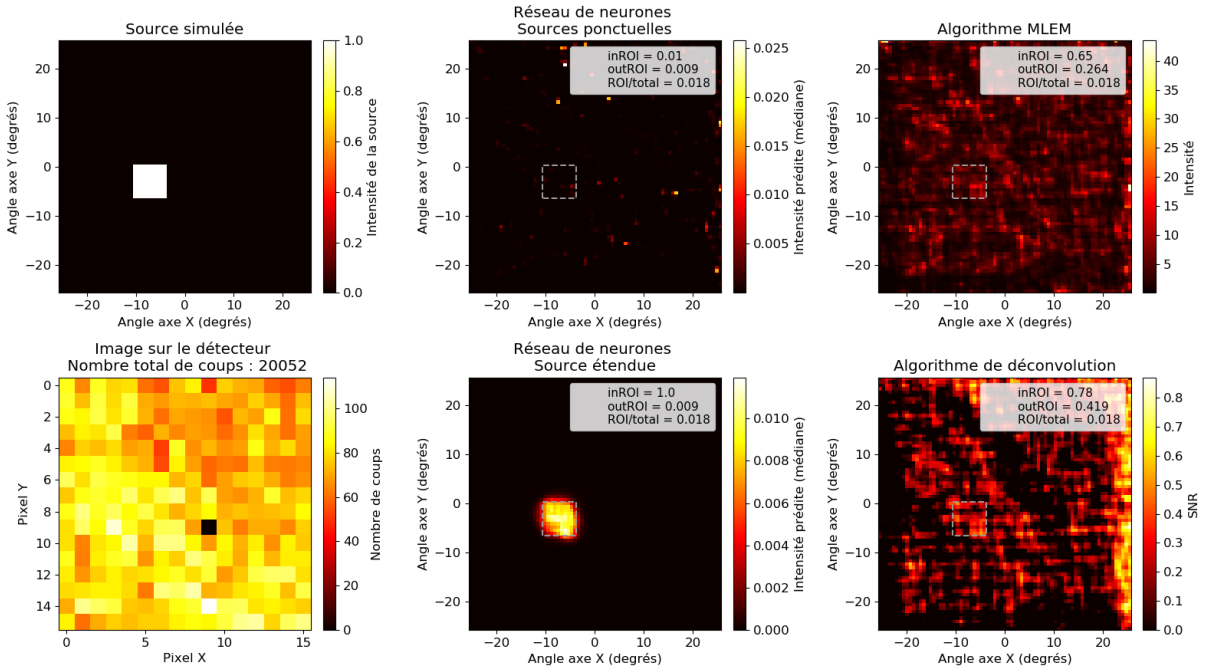


Figure 2-11 : Résultat de simulation d'une source carrée 7°x7° uniforme avec 20 000 photons, sans bruit. En haut à gauche : source simulée. En bas à gauche : image de la simulation sur le détecteur. Au milieu et à droite : reconstructions par les différents algorithmes.

2.2.4 Cas d'une source étendue de forme et taille arbitraires

Pour finir, je simule une source étendue d'un aspect complètement différent de ce qui a été utilisé dans la base d'entraînement, avec une intensité non uniforme, tirée aléatoirement dans la source. Les résultats sont présentés en Figure 2-12 et Figure 2-13, respectivement avec 9700 photons sans bruit de fond et avec 50 000 photons et un bruit de fond uniforme d'intensité égale à celle de la source.

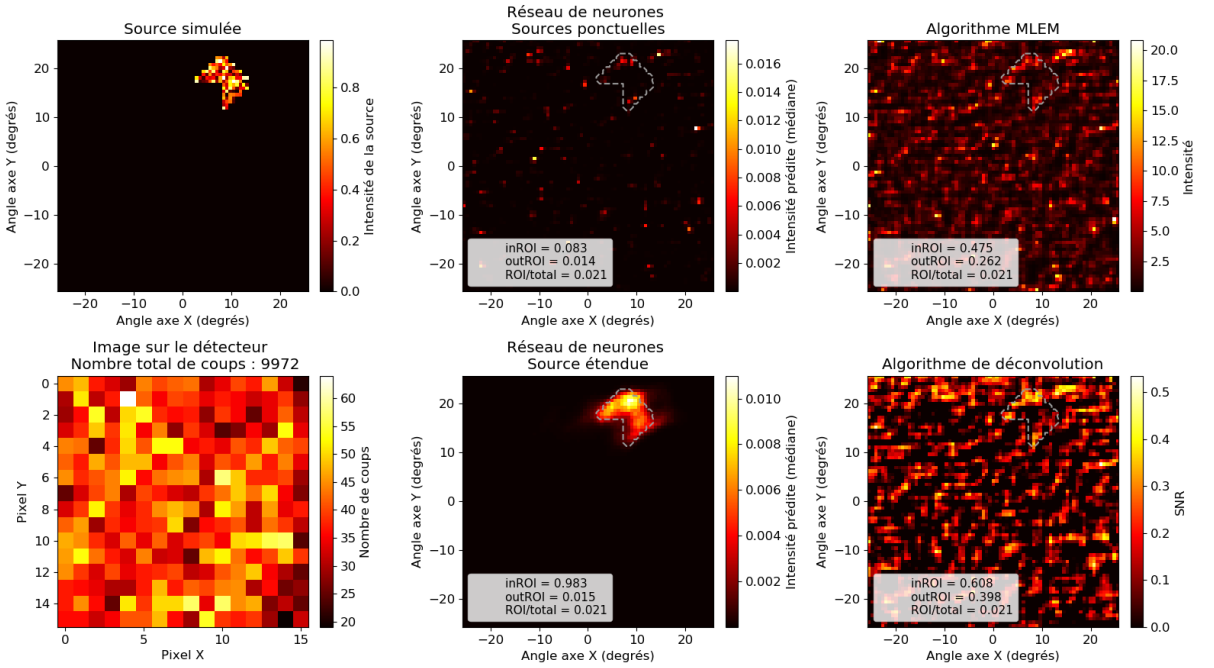


Figure 2-12 : Résultat de simulation d'une source d'intensité non uniforme avec 9700 photons, sans bruit. En haut à gauche : source simulée. En bas à gauche : image de la simulation sur le détecteur. Au milieu et à droite : reconstructions par les différents algorithmes.

Dans tous les cas, les algorithmes classiques et le réseau de neurones spécifique aux sources ponctuelles échouent à reconstruire les sources que j'ai simulées.

En l'absence de bruit, le réseau de neurones entraîné pour les sources étendues a besoin de 10 000 photons environ pour localiser la source et retrouver sa taille. La forme est peu précise mais reste assez proche de la source simulée : plus de 98 % de la source est reconstruite avec un outROI de 1,5 % alors que la source occupe 2,1 % du ciel. Les positions reconstruites à tort sont situées au maximum à une distance de 4 pixels de la zone attendue. En revanche, les variations d'intensité à l'intérieur de la source ne sont pas reconstruites.

Avec un bruit de fond uniforme, environ cinq fois plus de photons sont nécessaires pour obtenir une reconstruction cohérente avec la source simulée, avec un inROI de plus de 97 %, comme je le montre en Figure 2-13.

Cet exemple illustre la capacité de généralisation des réseaux de neurones. Bien que la forme que j'ai utilisée en exemple ne faisait pas partie des formes simulées, le réseau de neurones a été capable de renvoyer une information sur la position et la taille et une forme cohérente avec la source simulée.

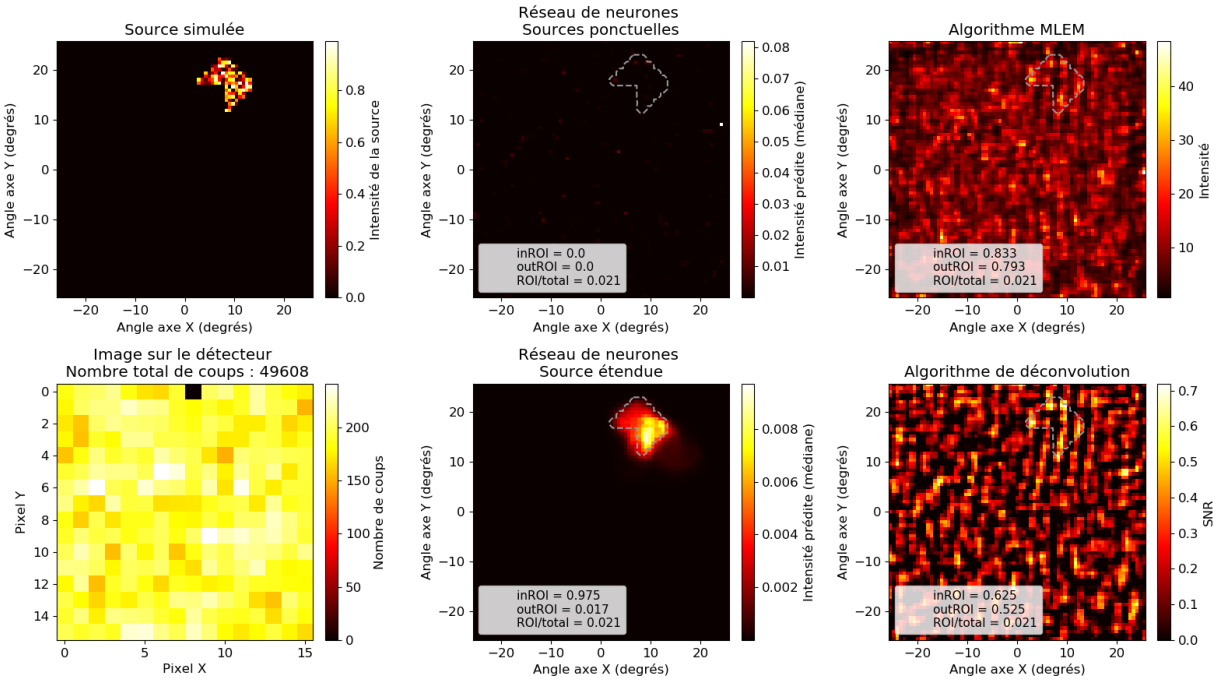


Figure 2-13 : Résultat de simulation d'une source d'intensité non uniforme avec 50 000 photons, avec bruit uniforme. En haut à gauche : source simulée. En bas à gauche : image de la simulation sur le détecteur. Au milieu et à droite : reconstructions par les différents algorithmes.

2.3 Conclusion sur l'application des réseaux de neurones à la reconstruction par masque codé

J'ai étudié la possibilité d'utilisation des réseaux de neurones pour l'application aux masques codés pour la reconstruction d'images gamma de sources ponctuelles, uniques ou multiples ou encore de sources étendues. En utilisant la génération continue d'une base de données reposant sur des simulations, j'ai entraîné deux réseaux de neurones indépendants, le premier étant dédié à la localisation de plusieurs sources ponctuelles, le second à la localisation d'une source étendue de taille et de forme arbitraires.

Les performances de ces réseaux sont caractérisées sur des données réelles pour des sources ponctuelles disponibles au laboratoire et en exploitant des données simulées pour établir la faisabilité et l'analyse des performances sur la reconstruction de sources étendues (en paragraphe 3.2, une reconstruction de source étendue sur des données réelles est présentée).

Sur les sources ponctuelles, j'ai pu montrer que le réseau de neurones spécifique aux sources ponctuelles présente une sensibilité similaire à l'algorithme MLEM. Autrement dit, le nombre de photons nécessaire pour localiser la source d'intérêt est similaire pour les deux algorithmes. Par ailleurs, sur la spécificité, c'est-à-dire la capacité à ne pas reconstruire de fantômes, le réseau de neurones montre de meilleures performances et reconstruit généralement moins de fantômes que l'algorithme MLEM. Un post-traitement adapté serait nécessaire à l'algorithme MLEM pour effacer ces fantômes sans retirer de vraies sources. Par ailleurs, lorsque le nombre de sources devient plus important, à partir de cinq sources dans le champ de vue par exemple, l'ensemble des algorithmes présente des difficultés pour reconstruire ces sources.

Le réseau de neurones entraîné sur des sources étendues peut aussi être appliqué au cas des sources ponctuelles. Cependant, il a une moins bonne sensibilité que les algorithmes précédents, et dans le cas de la présence de plusieurs sources éloignées, il reconstruit uniquement la source la plus intense, du fait que je n'ai conçu des exemples que dans le cas d'une seule source étendue pour son entraînement.

Dans le cas des sources étendues, les tests que j'ai effectués montrent que les algorithmes classiques et le réseau de neurones entraîné sur des sources ponctuelles sont incapables de fournir une information même partielle sur la localisation de la source, dès lors qu'elle est trop étendue, au-delà d'une taille angulaire typique de 3°x3° dans le champ de vue.

En revanche, le réseau de neurones entraîné sur des sources étendues est capable de traiter ce type de problème, en renvoyant la bonne position de la source avec une taille et une forme cohérentes à celles attendues. J'ai même montré que lorsque je présente un exemple de forme différente à ce que j'ai utilisé dans la base de données, avec une intensité variable à l'intérieur de la source, le réseau de neurones est capable de traiter ce type d'exemple. Par ailleurs, en matière de sensibilité, pour une source étendue et sans bruit de fond, il faut typiquement un millier à une dizaine de milliers de photons pour que l'algorithme renvoie un résultat satisfaisant.

Un problème que l'ensemble des algorithmes a des difficultés à surmonter est la reconstruction correcte de l'intensité relative de chaque source. Ce problème, comme je l'ai montré en Chapitre 4 est lié à l'effet de *vignetting* et à la fraction d'ouverture du masque, ainsi qu'à la redistribution des intensités dans les fantômes. Dans des études ultérieures, il pourra être intéressant d'inclure ces facteurs dans la base de données d'apprentissage des réseaux de neurones.

Un dernier point important est la rapidité d'exécution des différents algorithmes. Les réseaux de neurones et l'algorithme de déconvolution sont aussi rapides à exécuter qu'une seule itération de l'algorithme MLEM, de l'ordre de 0,03 s. Suivant, le nombre d'itérations de MLEM, le temps de calcul est modifié en conséquence, d'un facteur 100 dans mon cas, car je fais 100 itérations de cet algorithme. Cet aspect est important notamment pour le calcul en temps réel.

Le Tableau 2-2 résume les performances comparées des différents algorithmes.

Algorithme	Sensibilité source ponctuelle	Spécificité source ponctuelle (fantômes)	Plusieurs sources ponctuelles	Source étendue	Incertitude	Temps d'exécution	Intensité des sources
Déconvolution	+	–	–	NA	+	+	–
MLEM	+	–	+	NA	NA	–	–
CNN Sources Ponctuelles	+	+	+	NA	+	+	–
CNN Source Etendue	–	+	–	+	+	+	–

Tableau 2-2 : Avantages et inconvénients comparés des différents algorithmes.
+ : Avantage
– : Inconvénient
NA : Non géré par l'algorithme

Un des inconvénients de ma méthode est la distinction entre le réseau de neurones spécifique aux sources ponctuelles et celui conçu pour les sources étendues. Dans le cas d'applications réelles, on peut cependant utiliser l'un ou l'autre des réseaux de neurones en fonction du cas que l'on s'attend à rencontrer, spécifique à l'application, ou encore d'utiliser ces deux réseaux de neurones en parallèle et indépendamment. Par ailleurs, il est aussi envisageable d'effectuer un apprentissage combiné, prenant en compte tous les cas, au prix d'une diminution éventuelle des performances d'apprentissage ou d'une complexification du réseau de neurones.

3 Superposition optique et sources étendues réelles

3.1 Superposition image gamma et image optique

Dans le cas d'applications réelles, les caméras gamma sont équipées d'une caméra optique fournissant une image dans le domaine du visible sur laquelle est superposée l'image gamma. Dans cette partie, je montre comment je peux appliquer ce type de procédé, qui fait partie de la catégorie des « fusions de données » (*data fusion*) du fait de l'agrégation de données provenant de sources hétérogènes, ici une image optique et une image gamma.

J'ai équipé la caméra WIX d'une caméra stéréoscopique ZED (StereoLabs, 2020), comme présenté en Figure 3-1, qui présente l'intérêt de fournir une image optique avec une reconstruction stéréoscopique des distances entre les objets de la scène et la caméra. La caméra peut reconstruire des distances allant de 30 cm à 25 m, du fait de l'écartement de 12 cm entre les deux objectifs optiques. Le champ de vue de chaque objectif optique est de 90°x60°, soit au-delà du champ de vue couvert par le masque codé.

Figure 3-1 : Caméra gamma WIX équipée d'une caméra optique ZED.

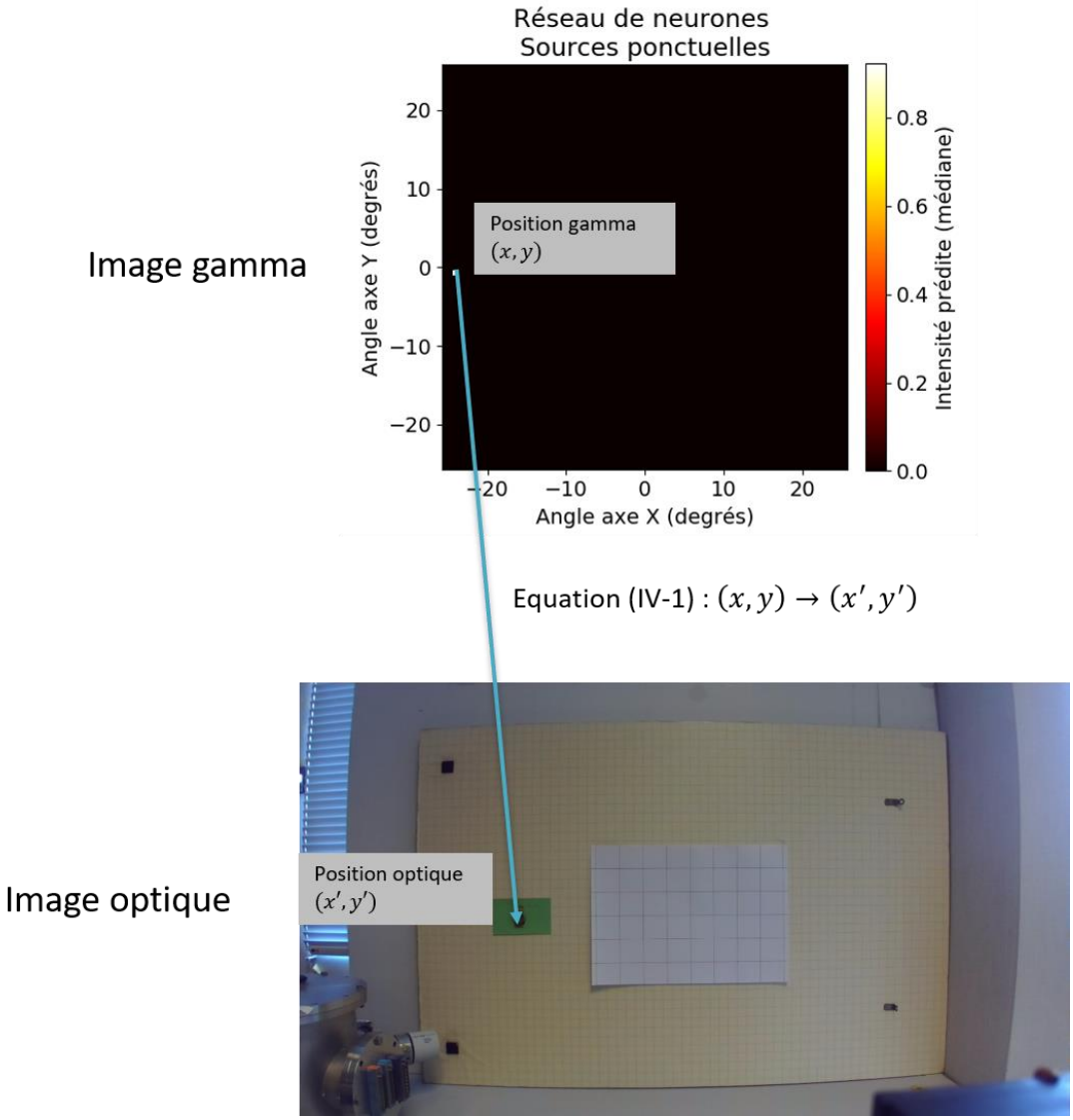


Figure 3-2 : En haut : Résultat du réseau de neurones sur une acquisition d'une source de ^{241}Am de 440 kBq située à 1 m pendant 30 minutes. En bas : Image optique de la scène.

Les deux objectifs optiques sont assez écartés de la position du Caliste, d'une dizaine de centimètres, ce qui implique un effet de parallaxe fort entre l'image gamma et les images optiques fournies par les deux objectifs, effet qui doit être corrigé.

Pour ce faire, j'utilise la méthode suivante. Je note (x, y) les coordonnées d'un point dans l'image gamma et (x', y') les coordonnées d'un point dans l'image optique. Le but est de trouver le lien entre ces deux coordonnées. Je recherche une transformation affine de la forme suivante :

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} m_{xx} & m_{xy} \\ m_{yx} & m_{yy} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} t_x \\ t_y \end{pmatrix} \tag{3-1}$$

Cette transformation est composée d'une matrice $M = \left(m_{ij}\right)_{ij}$ qui représente une transformation pouvant être une combinaison d'une homothétie, d'une rotation, d'une symétrie et d'une distorsion linéaire en x et y. Le vecteur $T = (t_i)_i$ est un vecteur caractérisant une translation. Je discuterai par la suite de la validité de cette transformation.

L'objectif est de déterminer les valeurs des quatre composantes de la matrice M et des deux composantes du vecteur T. Pour cela, j'utilise une méthode de calibration basée sur des mesures réelles. J'effectue une mesure avec une source ponctuelle à une position donnée. Je localise ensuite la source avec mon réseau de neurones spécialisé pour les sources ponctuelles ce qui me donne la position (x_1, y_1) de la source dans l'image gamma. Je fais une prise de vue de la scène avec la caméra ZED et je repère sur cette prise de vue les coordonnées des pixels correspondant à la source, (x_1', y_1'). Je représente un exemple d'une telle scène de calibration en Figure 3-2.

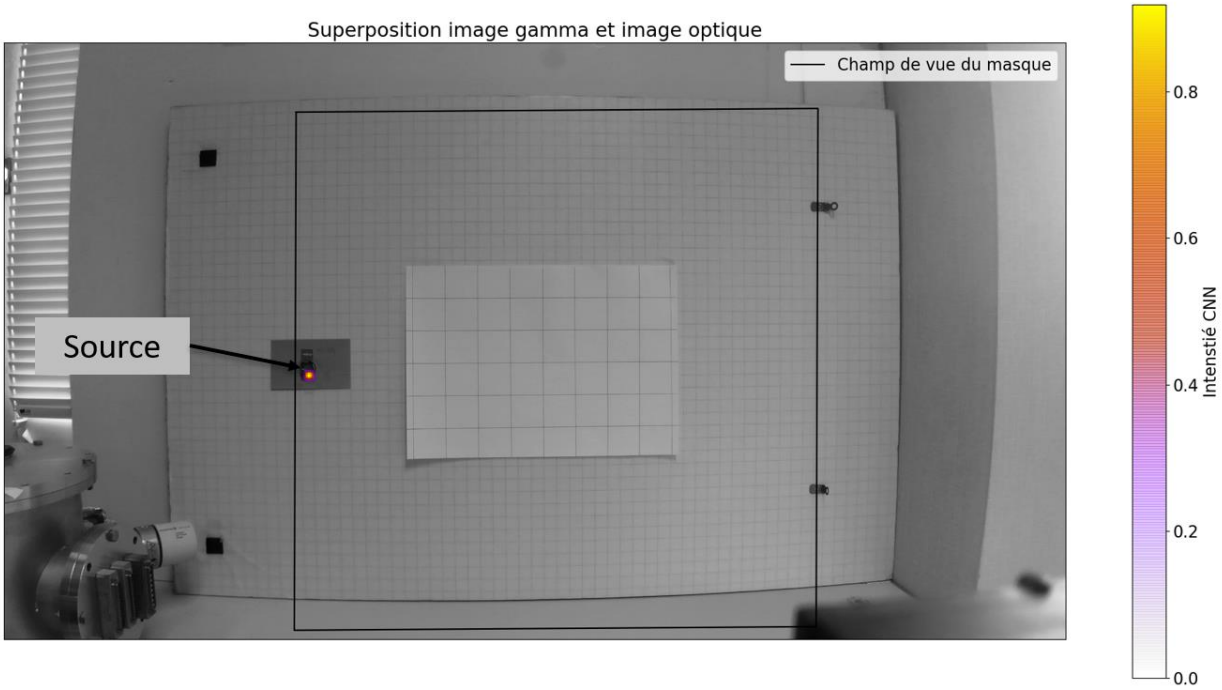


Figure 3-3 : Superposition gamma et optique pour l'exemple présenté en Figure 3-2

Avec les résultats (x_1, y_1) et (x_1', y_1') de cette mesure et l'équation (3-1), j'obtiens deux équations pour résoudre les six inconnues de la matrice M et du vecteur T. Pour obtenir six équations au total, j'effectue deux mesures supplémentaires, ce qui me permet d'aboutir à un système de six équations à

six inconnues. Les trois positions utilisées pour les trois mesures ne doivent pas être alignées, au risque d'avoir un système d'équations liées qui ne permet donc pas de déterminer les six inconnues.

Avec cette méthode, j'obtiens une matrice M et un vecteur T grâce à ces trois mesures. Pour résoudre les équations et obtenir les paramètres* de M et T, j'utilise la fonction *getAffineTransform* de la bibliothèque OpenCV (OpenCV, 2019). En appliquant la superposition obtenue dans le cas de l'exemple de la Figure 3-2, j'obtiens la Figure 3-3.

La Figure 3-4 et la Figure 3-5 présentent les résultats de la superposition optique et gamma pour une acquisition avec les trois sources de ^{241}Am du laboratoire, en utilisant respectivement le réseau de neurones dédié aux sources ponctuelles et l'algorithme MLEM. La sélection en énergie des événements utilisés est comprise entre 57 keV et 62 keV.

On constate en particulier que l'algorithme MLEM reconstruit mieux les intensités relatives attendues de chaque source mais fait apparaître des fantômes plus intenses que le réseau de neurones. Par exemple, la source à 393 kBq est présente avec le réseau de neurones, mais avec une intensité deux fois plus faible que la source à 440 kBq. Cela rejoint les conclusions que nous avions déjà tirées en paragraphe 2.3.

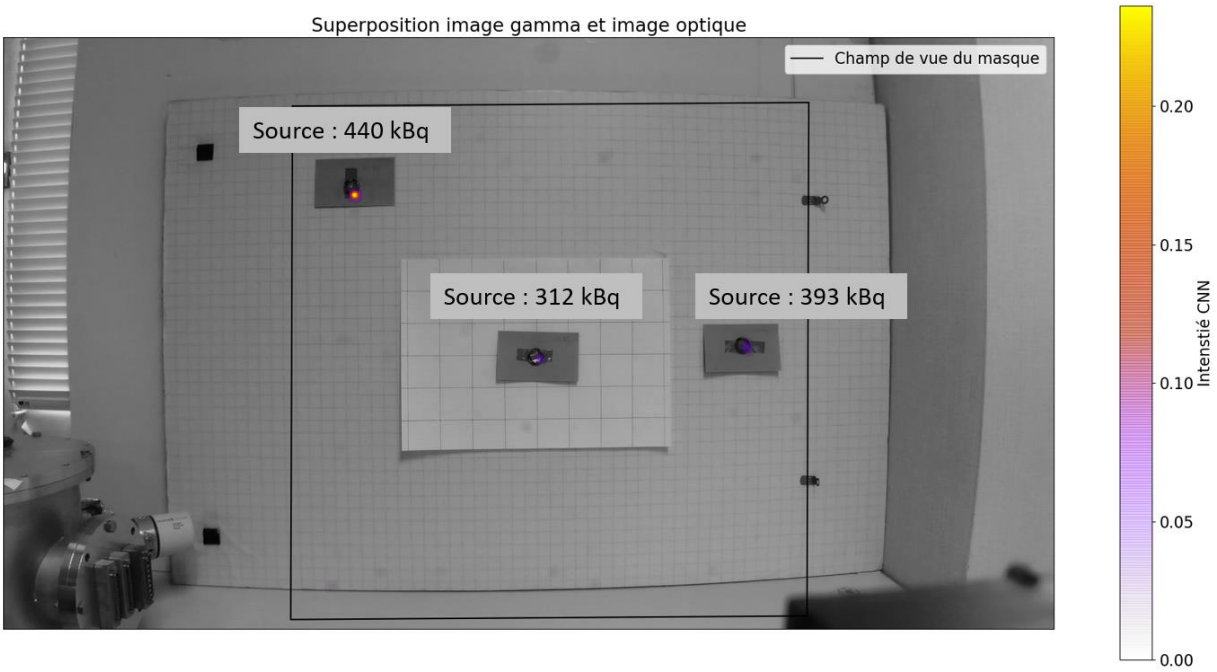


Figure 3-4 : Résultat du réseau de neurones sur l'acquisition avec 3 sources de ^{241}Am pendant 60 minutes.

* De sorte à avoir une approximation plus robuste de ces paramètres, j'ai en réalité effectué quatre mesures, avec des positions sur l'axe vertical du détecteur en haut et en bas, et sur l'axe horizontal du détecteur à gauche et à droite. J'ai appliqué la méthode décrite précédemment sur les quatre combinaisons de trois mesures possibles, de sorte à obtenir quatre matrices M et quatre vecteurs T. J'ai ensuite fait la moyenne de ces quatre matrices M et vecteurs T pour obtenir les paramètres finaux de l'équation (3-1).

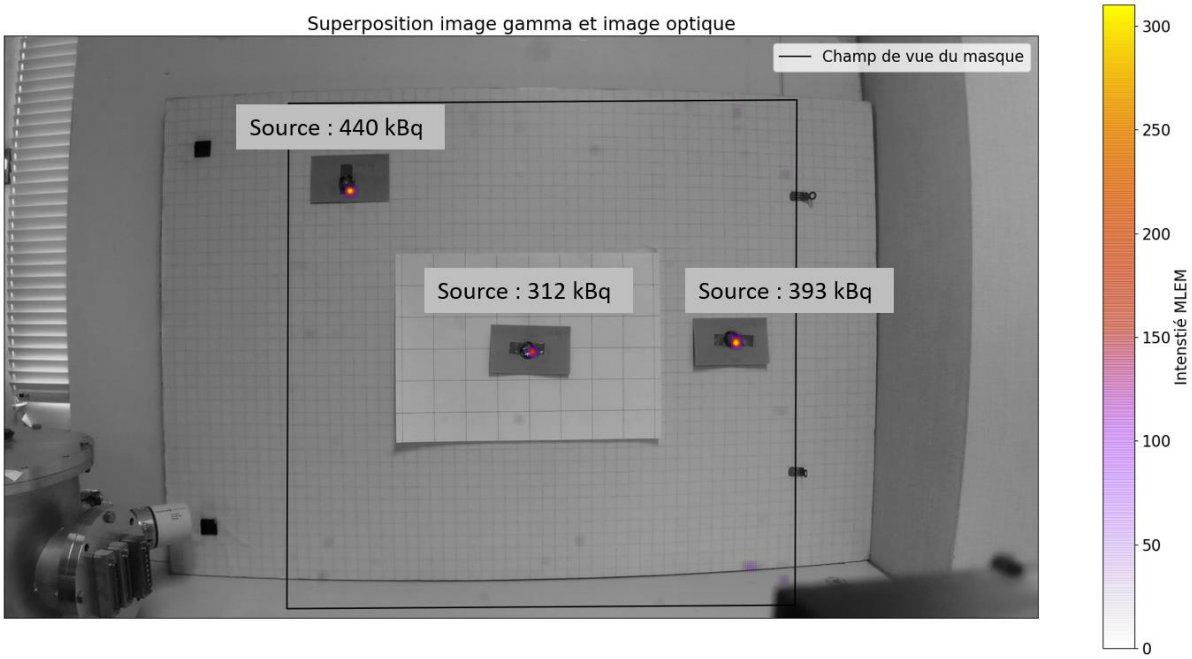


Figure 3-5 : Résultat de l'algorithme MLEM sur l'acquisition avec 3 sources de ^{241}Am pendant 60 minutes.

Par ailleurs, on peut noter que la superposition n'est pas parfaite, en particulier pour la source en haut à gauche. Le petit décalage observé provient en fait de la transformation utilisée en équation (3-1). En effet, en prêtant attention à l'image optique en Figure 3-2, on peut se rendre compte qu'il existe une distorsion qui donne un effet « bombé » de lentille à l'image : le haut et le bas de la planche quadrillée apparaissent courbes au lieu d'apparaître horizontaux. La transformation utilisée en équation (3-1) n'est pas capable de rendre compte de ce type de distorsion non linéaire, ce qui peut expliquer les légers décalages dans la superposition. L'utilisation d'un autre modèle de transformation incluant ce type de distorsion doit être étudiée. Un modèle idéal prendrait aussi en compte le fait que les centres optiques des deux caméras (image gamma et image optique) ne sont pas confondus.

Pour conclure sur cette superposition optique avec des sources ponctuelles, je souligne que la calibration a été effectuée avec des sources situées à 1 m de la caméra gamma, tout comme les tests que je montre ensuite. Du fait des effets de parallaxe, les matrices M et T à appliquer sont différentes si les sources sont situées à d'autres distances. Il faudrait en fait effectuer une calibration de ces paramètres à différentes distances. On peut d'ailleurs connaître cette distance en utilisant la méthode décrite en Chapitre 4. Idéalement, de façon à réduire au maximum cet effet de parallaxe, il faudrait que la caméra optique soit au plus proche du détecteur, à une distance typique de l'ordre du centimètre. Il faut cependant respecter les contraintes de structure du système et en particulier veiller à ce que le masque codé n'interfère pas dans le champ de vue de la caméra optique et réciproquement.

3.2 Cas de sources étendues

Le laboratoire ne possède que des sources ponctuelles, mais il est intéressant de pouvoir vérifier l'application de mes algorithmes sur des sources étendues avec des données réelles. Pour ce faire, j'ai réalisé un système qui « mime » des sources étendues à partir de sources ponctuelles.

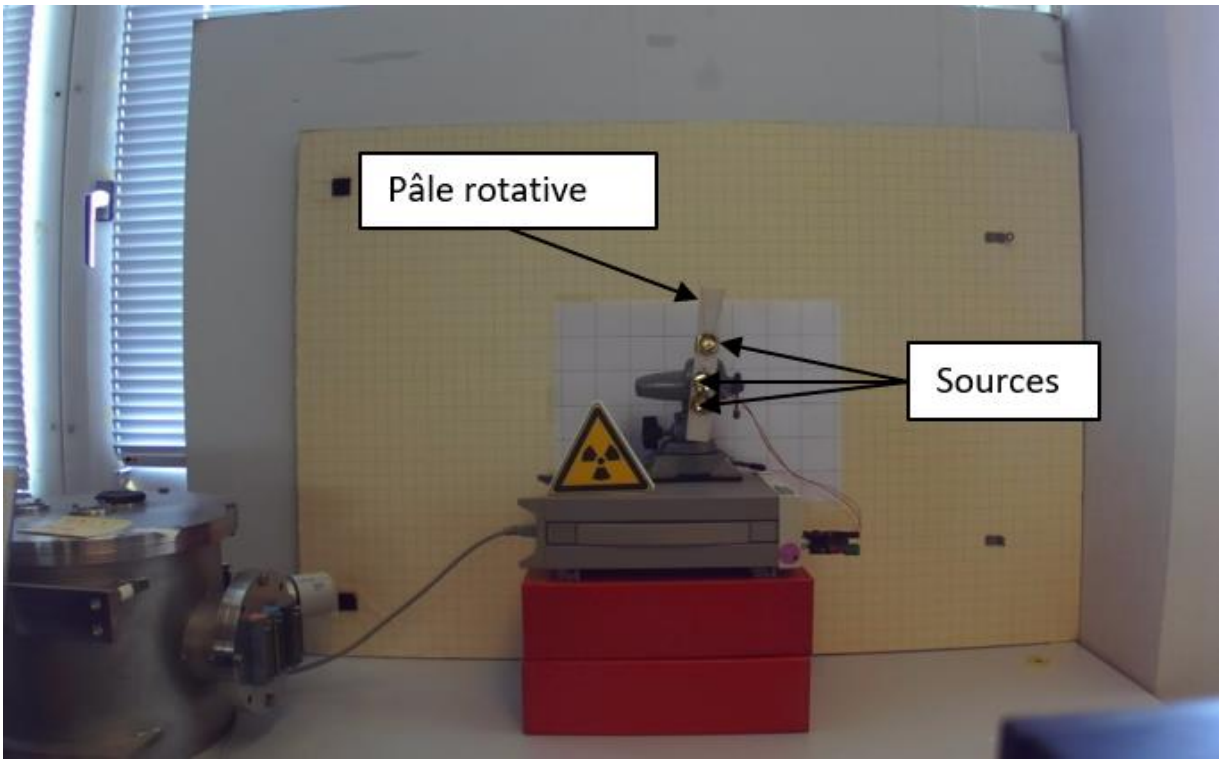


Figure 3-6 : Système rotatif pour mimer des sources étendues.

J'ai fixé les trois sources de ^{241}Am sur un système rotatif, qui tourne à l'aide d'un petit moteur, comme présenté en Figure 3-6. Le système rotatif tourne à une vitesse de 12 tours par minute, ce qui permet de mimer une source étendue en forme de disque, vu que la durée d'acquisition typique que j'utilise est au minimum de 30 minutes.

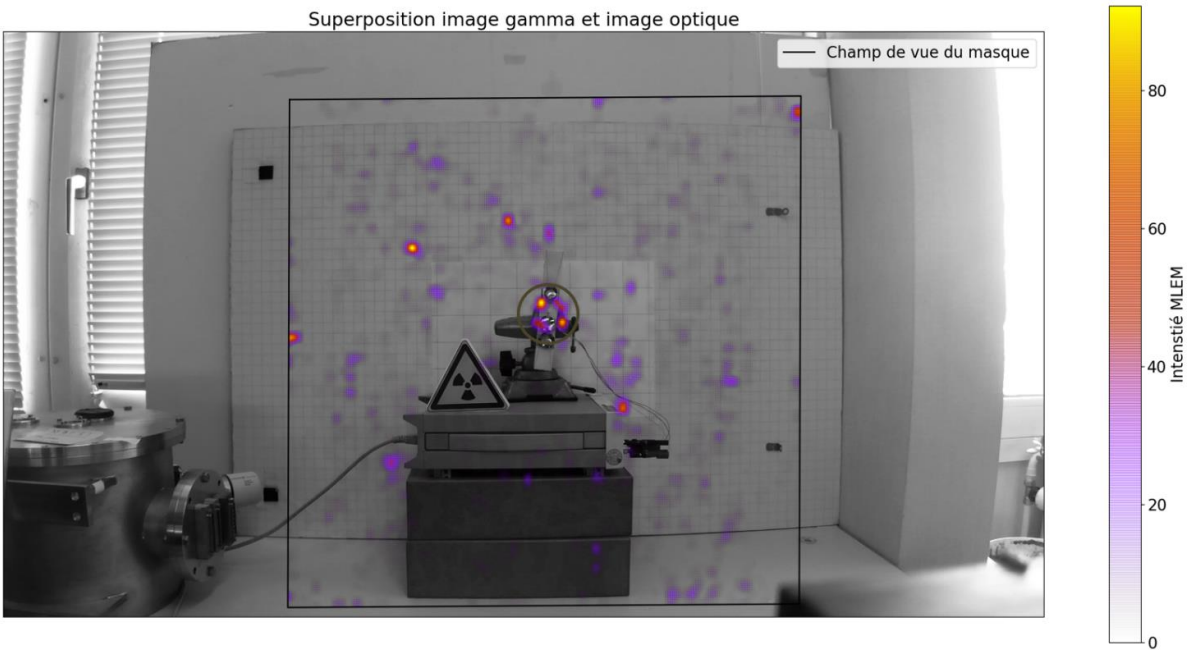


Figure 3-7 : Résultat de l'algorithme MLEM après 2 h 35 mn d'acquisition (configuration 1). Le cercle brun délimite le disque formé par la rotation des sources.

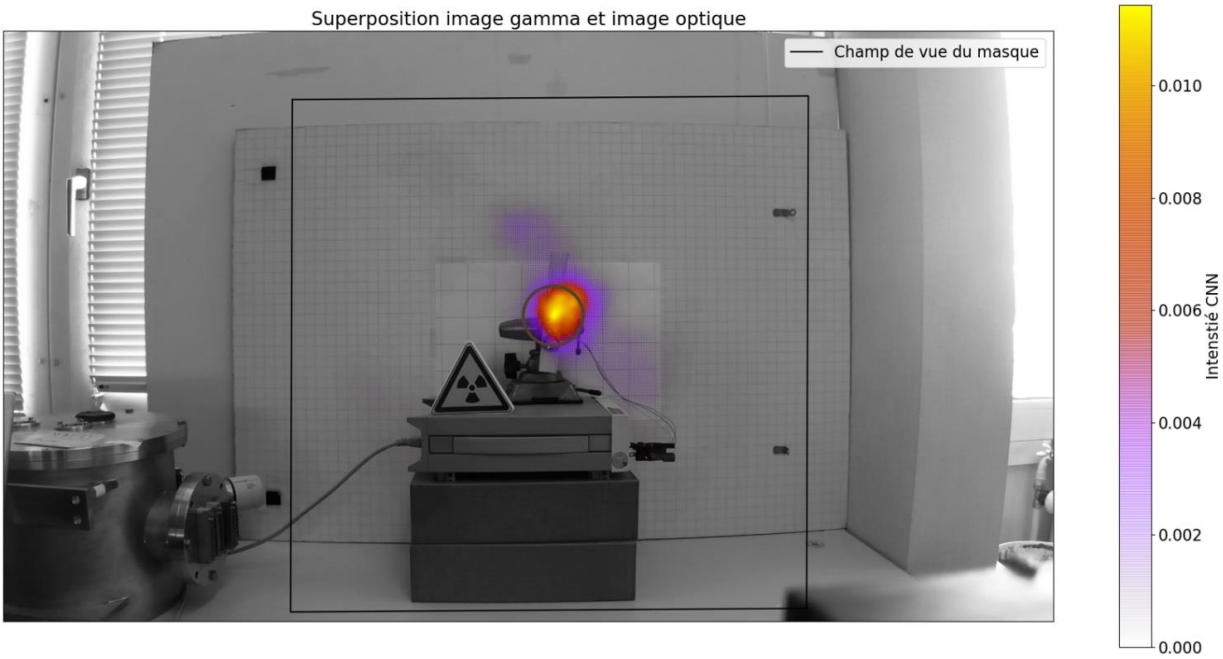


Figure 3-8 : Résultat du réseau de neurones après 30 minutes d'acquisition (configuration 1). Le cercle brun délimite le disque formé par la rotation des sources.

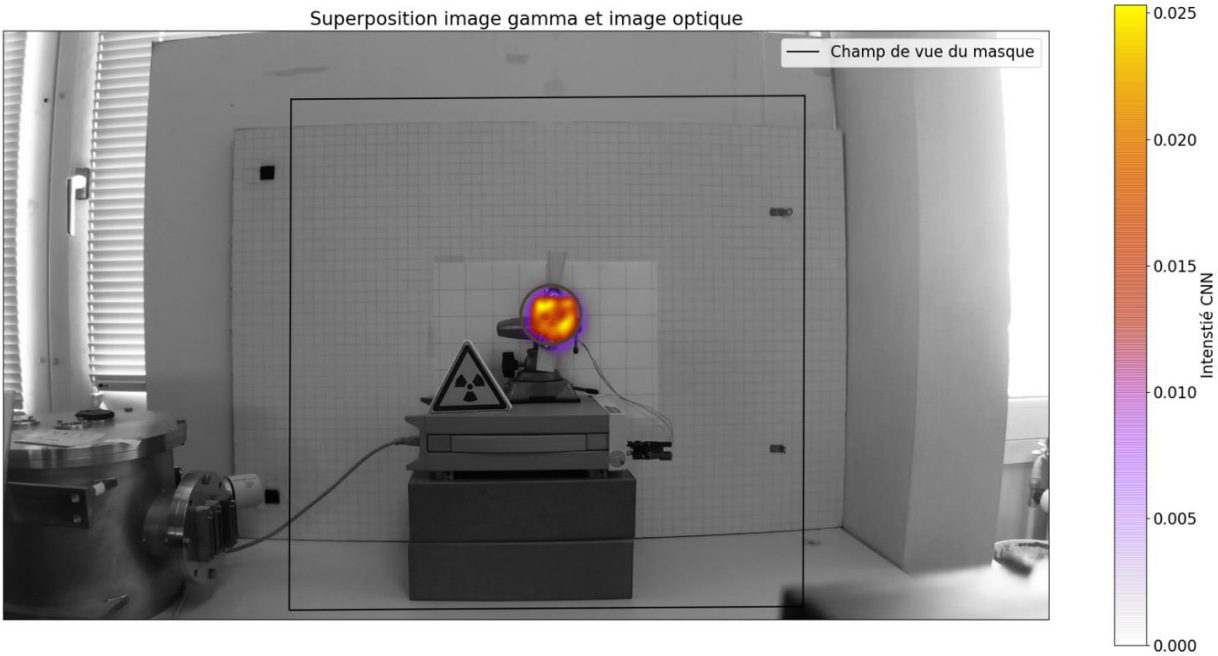


Figure 3-9 : Résultat du réseau de neurones après 2 h 35 mn d'acquisition (configuration 1). Le cercle brun délimite le disque formé par la rotation des sources.

Sur cette configuration, la Figure 3-7 montre la reconstruction effectuée par l'algorithme MLEM avec l'acquisition de 2 h 35 mn que j'ai effectuée, et les Figure 3-8 et Figure 3-9 montrent le résultat avec le réseau de neurones dédié aux sources étendues, pour des durées d'acquisition respectives de 30 mn et 2 h 35 mn. Sur l'ensemble de ces figures, la sélection en énergie est faite entre 15 keV et 62 keV, de sorte à avoir un plus grand nombre de photons, les sources à détecter étant uniquement du ^{241}Am.

On constate qu'avec l'ensemble de l'acquisition, l'algorithme MLEM n'est pas capable de localiser la zone où se situe la source. L'intensité dans cette zone est aussi grande que pour d'autres fantômes dans l'image. En revanche, avec seulement 30 minutes d'acquisition, le réseau de neurones localise la position de la source, avec la présence d'une zone fantôme autour de la zone. Avec les 2 h 35 mn d'acquisition, il est capable de retrouver la position, la forme et la taille de la source, taille légèrement plus petite de 10 % par rapport à la taille attendue.

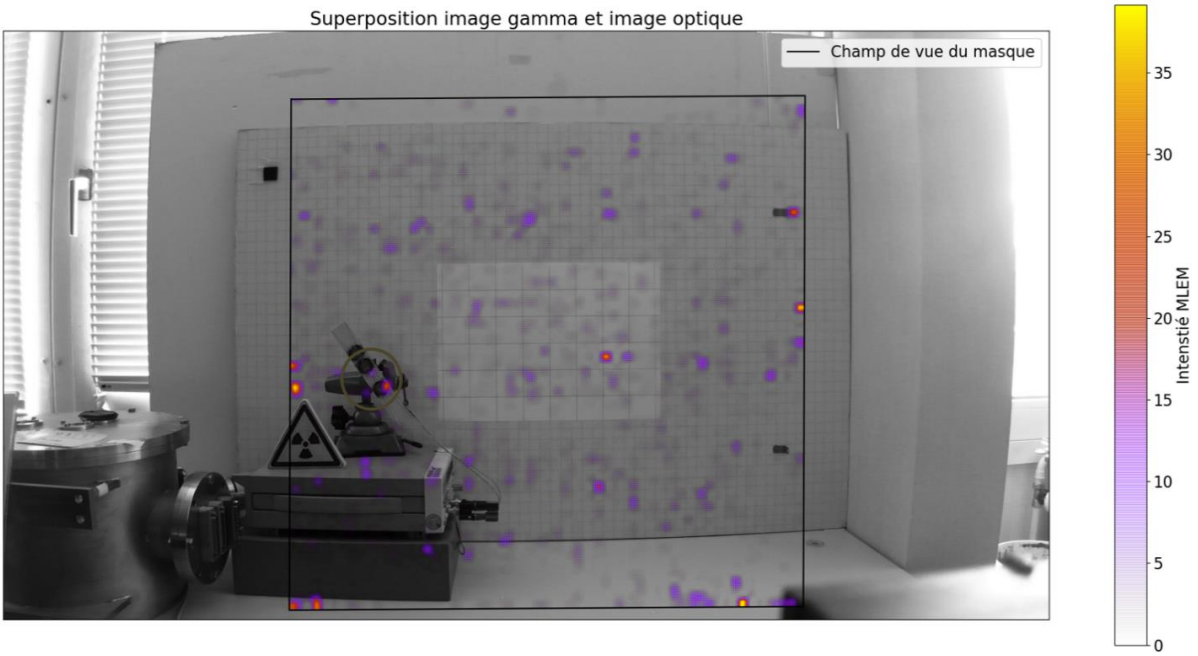


Figure 3-10 : Résultat de l'algorithme MLEM après 1 h d'acquisition (configuration 2). Le cercle brun délimite le disque formé par la rotation des sources.

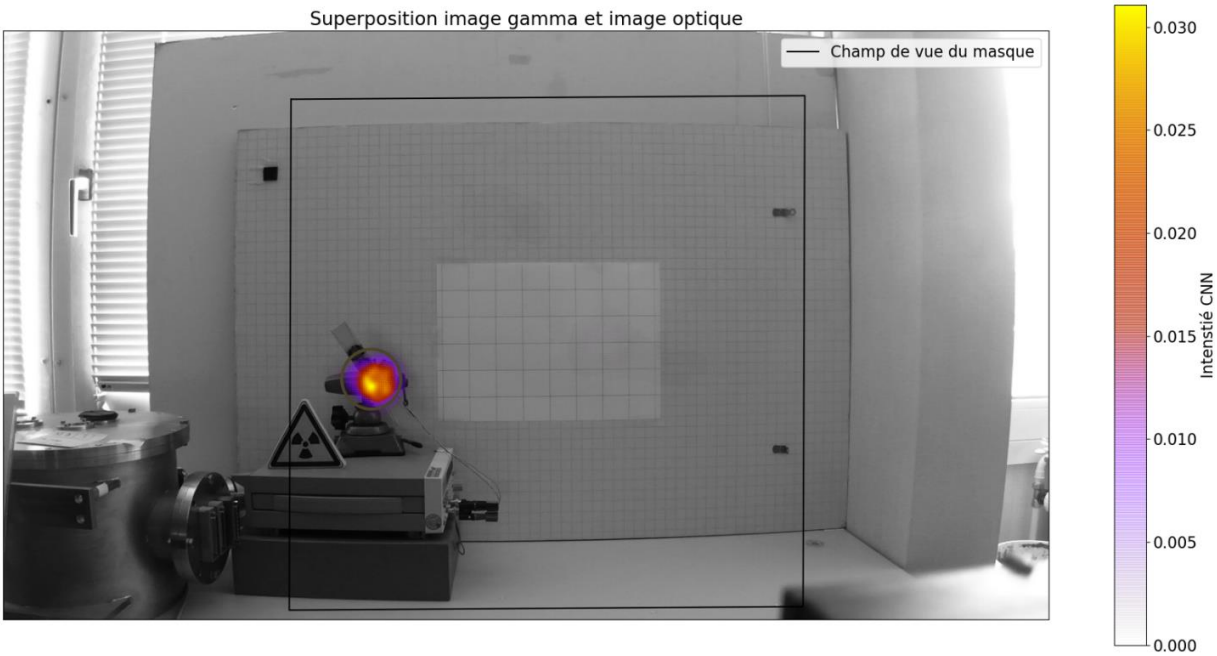


Figure 3-11 : Résultat du réseau de neurones après 1 h d'acquisition (configuration 2). Le cercle brun délimite le disque formé par la rotation des sources.

Les Figure 3-10 et Figure 3-11 montrent les résultats de l'algorithme MLEM et du réseau de neurones dans une configuration où les sources sont plus proches du bord du champ de vue, avec 1 h d'acquisition.

Comme précédemment, MLEM n'est pas capable de localiser la source, contrairement au réseau de neurones qui retrouve la position, la forme et la taille de la source, plus petite de 20 % devant celle attendue.

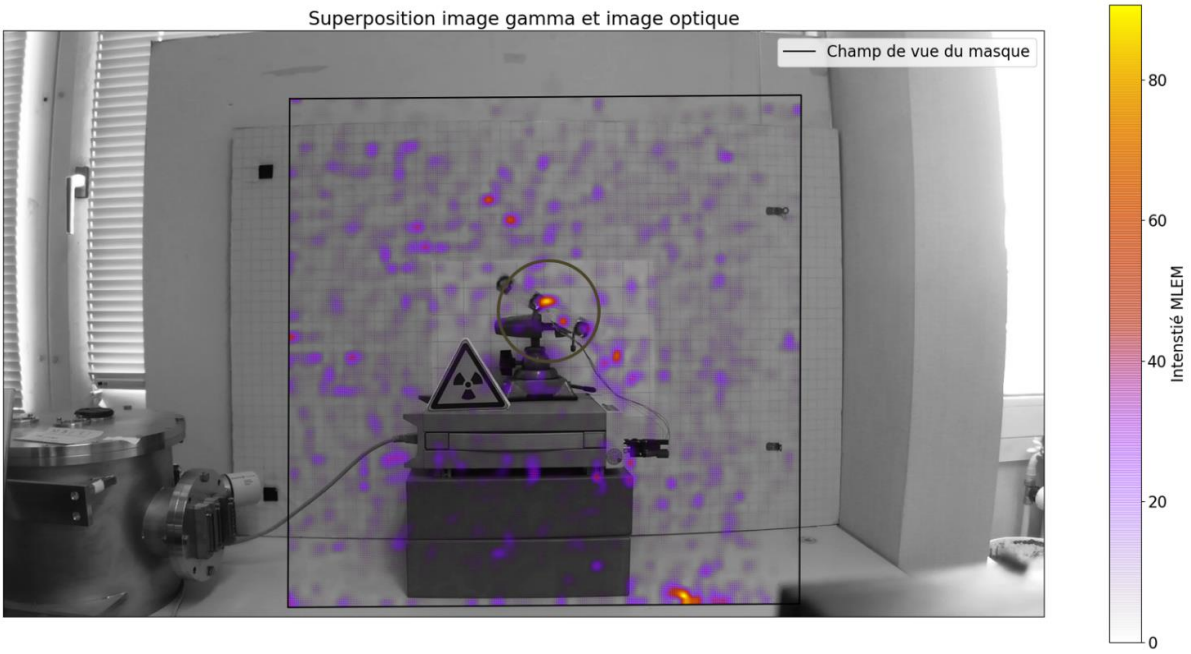


Figure 3-12 : Résultat de l'algorithme MLEM après 7 h d'acquisition (configuration 3). Le cercle brun délimite le disque formé par la rotation des sources.

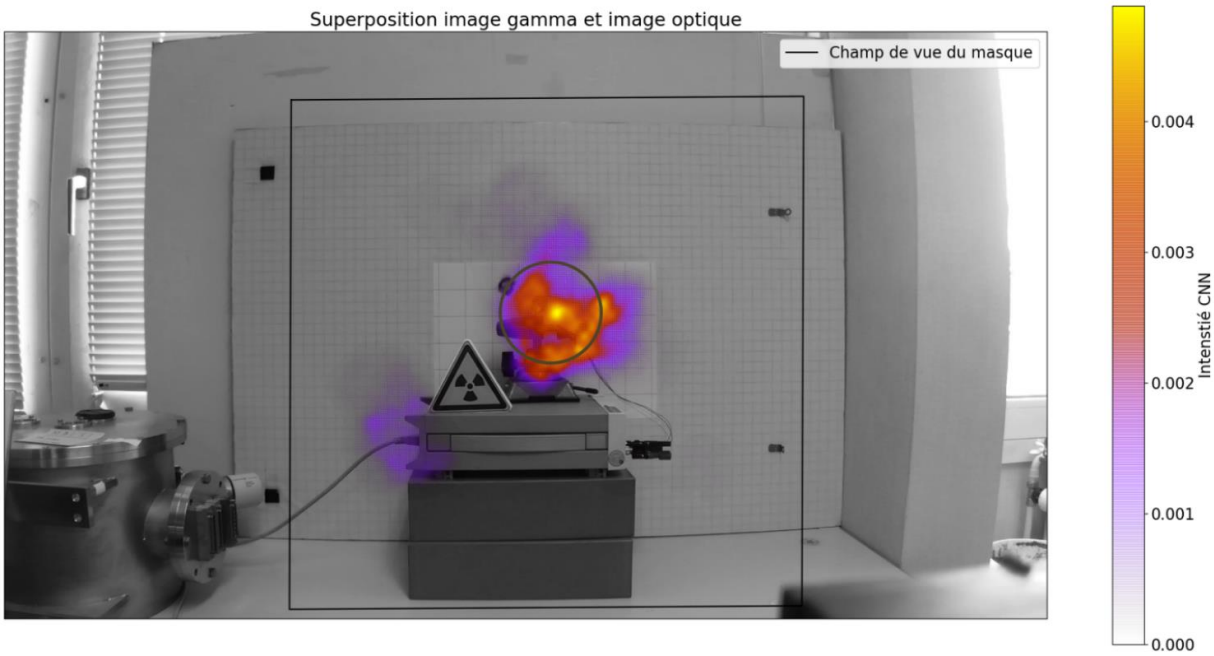


Figure 3-13 : Résultat du réseau de neurones après 7 h d'acquisition (configuration 3). Le cercle brun délimite le disque formé par la rotation des sources.

Les Figure 3-12 et Figure 3-13 montrent un cas où l'écartement des sources est plus important pour former un disque deux fois plus grand.

La durée d'acquisition nécessaire pour le réseau de neurones pour fournir une information sur la position de la source est plus longue, environ 7 h d'acquisition. On voit ici une limite du réseau de neurones car il ne renvoie plus une forme circulaire attendue. La position est aussi plus approximative et la forme renvoyée s'étend de 2° au-delà de la frontière attendue de la source pour la zone la plus intense. La taille renvoyée est 10 % plus grande que la taille attendue. On note aussi la présence d'un léger fantôme en bas à gauche de la source. Cette reconstruction est à mettre en perspective avec l'algorithme MLEM qui ne renvoie pas d'information, même partielle, sur la position, la taille ou la forme de la source.

J'ai testé une dernière configuration pour changer la forme de la source, en masquant une partie du disque avec une brique de plomb de 5 cm d'épaisseur, de forme triangulaire, qui absorbe la totalité des photons d'énergie inférieure à 60 keV. Les résultats des algorithmes MLEM et réseau de neurones après 1 h 30 mn d'acquisition sont présentés en Figure 3-14 et Figure 3-15.

De nouveau, l'algorithme MLEM n'est capable de retrouver ni la position de la source étendue, ni sa taille, ni sa forme. Le réseau de neurones renvoie la position de la source, en dépassant légèrement de 0,5° au-delà de la frontière délimitée par la brique de plomb, en bas à droite. La forme et la taille de la source sont fidèlement retrouvées, à 5 % près.

Il est important de noter que ce type de forme ne faisait pas partie de la base d'apprentissage, ce qui montre une certaine robustesse des réseaux de neurones à traiter des cas qui n'ont pas été présentés en apprentissage.

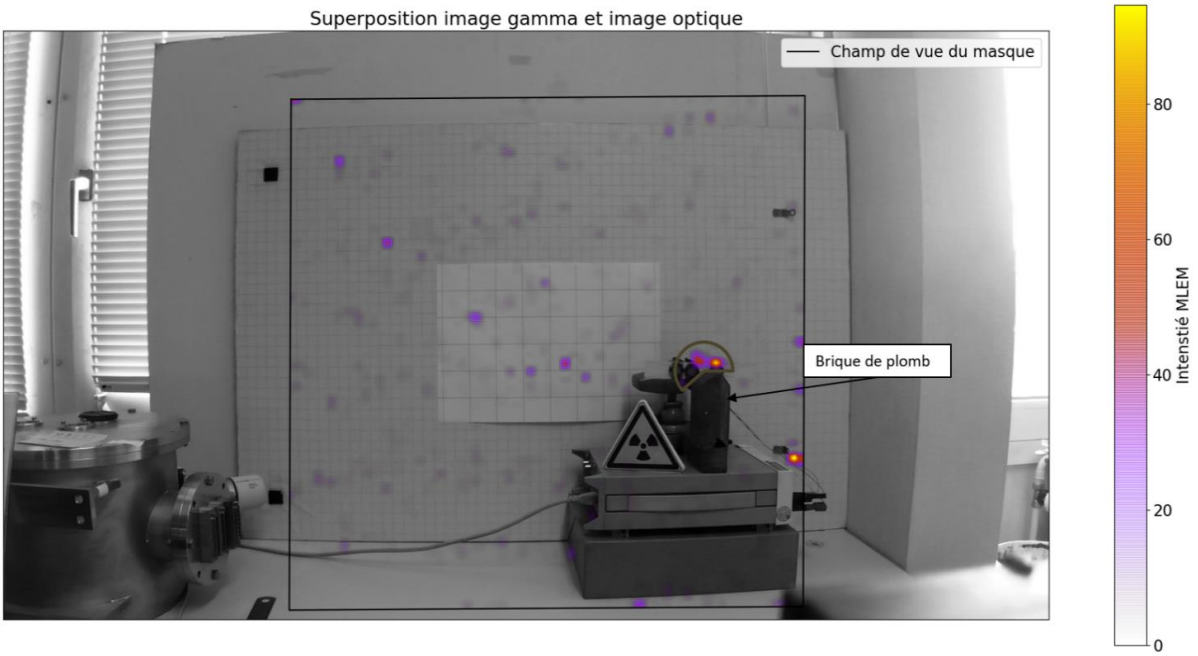


Figure 3-14 : Résultat de l'algorithme MLEM après 1 h 30 mn d'acquisition (configuration 4). La ligne brune délimite la forme donnée par la rotation des sources et masquée en partie par le plomb.

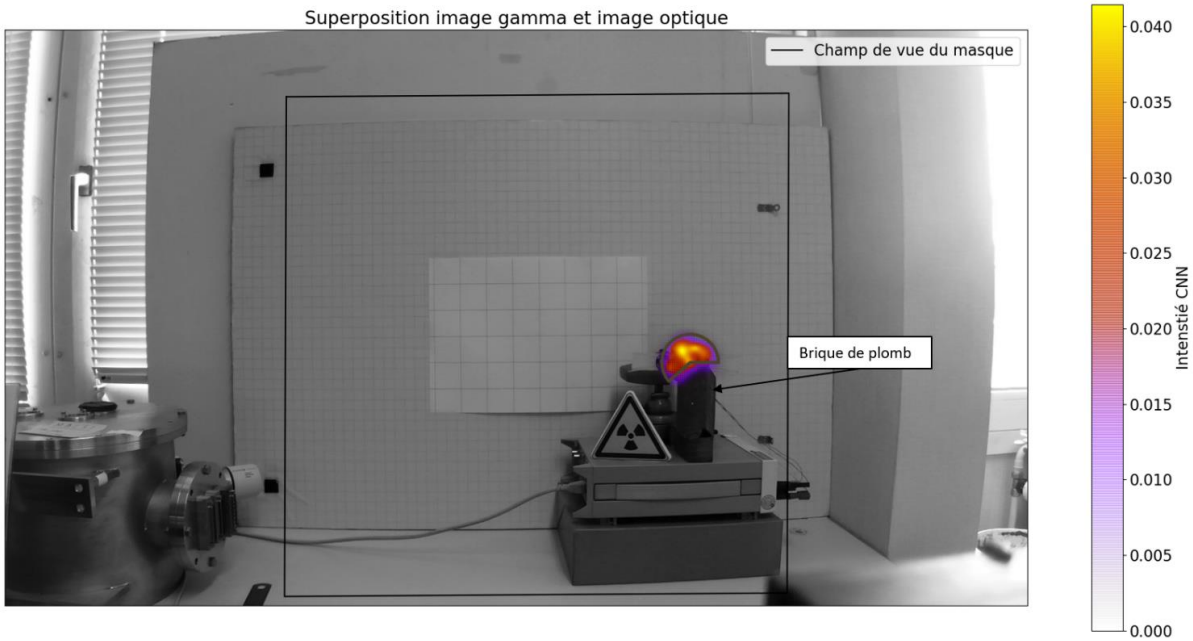


Figure 3-15 : Résultat du réseau de neurones après 1 h 30 mn d'acquisition (configuration 4). La ligne brune délimite la forme donnée par la rotation des sources et masquée en partie par le plomb.

4 Conclusion

À travers ce chapitre, j'ai présenté les méthodes d'imagerie à masque codé appliquées au détecteur Caliste et à la caméra gamma WIX. J'ai notamment montré le soin qui doit être apporté sur les algorithmes de traitement des données, en appliquant dans un premier temps deux algorithmes classiquement utilisés pour la problématique, puis en étudiant la possibilité d'utiliser les réseaux de neurones pour résoudre ce problème.

J'ai montré que, pour des sources ponctuelles, avec ce système de détection, les réseaux de neurones entraînés sur des simulations à repérer des sources ponctuelles obtiennent, sur des données réelles, des performances similaires aux algorithmes classiques, en particulier MLEM, en sensibilité, et de meilleurs résultats en spécificité. En temps d'exécution, les réseaux de neurones sont beaucoup plus rapides que l'algorithme MLEM, du fait qu'il doit effectuer un certain nombre d'itérations pour parvenir au résultat.

Par ailleurs, une limitation des algorithmes classiques concerne leur incapacité à imager des sources étendues, même partiellement, problème qui est surmonté par l'utilisation des réseaux de neurones avec un apprentissage dédié à la localisation de telles sources, dont j'ai pu montrer l'application sur des cas avec des mesures réelles. De plus, les réseaux de neurones, en utilisant le concept de réseaux de neurones bayésiens, fournissent une incertitude sur le résultat qui est renvoyé, contrairement à l'algorithme MLEM.

Il serait intéressant de comparer les performances de ces algorithmes en utilisant d'autres systèmes de détection, en particulier avec un détecteur possédant une plus grande densité de pixels. Pour ce type de système, les capacités d'imagerie attendues sont supérieures à celles de Caliste, au détriment

de leurs performances spectrométriques, mais il serait intéressant d'étudier si les réseaux de neurones améliorent le traitement des données, en particulier en utilisant leur aspect convolutif. De même, le masque codé que j'ai utilisé est un masque aléatoire, et on pourrait aussi appliquer la méthode de *deep learning* que j'ai mise en place à des masques de type MURA par exemple.

Enfin, le modèle du masque lui-même peut aussi être modifié et optimisé selon les applications et en fonction du détecteur utilisé. Pour ce faire, il serait nécessaire d'évaluer les performances attendues en imagerie en fonction du modèle de masque utilisé.

5 Bibliographie

Ballesteros, F., Muro, E. & Luque, B., 2001. *Speeding Up Image Reconstruction Methods in Coded Mask γ Cameras Using Neural Networks: Application to the EM Algorithm.* in: Experimental Astronomy 11, 207–222.

Kingma, D. & Ba, J., 2015. *Adam: A Method for Stochastic Optimization.* in: Conference paper at the 3rd International Conference for Learning Representations, San Diego.

OpenCV, 2019. *Geometric Image Transformations.* [En ligne]
Available at:
https://docs.opencv.org/2.4/modules/imgproc/doc/geometric_transformations.html?highlight=getaffinetransform

StereoLabs, 2020. *Meet Zed.* [En ligne]
Available at: https://www.stereolabs.com/zed/

Van der Walt, S. et al., 2014. *scikit-image: Image processing in Python.* in: PeerJ 2:e453.

Zhang, R. et al., 2019. *Reconstruction method for gamma-ray coded-aperture imaging based on convolutional neural network.* in: Nuclear Inst. and Methods in Physics Research, A, Volume 934, p. 41-51..

Zhang, R. et al., 2020. *Low-noise reconstruction method for coded-aperture gamma camera based on multi-layer perceptron.* in: Nuclear Engineering and Technology.

Ballesteros, F., Muro, E. & Luque, B., 2001. *Speeding Up Image Reconstruction Methods in Coded Mask γ Cameras Using Neural Networks: Application to the EM Algorithm.* in: Experimental Astronomy 11, 207–222.

Kingma, D. & Ba, J., 2015. *Adam: A Method for Stochastic Optimization.* in: Conference paper at the 3rd International Conference for Learning Representations, San Diego.

OpenCV, 2019. *Geometric Image Transformations.* [Online]
Available at:
https://docs.opencv.org/2.4/modules/imgproc/doc/geometric_transformations.html?highlight=getaffinetransform

StereoLabs, 2020. *Meet Zed.* [Online]
Available at: https://www.stereolabs.com/zed/

Van der Walt, S. et al., 2014. *scikit-image: Image processing in Python.* in: PeerJ 2:e453.

Zhang, R. et al., 2019. *Reconstruction method for gamma-ray coded-aperture imaging based on convolutional neural network.* in: Nuclear Inst. and Methods in Physics Research, A, Volume 934, p. 41-51..

Zhang, R. et al., 2020. *Low-noise reconstruction method for coded-aperture gamma camera based on multi-layer perceptron.* in: Nuclear Engineering and Technology.

Chapitre 6 : Localisation de sources radioactives par imagerie Compton

1 **Principe de la reconstruction Compton** _____ **244**

2 **Reconstruction Compton avec Caliste** _____ **249**

3 **Algorithmes de traitement Compton** _____ **265**

4 **Perspectives : vers des algorithmes de Deep Learning** _____ **291**

5 **Conclusion** _____ **296**

6 **Bibliographie** _____ **297**

Une des principales méthodes d'imagerie gamma repose sur le principe de la diffusion Compton, diffusion des photons par des électrons, induisant un changement de direction du photon et un transfert d'énergie du photon vers l'électron. À travers ce chapitre, je montre comment ce principe peut être utilisé pour localiser des sources radioactives.

Dans un second temps, je propose une méthode de sélection des événements d'intérêt parmi l'ensemble des événements enregistrés par le détecteur afin d'appliquer cette méthode d'imagerie. Puis, je montre que l'application directe de cette méthode induit un bruit intrinsèque qui empêche d'obtenir une localisation précise des sources, en tenant compte de la géométrie peu favorable à l'imagerie Compton du système de détection utilisé, qui consiste en l'unique détecteur plan que constitue le détecteur CdTe pixélisé de Caliste. J'ai développé, en tenant compte de cette contrainte de géométrie, des algorithmes de traitement à la fois issus de la littérature et de nouveaux algorithmes. Ces traitements permettent d'effectuer la reconstruction Compton en essayant de réduire au maximum le bruit intrinsèque correspondant.

Je teste la validité de ces algorithmes à l'aide d'acquisitions réelles avec Caliste et la caméra WIX, avec une source de ^{137}Cs et sa ligne principale à 662 keV. Je montre notamment les limitations en sensibilité, demandant des acquisitions longues afin de pouvoir localiser correctement des sources radioactives. Enfin, je propose des solutions afin de surpasser cette limitation.

1 Principe de la reconstruction Compton

1.1 Diffusion Compton et élargissement Doppler

L'imagerie Compton repose sur le phénomène de diffusion inélastique Compton, qui est l'un des principaux modes d'interaction des photons X et gamma de haute énergie avec la matière. C'est même le mode dominant d'interaction dans le CdTe à partir de 265 keV, comme montré en Annexe A.

Lors d'une diffusion Compton, un photon interagit avec un électron de l'atome et lui cède une partie de son énergie. Il part ensuite dans une autre direction, avec le reste de l'énergie : on parle de photon diffusé. Ce principe est illustré en Figure 1-1.

En notant E_0 l'énergie du photon initial, E_1 l'énergie du photon diffusé et E_2 l'énergie transférée à l'électron, la conservation de l'énergie totale donne directement la relation suivante :

$$E_0 = E_1 + E_2 \tag{1-1}$$

La conservation de l'impulsion totale du système permet d'aboutir à la relation entre l'angle de diffusion θ et l'énergie du photon diffusé E_1 :

$$E_1 = \frac{E_0}{1 + \frac{E_0}{m_e c^2}(1 - \cos(\theta))} \tag{1-2}$$

où $m_e c^2$ est l'énergie de masse au repos d'un électron, soit environ 511 keV.

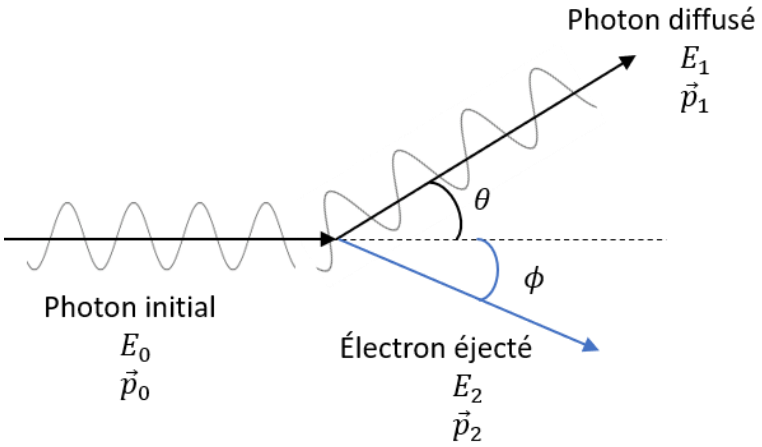


Figure 1-1 : Principe de la diffusion Compton - Cas où l'électron initial est au repos. E_0 et $\vec{p}_0$ sont respectivement l'énergie et l'impulsion du photon incident ; E_1 et $\vec{p}_1$ l'énergie et l'impulsion du photon diffusé ; E_2 et $\vec{p}_2$ l'énergie et l'impulsion transmises à l'électron.

L'équation (1-2), appelée équation de la cinématique Compton, est valable en considérant l'électron initialement au repos. Je démontre cette formule en Annexe A dans le cas général où l'électron n'est pas initialement au repos. La difficulté lorsque l'électron n'est pas initialement au repos est que sa configuration initiale n'est pas connue, en particulier l'impulsion initiale et la direction initiale de l'électron par rapport au photon incident. Or, ces deux paramètres sont nécessaires pour caractériser l'ensemble du système et appliquer la cinématique Compton dans le cas général mais il n'est pas possible de les mesurer. L'approximation usuelle consiste à considérer l'électron comme initialement

au repos et à utiliser l'équation (1-2), en induisant alors une incertitude sur la relation entre E_1 et θ appelée élargissement Doppler.

(Ordonez, et al., 1997) montre le calcul de la section efficace σ_{E_0} de diffusion en fonction de l'énergie de diffusion E_1 et de l'angle de diffusion θ dans le cas général où l'électron est initialement en mouvement :

$$\frac{d^2\sigma_{E_0}}{d\theta dE_1} = \sum_n \frac{mr_0^2}{2E_0}\left(\frac{E_{\text{repos}}(\theta)}{E_0} + \frac{E_0}{E_{\text{repos}}(\theta)} - \sin^2(\theta)\right)\frac{E_1}{\sqrt{E_0^2 + E_1^2 - 2E_0E_1\cos(\theta)}}J_n(p_z)2\pi\sin(\theta) \qquad (1\text{-}3)$$

où r_0 est le rayon classique de l'électron, $J_n(p_z)$ est le profil Compton pour la n-ième sous-couche de l'atome de diffusion*, donné par (Biggs, et al., 1975) et $E_{\text{repos}}(\theta)$ est l'énergie du photon diffusé en considérant l'électron initial au repos, donné par l'équation (1-2).

p_z est la projection de l'impulsion initiale de l'électron sur l'impulsion de l'électron après la collision, donnée par :

$$p_z = -m_e c\frac{\left(E_0 - E_1 - \frac{E_0E_1(1-\cos(\theta))}{m_ec^2}\right)}{\sqrt{E_0^2 + E_1^2 - 2E_0E_1\cos(\theta)}} \qquad (1\text{-}4)$$

J'ai utilisé ces équations en Chapitre 2 pour établir les structures spectrales Compton attendues en intégrant l'équation (1-3) sur l'ensemble des angles possibles. Ici, je vais utiliser l'équation (1-3) afin de prédire, pour différentes énergies E_1 données, la distribution en angle de diffusion θ attendue par élargissement Doppler. Je considère comme énergie E_0 la ligne d'émission principale du ^{137}Cs à 662 keV. J'applique l'équation (1-3) dans le cas de l'étain, Z = 50, matériau modèle utilisé pour représenter le CdTe (Z = 48 et 52 pour le Cd et le Te respectivement).

Le résultat de la distribution de probabilité attendue pour l'angle θ est représenté en Figure 1-2 pour quelques énergies de photon diffusé E_1.

Les distributions présentées en Figure 1-2 permettent de caractériser l'erreur que l'on commet sur la détermination de l'angle de diffusion en fonction de l'énergie du photon diffusé, lorsque l'on utilise la relation (1-2), du fait de l'élargissement Doppler. Ce calcul fournit, pour un angle de diffusion donné, l'incertitude attendue sur cet angle, illustrée en Figure 1-3. Pour ce faire, pour un angle de diffusion θ, je calcule l'énergie du photon diffusé E_1 correspondante avec la relation (1-2), je calcule la distribution associée à E_1 comme présenté précédemment en Figure 1-2. Sur cette distribution, je calcule la largeur totale à mi-hauteur du pic de cette distribution, illustrée ensuite en Figure 1-3. Ce calcul est effectué avec une énergie de photon incident de 662 keV.

L'incertitude liée à l'élargissement Doppler augmente avec l'angle de diffusion. Elle est de l'ordre de 2° pour un angle de diffusion à 90° et elle est inférieure à 5° si l'angle de diffusion est inférieur à 135°. Elle augmente jusqu'à atteindre 24° d'incertitude pour des angles de diffusion au-delà de 175°.

L'élargissement Doppler constitue une limite intrinsèque liée au matériau utilisé en tant que diffuseur et l'erreur commise du fait de cet élargissement montre que la limite que l'on aura sur la détermination

* $J_n = 0$ si l'énergie $E_0 - E_1$ déposée à l'électron est inférieure à l'énergie de liaison de cet électron de la sous-couche n.

de l'angle Compton et donc sur la reconstruction des sources par imagerie Compton est typiquement de l'ordre de 2° pour des angles de diffusion de 90°. C'est par conséquent la résolution angulaire ultime que l'on peut chercher à atteindre en imagerie Compton.

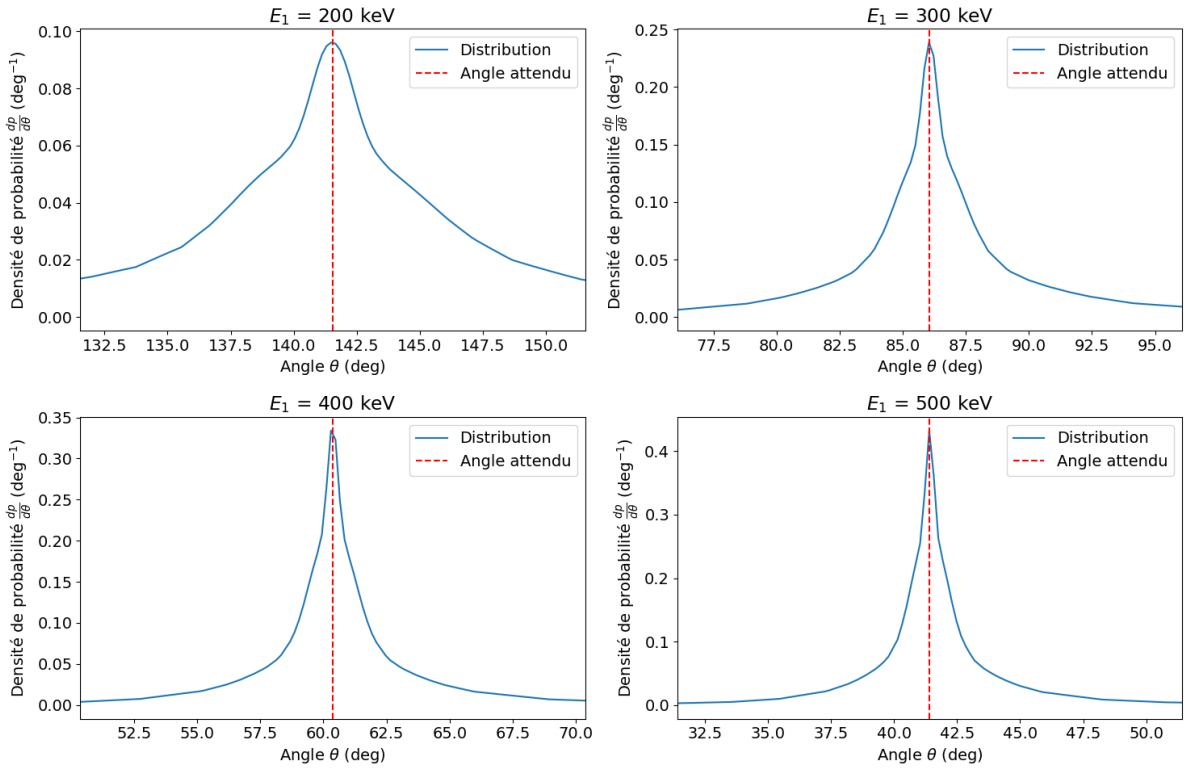


Figure 1-2 : Distribution de probabilité pour l'angle de diffusion. La ligne en pointillés représente l'angle de diffusion attendu lorsque l'électron est initialement au repos. L'énergie du photon incident est 662 keV.

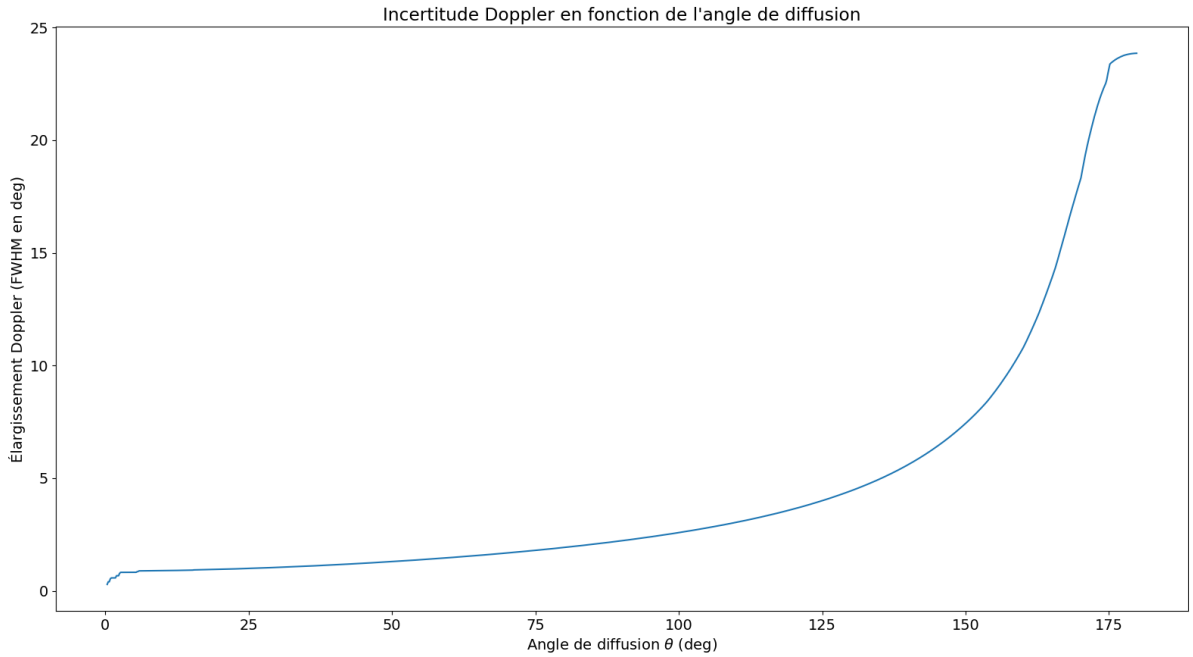


Figure 1-3 : Incertitude liée à l'élargissement Doppler en fonction de l'angle de diffusion, pour un photon d'énergie initiale de E_0 = 662 keV.

La Figure 1-4 montre un exemple de distribution pour une énergie E_1 de 185,4 keV, correspondant à un angle de diffusion de 170° d'un photon de E_0 = 662 keV. On voit effectivement que la distribution est particulièrement élargie, non symétrique et elle n'est pas centrée sur l'angle attendu de 170°. Cela illustre les limites de l'utilisation de l'équation (1-2) pour des grands angles de diffusion.

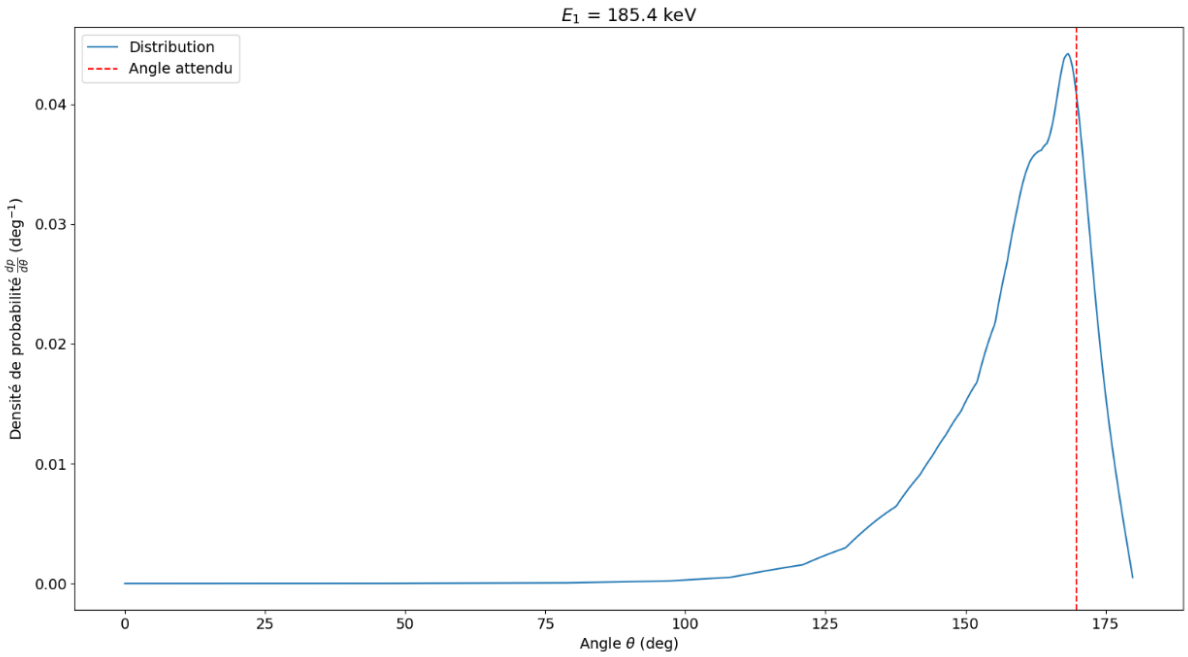


Figure 1-4 : Incertitude liée à l'élargissement Doppler pour un angle de diffusion de 170°.

1.2 Application de la diffusion Compton à la localisation de sources radioactives

Comme présenté au Chapitre 1, plusieurs caméras gamma utilisent le principe de la diffusion Compton pour effectuer la localisation de sources radioactives (Takeda, et al., 2012) (H3D, 2020) (NUVIATECH Instruments, 2020). Ce mode d'imagerie présente l'intérêt qu'il peut être appliqué à des photons de haute énergie, au-delà de 250 keV, pour des domaines d'énergie où la diffusion Compton est le mode d'interaction dominant. Les méthodes d'imagerie à masque codé nécessitent des photons de plus faible énergie, interagissant principalement par absorption photoélectrique. La reconstruction Compton est réputée sensible car elle repose sur la détection en coïncidence temporelle d'au moins deux événements dépendants. La probabilité pour deux événements indépendants en coïncidence temporelle de mimer une diffusion Compton est faible et vaut $n_1.n_2.\tau$, où n_1 et n_2 sont les taux de comptages des éléments de détection et τ la fenêtre temporelle considérée pour une coïncidence.

L'imagerie Compton est effectuée de la manière suivante. J'ai montré au paragraphe 1.1 que l'angle de diffusion θ et l'énergie du photon diffusé E_1 sont liés par la relation (1-2), à l'élargissement Doppler près. De même, l'énergie du photon diffusé E_1 et l'énergie transmise à l'électron E_2 lors de la diffusion Compton sont reliées par l'équation (1-1). L'angle de diffusion θ est donc connu par la mesure de l'une de ces deux énergies E_1 ou E_2. Notons que la précision sur l'angle θ est directement liée à la résolution spectrale du détecteur, dans la limite de l'élargissement Doppler.

La direction du photon diffusé peut aussi être connue dans le cas où ce photon diffusé interagit par la suite dans le système de détection. Le cas le plus simple est celui où le photon diffusé interagit par

absorption photoélectrique, cas que je vais expliciter ensuite. Certains systèmes sont basés sur l'utilisation de plusieurs interactions Compton successives dans le détecteur (Kroeger, et al., 2001).

Lorsque le photon interagit d'abord par diffusion Compton dans un élément détecteur puis par absorption photoélectrique dans un second élément détecteur, ou dans un autre détecteur du même système de détection, les positions des deux interactions sont connues. Je les note respectivement X_{Comp} et X_{phot}. Elles permettent d'identifier la direction de diffusion, donnée par la droite $(X_{\mathrm{Comp}}X_{\mathrm{phot}})$. L'angle de diffusion θ étant connu par les énergies déposées, la provenance du photon est identifiée par le cône de demi-axe $[X_{\mathrm{phot}}X_{\mathrm{Comp}})$ et d'angle au sommet θ. En Figure 1-5, à gauche, je présente ce principe qui permet de reconstruire un cône Compton.

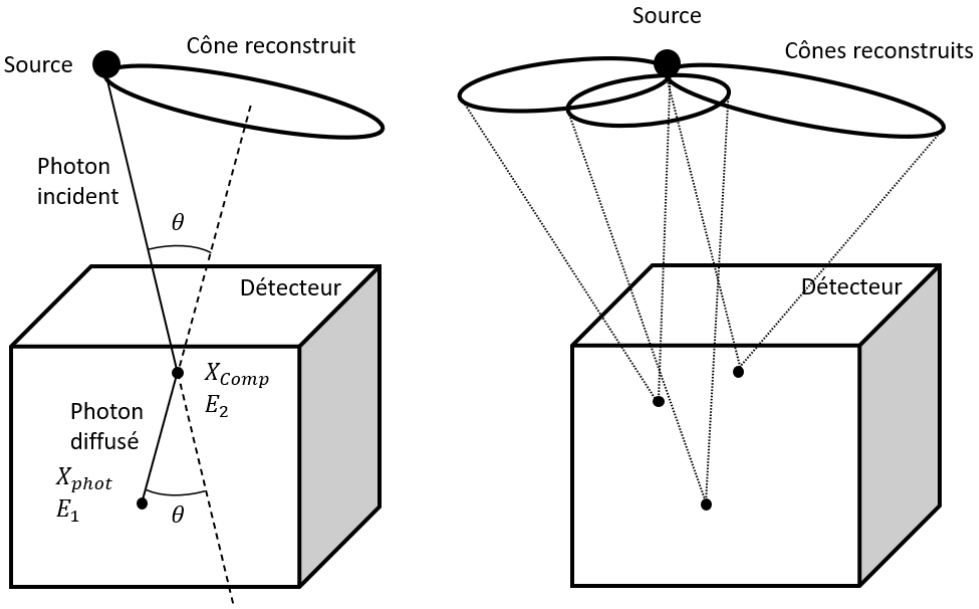


Figure 1-5 : À gauche : principe de la reconstruction Compton avec un seul photon. À droite : reconstruction Compton utilisant plusieurs cônes.

La détection de plusieurs photons permet de tracer les cônes associés à chaque photon, comme illustré en Figure 1-5, à droite. Avec un détecteur parfait, c'est-à-dire fournissant une connaissance parfaite des positions des interactions et énergies déposées, et sans élargissement Doppler, trois cônes sont théoriquement nécessaires pour connaître la position d'une source ponctuelle. Cependant, les effets liés à la résolution du détecteur empêchent d'atteindre ces performances et la sensibilité, c'est-à-dire le nombre minimal de cônes nécessaire pour effectivement localiser une source, est une problématique majeure de la reconstruction Compton.

Les cônes peuvent théoriquement être tracés en trois dimensions ce qui permettrait de localiser la source dans l'espace, avec la distance, mais de nouveau, les limitations liées au système de détection sont souvent trop importantes pour avoir cette évaluation en trois dimensions. De ce fait, la reconstruction des images se fait usuellement en deux dimensions, en recherchant la direction de la source dans l'espace et en projetant les cônes sur une sphère à une distance donnée. Je décris ce processus, appelée *backprojection* directe, en paragraphe 2.3 et je montre les limites de cette approche seule. En partie 3, je présente l'intérêt de développer des méthodes de traitement avancées, basées sur ce principe, dans le cas d'application à des données réelles.

Il existe deux types principaux de systèmes de détection adaptés à l'imagerie Compton. D'un côté, on trouve les détecteurs constitués d'un unique cristal épais, dans lequel les positions d'interaction peuvent être mesurées en trois dimensions, comme dans le cas de la caméra gamma Polaris (H3D, 2020). C'est ce type d'exemple que j'ai représenté en Figure 1-5. De l'autre, il existe les systèmes de détection constitués de plusieurs couches de détection, comme par exemple la caméra Astrocam (Takeda, et al., 2012). Dans ces systèmes, constitués de deux couches a minima, les premières couches de détection sont composées d'un matériau à faible numéro atomique, comme le Silicium, Z = 14, pour favoriser la diffusion Compton des photons. Les couches plus profondes sont constituées d'un matériau à plus haut numéro atomique, comme le CdTe, pour favoriser l'absorption photoélectrique.

Dans le cadre de mes recherches, j'applique la reconstruction Compton à la caméra WIX basée sur la technologie Caliste. Je me place ainsi dans le cas où le détecteur est constitué d'un unique cristal, pouvant mesurer la position des interactions en deux dimensions seulement, la profondeur d'interaction n'étant pas connue. L'énergie est mesurée avec de bonnes capacités spectrométriques, comme présenté en Chapitre 1. Cette configuration, sans connaissance de la profondeur des interactions, est a priori défavorable pour une reconstruction Compton précise, mais je montrerai à travers ce chapitre les performances qui peuvent être atteintes en utilisant des algorithmes de traitement avancés, algorithmes qui peuvent être adaptés à d'autres systèmes de détection que Caliste.

2 Reconstruction Compton avec Caliste

Dans cette partie, je montre la manière dont j'effectue la reconstruction Compton avec des données réelles acquises par la caméra gamma WIX. Dans un premier temps, je présente la sélection des événements d'intérêt que j'applique pour la reconstruction. Puis, je montre la problématique de la détermination de l'ordre d'interaction entre la diffusion Compton et l'absorption photoélectrique et la manière dont je la résous à travers des simulations du détecteur. Enfin, j'étudie l'application de la *backprojection* directe (DBP pour *Direct BackProjection*) et j'en montre les limites et la nécessité d'utiliser des algorithmes de traitement avancés.

2.1 Sélection des événements Compton

Un soin particulier doit être apporté à cette sélection pour éviter d'utiliser des événements qui ne sont pas exploitables pour la reconstruction Compton et ainsi éviter de générer du bruit supplémentaire dans l'image reconstruite. Cela améliore in fine la sensibilité liée à l'imagerie Compton.

La première règle pour sélectionner les événements d'intérêt est la coïncidence temporelle des événements. En effet, je dois rechercher un événement constitué d'une diffusion Compton et d'une absorption photoélectrique, ayant eu lieu chacun dans des pixels différents[*]. Vu la distance parcourue par le photon dans le détecteur, typiquement de l'ordre de quelques millimètres à un centimètre, le temps nécessaire au photon pour parcourir cette distance est typiquement de l'ordre de 0,03 ns, soit 10 000 fois plus petite que la résolution temporelle de Caliste. C'est la raison pour laquelle, je peux

[*] Si les deux interactions ont lieu dans le même pixel, il n'est pas possible de les distinguer et cet événement est alors inutilisable.

considérer que les deux interactions se produisent simultanément, et je recherche donc des événements multiples enregistrés par Caliste.

La deuxième règle de sélection des événements est une sélection spatiale. Les événements que je recherche ne doivent pas être confondus avec du partage de charge ou de la fluorescence X. C'est pourquoi je dois rechercher des événements multiples qui ont eu lieu en étant séparés d'au moins un pixel de sorte à éviter cette confusion. Les événements multiples peuvent éventuellement être constitués de deux clusters de pixels, non voisins. La Figure 2-1 présente les différents cas qui peuvent se présenter pour les événements multiples et ceux que je sélectionne.

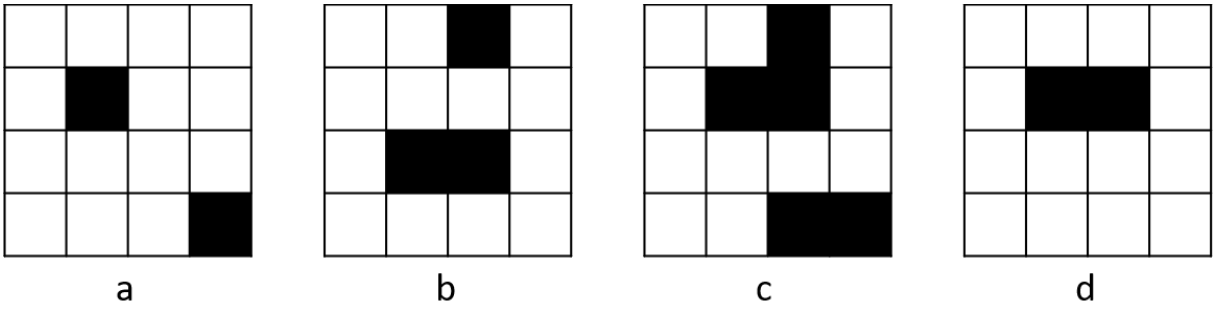


Figure 2-1 : a : Deux pixels non voisins, accepté. b : Un cluster et un pixel non voisins, accepté. c : Deux clusters non voisins, accepté. d : Deux pixels voisins ou un seul cluster, rejeté.

La troisième règle de sélection est une sélection en énergie. Sauf mention contraire, je suppose dans l'ensemble de ce chapitre que l'énergie E_0 initiale des photons avant interaction est connue. Par exemple, on peut avoir identifié les radioéléments dans la scène par spectrométrie, comme présenté au Chapitre 3, puis on essaie de les localiser par imagerie Compton. Comme je considère une diffusion Compton suivie d'une absorption photoélectrique, cela signifie que l'ensemble de l'énergie E_0 a été déposée à travers ces deux interactions, et la conservation de l'énergie (1-1) implique que la somme des deux énergies enregistrées doit être égale à E_0.

Notons qu'il est possible que le photon initial soit diffusé par interaction Compton dans l'environnement du détecteur, ce qui signifie que son énergie après cette interaction est modifiée et donc ce photon est rejeté par la sélection en énergie. Cela est justement intéressant car si le photon a déjà subi une interaction Compton avant d'entrer dans le détecteur, sa direction a préalablement été modifiée et elle ne correspond plus à celle de la source radioactive.

La Figure 2-2 présente cette sélection en énergie dans le cas d'une mesure réelle de ^{137}Cs avec Caliste. Je représente le spectre des sommes des énergies des événements multiples et je sélectionne les événements dont la somme est comprise dans une fenêtre autour de l'énergie E_0 d'intérêt, ici 662 keV. Je distingue notamment sur ce spectre les différentes configurations d'événements multiples non voisins décrites en Figure 2-1.

En réalisant le spectre directement, comme représenté à gauche en Figure 2-2, on note un décalage du pic qui devrait se situer à 662 keV. Ce décalage est d'ailleurs plus important pour les événements avec des clusters (type c) que pour les événements avec un seul cluster (type b) et encore plus important que pour les événements constitués de deux pixels non voisins (type a). Ce phénomène s'explique d'une part par la perte d'énergie due à l'efficacité d'induction de charges. La diffusion Compton et l'absorption photoélectrique correspondante peuvent s'effectuer dans des zones

profondes du détecteur, là où la perte de charge est plus importante, et cette perte de charge s'accumule dans la somme des énergies mesurées pour chaque pixel. D'autre part, lorsque l'événement est déposé dans un cluster de pixels voisins, cela signifie qu'il y a un partage de charges entre les pixels et que l'interaction a lieu dans la zone inter-pixel. Dans cette zone, il existe aussi un phénomène de perte de charge, mis en évidence par (Dubos, 2015), c'est pourquoi le décalage est plus marqué pour les événements contenant un ou plusieurs clusters (types b et c).

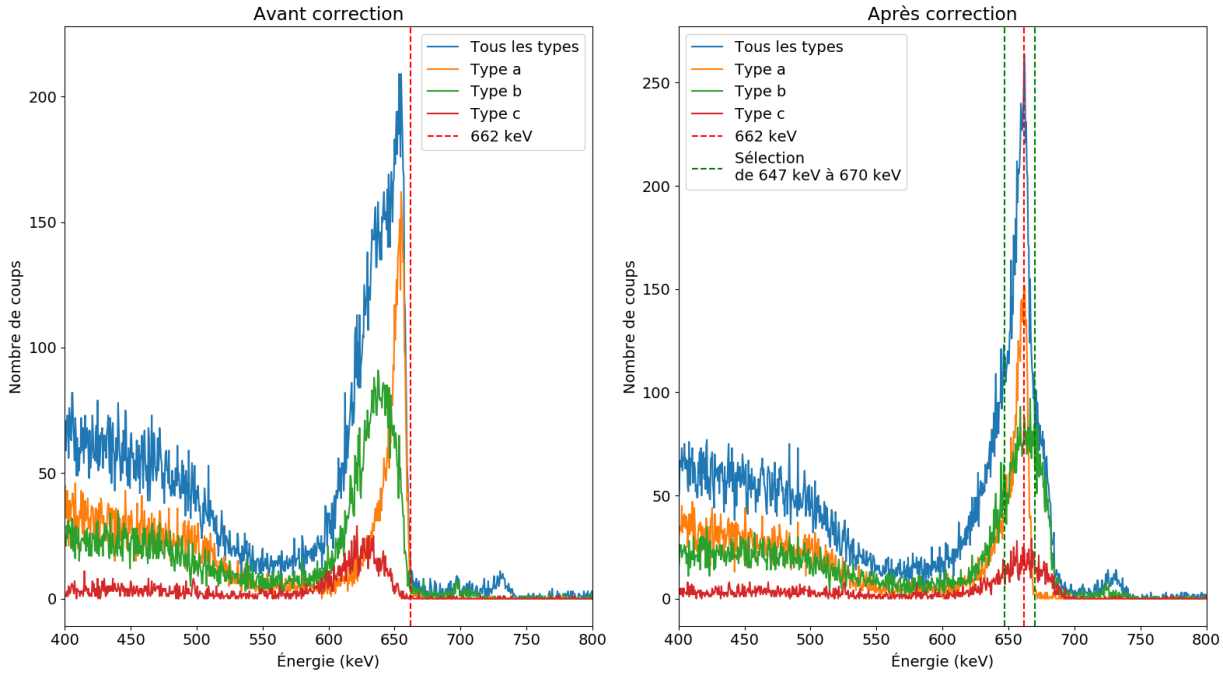


Figure 2-2 : Spectre des événements multiples d'une acquisition avec Caliste-HD d'une source de ^{137}Cs de 1,6 MBq située à 30 cm de la caméra pendant 57 h d'acquisition. Conditions d'acquisition : haute tension 300 V, température -10°C, temps de *peaking* : 6,07 µs. À gauche : spectre direct, selon les différents types distingués en Figure 2-1. À droite : spectre corrigé et recalé en énergie.

La correction en énergie que j'effectue consiste à appliquer un facteur multiplicatif sur le gain de la chaîne pour compenser la perte sur les énergies mesurées. Dans l'exemple présenté ici, un facteur 1,01 est appliqué pour les événements de type a, 1,04 pour les événements de type b et 1,06 pour ceux de type c. Une fois les événements corrigés en énergie, je sélectionne ceux compris dans la largeur à mi-hauteur du pic à 662 keV pour l'ensemble des événements, ici entre 647 keV et 670 keV.

L'ensemble de ces sélections (coïncidence temporelle, pixels non voisins et sélection en énergie) assure la meilleure sélection des événements en isolant ceux qui correspondent au critère d'être une interaction Compton suivie d'une absorption photoélectrique et ainsi cela minimise le bruit en rejetant les événements qui ne correspondent pas à ces critères.

2.2 Ordre d'interaction et simulations

2.2.1 Problématique de l'ordre d'interaction

Pour déterminer l'angle de diffusion θ, j'ai montré en partie 1.1 que nous pouvons utiliser la cinématique Compton (1-2). Dans le principe de la reconstruction Compton, E_1 correspond alors à l'énergie déposée par absorption photoélectrique du photon diffusé.

Afin d'appliquer cette méthode, il est nécessaire de savoir laquelle des deux interactions mesurées correspond à l'absorption photoélectrique et laquelle correspond à la diffusion Compton. Le problème est que cette distinction n'est pas toujours possible dans l'absolu. Si on reprend l'équation (1-2), l'énergie minimale d'un photon diffusé est donnée lorsque l'angle de diffusion θ vaut 180°, on parle alors de photon rétrodiffusé. Je note cette énergie E_c et elle est donnée par :

$$E_c = \frac{E_0}{1 + 2\frac{E_0}{m_e c^2}} \tag{2-1}$$

L'énergie déposée par interaction Compton correspondante est l'énergie du front Compton E_f :

$$E_f = E_0 - E_c \tag{2-2}$$

Par exemple, pour une énergie initiale E_0 de 662 keV, E_c vaut 184,3 keV et E_f vaut 477,7 keV. Ainsi, pour toutes les interactions déposant des énergies comprises entre E_c et E_f, il n'est pas possible dans l'absolu de déterminer laquelle des deux interactions correspond à la diffusion Compton et laquelle correspond à l'absorption photoélectrique.

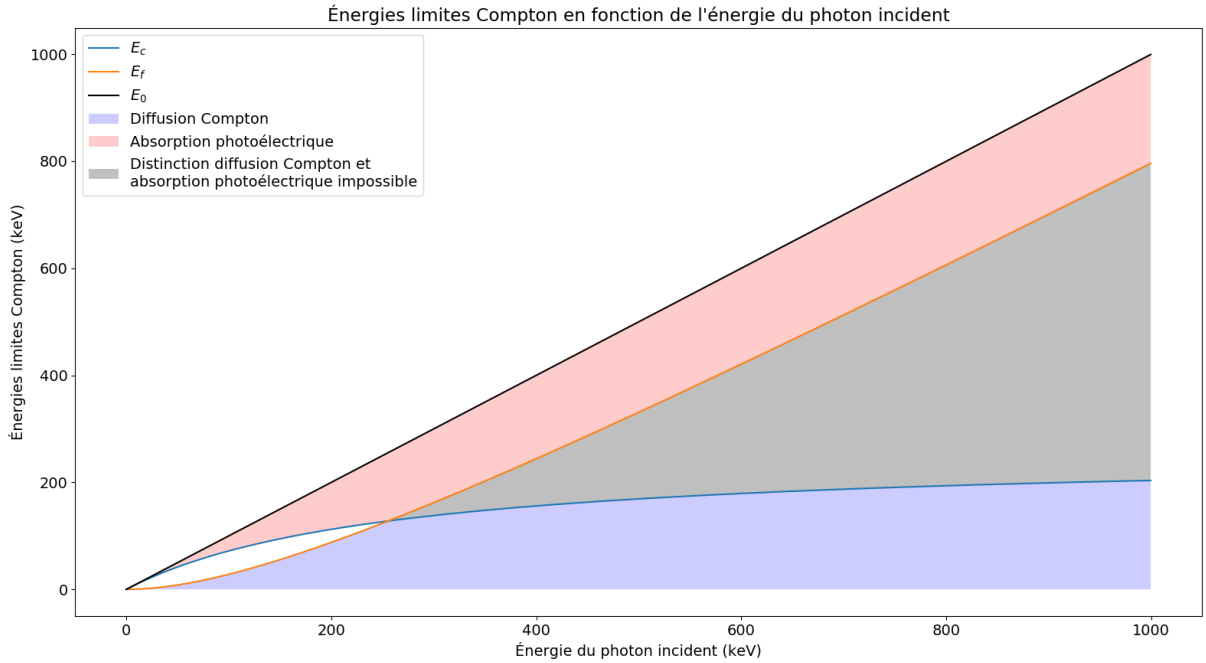


Figure 2-3 : Énergies limites Compton E_f et E_c, zones d'attribution des événements à la diffusion Compton, à l'absorption photoélectrique et zone « indécidable ».

En revanche, si les deux interactions déposent des énergies qui ne sont pas comprises entre E_c et E_f, dans ce cas, on peut associer* l'interaction qui a déposé une énergie inférieure à E_c à la diffusion Compton et celle qui a déposé l'énergie supérieure à E_f à l'absorption photoélectrique. La Figure 2-3 représente les énergies E_c et E_f en fonction de l'énergie du photon incident E_0, et les zones dans

* À l'élargissement Doppler près. En effet, l'équation (1-4) autorise des diffusions Compton dont l'énergie est supérieure à E_f, peu probables certes, mais pas impossibles.

lesquelles la règle de décision définie précédemment est applicable et celle où cette règle ne peut pas l'être.

En particulier, dans le cas où l'énergie initiale est inférieure à $\frac{m_e c^2}{2} = 205{,}5\ \mathrm{keV}$, il est toujours possible d'appliquer la discrimination entre la diffusion Compton et l'absorption photoélectrique. Au-delà de cette énergie, il existe une zone d'indécision, qui augmente avec l'énergie du photon incident.

Pour une énergie initiale E_0 de 662 keV, cette zone n'est pas négligeable et s'étend de 184,3 keV à 477,7 keV. Les cas où les énergies déposées se situent dans cette zone d'indécision sont même largement dominants pour des photons de plus haute énergie initiale. Il est donc nécessaire d'avoir une règle précise et optimale pour traiter ces événements dans le cas de notre détecteur plan unique.

Dans le cas où le système de détection est constitué d'au moins deux couches de détection, dont une joue le rôle de diffuseur et l'autre d'absorbeur, la règle de décision par défaut convient d'attribuer la diffusion Compton à la première couche et l'absorption photoélectrique à la deuxième couche, sans forcément passer par des a priori sur l'énergie déposée par chaque interaction. Cependant, dans mon cas avec Caliste, avec une seule couche de détection, je dois prendre en compte cette indécision.

Pour ce faire, je peux utiliser deux méthodes. La première consiste à appliquer la reconstruction en utilisant les deux ordres d'interaction possible, ce qui aboutit à la reconstruction de deux cônes. Cette méthode apporte en fait plus de bruit que de signal pertinent sur la reconstruction et ne s'avère pas utilisable en pratique.

La deuxième méthode, que je privilégie consiste à utiliser le fait que l'absorption photoélectrique est plus probable pour les photons de plus basse énergie. Le dépôt de plus faible énergie est alors associé à l'absorption photoélectrique et celui de plus haute énergie à la diffusion Compton, dans le cas où les deux énergies se situent dans la zone d'indécision entre E_c et E_f. Si l'une des deux énergies est plus grande que E_f, alors elle sera associée à l'absorption photoélectrique (cf. Figure 2-3).

2.2.2 Simulations d'événements Compton

Afin de valider la méthode de sélection de l'ordre des interactions Compton, il est nécessaire de recourir à des simulations, dans lesquelles il est possible de savoir quelle interaction est la diffusion Compton et laquelle est l'absorption photoélectrique. J'ai développé un algorithme de simulation qui, à partir d'une position donnée d'une source, simule une liste d'événements composés de deux interactions, une diffusion Compton et une diffusion photoélectrique.

Pour ce faire, je considère une source du ciel à une position donnée S. Je simule un photon d'énergie E_0 envoyé aléatoirement depuis S dans un cône contenant le détecteur. J'envoie les photons dans un tel cône pour éviter de simuler des photons qui de toute façon n'atteindraient pas le détecteur.

Si la trajectoire du photon intercepte le cristal CdTe, je calcule la longueur du segment d'intersection entre cette trajectoire et le cristal, que je note L_t. Je calcule la probabilité p_C d'effectuer une diffusion Compton dans ce segment à partir des coefficients d'atténuation totaux $\mu(E_0)$ et Compton $\mu_C(E_0)$ associés au CdTe pour l'énergie E_0 :

$$p_C = \frac{\mu_C(E_0)}{\mu(E_0)}\left(1 - e^{-\mu(E_0)L_t}\right) \tag{2-3}$$

Dans une simulation Monte-Carlo classique, il faudrait accepter l'interaction avec une probabilité p_c ou la rejeter et relancer un nouveau photon. Ici, pour éviter de rejeter un trop grand nombre de photons, j'accepte l'interaction dans tous les cas si elle est possible, et je conserve la valeur de p_c pour un coefficient de pondération que j'utiliserai par la suite.

Une fois l'interaction acceptée, je tire aléatoirement la position d'interaction dans le cristal. Pour ce faire, je considère la probabilité d'interagir à une longueur l_c, en sachant que $l_c \leq L_t$:

$$p(l_c|L_t) = \frac{1 - e^{-\mu(E_0)l_c}}{1 - e^{-\mu(E_0)L_t}} \tag{2-4}$$

Le facteur au dénominateur de l'équation (2-4) prend en compte le fait que l'interaction doit avoir lieu avant une longueur L_t. Je tire un nombre aléatoire α uniformément entre 0 et 1, correspondant à $p(l_c|L_t)$ et je calcule la longueur l_c correspondante :

$$l_c = \frac{\log\left(1 - \alpha\left(1 - e^{-\mu(E_0)L_t}\right)\right)}{-\mu(E_0)} \tag{2-5}$$

Une fois la position d'interaction Compton X_{Comp} sélectionnée, je tire l'angle de diffusion θ à l'aide de la densité de probabilité donnée par la formule de Klein-Nishina, explicitée en Annexe A. J'en déduis l'énergie du photon diffusé E_1 correspondante par la relation (1-2)[*]. La trajectoire du photon diffusé est alors située sur un cône d'axe $\left(SX_{\mathrm{Comp}}\right)$, de sommet X_{Comp} et d'angle au sommet θ. Je tire aléatoirement la direction azimutale de la trajectoire du photon diffusé sur ce cône, de manière uniforme[†].

Sur la trajectoire du photon diffusé, je calcule la longueur du segment de cette trajectoire dans le détecteur, à partir de la position X_{Comp}, je note cette longueur L_{t2}. La probabilité que le photon soit absorbé par effet photoélectrique p_{PE} est donnée en utilisant le coefficient d'absorption associé à l'effet photoélectrique $\mu_{\mathrm{PE}}(E_1)$ pour le CdTe à l'énergie E_1 :

$$p_{\mathrm{PE}} = \frac{\mu_{\mathrm{PE}}(E_1)}{\mu(E_1)}\left(1 - e^{-\mu(E_1)L_{t2}}\right) \tag{2-6}$$

De manière analogue au cas de la première diffusion Compton, je considère que j'accepte cette interaction dès lors qu'elle est possible, en associant un coefficient de pondération donné par p_{PE}, de nouveau pour éviter de rejeter trop de photons et donc d'alourdir les simulations. Le coefficient total de pondération P associé à l'événement simulé est alors donné par :

$$P = p_c \times p_{\mathrm{PE}} \tag{2-7}$$

[*] Je pourrais utiliser la densité de probabilité dans le cas où on considère l'élargissement Doppler, donné par la formule (1-3), mais cela rendrait le calcul plus complexe, car il faudrait tirer aléatoirement θ et E_1 en prenant en compte le fait que leurs densités de probabilité sont liées.

[†] La probabilité sur la direction azimutale n'est pas uniforme lorsque le photon incident est polarisé comme mis en évidence dans (Antier, et al., 2015), mais je fais l'hypothèse ici de photons non polarisés, cas auxquels on s'attend pour des sources radioactives.

Ce coefficient de pondération prend en compte le fait que j'ai accepté la diffusion Compton dès lors qu'elle était possible et de même pour l'absorption photoélectrique. Ce coefficient pondérera l'événement lorsque je l'utiliserai par la suite pour la reconstruction.

Pour finir, je tire la position de l'absorption photoélectrique. Je procède de manière analogue au cas de la diffusion Compton, en tirant aléatoirement la longueur parcourue par le photon dans la trajectoire avec l'interaction photoélectrique l_{PE}, en tirant un nombre α_2 entre 0 et 1 et en calculant l_{PE} par :

$$l_{\mathrm{PE}} = \frac{\log\left(1 - \alpha_2\left(1 - e^{-\mu(E_1)L_{t2}}\right)\right)}{-\mu(E_1)} \tag{2-8}$$

La Figure 2-4 montre le déroulement de cette simulation.

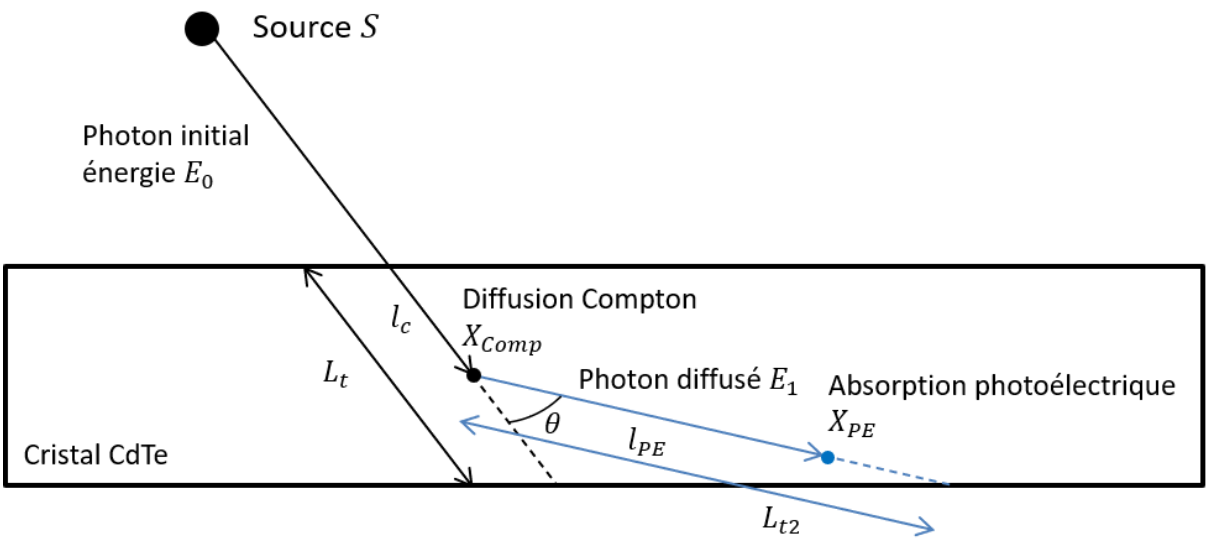


Figure 2-4 : Simulation d'un photon dans un cristal CdTe pixélisé, monolithique et plan au format de Caliste.

À la fin de la simulation, pour chaque événement je relève :

- la position d'interaction Compton X_{Comp} ;
- la position du centre du pixel du détecteur dans lequel a lieu cette interaction X_1 que je calcule à partir de X_{Comp} ;
- l'énergie déposée dans le détecteur par diffusion Compton, $E_2 = E_0 - E_1$;
- la position de l'absorption photoélectrique X_{PE} ;
- la position du centre du pixel du détecteur dans lequel a lieu cette interaction X_2 que je calcule à partir de X_{PE} ;
- l'énergie déposée par absorption photoélectrique E_1 ;
- le coefficient de pondération P qui prend en compte les probabilités que la diffusion Compton et l'absorption photoélectrique aient lieu dans la longueur de cristal traversée.

Par ailleurs, si les deux pixels dans lesquels ont lieu les interactions sont voisins, alors l'événement est rejeté car il peut être confondu avec un partage de charge, comme décrit au paragraphe 2.1. Dans la

suite, je teste préférentiellement mes algorithmes sur des données réelles, mais dans certains cas, je me référerai à ces simulations.

2.2.3 Validation de la méthode

Pour revenir à la détermination de l'ordre d'interaction, j'utilise ce simulateur pour valider la règle de décision que j'ai donnée en fin de paragraphe 2.1, en fonction de la position de la source. J'ai simulé une liste de 35 000 événements pour des sources à des positions couvrant l'ensemble de la demi-sphère devant le détecteur. Les sources sont repérées par leurs coordonnées sphériques, longitude et latitude, par rapport au détecteur. Je donne plus de détails sur ce système de coordonnées en paragraphe 2.3.1.

Pour une position donnée, je calcule le nombre N_b d'événements i simulés, pondérés par le coefficient P_i, pour lesquels la règle de décision donne la bonne séquence. Ce sont donc les événements pour lesquels l'une des conditions suivantes est respectée $E_2 \leq E_c$ ou $E_c \leq E_1 \leq E_2$, je note R l'ensemble de ces événements. Je note N_n le nombre d'événements i pondérés par le coefficient P_i qui ne respectent pas cette règle, je note $\bar{R}$ l'ensemble de ces événements.

$$N_b = \sum_{i \in R} P_i \tag{2-9}$$

$$N_n = \sum_{i \in \bar{R}} P_i \tag{2-10}$$

La probabilité que la règle de décision que j'ai fixée conduise à déterminer la bonne séquence d'interaction est donnée par :

$$p_b = \frac{N_b}{N_n + N_b} \tag{2-11}$$

L'application de cette méthode avec une énergie initiale E_0 à 662 keV donne le résultat présenté en Figure 2-5.

Pour toutes les positions, la règle de décision que j'utilise permet d'obtenir la bonne séquence d'interactions dans 75 % des cas au minimum. Pour les sources situées dans l'axe du détecteur, cette probabilité atteint même plus de 92 %. À ces positions, les seuls angles de diffusion Compton autorisés sont ceux autour de 90°, ce qui correspond à une énergie de photon diffusé de 288 keV, donc inférieure à l'énergie déposée lors de la diffusion Compton de 374 keV, ce qui respecte la règle de sélection.

D'un autre côté, pour les sources proches du plan de détection, sur les bords de la Figure 2-5, la probabilité d'avoir la bonne séquence est aussi élevée (90 %). Cela provient principalement du fait que pour ces positions, la plupart des diffusions Compton se font soit vers l'avant avec un angle de diffusion proche de 0°, soit vers l'arrière, à 180°. Dans le premier cas, si l'angle de diffusion est proche de 0°, cela signifie que l'énergie déposée par diffusion Compton est faible et donc inférieure à E_c, ce qui valide forcément la règle de décision. Dans le second cas, pour des angles proches de 180°, cela signifie que l'énergie du photon diffusé est proche de E_c et donc inférieure à l'énergie déposée lors de la diffusion Compton, ce qui valide aussi la règle de décision.

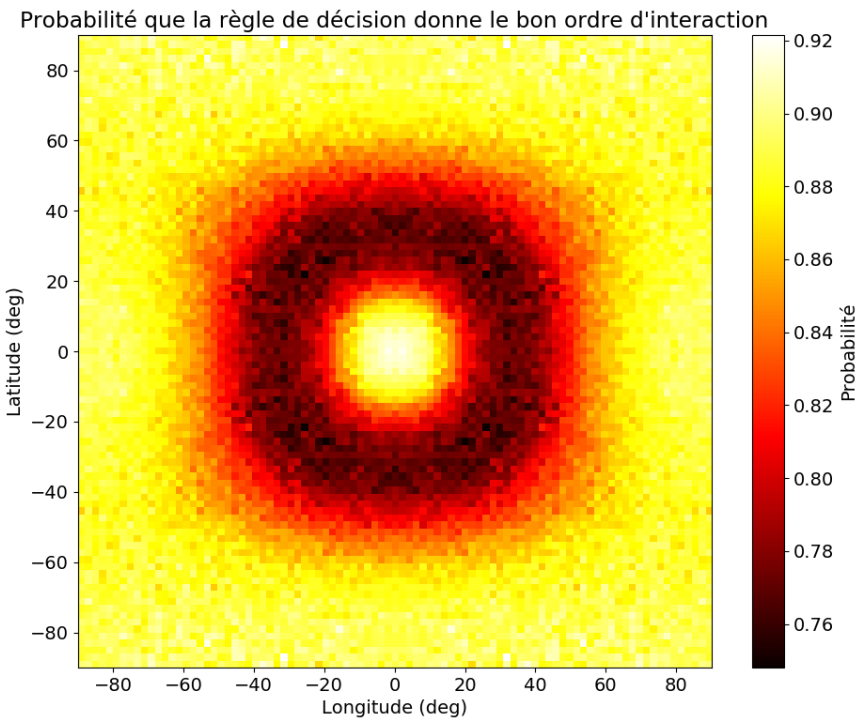


Figure 2-5 : Probabilité que la règle de décision donne la bonne séquence d'interaction. Le centre de la figure représente les positions des sources dans l'axe du détecteur, les côtés représentent des sources situées dans le plan du détecteur.

Cette étude montre que la méthode que j'utilise pour valider la séquence d'interaction est valable sur l'ensemble des positions, permettant de déterminer la bonne séquence dans plus de 75 % des cas, au minimum. Les photons utilisés dans cette étude ont une énergie initiale à 662 keV. Si on considère d'autres énergies, il est nécessaire d'effectuer de nouvelles simulations pour s'assurer de la validité de cette règle de décision.

2.3 Backprojection directe

2.3.1 Système de coordonnées

Afin d'appliquer la reconstruction Compton, je présente dans un premier temps le référentiel dans lequel je me place pour effectuer les calculs liés à la projection des cônes. Soit un repère cartésien (O, x, y, z) dont l'origine O est placée au centre de la surface du détecteur de sorte que la surface du détecteur soit contenue dans le plan (Oxz), les bords du détecteur étant parallèles aux axes (Ox) et (Oz). L'axe (Oy) est orthogonal au plan de détection et dirigé vers l'extérieur.

Étant donné que je vais réaliser les projections des cônes sur une sphère de centre O et de rayon r , je considère aussi un système de coordonnées sphériques (O, r, δ, φ) où r est le rayon de la sphère, δ est la colatitude d'un point de la sphère, mesurée à partir de l'axe (Ox) et φ est la longitude du point considéré, mesurée à partir de l'axe (Oy). La Figure 2-6 représente ce système de coordonnées. Pour effectuer les calculs des projections des cônes sur la sphère, je me placerai dans le plan (φ, δ). Pour effectuer les représentations des reconstructions, j'utiliserai la latitude au lieu de la colatitude, la latitude valant directement $90° - \delta$.

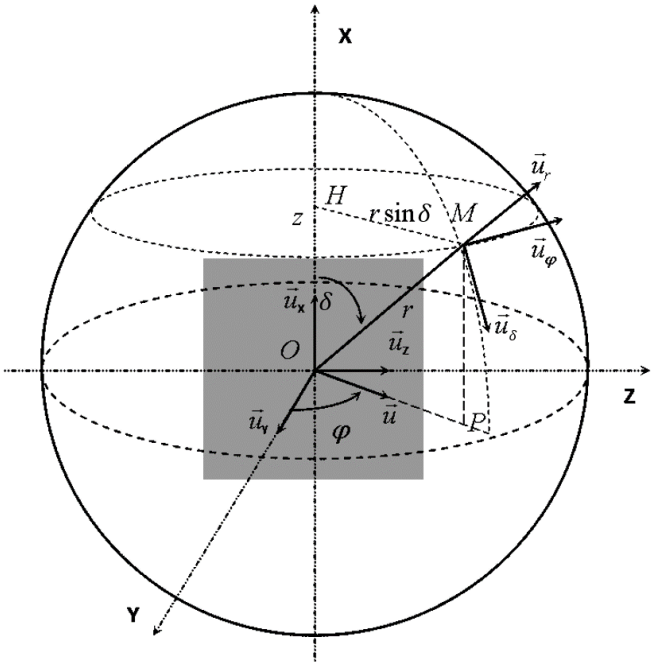


Figure 2-6 : Système de coordonnées utilisé, le carré représente le plan de détection, i.e. le cristal CdTe de Caliste.

2.3.2 Calcul et projection des cônes sur la sphère

Considérons un événement Compton : soit une diffusion Compton à la position X_{Comp} de coordonnées cartésiennes (x_c, y_c, z_c) avec un dépôt d'énergie E_2 et une absorption photoélectrique à la position X_{PE} de coordonnées (x_p, y_p, z_p) avec un dépôt d'énergie E_1. L'angle de diffusion θ est calculé à partir de E_1 avec l'équation (2-2).

La *backprojection* directe consiste à calculer l'intersection du cône d'axe $\overrightarrow{X_{\mathrm{PE}}X_{\mathrm{Comp}}}$, de sommet X_{Comp} et d'angle au sommet θ avec la sphère de centre O et de rayon R, qui représente idéalement la distance à la source.

Soit un point M de coordonnées (x, y, z) de ce cône. Dans le cas où l'axe du cône coïncide avec l'axe (Ox), ces coordonnées sont données directement par :

$$\begin{cases} x = \rho \cot(\theta) \\ y = \rho \cos(\alpha) \\ z = \rho \sin(\alpha) \end{cases} \tag{2-12}$$

où ρ est un paramètre réel strictement positif qui représente la distance entre l'axe du cône et le point M et $\alpha \in [0,2\pi]$ correspond à la position azimutale du point M sur le cône.

Dans le cas général où l'axe $\overrightarrow{X_{\mathrm{PE}}X_{\mathrm{Comp}}}$ ne coïncide pas avec l'axe (Ox), je note φ, la longitude de ce vecteur et δ sa colatitude. Je dois appliquer à l'équation (2-12) deux rotations successives pour passer de l'axe (Ox) à l'axe $\overrightarrow{X_{\mathrm{PE}}X_{\mathrm{Comp}}}$. La première rotation est alors une rotation d'axe (Oz) et d'angle δ, donnée par la matrice :

$$R_z(\delta) = \begin{pmatrix} \cos(\delta) & -\sin(\delta) & 0 \\ \sin(\delta) & \cos(\delta) & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{2-13}$$

La seconde est la rotation d'axe (Ox) et d'angle φ donnée par la matrice :

$$R_x(\varphi) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\varphi) & -\sin(\varphi) \\ 0 & \sin(\varphi) & \cos(\varphi) \end{pmatrix} \tag{2-14}$$

Enfin, je dois appliquer une translation de vecteur $\overrightarrow{OX_{\text{Comp}}}$ pour que le sommet du cône soit bien en X_{Comp}. L'application de ces deux rotations et de cette translation aux coordonnées définies dans l'équation (2-12) donnent alors les coordonnées suivantes :

$$\begin{cases} x = \rho(\cos(\delta)\cot(\theta) - \sin(\delta)\cos(\alpha)) + x_c \\ y = \rho(\cos(\varphi)\,(\sin(\delta)\cot(\theta) + \cos(\delta)\cos(\alpha)) - \sin(\varphi)\sin(\alpha)) + y_c \\ z = \rho(\sin(\varphi)\,(\sin(\delta)\cot(\theta) + \cos(\delta)\cos(\alpha)) + \cos(\varphi)\sin(\alpha)) + z_c \end{cases} \tag{2-15}$$

Calculons maintenant les points M se situant sur l'intersection entre ce cône et la sphère de rayon R. Pour ce faire, je dois trouver la relation entre les paramètres ρ et α pour les points situés sur cette intersection. Un tel point M de coordonnées (x, y, z) doit obéir à l'équation (2-15) pour être sur le cône et à l'équation suivante pour être sur la sphère :

$$x^2 + y^2 + z^2 - R^2 = 0 \tag{2-16}$$

Pour simplifier, je réécris l'équation (2-15) sous la forme suivante :

$$\begin{cases} x = \rho k_x + x_p \\ y = \rho k_y + y_p \\ z = \rho k_z + z_p \end{cases} \tag{2-17}$$

Les coefficients k_x, k_y et k_z ne dépendent que de α. En introduisant les relations (2-17) dans l'équation (2-16), on obtient :

$$\rho^2\left(k_x^2 + k_y^2 + k_z^2\right) + 2\rho\left(k_x x_p + k_y y_p + k_z z_p\right) + x_p^2 + y_p^2 + z_p^2 - R^2 = 0 \tag{2-18}$$

Cette équation du second degré se résout directement avec l'utilisation du discriminant et la détermination des racines du polynôme correspondant. Ainsi, pour α fixé, je détermine les coefficients k_x, k_y et k_z correspondants et je résous l'équation (2-18). Si l'équation ne possède pas de solution, cela signifie qu'il n'existe pas d'intersection entre la sphère et le cône, soit donc que la sphère est trop petite. Si l'équation possède deux solutions ρ, il faut choisir la solution pour laquelle ρ est positif.

Le paramètre ρ se calcule alors directement à partir du paramètre α et de la résolution de (2-18), ce qui permet uniquement de paramétriser l'intersection du cône avec α et de trouver l'ensemble des points de cette intersection. Pour ce faire, je calcule l'ensemble des points $M(x, y, z)$ avec l'équation (2-15) en faisant varier α de 0 à 2π. Enfin, pour construire la carte en longitude et latitude de ces intersections, je calcule la longitude et la latitude de chacun de ces points M.

Dans cette méthode, il est nécessaire de connaître les positions des interactions, les coordonnées (x_c, y_c, z_c) et (x_p, y_p, z_p) et les énergies déposées. En ce qui concerne les énergies déposées, j'ai décrit les méthodes de sélection en énergie, données en paragraphes 2.1 et 2.2.3. Pour la profondeur d'interaction dans le détecteur, les coordonnées y_c et y_p dans le système de coordonnées défini au paragraphe 2.3.1 ne sont pas mesurées par le détecteur. Ces données sont inconnues et j'ai fait le choix de les fixer à la surface du détecteur, c'est-à-dire de prendre $y_c = y_p = 0$ pour tous les événements. Cette inconnue est en fait le facteur créant la plus grande incertitude sur la reconstruction Compton avec le système que j'utilise. Elle est dominante sur l'incertitude en énergie ou l'incertitude sur la position sur les coordonnées (x, z).

La Figure 2-7 montre le décalage angulaire que crée cette incertitude maximale sur l'axe du cône à reconstruire, en fonction de la distance entre les deux interactions. L'erreur maximale est commise lorsque l'une des interactions a lieu en surface du cristal et la seconde a lieu au fond du cristal, alors que je considère par ma méthode que les deux interactions ont lieu à la même profondeur. Ainsi, plus les interactions sont proches, plus l'incertitude sur la profondeur crée un bras de levier important, pouvant donner une erreur allant jusque 90°. En tenant compte du fait que les interactions sont sélectionnées dans des pixels non voisins, l'erreur maximale attendue est de l'ordre de 57°. Ainsi, on voit que cette erreur est potentiellement dominante dans la reconstruction devant les autres sources d'erreurs, notamment par rapport à l'incertitude liée à l'élargissement Doppler qui est de 24° au maximum. Notons que cette erreur est calculée dans le pire des cas, où la différence de profondeur entre les interactions est de 1 mm. Cette erreur n'empêche pas le système d'effectuer des reconstructions Compton, comme je le montre en paragraphe 2.3.3 puis en partie 3.

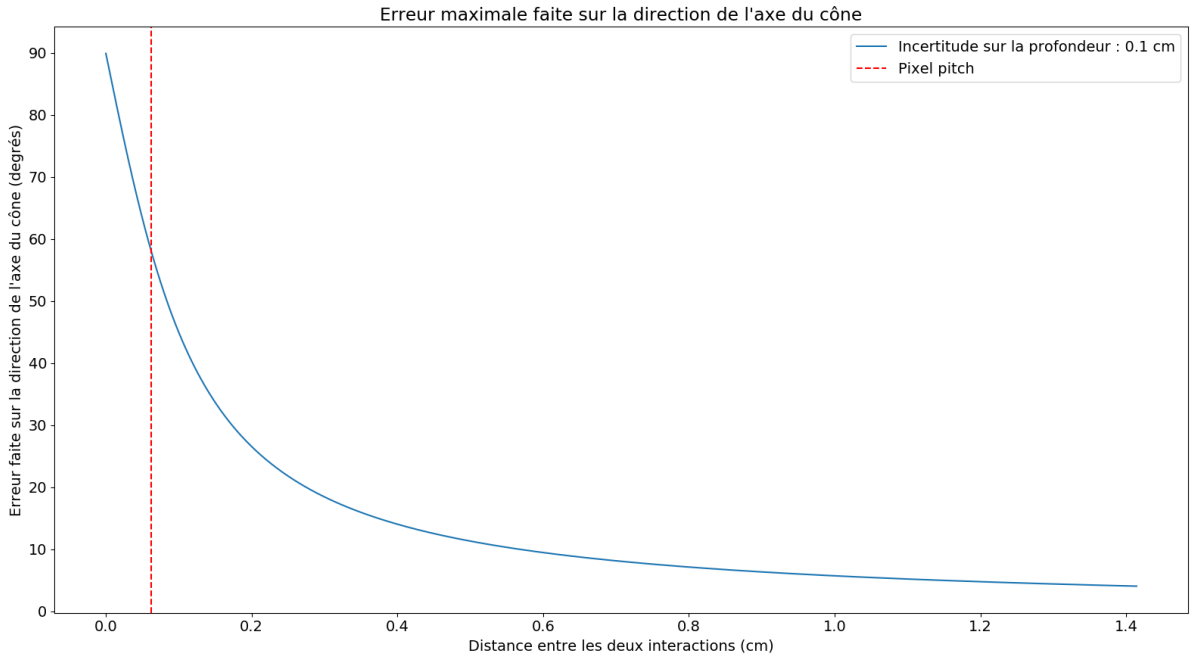


Figure 2-7 : Erreur maximale sur la direction de l'axe du cône, soit pour une incertitude de 1 mm. La ligne rouge représente la taille d'un pixel, soit la distance minimale entre deux interactions, du fait de la sélection d'interactions ayant lieu dans deux pixels non voisins.

La deuxième incertitude en position concerne les coordonnées (x, z) des deux interactions. Pour cela, j'associe ces coordonnées au centre du pixel qui a été touché par l'interaction. S'il existe un cluster de

pixels touchés par l'interaction, comme dans le cas des types b ou c représentés en Figure 2-1, j'associe le barycentre des pixels du cluster, pondéré par les énergies déposées dans chaque pixel. L'erreur que je commets à cause de cette incertitude est inférieure à celle commise par l'incertitude en profondeur, elle est typiquement de l'ordre de 5°, comparable à l'incertitude angulaire du fait de la résolution en énergie.

En matière de représentation, je discrétise la carte du « ciel » à reconstruire par zones de 2°x2° (2° en longitude et en latitude), sauf mention contraire, du fait que dans la configuration du détecteur, je n'attends pas une précision meilleure que 2° sur la reconstruction de l'angle diffusé Compton à cause de l'élargissement Doppler. J'effectue le calcul en discrétisant le cône avec 5000 points, c'est-à-dire que j'utilise 5000 valeurs pour le paramètre α uniformément réparties entre 0 et 2π. Ces 5000 points sont suffisants pour obtenir au moins un point du cône dans chaque zone de 2°x2° de la carte du ciel, quel que soit le cône. S'il existe au moins un point du cône dans une zone de 2°x2° du ciel, j'affecte 1 à cette position et 0 sinon. Le résultat final sur un ensemble d'événements Compton consiste à représenter l'accumulation de toutes les projections des cônes.

Une autre méthode consiste à « compter » le nombre de points dans chaque zone, de façon à tenir compte de la « longueur de cône » présente dans la zone. Cependant, cette méthode présente des difficultés de reconstruction au niveau des « pôles » de l'image, pour une latitude proche de 90° ou de -90°, du fait que la diminution de l'angle solide défini par la zone de 2°x2° doit être prise en compte.

2.3.3 Application de la backprojection directe à des données réelles acquises avec Caliste

J'ai effectué des mesures avec Caliste en utilisant une source de ^{137}Cs de 1,6 MBq, dont la ligne d'émission principale à 661,7 keV est intéressante pour appliquer l'imagerie Compton. La source est placée à 30 cm du détecteur. Le débit de dose correspondant est de 1,4 µSv/h environ au niveau du détecteur.

J'ai effectué trois acquisitions à trois positions différentes : une aux coordonnées $(0°, 0°)$ soit dans l'axe du détecteur, une aux coordonnées $(10°, 0°)$ et une aux coordonnées $(30°, 0°)$. Il faut noter que, dans les conditions expérimentales, l'incertitude sur placement des sources est typiquement de l'ordre de $\pm$ 1,5 cm à 30 cm soit une incertitude angulaire de $\pm$ 3°. Cette incertitude est liée d'une part au fait que la source n'est pas absolument ponctuelle et d'autre part que nous n'avons pas à ce stade mis en œuvre de dispositif de positionnement plus précis, tel qu'un système de translation manuel ou robotisé. La boîte grise représentée pour la position des sources sur chaque image de reconstruction par la suite correspond à cette incertitude. Pour chaque acquisition, le temps de *peaking* du détecteur est 6,07 µs, la haute-tension appliquée est de 300 V et la température du détecteur est de -10°C environ.

Le Tableau 2-1 récapitule les caractéristiques de chaque acquisition. Les durées d'acquisition sont très longues car le taux de comptage est faible (de l'ordre de 5 à 5,5 coups/s). Le nombre total d'événements est donné en tenant compte de la multiplicité, c'est-à-dire que si plusieurs événements sont arrivés en coïncidence temporelle, je ne compte qu'une seule fois ce groupe d'événements. Le nombre d'événements Compton correspond au nombre d'événements sélectionnés par ma méthode explicitée en paragraphe 2.1. En moyenne, entre 7 et 8 événements pour 1000 sont des événements Compton d'intérêt pour la reconstruction d'une image avec Caliste.

Mesure	Durée d'acquisition (h)	Nombre total d'événements	Taux de comptage (cps/s)	Nombre d'événements Compton	Ratio événements Compton vs total d'événements
$(0°, 0°)$	35,7	650921	5	4855	0,75 %
$(10°, 0°)$	57,6	1146230	5,5	8065	0,7 %
$(30°, 0°)$	84,1	1537473	5,1	12222	0,79 %

Tableau 2-1 : Caractéristiques des acquisitions utilisées pour les reconstructions.

Je présente d'une part les résultats pour la source seule aux positions $(10°, 0°)$ et $(30°, 0°)$. Puis, par un processus d'augmentations de données consistant à mélanger les listes d'événements acquises avec les acquisitions à $(0°, 0°)$ et $(10°, 0°)$, j'ai créé une nouvelle liste d'événements qui fusionne les deux acquisitions et qui représente donc une acquisition avec deux sources. Cette approche est valable du fait du faible taux de comptage sur les détecteurs. En effet, si le taux de comptage était très élevé, supérieur à 20 000 coups par seconde typiquement, il aurait fallu tenir compte du phénomène de *pile-up* ou empilement de photons sur un même pixel lors de la fusion des données. Sur les images optiques, comme en Figure 2-12, j'ai recréé une scène équivalente à celle correspondant à la fusion de données à titre de représentation.

Pour ces données, j'effectue la *backprojection* directe sur une sphère de rayon $R = 30$ cm. Je projette à une distance qui correspond à la distance réelle qui sépare le détecteur de la source, donnée qui est en pratique inconnue. Cependant, dans le cadre de l'étude des performances de la *backprojection* directe, ainsi que des algorithmes de traitement que je développe dans la partie 3, je retire cette inconnue pour caractériser les performances des algorithmes dans une situation où l'on connaît déjà la distance à la source, et je montre justement les limites des algorithmes même avec cette connaissance. Par ailleurs, cette distance est 30 fois plus grande que la taille du détecteur, ce qui permet de considérer qu'elle est proche de « l'infini » devant le détecteur, et projeter à des distances plus grandes ne modifie pas les résultats de manière significative.

En Figure 2-8, je montre le résultat de *backprojection* directe pour la source à la position $(10°, 0°)$. La figure finale représente le nombre de cônes qui passent par chaque position du ciel. On note déjà que le maximum d'intersections se situe bien à la position de la source. Par ailleurs, on remarque une symétrie de la figure, autour de l'axe à 90° en longitude, ainsi qu'autour de l'axe à -90°/270° en longitude. Cette symétrie correspond en fait au plan de détection. Elle implique que tout ce qui est reconstruit vers l'avant du détecteur est aussi reconstruit de manière similaire derrière le détecteur. Pour cette raison, dans la suite de ce chapitre, je me restreins seulement à la zone entre -90° et 90° en longitude. Cette symétrie implique aussi qu'il n'est pas possible de distinguer l'avant de l'arrière du détecteur, du fait de sa géométrie planaire, ce qui s'ajoute à la méconnaissance de la profondeur d'interaction. De ce fait, le champ de vue atteignable avec Caliste est en réalité de 180° en longitude. Si une source se situe derrière le détecteur, elle pourra être imagée, mais elle risque d'être considérée à tort vers l'avant du détecteur au lieu de l'arrière.

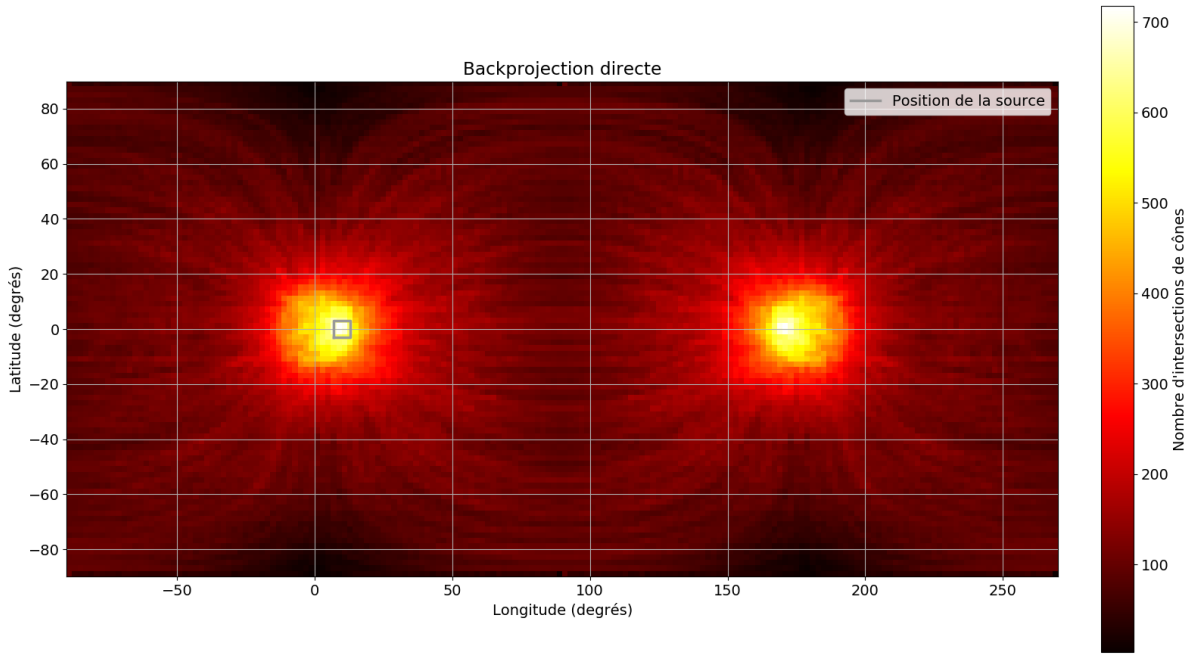


Figure 2-8 : *Backprojection* directe pour la mesure d'une source de ^{137}Cs (8065 cônes) en $(10°, 0°)$.

La Figure 2-9 représente la *backprojection* pour cette même mesure, mais uniquement sur le champ de vue vers l'avant du détecteur. La largeur à mi-hauteur de la zone la plus intense est très large et vaut 36,3°. J'applique aussi un seuil sur cette image afin de mieux localiser les zones les plus intenses. Le seuil que j'ai choisi ne retient que les zones où l'intensité est supérieure ou égale à 90 % de l'intensité maximale sur l'image*. On note qu'avec ce seuil, la source est effectivement retrouvée à la bonne position, mais la zone identifiée s'étale sur 6° en longitude et entre 4° et 6° en latitude, alors que la source est quasiment ponctuelle (phi ~ 5 mm à 30 cm).

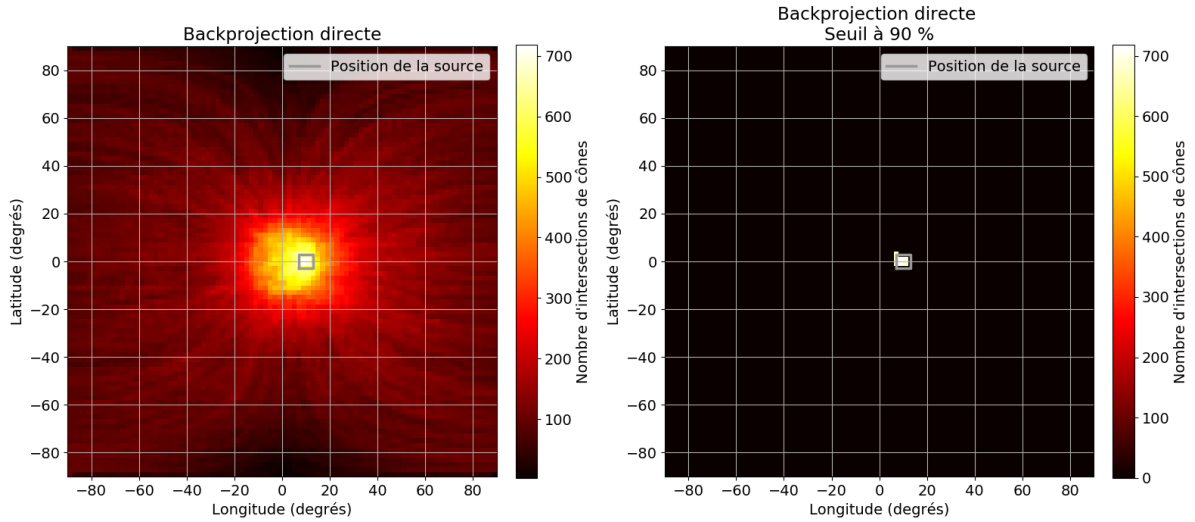


Figure 2-9 : À gauche : *Backprojection* directe pour la mesure d'une source de ^{137}Cs (8065 cônes) en $(10°, 0°)$ sur le champ de vue restreint vers l'avant du détecteur. À droite : image seuillée à 90 %.

* L'optimisation de ce seuil peut faire l'objet d'une étude selon les sources à détecter, qui n'a pas été menée dans le cadre de mon travail de thèse.

La Figure 2-10 représente la *backprojection* pour la source à la position $(30°, 0°)$. De nouveau, l'intensité maximale est bien localisée au niveau de la position attendue de la source. La largeur à mi-hauteur de la reconstruction est de 71,7°, ce qui est très imprécis. Le seuil à 90 % rend une zone étalée, jusqu'à 8° en longitude.

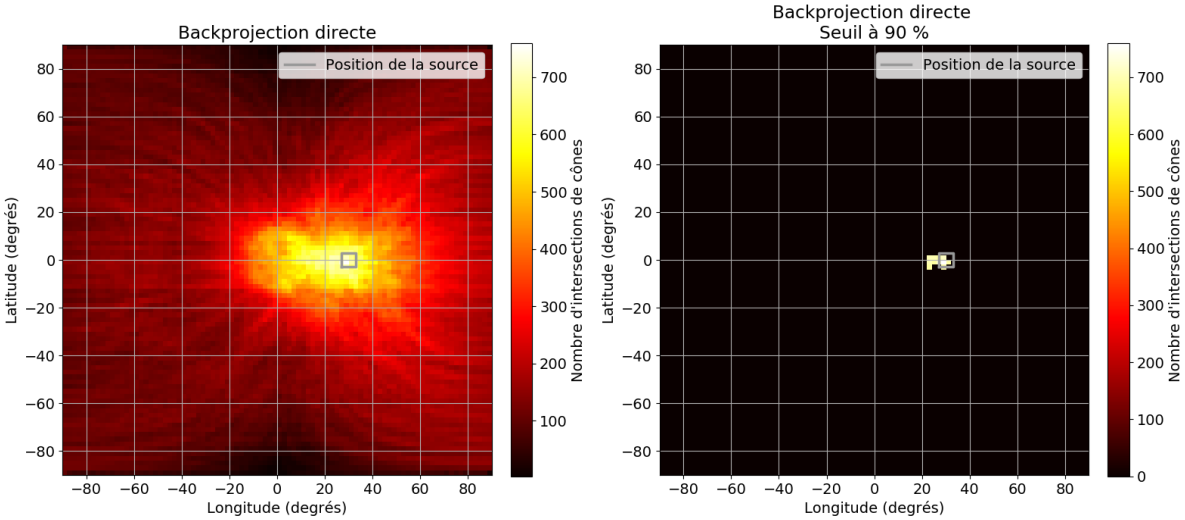


Figure 2-10 : À gauche : *Backprojection* directe pour la mesure d'une source de ^{137}Cs (12 222 cônes) en $(30°, 0°)$. À droite : image seuillée à 90 %.

La Figure 2-11 montre le résultat de la *backprojection* appliquée aux données combinant les mesures de sources situées en $(0°, 0°)$ et $(10°, 0°)$. On constate avec le seuil à 90 % que seule la source en $(10°, 0°)$ est retrouvée, du fait notamment qu'elle apparaît comme plus intense vu qu'elle contient les deux tiers des cônes reconstruits d'après le Tableau 2-1. La *backprojection* seule ne retrouve pas la source positionnée en $(0°, 0°)$ ici.

La Figure 2-12 montre la superposition optique de la reconstruction gamma et de l'image de la scène correspondant à cette mesure. La photographie est prise au niveau de la tête de détection. Les images gamma sont interpolées par la fonction *spline16* de Python pour obtenir un rendu plus lisse.

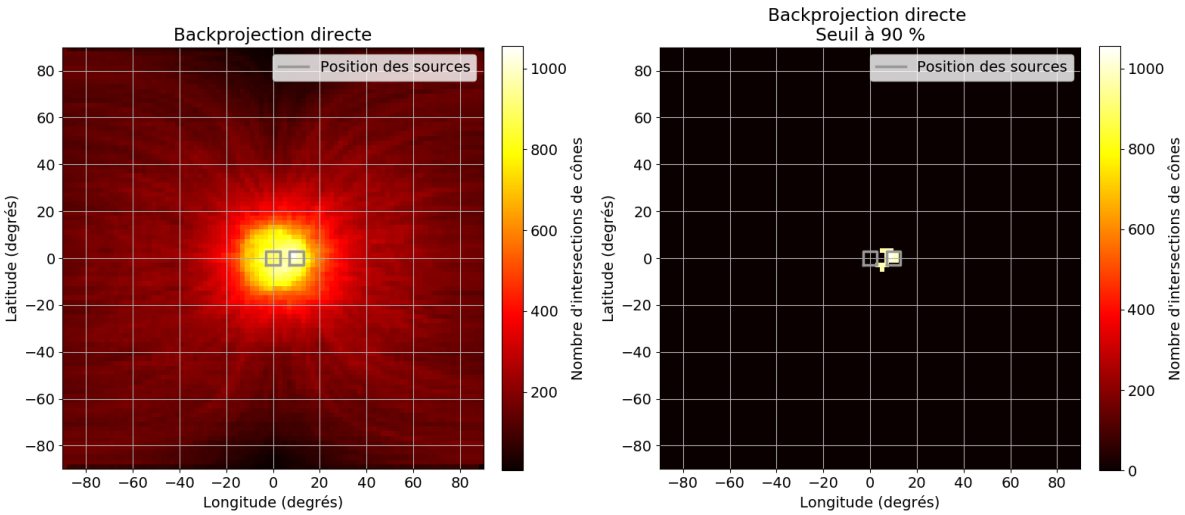


Figure 2-11 : À gauche : *Backprojection* directe pour la mesure de deux sources de ^{137}Cs (12 920 cônes), à $(0°, 0°)$ et à $(10°, 0°)$. À droite : image seuillée à 90 %.

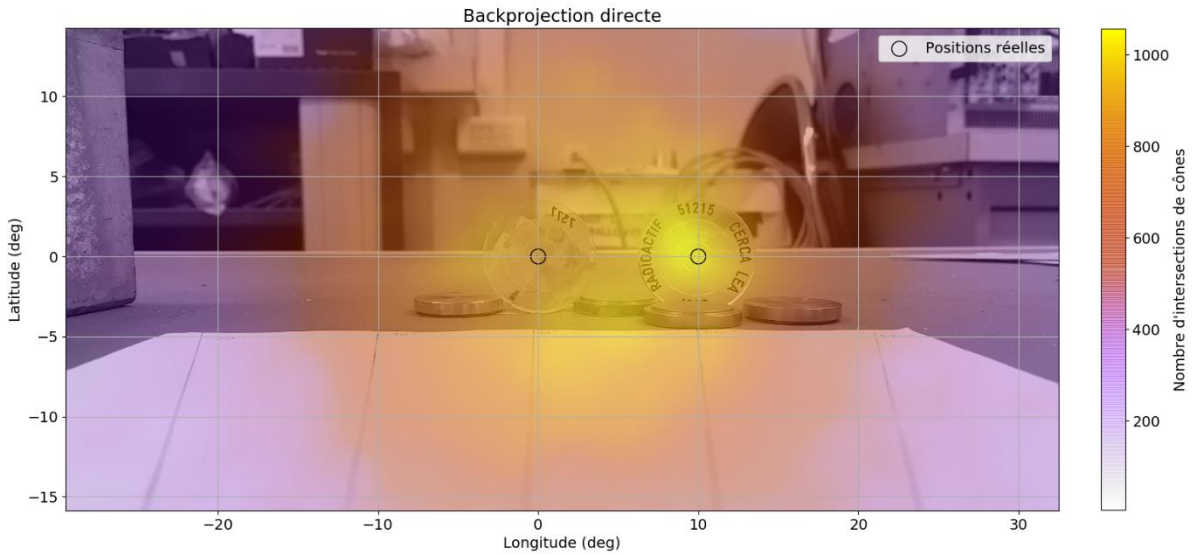


Figure 2-12 : Superposition optique de la scène et de l'image gamma de la *backprojection* correspondante.

Ces exemples montrent que l'application de la reconstruction Compton avec Caliste est possible et qu'elle retrouve correctement la position d'une unique source ponctuelle dans le champ de vue. Ils montrent aussi les limites de la reconstruction Compton, liées au bruit intrinsèque généré par la méthode : le bruit généré par les traces des cônes dans les zones où ne se situent pas les sources.

Ce bruit intrinsèque est le principal inconvénient de la reconstruction Compton appliquée avec la *backprojection* directe. On constate notamment qu'elle ne retrouve pas correctement la position de plusieurs sources ponctuelles, en particulier si ces sources ponctuelles de même nature sont proches l'une de l'autre, séparées par une distance angulaire de moins de 10°. En effet, le bruit généré par les cônes d'une source se répercute au niveau de la position des autres sources, et réciproquement. Dans le cas où l'on cherche à imager des radioisotopes différents, que l'on aurait préalablement identifiés par un algorithme de spectro-identification par exemple, on peut effectuer une sélection en énergie propre aux lignes d'émission d'intérêt de chacun des radioisotopes.

Ainsi, ces exemples illustrent la nécessité d'utiliser des algorithmes de traitement avancés, que je présente dans la partie 3 suivante.

Pour finir, en ce qui concerne le temps de calcul, l'algorithme de *backprojection* directe nécessite 22 secondes de calcul pour traiter 8065 événements, le nombre d'opérations à effectuer étant linéaire suivant le nombre d'événements.

3 Algorithmes de traitement Compton

La problématique du traitement de données Compton a abouti au développement de nombreuses techniques de reconstruction d'images. Une revue de l'état de l'art de ces techniques a été effectuée par (Frandes, et al., 2016) où plusieurs catégories d'algorithmes sont distinguées.

Parmi ces techniques, l'une des plus utilisées est l'algorithme LM-MLEM (*List-Mode Maximum Likelihood Expectation Maximisation*) qui fait partie de la catégorie des algorithmes utilisant la vraisemblance. Je l'applique dans un premier temps sur les données issues de Caliste, mais je vais montrer que cette méthode semble peu adaptée à notre système de détection, au vu de ses incertitudes.

J'ai choisi aussi de m'intéresser à l'approche stochastique appelée SOE (*Stochastic Origin Ensemble*) et plus particulièrement à la version dite RR (*Resolution Recovery*) qui prend en compte les incertitudes sur le détecteur et les inclut dans le traitement.

Enfin, j'aborderai deux méthodes nouvelles que j'ai développées. D'une part, j'ai développé une méthode basée sur la formule de Bayes, qui peut être vue comme une version bayésienne de la *backprojection* directe. D'autre part, j'ai implémenté un algorithme d'inversion Compton basé sur la recherche de la position optimale des sources permettant, à partir de simulations, d'expliquer la *backprojection* obtenue sur des données réelles.

Dans cette partie, je présente les différents algorithmes et j'évalue leurs performances à partir des mesures que j'ai utilisées en paragraphe 2.3.3 pour des sources ponctuelles. Je montre les avantages et les limites de chacun de ces algorithmes. Dans un second temps, je teste ces algorithmes sur des simulations d'une source étendue, pour évaluer leur capacité à traiter ce type de cas.

J'ai présenté les méthodes que j'ai utilisées et leurs applications à des données réelles avec Caliste à la conférence ANIMMA 2019, dont le *proceedings* est disponible ici (Daniel, et al., 2020).

Je souligne aussi qu'il existe des algorithmes de traitement analytiques comme la *Filtered Back Projection* (Haefner, et al., 2015) ou une approche analytique de l'inversion Compton (Maxim, et al., 2009). Dans ce type d'approche, les projections des cônes sont transformées en utilisant des harmoniques sphériques afin d'amener les données dans un espace dans lequel elles peuvent être traitées plus efficacement, en utilisant par exemple la transformée inverse de Radon comme c'est le cas dans l'utilisation des CT-scan (*Computerized Tomography*) (Herman, 2009). Je n'ai pas implémenté ces différents algorithmes bien qu'il serait naturellement intéressant de tester leurs applications avec Caliste.

3.1 Algorithme LM-MLEM : *List-Mode Maximum Likelihood Expectation Maximization*

3.1.1 Principe de la méthode

L'algorithme LM-MLEM (Xu, et al., 2004) reprend l'algorithme MLEM que j'ai présenté au Chapitre 4 avec des adaptations liées à la spécificité des données Compton. En effet, cet algorithme consiste à maximiser la vraisemblance, c'est-à-dire la probabilité d'observer les données i effectivement acquises sous l'hypothèse d'une distribution de position des sources j. Pour ce faire, l'approche MLEM s'applique au cas de données suivant des lois de Poisson indépendantes.

Pour la reconstruction Compton, on recherche le nombre de photons émis par chaque position j du ciel, qui ont été détectés par le système de détection. On suppose que ce nombre de photons suit une loi de Poisson. Un événement Compton est caractérisé par trois paramètres* : la position p_1 à laquelle

* Dans le cas du masque codé présenté en Chapitre 4, seul un paramètre est utilisé : la position du pixel dans lequel a lieu l'interaction.

a lieu l'interaction Compton, la position p_2 à laquelle a lieu l'absorption photoélectrique et l'énergie déposée par effet photoélectrique E_1. Le nombre de possibilités pour un événement Compton est donc très grand : si on discrétise les positions p_1 et p_2 suivant les 256 pixels* du détecteur, et les énergies par canal de 1 keV, soit $E_0 - E_c = 484$ keV possibles pour une énergie initiale E_0 de 662 keV, il existe 31 719 424 configurations i possibles. Sachant que l'ordre de grandeur du nombre d'événements Compton que je détecte est de quelques milliers à une dizaine de milliers, je ne peux pas couvrir l'ensemble de l'espace des configurations possibles, ce qui rend l'algorithme MLEM inapplicable dans la version que j'ai présentée en Chapitre 4.

Pour rappel, l'algorithme MLEM donne une estimation λ_j du nombre de photons émis par la position j du ciel, en fonction du nombre d'événements Y_i détectés dans le canal i de la discrétisation de l'espace des événements (positions p_1, p_2 et énergie E_1) et la probabilité $p(i|j)$ d'observer un événement dans ce canal i sachant que le photon est émis depuis la position j :

$$\lambda_j^{\text{new}} = \lambda_j^{\text{old}} \sum_i \frac{Y_i p(i|j)}{\sum_{j'} p(i|j')\lambda_{j'}^{\text{old}}} \tag{3-1}$$

La version List-Mode a été développée pour répondre à ce problème (Parra & Barrett, 1998). Elle consiste à appliquer l'algorithme MLEM uniquement sur les événements observés, et non en considérant l'ensemble des événements possibles. Cela revient à considérer une discrétisation infinitésimale et à appliquer l'équation (3-1) avec $Y_i = 1$ pour chaque événement i observé et $Y_i = 0$ sinon.

L'équation (3-1) devient donc† :

$$\lambda_j^{\text{new}} = \lambda_j^{\text{old}} \sum_{i \text{ observé}} \frac{p(i|j)}{\sum_{j'} p(i|j')\lambda_{j'}^{\text{old}}} \tag{3-2}$$

Le terme qui doit être calculé est le terme $p(i|j)$. C'est la probabilité que l'événement i soit produit à partir de la position j du ciel. Si la position j du ciel n'appartient pas au cône défini par l'événement i, cette probabilité vaut directement 0 : l'événement i ne peut pas être produit par un photon émis depuis la position j. Sinon, si la position j se situe sur le cône correspondant, on peut décomposer cette probabilité suivant trois termes :

$$p(i|j) \propto \exp\left(-\mu_t(E_0) l_{j1}\right) \left.\frac{d\sigma_c}{d\theta}\right|_{\theta(E_1)} \exp(-\mu_t(E_1) l_{12}) \tag{3-3}$$

où $\mu_t(E)$ est le coefficient d'absorption linéique total dans le CdTe à l'énergie E, l_{j1} est la longueur parcourue par le photon dans le CdTe avant la position p_1 s'il a été émis depuis la position j du ciel, l_{12}

* Ceci n'est pas forcément automatique : la position peut être différente du centre d'un seul pixel. Dans le cas où l'interaction produit un cluster de pixels, j'utilise le barycentre de ce cluster.

† L'équation (3-2) est parfois écrite en ajoutant un terme s_j de la manière suivante :

$$\lambda_j^{\text{new}} = \frac{\lambda_j^{\text{old}}}{s_j} \sum_{i \text{ observé}} \frac{p(i|j)}{\sum_{j'} p(i|j')\lambda_{j'}^{\text{old}}}$$

Ce terme s_j représente la probabilité qu'un photon émis depuis une position j soit effectivement détecté, correspondant à l'efficacité de détection à cette position. Ce terme est un terme de redressement qui peut être pris en compte a posteriori, après application de l'algorithme LM-MLEM.

est la distance entre les positions p_1 et p_2 et $\left.\frac{d\sigma_c}{d\theta}\right|_{\theta(E_1)}$ est la section efficace différentielle Compton à l'angle $\theta(E_1)$, angle de diffusion correspondant à l'énergie E_1.

Les trois termes de l'équation (3-3) prennent ainsi en compte la probabilité que le photon n'interagisse pas entre son entrée dans le détecteur et la position p_1, la probabilité qu'il soit diffusé à l'angle $\theta(E_1)$ c'est-à-dire avec une énergie E_1 et la probabilité que le photon n'interagisse pas entre les positions p_1 et p_2.

Il est intéressant de noter que les termes $\left.\frac{d\sigma_c}{d\theta}\right|_{\theta(E_1)}$ et $\exp(-\mu_t(E_1)l_{12})$ sont indépendants de la position j du ciel. Comme ils apparaissent à la fois au numérateur et au dénominateur de l'équation (3-2), ils se simplifient, et il n'est donc pas nécessaire de les prendre en compte.

Par ailleurs, (Xu, et al., 2004) a montré à travers des simulations que pour leur détecteur CZT de 15x15x10 mm^3, les distances parcourues dans le détecteur sont suffisamment petites pour considérer le terme $\exp\left(-\mu_t(E_0)l_{j1}\right)$ comme indépendant de la position j. Le détecteur que j'utilise ayant un volume plus petit, de 10x10x1 mm^3, je peux aussi utiliser cette approximation. De ce fait, ce dernier terme se simplifie aussi dans l'équation (3-2). Le calcul exact de cette probabilité exigerait de lourdes simulations pour pouvoir couvrir l'ensemble des événements i émis par une position j.

Ainsi, finalement, du fait que $p(i|j)$ est considéré comme indépendant de la position j d'émission du photon, il suffit de remplacer ce terme par 1 dans l'équation (3-2) si la position j appartient au cône défini par l'événement i, le facteur $\sum_{j'} p(i|j')\lambda_{j'}^{\mathrm{old}}$ assurant la normalisation. Cette approximation évite d'effectuer le calcul de l'ensemble des termes de l'équation (3-3) et elle est par exemple utilisée avec succès dans (Xu, et al., 2004).

3.1.2 Test LM-MLEM avec des données Caliste

J'ai appliqué la méthode LM-MLEM aux données acquises avec Caliste, que j'ai présentées au paragraphe 2.3.3. J'ai discrétisé l'espace en pixels de taille de 4°x4°, contrairement à ce que j'avais présenté pour la *backprojection* directe, car cela augmente le nombre de cônes passant par chacune des positions du ciel. J'ai utilisé 50 itérations de l'algorithme LM-MLEM. Le nombre d'itérations peut être optimisé suivant le cas d'application, mais tout comme l'algorithme MLEM, l'absence de critère d'arrêt rigoureusement défini est un des principaux inconvénients de la méthode. Les résultats que j'ai obtenus sont présentés en Figure 3-1 et avec la superposition optique dans le cas des deux sources en Figure 3-2.

On constate en particulier que l'algorithme retrouve la position des sources où l'intensité renvoyée est maximale lorsqu'il n'y a qu'une seule source ponctuelle. En revanche, la largeur à mi-hauteur sur la zone la plus intense pour une source ponctuelle unique est de 8° pour la source en $(10°, 0°)$ et de 24° pour la source à $(30°, 0°)$.

On note aussi que l'algorithme n'efface pas de manière efficace le bruit Compton, et de nombreuses zones fantômes sont présentes dans l'image reconstruite. En particulier, pour l'exemple avec les deux sources, l'algorithme ne permet pas de déterminer la présence des deux sources de manière indépendante.

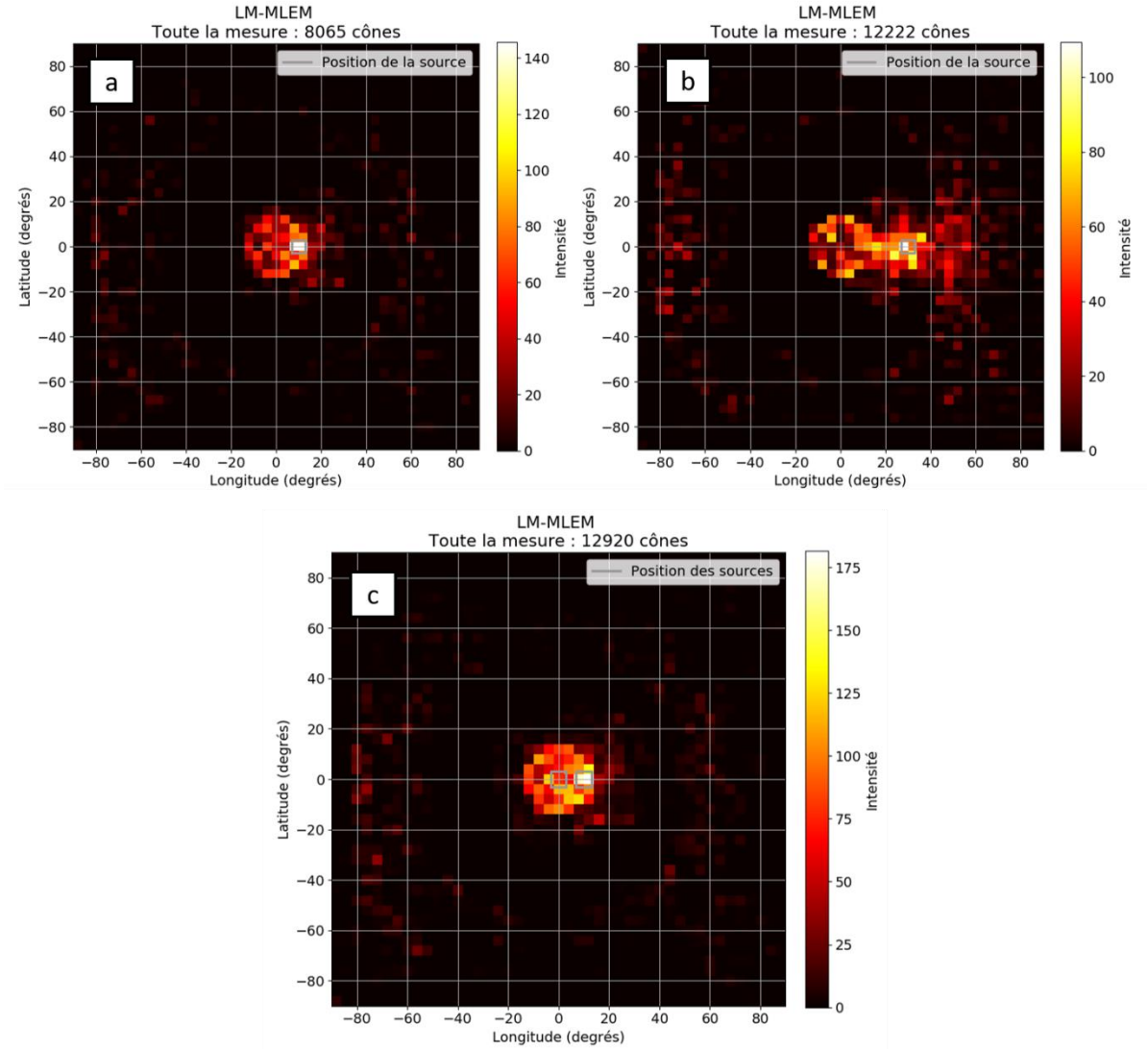


Figure 3-1 : Résultats LM-MLEM avec des sources de ^{137}Cs. a : Source en $(10°,0°)$ avec 8065 cônes. b : Source en $(30°,0°)$ avec 12 222 cônes. c : Sources en $(0°,0°)$ et $(10°, 0°)$ avec 12 920 cônes.

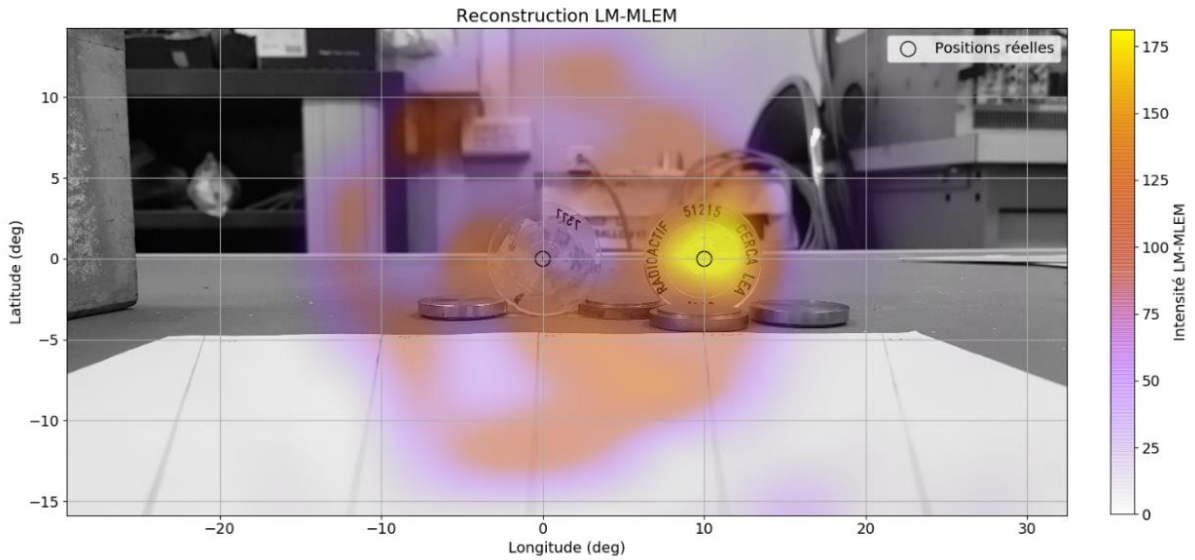


Figure 3-2 : Superposition optique de la scène avec les deux sources et du résultat de l'algorithme LM-MLEM.

Par ailleurs, un dernier inconvénient de cet algorithme est la vitesse d'exécution, étant donné qu'il faut exécuter une *backprojection* directe sur l'ensemble des événements Compton à chaque itération. De ce fait, le temps de calcul pour 8065 événements est d'environ 18 minutes et il est proportionnel au nombre d'événements Compton à traiter. Ce temps de calcul peut aussi être amélioré en optimisant la parallélisation des calculs par exemple.

3.2 SOE-RR : *Stochastic Origin Ensemble Resolution Recovery*

3.2.1 Principe de la méthode

L'algorithme SOE est un algorithme stochastique basé sur les méthodes de Monte-Carlo par chaînes de Markov (MCMC pour *Markov Chains Monte-Carlo*) (Andreyev, et al., 2010) afin d'effectuer la localisation par effet Compton. L'idée de cet algorithme est d'associer chaque événement à une seule position sur le ciel en utilisant le schéma suivant.

À l'initialisation, pour chaque événement i, une position aléatoire x_i^0 est tirée sur le cône correspondant à l'événement i. Pour ce faire, lorsque j'effectue la *backprojection* telle que décrite en paragraphe 2.3.2, je ne discrétise pas l'ensemble de la projection du cône avec 5000 éléments, mais je tire une seule valeur azimutale α aléatoirement entre 0 et 2π.

Puis, à chaque étape $n > 0$ de l'algorithme, pour chaque événement i, je tire une nouvelle position temporaire sur le cône correspondant à cet événement i, que je note $\hat{x}_i^n$. Cette nouvelle position est acceptée suivant une probabilité dite de « transition » $p\left(x_i^{n-1} \to \hat{x}_i^n\right)$, que je décris ensuite. Si la nouvelle position est acceptée, alors elle devient la nouvelle position associée à l'événement i : $x_i^n \coloneqq \hat{x}_i^n$. Si elle est rejetée, alors l'ancienne position est conservée pour l'événement i : $x_i^n \coloneqq x_i^{n-1}$. C'est à cette étape qu'intervient l'aspect Chaînes de Markov de la méthode, et en particulier à travers l'application de l'algorithme de Metropolis-Hastings (Metropolis, et al., 1953) (Hastings, 1970) qui est l'un des principaux algorithmes utilisés pour les MCMC.

Le résultat de l'algorithme consiste à faire l'histogramme des positions x_i^n à l'étape finale ou alors de considérer la moyenne de ces histogrammes en prenant aléatoirement certaines étapes parmi les dernières itérations. Pour ma part, j'utilise 20 000 étapes et j'effectue la moyenne de 100 étapes tirées aléatoirement parmi les 5000 dernières étapes. J'ai en effet pu constater à travers les tests que 20 000 étapes étaient suffisantes pour atteindre la convergence de l'algorithme avec les données que j'utilisais. Cela peut être modifié suivant les données et le système de détection utilisé. Actuellement, il n'existe pas véritablement de critère pour optimiser ce choix.

La probabilité de transition $p\left(x_i^{n-1} \to \hat{x}_i^n\right)$ se calcule en comparant la pertinence de la position $\hat{x}_i^n$ par rapport à la position x_i^{n-1} (Sitek, 2008) en considérant le fait que passer à la position $\hat{x}_i^n$ augmente de 1 le nombre d'événements à cette position et diminue de 1 celui à la position x_i^{n-1}. La pertinence de cette opération est évaluée en équation (3-4), en comparant l'impact sur le nombre d'événements $D^n(x)$ associés à la position x à l'étape n :

$$p(x_i^{n-1} \to \hat{x}_i^n) = \min\left(1, \frac{\alpha_i(\hat{x}_i^n)}{\alpha(x_i^{n-1})} \frac{(D^n(\hat{x}_i^n)+1)^{D^n(\hat{x}_i^n)+1}}{D^n(\hat{x}_i^n)^{D^n(\hat{x}_i^n)}} \frac{D^n(x_i^{n-1}-1)^{D^n(x_i^{n-1})-1}}{D^n(x_i^{n-1})^{D^n(x_i^{n-1})}} \frac{\varepsilon(x_i^{n-1})}{\varepsilon(\hat{x}_i^n)}\right) \quad (3\text{-}4)$$

Dans cette équation, je note $\varepsilon(x)$ l'efficacité du détecteur à la position x, c'est-à-dire, la probabilité

qu'un photon émis à la position x soit effectivement détecté. Je note $\alpha_i(x)$ la probabilité que l'événement i provienne effectivement de la position x.

Je simplifie l'équation (3-4) sous quelques hypothèses. D'une part, je considère l'efficacité ε comme uniforme pour toutes les positions, vu qu'il n'y a pas de raison a priori pour qu'une position du ciel soit privilégiée plutôt qu'une autre. Cela peut être différent si on considère d'autres géométries de détecteur ou la présence d'un blindage. D'autre part, du fait que je tire les positions x_i aléatoirement sur le cône lié à l'événement i, la probabilité α_i liée à cette position est uniforme sur l'ensemble du cône a priori. Ainsi, avec ces deux remarques, les deux rapports $\frac{\alpha_i(\hat{x}_i^n)}{\alpha(x_i^{n-1})}$ et $\frac{\varepsilon(x_i^{n-1})}{\varepsilon(\hat{x}_i^n)}$ valent 1 et peuvent être simplifiés.

J'obtiens alors l'expression suivante de la probabilité de transition :

$$p\left(x_i^{n-1} \rightarrow \hat{x}_i^n\right) = \min\left(1, \frac{(D^n(\hat{x}_i^n)+1)^{D^n(\hat{x}_i^n)+1}}{D^n\left(\hat{x}_i^n\right)^{D^n(\hat{x}_i^n)}} \; \frac{D^n\left(x_i^{n-1}-1\right)^{D^n(x_i^{n-1})-1}}{D^n\left(x_i^{n-1}\right)^{D^n(x_i^{n-1})}}\right) \tag{3-5}$$

Cette probabilité illustre le fait que la nouvelle position $\hat{x}_i^n$ sera acceptée avec une probabilité de 100 % si cette nouvelle position $\hat{x}_i^n$ correspond déjà à un plus grand nombre d'événements $D^n(\hat{x}_i^n)$ par rapport à l'ancienne position x_i^{n-1}, c'est-à-dire si $D^n(\hat{x}_i^n) > D^n\left(x_i^{n-1}\right)$. Sinon, la nouvelle position sera acceptée avec une probabilité donnée par le terme de droite de l'équation (3-5).

Pour des raisons de performances et de parallélisation du calcul, je traite l'ensemble des événements i à chaque étape n. Plusieurs variantes peuvent exister en ne traitant qu'un seul événement i à chaque étape n ou alors un sous-ensemble des événements à chaque étape. Ces différentes pistes doivent être comparées au niveau des performances.

L'intérêt de cet algorithme réside dans le fait particulier que l'on peut appliquer une variante tenant compte de la résolution spatiale et spectrale du détecteur, variante appelée SOE *Resolution Recovery* SOE-RR (Andreyev, et al., 2011). À chaque étape n de l'algorithme, pour chacun des événements i, on peut tenir compte de la résolution du détecteur de la manière suivante : les positions des interactions pour l'événement i sont (x, y, z) associées à une incertitude $(\Delta x, \Delta y, \Delta z)$, et de même, l'énergie mesurée pour l'absorption photoélectrique E est associée à une incertitude ΔE. Ainsi, au lieu d'utiliser la position (x, y, z) et l'énergie E, on peut tirer aléatoirement la position et l'énergie suivant leurs résolutions respectives à chaque étape n de sorte à projeter un nouveau cône pour l'événement i.

Pour ma part, j'applique cette variante en ne tenant compte que de l'incertitude sur la profondeur d'interaction y, qui contribue le plus à l'incertitude totale sur la reconstruction. Pour ce faire, pour chaque événement i, je tire aléatoirement la profondeur de l'interaction Compton et de l'absorption photoélectriques indépendamment l'une de l'autre, aléatoirement entre 0 et 1 mm. En effet, à une énergie de 662 keV, la probabilité d'interaction en fonction de la profondeur est proche d'être uniforme.

3.2.2 Tests SOE-RR sur des données réelles acquises avec Caliste

Je teste ici l'algorithme SOE-RR avec les données de sources ponctuelles acquises avec Caliste. Je représente les résultats avec deux discrétisations, une en 2°x2° et une en 4°x4°. La discrétisation est effectivement un paramètre important car elle a une influence directe sur la probabilité de transition

calculée en équation (3-5), à travers le nombre $D^n(x)$ d'événements associés à la position x. Une discrétisation plus fine va diminuer ce nombre $D^n(x)$, ce qui peut empêcher la convergence de l'algorithme. Par ailleurs, la limite liée à l'élargissement Doppler empêche d'atteindre une résolution meilleure que 2°, d'où le fait que je me limite au mieux à une grille de 2°x2°. Il est aussi possible d'appliquer la méthode *Resolution Recovery* en prenant en compte cette incertitude sur l'élargissement Doppler. Par ailleurs, il est à noter que, dans l'état de l'art, les développements de cet algorithme et les tests sont faits systématiquement à travers de données simulées, et que la mise en œuvre sur des données réelles est nouvelle à ma connaissance.

Je présente aussi les résultats en utilisant l'ensemble de la mesure, ainsi qu'en utilisant le minimum de cônes requis pour obtenir une localisation correcte de la source, ce qui est représentatif de la sensibilité. L'intensité qui est donnée correspond au nombre d'événements associés à la position par l'algorithme SOE-RR.

La Figure 3-3 correspond à la mesure à $(10°, 0°)$ et la Figure 3-4 correspond à la mesure à $(30°, 0°)$.

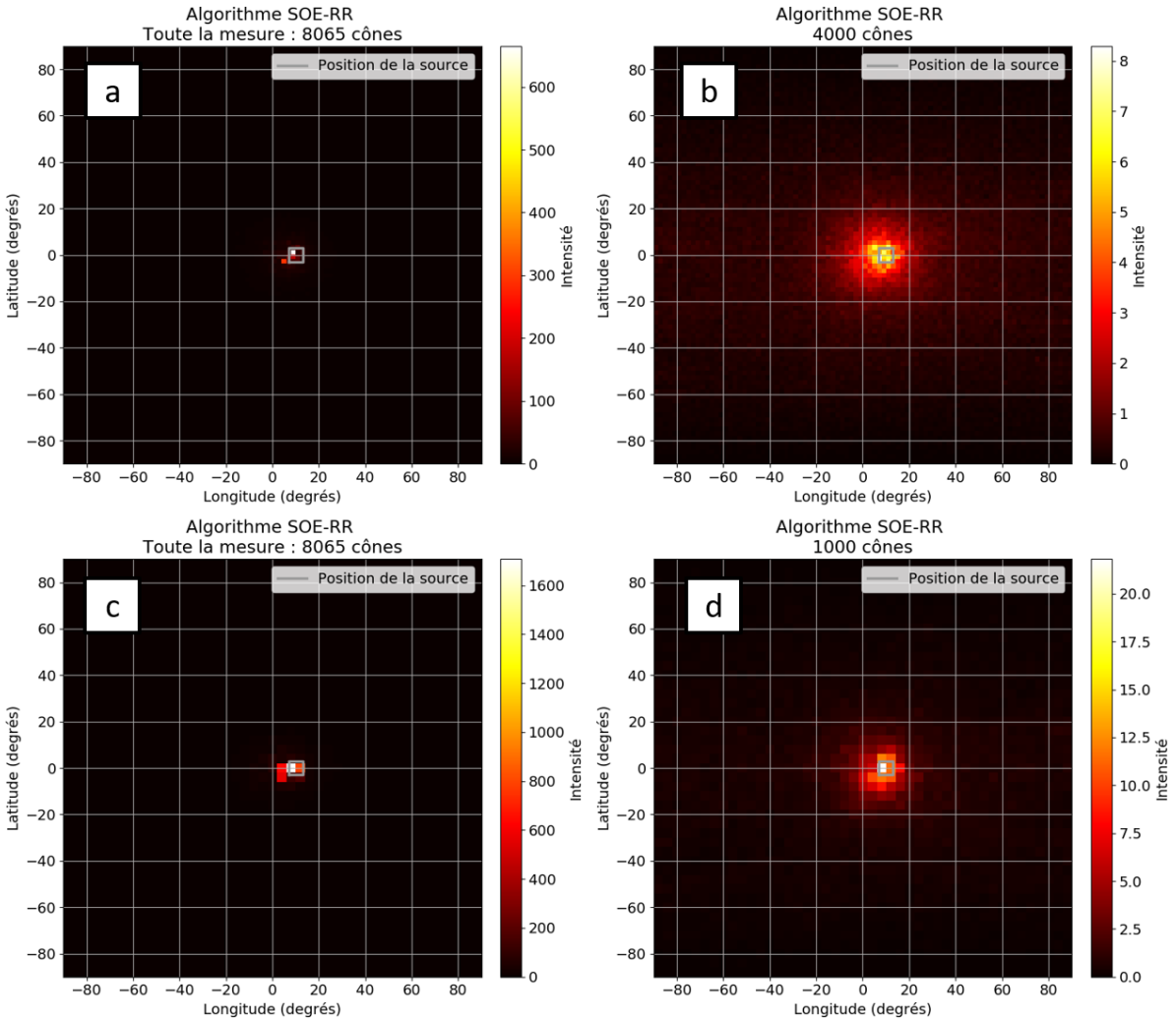


Figure 3-3 : SOE-RR – Mesure de ^{137}Cs à $(10°, 0°)$. a : Toute la mesure (8065 cônes) avec discrétisation 2°x2°. b : Utilisation des 4000 premiers cônes avec discrétisation 2°x2°. c : Toute la mesure (8065 cônes) avec discrétisation 4°x4°. d : Utilisation des 1000 premiers cônes avec discrétisation 4°x4°.

Source en $(10°, 0°)$	2°x2°	4°x4°
FWHM sur toute la mesure	2,6°	7,3°
FWHM sur la mesure tronquée (nombre de cônes)	14,4° (4000 cônes)	10,4° (1000 cônes)

Tableau 3-1 : Résolution angulaire pour l'algorithme SOE-RR, source en $(10°, 0°)$.

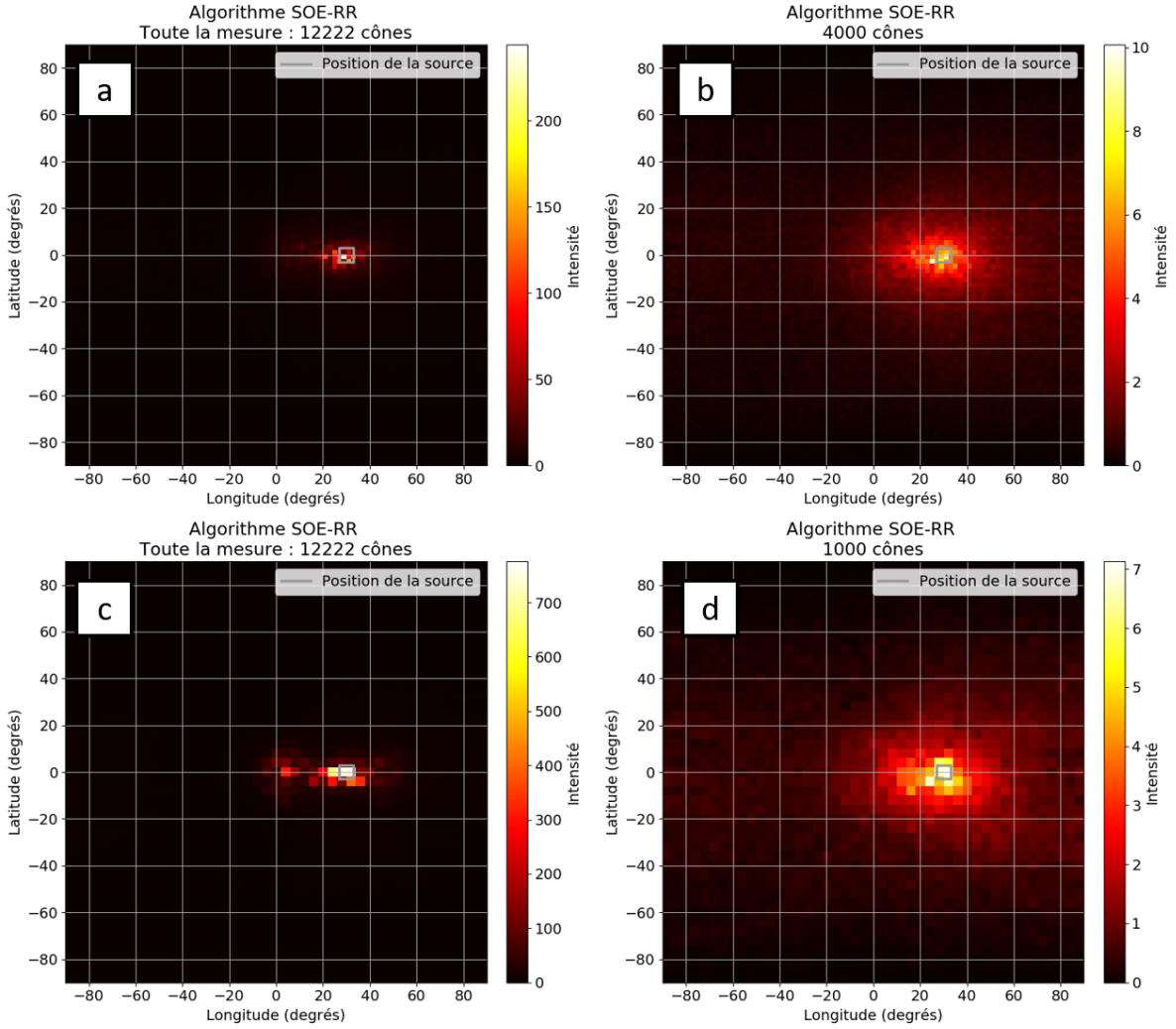


Figure 3-4 : SOE-RR – Mesure de ^{137}Cs à $(30°, 0°)$. a : Toute la mesure (12 222 cônes) avec discrétisation 2°x2°. b : Utilisation des 4000 premiers cônes avec discrétisation 2°x2°. c : Toute la mesure (12 222 cônes) avec discrétisation 4°x4°. d : Utilisation des 1000 premiers cônes avec discrétisation 4°x4°.

Source en $(30°, 0°)$	2°x2°	4°x4°
FWHM sur toute la mesure	7°	13,6°
FWHM sur la mesure tronquée (nombre de cônes)	14° (4000 cônes)	22° (1000 cônes)

Tableau 3-2 : Résolution angulaire pour l'algorithme SOE-RR, source en $(30°, 0°)$.

Avec l'ensemble de la mesure, l'algorithme reconstruit la position de la source, l'élément le plus intense de la reconstruction étant effectivement situé à la position attendue de la source. Les résultats sur la résolution sont synthétisés en Tableau 3-1 et Tableau 3-2.

Par ailleurs, afin de reconstruire la position de la source avec une précision de l'ordre de 15°, il est nécessaire d'avoir environ 1000 cônes Compton pour une discrétisation de 4°x4°. De même pour une discrétisation de 2°x2°, 4000 cônes Compton sont nécessaires. La dégradation d'un facteur 4 de la sensibilité entre les deux discrétisations provient principalement du changement de taille d'un facteur 4 pour les pixels de l'image reconstruite. On note aussi que, dans les deux cas, le changement de résolution ne va pas dans le même sens, elle s'améliore avec une discrétisation plus fine pour la source en $(30°, 0°)$ et se dégrade pour la source en $(10°, 0°)$. Ce constat peut provenir de l'aspect stochastique de l'algorithme, et on pourrait relancer l'algorithme plusieurs fois pour vérifier si la résolution est modifiée ou non à chaque application de l'algorithme.

Les Figure 3-5 et Figure 3-6 montrent le résultat pour la mesure avec les deux sources à $(0°, 0°)$ et à $(10°, 0°)$ avec une discrétisation de 2°x2° et de 4°x4° respectivement.

On voit sur la Figure 3-5 que l'algorithme n'est pas capable de localiser les deux sources avec une discrétisation de 2°x2°, même en utilisant toute la mesure, c'est pourquoi je n'ai pas étudié le résultat avec moins de cônes. Globalement l'algorithme identifie une zone étendue sur 10° dans laquelle se situent effectivement les sources.

Sur la Figure 3-6, on voit qu'avec l'ensemble de la mesure, l'algorithme est capable de retrouver la zone dans laquelle se situent les deux sources, zone d'une taille angulaire de 12°. L'intensité de la source à $(0°, 0°)$ est comparable à celle de de la source en $(10°, 0°)$. En utilisant seulement 5000 cônes, l'algorithme converge sur la position entre les deux sources, ce qui ne couvre pas correctement les deux sources.

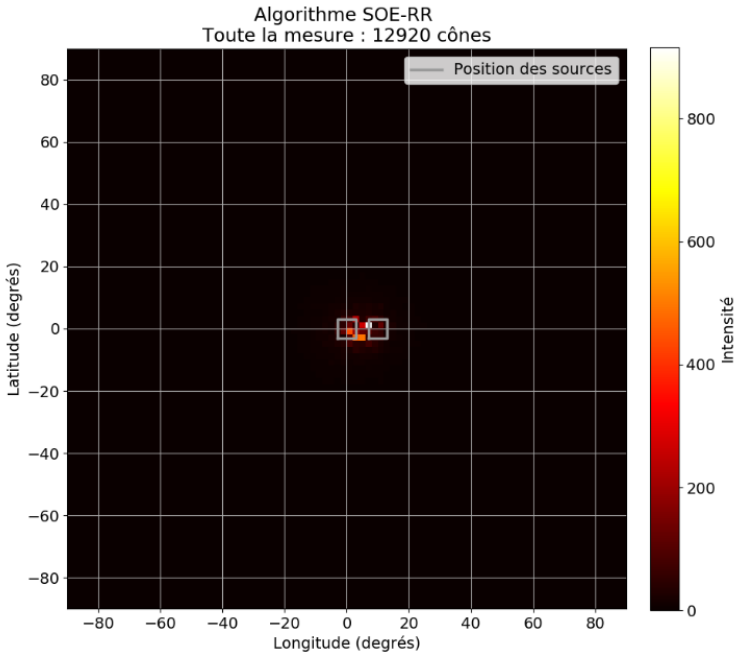


Figure 3-5 : SOE-RR – Mesure pour deux sources de ^{137}Cs à $(0°, 0°)$ et $(10°, 0°)$, toute la mesure (12 920 cônes) avec discrétisation 2°x2°.

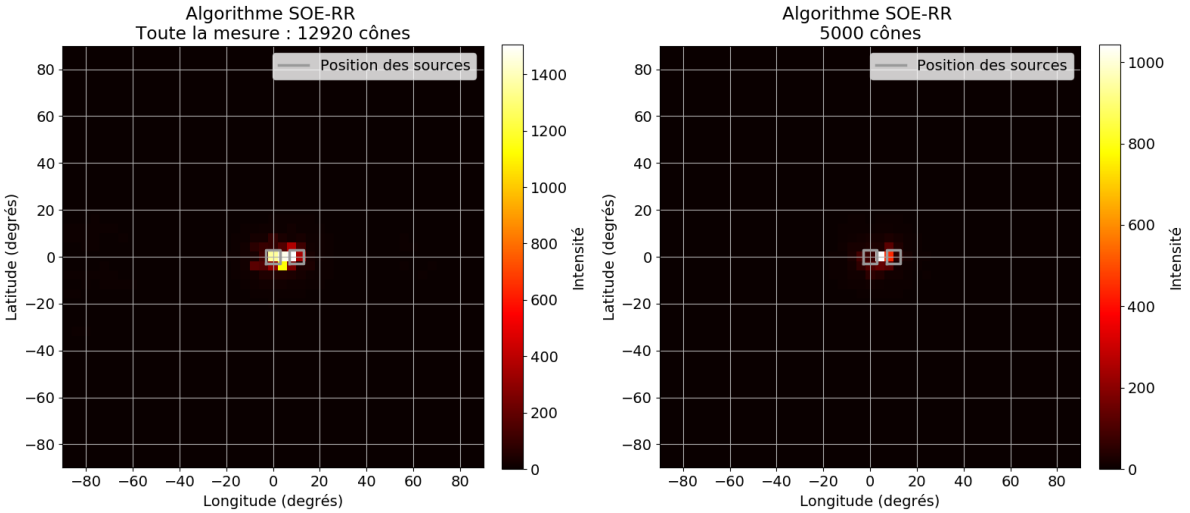


Figure 3-6 : SOE-RR – Mesure pour deux sources de ^{137}Cs à $(0°, 0°)$ et $(10°, 0°)$. À gauche : Toute la mesure (12 920 cônes) avec discrétisation 4°x4°. À droite : Utilisation des 5000 premiers cônes avec discrétisation 4°x4°.

J'ai représenté en Figure 3-7 la superposition optique avec le résultat issu de l'ensemble de la mesure avec une discrétisation de 4°x4°. Cela permet de rendre compte de la précision de l'algorithme, qui est de 7°, ce qui fait que l'on peut s'attendre à une superposition gamma dans la zone entre les deux sources, avec la possibilité quand même de distinguer la présence des deux sources.

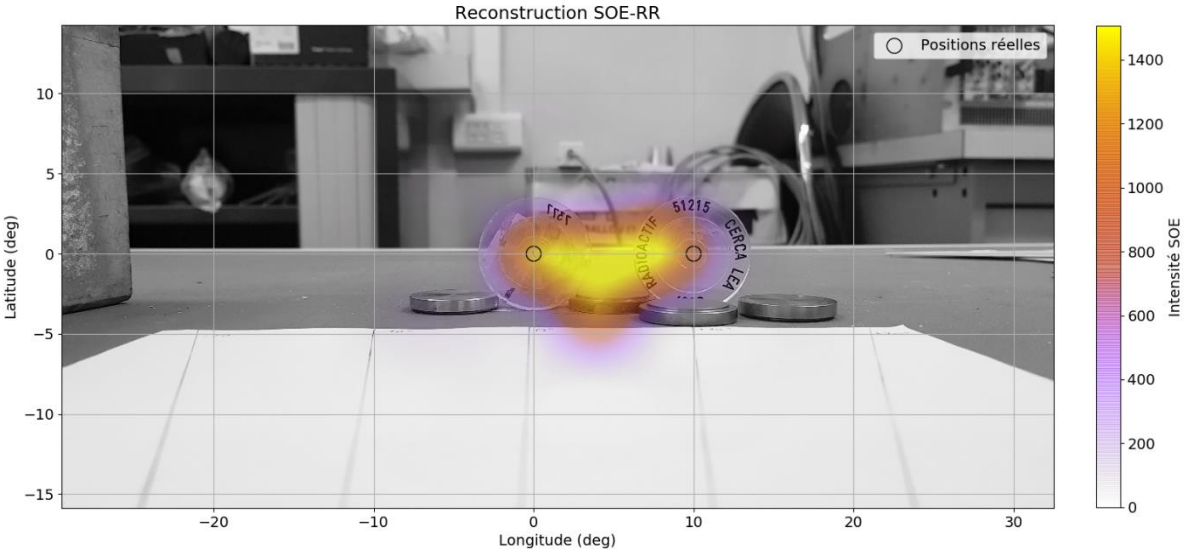


Figure 3-7 : Superposition optique de la scène avec les deux sources et du résultat de l'algorithme SOE-RR pour une discrétisation de 4°x4°

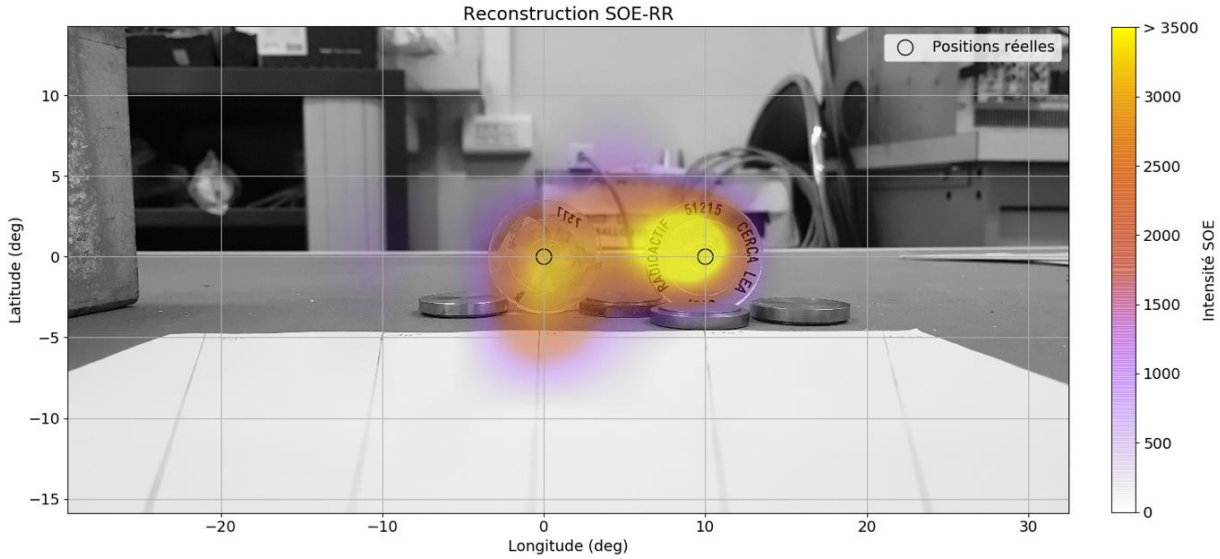


Figure 3-8 : Superposition optique de la scène avec les deux sources et du résultat de l'algorithme SOE-RR pour une discrétisation de 4°x4° avec un décalage de 2°.

J'ai effectué une reconstruction SOE avec une discrétisation de 4°x4° mais au lieu de l'effectuer entre -90° et +90°, j'ai appliqué un décalage de 2° de sorte à ce qu'elle soit entre -88° et +92°. De ce fait, la source à reconstruire à 10° en longitude est située au centre d'un pixel de discrétisation et non plus sur le bord d'un pixel, comme c'était le cas sans le décalage. En appliquant ainsi l'algorithme SOE et en superposant avec l'image optique, on obtient la Figure 3-8 avec une séparation des deux sources. Cette technique dite de *dithering* (Encyclopedia of Multimedia, 2008) qui consiste à effectuer ce décalage d'un demi-pixel de discrétisation peut être envisagée pour obtenir une évaluation de l'erreur de l'algorithme SOE due à la discrétisation utilisée.

En ce qui concerne le temps de calcul, pour traiter 8065 événements, le temps de calcul est de 197 s, soit un peu plus de 3 minutes, et donc 6 fois plus rapide que l'algorithme LM-MLEM. Le temps de calcul augmente linéairement avec le nombre d'événements à traiter et le nombre d'itérations utilisées dans l'algorithme.

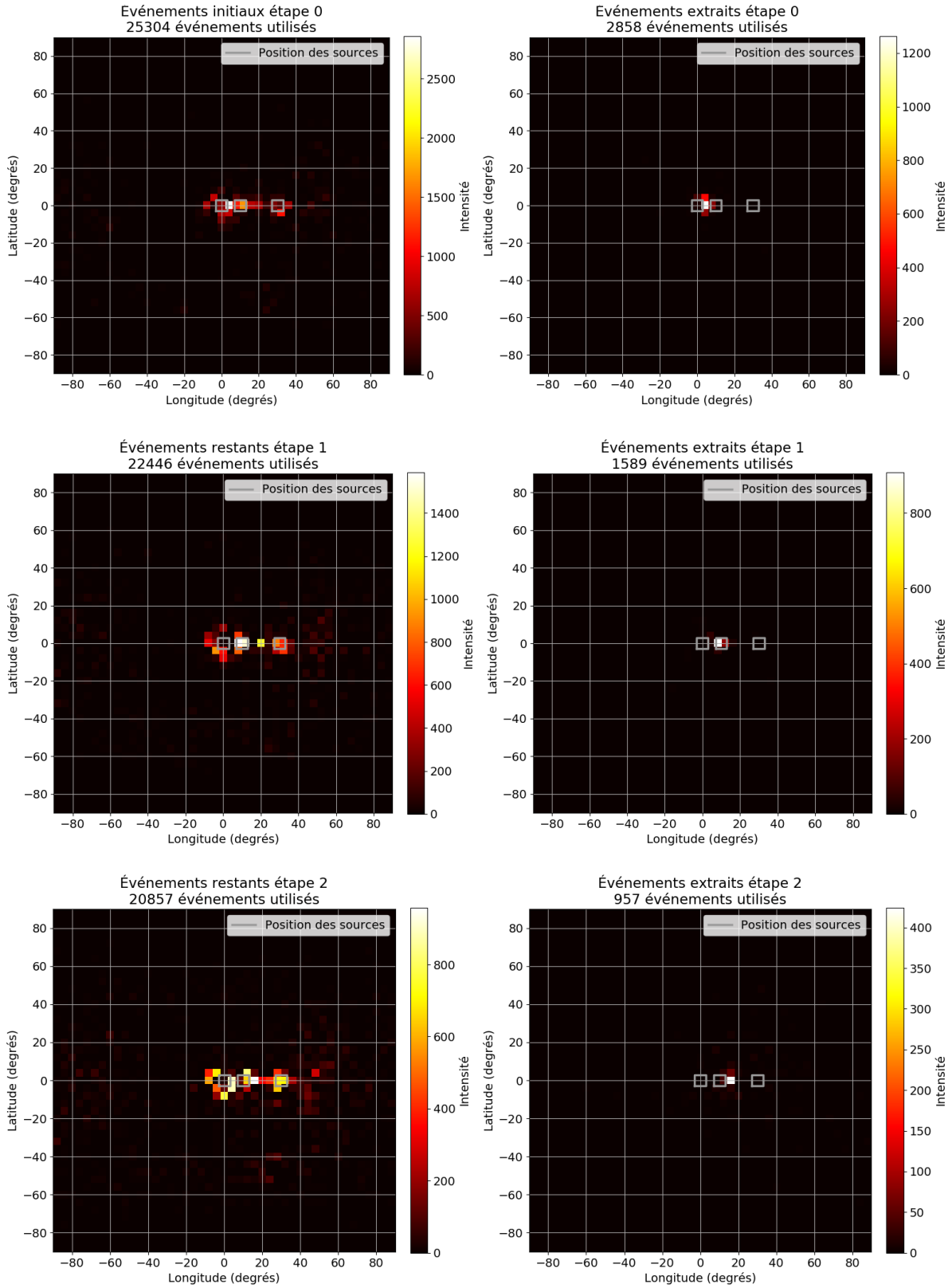
3.2.3 Soustraction itérative de sources avec l'algorithme SOE-RR

L'algorithme SOE-RR présente un avantage considérable du fait que, à la fin des itérations, en prenant en compte le dernier résultat de l'algorithme[*], on peut associer chaque événement Compton à une position particulière.

J'ai ainsi mis en place une nouvelle méthode qui soustrait itérativement les sources en procédant de la manière suivante. La première étape consiste à appliquer l'algorithme SOE-RR et à prendre le résultat à la dernière itération. J'identifie alors la position contenant le maximum d'événements. J'extrais les événements correspondant à cette position et je leur applique l'algorithme SOE-RR afin de vérifier qu'ils correspondent bien à la position trouvée. Vérification faite, j'applique l'algorithme SOE-RR à l'ensemble des événements restants de la mesure. Je poursuis itérativement cette méthode

[*] Et non la moyenne sur un ensemble de résultats, comme je l'ai présenté précédemment.

jusqu'à ce qu'il n'y ait plus suffisamment d'événements pour que l'algorithme SOE-RR puisse converger.



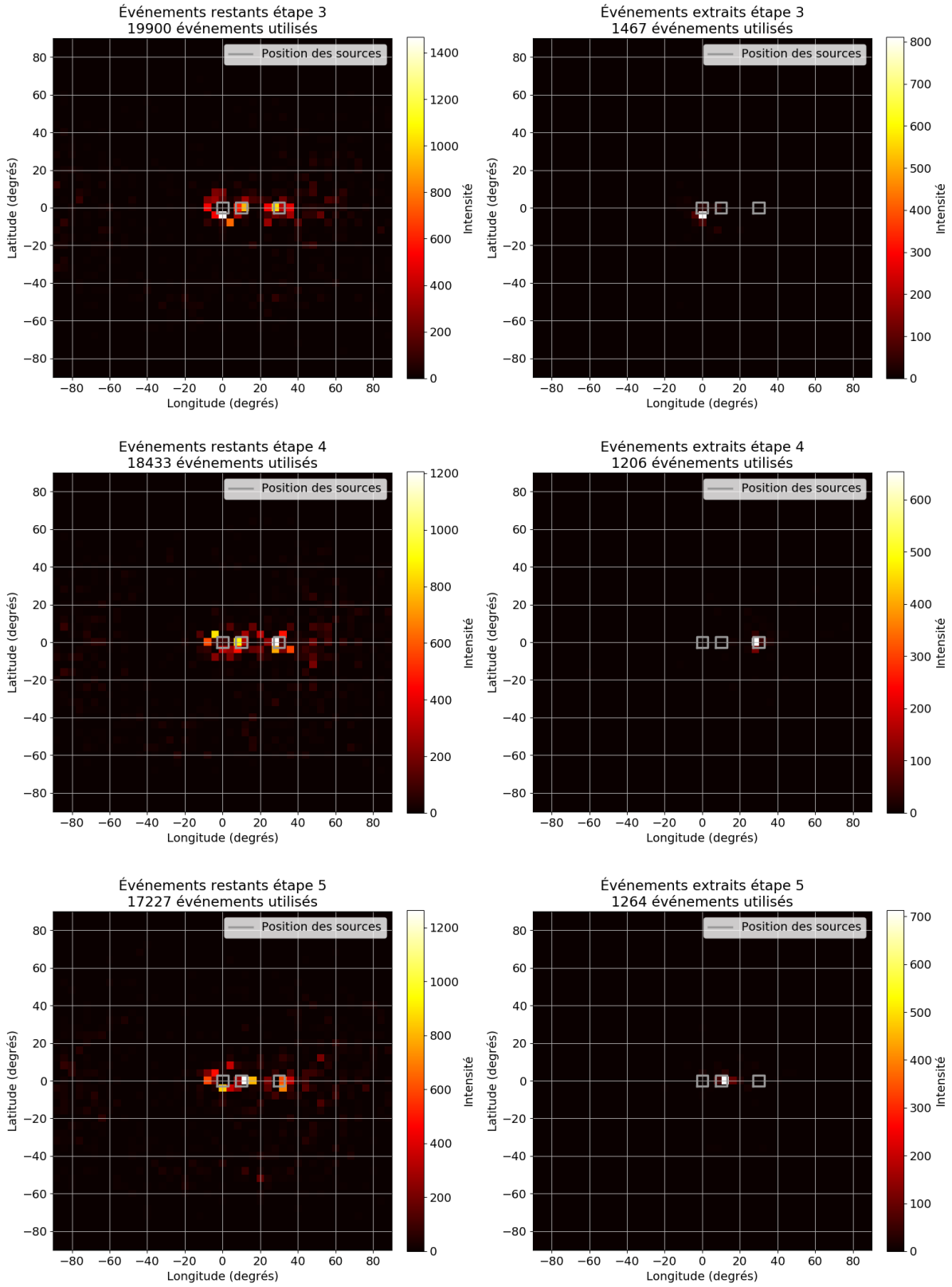


Figure 3-9 : Application de l'algorithme SOE-RR avec trois sources et six itérations soustractives. La colonne de gauche correspond à l'application de l'algorithme aux événements restants de l'étape n. La colonne de droite correspond à son application aux événements extraits à l'étape n.

Pour illustrer cette méthode, je l'applique avec un mélange des trois mesures, à $(0°, 0°)$, $(10°, 0°)$ et $(30°, 0°)$, afin d'avoir un mélange avec trois sources. La Figure 3-9 suivante montre les six premières étapes de cette soustraction itérative, en utilisant une discrétisation de 4°x4°. La colonne de gauche correspond à l'algorithme SOE-RR appliqué à l'ensemble des événements restants, c'est-à-dire l'ensemble des événements à l'étape initiale et pour une étape n suivante, les événements de l'étape $n-1$ auxquels j'ai retiré les événements extraits. La colonne de droite montre le résultat de l'algorithme SOE-RR appliqué aux événements extraits à cette étape, c'est-à-dire, les événements qui ont été identifiés dans la position d'intensité maximale à cette étape sur l'image de gauche.

Dans les événements extraits dans la colonne de droite, on voit que, dans la limite de la précision de 7° que j'ai établie en paragraphe 3.2.2, les étapes 0 et 3 ont extrait des événements correspondant à la source en $(0°, 0°)$, les étapes 1, 2 et 5 à la source en $(10°, 0°)$ et l'étape 4 à la source en $(30°, 0°)$.

Ce processus est long à exécuter car il faut compter environ 5 minutes de calcul à chaque étape pour appliquer l'algorithme SOE-RR sur les 17 000 à 25 000 événements utilisés. Cependant, cette méthode montre l'intérêt de l'algorithme SOE-RR pour avoir une information plus précise sur la localisation de plusieurs sources dans un mélange. En particulier, cela permet de confirmer la présence de plusieurs sources dans un post traitement par exemple.

3.3 Algorithme bayésien de reconstruction Compton

3.3.1 Principe de la méthode

Afin de traiter la reconstruction Compton, j'ai développé un algorithme basé sur la formule de Bayes, reprenant l'idée de l'algorithme donné par (Bobin, et al., 2016), dédié à la spectro-identification de sources gamma. J'avais déjà donné une application de cet algorithme dans le cas de l'inversion spectrale Compton en Chapitre 3.

L'algorithme que j'ai mis en place ici repose sur l'idée d'effectuer la *backprojection*, mais en tenant compte des cônes qui ont été déjà tracés, par application de la formule de Bayes. Pour ce faire, sur l'image à reconstruire, j'attribue un compteur $k(j)$ à chaque position j. J'initialise ce compteur à 0.

J'attribue une probabilité $p(j)$ qui constitue une probabilité a priori sur la localisation des sources dans le « ciel ». Je choisis cette probabilité comme étant uniforme sur le ciel, normalisée en somme à 1, mais il est aussi envisageable, et même encouragé, d'ajouter de l'information si elle est disponible concernant la géométrie du détecteur ou sur l'environnement à analyser. Par exemple, si certaines parties du détecteur sont blindées, certaines directions peuvent être enlevées ou atténuées.

Pour chaque événement i, je calcule la probabilité a posteriori $p(j|i)$ pour chaque position j. C'est la probabilité que cette position soit à l'origine de l'événement i. Pour cela, j'applique le théorème de Bayes :

$$p(j|i) = \frac{p(j)p(i|j)}{\sum_{j'} p(j')p(i|j')} \tag{3-6}$$

Dans cette équation, $p(j)$ est le *prior* actuel sur la localisation des sources.

$p(i|j)$ est la probabilité d'observer l'événement i parmi tous les événements originaires de la position j. Il correspond à la probabilité donnée par l'équation (3-3), aussi présente pour l'algorithme LM-

MLEM. De la même manière que pour cet algorithme, les termes se simplifient dans le ratio donné par l'équation (3-6) et je considère cette probabilité comme uniforme sur les positions j appartenant au cône défini par l'événement i, et nulle sur le reste des positions j en dehors du cône.

Puis, je mets à jour le compteur k associé à chaque position par :

$$k(j) \coloneqq k(j) + p(j|i) \tag{3-7}$$

Enfin, je mets à jour la probabilité *prior*, suite à chaque événement i de la manière suivante :

$$p(j) = \frac{k(j)}{\sum_{j'} k(j')} \tag{3-8}$$

La mise à jour du *prior* doit être faite après avoir accumulé des événements, pour éviter que l'utilisation seule des premiers événements ne fixe à 0 la probabilité $p(j)$ à certaines positions j. Avec l'équation (3-6), si $p(j)$ vaut 0 à une étape, cela amènerait à conserver cette probabilité de 0 tout au long du calcul. J'ai empiriquement choisi de mettre à jour le *prior* à partir du moment où 10 % des événements ont été traités.

Après avoir traité l'ensemble des événements i, les valeurs de $p(j)$ donnent une carte de probabilité de présence des sources dans le ciel à chaque position j.

Pour finir sur la description de l'algorithme, (Bobin, et al., 2016) propose plusieurs variantes de l'algorithme bayésien dans le cas de la spectro-identification, variantes qui peuvent être aussi implémentées dans le cas d'une application à la reconstruction Compton.

3.3.2 Tests sur des données réelles acquises avec Caliste

Tout comme pour l'algorithme SOE-RR, je teste l'algorithme bayésien sur les mesures que j'ai acquises avec Caliste, avec des discrétisations du ciel à reconstruire de 2°x2° et de 4°x4°, en étudiant le nombre de cônes nécessaires pour pouvoir effectuer cette reconstruction. La Figure 3-10 présente le résultat pour la source située à la position $(10°, 0°)$, la Figure 3-11 pour la source à la position $(30°, 0°)$.

Comme pour le cas de l'algorithme SOE, le changement de discrétisation influe de manière différente sur la résolution en fonction de la position des sources. Il est possible que cela provienne de l'ordre dans lequel les événements sont traités, qui correspond à leur ordre dans la mesure. Afin de vérifier cela, il serait envisageable de traiter les mesures en appliquant plusieurs fois l'algorithme bayésien, mais en modifiant aléatoirement l'ordre des événements, ou alors en évaluant les performances de l'algorithme sur différentes mesures dans les mêmes conditions.

Par ailleurs, pour pouvoir reconstruire la position de la source, il est nécessaire d'utiliser de l'ordre de 1000 cônes Compton pour une discrétisation de 4°x4° pour les sources à $(10°, 0°)$ et à $(30°, 0°)$. De même pour une discrétisation de 2°x2°, il faut avoir 1000 cônes Compton pour la source à $(10°, 0°)$ et 2000 cônes pour la source à $(30°, 0°)$.

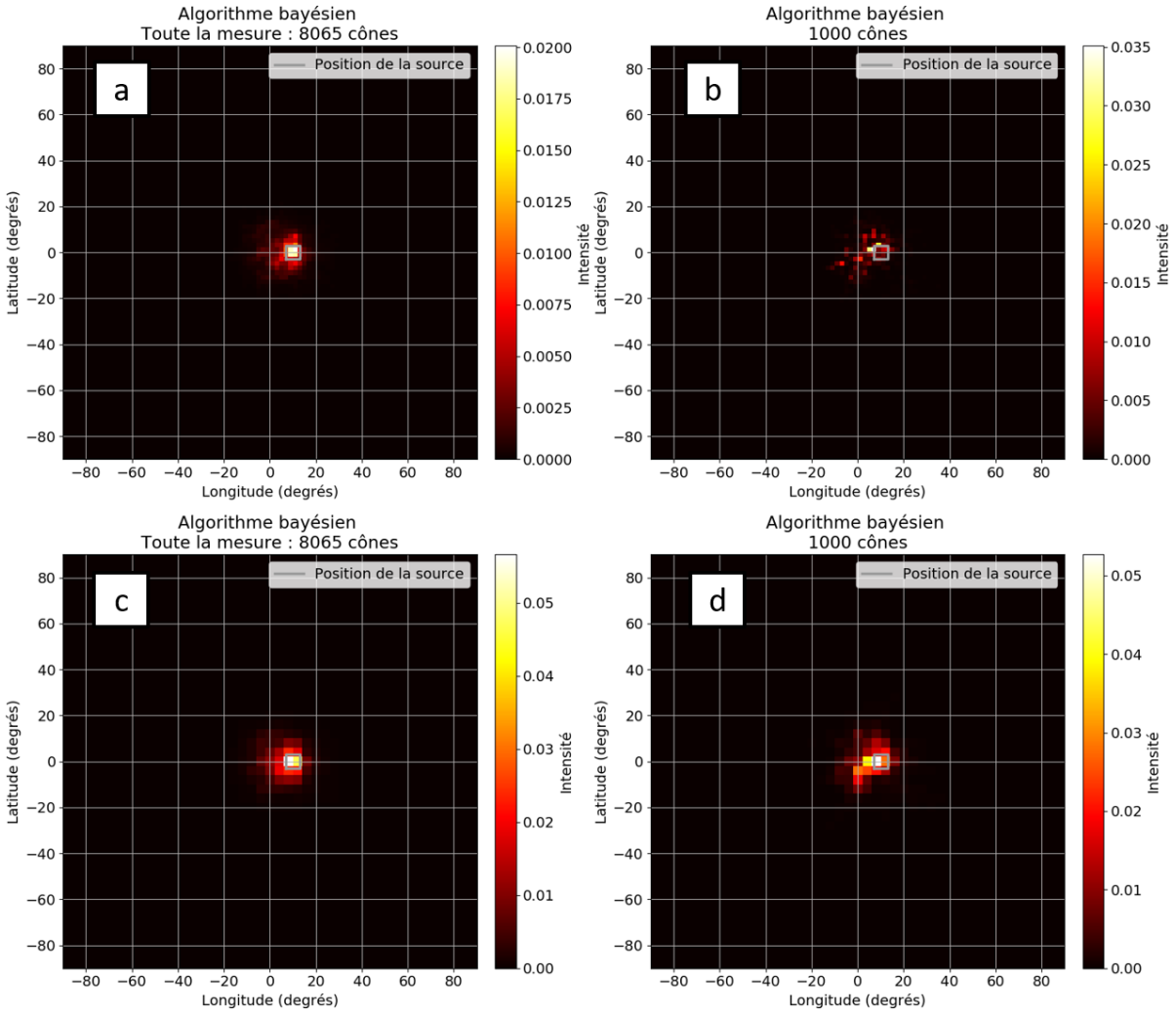


Figure 3-10 : Algorithme bayésien – Mesure de ^{137}Cs à $(10°, 0°)$. a : Toute la mesure (8065 cônes) avec discrétisation 2°x2°. b : Utilisation des 1000 premiers cônes avec discrétisation 2°x2°. c : Toute la mesure (8065 cônes) avec discrétisation 4°x4°. d : Utilisation des 1000 premiers cônes avec discrétisation 4°x4°.

Source en $(10°, 0°)$	2°x2°	4°x4°
FWHM sur toute la mesure	5,4°	8,3°
FWHM sur la mesure tronquée (nombre de cônes)	6,2° (1000 cônes)	14° (1000 cônes)

Tableau 3-3 : Résolution angulaire pour l'algorithme bayésien, source en $(10°, 0°)$.

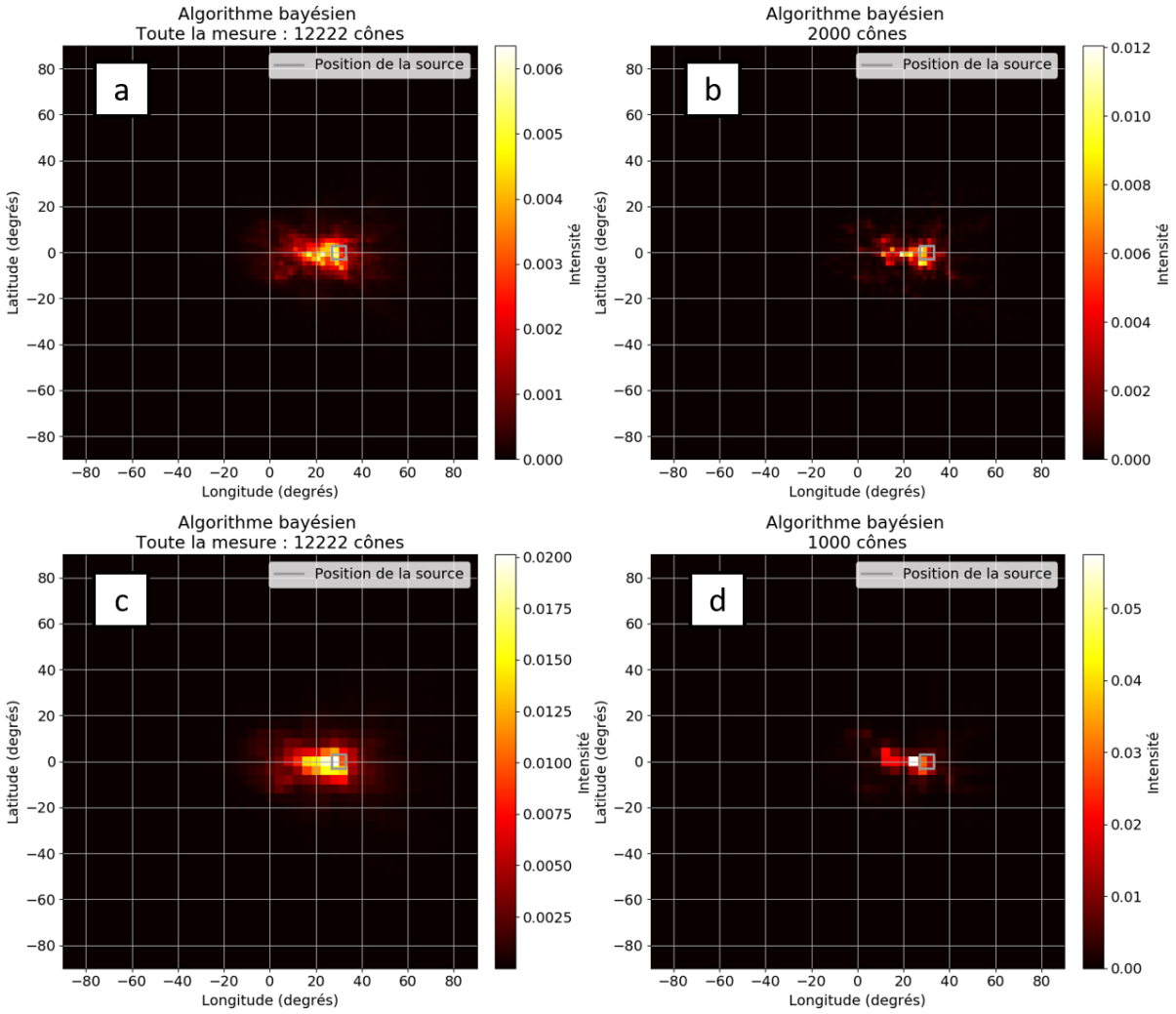


Figure 3-11 : Algorithme bayésien – Mesure de ^{137}Cs à $(30°, 0°)$. a : Toute la mesure (12 222 cônes) avec discrétisation 2°x2°. b : Utilisation des 2000 premiers cônes avec discrétisation 2°x2°. c : Toute la mesure (12 222 cônes) avec discrétisation 4°x4°. d : Utilisation des 1000 premiers cônes avec discrétisation 4°x4°.

Source en $(30°, 0°)$	2°x2°	4°x4°
FWHM sur toute la mesure	15,9°	20,4°
FWHM sur la mesure tronquée (nombre de cônes)	13° (2000 cônes)	8,1° (1000 cônes)

Tableau 3-4 : Résolution angulaire pour l'algorithme bayésien, source en $(30°, 0°)$.

La Figure 3-12 montre le résultat pour le mélange des deux sources aux positions $(0°, 0°)$ et $(10°, 0°)$ et la Figure 3-13 montre la superposition optique avec le résultat pour une discrétisation de 4°x4°.

On remarque notamment que l'algorithme a des difficultés à distinguer les deux sources de même énergie, même s'il permet de voir la présence potentielle de deux sources, dans la résolution de 8° environ, donnée par le Tableau 3-3.

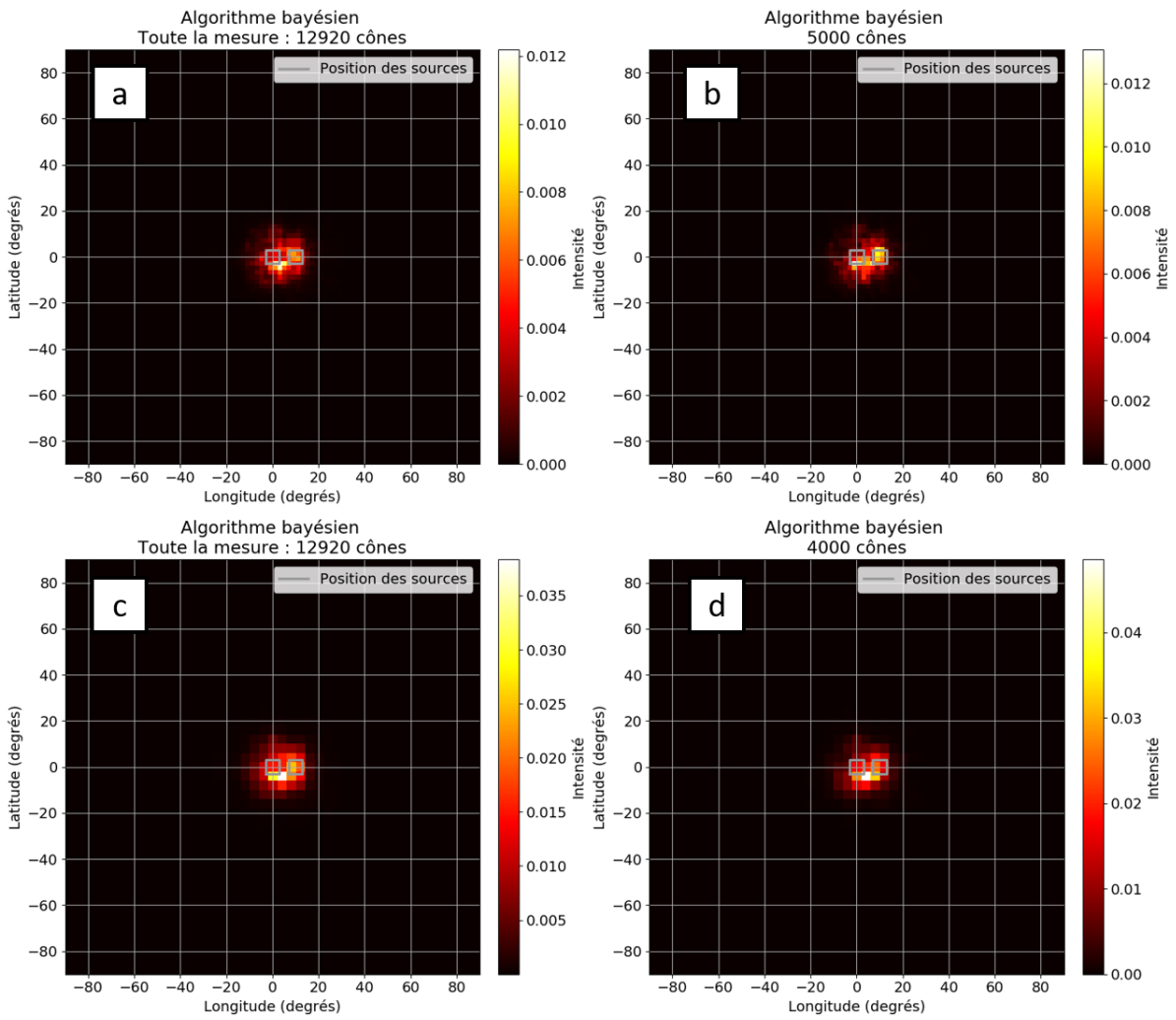


Figure 3-12 : Algorithme bayésien – Mesure de deux sources de ^{137}Cs à $(0°, 0°)$ et $(10°, 0°)$. a : Toute la mesure (12 920 cônes) avec discrétisation 2°x2°. b : Utilisation des 5000 premiers cônes avec discrétisation 2°x2°. c : Toute la mesure (12 920 cônes) avec discrétisation 4°x4°. d : Utilisation des 4000 premiers cônes avec discrétisation 4°x4°.

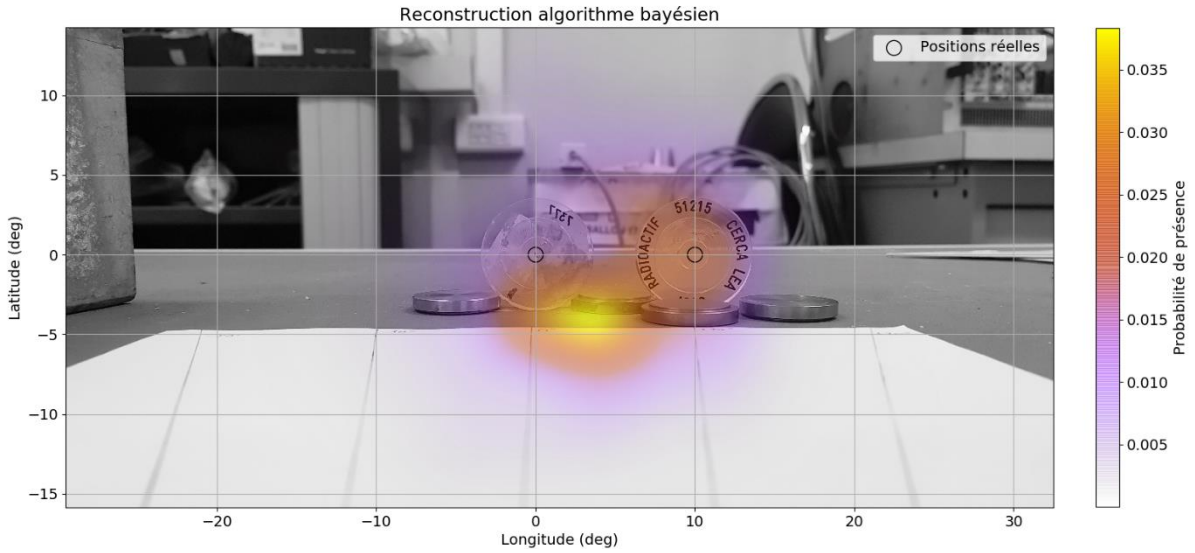


Figure 3-13 : Superposition optique de la scène avec les deux sources et du résultat de l'algorithme bayésien pour une discrétisation de 4°x4°

Pour finir, un des avantages de cet algorithme est son temps d'exécution, qui est égal à celui de la *backprojection* directe puisque l'ensemble des événements est traité une fois uniquement en appliquant les mêmes calculs. Il est de 22 secondes pour 8065 événements. Ce temps est systématiquement court devant le temps d'acquisition.

3.4 Reconstruction par Inversion Compton

3.4.1 Principe de la méthode d'inversion Compton

L'inversion Compton que j'ai développée se base sur une formulation de la problématique de la reconstruction comme un problème inverse élémentaire. Considérons Y, une *backprojection* directe à travers laquelle on cherche à localiser une source. Je représente Y non pas sous forme d'un tableau en 2D de taille $n_1 \times n_2$ mais sous forme d'un vecteur en 1D de $N_{\text{back}} = n_1 n_2$ éléments.

Notons M, la matrice de taille $N_{\text{back}} \times N_{\text{sky}}$ qui contient, sur chacune de ses colonnes, la *backprojection* théorique associée à chaque position du ciel : elle représente ce que l'on peut appeler une fonction d'étalement du point spatialement dépendante ou PSF (*Point Spread Function*) *space variant*. Cette *backprojection* théorique est celle que l'on aurait obtenue avec une infinité de photons émis par chaque position du ciel, en supposant que l'on discrétise le ciel en N_{sky} éléments*. Ainsi, en considérant un vecteur X, de taille N_{sky}, qui contient l'intensité de sources sur le ciel, le vecteur $M.X$ représente la *backprojection* associée aux sources représentées par le vecteur X.

En pratique, je construis la matrice M à partir de simulations en utilisant la méthode que j'ai mise en place et décrite au paragraphe 2.2.2. Afin de produire des simulations sur une durée de l'ordre de cinq jours au maximum, j'ai limité le nombre de photons simulés à 35 000 parmi lesquels typiquement 10 % de ces photons donnent une paire exploitable d'interactions Compton et photoélectrique, c'est-à-dire, dans des pixels non voisins. Il est envisageable d'utiliser un plus grand nombre de photons au risque d'augmenter le temps de simulation.

Le problème que je cherche à résoudre est de trouver le vecteur X tel que :

$$Y = M.X \tag{3-9}$$

La matrice M que j'ai composée avec ces simulations n'étant pas inversible, la résolution de cette équation n'est pas triviale. Pour ce faire, je recherche une solution en minimisant la norme suivante :

$$L(X) = \|Y - M.X\|_2^2 \tag{3-10}$$

Cette minimisation doit être faite sous la contrainte que l'ensemble des composantes de X sont positives, car il ne peut pas y avoir de positions de l'espace qui « retirent » des photons. J'effectue cette minimisation par un algorithme de descente de gradient, le problème d'optimisation à traiter étant convexe.

Une fois la solution optimale X_{opt} trouvée, j'applique une régularisation dite régularisation TV (*Total Variation* ou *TV-denoising*) (Chambolle, et al., 2009), qui consiste à débruiter l'image par minimisation du gradient de X_{opt} : $\|\nabla X_{\text{opt}}\|_1$. Cette régularisation a déjà été appliquée avec succès dans le cadre

* En pratique, je prendrai $N_{\text{sky}} = N_{\text{back}}$, de sorte à discrétiser de la même manière le ciel à reconstruire et la *backprojection*.

d'autres méthodes de reconstruction Compton (Maxim, 2018). Plus précisément, la régularisation TV que j'applique cherche à minimiser la fonction suivante :

$$L_{\mathrm{TV}}(X) = \frac{1}{2}\left\|X_{\mathrm{opt}} - X\right\|_2^2 + \mu\|\nabla X\|_1 \tag{3-11}$$

μ est un paramètre de régularisation, qui donne plus ou moins de force à la régularisation. Si $\mu \ll \max(X_{\mathrm{opt}})$, la régularisation aura peu de poids et si $\mu \gg \max(X_{\mathrm{opt}})$, l'ensemble de l'image sera fortement régularisé, et cherchera à mettre toutes les composantes de ∇X à 0, ce qui revient à une image uniforme. J'ai choisi empiriquement un paramètre $\mu = 0{,}3.\max(X_{opt})$ qui est un compromis dans la régularisation.

Le minimum de la fonction L_{TV} est donné par l'algorithme fourni par (Chambolle, et al., 2009). Pour finir, je noterai qu'idéalement, il faudrait combiner l'inversion Compton et la régularisation TV en minimisant la fonction suivante :

$$L_{\mathrm{tot}}(X) = \frac{1}{2}\|Y - M.X\|_2^2 + \mu\|\nabla X\|_1 \tag{3-12}$$

La difficulté intervient dans la minimisation de cette fonction par descente de gradient car ∇L_{tot} est complexe à calculer par la présence du terme* $\|\nabla X\|_1$. L'algorithme de Chambolle donne une méthode pour calculer la minimisation de cette fonction, dans le cas où la matrice M est la matrice identité, ce qui ne correspond pas au cas que je traite, d'où l'approche en deux temps que je propose et qui correspond à celle utilisée par (Maxim, 2018).

3.4.2 Test de la reconstruction par inversion Compton avec des données réelles issues de Caliste

Pour tester mon algorithme d'inversion Compton, j'applique dans un premier temps la *backprojection* aux données telle que décrite au paragraphe 2.3.3. J'utilise une discrétisation de 2°x2° pour effectuer les *backprojections* directes que j'ai notées Y dans l'équation (3-10) et de même une discrétisation de 2°x2° pour rechercher les solutions optimales X de l'équation (3-10). Puis, je recherche par régularisation TV la solution optimale de la fonction définie par l'équation (3-11).

Les Figure 3-14 et Figure 3-15 représentent les deux résultats, inversion sans et avec régularisation TV, pour la mesure avec la source en $(10°, 0°)$ et avec la source en $(30°, 0°)$ respectivement. Le Tableau 3-5 rassemble les mesures de résolution associées sur le résultat avec régularisation TV.

L'utilisation de l'ensemble de la mesure permet bien de retrouver les sources, dans une résolution de l'ordre de 10° à 20° FWHM. Avec seulement 1000 cônes, pour la source à $(10°, 0°)$, la source est bien retrouvée avec une résolution de 11,4°. En revanche, la source à $(30°, 0°)$ est localisée dans la bonne zone, mais pas exactement à la bonne position, avec une résolution qui est pourtant assez précise, de 6,4°. Le décalage est de 3° en longitude et de 5° en latitude dans cette image par rapport à la position attendue.

* Il faut noter que ∇ n'a pas exactement le même sens dans ∇L_{tot} et ∇X. En effet, ∇X est le gradient sur l'image X, c'est-à-dire, les variations entre deux pixels voisins de l'image. ∇L_{tot} est le gradient dans l'espace vectoriel défini par l'ensemble des images X possibles.

Position de la source	(10°,0°)	(30°,0°)
FWHM toute la mesure	11,7°	18,1°
FWHM 1000 cônes	11,4°	6,4°

Tableau 3-5 : Résolution angulaire pour l'inversion Compton avec régularisation TV.

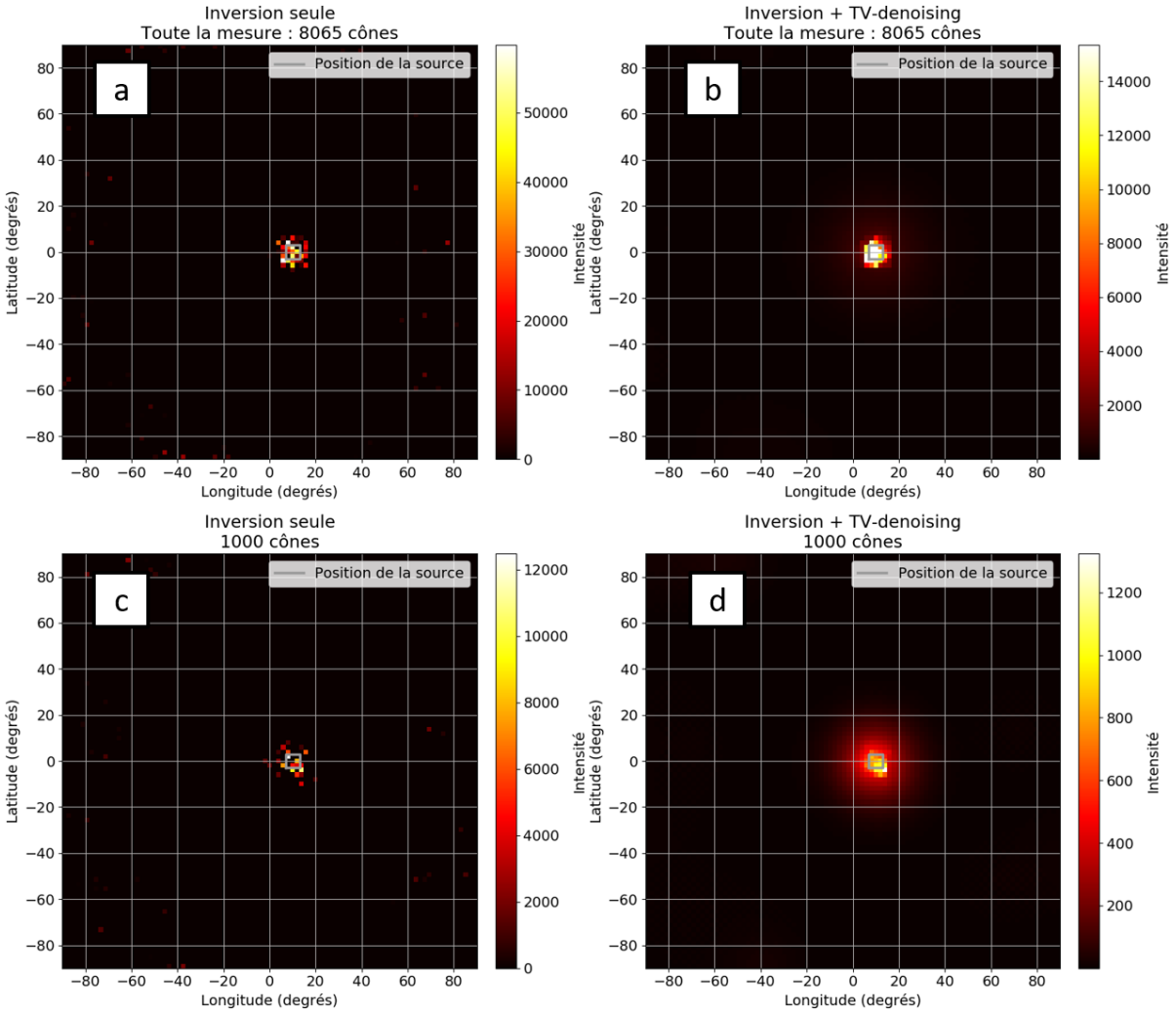


Figure 3-14 : Inversion Compton – Mesure de ^{137}Cs à $(10°, 0°)$. a : Inversion seule avec toute la mesure (8065 cônes). b : Avec régularisation TV et toute la mesure (8065 cônes). c : Inversion seule avec 1000 cônes. d : Avec régularisation TV et 1000 cônes.

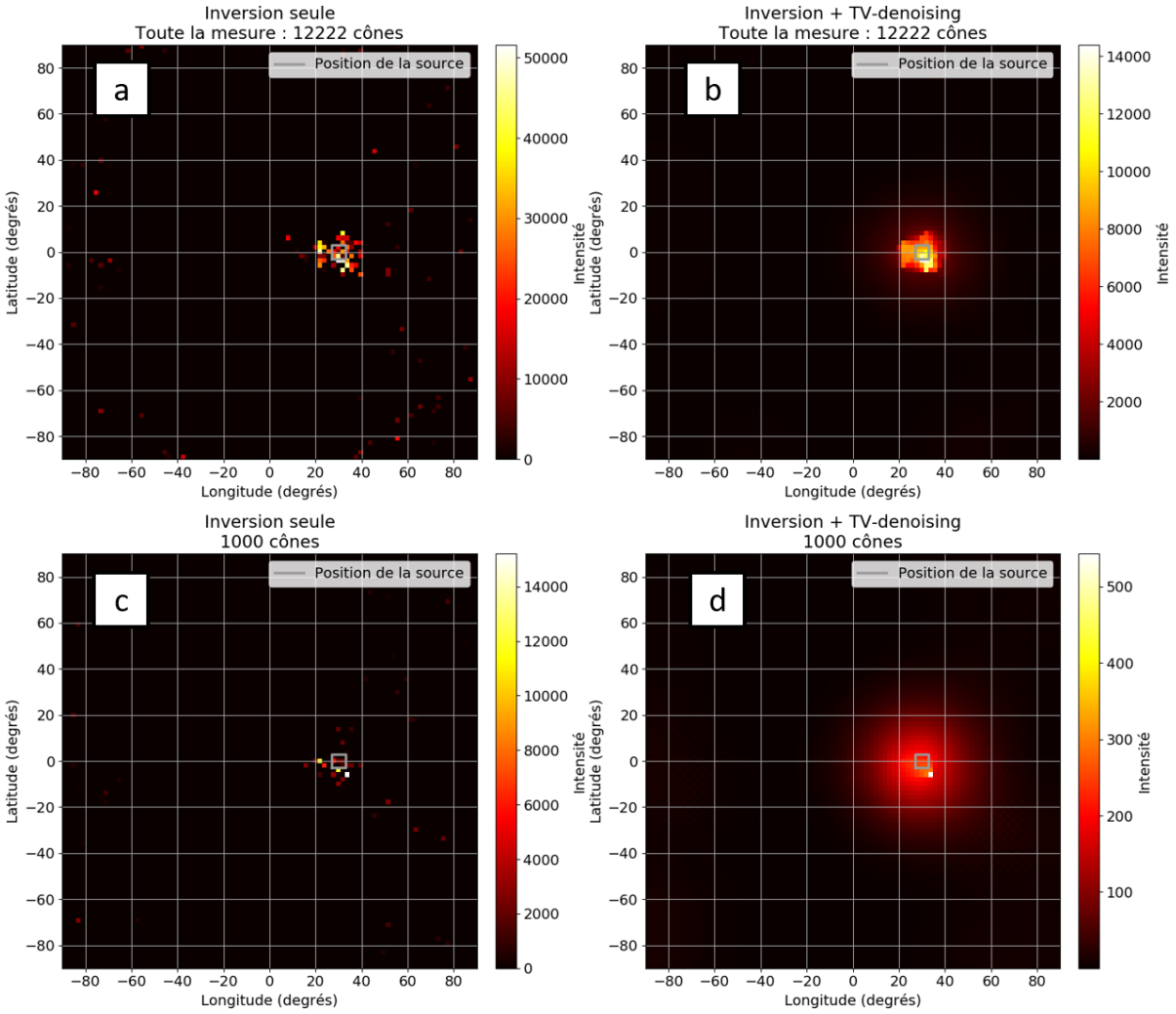


Figure 3-15 : Inversion Compton – Mesure de ^{137}Cs à $(30°, 0°)$. a : Inversion seule avec toute la mesure (12 222 cônes). b : Avec régularisation TV et toute la mesure (12 222 cônes). c : Inversion seule avec 1000 cônes. d : Avec régularisation TV et 1000 cônes.

La Figure 3-16 montre le résultat avec les deux sources en $(0°, 0°)$ et $(10°, 0°)$ en utilisant l'ensemble de la mesure et la Figure 3-17 montre la superposition des images optique et gamma.

On note que la zone dans laquelle se situent les deux sources est bien localisée, mais que les deux sources ne sont pas séparées correctement par l'algorithme, ce qui est attendu aussi vu la résolution obtenue de plus de 10° FWHM pour une source seule.

Pour le temps de calcul, cet algorithme a besoin d'une part d'effectuer la *backprojection* directe et il faut donc prendre en compte le temps de calcul associé, qui est de 22 secondes pour 8065 cônes. Ensuite, pour effectuer 10 000 itérations dans le cas de la descente de gradient, il faut 53 secondes. Pour cette étape, le temps de calcul est constant et ne dépend pas du nombre d'événements à traiter, seule la taille des images de *backprojection* est importante et elle n'est pas modifiée. Au total, pour 8065, il faut donc compter 75 secondes de calcul.

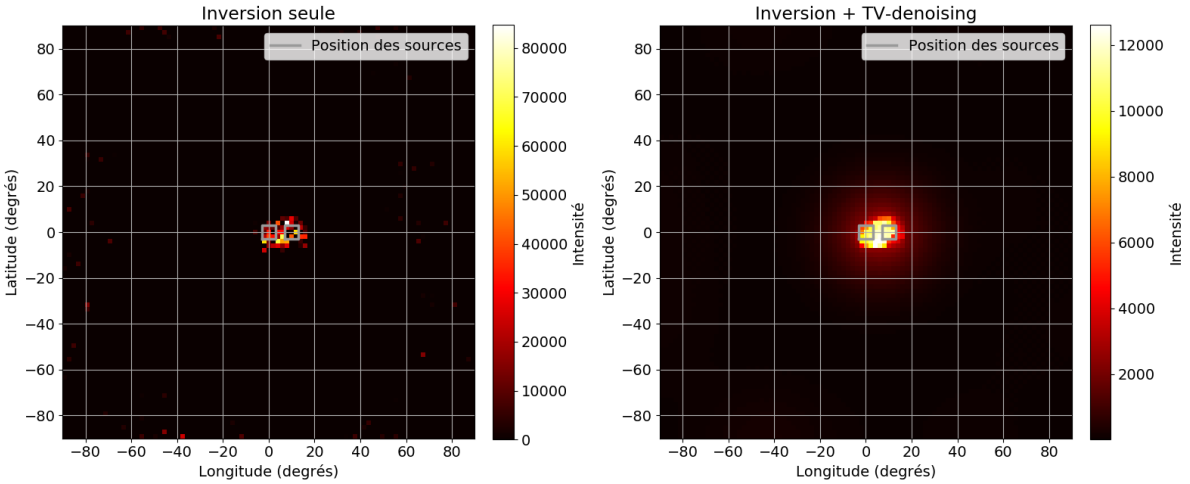


Figure 3-16 : Inversion Compton – Résultat pour les deux sources en $(0°, 0°)$ et $(10°, 0°)$. À gauche : inversion Compton seule. À droite : Inversion avec régularisation TV.

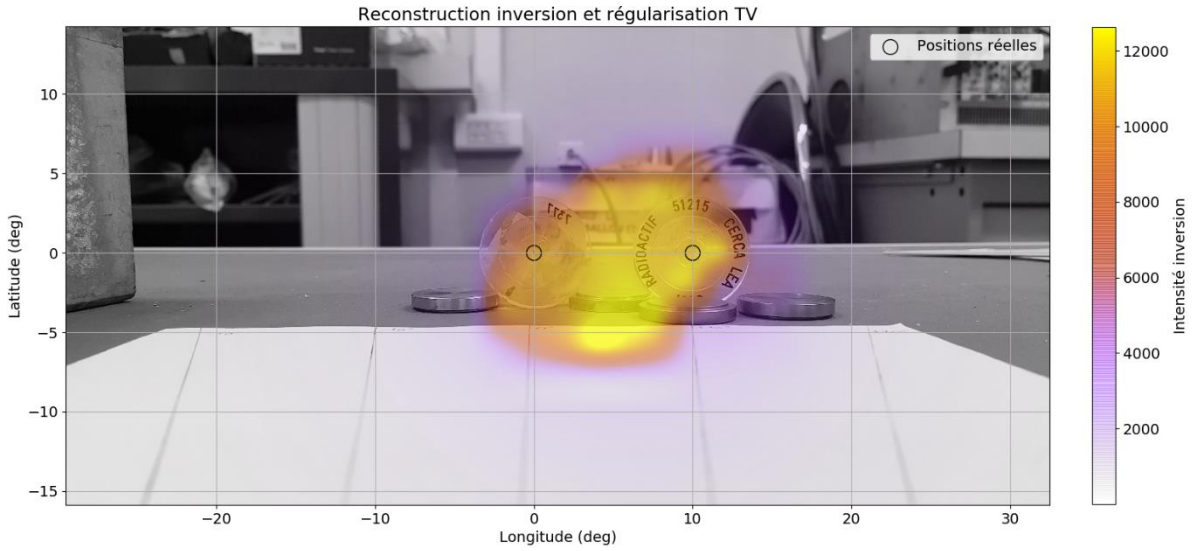


Figure 3-17 : Superposition optique de la scène avec les deux sources et du résultat de l'inversion Compton avec régularisation TV.

Une contrainte liée à cette méthode de reconstruction est qu'il faut pouvoir calculer la matrice M en amont, que j'ai obtenue à partir de simulations. Ces simulations sont faites à une énergie donnée, et dans mon cas, je ne les ai effectuées que pour le ^{137}Cs et sa ligne d'émission à 662 keV. Pour localiser d'autres radioéléments avec d'autres énergies ou avec un autre système de détection, il est nécessaire d'effectuer de nouvelles simulations à l'avance. Cependant, les simulations ne doivent être réalisées qu'une seule fois pour obtenir la matrice M et non pas à chaque mesure ce qui ne constitue pas un frein à cette méthode.

3.4.3 Test de l'inversion Compton sur une simulation de source étendue

Dans ce paragraphe, je m'intéresse aux capacités de reconstruction de l'algorithme pour des sources étendues, en le comparant aux performances des autres algorithmes que j'ai développés dans les paragraphes précédents.

Afin d'avoir des données sur des sources étendues suffisamment rapidement, j'ai simulé* une source étendue sous la forme d'un anneau ouvert, comme représenté en Figure 3-18. Pour ce faire, j'ai simulé 1 million de photons qui ont effectué une interaction Compton dans le cristal CdTe suivie d'une absorption photoélectrique, aussi dans le cristal. Parmi ce million de photons, les règles de sélections décrites au paragraphe 2.1 me donnent 81 311 cônes Compton conformes et exploitables.

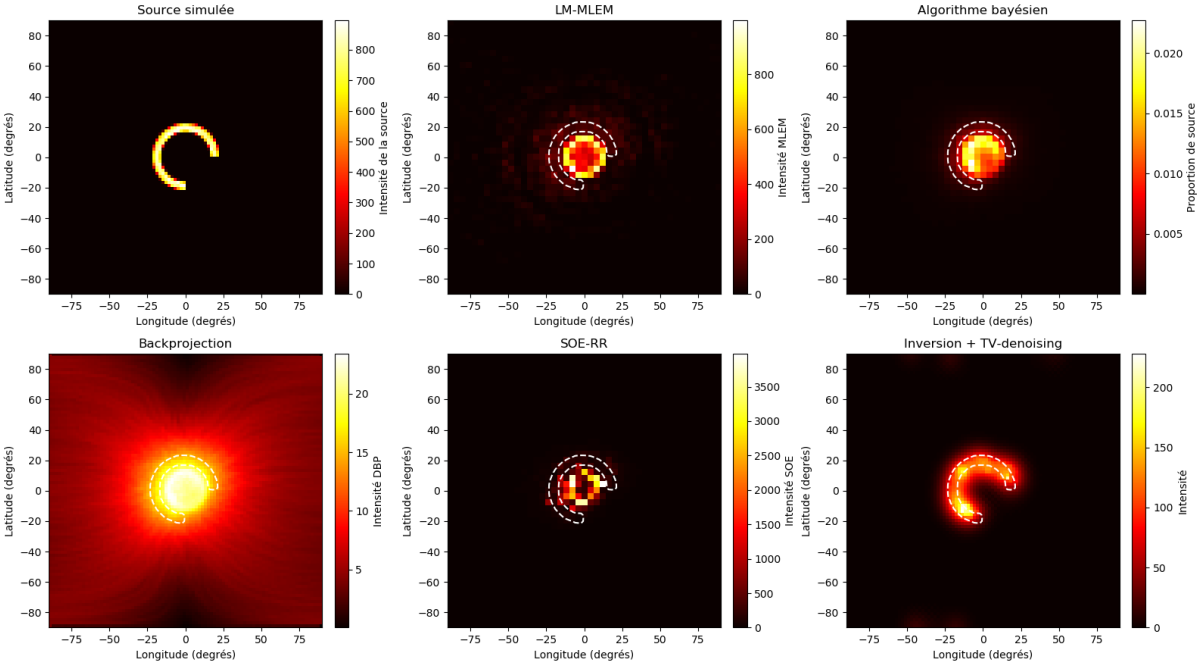


Figure 3-18 : Simulation d'un anneau ouvert avec 83 311 cônes et résultats des algorithmes sur cette simulation. Sur les résultats, le contour en traits blancs discontinus est celui de la source simulée.

J'utilise ensuite l'ensemble de ces cônes avec chaque algorithme que j'ai développé pour tenter de reconstruire la source en forme d'anneau ouvert. L'ensemble des résultats est montré en Figure 3-18.

La *backprojection* directe voit sa zone la plus intense dans une région à l'intérieur de la vraie source simulée. De la même manière, les algorithmes LM-MLEM, SOE-RR et bayésien retrouvent la source dans cette même zone trop réduite, car ces algorithmes se basent principalement sur les zones dans lesquelles se trouvent le maximum d'intensité sur les intersections de cônes.

Seul l'algorithme d'inversion retrouve la source dans une zone superposée avec la région réelle de simulation. Cela s'explique par le fait que l'inversion ne prend pas en compte seulement les intersections, mais la *backprojection* dans sa globalité afin de retrouver la position de la source.

3.5 Conclusion sur les algorithmes

La localisation de sources par imagerie Compton représente un défi du fait de l'auto-bruit intrinsèque généré par la méthode à travers les traces des cônes Compton. De ce fait, la résolution attendue est souvent très peu satisfaisante en utilisant la *backprojection* directe, de l'ordre de 35° à

* Le système de « ventilateur » que j'ai utilisé au Chapitre 4 ne permet pas d'avoir des sources étendues d'une taille angulaire plus grande que 10°, qui est l'ordre de grandeur de la résolution de l'ensemble des algorithmes que j'ai développés ; c'est pour cela que je ne l'ai pas utilisé.

plus de 70° FWHM pour notre système de détection, dépendant de la position de la source. La séparation de deux sources proches à la même énergie est très complexe.

Pour surpasser ce problème, j'ai présenté différents algorithmes, à la fois issus de l'état de l'art, LM-MLEM et SOE-RR, et de nouveaux algorithmes que j'ai développés, algorithmes de reconstruction bayésienne Compton et algorithme d'inversion, afin de traiter plus finement les données Compton acquises avec Caliste. Le Tableau 3-6 résume l'ensemble des caractéristiques des algorithmes que j'ai développés. Les résolutions angulaires FWHM sont présentées dans le cas utilisant l'ensemble des événements présents dans les mesures que j'ai effectuées.

	Backprojection **Directe (DBP)**	**MLEM**	**SOE-RR (Discrétisation)**	**Bayes (Discrétisation)**	**Inversion + TV régularisation**
FWHM à (10°,0°)	36,3°	8°	**2,6° (2°x2°)** **7,3° (4°x4°)**	5,4° (2°x2°) 8,3°(4°x4°)	11,7°
FWHM à (30°,0°)	71,7°	24°	**7° (2°x2°)** **13,6° (4°x4°)**	15,9° (2°x2°) 20,4° (4°x4°)	18,1°
Séparation de deux sources (0°,0°) et (10°,0°)	Non	Non	**Oui (via soustraction itérative de sources)**	**Oui mais imprécise**	Non
Temps de calcul (8065 cônes)	**22 s ± 1 s**	50 itérations : 1100 s ± 10 s	197 s ± 4 s	**22 s ± 1 s**	75 s ± 3 s (22 s DBP + 53 s descente de gradient)
Complexité du calcul	Linéaire : nombre d'événements	Linéaire : nombre d'événements + nombre d'itérations	Linéaire : nombre d'événements + nombre d'itérations	Linéaire : nombre d'événements	Linéaire : DBP Constant : descente de gradient
Prise en compte des incertitudes du détecteur	Non	Non	**Oui via *Recovery Resolution***	Non	**Oui via simulations**
Source étendue	Non	Non	Non	Non	**Oui**

Tableau 3-6 : Synthèse des performances des différents algorithmes présentés.

Au regard des performances de ces algorithmes, dans le cas de sources ponctuelles, l'algorithme SOE-RR présente les meilleures performances à la fois sur la résolution angulaire et sur le pouvoir de séparation de sources. Cela en fait le meilleur candidat parmi les algorithmes proposés pour effectuer la reconstruction Compton dans ces différents cas, dans les conditions d'un détecteur plan unique pour des sources de même énergie.

En revanche, si le temps de calcul est une contrainte forte, les deux algorithmes les plus rapides sont la *backprojection* directe et l'algorithme bayésien, mais en prenant en compte le reste des performances, l'algorithme bayésien est à privilégier.

Dans les deux cas, les algorithmes ont été capables de reconstruire les positions des sources avec une résolution compatible avec l'incertitude de 3° sur la mesure réelle de la position des sources, sachant que la résolution limite due à l'élargissement Doppler est de 2°.

L'algorithme d'inversion Compton a de meilleures capacités que les autres algorithmes pour traiter le cas des sources étendues, d'après mes tests. Il serait intéressant d'approfondir ces résultats, que j'ai obtenus par simulations, pour des sources étendues en les appliquant à des données réelles.

Un des principaux inconvénients de la reconstruction Compton avec Caliste-HD concerne la sensibilité du détecteur. Comme je l'ai montré au Tableau 2-1, moins de 0,9 % des événements enregistrés au total sont des événements Compton sélectionnés. Par ailleurs, l'ensemble des algorithmes a besoin d'au minimum de l'ordre de 1000 cônes pour pouvoir fournir des résultats exploitables. De ce fait, cela correspond à plus de 7 h d'acquisition pour un débit de dose de 1,4 µSv/h au niveau de la tête de détection. Une étape nécessaire sera de pouvoir augmenter la sensibilité d'un facteur 10 au moins avec les algorithmes, c'est-à-dire, pouvoir effectuer des reconstructions Compton avec une centaine de cônes. Pour des applications pratiques utilisant des mesures en temps réel, de l'ordre de quelques minutes, un travail supplémentaire sur le système de détection, en plus des algorithmes, est indispensable pour augmenter la sensibilité d'un facteur 100 au total, à travers par exemple la mesure de la position des interactions en trois dimensions dans le cristal.

4 Perspectives : vers des algorithmes de Deep Learning

Dans cette partie, j'élabore une méthode basée sur des algorithmes de *Deep Learning* pour répondre à la question de la reconstruction Compton. Aux Chapitres 3 et 5, j'ai montré leur utilisation et leurs avantages sur la problématique de la spectro-identification et de l'imagerie à masque codé.

Dans un premier temps, j'étudie l'application des algorithmes de *Deep Learning* de manière générale à la reconstruction Compton, sans limite de sensibilité. Le but, dans un second temps qui sera à poursuivre au-delà de mon travail de thèse, sera de voir si ces méthodes permettent de répondre à la nécessité d'un gain de sensibilité d'un facteur 10.

Il est à noter que l'application des méthodes de *Deep Learning* à la reconstruction Compton n'a, à ma connaissance, fait l'objet d'aucune étude publiée.

4.1 Mise en place du problème à traiter et développement de réseaux de neurones

Pour mettre en place le problème à traiter, il est nécessaire d'identifier les données d'entrée dont on dispose et les données de sortie qui doivent être obtenues. Pour l'utilisation de réseaux de neurones à couches de convolution, les données d'entrée sont contraintes à être toutes de la même dimension. C'est la raison pour laquelle j'ai choisi la *backprojection* directe comme représentation des données Compton, qui a le mérite essentiel d'avoir toujours la même dimension[*].

Afin de constituer une base d'entraînement, j'ai utilisé mon algorithme de simulation pour produire des *backprojections* directes à différentes positions de l'espace. Pour une discrétisation de 2°x2°, j'ai simulé 90x90 positions possibles dans l'espace du ciel discrétisé, ce qui fait 8100 positions possibles. Pour chaque position, j'ai créé 6 *backprojections* possibles en utilisant tous les événements Compton

[*] On pourrait penser utiliser directement la liste non traitée des événements, mais elle dépend forcément du nombre d'événements acquis, et elle est donc de taille variable.

simulés, soit de l'ordre de 2000 à 3000 événements en fonction des positions sources, mais en tirant aléatoirement les énergies déposées selon la résolution du détecteur.

Les données de sortie sont les 8100 classes correspondant aux 8100 positions du ciel possibles. Pour l'instant, je n'ai effectué des simulations que pour une source ponctuelle unique, afin de vérifier la validité de l'approche par réseaux de neurones. L'idée sous-jacente est que, si cette approche ne fonctionne pas pour des sources ponctuelles uniques avec un grand nombre d'événements Compton, il ne sera pas possible de traiter le cas de plusieurs sources ponctuelles, de sources étendues ou encore des cas avec moins d'événements Compton.

Le défaut de cette base de données d'apprentissage est le fait que chaque classe ne possède que six exemples, soit 48 600 exemples au total, car je suis limité par les capacités de mémoire vive de l'ordinateur. Chaque exemple possédant 90x90 valeurs pour les données d'entrée, la base de données représente au total environ 3 Go de stockage. Une solution intéressante serait de générer une base de données de manière continue, comme je l'ai réalisé pour l'imagerie à masque codé (cf. Chapitre 4). Cependant, les simulations pour l'imagerie à masque codé étaient très rapides à produire, de l'ordre de 0,5 secondes pour 1000 exemples, alors que dans le cadre de l'imagerie Compton, il faut de l'ordre de 5 secondes pour produire une seule *backprojection* avec 2000 cônes. Ce temps de calcul rend l'approche de génération continue de la base de données moins efficacement réalisable en pratique.

Je normalise les données de *backprojection* par leur maximum sur la *backprojection*. Pour l'architecture du réseau de neurones, comme les données d'entrée ont une dimension assez grande, de l'ordre de 90x90, je me suis intéressé aux réseaux de neurones existant dans la littérature pour traiter ce type d'image. Plus particulièrement, j'ai travaillé sur une adaptation du réseau de neurones LeNet (Lecun, et al., 1998) qui présente l'intérêt d'avoir été généralisé à différents types de problèmes de traitements d'images à un seul canal.

La Figure 4-1 présente l'architecture que j'ai mise en place.

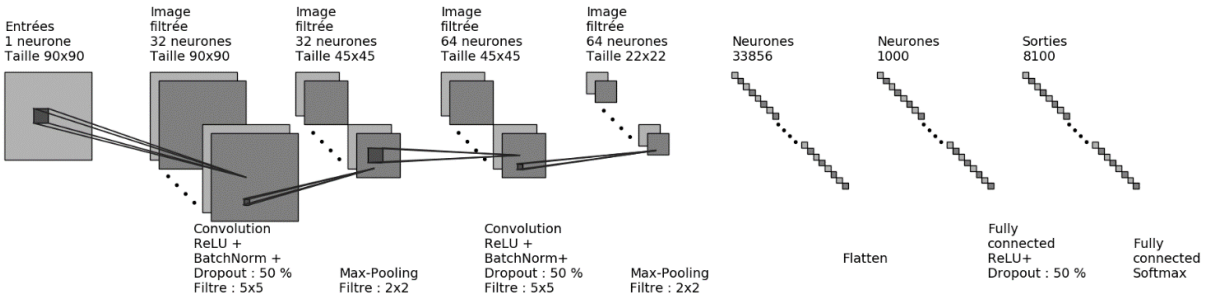


Figure 4-1 : Architecture du réseau de neurones mise en place pour traiter le problème de la reconstruction Compton.

J'ai appliqué l'idée des réseaux de neurones bayésiens, en utilisant un *dropout* sur chaque couche, avec un taux que j'ai pu mettre à 50 % sans que cela n'empêche l'apprentissage de converger.

J'utilise la fonction de coût d'entropie croisée binaire et l'algorithme d'optimisation ADAM, avec un taux d'apprentissage de 10^{-4}. Par ailleurs, vu la taille des données à traiter, j'utilise des *mini-batchs* de taille 32 de sorte à pouvoir effectuer l'apprentissage sur GPU. J'ai mis en place un jeu de validation constitué d'un exemple par classe, soit 8100 exemples au total. J'effectue l'apprentissage en

surveillant l'évolution de la fonction de coût sur le jeu de validation et en arrêtant l'apprentissage si elle ne s'améliore pas au bout de 5 itérations successives.

La Figure 4-2 représente l'évolution de l'apprentissage. Sur cette figure, on voit une première étape de l'apprentissage dans laquelle la fonction de coût diminue fortement, jusqu'à l'itération 30 environ, puis une seconde étape où elle se stabilise au-delà de l'itération 30.

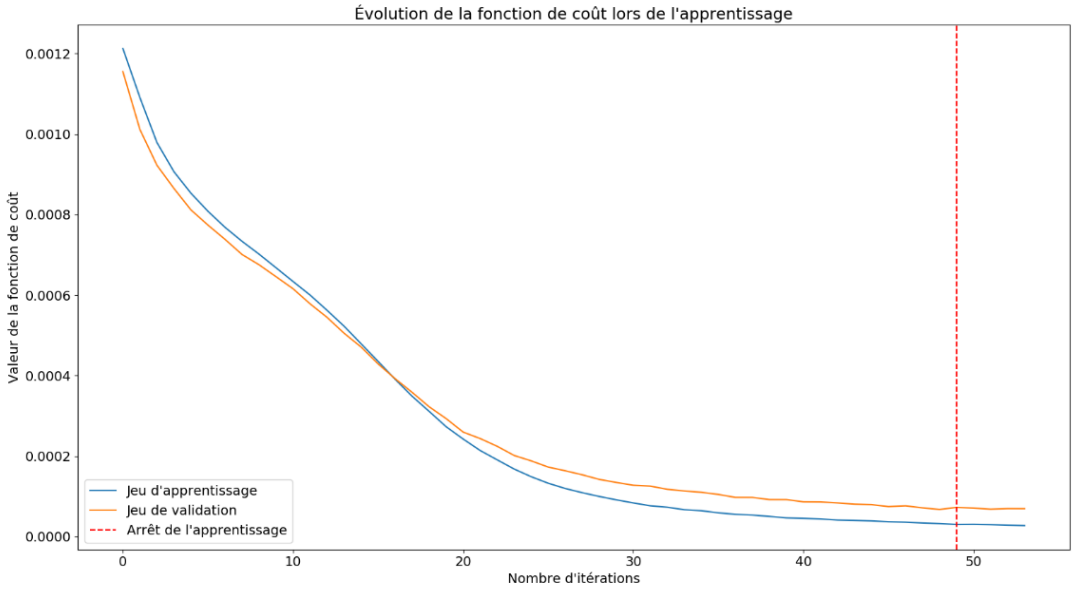


Figure 4-2 : Évolution de la fonction de coût lors de l'apprentissage du réseau de neurones pour la reconstruction Compton.

4.2 Test du réseau de neurones sur des données de simulation

Pour tester le réseau de neurones, j'ai constitué une base de tests contenant une simulation pour chacune des 8100 positions du ciel. Pour chaque position, j'ai exécuté le réseau bayésien 100 fois et j'ai utilisé la valeur médiane des résultats. Le temps d'exécution du réseau de neurones appliqué 100 fois est seulement de 3 ms.

Pour chaque test, je repère la valeur maximale renvoyée par le réseau de neurones et je calcule l'erreur angulaire entre la position correspondant à cette valeur maximale et la position attendue par la simulation. Si la prédiction par le réseau de neurones correspond à la valeur attendue, cette erreur angulaire est de 0 : la classe prédite et la classe attendue sont égales. Les résultats sont représentés en Figure 4-3.

On constate que le réseau de neurones n'a commis des erreurs que pour quatre positions sur les 8100 testées, soit un taux d'erreur de 0,05 %. Les erreurs angulaires sur chaque mauvaise position sont de 1,9°, 2,8°, 4,2° et 8°. Les erreurs à 1,9° et 2,8° sont proches de la limite de la discrétisation utilisée, de 2°x2°.

Ces performances sont à mettre en perspective avec les données utilisées pour l'entrainement et pour les tests. Du fait que j'utilise aussi des données simulées pour ces tests, la différence entre ces simulations et celles de la base d'apprentissage est le tirage aléatoire de l'énergie déposée dans le détecteur, suivant la résolution du détecteur. Dans le paragraphe suivant, j'effectue le test sur des données réelles pour évaluer les capacités de généralisation du réseau de neurones.

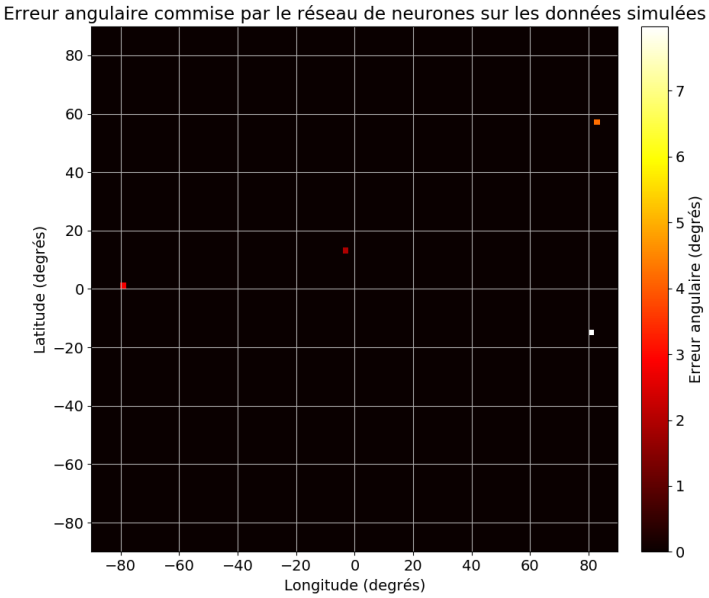


Figure 4-3 : Erreur angulaire commise par le réseau de neurones sur le jeu de test.

4.3 Résultats sur des données réelles de sources ponctuelles uniques par deep learning

Je teste ici le réseau de neurones, entraîné avec des sources simulées, sur des données réelles. Je reprends les mesures présentées en partie 2.3.3, la mesure avec la source en $(10°, 0°)$ et celle en $(30°, 0°)$.

La Figure 4-4 et la Figure 4-5 présentent les résultats des réseaux de neurones, avec la *backprojection* utilisée comme donnée d'entrée, et le résultat de la médiane des résultats, testés 10 000 fois avec la méthode des réseaux de neurones bayésiens, le temps de calcul étant de 0,3 secondes sur ces 10 000 tests.

J'ai appliqué le réseau de neurones sur la *backprojection* contenant l'ensemble de la mesure et sur celle ne contenant que 1000 cônes, pour tester le réseau de neurones face à une diminution de la statistique de comptage.

En utilisant l'ensemble de la statistique de comptage de l'acquisition, le réseau de neurones retrouve les positions des sources, avec un écart de 3° par rapport à la position attendue, ce qui est conforme avec l'incertitude due au placement des sources dans la scène (de l'ordre de $\pm$ 1,5 cm à 30 cm, soit environ $\pm$ 3°). Cela montre que le réseau de neurones est capable de renvoyer, sur des données réelles, une information de localisation correcte au regard de l'incertitude sur le placement des sources.

En revanche, en utilisant seulement 1000 cônes, les erreurs sont plus grandes, de 6° pour la source en (10°,0°) à 10° pour la source en (30°,0°). Le réseau de neurones a besoin d'un nombre de cônes plus élevé pour effectuer la reconstruction en utilisant cet apprentissage. Il sera intéressant d'adapter l'apprentissage avec une base de données contenant des exemples avec moins de cônes.

En ce qui concerne le temps de calcul, l'application des réseaux de neurones nécessite le temps de calcul de la *backprojection* soit 22 secondes pour 8065 cônes, puis le temps de calcul du réseau de neurones est de l'ordre de seulement 3 ms pour 100 exemples à 0,3 secondes pour 10 000 exemples.

Cette première étude exploratoire montre l'applicabilité des réseaux de neurones à la problématique de la reconstruction Compton avec des résultats prometteurs qui ne pouvaient pas être anticipés a priori. Elle valide l'approche de l'utilisation de la *backprojection* produite sur des simulations en tant que base d'apprentissage, puis son application sur des données réelles.

Plusieurs défis doivent encore être relevés à présent : il faut être capable d'effectuer l'apprentissage dans le cas de l'utilisation de plusieurs sources ponctuelles, d'un nombre de cônes réduits ou encore de sources étendues. Pour ce faire, il semble intéressant de pouvoir optimiser le temps de calcul des *backprojections* de sorte à pouvoir produire une base de données de manière continue dans la phase d'apprentissage, comme je l'ai fait en Chapitre 4 pour les masques codés. En fonction de l'objectif visé, l'architecture du réseau de neurones et son apprentissage pourront aussi être révisés et optimisés.

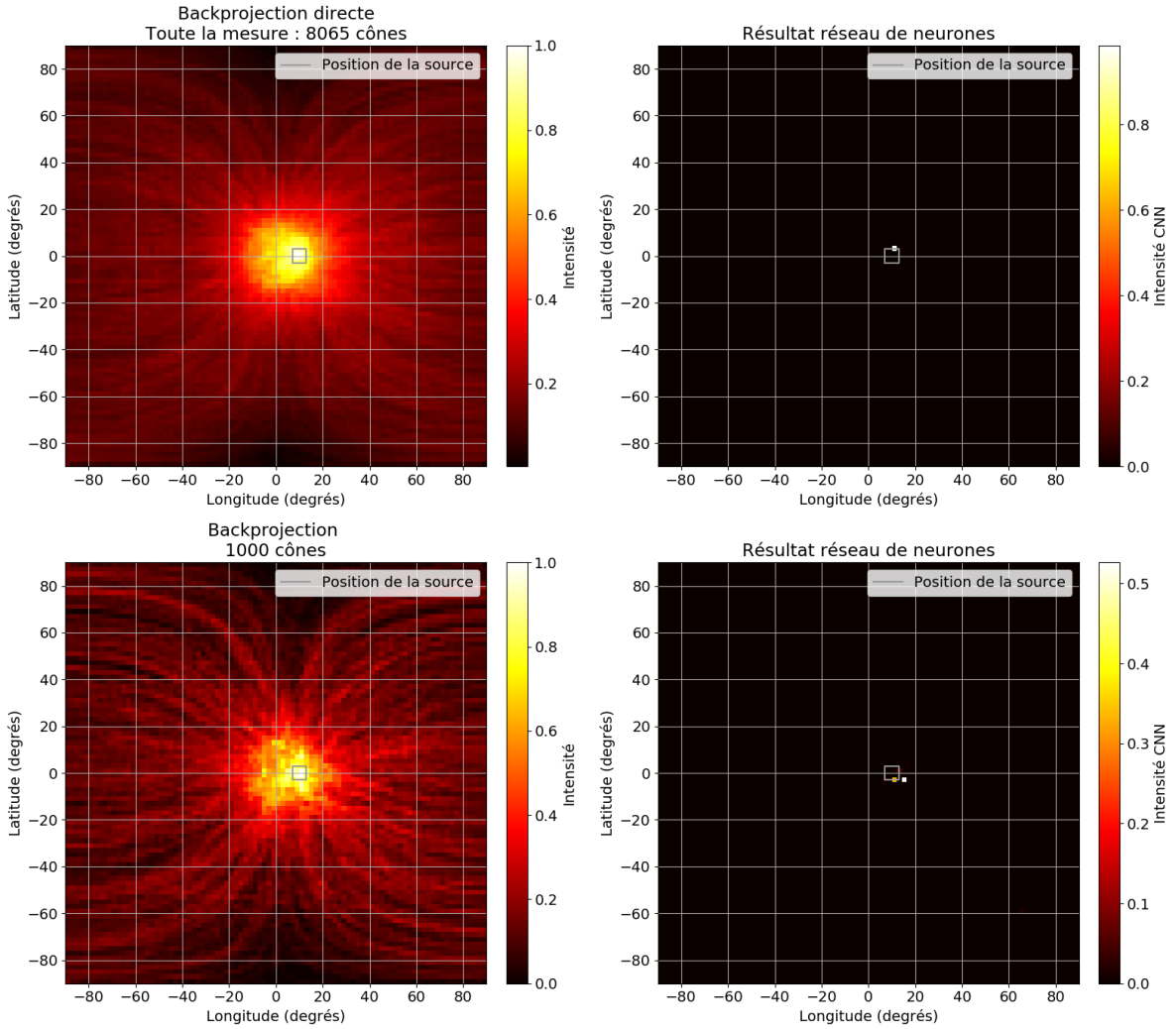


Figure 4-4 : a : *Backprojection* directe utilisée en donnée d'entrée, normalisée, source en $(10°, 0°)$ avec toute la mesure (8065 cônes). b : Résultat du réseau de neurones correspondant à toute la mesure. c : *Backprojection* directe avec 1000 cônes. d : Résultat du réseau de neurones avec 1000 cônes.

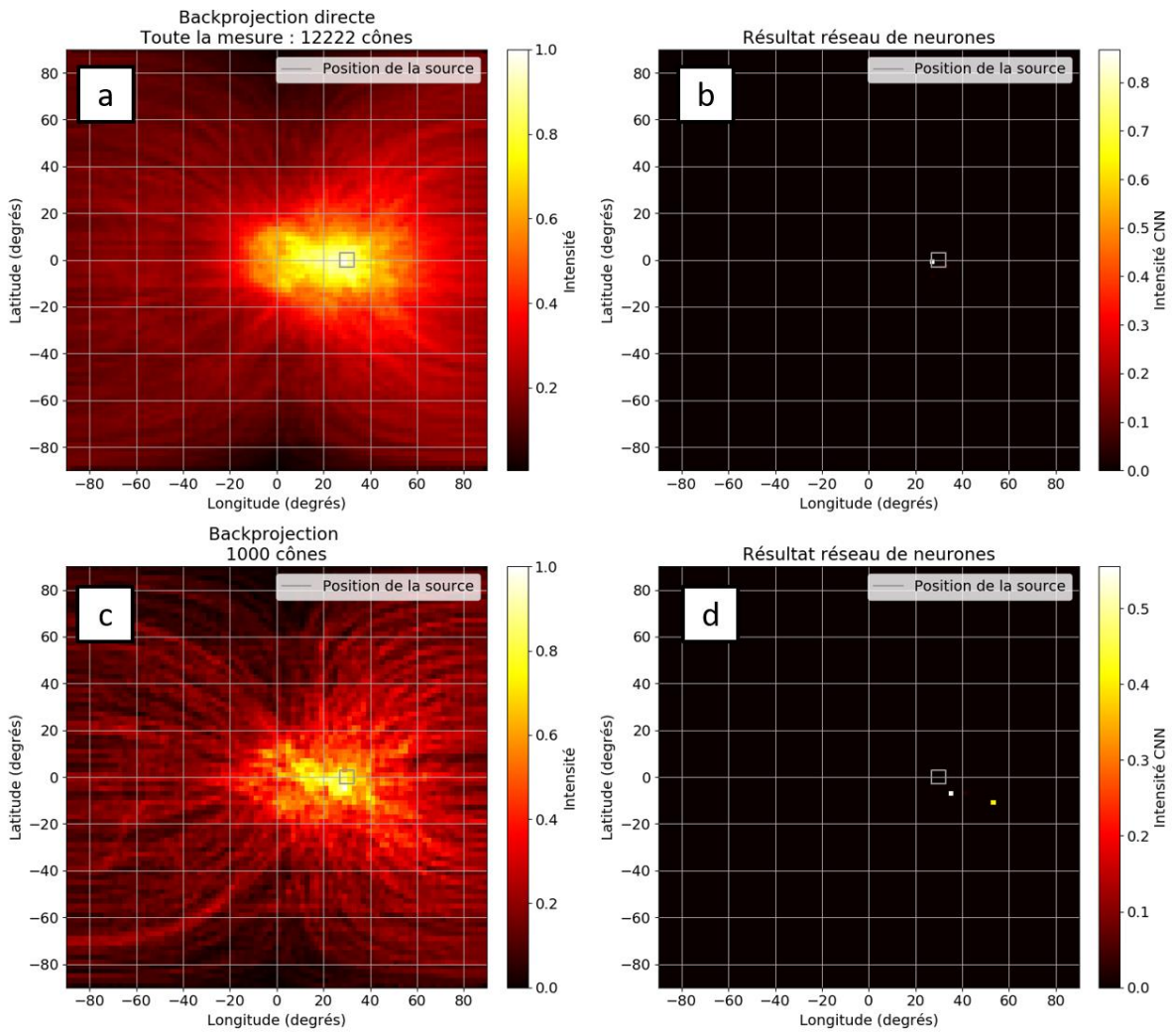


Figure 4-5 : a : *Backprojection* directe utilisée en donnée d'entrée, normalisée, source en $(30°, 0°)$ avec toute la mesure (8065 cônes). b : Résultat du réseau de neurones correspondant à toute la mesure. c : *Backprojection* directe avec 1000 cônes. d : Résultat du réseau de neurones avec 1000 cônes.

5 Conclusion

À travers ce chapitre, j'ai présenté la manière dont le processus physique de la diffusion Compton peut être utilisé pour effectuer la localisation de sources radioactives, émettant des photons de haute énergie qui interagissent avec la matière principalement par ce mode d'interaction, rendant plus délicate l'utilisation de masques codés. J'ai montré en particulier le soin qui doit être apporté à la sélection des événements Compton pour éviter d'utiliser des événements non pertinents de ce point de vue qui généreraient un bruit supplémentaire sur les images reconstruites.

J'ai notamment mis en avant la principale difficulté liée à la reconstruction Compton qui consiste en la génération d'un auto-bruit dû à la projection des cônes Compton, qui affecte les capacités de localisation, particulièrement dans le cas d'un détecteur plan unique comme Caliste. Avec d'autres géométries de détecteurs, comme par exemple avec la superposition de plusieurs couches de

détection, le bras de levier dû à l'incertitude sur la position des interactions est moins important, ce qui réduit en partie cet auto-bruit Compton. Afin de surmonter ce problème, j'ai présenté le développement de différents algorithmes et leurs applications à des données réelles issues du détecteur Caliste, en mettant en avant les avantages et les inconvénients de chaque algorithme.

En particulier, l'algorithme SOE-RR, du fait de sa prise en compte des incertitudes sur la position des interactions dans le cristal, est le mieux adapté à la reconstruction de sources ponctuelles du point de vue de la précision de localisation. L'algorithme bayésien possède de moins bonnes performances de localisation, mais il est bien meilleur en temps de calcul et peut être préférable si cela représente une contrainte forte pour l'utilisateur. Enfin, l'algorithme d'inversion Compton est à ce stade le meilleur candidat dans le cas où on souhaite localiser des sources étendues, ce que je n'ai cependant pas vérifié sur des données réelles.

Par ailleurs, je souligne que j'ai développé ces algorithmes en les testant dans le cas du détecteur Caliste-HD, mais ils peuvent être adaptés à d'autres systèmes de détection sans modification majeure[*] dans leur fonctionnement. En particulier, ils peuvent être adaptés à des instruments utilisés pour des observations astrophysiques et je présente en Annexe E une application de l'algorithme de reconstruction bayésienne à des données issues du satellite INTEGRAL.

Pour finir, j'ai montré qu'une des principales limitations de ces différentes méthodes concernait la sensibilité, c'est-à-dire, le nombre nécessaire de cônes pour pouvoir appliquer ces algorithmes et obtenir la localisation des sources. En particulier, il faut de l'ordre de 1000 cônes Compton au minimum pour appliquer les différents algorithmes, ce qui représente un inconvénient vu que les événements Compton d'intérêt représentent moins de 0,9 % des événements au total mesurés par Caliste. Cette limitation est indissociable du système de détection utilisé, un détecteur plan unique, qui limite intrinsèquement la détection des événements Compton d'intérêt.

Une voie nouvelle que j'ai commencé à étudier pour relever ce défi consiste en l'application des méthodes de *Deep Learning*. Bien que ces méthodes appliquées à la reconstruction Compton ne se retrouvent pas dans la littérature, j'ai montré une manière de les utiliser, dans un premier temps dans un cas avec des *backprojections* construites à partir de 2000 à 3000 cônes, en testant avec succès cette méthode sur des données réelles. Ces résultats encourageants montrent la faisabilité de l'utilisation des réseaux de neurones pour cette problématique ainsi que l'intérêt de poursuivre dans cette direction pour évaluer la possibilité de relever l'objectif d'un gain d'un facteur 10 en sensibilité.

6 Bibliographie

Andreyev, A., Celler, A. & Sitek, A., 2011. *Resolution recovery for Compton camera using origin ensemble algorithm.* in: IEEE Nuclear Science Symposium Conference Record.

[*] Les principales modifications seraient dans la prise en compte des incertitudes pour l'algorithme SOE-RR, qui doivent correspondre à celles du système utilisé, ainsi que pour l'inversion Compton, qui doit utiliser des simulations adaptées au système de détection.

Andreyev, A., Sitek, A. & Celler, A., 2010. *Fast image reconstruction for Compton camera using stochastic origin ensemble approach.* in: Nuclear medicine physics.

Antier, S. et al., 2015. *Hard X-ray polarimetry with Caliste, a high performance CdTe based imaging spectrometer.* in: Experimental Astronomy 39, Pages 233–258.

Biggs, F., Mendelsohn, L. & Mann, J., 1975. *Hartree-Fock Compton profiles for the elements.* in: Atomic Data and Nuclear Data Tables, Volume 16, Issue 3, Pages 201-309.

Bobin, C. et al., 2016. *Real-time radionuclide identification in γ-emitter mixtures based on spiking neural network.* in: Applied Radiation and Isotopes, Volume 109, Pages 405-409.

Chambolle, A. et al., 2009. *An introduction to Total Variation for Image Analysis.* in: hal-00437581.

Daniel, G. et al., 2020. *Compton imaging reconstruction methods: a comparative performance study of direct back-projection, SOE, a new Bayesian algorithm and a new Compton inversion method applied to real data with Caliste.* in: EPJ Web Conf., Volume 225, ANIMMA 2019.

Dubos, S., 2015. *Nouveau spectro-imageur CdTe à très haute résolution spatiale et spectrale pour l'astronomie X et gamma.* at: Université Paris 7 - Denis Diderot ; Commissariat à l'Energie Atomique et aux Energies Alternatives (CEA); Centre National d'Etudes Spatiales (CNES)..

Encyclopedia of Multimedia, 2008. *Dithering.* [Online]
Available at: https://link.springer.com/referenceworkentry/10.1007%2F978-0-387-78414-4_18

Frandes, M., Timar, B. & Lungeanu., D., 2016. *Image Reconstruction Techniques for Compton Scattering Based Imaging: An Overview [Compton Based Image Reconstruction Approaches].* in: Current Medical Imaging, Volume 12 , Issue 2.

H3D, 2020. *H400 - Gamma-Ray Imaging Spectrometer (datasheet).* [Online]
Available at: https://h3dgamma.com/H400Specs.pdf?

Haefner, A., Gunter, D., Barnowski, R. & Vetter, K., 2015. *A Filtered Back-Projection Algorithm for 4π Compton Camera Data.* in: IEEE Transactions on Nuclear Science, Volume: 62, Issue: 4, Pages : 1911 - 1917.

Hastings, W., 1970. *Monte Carlo Sampling Methods Using Markov Chains and Their Applications.* in: Biometrika, Vol. 57, No. 1, pp. 97-109.

Herman, G., 2009. *Fundamentals of computerized tomography: image.* in: Springer Science & Business Media.

Kroeger, R. et al., 2001. *Three-Compton telescope: theory, simulations, and performance.* in: IEEE Nuclear Science Symposium Conference Record.

Lecun, Y., Bottou, L., Bengio, Y. & Haffner, P., 1998. *Gradient-based learning applied to document recognition.* in: Proceedings of the IEEE 86(11):2278 - 2324.

Maxim, V., 2018. *Enhancement of Compton camera images reconstructed by inversion of a conical Radon transform.* in: Inverse Problems, Volume 35, Number 1.

Maxim, V., Frandes, M. & Prost, R., 2009. *Analytical inversion of the Compton transform using the full set of available projections.* in: Inverse Problems, Volume 25, Number 9.

Metropolis, N., Rosenbluth, A., Rosenbluth, M. & Teller, A., 1953. *Equation of State Calculations by Fast Computing Machines.* in: J. Chem. Phys. 21, 1087.

NUVIATECH Instruments, 2020. *NuVISION - Portable spectrometric gamma imaging system (specification sheet).* [Online]
Available at: https://www.nuviatech-instruments.com/wp-content/uploads/sites/3/2019/09/NVG-375022-FicheProduit-Sept2019-06-NuVISION-3.pdf

Ordonez, E., Bolozdynya, A. & Chang, W., 1997. *Doppler broadening of energy spectra in Compton cameras.* in: 1997 IEEE Nuclear Science Symposium Conference Record, Albuquerque, NM, USA, pp. 1361-1365 vol.2.

Parra, L. & Barrett, H., 1998. *List-mode likelihood: EM algorithm and image quality estimation demonstrated on 2-D PET.* in: IEEE Transactions on Medical Imaging, Volume: 17, Issue: 2.

Sitek, A., 2008. *Representation of photon limited data in emission tomography using origin ensembles.* in: Phys Med Biol.; 53(12): 3201–3216..

Takeda, S. et al., 2012. *Applications and Imaging Techniques of a Si/CdTe Compton Gamma-Ray Camera.* in: Physics Procedia, Volume 37, Pages 859-866.

Xu, D., He, Z., Lehner, C. & Zhang, F., 2004. *4-pi Compton imaging with single 3D position-sensitive CdZnTe detector.* in: Proceedings Volume 5540, Hard X-Ray and Gamma-Ray Detector Physics VI.

Conclusion et perspectives

La spectro-identification et la localisation de radioisotopes X et gamma représentent un enjeu majeur dans le contexte de la sûreté nucléaire, pour l'analyse d'environnements radiatifs, comme en situations accidentelles et post-accidentelles, ou dans des cas d'assainissement-démantèlement. Les caméras gamma ont été développées pour répondre à cet enjeu. Les détecteurs semi-conducteur CdTe Caliste représentent un choix pertinent pour la conception de ce type de système, de par leurs caractéristiques comme la miniaturisation, les performances spectrales et leurs capacités d'imagerie.

À travers mon travail de thèse, j'ai montré que le détecteur lui-même doit être associé à des algorithmes de traitement des données adaptés pour tirer profit de ses performances. J'ai étudié dans un premier temps les données issues du détecteur à la lumière des connaissances sur la physique de l'interaction rayonnement-matière. J'ai pu notamment comprendre et expliquer l'ensemble des structures visibles dans les spectres acquis par le détecteur, ainsi que la forme de ces structures à travers une modélisation fine de la réponse du système. Cette compréhension est nécessaire pour comprendre les caractéristiques spectrales qui sont les signatures des radioéléments que je cherche à reconnaître, et qui doivent être utilisées par un algorithme de spectro-identification. De plus, cette étape m'a permis de créer un algorithme de simulations de données qui génère des spectres de radioéléments sur lesquels j'ai pu me fonder pour le développement de méthodes d'analyse spectrométrique.

J'ai ensuite abordé l'axe de la spectro-identification par la mise en place du problème qui consiste en trois questions principales : identifier les radioéléments qui ont été à l'origine d'un spectre mesuré, quantifier la présence de chaque radioélément identifié et donner une incertitude pour chacun de ces résultats. Ces questions posent le problème d'un point de vue de méthodes de *Deep Learning*, en particulier en utilisant des réseaux de neurones, méthodes qui ont fait leurs preuves dans de nombreux domaines. J'ai développé une méthodologie basée sur l'apprentissage de tels algorithmes à partir de données entièrement synthétiques. Cette approche est motivée par le fait que je peux effectuer un apprentissage sur des radioéléments qui ne se limitent pas à ceux disponibles en laboratoire, et d'autre part, sur le fait que je peux introduire des modifications dans les simulations, comme des dérives de calibration.

J'ai caractérisé les performances de ces algorithmes sur des données réelles et j'ai pu montrer qu'ils atteignaient des précisions d'identification supérieures à 90 %, avec des statistiques de comptage de l'ordre de 1000 photons. Il est à noter que ces statistiques de comptage sont atteignables en une seconde d'acquisition sans que l'environnement radiatif ne soit dangereux pour un opérateur. Ensuite, j'ai montré que ces algorithmes de *Deep Learning* estiment la proportion de chaque radioélément identifié dans le signal, afin de répondre à la problématique de quantification. De plus, le temps d'exécution des calculs est d'une fraction de seconde, ce qui permet des applications en temps réel de ces méthodes, tout en renvoyant une incertitude auto-estimée par l'algorithme sur le résultat, et ce, de manière complètement automatique. La validité de ces incertitudes doit encore être étudiée. Ces développements permettent ultimement d'effectuer une mesure du débit de dose reçue au niveau du détecteur, indépendamment par radioélément présent dans la scène à analyser.

En ce qui concerne l'imagerie, je me suis intéressé dans un premier temps à l'application des méthodes d'imagerie à masque codé avec Caliste. J'ai analysé les performances des algorithmes classiquement utilisés pour traiter ce problème, algorithme de déconvolution et MLEM. J'ai notamment montré que leur sensibilité permettait de localiser des sources ponctuelles avec une vingtaine de photons seulement, et plusieurs sources ponctuelles avec une centaine de photons.

Cependant, j'ai aussi pu mettre en évidence les limites de ces algorithmes et notamment leur incapacité à reconstruire des sources étendues. Afin de surpasser ce problème, j'ai développé des algorithmes basés sur des réseaux de neurones, entraînés sur des données issues de simulations, et je les ai testés sur des données réelles. J'ai notamment mis en avant le fait qu'ils atteignaient des performances similaires à celles des algorithmes classiques sur des sources ponctuelles du point de vue de la sensibilité pour une source ponctuelle unique, tout en ayant un temps de calcul réduit d'un facteur 100 par rapport à MLEM. Les réseaux de neurones peuvent aussi renvoyer une incertitude autoévaluée sur le résultat, ce qui n'est pas permis par l'algorithme MLEM. J'ai montré qu'ils surmontent en outre l'obstacle lié aux sources étendues. J'ai pu appliquer ces algorithmes sur des données réelles afin de reconstruire une source étendue, et même dans des cas qui étaient sensiblement différents de ceux utilisés pour l'entraînement de l'algorithme.

Enfin, je me suis intéressé à la thématique de l'imagerie Compton, pour laquelle le système de détection que j'utilise avec un seul détecteur plan est défavorable, à la fois pour la sensibilité et la précision de la localisation des interactions au sein du cristal. Ces contraintes demandent déjà à porter un soin particulier sur la sélection des événements utilisés pour la reconstruction Compton, puis de développer des algorithmes de traitement adaptés. J'ai évalué les performances de plusieurs algorithmes issus de la littérature et je les ai comparées à celles de nouveaux algorithmes que j'ai développés, à partir d'acquisitions réelles.

J'ai notamment montré que l'algorithme SOE-RR a un meilleur pouvoir séparateur que les autres algorithmes, capable de localiser indépendamment deux sources séparées de 10°, notamment du fait que cet algorithme prend en compte les incertitudes connues sur le détecteur. En revanche, l'algorithme bayésien que j'ai développé obtient de meilleures performances sur le temps de calcul, d'un facteur 9, tout en ayant des performances proches de celles de SOE-RR en matière de résolution. Enfin, sur des simulations de sources étendues, j'ai montré que l'algorithme d'inversion Compton que j'ai développé était le seul à avoir la capacité d'effectuer la reconstruction correcte dans ces cas.

Il reste des progrès à effectuer sur la sensibilité. En effet, il est nécessaire de disposer de quelques milliers d'événements Compton pour effectuer une reconstruction. Or, le taux d'événements Compton exploitables étant très faible, de moins de 1 %, il est nécessaire d'attendre plusieurs dizaines d'heures d'acquisition pour un débit de dose de l'ordre de 1,4 µSv/h de ^{137}Cs au niveau de la tête de détection. De ce fait, gagner un facteur 100 en sensibilité est un objectif à atteindre. Pour ce faire, il est indispensable de travailler sur la géométrie du détecteur, de façon à la rendre plus favorable à cette méthode d'imagerie. On peut notamment penser à augmenter l'épaisseur du cristal, mais à condition aussi d'être capable de mesurer la profondeur des interactions, c'est-à-dire, avoir une information en 3D sur la position de l'interaction dans le cristal. D'autre part, concernant les algorithmes de traitement, j'ai commencé à étudier la possibilité d'appliquer les réseaux de neurones, qui ont déjà fait leurs preuves pour relever des défis dans le cas des masques codés. J'ai montré qu'ils peuvent effectuer la localisation dans le cas d'une source ponctuelle unique, en utilisant quelques milliers

d'événements Compton, à partir de la *backprojection* directe des cônes. Cette méthode est entièrement nouvelle du point de vue de l'état de l'art du domaine. Maintenant, avec un apprentissage adapté, il sera intéressant de tester leurs capacités à traiter des cas avec plusieurs sources ponctuelles et des cas avec une statistique moindre d'événements Compton.

Du point de vue de la sûreté nucléaire, l'ensemble des méthodes que j'ai mises en œuvre peuvent être intégrées dans un *pipeline* de traitement entièrement automatisé. Dans un premier temps, les algorithmes de spectro-identification identifient et quantifient la présence des radioéléments. Selon les radioéléments identifiés, le spectre acquis peut être utilisé pour éventuellement corriger automatiquement l'étalonnage du détecteur. Puis, une des deux méthodes d'imagerie[*], masque codé ou Compton, est sélectionnée selon les lignes d'émission de ces radioéléments et fournit une image gamma superposée à une image optique. Un tel système complètement automatisé serait déployable pour un grand nombre d'opérateurs non experts du domaine.

D'autres applications dans le cadre de la sûreté nucléaire peuvent être envisagées en utilisant la méthodologie présentée dans ce travail, avec quelques modification adéquates. Si le système de détection est différent, comme par exemple l'utilisation de détecteurs à scintillation, ou en modifiant la pixellisation du détecteur, l'ensemble des algorithmes de *Deep Learning* peuvent être adaptés à travers la prise en compte de ces différences dans la création de la base de données, notamment dans les simulations.

Par ailleurs, on peut penser à viser des applications plus spécifiques. Par exemple, si on souhaite être capable d'effectuer de la quantification en isotopes d'Uranium et de Plutonium, on peut utiliser des réseaux de neurones dédiés à l'étude des régions d'intérêt pour ces isotopes pour répondre à cette problématique : on ne donne que la ou les parties du spectre avec les canaux de cette région d'intérêt. On peut aussi envisager des applications dans lesquelles les radioéléments à détecter seraient dans des blindages particuliers, comme dans des fûts de déchets ou encore dans des mallettes en plomb du point de vue de la sécurité nucléaire. Dans ce cas, on peut effectuer des entraînements spécifiques des réseaux de neurones, dans lesquels la base de données est formée par des simulations où des éléments absorbeurs ou diffuseurs sont placés entre la source et le détecteur. Cela peut par exemple forcer le réseau de neurones à utiliser des raies à plus haute énergie pour discriminer les radioéléments.

Des applications peuvent aussi être considérées dans d'autres domaines, comme la médecine nucléaire par exemple. L'imagerie Compton est très utilisée avec des applications dans ce cadre pour le suivi de traceurs radioactifs. En particulier, les caméras Compton[†] présentent des avantages pour la sensibilité par rapport aux techniques de tomographie par émission de photons uniques, ou caméra SPECT (*Single Photon Emission Computed Tomography*) qui utilisent des collimateurs et ont besoin d'un mouvement mécanique autour du patient. De ce fait, les techniques de reconstruction Compton sont très pertinentes dans ce cadre. Les méthodes que j'ai mises en en place peuvent être appliquées dans ce cadre, notamment l'algorithme de reconstruction bayésienne ou l'algorithme d'inversion Compton, en adaptant les simulations aux énergies des photons des traceurs et à la géométrie du

[*] Éventuellement les deux méthodes d'imagerie sont complémentaires si plusieurs radioéléments sont détectés, ou si on souhaite obtenir une première estimation de la position avec l'imagerie Compton dans un champ de vue large, pour ensuite être plus précis avec l'imagerie à masque codé.
[†] L'algorithme SOE-RR a par exemple été développé dans cet objectif.

système de détection. De même, les méthodes de *Deep Learning* doivent être étudiées pour ces applications avec le même type d'adaptation à travers la création de la base de données.

Enfin, on peut naturellement penser à des applications en astronomie gamma, domaine qui a donné naissance à mes recherches. Je montre en Annexe E une application des méthodes de reconstruction Compton aux données issues de l'instrument IBIS du satellite INTEGRAL, notamment avec l'algorithme bayésien. Des méthodes basées sur des réseaux de neurones peuvent aussi être appliquées, en prenant en compte le fait que, dans ce cas, les énergies initiales des photons ne sont pas connues et ne sont pas issues d'une ligne unique d'un radioélément. Il sera intéressant d'étudier les capacités de ces méthodes à surpasser ces difficultés.

Concernant l'imagerie à masque codé, les sources à reconstruire en astrophysique sont principalement ponctuelles, mais la problématique du bruit est la principale composante à prendre en compte. En effet, ce bruit provient de deux sources différentes. D'une part, il existe le bruit provenant de l'environnement du détecteur, qui n'est pas forcément uniforme sur le détecteur, suivant le blindage qui est utilisé ou du fait des rayonnements induits par l'interaction des rayons cosmiques avec le satellite. D'autre part, selon le pointage du satellite, plusieurs sources X ou gamma connues peuvent être présentes dans le champ de vue et donc être sources de photons qui constituent in fine du bruit pour les sources inconnues d'intérêt que l'on cherche à localiser. Dans ce cas, les méthodes de *Deep Learning* que j'ai mises en place sont intéressantes pour résoudre cette problématique, en adaptant de nouveau l'apprentissage en prenant en compte les composantes de bruit que j'ai décrites. Par exemple, en fonction de la position des sources connues, il est envisageable d'utiliser des modèles dits adversariaux qui génèrent des images sur le détecteur en prenant en compte ces sources connues, puis qui comparent avec des cas où d'autres sources sont présentes à des positions inconnues. Plus simplement, il est aussi possible de prendre en compte différentes formes de bruits directement dans un apprentissage supervisé.

Par exemple, les algorithmes que j'ai développés peuvent être adaptés au cas de l'instrument ECLAIRs de la mission SVOM pour la détection de sursauts gamma. J'ai montré leurs capacités à reconstruire des sources ponctuelles avec la caméra WIX équipée du masque initialement prévu pour ECLAIRs. Au niveau des adaptations, il faudrait effectuer un apprentissage des réseaux de neurones en tenant compte de la nouvelle version du masque et de la structure du détecteur utilisée. La possibilité d'exécuter ces algorithmes sur des systèmes embarqués et leurs performances attendues sur la localisation de sources ponctuelles en font des méthodes de choix pour ce type d'application.

Sur la spectrométrie gamma, les méthodes de redressement Compton que j'ai montrées en Chapitre 6 sur l'étude de la dosimétrie peuvent être étudiées pour des applications en astronomie gamma. En effet, contrairement aux cas que j'ai étudiés dans mon travail de thèse, les spectres observés dans ce secteur d'application sont des spectres continus, où les pics photoélectriques n'apparaissent pas ou sont au mieux difficilement identifiables. La méthode d'inversion spectrale Compton que j'ai décrite peut s'appliquer ici pour redresser le spectre et retrouver ses composantes avant interaction dans le détecteur.

De même, les algorithmes de spectro-identification peuvent être utilisés dans le domaine de l'astronomie X, comme par exemple pour la mission ATHENA (*Advanced Telescope for High Energy Astrophysics*) dont le lancement est prévu en décembre 2031. L'instrument X-IFU de cette mission prévoit d'observer le ciel dans le domaine X entre 0,2 et 12 keV, avec une résolution de 2,5 eV FWHM

jusqu'à 7 keV, et de 0,03 % FWHM au-delà de 7 keV. Des algorithmes de spectro-identification ou de redressement de spectres trouvent leur intérêt pour ce type de mission, en fournissant rapidement des informations à partir du spectre. En particulier, il serait intéressant d'utiliser ces algorithmes pour identifier les structures spectrales connues et leur origine, mais aussi de distinguer des structures spectrales qui ne peuvent pas être directement expliquées par les modèles utilisés pour l'apprentissage des réseaux de neurones.

Pour finir, je souligne que sur de nouvelles applications, les méthodes que j'ai mises en place peuvent être prometteuses en effectuant les modifications adéquates. Cependant, il est essentiel de tester ces méthodes, dans ces nouvelles applications, avec des données réelles. D'une part, cela permet de vérifier si les données réelles ont besoin d'un prétraitement particulier, comme par exemple une sélection des événements Compton, ou une sélection en énergie sur l'imagerie à masque codé. D'autre part, ces tests permettent aussi d'évaluer la robustesse des algorithmes car les données réelles peuvent comprendre des éléments qui ne sont pas forcément pris en compte dans les simulations, comme des diffusions Compton dans l'environnement ou du bruit supplémentaire dans la détection. C'est pourquoi il est toujours indispensable de confronter les méthodes et les modèles avec les vraies mesures et l'observation.

Annexe A : Interactions photon-matière

Dans cette annexe, je décris les principaux processus d'interaction photon-matière, phénomènes qui sont au cœur de la détection des photons X et gamma. Dans un second temps, je montre comment on quantifie la probabilité d'interaction des photons dans un matériau.

1 Les interactions photon-matière dans le domaine X et gamma

Cette partie s'intéresse aux interactions photon-matière qui entraînent un dépôt d'énergie dans une dynamique en énergie allant jusque 1 MeV. De ce fait, je ne considère pas les diffusions élastiques, aussi appelées diffusion de Rayleigh, qui ne déposent pas d'énergie dans les détecteurs, ce qui les rend invisibles du point de vue de la détection. Je ne considère pas non plus les processus de créations de paires électron-positon, qui ont lieu lorsque l'énergie du photon incident est supérieure à 1,022 MeV, soit deux fois l'énergie de masse d'un électron ou d'un positon au repos qui est de 511 keV. Cette énergie est au-delà du domaine que j'étudie.

1.1 L'absorption photoélectrique

Phénomène expliqué pour la première fois par Albert Einstein, en 1905, en utilisant le concept de photon, ce qui lui a valu le prix Nobel de physique en 1921, l'effet photoélectrique ou absorption photoélectrique correspond à l'absorption du photon par un électron d'un des atomes du matériau dans lequel a lieu l'interaction.

En notant E_0 l'énergie du photon, E_L l'énergie de liaison de l'électron et E_e l'énergie cinétique de l'électron après interaction, la conservation de l'énergie donne :

$$E_e = E_0 - E_L \tag{1-1}$$

L'électron est libéré de l'atome et emporte sous forme d'énergie cinétique l'énergie que lui a transféré le photon. Sur son parcours, le photoélectron ionise d'autres atomes et cède progressivement toute son énergie au milieu dans lequel il se propage.

Dans le cas des matériaux semi-conducteurs, l'ionisation des atomes du cristal par interaction du photoélectron avec la matière conduit à la création de paires de charges électrons-trous. J'étudie la manière dont ces paires de charge induisent un signal représentatif de l'énergie E_e dans les détecteurs semi-conducteurs en Chapitre 2.

En général, on a $E_L \ll E_0$, de sorte que l'on considère que $E_e \cong E_0$. Par exemple, l'énergie de bande interdite dans le CdTe est de l'ordre 1,5 eV, alors que les énergies E_0 que l'on cherche à détecter sont supérieures à 2 keV.

L'atome ayant subi l'interaction initiale possède une lacune qui est ensuite comblée par un électron des couches supérieures par un des processus suivants :

- par transition Auger, qui est une transition non radiative correspondant à la libération d'un électron des couches externes, appelé électron d'Auger, qui sera totalement réabsorbé par le matériau par la suite ;
- par transition radiative, en émettant un photon de fluorescence X dont l'énergie E_X correspond à la différence d'énergie entre celle des électrons de la couche externe et celle de la couche contenant la lacune, à condition que l'énergie du photon initial E_0 soit suffisante*. Ce photon peut ensuite être réabsorbé par le détecteur puis créer un nouveau photoélectron d'énergie E_X qui va à son tour convertir son énergie en paires électrons-trous. En revanche, le photon de fluorescence peut aussi s'échapper du détecteur, auquel cas l'énergie enregistrée est de $E_0 - E_X$. Ce phénomène fait par exemple apparaître des raies d'échappement sur les spectres, comme on peut le voir en Chapitre 2.

La Figure 1-1 représente le phénomène d'absorption photoélectrique dans le cas d'un semi-conducteur et le phénomène d'émission électronique ou X associé.

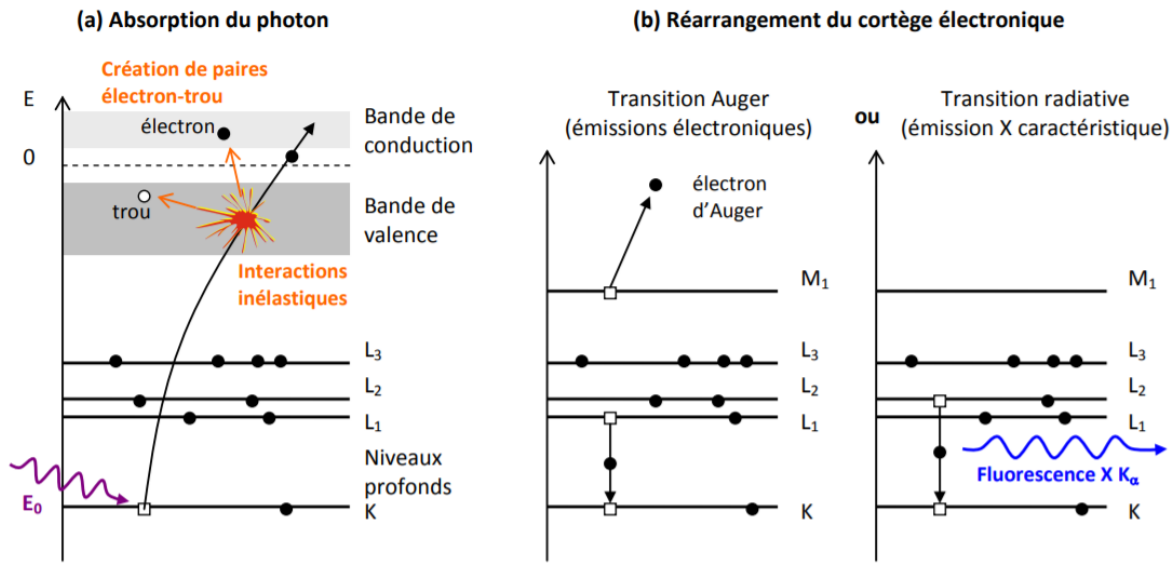


Figure 1-1 : Absorption photoélectrique dans un semi-conducteur (a) Absorption du photon par un électron d'une couche interne, le photoélectron créé ionise les atomes environnant en créant des paires électrons trous. (b) La lacune dans l'atome est comblée par transition Auger ou transition radiative (Meuris, 2009).

1.2 La diffusion Compton

Mise en évidence par Arthur Compton, qui a reçu le prix Nobel en 1927 pour cette découverte qui permet de valider la dualité onde-corpuscule de la lumière, la diffusion Compton est un phénomène d'interaction amenant à un dépôt partiel de l'énergie d'un photon sur un électron du matériau. L'électron est alors éjecté et, dans le cas d'un semi-conducteur, il se comporte comme le photoélectron dans le cas d'une absorption photoélectrique, en créant une cascade de paires de charge par ionisation de la matière sur sa trajectoire.

* Par exemple, si la lacune à combler se situe sur une couche K, il est nécessaire que E_0 ait une énergie supérieure au K-*edge* de l'atome.

La principale différence avec l'absorption photoélectrique est le fait que le photon est toujours présent après interaction, avec une énergie plus faible, on parle de photon diffusé. Par conservation de l'énergie du système, en notant E_0 l'énergie du photon initial, E_1 l'énergie du photon diffusé et E_2 l'énergie transférée à l'électron :

$$E_0 = E_1 + E_2 \tag{1-2}$$

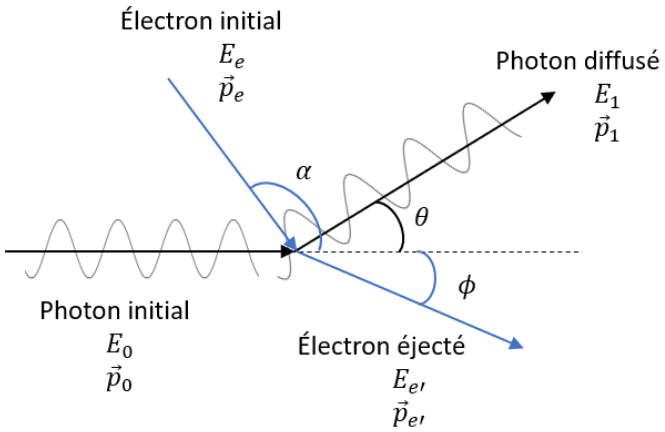


Figure 1-2 : Description de l'effet Compton, un photon initial est diffusé sur un électron d'un atome du matériau et lui dépose une partie de son énergie.

Par conservation de l'énergie et de la quantité de mouvement, l'angle de diffusion θ et l'énergie du photon diffusé E_1 sont reliés. Pour établir cette relation dans un cas général où l'électron n'est pas au repos initialement, j'utilise le formalisme de la relativité restreinte, en particulier les quadrivecteurs impulsion (Langlois, 2011).

Pour une particule d'énergie E et d'impulsion classique $\vec{p}$, en notant c la célérité de la lumière, le quadrivecteur impulsion est :

$$\underline{P} = \left\{\frac{E}{c}; \vec{p}\right\} \tag{1-3}$$

Avec ce formalisme, le produit scalaire de Minkowski pour deux quadrivecteurs impulsion $\underline{P_a}$ et $\underline{P_b}$ vaut :

$$\underline{P_a}.\underline{P_b} = \frac{E_a}{c}.\frac{E_b}{c} - \overrightarrow{p_a}.\overrightarrow{p_b} \tag{1-4}$$

Notons que pour une particule massique de masse m, la norme au carré du quadrivecteur impulsion $\underline{P}$ associé à cette particule vaut :

$$\underline{P^2} = mc^2 \tag{1-5}$$

Cela correspond à l'énergie de masse de la particule.

Pour une particule de masse nulle, soit pour un photon, cette norme est nulle :

$$\underline{P^2} = 0 \tag{1-6}$$

Cela traduit notamment la relation entre l'impulsion et l'énergie d'un photon $E = pc$.

En reprenant les notations de la Figure 1-2, la conservation des quadrivecteurs impulsion s'écrit :

$$\underline{P_{e'}} = \underline{P_0} + \underline{P_e} - \underline{P_1} \tag{1-7}$$

En mettant cette relation au carré, on obtient :

$$\underline{P_{e'}^2} = \underline{P_0^2} + \underline{P_e^2} + \underline{P_1^2} - 2\underline{P_0}.\underline{P_1} - 2\underline{P_e}.\underline{P_1} + 2\underline{P_0}.\underline{P_e} \tag{1-8}$$

Par l'équation (1-6), $\underline{P_0^2} = \underline{P_1^2} = 0$ et par l'équation (1-5), $\underline{P_e^2} = \underline{P_{e'}^2} = m_e c^2$, où m_e est la masse d'un électron. En simplifiant, l'équation (1-8) devient :

$$\underline{P_0}.\underline{P_1} + \underline{P_e}.\underline{P_1} - \underline{P_0}.\underline{P_e} = 0 \tag{1-9}$$

En appliquant le produit scalaire de Minkowski pour chacun de ces termes, on obtient, en comptant les angles α et θ comme des angles orientés :

$$\underline{P_0}.\underline{P_1} = \frac{E_0 E_1}{c^2}(1 - \cos(\theta)) \tag{1-10}$$

$$\underline{P_e}.\underline{P_1} = \frac{E_1}{c^2}(E_e + cp_e \cos(\alpha - \theta)) \tag{1-11}$$

$$\underline{P_0}.\underline{P_e} = \frac{E_0}{c^2}(E_e + cp_e \cos(\alpha)) \tag{1-12}$$

En reportant dans l'équation (1-9), on obtient :

$$E_0 E_1(1 - \cos(\theta)) + E_1(E_e + cp_e \cos(\alpha - \theta)) - E_0(E_e + cp_e \cos(\alpha)) = 0 \tag{1-13}$$

Enfin, en isolant E_1, on obtient :

$$E_1 = \frac{E_0\left(1 + \frac{cp_e}{E_e}\cos(\alpha)\right)}{\frac{E_0}{E_e}(1 - \cos(\theta)) + 1 + \frac{cp_e}{E_e}\cos(\alpha - \theta)} \tag{1-14}$$

L'équation (1-14) relie l'énergie du photon diffusé à l'énergie du photon initial et à l'angle de diffusion, mais prend aussi en compte l'énergie initiale de l'électron ainsi que son impulsion p_e et la direction α de son impulsion. Cependant, les quantités liées à l'électron initial ne peuvent généralement pas être connues.

On considère souvent le cas d'un électron initial au repos, c'est-à-dire $p_e = 0$. Dans ce cas E_e ne contient que l'énergie de masse de l'électron : $E_e = m_e c^2 = 511$ keV. L'équation (1-14) devient :

$$E_1 = \frac{E_0}{1 + \frac{E_0}{m_e c^2}(1 - \cos(\theta))} \tag{1-15}$$

Cette forme est généralement celle donnée pour l'énergie du photon diffusé par diffusion Compton. Avec la formulation (1-15) et la connaissance de l'énergie du photon initial E_0 et celle du photon diffusé E_1*, il n'existe qu'un seul angle de diffusion correspondant.

Dans le cas général, l'angle de diffusion dépend aussi des paramètres de l'électron initial, paramètres qui sont inconnus. Cela aboutit à plusieurs angles de diffusion θ possibles pour une même énergie E_2 déposée. Ce phénomène est appelé élargissement Doppler et il constitue une limitation intrinsèque aux systèmes d'imagerie Compton.

En Chapitre 2, j'étudie la distribution des angles θ et des énergies du photon diffusé E_1, donné par la section efficace doublement différentielle suivante (Ordonez, et al., 1997) qui prend en compte l'élargissement Doppler ainsi que le fait que l'électron est initialement attaché à une sous-couche n de l'atome :

$$\left(\frac{d^2\sigma_{E_0}}{d\theta dE_1}\right)_n = \frac{m_e r_0^2}{2E_0}\left(\frac{E_{\text{repos}}(\theta)}{E_0} + \frac{E_0}{E_{\text{repos}}(\theta)} - \sin^2(\theta)\right)\frac{E_1}{\sqrt{E_0^2 + E_1^2 - 2E_0E_1\cos(\theta)}}J_n(p_z)2\pi\sin(\theta) \quad (1\text{-}16)$$

où r_0 est le rayon classique de l'électron, E_{repos} est donné par l'équation (1-15), p_z est la projection de l'impulsion initiale de l'électron sur celui de l'électron après la collision et $J_n(p_z)$ est le profil Compton de la n-ième sous-couche de l'atome (Biggs, et al., 1975).

Dans le cas où on considère un électron libre au repos, pour une diffusion sur un matériau constitué d'atomes de numéro atomique Z, (Klein & Nishina, 1994) a établi une formule plus souvent utilisée du fait de sa simplicité par rapport à l'équation (1-16) :

$$\frac{d\sigma_{E_0}}{d\theta} = \pi Z r_0^2 \left(\frac{E_1}{E_0}\right)^2 \left(\frac{E_1}{E_0} + \frac{E_0}{E_1} - \sin^2(\theta)\right)\sin(\theta) \quad (1\text{-}17)$$

Je compare en Chapitre 2 les prédictions spectrales issues de l'équation (1-16) et celles issues de l'équation (1-17).

2 Probabilité d'interaction

Lorsqu'un photon pénètre dans un matériau, il a une certaine probabilité d'effectuer une interaction. Cette probabilité dépend de l'énergie du photon E et de la longueur du matériau à traverser. Elle est caractérisée par le coefficient d'absorption linéique $\mu(E)$ qui dépend de l'énergie E du photon et du matériau considéré.

Ce coefficient correspond à la probabilité d'interaction par unité de longueur. Si on considère un nombre N de photons d'énergie E qui traversent une longueur infinitésimale dx de matériau, la variation dN du nombre de photons attendu après cette longueur est :

$$dN = -\mu(E)Ndx \quad (2\text{-}1)$$

* Ou de l'énergie déposée dans le détecteur E_2 donnée par l'équation (1-2).

Cela donne une équation différentielle simple, dont la résolution fournit le nombre de photons $N(L)$ attendus après une longueur L de matériau, en considérant un nombre initial N_0 de photons entrants :

$$N(L) = N_0 e^{-\mu(E)L} \tag{2-2}$$

La probabilité d'interagir dans un matériau de longueur L est alors :

$$p(L) = 1 - e^{-\mu(E)L} \tag{2-3}$$

Par ailleurs, il existe un coefficient d'absorption linéique $\mu_i(E)$ pour chacune des interactions i* décrites en partie 1 qui caractérise la probabilité par unité de longueur d'effectuer cette interaction i. Cela se répercute dans l'équation (2-3) de la manière suivante : la probabilité qu'un photon d'énergie E effectue une interaction de type i dans le matériau de longueur L vaut :

$$p_i(L) = \frac{\mu_i(E)}{\mu(E)}(1 - e^{-\mu(E)L}) \tag{2-4}$$

Les données μ sont disponibles à partir de (NIST, 1998) ou encore à travers la bibliothèque Python *Xraylib* (Schoonjans, et al., 2011), bibliothèque que j'ai principalement utilisée au cours de ma thèse. Les données de *Xraylib* ne sont disponibles que pour des photons d'énergie inférieure à 800 keV, ce qui correspond à la majorité des cas que j'ai traités.

En réalité, la donnée qui est fournie par ces deux bases de données est le coefficient d'absorption massique μ/ρ en cm²/g où ρ représente la masse volumique du matériau considéré, de 5,85 g.cm^{-3} pour le CdTe par exemple.

En Figure 2-1, je représente les coefficients d'atténuation totaux pour le CdTe ainsi que ceux associés à l'effet photoélectrique et à la diffusion Compton. On voit sur cette figure que la diffusion Compton devient plus probable qu'une absorption photoélectrique à partir de 265 keV.

À partir des équations (2-3) et (2-4), je calcule la probabilité d'interaction totale pour chaque type d'interaction indépendamment pour une épaisseur L de CdTe traversée de 1 mm. Le résultat est représenté en Figure 2-2.

On voit par exemple que, à 662 keV, la probabilité qu'un photon effectue une interaction est de 4,1 %, dont 3,6 % de chance d'effectuer une interaction Compton et 0,5 % de chance d'effectuer une interaction photoélectrique. Cet exemple illustre bien l'intérêt de la prise en compte des structures Compton pour effectuer une spectro-identification rapide, comme décrit en Chapitre 3.

Pour finir, je note que les données d'absorption sont souvent représentées par leur section efficace σ, exprimée en cm², ou encore en barns = 10^{-24} cm². La section efficace est liée au coefficient d'atténuation linéique à travers la densité électronique (en cm^{-3}) :

$$\sigma \times N_e = \mu \tag{2-5}$$

* i correspond à l'absorption photoélectrique ou à la diffusion Compton. Plus généralement, il existe aussi les processus de créations de paires électron-positon et les diffusions de Rayleigh.

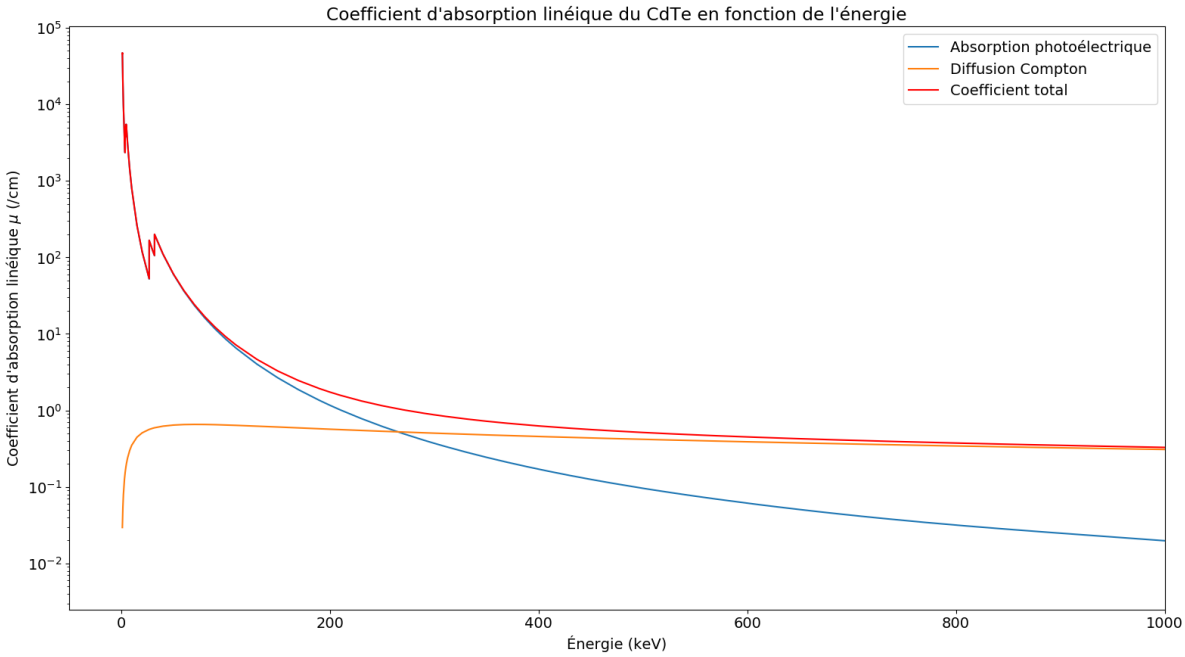


Figure 2-1 : Coefficient d'absorption linéique du CdTe en fonction de l'énergie.

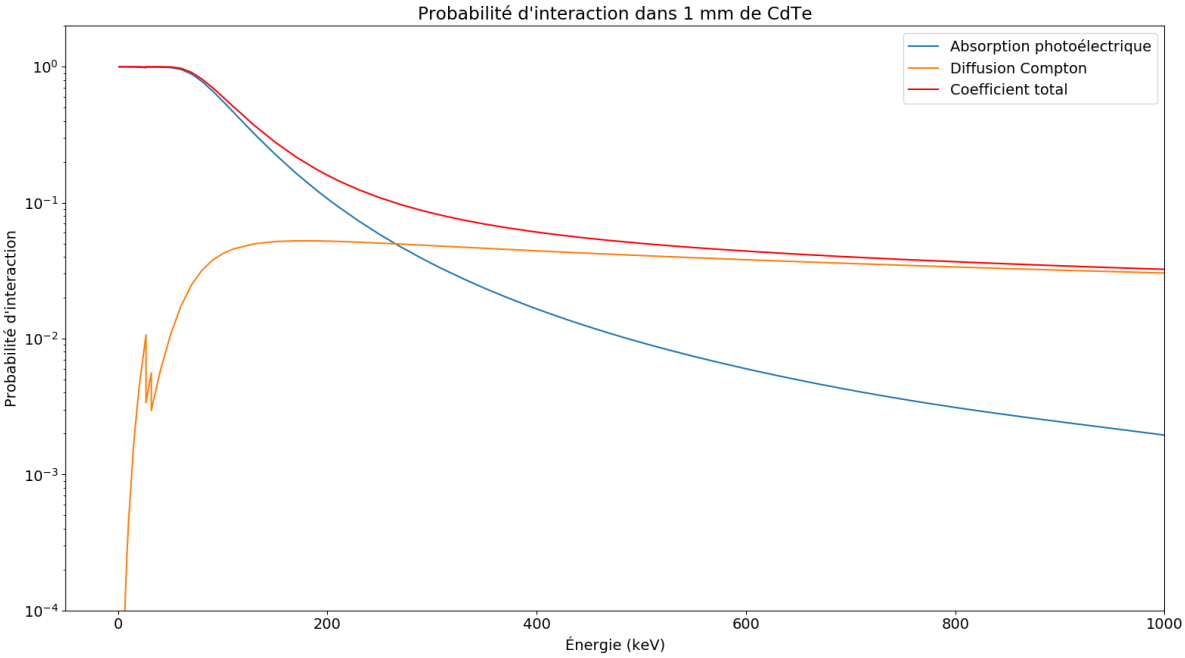


Figure 2-2 : Probabilité d'interaction des photons dans 1 mm de CdTe.

3 Bibliographie

Biggs, F., Mendelsohn, L. & Mann, J., 1975. *Hartree-Fock Compton profiles for the elements.* in: Atomic Data and Nuclear Data Tables, Volume 16, Issue 3, Pages 201-309.

Klein, O. & Nishina, Y., 1994. *On the Scattering of Radiation by Free Electrons According to Dirac's New Relativistic Quantum Dynamics.* in: The Oskar Klein Memorial Lectures (Vol 2), EKSPONG G., World Scientific Publishing Co. Pte. Ltd., pp 113-129.

Langlois, D., 2011. *Introduction à la relativité.* by: École polytechnique, Département de physique.

Meuris, A., 2009. *Etude et optimisation du plan de détection de haute énergie en Cd(Zn)Te pour la mission spatiale d'observation astronomie X et gamma SIMBOL-X.* at: Université Paris-Diderot.

NIST, 1998. *XCOM.* [En ligne]
Available at: https://physics.nist.gov/PhysRefData/Xcom/html/xcom1.html

Ordonez, E., Bolozdynya, A. & Chang, W., 1997. *Doppler broadening of energy spectra in Compton cameras.* in: 1997 IEEE Nuclear Science Symposium Conference Record, Albuquerque, NM, USA, pp. 1361-1365 vol.2.

Schoonjans, T. et al., 2011. *The xraylib library for X-ray–matter interactions. Recent developments.* in: Spectrochimica Acta Part B: Atomic Spectroscopy, Volume 66, Issues 11–12, Pages 776-784.

Annexe B : Réseaux de neurones et *backpropagation*

Cette annexe vise à montrer la manière dont est effectué l'apprentissage des réseaux de neurones à travers ce qui est appelé la *backpropagation* de l'erreur de prédiction sur les paramètres du réseau. Au Chapitre 3, j'ai présenté de manière générale la modification des paramètres d'un réseau de neurones par le calcul du gradient de cette fonction de coût. Dans cette annexe, je détaille le calcul du gradient dans le cas d'un réseau de neurones du type perceptron, en montrant l'intérêt de la structure des réseaux de neurones pour ce type calcul, qui s'effectue à partir d'opérations élémentaires.

1 Rappel du problème et notations

Je considère un réseau de neurones de type perceptron composé de N couches. Considérons une couche l du réseau de neurones, contenant n_l neurones. Pour un neurone j de la couche l, je note $W_j^{[l]}$ et $b_j^{[l]}$, le vecteur de poids et le biais associés à ce neurone.

Je note $A^{[l-1]}$ le vecteur correspondant aux sorties de la couche précédente $l-1$, vecteur de taille n_{l-1}, de sorte que $W_j^{[l]}$ soit un vecteur de taille n_{l-1}.

Je note $f^{[l]}$ la fonction d'activation $\mathbb{R} \to \mathbb{R}$ associée à la couche l. Cette fonction peut être différente pour chaque neurone j et chaque couche l. En général, elle est souvent choisie identique sur tous les neurones d'une même couche.

Le résultat $a_j^{[l]}$ du neurone j s'obtient alors à partir des deux opérations suivantes :

$$z_j^{[l]} = {W_j^{[l]}}^T . A^{[l-1]} + b \tag{1-1}$$

$$a_j^{[l]} = f^{[l]}\left(z_j^{[l]}\right) \tag{1-2}$$

Pour la couche $l = 0$, les valeurs du vecteur $A^{[0]}$ sont les valeurs des données d'entrée fournies au réseau de neurone pour l'exemple i, que je note X_i.

La dernière couche du réseau contient les valeurs de sorties $A^{[N]}$ associées à l'exemple i, que j'appelle aussi prédiction du réseau $\hat{Y}_i$. On obtient l'ensemble des valeurs $z_j^{[l]}$ et $a_j^{[l]}$ du réseau en partant de la couche d'entrée et en appliquant les équations (1-1) et (1-2) sur les couches l successives. Cette opération est appelée *forward-pass* ou propagation vers l'avant.

Ainsi, je note $\hat{Y}_i(\theta)$ la prédiction du réseau de neurones pour l'exemple i effectuée avec les données d'entrée X_i et les paramètres θ du réseau. θ regroupe l'ensemble des poids W et biais b de toutes les couches du réseau.

Afin d'évaluer l'écart entre la prédiction $\hat{Y}_i(\theta)$ et la valeur réelle attendue en sortie Y_i, j'utilise une fonction de coût que je note $L_i\left(\hat{Y}_i(\theta), Y_i\right)$. La fonction de coût totale du réseau L sur l'ensemble de la base de données considérée est la moyenne des fonctions de coût sur les exemples i :

$$L(\theta) = \frac{1}{N_{\text{exemples}}} \sum_{i \in \text{exemples}} L_i\left(\hat{Y}_i(\theta), Y_i\right) \tag{1-3}$$

La fonction de coût peut par exemple être l'entropie croisée binaire ou encore la distance euclidienne au carré entre la prédiction et la valeur attendue.

L'objectif d'un réseau de neurones est de trouver les paramètres θ_{opt} qui minimisent la fonction de coût $L(\theta)$.

$$\theta_{\text{opt}} = \underset{\theta}{\operatorname{argmin}}\, L(\theta) \tag{1-4}$$

Pour ce faire, l'algorithme principalement utilisé est l'algorithme de descente de gradient, qui consiste à partir de paramètres initiaux $\theta^{(0)}$ et à chaque étape n, de calculer le gradient de la fonction de coût avec ces paramètres $\nabla L\left(\theta^{(n)}\right)$ et de mettre à jour les paramètres suivant ce gradient, pondéré par un taux d'apprentissage λ choisi par le concepteur du réseau de neurones. La mise à jour des paramètres se fait à travers la relation suivante :

$$\theta^{(n+1)} \coloneqq \theta^{(n)} - \lambda \nabla L\left(\theta^{(n)}\right) \tag{1-5}$$

L'objectif de la suite de cette annexe est de montrer comment s'effectue le calcul du gradient $\nabla L\left(\theta^{(n)}\right)$. En particulier, cela revient à calculer chaque composante du gradient $\frac{\partial L}{\partial W_j^{[l]}}$ et $\frac{\partial L}{\partial b_j^{[l]}}$ pour chaque couche l et chaque neurone j de la couche. L'équation (1-5) est alors plus précisément :

$$W_j^{[l](n+1)} \coloneqq W_j^{[l](n)} - \lambda \frac{\partial L}{\partial W_j^{[l](n)}} \tag{1-6}$$

$$b_j^{[l](n+1)} \coloneqq b_j^{[l](n)} - \lambda \frac{\partial L}{\partial b_j^{[l](n)}} \tag{1-7}$$

2 Calcul des dérivées partielles de la fonction de coût suivant les paramètres

Pour effectuer les calculs de $\frac{\partial L}{\partial W_j^{[l]}}$ et $\frac{\partial L}{\partial b_j^{[l]}}$, je m'intéresse à un exemple particulier i de la base de données, de sorte à calculer les dérivées partielles $\frac{\partial L_i}{\partial W_j^{[l]}}$ et $\frac{\partial L_i}{\partial b_j^{[l]}}$ pour cet exemple.

Les dérivées partielles pour la fonction de coût totale L s'obtiennent à partir de l'équation (1-3), en moyennant l'ensemble des dérivées partielles calculées sur l'ensemble des exemples i :

$$\frac{\partial L}{\partial W_j^{[l]}} = \frac{1}{N_{\text{exemples}}} \sum_{i \in \text{exemples}} \frac{\partial L_i}{\partial W_j^{[l]}} \tag{2-1}$$

$$\frac{\partial L}{\partial b_j^{[l]}} = \frac{1}{N_{\text{exemples}}} \sum_{i \in \text{exemples}} \frac{\partial L_i}{\partial b_j^{[l]}} \tag{2-2}$$

2.1 Calcul sur la dernière couche du réseau de neurones

Dans un premier temps, j'effectue le calcul de ces dérivées partielles pour un neurone j de la dernière couche N du réseau. Le théorème de dérivation sur les fonctions composées donne :

$$\frac{\partial L_i}{\partial W_j^{[N]}} = \frac{\partial L_i}{\partial \hat{y}_{i\,j}} \frac{\partial \hat{y}_{i\,j}}{\partial W_j^{[N]}} \tag{2-3}$$

$$\frac{\partial L_i}{\partial b_j^{[N]}} = \frac{\partial L_i}{\partial \hat{y}_{i\,j}} \frac{\partial \hat{y}_{i\,j}}{\partial b_j^{[N]}} \tag{2-4}$$

$\hat{y}_{i\,j}$ représente la j-ième composante du vecteur de prédiction $\hat{Y}_i$, correspondant au neurone j de la couche de sortie.

Il faut maintenant calculer chaque terme du membre de droite des équations (2-3) et (2-4).

Le terme $\frac{\partial L_i}{\partial \hat{y}_{i\,j}}$ se calcule directement selon la fonction de coût utilisée. Par exemple, si on considère la distance euclidienne au carré comme fonction de coût :

$$L_i = \sum_j \left(\hat{y}_{i_j} - y_{i_j}\right)^2 \tag{2-5}$$

La dérivée partielle correspondante est :

$$\frac{\partial L_i}{\partial \hat{y}_{i\,j}} = 2\hat{y}_{i_j}\left(\hat{y}_{i_j} - y_{i_j}\right) \tag{2-6}$$

Pour calculer les autres termes $\frac{\partial \hat{y}_{i\,j}}{\partial W_j^{[N]}}$ et $\frac{\partial \hat{y}_{i\,j}}{\partial b_j^{[N]}}$, j'utilise les équations (1-1) et (1-2). En effet, d'après ces deux équations :

$$\hat{y}_{i_j} = f^{[N]}\left(z_j^{[N]} = W_j^{[N]^T} A^{[N-1]} + b_j^{[N]}\right) \tag{2-7}$$

De nouveau en utilisant les dérivées de fonction composées, en notant $f^{[N]'}$ la dérivée de $f^{[N]}$:

$$\frac{\partial \hat{y}_{i_j}}{\partial W_j^{[N]}} = \frac{\partial \hat{y}_{i_j}}{\partial z_j^{[N]}} \frac{\partial z_j^{[N]}}{\partial W_j^{[N]}} = f^{[N]'}\left(z_j^{[N]}\right) a^{[N-1]} \tag{2-8}$$

$$\frac{\partial \hat{y}_{i_j}}{\partial b_j^{[N]}} = \frac{\partial \hat{y}_{i_j}}{\partial z_j^{[N]}} \frac{\partial z_j^{[N]}}{\partial b_j^{[N]}} = f^{[N]'}\left(z_j^{[N]}\right) \tag{2-9}$$

Ainsi, les valeurs de $z_j^{[N]}$ et $a^{[N-1]}$, qui s'obtiennent à travers le calcul *forward pass* décrit en partie 1,

permettent de calculer les valeurs des dérivées partielles. On note ici l'intérêt de choisir des fonctions d'activation f « facilement » dérivables, du fait que leurs dérivées interviennent dans le calcul.

Les équations (2-8) et (2-9) montrent comment calculer les dérivées partielles pour la dernière couche et comment mettre à jour les paramètres de cette couche avec les équations (1-6) et (1-7).

2.2 Calcul des dérivées partielles pour les autres couches

Considérons maintenant une couche l quelconque du réseau de neurones et un neurone j de cette couche. En appliquant de nouveau le théorème sur la dérivée d'une fonction composée, ainsi qu'en utilisant les équations (1-1) et (1-2) :

$$\frac{\partial L_i}{\partial W_j^{[l]}} = \frac{\partial L_i}{\partial a_j^{[l]}} \frac{\partial a_j^{[l]}}{\partial W_j^{[l]}} = \frac{\partial L_i}{\partial a_j^{[l]}} f^{[l]\prime}\left(z_j^{[l]}\right) A^{[l-1]} \tag{2-10}$$

$$\frac{\partial L_i}{\partial b_j^{[l]}} = \frac{\partial L_i}{\partial a_j^{[l]}} \frac{\partial a_j^{[l]}}{\partial b_k^{[l]}} = \frac{\partial L_i}{\partial a_j^{[l]}} f^{[l]\prime}\left(z_j^{[l]}\right) \tag{2-11}$$

Comme pour la dernière couche N, les valeurs de $z_j^{[l]}$ et de $A^{[l-1]}$ sont obtenues par le calcul *forward pass* décrit au paragraphe 1.

Il ne reste qu'à calculer le terme pour l'instant inconnu $\frac{\partial L_i}{\partial a_j^{[l]}}$. Pour ce faire, j'applique de nouveau le théorème des fonctions composées :

$$\frac{\partial L_i}{\partial a_j^{[l]}} = \sum_k \frac{\partial L_i}{\partial a_k^{[l+1]}} \frac{\partial a_k^{[l+1]}}{\partial a_j^{[l]}} \tag{2-12}$$

Le terme $\frac{\partial a_k^{[l+1]}}{\partial a_j^{[l]}}$ s'obtient par les équations (1-1) et (1-2). En effet, d'après ces équations, en explicitant le produit scalaire ${W_k^{[l]}}^T . A^{[l-1]}$:

$$a_k^{[l+1]} = f^{[l+1]}\left(z_k^{[l+1]} = \sum_{j'} \left(W_k^{[l+1]}\right)_{j'} a_{j'}^{[l]} + b_k^{[l+1]}\right) \tag{2-13}$$

On obtient alors :

$$\frac{\partial a_k^{[l+1]}}{\partial a_j^{[l]}} = \left(W_k^{[l+1]}\right)_j f^{[l+1]\prime}\left(z_k^{[l+1]}\right) \tag{2-14}$$

L'équation (2-14) rend compte du deuxième terme de la somme dans l'équation (2-12).

Pour le premier terme apparaissant dans la somme de l'équation (2-12), $\frac{\partial L_i}{\partial a_k^{[l+1]}}$, celui-ci se calcule en fait de manière récursive en utilisant l'équation (2-12) elle-même. En effet, ce terme est déjà calculé pour la dernière couche N comme je l'ai montré au paragraphe 2.1 à travers $\frac{\partial L_i}{\partial \hat{y}_{ij}}$, avec un exemple dans le cas de la distance euclidienne en équation (2-6).

Ensuite, connaissant le terme $\frac{\partial L_i}{\partial a_k^{[l+1]}}$ pour la couche $l+1$, l'équation (2-12) permet récursivement de calculer le terme correspondant pour la couche l.

De ce fait, tous les termes des équations (2-10) et (2-11) sont connus.

2.3 Conclusion sur l'algorithme de *backpropagation*

L'appellation *backpropagation* prend tout son sens à travers les deux étapes que j'ai décrites : on effectue d'abord le calcul du gradient sur la dernière couche puis on « propage » ce calcul sur les autres couches du réseau à travers notamment l'équation (2-12).

On voit ici l'intérêt de la structure des réseaux de neurones qui permet d'effectuer ce calcul à l'aide d'opérations élémentaires. Si les fonctions d'activation sont dérivables « facilement », l'ensemble des calculs pour la mise à jour des paramètres réside principalement sur des sommes et des multiplications, opérations élémentaires qui sont parallélisables avec le matériel adapté, comme par exemple des processeurs GPU.

Pour finir, la description que j'ai effectuée ici concerne les réseaux de neurones de type perceptron. Le calcul est similaire pour des réseaux de neurones à couches de convolution, en traitant de manière adaptée les dérivées partielles liées aux opérations de convolution, comme décrit par (Jefkine, 2016).

3 Bibliographie

Jefkine, 2016. *Backpropagation In Convolutional Neural Networks.* [En ligne]
Available at: https://www.jefkine.com/general/2016/09/05/backpropagation-in-convolutional-neural-networks/

Annexe C : Utilisation de deux sorties complémentaires *softmax*

Cette annexe vise à justifier et discuter l'utilisation de deux sorties complémentaires *softmax* par rapport à l'utilisation d'une unique sortie sigmoïde pour les réseaux de neurones développés en Chapitre 3.

Dans le Chapitre 3 consacré à la spectro-identification de radioéléments, j'ai développé une méthode basée sur l'utilisation de réseaux de neurones spécifiques à l'identification de chaque radioélément. La sortie attendue d'un de ces réseaux de neurones est un nombre compris entre 0 et 1, pouvant être interprété comme une probabilité de présence de la source où 1 indique la présence du radioélément et 0 son absence.

Pour ce faire, en utilisant une sortie unique, la fonction sigmoïde définie par l'équation (1) répond à ce critère, étant définie sur $\mathbb{R}$, à valeurs dans $]0,1[$ par :

$$\sigma(z) = \frac{1}{1 + e^{-z}} \tag{1}$$

Cette fonction peut aussi être utilisée dans le cas où le réseau de neurones possède plusieurs sorties, qui ne sont pas mutuellement exclusives, c'est-à-dire, où plusieurs neurones de sortie peuvent présenter la valeur 1.

En revanche, dans le cas où l'on souhaite effectuer une classification exclusive, c'est-à-dire, pour laquelle seul 1 neurone peut avoir pour valeur 1, on utilise généralement la fonction *softmax* définie pour un neurone j par :

$$\mathrm{softmax}(Z)_j = \frac{e^{z_j}}{\sum_{\text{neurone } i} e^{z_i}} \tag{2}$$

où $Z = (z_i)_i$ représente le vecteur des valeurs de sortie du réseau de neurones avant application de la fonction d'activation. Par définition, les valeurs prises par cette fonction sont comprises entre 0 et 1 et la somme sur toutes les sorties du réseau vaut 1, d'où le fait que je parle de sorties complémentaires.

Dans le cas où un réseau de neurones possède deux sorties avec une fonction d'activation *softmax*, le fait de connaître la première sortie implique la connaissance directe de la deuxième sortie, du fait qu'elles sont complémentaires à 1. Ainsi, on pourrait imaginer qu'il n'est pas nécessaire d'utiliser un réseau de neurones avec deux sorties complémentaires *softmax* et que l'on peut les remplacer par une seule sortie de type sigmoïde. Par ailleurs, je précise qu'il n'est pas possible d'utiliser une fonction *softmax* avec une seule sortie, puisque par définition, la somme des sorties doit être égale à 1, ce qui forcerait à mettre la valeur 1 pour l'unique neurone utilisé.

Cependant, j'ai expérimentalement constaté de meilleures performances globales en utilisant deux sorties *softmax* plutôt qu'une seule sortie sigmoïde. En Figure 3-1, je compare les performances d'un réseau de neurones à couches de convolution, d'architecture présentée en Chapitre 3 avec une sortie sigmoïde ou deux sorties *softmax*. Les réseaux de neurones ont été entraînés sur une base de données

simulée composée des six radioéléments* et testés sur des données réelles. Les performances sont évaluées en utilisant l'*accuracy*, c'est-à-dire, en taux de réponses correctes.

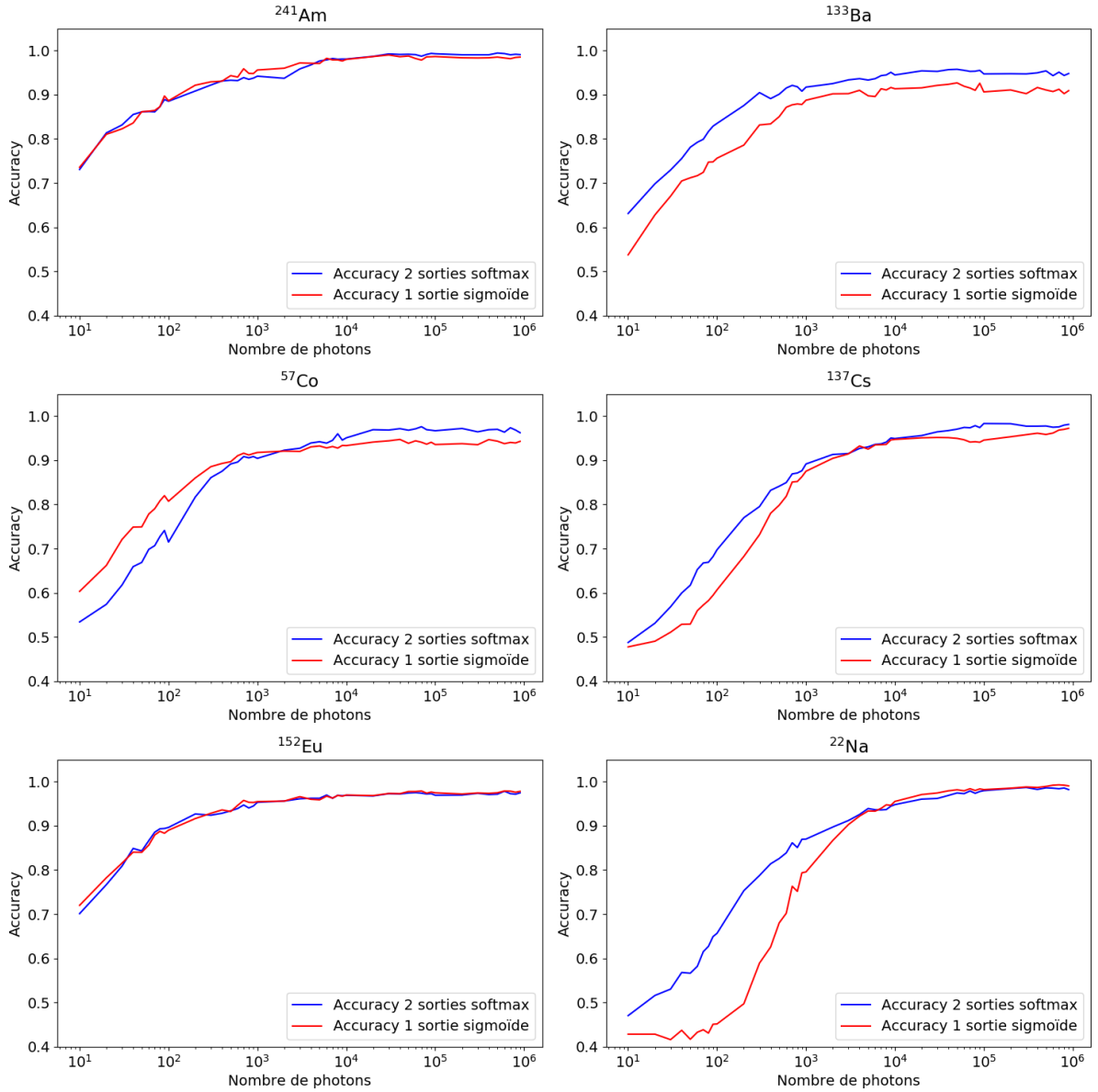


Figure 3-1 : Comparaison de l'*accuracy* pour un réseau de neurones utilisant une seule sortie ou deux sorties complémentaires. En utilisant 1000 exemples pour chaque statistique de comptage, la marge d'erreur des calculs effectués est de 3 %. Je n'ai pas représenté ces marges d'erreur pour des raisons de lisibilité du graphique.

On voit sur ce graphique que les performances des deux réseaux de neurones sont équivalentes pour le ^{241}Am et le ^{152}Eu. En revanche, le réseau de neurones avec deux sorties complémentaires a de meilleures performances globalement pour le ^{133}Ba et le ^{137}Cs, de l'ordre de 5 % supplémentaires d'*accuracy*. Pour le ^{57}Co, à basse statistique, moins de 1000 photons, le réseau à une sortie a de meilleures performances et inversement, à haute statistique, c'est le réseau à deux sorties qui a de meilleures performances. Pour le ^{22}Na, le réseau à deux sorties a de meilleures performances que

* Sur ces premiers essais, je n'avais utilisé que les 6 radioéléments dont le laboratoire disposait pour des tests réels, et non les 14 que j'ai pris en compte en Chapitre 3.

l'autre réseau à basse statistique, moins de 2000 photons, pouvant aller jusqu'à 25 % d'*accuracy*, puis les deux réseaux ont des performances équivalentes.

Cette observation a principalement motivé l'utilisation de deux sorties *softmax* complémentaires. J'ai réalisé cette étude à la suite d'un seul entraînement pour chacun des réseaux de neurones. Il est possible que ces différences soient dues à l'initialisation de chaque réseau de neurones, et qu'un autre entraînement, avec éventuellement l'utilisation d'un autre algorithme d'optimisation pour l'apprentissage, donne des résultats différents, mais il faudrait effectuer de nouveaux tests pour cela.

Par ailleurs, cette étude a été faite avec des réseaux de neurones non bayésiens, sur une bibliothèque ne contenant que six radioéléments. On pourrait envisager de reconduire cette étude comparant l'utilisation de deux sorties complémentaires à une seule sortie, dans le cas d'une bibliothèque de 14 sources avec des réseaux de neurones bayésiens pour vérifier si les résultats sont encore valides ou non.

Annexe D : L'algorithme MLEM

Cette annexe vise à compléter le Chapitre 4. Dans un premier temps, je montre que l'on peut donner une interprétation un peu différente à l'algorithme MLEM que j'ai développé dans le Chapitre 4, en partant d'un point de vue bayésien. Dans un second temps, je reprends l'algorithme tel que je l'ai décrit dans le Chapitre 4 et je donne les arguments qui prouvent sa convergence.

1 Problème et notations

Je rappelle ici le problème lié à l'imagerie à masque codé que je résous en utilisant l'algorithme MLEM et je redonne les notations que je vais utiliser tout au long de cette annexe. Dans le cas de l'imagerie à masque codé, on observe un nombre de coups Y_i dans chaque pixel i du détecteur, et on cherche à retrouver le nombre de photons X_j émis par chaque position du ciel j qui ont interagi dans le détecteur. On a mesuré au total N_{tot} coups.

On réattribue chaque photon détecté à une position du ciel, de sorte que :

$$\sum_j X_j = \sum_i Y_i = N_{\text{tot}} \tag{1-1}$$

Je note N_{ij} le nombre de photons détectés dans le pixel i qui proviennent de la position j du ciel, N_{ij} est une variable aléatoire. Dans l'ensemble de ce problème, la seule donnée qui est mesurée, et à laquelle on a donc accès, est Y_i, le nombre de photons détectés dans chaque pixel i du détecteur.

2 Approche bayésienne

Une approche fondée sur un point de vue bayésien a été introduite dans (Willingale, et al., 1984). Elle consiste à retrouver la probabilité qu'une position j du ciel émette un photon. On note cette probabilité $p(j)$. D'après la règle des probabilités totales, pour une position j du ciel :

$$p(j) = \sum_i p(j|i)p(i) \tag{2-1}$$

où $p(i)$ est la probabilité qu'un photon soit détecté dans le pixel i du détecteur. Pour une mesure avec un masque codé et un détecteur, cette probabilité vaut :

$$p(i) = \frac{Y_i}{\sum_{i'} Y_{i'}} = \frac{Y_i}{N_{\text{tot}}} \tag{2-2}$$

$p(j|i)$ est la probabilité qu'un photon détecté provienne de la position j du ciel, sachant que ce photon a été détecté dans le pixel i du détecteur.

Cette probabilité se calcule à l'aide de la formule de Bayes :

$$p(j|i) = \frac{p(i|j)p(j)}{\sum_{j'} p(i|j')p(j')} \tag{2-3}$$

Dans cette formule, $p(i|j)$ est calculable : c'est la probabilité qu'un photon soit détecté en i sachant qu'il a été émis en j. Elle correspond à la matrice de réponse du masque codé telle que je l'ai décrite en Chapitre 4.

En substituant dans l'équation (2-2), on obtient :

$$p(j) = \sum_i \frac{p(i|j)p(j)}{\sum_{j'} p(i|j')p(j')} p(i) \tag{2-4}$$

Soit donc :

$$p(j) = p(j) \sum_i \frac{p(i|j)p(i)}{\sum_{j'} p(i|j')p(j')} \tag{2-5}$$

On pourrait être tenté de simplifier par $p(j)$ mais en fait, l'astuce consiste à résoudre cette équation par la méthode du point fixe. Cette méthode correspond à la résolution de $f(x) = x$ par itération. On fixe x^0 à une certaine valeur et on calcule $x^{n+1} = f(x^n)$ qui converge vers la solution x_{sol} de l'équation précédente, sous la condition que f est lipschitzienne, de constante de Lipschitz strictement plus petite que 1. Cette condition est cependant difficile à prouver dans le cas de l'équation (2-5) du fait que toutes les $p(j)$ pour toutes les positions du ciel sont liées dans cette équation.

Dans le cas présent, on peut choisir $p^0(j) = \frac{1}{N_j}$ où N_j est le nombre de positions du ciel discrétisé. C'est un *prior* uniforme.

L'itération se fait ensuite sur l'ensemble des positions du ciel, où on calcule pour tous les j, à chaque itération n :

$$p^{n+1}(j) = p^n(j) \sum_i \frac{p(i|j)p(i)}{\sum_{j'} p(i|j')p^n(j')} \tag{2-6}$$

On calcule ensuite le nombre de photons émis à la position j du ciel X_j avec :

$$X_j = p(j)N_{\text{tot}} \tag{2-7}$$

On peut appliquer aussi l'algorithme précédent à X_j directement, en notant : $p^n(j) = \frac{X_j^n}{N_{\text{tot}}}$. Cela donne :

$$\frac{X_j^{n+1}}{N_{\text{tot}}} = \frac{X_j^n}{N_{\text{tot}}} \sum_i \frac{p(i|j)\frac{Y_i}{N_{\text{tot}}}}{\sum_{j'} p(i|j')\frac{X_{j'}^n}{N_{\text{tot}}}} \tag{2-8}$$

Soit en simplifiant par N_{tot} :

$$X_j^{n+1} = X_j^n \sum_i \frac{p(i|j)Y_i}{\sum_{j'} p(i|j')X_{j'}^n} \tag{2-9}$$

L'équation (2-9) correspond exactement à l'algorithme MLEM tel que décrit dans le Chapitre 4, je l'ai établie directement à partir de la formule de Bayes. La convergence de cette méthode reste à prouver. Cependant, elle garantit que, si la méthode donnée par l'équation (2-6) ou l'équation (2-9) converge, c'est vers un point fixe de l'équation (2-5) qui constitue alors une solution de cette équation.

3 Algorithme MLEM et convergence

L'approche MLEM classique est celle que j'ai développée dans le Chapitre 4, basée sur (Shepp & Vardi, 1982). Elle consiste à maximiser la vraisemblance des données observées sur le nombre de photons X_j émis dans chaque position du ciel j :

$$\max_{(X_j)_j} p\left((Y_i)_i | (X_j)_j\right) \tag{3-1}$$

Je note $L\left((X_j)_j\right) = p\left((Y_i)_i | (X_j)_j\right)$ la vraisemblance associée à $(X_j)_j$.

Comme je l'ai précisé en paragraphe 1, je note N_{ij} le nombre de photons émis en j qui ont été détectés dans le pixel i du détecteur. N_{ij} est une variable aléatoire qui suit une loi de Poisson, dont je note la moyenne $\lambda(i,j)$.

$\lambda(i,j)$ est relié à X_j par :

$$\lambda(i,j) = p(i|j)X_j \tag{3-2}$$

Comme en paragraphe 2, $p(i|j)$ représente la probabilité qu'un photon émis par la position j du ciel et détecté par le détecteur, soit détecté dans le pixel i du détecteur.

La probabilité associée à N_{ij} vaut :

$$p(N_{ij}|X_j) = \frac{e^{-\lambda(i,j)}\lambda(i,j)^{N_{ij}}}{N_{ij}\,!} \tag{3-3}$$

Je peux alors écrire la vraisemblance qui doit être maximisée dans l'équation (3-1), sachant que le nombre de photons détectés dans chaque pixel i provenant de chaque position du ciel j sont des variables de Poisson indépendantes :

$$L\left((X_j)_j\right) = p\left((Y_i)_i | (X_j)_j\right) = \prod_{i,j} p(N_{ij}|X_j) = \prod_{i,j} \frac{e^{-\lambda(i,j)}\lambda(i,j)^{N_{ij}}}{N_{ij}\,!} \tag{3-4}$$

3.1 Calcul de la dérivée de la vraisemblance

Pour un pixel du ciel j_0, je calcule $\frac{\partial L}{\partial X_{j_0}}$.

En substituant l'équation (3-2) dans l'équation (3-4) et en isolant les termes qui concernent X_{j_0} :

$$L\left(\left(X_j\right)_j\right) = \prod_{i,j} \frac{e^{-\lambda(i,j)}\lambda(i,j)^{N_{ij}}}{N_{ij}\,!} = \prod_{i,j\neq j_0} \frac{e^{-\lambda(i,j)}\lambda(i,j)^{N_{ij}}}{N_{ij}\,!} \times \prod_i \frac{e^{-p(i|j_0)X_{j_0}}(p(i|j_0)X_{j_0})^{N_{ij_0}}}{N_{ij_0}\,!}$$
$$= \prod_{i,j\neq j_0} \frac{e^{-\lambda(i,j)}\lambda(i,j)^{N_{ij}}}{N_{ij}\,!} \times \prod_i \frac{p(i|j_0)^{N_{ij_0}}}{N_{ij_0}\,!} \times e^{-\sum_i p(i|j_0)X_{j_0}} X_{j_0}^{\sum_i N_{ij_0}} \tag{3-5}$$

Les deux premiers produits ne dépendent pas de X_{j_0}. La dérivée donne donc :

$$\frac{\partial L}{\partial X_{j_0}} = \prod_{i,j\neq j_0} \frac{e^{-\lambda(i,j)}\lambda(i,j)^{N_{ij}}}{N_{ij}\,!} \times \prod_i \frac{p(i|j_0)^{N_{ij_0}}}{N_{ij_0}\,!} \times \left(-\sum_i p(i|j_0)\, e^{-\sum_i p(i|j_0)X_{j_0}} + e^{-\sum_i p(i|j_0)X_{j_0}} \sum_i N_{ij_0}\, X_{j_0}^{(\sum_i N_{ij_0})-1}\right) \tag{3-6}$$

Or $\sum_i p(i|j_0) = 1$, ce qui donne donc :

$$\frac{\partial L}{\partial \lambda_{j_0}} = \prod_{i,j\neq j_0} \frac{e^{-\lambda(i,j)}\lambda(i,j)^{N_{ij}}}{N_{ij}\,!} \times \prod_i \frac{p(i|j_0)^{N_{ij_0}}}{N_{ij_0}\,!} \times e^{-\sum_i p(i|j_0)X_{j_0}} \times X_{j_0}^{\sum_i N_{ij_0}} \times \left(-1 + \frac{1}{X_{j_0}} \sum_i N_{ij_0}\right) \tag{3-7}$$

Je rassemble de nouveau les termes dans un même produit :

$$\frac{\partial L}{\partial X_{j_0}} = \prod_{i,j} \frac{e^{-\lambda(i,j)}\lambda(i,j)^{N_{ij}}}{N_{ij}\,!} \left(-1 + \frac{1}{X_{j_0}} \sum_i N_{ij_0}\right) = L \times \left(-1 + \frac{1}{X_{j_0}} \sum_i N_{ij_0}\right) \tag{3-8}$$

Je note $l = \log(L)$, appelée log-vraisemblance. Par stricte croissance de la fonction logarithmique et du fait que L ne prend que des valeurs strictement positives, maximiser L revient à maximiser l. Je peux donc calculer la dérivée de l par rapport à X_{j_0} :

$$\frac{\partial l}{\partial X_{j_0}} = \frac{1}{L}\frac{\partial L}{\partial X_{j_0}} = -1 + \frac{1}{X_{j_0}} \sum_i N_{ij_0} \tag{3-9}$$

On considère l'espérance conditionnelle suivante :

$$\mathbb{E}\left(\frac{\partial l}{\partial X_{j_0}} \middle| (Y_i)_i, \left(X_j\right)_j\right) = \mathbb{E}\left(-1 + \frac{1}{X_{j_0}} \sum_i N_{ij_0} \middle| (Y_i)_i, \left(X_j\right)_j\right) \tag{3-10}$$

Dans le terme de gauche, à $\left(X_j\right)_j$ fixé, $\frac{\partial l}{\partial X_{j_0}}$ est une constante, donc :

$$\mathbb{E}\left(\frac{\partial l}{\partial X_{j_0}} \middle| (Y_i)_i, \left(X_j\right)_j\right) = \frac{\partial l}{\partial X_{j_0}} \tag{3-11}$$

En reportant dans l'équation (3-10), et en utilisant la linéarité de l'espérance :

$$\frac{\partial l}{\partial X_{j_0}} = -1 + \frac{1}{X_{j_0}} \sum_i \mathbb{E}\left(N_{ij_0} \middle| (Y_{i'})_{i'}, (X_j)_j\right) \tag{3-12}$$

Or, N_{ij_0} ne dépend pas des $Y_{i'}$ si $i' \neq i$ dans les autres pixels du détecteur, donc on peut réécrire :

$$\mathbb{E}\left(N_{ij_0} \middle| (Y_{i'})_{i'}, (X_j)_j\right) = \mathbb{E}\left(N_{ij_0} \middle| Y_i, (X_j)_j\right) \tag{3-13}$$

Y_i est le nombre de photons détectés au total dans le pixel i donc $Y_i = \sum_j N_{ij}$ et en remplaçant dans l'équation (3-13) :

$$\mathbb{E}\left(N_{ij_0} \middle| Y_i, (X_j)_j\right) = \mathbb{E}\left(N_{ij_0} \middle| \sum_j N_{ij}, (X_j)_j\right) \tag{3-14}$$

Je rappelle que les N_{ij} sont des variables aléatoires suivant une loi de Poisson de paramètre $\lambda(i,j)$. Or, d'après les propriétés sur la loi de Poisson, si A et B sont des variables indépendantes suivant une loi de Poisson de paramètres λ_A et λ_B respectivement, alors :

$$\mathbb{E}(A|A+B) = \frac{(A+B)\lambda_A}{(\lambda_A + \lambda_B)} \tag{3-15}$$

En appliquant cela avec les sommes $A = N_{ij_0}$ et $B = \sum_{j \neq j_0} N_{ij}$, on obtient :

$$\mathbb{E}\left(N_{ij_0} \middle| \sum_j N_{ij}, (X_j)_j\right) = \mathbb{E}\left(\frac{\lambda(i,j_0)\sum_j N_{ij}}{\sum_j(\lambda(i,j))} \middle| (X_j)_j\right) \tag{3-16}$$

Avec $Y_i = \sum_j N_{ij}$ et en utilisant l'équation (3-2), on obtient :

$$\mathbb{E}\left(\frac{\lambda(i,j)\sum_j N_{ij}}{\sum_j(\lambda(i,j))} \middle| (X_j)_j\right) = \mathbb{E}\left(\frac{p(i|j_0)X_{j_0}Y_i}{\sum_j p(i|j)X_j} \middle| (X_j)_j\right) \tag{3-17}$$

À $(X_j)_j$ fixé, $\frac{p(i|j_0)X_{j_0}Y_i}{\sum_j p(i|j)X_j}$ est une constante, donc on peut écrire directement :

$$\mathbb{E}\left(\frac{p(i|j_0)X_{j_0}Y_i}{\sum_j p(i|j)X_j} \middle| (X_j)_j\right) = \frac{p(i|j_0)X_{j_0}Y_i}{\sum_j p(i|j)X_j} \tag{3-18}$$

Finalement, en introduisant dans l'équation (3-12) la valeur calculée dans l'équation (3-18) :

$$\frac{\partial l}{\partial X_{j_0}} = -1 + \frac{1}{X_{j_0}} \sum_i \frac{p(i|j_0)X_{j_0}Y_i}{\sum_j p(i|j)X_j} \tag{3-19}$$

Soit en simplifiant :

$$\frac{\partial l}{\partial X_{j_0}} = -1 + \sum_i \frac{p(i|j_0)Y_i}{\sum_j p(i|j)X_j} \tag{3-20}$$

3.2 Concavité de la log-vraisemblance et existence d'un maximum

Pour maximiser la log-vraisemblance l, on doit trouver $\hat{X}_{j_0}$ tel que $\frac{\partial l}{\partial X_{j_0}}\left(\hat{X}_{j_0}\right) = 0$ et $\widehat{X_{j_0}}$ donne un maximum global de l si l est concave.

Je démontre dans un premier temps que l est concave. Pour cela, on doit montrer que la hessienne de l est définie négative. Autrement dit, quel que soit le vecteur Z de $\mathbb{R}^{N_j}$, où N_j est le nombre de positions possibles du ciel :

$$Z^T H Z \leq 0 \tag{3-21}$$

avec $H = \left(\frac{\partial^2 l}{\partial X_{j_0} \partial X_{j_1}}\right)_{j_0, j_1}$.

On calcule dans un premier temps $\frac{\partial^2 l}{\partial X_{j_0} \partial X_{j_1}}$. En dérivant $\frac{\partial l}{\partial X_{j_0}}$ par rapport à X_{j_1} et en utilisant la formule (3-20), on a :

$$\frac{\partial^2 l}{\partial X_{j_0} \partial X_{j_1}} = -\sum_i \frac{p(i|j_0) Y_i p(i|j_1)}{\left(\sum_j p(i|j) X_j\right)^2} \tag{3-22}$$

On calcule alors pour un vecteur $Z = \left(Z_j\right)_j$:

$$\begin{aligned} Z^T H Z &= \sum_{j_0} \sum_{j_1} z_{j_0} z_{j_1} \left(\frac{\partial^2 l}{\partial X_{j_0} \partial X_{j_1}}\right) = \sum_{j_0} \sum_{j_1} z_{j_0} z_{j_1} \left(-\sum_i \frac{p(i|j_0) Y_i p(i|j_1)}{\left(\sum_j p(i|j) X_j\right)^2}\right) \\ &= -\sum_i \frac{Y_i}{\left(\sum_j p(i|j) X_j\right)^2} \sum_{j_0} \sum_{j_1} z_{j_0} z_{j_1} p(i|j_0) p(i|j_1) \end{aligned} \tag{3-23}$$

La dernière ligne permute les sommes en isolant les éléments qui ne dépendent pas de j_0 et j_1.

On remarque avec le binôme de Newton que :

$$\sum_{j_0} \sum_{j_1} z_{j_0} z_{j_1} p(i|j_0) p(i|j_1) = \left(\sum_j z_j p(i|j)\right)^2 \tag{3-24}$$

En reportant dans l'équation (3-23) :

$$Z^T H Z = -\sum_i \frac{Y_i}{\left(\sum_j p(i|j) X_j\right)^2} \left(\sum_j z_j p(i|j)\right)^2 \tag{3-25}$$

Tous les termes à l'intérieur de la somme sont positifs : Y_i est un nombre de coups donc positif et les deux autres termes sont des carrés.

Donc finalement : $Z^T HZ \leq 0$ quel que soit Z, et la hessienne de l est semi-définie négative. l est alors concave, admet un maximum global et pas de maxima locaux. Cependant, l n'est pas forcément strictement concave donc ce maximum global peut être atteint sur un ensemble convexe de points, on peut penser à un plateau par exemple.

La concavité de l donne une condition suffisante pour qu'un vecteur $\hat{X}$ maximise l à partir des conditions de Kuhn-Tucker et de (Zangwill, 1969), Théorème 2.19, tel que décrit dans (Shepp & Vardi, 1982) :

Si $\hat{X} = \left(\hat{X}_j\right)_j$ vérifie les conditions suivantes :

$$\hat{X}_j \frac{\partial l}{\partial X_j}\left(\hat{X}\right) = 0 \text{ si } \hat{X}_j \neq 0 \tag{3-26}$$

$$\frac{\partial l}{\partial X_j}\left(\hat{X}\right) \leq 0 \text{ si } \hat{X}_j = 0 \tag{3-27}$$

Alors $\hat{X}$ maximise la log-vraisemblance l et donc la vraisemblance L.

3.3 Algorithme de recherche du maximum de vraisemblance

(Shepp & Vardi, 1982) propose un algorithme EM (*Expectation-Maximization* ou Espérance-Maximisation), au sens de (Dempster, et al., 1977), qui va converger vers une solution vérifiant ces conditions. C'est cet algorithme que j'ai donné dans le Chapitre 4.

L'étape Espérance n est donnée de la manière suivante :

Si on suppose que $\hat{X}^{n-1}$ est la valeur de $\hat{X}$, on note X_j la variable aléatoire représentant le nombre de photons effectivement émis par la position j du ciel. L'espérance sur X_j est alors :

$$\mathbb{E}\left(X_j | \hat{X}^{n-1}, (Y_i)_i\right) = \mathbb{E}\left(\sum_i N_{ij} \,|\hat{X}^{n-1}, (Y_i)_i\right) = \sum_i \mathbb{E}\left(N_{ij} | \hat{X}^{n-1}, (Y_i)_i\right) \tag{3-28}$$

L'étape de Maximisation est directe puisque, par construction, $\hat{X}_j^n$ représente le nombre moyen de photons émis à la position j du ciel. En identifiant $\hat{X}_j^n$ à l'équation (3-28) et en se reportant aux équations (3-14) et (3-18), on obtient le schéma itératif suivant :

$$\hat{X}_j^n = f\left(\hat{X}_j^{\,n-1}\right) = \hat{X}_j^{\,n-1} \sum_i \frac{p(i|j) Y_i}{\sum_{j'} p(i|j') \hat{X}_{j'}^{\,n-1}} \tag{3-29}$$

Où f est la fonction décrivant le schéma itératif qui passe de $\hat{X}_j^{n-1}$ à $\hat{X}_j^n$.

(Dempster, et al., 1977) a démontré dans son théorème 1 qu'en suivant ce schéma itératif pour un algorithme EM, on a effectivement :

$$l\left(\hat{X}^{n-1}\right) \leq l\left(\hat{X}^n\right) \tag{3-30}$$

Avec égalité seulement si $\hat{X}^{n-1} = \hat{X}^n$.

L'algorithme EM présenté renvoie donc des valeurs dont la vraisemblance augmente à chaque itération. De ce fait, la suite $l(\hat{X}^n)$ est croissante et de plus, elle est majorée par 1, par définition de la vraisemblance, qui est une probabilité, donc la suite $l(\hat{X}^n)$ converge vers une valeur que je note l^∞.

Par ailleurs, on remarquera que, pour $n > 1$, $\|\hat{X}^n\|_1 = N_{\text{tot}}$ où N_{tot} est le nombre total de photons enregistrés. En effet, pour $n > 1$, on a :

$$\begin{aligned}\|\hat{X}^n\|_1 &= \sum_j \hat{X}_j^n = \sum_j \hat{X}_j^{n-1} \sum_i \frac{p(i|j)Y_i}{\sum_{j'} p(i|j')\hat{X}_{j'}^{n-1}} = \sum_i \sum_j \frac{\hat{X}_j^{n-1} p(i|j)Y_i}{\sum_{j'} p(i|j')\hat{X}_{j'}^{n-1}} \\ &= \sum_i \frac{Y_i}{\sum_{j'} p(i|j')\hat{X}_{j'}^{n-1}} \sum_j \hat{X}_j^{n-1} p(i|j) = \sum_i Y_i = N_{\text{tot}}\end{aligned} \tag{3-31}$$

De ce fait, la suite $\hat{X}^n$ est bornée, et évolue dans l'espace compact constitué par les éléments dont la norme 1 vaut N_{tot}. Par le théorème de Bolzano-Weierstrass, je peux donc extraire une suite $\hat{X}^{\varphi(n)}$ de $\hat{X}^n$ qui converge, où φ est une fonction strictement croissante de $\mathbb{N}$ dans $\mathbb{N}$. Je note $\hat{X}^\infty$ la limite de la suite $\hat{X}^{\varphi(n)}$.

Par ailleurs, il a été démontré par (Salvo & Defrise, 2019) que $\hat{X}^\infty$ est en réalité la limite globale de la suite $\hat{X}^n$. L'argument repose sur deux points. D'une part, il est montré dans (Salvo & Defrise, 2019) que la divergence généralisée de Kullback-Leibler d^n entre $\hat{X}^\infty$ et $\hat{X}^n$ est décroissante et minorée par 0 et elle est donc convergente. D'autre part, comme $\hat{X}^{\varphi(n)}$ tend vers $\hat{X}^\infty$, alors $d^{\varphi(n)}$ tend vers 0, ce qui implique par conséquent que la limite globale de d^n est aussi 0, ce qui montre que $\hat{X}^n$ converge vers $\hat{X}^\infty$.

En ce point $\hat{X}^\infty$, on obtient de fait, par continuité de l :

$$l(\hat{X}^n) \xrightarrow[n\to+\infty]{} l(\hat{X}^\infty) \tag{3-32}$$

Et donc, par identification :

$$l^\infty = l(\hat{X}^\infty) \tag{3-33}$$

Les équations (3-30) et (3-32) impliquent aussi que :

$$l\left(f(\hat{X}^\infty)\right) = l(\hat{X}^\infty) \tag{3-34}$$

Cela implique donc, par la condition d'égalité de l'équation (3-30) :

$$\hat{X}^\infty = f(\hat{X}^\infty) \tag{3-35}$$

3.4 Démonstration que le point de convergence maximise la vraisemblance

Il faut maintenant montrer qu'en ce point $\hat{X}^\infty$, les conditions de Kuhn-Tucker sont vérifiées. Cela prouve que $\hat{X}^\infty$ donne bien le maximum global de la vraisemblance l, ce qui montrera que l^∞ est le maximum de log-vraisemblance.

D'une part, on doit montrer la condition (3-26).

On peut calculer, en utilisant les équations (3-20), (3-29) et (3-32) :

$$\hat{X}_j^{\infty} \frac{\partial l}{\partial X_j}(\hat{X}^{\infty}) = \hat{X}_j^{\infty}\left(-1 + \sum_i \frac{p(i|j)Y_i}{\sum_{j'} p(i|j')\hat{X}_{j'}^{\infty}}\right) = -\hat{X}_j^{\infty} + \hat{X}_j^{\infty}\sum_i \frac{p(i|j)Y_i}{\sum_{j'} p(i|j')\hat{X}_{j'}^{\infty}} = -\hat{X}_j^{\infty} + f(\hat{X}_j^{\infty}) = 0 \tag{3-36}$$

Cela vérifie la condition (3-26).

D'autre part, on doit montrer que la condition (3-27) est aussi vérifiée.

Pour cela, on raisonne par l'absurde : s'il existe j tel que $\hat{X}_j^{\infty} = 0$ et que $\frac{\partial l}{\partial X_j}(\hat{X}^{\infty}) > 0$, alors comme $\hat{X}^{\infty}$ est la limite de la suite $\hat{X}^n$ et que $\frac{\partial l}{\partial X_j}$ est une fonction continue, pour tout $\varepsilon > 0$, il existe n_0 tel que pour tout $n > n_0$:

$$\left|\frac{\partial l}{\partial X_j}(\hat{X}^n) - \frac{\partial l}{\partial X_j}(\hat{X}^{\infty})\right| < \varepsilon \tag{3-37}$$

Or, on peut appliquer cela avec $\varepsilon = \frac{1}{2} \times \frac{\partial l}{\partial X_j}(\hat{X}^{\infty}) > 0$, ce qui donne l'existence de $n_0 > 0$ tel que pour tout $n > n_0$:

$$\frac{\partial l}{\partial X_j}(\hat{X}^n) - \frac{\partial l}{\partial X_j}(\hat{X}^{\infty}) < \frac{1}{2} \times \frac{\partial l}{\partial X_j}(\hat{X}^{\infty}) \tag{3-38}$$

Soit donc :

$$\frac{\partial l}{\partial X_j}(\hat{X}^n) > \frac{1}{2} \times \frac{\partial l}{\partial X_j}(\hat{X}^{\infty}) \tag{3-39}$$

Donc, en utilisant l'équation (3-20) :

$$-1 + \sum_i \frac{p(i|j)Y_i}{\sum_{j'} p(i|j')\hat{X}_{j'}^{n}} > \frac{1}{2} \times \frac{\partial l}{\partial X_j}(\hat{X}^{\infty}) \tag{3-40}$$

Donc on obtient, pour tout $n > n_0$:

$$\hat{X}_j^n \sum_i \frac{p(i|j)Y_i}{\sum_{j'} p(i|j')\hat{X}_{j'}^{n}} > \hat{X}_j^n\left(1 + \frac{1}{2} \times \frac{\partial l}{\partial X_j}(\hat{X}^{\infty})\right) \tag{3-41}$$

En utilisant l'équation de récurrence (3-28), on obtient :

$$\hat{X}_j^{n+1} > \hat{X}_j^n\left(1 + \frac{1}{2} \times \frac{\partial l}{\partial X_j}(\hat{X}^{\infty})\right) \tag{3-42}$$

On peut appliquer cette formule de récurrence à partir de n_0, et donc, pour tout $m > 0$, on a :

$$\hat{X}_j^{n_0+m} > \hat{X}_j^{n_0}\left(1+\frac{1}{2}\times\frac{\partial l}{\partial X_j}\left(\hat{X}^{\infty}\right)\right)^m \tag{3-43}$$

Si $\hat{X}_j^{n_0} > 0$, alors, quand m tend vers $+\infty$, on a nécessairement $\hat{X}_j^{n_0+m}$ qui tend aussi vers $+\infty$ car $1+\frac{1}{2}\times\frac{\partial l}{\partial \lambda_j}\left(\hat{X}^{\infty}\right) > 1$, par l'hypothèse que j'ai faite au début de ce raisonnement par l'absurde. Cela est en contradiction avec le fait que $\hat{X}_j^{n_0+m}$ tende vers $\hat{X}_j^{\infty} = 0$, ce qui conclut le raisonnement par l'absurde. Donc dans le cas où $\hat{X}_j^{n_0} > 0$, on a démontré que $\frac{\partial l}{\partial \lambda_j}\left(\hat{X}^{\infty}\right)$ est nécessairement négatif, ce qui démontre la condition (3-27) dans ce cas.

Si $\hat{X}_j^{n_0} = 0$, cela signifie d'une part que l'algorithme a convergé et donc que $\hat{X}_j^{n_0} = \hat{X}_j^{\infty} = 0$. Si on choisit n_0 comme étant le premier indice où $\hat{X}_j^{n_0} = 0$, alors nécessairement, par l'équation (3-29) :

$$\sum_i \frac{p(i|j)Y_i}{\sum_{j'} p(i|j')\hat{X}_{j'}^{n_0-1}} = 0 \tag{3-44}$$

Or, comme chaque terme de la somme est positif, pour que la somme soit nulle, il faut que chaque terme soit nul individuellement et donc que pour tout i, on a :

$$\frac{p(i|j)Y_i}{\sum_{j'} p(i|j')\hat{X}_{j'}^{n_0-1}} = 0 \tag{3-45}$$

Donc on a pour tout i :

$$p(i|j)Y_i = 0 \tag{3-46}$$

Cela signifie que, pour tout n,

$$\frac{\partial l}{\partial X_j}\left(\hat{X}^{n}\right) = -1 + \sum_i \frac{p(i|j)Y_i}{\sum_{j'} p(i|j')\hat{X}_{j'}^{n}} = -1 \tag{3-47}$$

Et donc de fait :

$$\frac{\partial l}{\partial X_j}\left(\hat{X}^{\infty}\right) = -1 < 0 \tag{3-48}$$

Donc, s'il existe n_0 tel que $\hat{X}_j^{n_0} = 0$, on a aussi $\frac{\partial l}{\partial X_j}\left(\hat{X}^{\infty}\right) \leq 0$, ce qui démontre la condition (3-27) dans ce cas.

Les conditions de Kuhn-Tucker sont vérifiées, et $\hat{X}^{\infty}$ fournit bien un maximum global de la log-vraisemblance l.

4 Conclusion

À travers cette annexe, j'ai montré que l'algorithme MLEM peut être obtenu par deux approches différentes.

Une première approche repose sur l'application de la formule de Bayes et la recherche d'un point fixe. Cette approche ne présuppose pas de forme particulière des données. Par exemple, on ne suppose pas que les données suivent une loi de Poisson. Si l'algorithme défini converge, alors il fournit un point fixe solution. Cependant, le défaut de cette première approche est qu'elle ne garantit pas la convergence de l'algorithme dans tous les cas.

Une deuxième approche consiste à maximiser la log-vraisemblance, en supposant cette fois que les données suivent une loi de Poisson. On obtient le même algorithme itératif que la première approche à travers un algorithme de type EM. L'intérêt de cette deuxième approche est qu'on peut prouver de manière solide la convergence de l'algorithme vers un maximum de la vraisemblance.

5 Bibliographie

Dempster, A., Laird, N. & Rubin, D., 1977. *Maximum Likelihood from Incomplete Data via the EM Algorithm.* in: Journal of the Royal Statistical Society. Series B (Methodological), Vol. 39, No. 1, pp. 1-38.

Salvo, K. & Defrise, M., 2019. *A Convergence Proof of MLEM and MLEM-3 With Fixed Background.* in: IEEE Transactions on Medical Imaging, Volume: 38, Issue: 3.

Shepp, L. & Vardi, Y., 1982. *Maximum Likelihood Reconstruction for Emission Tomography.* in: IEEE Transactions on Medical Imaging, Vol. 1, Issue 2.

Willingale, R., Sims, M. & Turner, M., 1984. *Advanced deconvolution technique for coded aperture imaging.* in: Nuclear Instruments and Methods in Physics Research, Volume 221, Issue 1, Pages 60-66.

Zangwill, W., 1969. *Nonlinear programming: a unified approach.* in: Prentice-Hall.

Annexe E : Reconstruction Compton sur des données IBIS/INTEGRAL

Cette annexe présente l'application des méthodes de reconstruction Compton que j'ai étudiées et développées au Chapitre 6 à des données issues de la mission INTEGRAL et plus particulièrement de l'instrument IBIS.

Je présente dans un premier temps l'instrument IBIS, que j'avais brièvement décrit dans le Chapitre 1 puis je montre les données que je dois traiter et leurs particularités. Enfin, j'applique des méthodes de reconstruction Compton sur ces données.

1 L'instrument IBIS

La mission INTEGRAL (ESA, 2019) est issue d'une coopération entre la NASA, l'ESA et la Russie dans l'objectif d'observer l'Univers dans la gamme des rayonnements électromagnétiques de haute énergie, et spécifiquement les sources de photons gamma.

Le satellite INTEGRAL embarque pour ce faire deux instruments principaux. D'une part, l'instrument SPI (*SPectrometer on Integral*) est un spectromètre constitué de détecteurs Germanium Haute Pureté (HPGe). Il est principalement utilisé pour la mesure des énergies des photons gamma dans la bande 18 keV – 8 MeV, avec une très bonne précision grâce à l'utilisation du Germanium, garantissant une résolution de 2,2 keV FWHM à 1,3 MeV. L'instrument est équipé d'un masque codé de sorte à lui fournir une capacité d'imagerie avec une résolution angulaire de 2,5° pour un champ de vue total de 30° environ.

D'autre part, l'instrument IBIS (*Imager on Board the Integral Satellite*) est un imageur à masque codé dont l'objectif est de fournir des images avec une résolution de 12' FWHM, soit 12,5 fois meilleure que l'instrument SPI, pour un champ de vue total de 29,1°x29,4° et un champ de vue entièrement codé de 8,3°x8°.

L'imageur est composé d'un masque codé de type URA (*Uniformly Redondant Array*) et de deux plans de détection parallèles séparés de 90 mm : le détecteur ISGRI et le détecteur PICsIT. La Figure 1-1 montre un schéma de cet instrument.

Le détecteur ISGRI correspond au premier détecteur, constitué d'une matrice de 8 modules de 64x32 pixels CdTe, soit un total de 128x128 pixels. Chaque pixel mesure 4x4 mm² pour 2 mm d'épaisseur, soit une surface de détection de 2621 cm². Du point de vue des performances spectrométriques, la dynamique en énergie couverte par ISGRI va de 15 keV à 950 keV, pour une résolution de 8 % FWHM à 122 keV (ESA, 2019).

Le détecteur PICsIT, situé 9 cm sous ISGRI, est un détecteur constitué d'une matrice de 8 modules de 32x16 pixels en Iodure de Césium (CsI), soit un total de 64x64 pixels. La surface d'un pixel est de 9x9 mm² et leur épaisseur est de 30 mm, pour couvrir une surface de détection de 3318 cm². Ce

détecteur est capable d'observer les photons gamma dans une bande plus large qu'ISGRI, allant de 100 keV à 10 MeV, avec une résolution de 12 % FWHM à 662 keV (Labanti, et al., 2003).

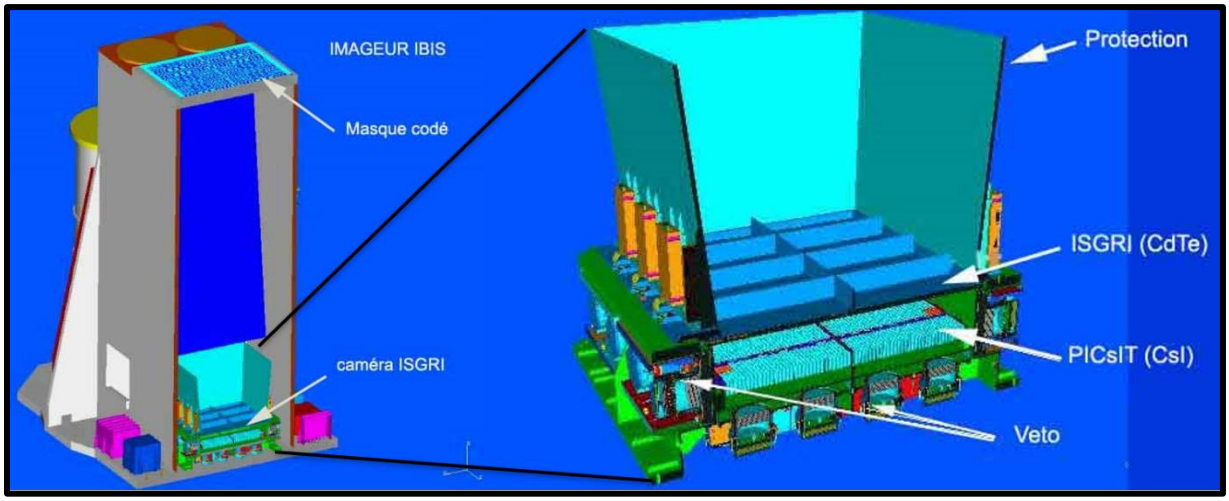


Figure 1-1 : À gauche : Schéma de l'imageur IBIS. À droite : Zoom sur les deux plans de détection ISGRI et PICsIT (CEA/IRFU/DAp, 2001).

L'utilisation de deux plans de détection parallèles est une configuration favorable à la reconstruction d'images Compton, le détecteur ISGRI servant de plan de diffusion et le détecteur PICsIT de plan d'absorption. J'étudie ici cette possibilité au regard des algorithmes que j'ai développés au Chapitre 6, sur un exemple de sursaut gamma ou *Gamma-Ray Burst* (GRB).

2 Application de la reconstruction Compton à un GRB observé par IBIS

2.1 Étude spectrométrique

Les données que j'utilise sont issues d'une acquisition faite par l'imageur IBIS lors de l'observation du sursaut gamma GRB181201A observé le 01/12/2018 vers 02:40 UTC (Laskar, et al., 2019). L'intérêt de cet événement est qu'il a été localisé avec la méthode à masque codé, ce qui me permettra de comparer la position que je retrouve à l'aide de la localisation Compton avec celle qui a été effectivement retrouvée par la méthode à masque codé.

L'ensemble des événements dont je dispose correspond aux événements doubles qui ont eu lieu en coïncidence temporelle dans ISGRI et dans PICsIT, soit un événement dans ISGRI et un dans PICsIT. La première étape est, parmi ces événements, de retrouver ceux qui ont eu lieu au moment du sursaut gamma. Pour ce faire, je trace en Figure 2-1 la date d'arrivée des événements en *Integral Julian Day* (IJD) ainsi que le nombre d'événements ayant lieu dans les mêmes fenêtres temporelles renvoyées par l'instrument.

Sur cette figure, on observe deux ruptures dans la pente des dates d'arrivée des événements, j'ai sélectionné une fenêtre des événements situés entre ces deux ruptures de pente. Cette fenêtre

correspond à des dates pour lesquelles les fenêtres temporelles comptent jusqu'à 4 fois plus d'événements que dans la période hors du sursaut gamma.

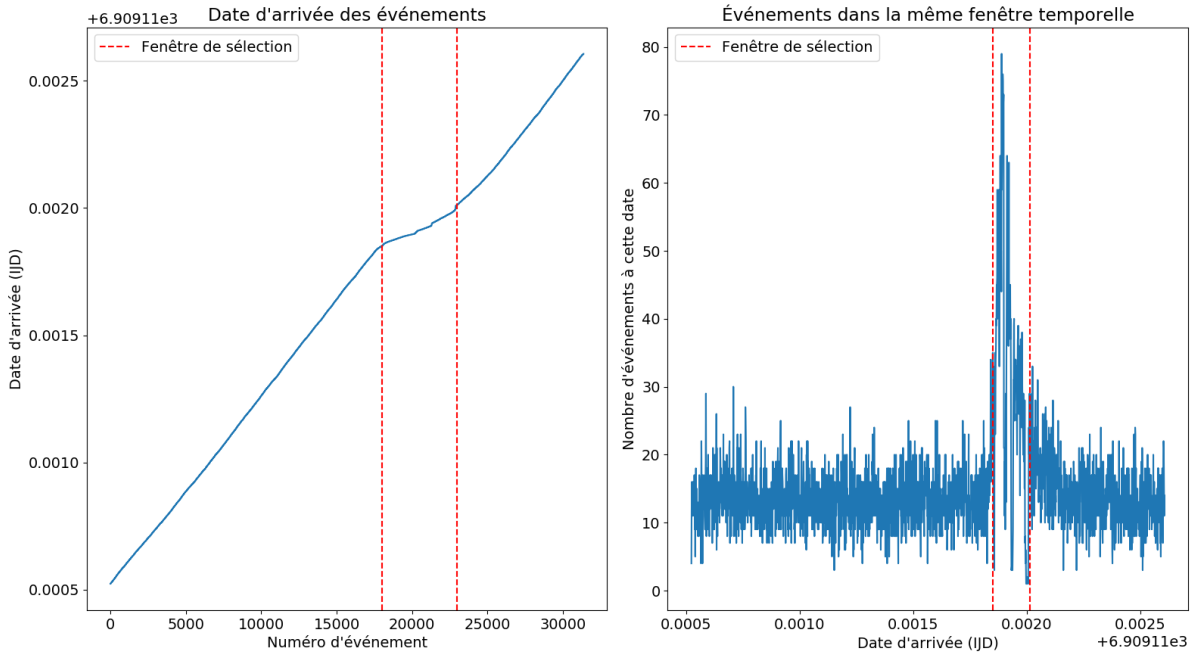


Figure 2-1 : À gauche : Date d'arrivée des événements détectés lors du sursaut gamma. À droite : Nombre d'événements détectés par fenêtre temporelle en fonction de la date d'arrivée.

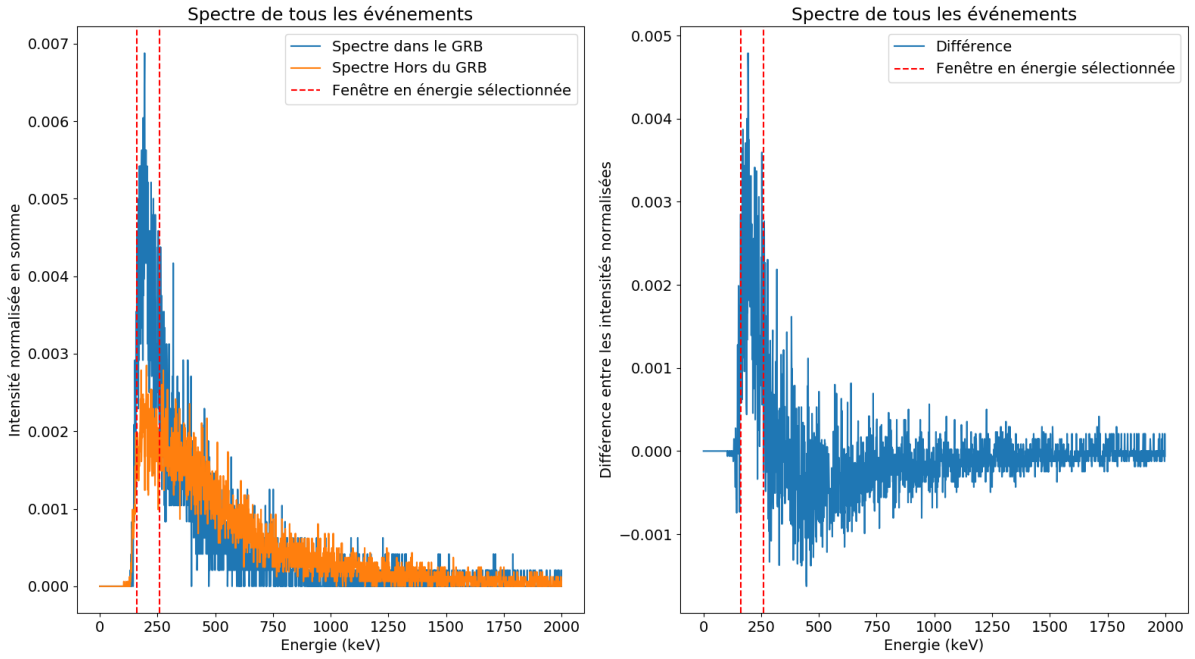


Figure 2-2 : À gauche : Spectre de la somme des énergies des événements en coïncidence dans ISGRI et PICsIT, avant et pendant le GRB. À droite : Différence entre les deux spectres.

Pour chacun des événements, je dispose des informations des interactions dans ISGRI et PICsIT : les deux pixels touchés et les deux énergies déposées dans chaque détecteur. J'effectue dans un premier temps le spectre de la somme des deux énergies déposées, à la fois dans la fenêtre temporelle que j'ai

identifiée, et pour les événements avant cette fenêtre temporelle. Je représente ce spectre normalisé en somme ainsi que la différence entre les deux spectres en Figure 2-2**.**

Je sélectionne les événements par rapport à leur énergie en fonction du spectre, dans la bande d'énergie dans laquelle le spectre du GRB domine par rapport au bruit de fond, soit entre 160 keV et 260 keV, cela correspond à 1851 événements.

2.2 Reconstruction Compton

Pour effectuer la reconstruction Compton, j'ai présenté dans le Chapitre 6 le cas où on connaît l'énergie initiale des photons avant interaction dans le détecteur. Du point de vue des applications, cela correspond aux cas où un radioélément est présent, avec une ligne d'émission bien précise.

Dans le cas des données dont je dispose, comme on le voit sur le spectre en Figure 2-2, il n'existe pas de pic dans ce spectre sur lequel je peux me concentrer, et l'énergie initiale des photons est inconnue. Dans ce cas, je fais la reconstruction dans l'hypothèse que l'ensemble de l'énergie des photons est déposé dans les deux détecteurs, soit par conservation de l'énergie :

$$E_0 = E_1 + E_2 \qquad \text{(II-1)}$$

où E_1 est l'énergie déposée par absorption photoélectrique, E_2 l'énergie déposée lors de la diffusion Compton et E_0 est l'énergie initiale du photon, déduite de E_1 et E_2. Cette hypothèse est a priori la principale source d'erreur pour ce système de détection pour effectuer une reconstruction Compton correcte.

Dans le Chapitre 6, avec des applications avec Caliste, j'ai décrit la problématique liée à l'identification de la diffusion Compton et de l'absorption photoélectrique parmi les deux interactions mesurées. Ici, vu la configuration des deux détecteurs, je fais l'hypothèse que la diffusion Compton a lieu dans ISGRI et que l'absorption photoélectrique a lieu dans PICsIT, soit donc E_2 est l'énergie mesurée dans ISGRI et E_1 est mesurée dans PICsIT.

En utilisant ces différentes hypothèses, je réalise en Figure 2-3 la *backprojection* directe des cônes sur le ciel. Contrairement au Chapitre 6, du fait que j'utilise deux plans de détections différents ici, je peux couvrir un champ de vue de 360° sans avoir de symétrie due à celle du plan de détection. Sur cette reconstruction, l'axe $(0°, 0°)$ correspond à l'axe perpendiculaire aux deux détecteurs, vers l'avant du détecteur ISGRI et l'axe $(180°, 0°)$ correspond au même axe, dans le sens vers l'arrière du détecteur PICsIT.

La position de la source observée est connue par l'utilisation du masque codé : la longitude est de $-8{,}23° \pm 0{,}2°$ et la latitude est de $+8{,}17° \pm 0{,}2°$ dans le système de coordonnées que j'utilise. Elle permet de vérifier la localisation effectuée par reconstruction Compton.

Sur la *backprojection* directe, le maximum d'intersections est retrouvé à deux degrés d'écart par rapport à la position attendue.

J'ai aussi appliqué l'algorithme bayésien présenté au Chapitre 6, le résultat est donné en Figure 2-4. La largeur à mi-hauteur de la reconstruction est de 20°. On peut aussi prendre un critère plus restrictif, par exemple la largeur à 75 % de la hauteur vaut 6,2°. La position attendue se situe bien dans cette marge d'erreur, à 2° du maximum renvoyé par l'algorithme bayésien. Il est intéressant de noter que le

premier détecteur servant de couche de diffusion est composé de CdTe et, en utilisant les calculs de l'incertitude liée à l'élargissement Doppler que j'ai effectués en Chapitre 6 pour ce matériau, on en déduit que l'incertitude typique que l'on peut attendre sur la reconstruction Compton est de l'ordre de 2°.

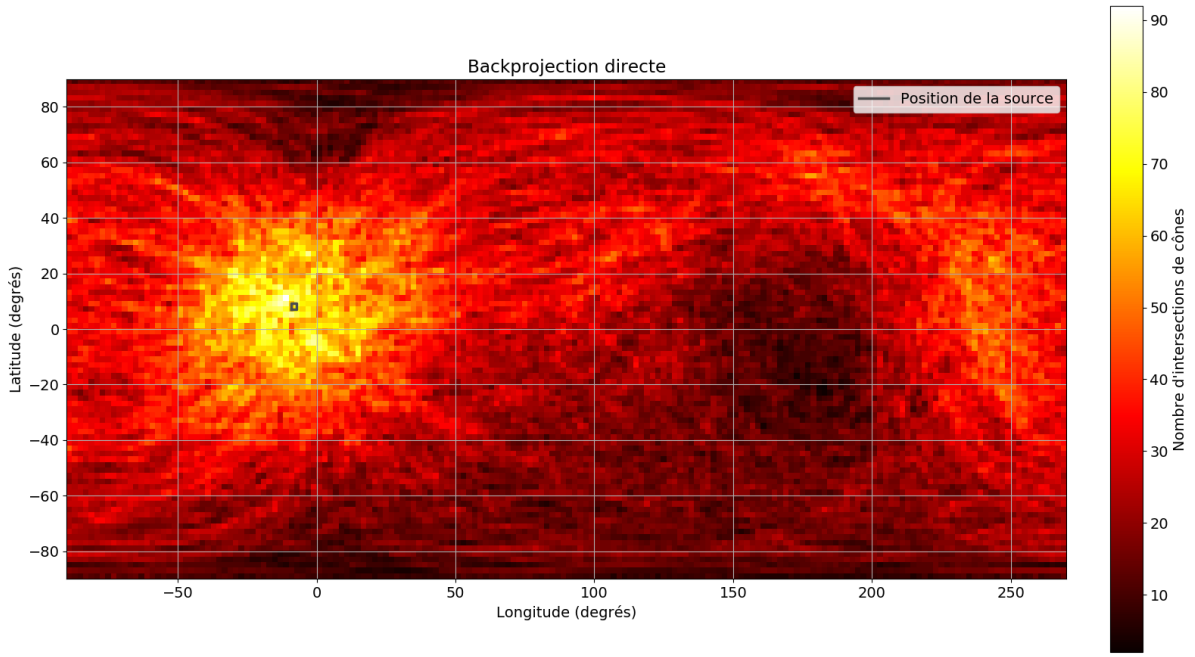


Figure 2-3 : Résultat de la *backprojection* directe.

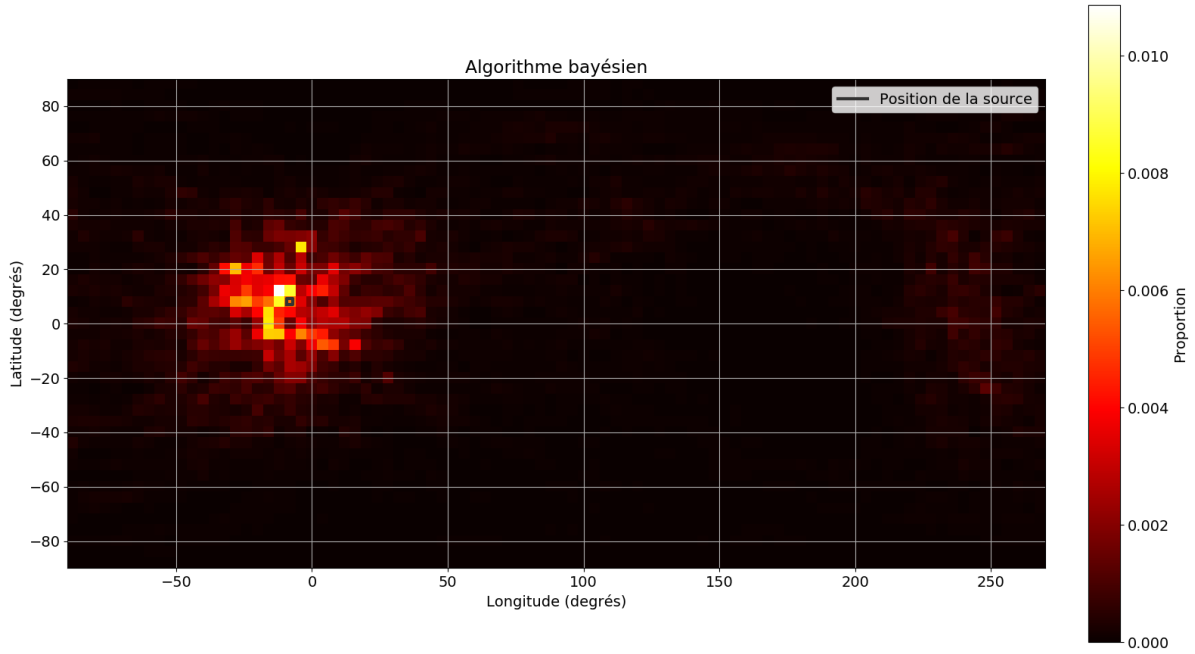


Figure 2-4 : Résultat de l'algorithme bayésien.

J'ai cherché à appliquer l'algorithme SOE-RR, mais il ne converge pas et l'image renvoyée est une répartition aléatoire sur le ciel, sans identifier de zone d'accumulation des événements. L'image qui en est issue n'est pas exploitable. Cela provient principalement du fait que le nombre d'événements

est trop faible devant un champ de vue de 360°x180°, en comparaison au champ de vue réduit à 180°x180° que j'avais en Chapitre 6 par la symétrie par rapport au plan de détection. Afin d'appliquer l'algorithme SOE-RR, il faudrait deux fois plus d'événements. Par ailleurs, je ne peux pas appliquer l'inversion Compton rapidement sans effectuer des simulations pour chaque position du ciel pour ce système de détection, ce qui est toutefois possible.

Pour conclure sur l'application de la reconstruction Compton avec l'instrument IBIS à bord d'INTEGRAL, on voit ici l'intérêt de cette reconstruction par le fait qu'elle est capable de localiser la source du GRB avec une erreur de 6,2° pour l'algorithme bayésien. Cette erreur est plus grande que celle attendue avec l'imagerie à masque codé, dont la résolution est de 12', soit 116 fois meilleure.

En revanche, une combinaison avec l'information Compton est particulièrement intéressante dans le cas où un phénomène d'intérêt serait localisé en dehors du champ de vue du masque. Le télescope pourrait ainsi être pointé dans la direction de ce phénomène dans une boîte d'erreur de l'ordre d'un peu plus de 6°, erreur qui est contenue dans le champ de vue totalement codé du masque d'IBIS. Puis, une localisation plus précise peut être effectuée par l'imagerie à masque codé. Cela demande aussi à ce que le phénomène soit suffisamment long ou intense pour que, d'une part, suffisamment d'événements Compton soient détectés pour localiser la source, et d'autre part, le satellite ait le temps d'être pointé dans la direction trouvée par localisation Compton.

3 Bibliographie

CEA/IRFU/DAp, 2001. *ISGRI : a camera for invisible light.* [En ligne]
Available at: http://irfu.cea.fr/dap/Phocea/Vie_des_labos/Ast/ast_visu.php?id_ast=1371

ESA, 2019. *Integral.* [En ligne]
Available at: https://sci.esa.int/web/integral

Labanti, C. et al., 2003. *The Ibis-Picsit detector onboard Integral.* in: A&A 411, L149-L152.

Laskar, T. et al., 2019. *A Reverse Shock in GRB 181201A.* in: The Astrophysical Journal, 884:121.