



HAL
open science

Amélioration des détecteurs CdZnTe pour l'imagerie gamma par apprentissage

Alexandre Delcourt

► **To cite this version:**

Alexandre Delcourt. Amélioration des détecteurs CdZnTe pour l'imagerie gamma par apprentissage. Physique mathématique [math-ph]. Université Grenoble Alpes [2020-..], 2023. Français. NNT : 2023GRALM056 . tel-04523963

HAL Id: tel-04523963

<https://theses.hal.science/tel-04523963>

Submitted on 27 Mar 2024

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.

THÈSE

Pour obtenir le grade de

DOCTEUR DE L'UNIVERSITÉ GRENOBLE ALPES

École doctorale : MSTII - Mathématiques, Sciences et technologies de l'information, Informatique

Spécialité : Mathématiques Appliquées

Unité de recherche : Laboratoire Jean Kuntzmann

Amélioration des détecteurs CdZnTe pour l'imagerie gamma par apprentissage

Improved CdZnTe gamma imager performance through model learning

Présentée par :

Alexandre DELCOURT

Direction de thèse :

Roland HILDEBRAND

CHARGE DE RECHERCHE HDR, CNRS DELEGATION ALPES

Guillaume MONTÉMONT

INGENIEUR DE RECHERCHE, CEA CENTRE DE GRENOBLE

Directeur de thèse

Co-encadrant de thèse

Rapporteurs :

FRANÇOIS GLINEUR

FULL PROFESSOR, UNIVERSITE CATHOLIQUE DE LOUVAIN

VINCENT MÉTIVIER

PROFESSEUR, IMT ATLANTIQUE

Thèse soutenue publiquement le **21 décembre 2023**, devant le jury composé de :

SANA LOUHICHI,

PROFESSEURE DES UNIVERSITES, UNIVERSITE GRENOBLE ALPES

Présidente

ROLAND HILDEBRAND,

CHARGE DE RECHERCHE HDR, CNRS DELEGATION ALPES

Directeur de thèse

FRANÇOIS GLINEUR,

FULL PROFESSOR, UNIVERSITE CATHOLIQUE DE LOUVAIN

Rapporteur

VINCENT MÉTIVIER,

PROFESSEUR, IMT ATLANTIQUE

Rapporteur

DAVID BRASSE,

DIRECTEUR DE RECHERCHE HDR, CNRS DELEGATION ALSACE

Examineur

Invités :

GUILLAUME MONTÉMONT

INGENIEUR DE RECHERCHE, CEA CENTRE DE GRENOBLE



Table des matières

Introduction	3
Motivations et démarche	5
1 Etat de l'art	9
1.1 Imagerie γ et détecteurs	9
1.1.1 Interaction rayonnement γ -matière	9
1.1.2 Détection des rayonnements γ	13
1.1.3 Différents types de détecteurs	17
1.1.4 Applications	19
1.2 Description des détecteurs CZT	21
1.2.1 Caractéristiques du CZT	21
1.2.2 Physique des semi-conducteurs	21
1.3 Apparition et caractérisation des défauts cristallins du CZT	25
1.3.1 Structure cristalline du CZT	25
1.3.2 Méthodes de croissance usuelles	26
1.3.3 Caractérisation du détecteur	27
1.3.4 Observation des principaux types de défauts	30
1.4 Modélisation des détecteurs	35
1.4.1 Évolution du développement de la modélisation des détecteurs semi-conducteurs	35
1.4.2 Modélisation et calcul du champ électrique	38
1.4.3 Création et transport des porteurs de charge	42
1.5 Conclusion	44
2 Développement de simulations 3D de détecteurs CZT avec défauts	45
2.1 Modélisation du transport	45
2.1.1 Équation de Hecht analytique	45
2.1.2 Simulation 1D	47
2.1.3 Simulation 3D	49
2.2 Approche de modélisation avec défauts	53
2.2.1 Modélisation détecteur	53
2.2.2 Modélisation des défauts	54
2.3 Utilisation de l'accélération GPU avec OpenCL	58
2.4 Impacts des défauts sur le signal de réponse	61
2.5 Conclusion	65
3 Apprentissage supervisé de la position par réseau de neurones	67

3.1	Localisation par sous-pixellisation	67
3.1.1	Théorie mathématique	67
3.1.2	Application à la localisation 3D	69
3.2	Structure du réseau de neurones	73
3.2.1	Fonctionnement d'un réseau de neurones pour l'apprentissage supervisé	74
3.2.2	Architecture du réseau utilisé	76
3.3	Test de l'apprentissage sur simulation	80
3.3.1	Etude avec quatre voisins	80
3.3.2	Étude avec huit voisins	83
3.3.3	Comparaison pour d'autres images simulées	85
3.4	Application aux données expérimentales	90
3.4.1	Utilisation des données de l'ESRF	90
3.4.2	Utilisation des données du banc expérimental	92
3.5	Conclusion	94
4	Apprentissage non supervisé des paramètres physiques du détecteur par équation différentielle adjointe	97
4.1	Modèle Neural ODE	97
4.1.1	Théorie mathématique	97
4.1.2	Apprentissage par équation adjointe	99
4.1.3	Application à notre problème	100
4.1.4	Propagation avant	103
4.1.5	Rétropropagation	104
4.2	Cas à une dimension	105
4.2.1	Équations du problème	105
4.2.2	Mise à jour avec le calcul direct	107
4.3	Cas à trois dimensions	109
4.3.1	Équations du problème	109
4.3.2	Résolution directe	111
4.3.3	Rétropropagation	111
4.4	Développement de l'algorithme	112
4.4.1	Implémentation et fonctionnement	113
4.4.2	Résultats sur simulation	114
4.4.3	Exploitation sur détecteur réel	115
4.5	Conclusion	117
	Conclusion	119
	Bibliographie	123
A	Intégration analytique en une dimension	131
A.1	Calcul direct des gradients	131
A.2	Rétropropagation	133
A.2.1	Charge du nuage électronique	133
A.2.2	Position initiale	134
A.2.3	Paramètres de champ	134

Acronymes

API : Application Programming Interface

CIE : Charge Induction Efficiency

CPU : Central Processing Unit

CdTe : Tellurure de Cadmium

CZT / CdZnTe : Tellurure de Cadmium/Zinc

eV : electron-volt

FPGA : Field-Programmable Gate Array

GPU : Graphical Processing Unit

HgI₂ : Iodure de Mercure

IR : Infrarouge

MLEM : Maximum-Likelihood Expectation Maximisation

NaI(Tl) : Iodure de Sodium dopé au Thallium

ODE : Ordinary Differential Equation

PET / TEP : Positron Emission Tomography / Tomographie à émission de positions

RMS : Root-Mean Square

SPECT / TEMP : Single-Photon Emission Computed Tomography / Tomographie par émission monophotonique

SNR : Signal-to-noise ratio

TlBr : Bromure de Thallium

THM : Travelling Heater Method

VBM : Vertical Bridgman Method

Opérateurs mathématiques

$\tilde{\nabla}$: Opérateur gradient

Δ : Opérateur Laplacien

\wedge : Opérateur produit vectoriel

$\frac{\partial}{\partial t} = \delta_t$: Opérateur dérivée partielle temporelle

Σ : Opérateur somme

\int : Opérateur intégrale

L'imagerie gamma est un procédé de localisation de substance radioactive utilisé dans plusieurs domaines. Dans le domaine médical, les caméras gamma permettent de localiser la substance radiopharmaceutique injectée dans l'organisme afin de détecter la présence d'une pathologie de manière non invasive. On utilise également ce type d'imagerie dans le domaine de la sécurité, notamment pour le contrôle des bagages et marchandises, et de la radioprotection, par exemple pour détecter d'éventuels points chauds à l'intérieur du bâtiment réacteur d'une centrale nucléaire. C'est dans les années 2000 que les matériaux semi-conducteurs à température ambiante ont commencé à être utilisés pour ces applications en imagerie X et γ . Ces matériaux ont apporté des améliorations de performances par rapport à ce qui se faisait avec les détecteurs scintillateurs, utilisant la conversion de l'énergie des radiations des sources radioactives en lumière, beaucoup plus répandus auparavant. L'utilisation des semi-conducteurs a par exemple amené le comptage des photons sur les applications en imagerie X, et en imagerie γ ils ont permis d'améliorer les performances des dispositifs, notamment les résolutions spatiale et énergétique.

Le CdZnTe, ou CZT, est un de ces matériaux semi-conducteurs dont les propriétés physiques sont des arguments pour l'exploiter comme matériau de détection dans les dispositifs présentés précédemment. De nombreux travaux ont été menés au fil des années afin d'améliorer la composition, la croissance des cristaux, la géométrie et l'électronique de traitement autour de ces détecteurs dans le but d'optimiser leurs performances.

Néanmoins l'utilisation du CZT présente une difficulté principale, qui est la présence de différents types de défauts dans la maille cristalline qui modifient les propriétés du matériau et ainsi impactent les signaux de sortie du détecteur. Ces défauts peuvent être de différentes natures et tailles : des lacunes Cd à l'échelle atomique au sous-joints de grains pouvant mesurer plusieurs cm de long. L'apparition de ces défauts lors de la croissance des cristaux de CZT freine le développement de détecteurs de volume plus important et de manière générale dégrade leurs performances.

L'objectif de cette thèse est de développer un protocole expérimental qui va permettre de corriger le signal de sortie en prenant en compte la présence et l'influence des défauts afin d'améliorer les performances du détecteur. Pour cela nous allons nous baser sur une approche d'optimisation par apprentissage à l'aide de simulations. L'idée est d'utiliser les recherches récentes sur l'apprentissage automatique, ou machine learning, et de les

appliquer aux équations de modélisation des détecteurs CZT. L'utilisation de simulations réalistes va permettre de créer un jeu de données avec différents types de défauts et de comparer ces derniers avec des données expérimentales afin de faire coïncider simulation et données réelles pour corriger les mesures de sortie et améliorer la localisation des interactions.

Le premier chapitre introduit l'état de l'art sur les différentes problématiques du sujet, tels que les détecteurs CZT, l'étude des défauts, la modélisation des détecteurs ainsi que la théorie de l'apprentissage automatique.

Le deuxième chapitre sera consacré au développement des simulations tridimensionnelles réalistes avec l'implémentation des défauts pour créer un jeu de données modulable comportant différents modèles de détecteurs. Il sera question de la physique derrière la simulation, de l'utilisation de l'accélération des calculs par GPU (Graphics Processing Units ou unité de traitement graphique, autrement dit la carte graphique) et des observations de quelques défauts.

Le troisième chapitre traitera de l'apprentissage de données supervisé par réseau de neurones appliqué à la localisation. En effet nous verrons que l'utilisation d'un réseau de neurones permet d'égaliser voire surclasser les performances de la méthode de localisation par maximum de vraisemblance utilisée aujourd'hui. On traitera également des résultats obtenus sur des données expérimentales.

Le quatrième et dernier chapitre présentera les travaux réalisés grâce à l'apprentissage non supervisé. On expliquera pourquoi il est nécessaire, la théorie qui est utilisée, son application sur les équations physiques et les résultats dans les cas à une et trois dimensions.

Motivations et démarche

La croissance des cristaux CZT pour l'imagerie est réalisée de manière artificielle selon plusieurs méthodes. Néanmoins ces techniques n'étant pas parfaites différentes contraintes sont appliquées sur le cristal lors de sa croissance, favorisant l'apparition de défauts chimiques et structurels, ayant un impact sur le bon fonctionnement de détection.

Un détecteur CZT est composé d'un cristal CZT ainsi que d'une électronique permettant le traitement du signal et la mesure des signaux issus de l'interaction des photons dans le détecteur. Les objectifs de ce type de détecteur peuvent être la localisation d'une source radioactive ainsi que la détermination de sa nature. Les performances de ce type de détecteur sont notamment caractérisées par les résolutions spatiale et en énergie, déterminant la capacité du détecteur à mesurer précisément la position et l'énergie des photons respectivement.

La présence de défauts dans un détecteur CZT entraîne une variabilité sur les signaux induits par les porteurs de charge et donc sur les signaux collectés sur les électrodes qui perturbent le courant mesuré en sortie. Pour s'affranchir de ces erreurs une des solutions est de connaître la physique interne du détecteur afin de pouvoir corriger ces erreurs d'après la nature et la répartition des défauts dans ce dernier. Néanmoins la fonction de réponse altérée de ces détecteurs du fait de la présence défauts empêche de réaliser des caractérisations précises d'un cristal de détection, qui peuvent nécessiter plusieurs heures si on veut une carte très précise à un pas sub-millimétrique. Ce type de calibration n'est donc pas optimale dans un cadre d'utilisation industriel ou médical où la durée de calibration est importante. L'objectif de cette thèse est de pouvoir élaborer un procédé de calibration des détecteurs CZT rapide, en utilisant des notions d'apprentissage à l'aide d'une simulation qui permettent de réduire drastiquement les temps mis en jeu. Le travail se déroule en quatre étapes distinctes permettant, en les imbriquant, d'aboutir à une nouvelle méthode de calibration du détecteur et donc d'améliorer ses performances en diminuant les incertitudes sur les mesures d'interaction.

1. Dans un premier temps on développe des simulations 3D réalistes de détecteurs permettant d'ajouter la présence de différents types de défauts et d'observer leur influence sur la collecte (Figure 1). A l'aide d'une source de photons gamma virtuelle on simule des interactions dans le détecteur dont les positions sont connues (X_{sim}). On paramétrise la simulation par la forme du champ électrique prenant en compte la

présence de défauts (θ), puis on calcule le mouvement des charges dans le détecteur pour obtenir des signaux induits correspondant aux signaux induits sur les anodes (Y_{sim}).

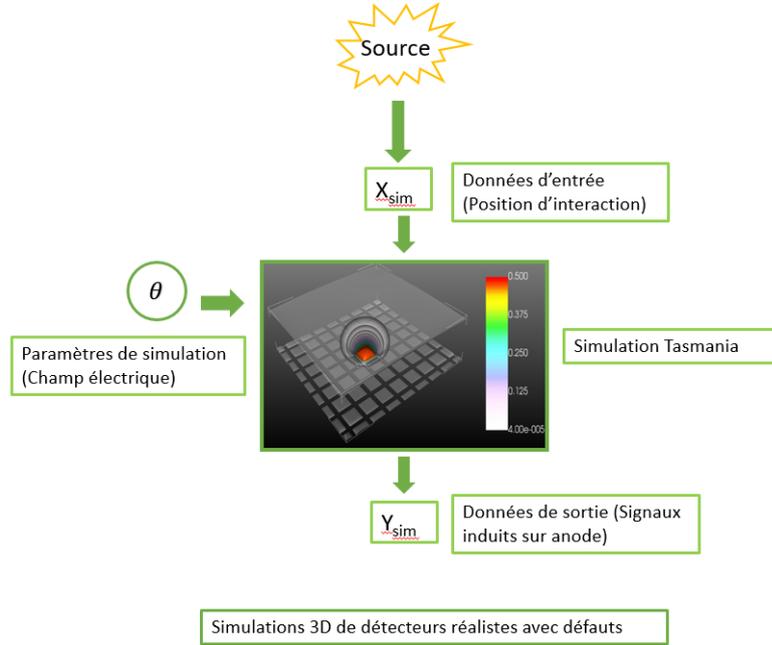


FIGURE 1 – Schéma du fonctionnement des simulations 3D.

2. Ensuite on s'intéresse à l'apprentissage des paramètres du détecteur étudié à partir des données expérimentales. Les positions d'interaction ainsi que ces paramètres sont inconnus pour un vrai détecteur (X_{exp}), et les seules données dont nous disposons sont les signaux d'induction sur les électrodes (Y_{exp}). Afin d'approcher la structure du champ électrique de notre détecteur nous allons mettre en place un algorithme d'apprentissage non supervisé des paramètres θ du détecteur simulé (Figure 2). On calcule d'abord l'erreur obtenue entre les signaux expérimentaux et simulés (ϵ_Y), et à l'aide d'une méthode d'optimisation basée sur les équations différentielles du système physique on calcule le gradient de cette erreur (∇Y) que l'on rétropropage dans la simulation pour obtenir le gradient d'erreur sur la position (∇X) pour mettre à jour cette dernière pour chaque événement. Puis de manière globale on met à jour les paramètres de la simulation à l'aide du gradient d'erreur sur ces derniers ($\nabla \theta$).
3. Parallèlement à l'étape précédente, on utilise la simulation pour estimer les positions d'interaction à l'aide d'un réseau de neurones (Figure 3). L'objectif est d'entraîner ce réseau de neurones sur un jeu de données simulées afin d'accélérer le processus. En effet la simulation permet d'obtenir très rapidement des données de détecteurs de structures différentes et ce réseau de neurones étant facilement implémentable de manière électronique il pourra être utilisé pour apprendre les positions d'après les données expérimentales. On utilise un modèle de réseau à plusieurs couches cachées, avec en entrée les données des signaux induits simulés (Y_{sim}), et en sortie la position estimée (\widehat{X}) par le réseau grâce aux poids (W). Puis on calcule l'erreur ϵ_X entre les positions connues (X_{sim}) de la simulation et les positions estimées. On met à jour les poids du réseau en rétropropageant le gradient de cette erreur $\nabla \widehat{X}$ à travers

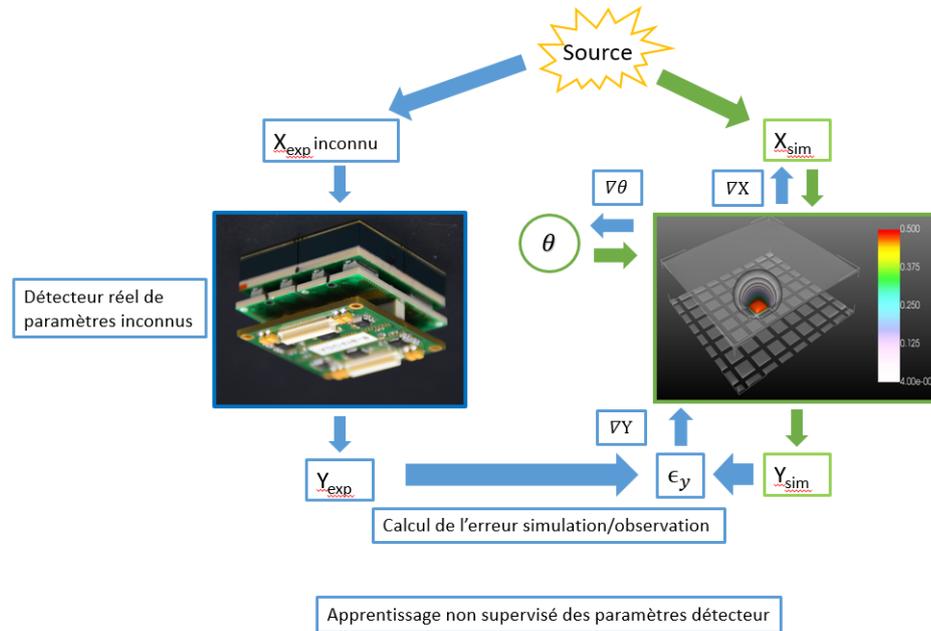


FIGURE 2 – Schéma du fonctionnement de l'apprentissage non supervisé des paramètres détecteur.

celui-ci. On itère ce processus jusqu'à obtenir une précision suffisante sur la position estimée.

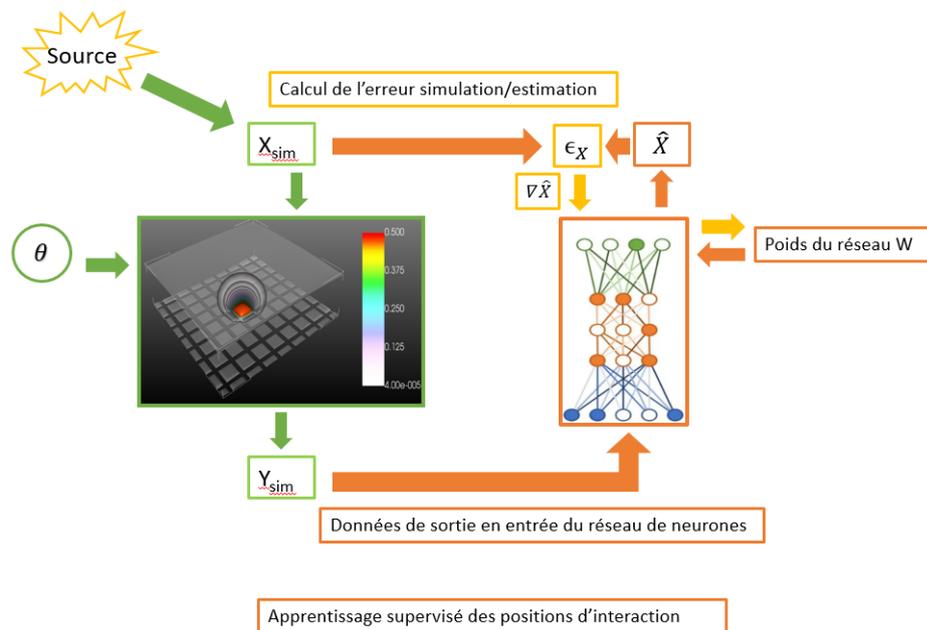


FIGURE 3 – Schéma du fonctionnement de l'apprentissage supervisé des positions d'interaction.

- La dernière étape consiste à fusionner les deux précédentes (Figure 4). La simulation nous permet à la fois d'obtenir une structure réaliste du détecteur étudié

par l'apprentissage non supervisé et d'entraîner un réseau de neurones à déterminer les positions d'interaction d'après les signaux de sortie. Une fois ces deux étapes réalisées, on peut mettre en place l'apprentissage des positions d'interaction expérimentales inconnues (X_{exp}) à partir des signaux de sortie (Y_{exp}) directement via le réseau de neurones pré-entraîné par simulation. En comparant de la même manière les positions estimées par le réseau et par le traitement du signal on améliore la précision sur la localisation.

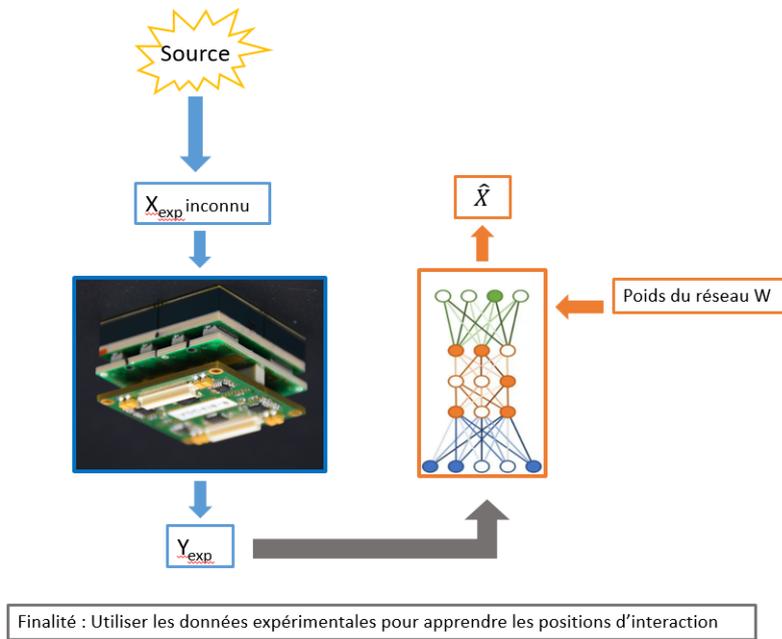


FIGURE 4 – Schéma du fonctionnement de l'optimisation de la localisation des interactions par réseau de neurones.

L'imagerie gamma est une technologie aujourd'hui utilisée dans des domaines tels que la santé, la sécurité ou la sûreté des installations nucléaires. Le principe est relativement simple : détecter des interactions de photons γ émis par une source radioactive dans l'espace, traiter les informations transmises par ces interactions et remonter à la localisation de la source par un système de reconstruction d'images. Différents dispositifs peuvent être utilisés, comportant chacun leurs avantages et inconvénients, suivant le type de détection recherché. Dans cette partie nous allons aborder les différents sujets liés au cadre de l'étude, du principe de détection des particules γ jusqu'à la modélisation des détecteurs CZT.

1.1 Imagerie γ et détecteurs

1.1.1 Interaction rayonnement γ -matière

Le rayonnement γ est un type de rayonnement ionisant provoqué par la désexcitation d'un noyau atomique provoquée par une désintégration. En effet la désintégration nucléaire amène le noyau de l'atome dans un état excité qui émet un ou plusieurs photons gamma afin de repasser dans un état stable [1]. Le transfert d'énergie se fait d'après l'équation



Avec A le nombre de nucléons (protons et neutrons), Z le nombre de protons et n le nombre de photons émis lors de la désexcitation. L'énergie libérée dépend des niveaux d'énergie de transition du noyau (Figure 1.1).

Les modes d'absorption de ces photons γ sont de différentes natures selon les énergies mises en jeu. En effet selon l'énergie du photon émis par le noyau celui-ci peut interagir de différente manière avec la matière du milieu avec des probabilités dépendant de ce dernier (Figure 1.2). La physique de ces interactions est détaillée dans les paragraphes suivants.

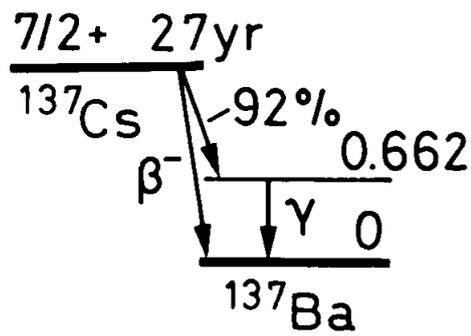


FIGURE 1.1 – Diagramme de désintégration du noyau de l'atome de césium 137. Celui-ci se désintègre dans 92% des cas par une émission β^- en baryum 137 excité qui émet un photon γ de 662 keV pour rejoindre son état fondamental [2].

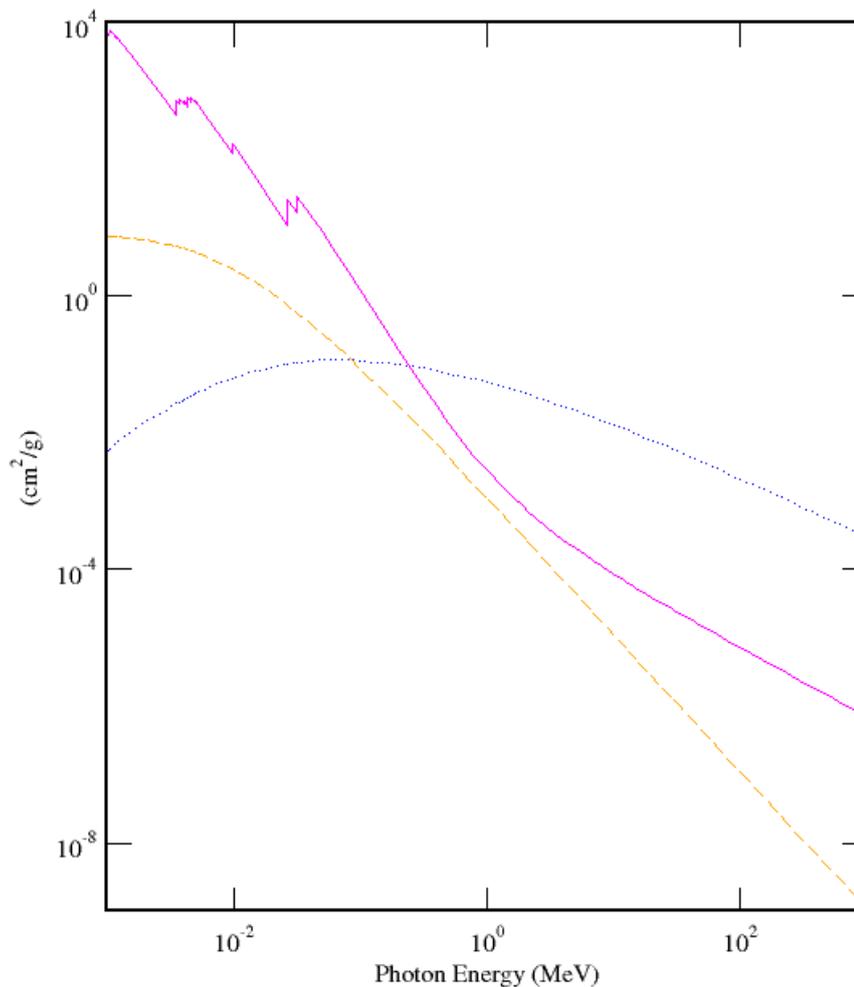


FIGURE 1.2 – Graphique du coefficient d'atténuation massique de l'effet photoélectrique (rose), de la diffusion Rayleigh (bleu) et de l'effet Compton (orange) en fonction de l'énergie du photon dans le CZT.

Diffusion Rayleigh

La diffusion Rayleigh correspond à une diffusion élastique d'un photon incident, sans transmission d'énergie. Le photon, en tant que rayonnement électromagnétique va déformer le nuage électronique d'un atome. Son barycentre va osciller et donc être décalé par rapport au barycentre des charges totales de l'atome (incluant les protons du noyau). De ce fait on a création d'un dipôle qui va rayonner de manière cohérente et le photon incident va être réémis avec la même longueur d'onde mais dans une direction différente. La diffusion Rayleigh permet l'analyse de matériaux grâce aux signatures de diffraction mais ne contribue pas à la mesure dans un détecteur puisqu'elle ne dépose pas d'énergie.

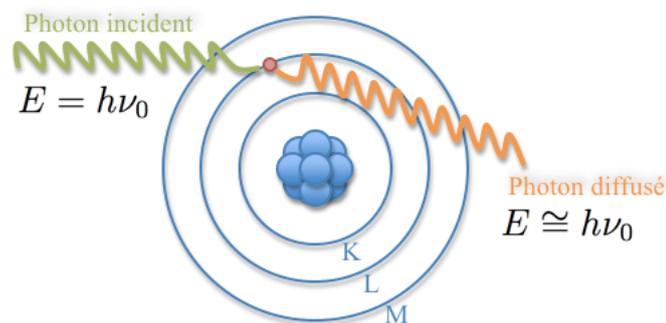


FIGURE 1.3 – Schéma de principe de la diffusion Rayleigh [3].

Effet Photoélectrique

L'effet photoélectrique est le phénomène prédominant pour les photons de basse et moyenne énergie (dans le cas du CZT cela correspond aux photons dont l'énergie est inférieure à 200 keV). Ce processus se déroule en deux étapes : d'abord le photon incident interagit avec un électron du cortège électronique d'un atome possédant une énergie de liaison inférieure à son énergie cinétique et lui transfère cette dernière. Ceci fait passer l'atome dans un état d'énergie instable et qui retrouve sa stabilité lorsqu'un électron d'une couche extérieure vient combler cette lacune. On a alors émission d'un rayonnement X de fluorescence lors de la transition de l'électron dont l'énergie correspond à la différence d'énergie de liaison entre les deux couches concernées. Dans la pratique on exploite cet effet en spectrométrie XRF (X-ray Fluorescence Spectroscopy) pour tracer des spectres en énergie afin de réaliser par exemple de l'identification d'espèces chimiques. Le photoélectron éjecté de l'atome va quant à lui va continuer de diffuser et d'interagir avec d'autres atomes, et au fur et à mesure se thermaliser en entraînant une cascade d'ionisations.

Effet Compton

L'effet Compton est une diffusion inélastique : un photon incident interagit avec un électron lié à un atome et transmet une partie de son énergie cinétique à ce dernier. On a alors éjection de cet électron qui devient un photoélectron et diffusion d'un photon [5]. Si on considère l'électron initialement au repos on a :

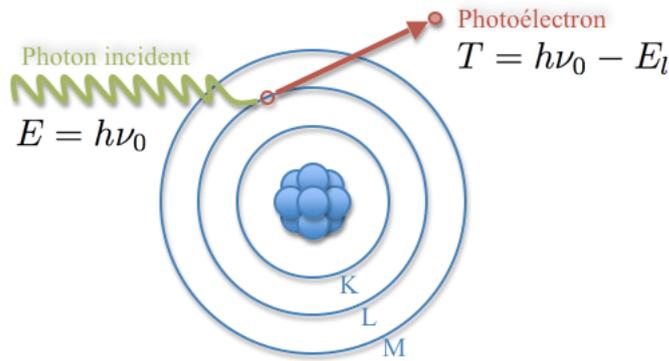
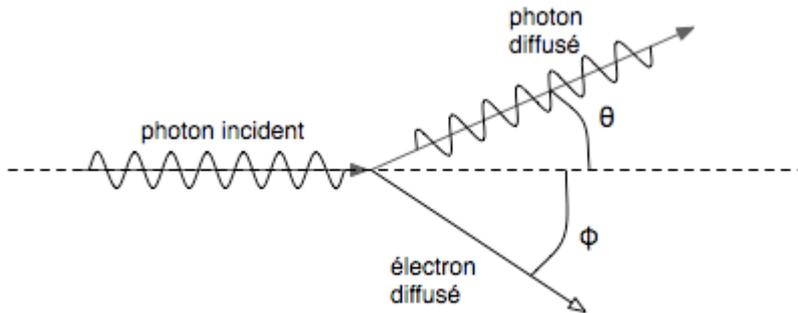


FIGURE 1.4 – Schéma de principe de l'effet Photoélectrique [4].

$$\vec{p} = \vec{p}' + \vec{q} \quad (1.2)$$

$$E + mc^2 = E' + \sqrt{m^2c^4 + q^2c^2} \quad (1.3)$$

Avec $E = pc$ l'énergie du photon incident, $E' = p'c$ l'énergie du photon diffusé, \vec{q} l'impulsion du photoélectron et m la masse de l'électron



Diffusion Compton: Collision d'un photon avec un électron au repos

FIGURE 1.5 – Schéma de principe de l'effet Compton [6].

Grâce à ces lois de conservation, on aboutit à la relation :

$$1 - \cos \theta = mc^2 \cdot \left(\frac{1}{E'} - \frac{1}{E} \right) \quad (1.4)$$

Avec E et E' les énergies des photons incident et diffusé, m_0 la masse de l'électron, θ l'angle de diffusion du photon et c la célérité de la lumière dans le vide. On utilise la physique de l'effet Compton afin de réaliser un type d'imagerie exploitant le photon diffusé suite à l'interaction afin d'utiliser un plan de diffusion et un plan de détection permettant de localiser la position des interactions.

Effet de production de paires

L'effet de production peut se dérouler lorsqu'un photon d'une énergie supérieure à 1022 keV entre dans le champ coulombien d'un noyau et qu'il se transforme en deux antiparticules de charges opposées : un positon et un électron.

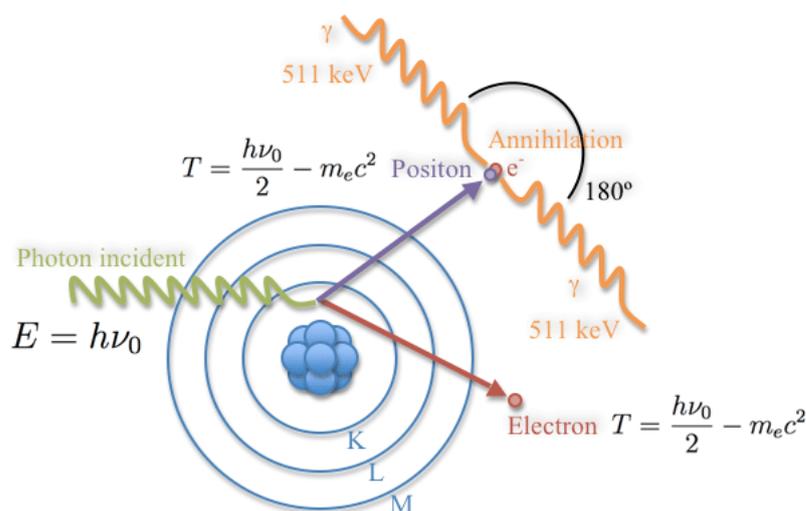


FIGURE 1.6 – Schéma de principe de l'effet de production de paires [7].

Suite à cette réaction, l'électron créé est ralenti et arrêté dans la matière, et le positon s'associe avec un électron du milieu pour former un positronium, qui est l'état lié d'un électron avec son antiparticule le positron. Ce dernier s'annihile ensuite en produisant deux photons d'énergie 511 keV.

1.1.2 Détection des rayonnements γ

Spectrométrie

La spectrométrie gamma consiste en l'identification et éventuellement la quantification d'une ou plusieurs sources radioactives émettrices de photons γ grâce à la mesure de l'énergie déposée par ceux-ci dans un dispositif composé d'un matériau de détection et d'une chaîne d'acquisition permettant de récupérer le signal correspondant à ces interactions (Figure 1.7). Les paramètres de ces détecteurs peuvent être très variables, passant de la nature du matériau utilisé à la géométrie selon l'application et la gamme d'énergie visée.

Voici le fonctionnement de cette technique pour un photon :

1. Un photon issu d'une désintégration γ est émis par une source et interagit dans le matériau de détection en déposant une certaine quantité d'énergie.
2. L'énergie déposée est convertie en signal électrique soit de manière directe (le photon génère un ou deux porteurs de charges selon les interactions physiques décrites précédemment) ou indirecte (le photon éjecte un électron d'un atome, ce dernier générant par la suite un certain nombre de porteurs de charges par ionisation de la matière en transférant son énergie cinétique au cours du mouvement) pour un photo-détecteur.

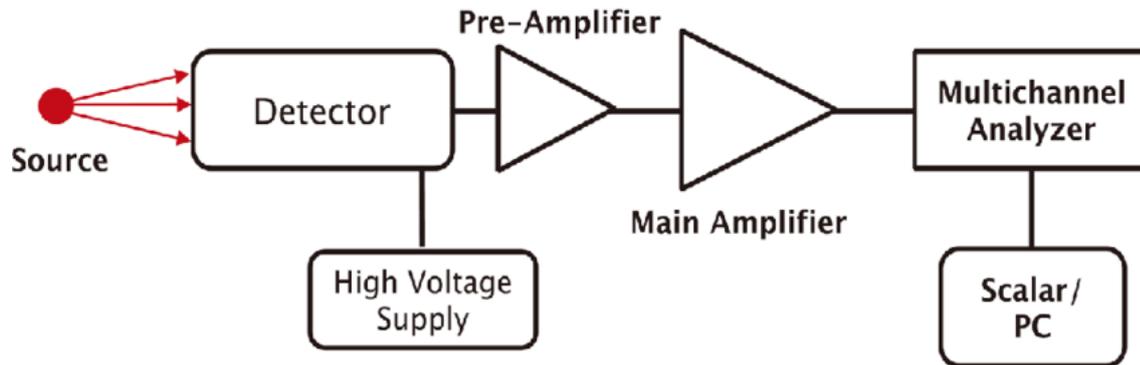


FIGURE 1.7 – Diagramme du fonctionnement général d'une chaîne de détection de rayonnement γ pour la spectrométrie [8].

3. Les électrodes récupèrent le signal électrique produit par les charges.
4. Un circuit de lecture permet de mesurer l'amplitude du signal électrique mesuré, l'amplifier et le filtrer pour améliorer le rapport signal sur bruit (SNR ou *signal to noise ratio* en anglais).
5. On trace l'histogramme des amplitudes en sortie qui permettent de tracer un spectre en énergie et l'étude de celui-ci permet l'identification des espèces présentes dans l'échantillon étudié (Figure 1.8).

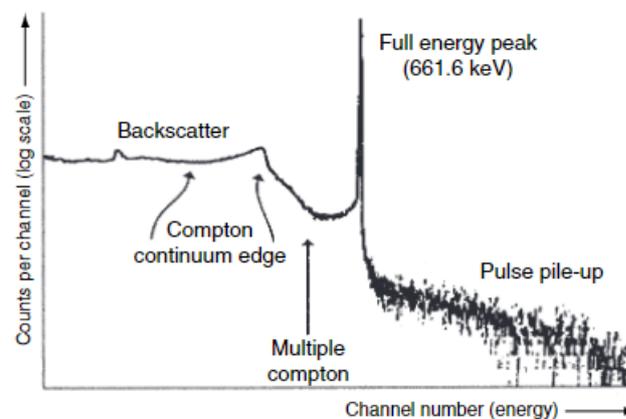


FIGURE 1.8 – Spectre du Césium 137 obtenu par spectrométrie gamma. [9].

Il y a plusieurs avantages à utiliser la spectrométrie. Tout d'abord il s'agit d'une technique d'identification non destructive, ainsi on peut analyser un échantillon sans le dégrader. C'est une méthode qui possède une sensibilité élevée, permettant de quantifier et de discriminer plus facilement les différents isotopes même présents en faible quantité. Il y a néanmoins certains facteurs dégradant la qualité de cette technique, tels que les interférences du fond radioactif ou la diffusion Compton affectant la résolution en énergie.

Localisation des interactions

Au delà de la simple identification de sources, les détecteurs de rayonnements gamma peuvent également permettre de localiser l'échantillon radioactif dans un environnement

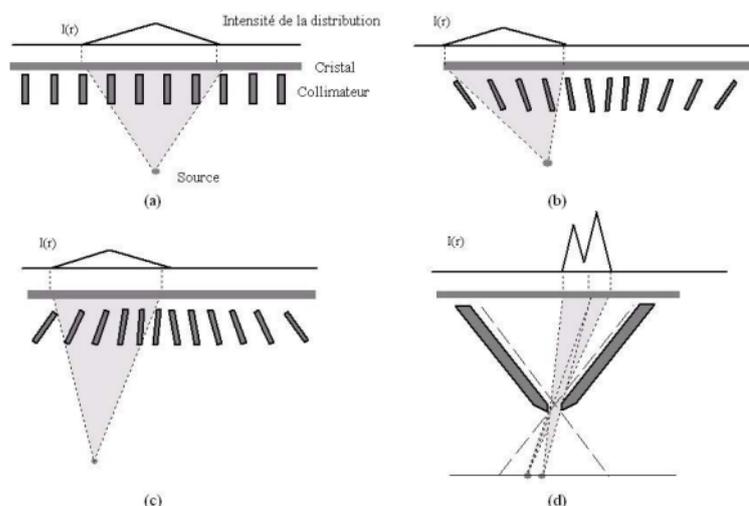


FIGURE 1.9 – Différents types de collimation : a) collimateur à trous parallèles. b) collimateur à trous convergents. c) collimateur à trous divergents. d) collimateur pinhole. [11].

(organisme pour le domaine médical, bagage dans le domaine de la sécurité, détection de points chauds en radioprotection). Les techniques utilisées pour la localisation ont évolué au fil du temps afin d'améliorer la résolution spatiale des détecteurs tout en conservant une résolution en énergie suffisamment précise.

Les détecteurs à base de scintillateurs ont fait partie des premiers détecteurs à être utilisés en localisation [10]. Le fonctionnement de ces derniers est présenté dans la partie suivante, de même que les autres types de détecteurs usuels. Aux énergies considérées, autour de 100 keV, on ne dispose pas de lentilles ou de miroirs assez efficaces pour permettre la formation d'image sur un détecteur. On utilise donc des collimateurs qui absorbent les rayons gamma n'ayant pas l'incidence désirée, ce qui permet de déterminer la direction du faisceau incident grâce à une géométrie spécifique (Figure 1.9).

Dans les années 1960 une invention a permis de révolutionner l'imagerie gamma en combinant notamment les technologies citées précédemment. Il s'agit de la caméra d'Anger, portant le nom de son inventeur et précurseur des gamma-caméras utilisées aujourd'hui pour réaliser de la scintigraphie. Cette caméra est composée d'un collimateur, d'un cristal scintillateur couplé à un guide de lumière et de photo-multiplicateurs permettant de convertir les photons de scintillation en signal électrique [12] (Figure 1.10).

À l'origine Anger a développé un système de réseau résistif afin d'effectuer la localisation en X et Y de l'interaction dans le détecteur. Chaque valeur en sortie de photo-multiplicateur est séparée en quatre partagée sur quatre résistances permettant de calculer les positions de X et Y selon un barycentre $X = \frac{X^+ - X^-}{X^+ + X^-}$ et $Y = \frac{Y^+ - Y^-}{Y^+ + Y^-}$ (Figure 1.11).

Certains dispositifs ont été conçus afin d'utiliser l'effet de diffusion Compton présenté dans la partie précédente. En effet la détection des photons γ diffusés permettent de déterminer l'angle de diffusion et de remonter à la localisation de la source grâce à la mesure du dépôt d'énergie. Généralement ce sont des imageurs à 2 couches, la première permettant la diffusion du photon incident par effet Compton ce qui produit le photon diffusé qui va aller déposer son énergie dans la seconde couche (Figure 1.12). Grâce à l'équation de l'effet Compton reliant les énergies déposées avec l'angle de diffusion et des algorithmes de localisation on peut remonter à la localisation de la source.

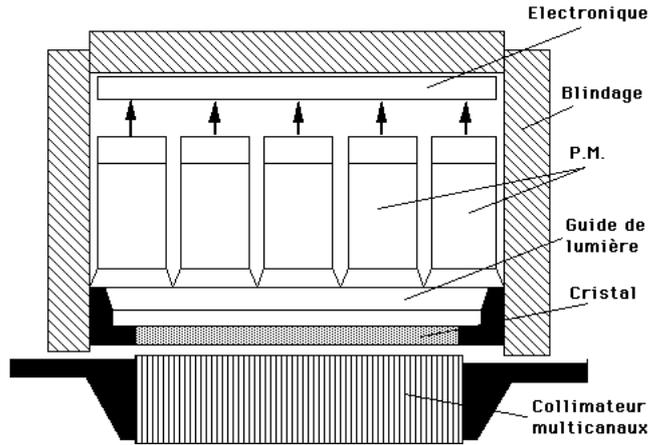


FIGURE 1.10 – Schéma de la structure d'une caméra d'Anger [13].

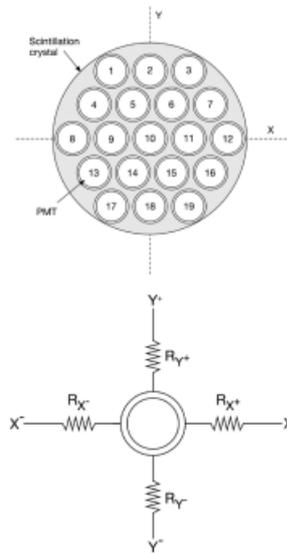


FIGURE 1.11 – Schéma d'une caméra d'Anger (haut), et du système de réseau resistif pour la localisation (bas) [14].

Avec un principe similaire à la caméra Compton, des détecteurs multi-couches ont été développés pour la localisation de sources de haute énergie. Dans ce cas on empile des couches de détecteur avec une haute résolution en énergie afin de reconstituer le parcours des particules et leurs différents dépôts d'énergie pour améliorer la localisation de la source (ce qu'on appelle *tracking* en anglais). Ce type de détecteur est notamment utilisé en physique des hautes énergies, pour des applications dans le domaine spatial [16], ou dans les accélérateurs de particules. L'analyse de donnée associée peut être particulièrement complexe du fait de la cascade d'interactions successives devant être considérées. Les algorithmes d'inférence Bayésienne peuvent ainsi être appliqués à l'imagerie gamma [17]. Ces méthodes utilisent l'analyse statistique ainsi que des modèles calculatoires basés sur la physique pour estimer la position de la source. Ces méthodes sont puissantes car elles permettent de combiner plusieurs modèles statistiques entre eux et ainsi tirer parti de différentes sources d'information et d'introduire des a priori évitant la sous-détermination

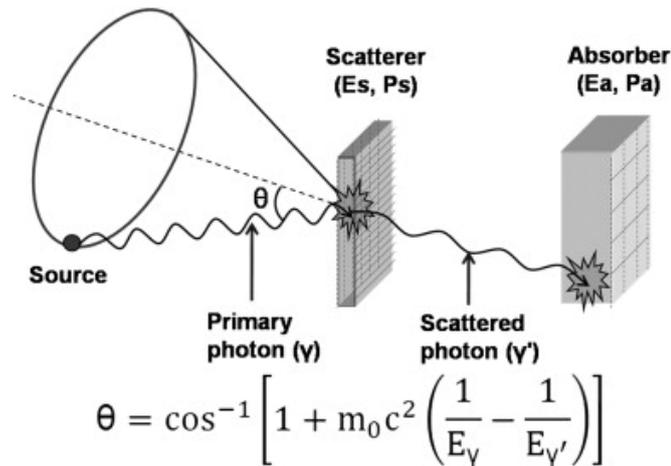


FIGURE 1.12 – Schéma du fonctionnement d'une caméra Compton à double couche [15].

de l'inférence.

1.1.3 Différents types de détecteurs

Détecteurs à scintillation

Les scintillateurs, ou détecteurs à scintillation, sont comme leur nom l'indique des dispositifs de détection utilisant le principe de scintillation de certains matériaux suite à l'interaction d'un rayonnement. Ces détecteurs sont généralement composés d'un matériau scintillateur où sont détectées les interactions, d'un photo-multiplicateur transformant la lumière de scintillation en signal électrique et l'amplifiant pour la collecter sur une anode qui transmet le signal et d'une électronique de traitement (Figure 1.13). Parmi les applications de ces détecteurs on trouve par exemple les détecteurs NaI pour l'imagerie SPECT, le LSO pour le PET ou encore le CsI pour la radiographie. Ces procédés d'imagerie seront expliqués plus en détail dans la suite.

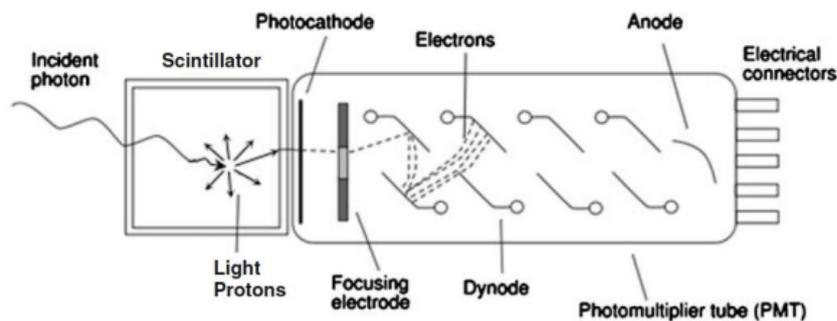


FIGURE 1.13 – Schéma d'un détecteur scintillateur [18]

Il existe 2 types de matériaux scintillateurs :

- Scintillateurs organiques : sous forme plastique, liquide ou cristalline. Leurs principaux avantages sont leur vitesse de réponse ainsi que leur capacité à détecter efficacement les neutrons. En revanche du fait de leur numéro atomique moyen leur efficacité de détection des photons est assez faible [19].

- Scintillateurs inorganiques : ce sont essentiellement des atomes du groupe alcalin dopés par une impureté (NaI(Tl) par exemple). Ils ont souvent un numéro atomique élevé et sont donc efficaces pour la détection de particules chargées sur de faibles épaisseurs et des photons. Ils sont également plus lumineux, en revanche leur prix est beaucoup plus important que les matériaux organiques [20].

Détecteurs gazeux

Les détecteurs gazeux mettent en jeu le processus d'ionisation de la matière, ici du gaz. Lorsqu'une particule ionisante comme un photon γ interagit dans le dispositif, elle va ioniser le gaz localement, éjectant des électrons des noyaux du milieu entraînant la création d'un certain nombre de paires électrons-ions proportionnellement à son énergie. Sous l'effet d'un champ électrique appliqué entre l'anode et la cathode du détecteur, les électrons et les ions vont migrer afin de créer un courant induit aboutissant à la mesure d'un signal électrique. Toutefois du fait de la densité faible du milieu dû à l'état gazeux, un étage d'amplification est nécessaire pour multiplier et amplifier les charges. Une solution est donc d'appliquer un second champ électrique d'intensité supérieure au premier. De ce fait les électrons se retrouvent accélérés et on obtient un phénomène d'avalanche dont la nature (proportionnelle ou non) dépend de l'intensité du champ électrique, grâce auquel le nombre de paires de porteurs de charges est multiplié de plusieurs milliers. Le déplacement de cette avalanche et la collecte des charges permet de remonter à la localisation de la position d'interaction de la particule [21].

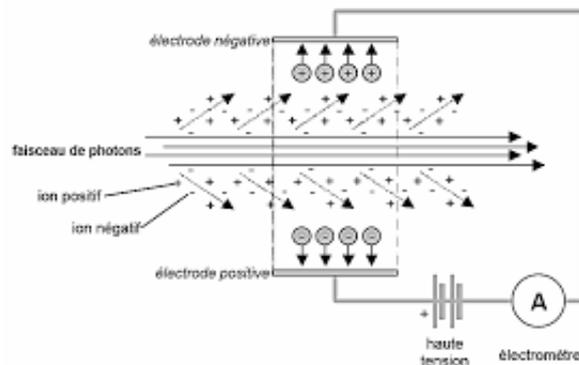


FIGURE 1.14 – Illustration d'un détecteur gazeux (Chambre d'ionisation) [22]

Détecteurs à semi-conducteurs

Un semi-conducteur est un matériau périodique cristallin présentant une variabilité sur le nombre de porteurs de charge libres. Les propriétés électriques des semi-conducteurs proviennent de leur structure électronique, celle-ci peut être expliquée par la théorie des bandes. La structure cristalline d'un semi-conducteur implique un recouvrement des orbitales atomiques des atomes qui la composent. Ceci crée un continuum d'états d'énergies accessibles aux électrons, on parle alors de bande d'énergie. La présence d'une bande interdite est liée à la nature cristalline du matériau. Ce dernier possède un potentiel périodique et si un électron est soumis à une équation avec un vecteur en phase avec ce potentiel,

on a une impossibilité pour cet électron de traverser la bande sans un apport d'énergie supérieur à sa largeur.

La bande interdite représente un continuum d'états d'énergies inaccessibles par les porteurs de charge. C'est l'excitation des électrons de la bande de valence qui permet de les faire rejoindre la bande de conduction et ainsi contribuer aux transport des charges à travers le matériau. Le principe de détection est le suivant : lors de l'interaction d'un photon γ dans le cristal semi-conducteur, ce dernier crée un photoélectron, qui est un électron arraché au noyau de l'atome subissant l'interaction à énergie cinétique élevée. Ce dernier perd son énergie en générant d'autres ionisations en cascade jusqu'à thermalisation. C'est de cette manière que sont créées les paires électron-trou en nombre proportionnel à l'énergie déposée. Puis leur dérive vers les électrodes donne un signal électrique.

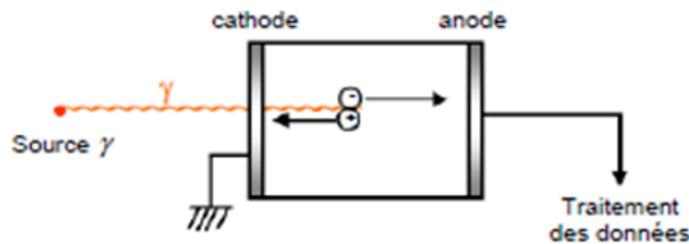


FIGURE 1.15 – Schéma de fonctionnement d'un détecteur à semi-conducteur [23].

Ces détecteurs présentent tout de même certaines limites. L'une est liée à la durée de vie des électrons et des trous, qui peuvent être piégés en raison des défauts du matériau, une autre aux fluctuations dans les pertes de collecte de charge dues aux impuretés, ce qui est en partie responsable du bruit et de la dégradation de la résolution en énergie. On remarque également une faible efficacité de détection pour les photons d'énergie élevée due à la faible épaisseur du matériau, on doit dans ce cas augmenter l'épaisseur du détecteur au détriment des résolutions en énergie et spatiale. Il faut donc trouver un compromis entre la plage d'énergie visée et les performances du détecteur.

1.1.4 Applications

Domaine médical

Un des secteurs où l'imagerie gamma est la plus présente est le milieu médical. En effet, l'utilisation de substances radioactives est commune pour le diagnostic et le traitement de certaines maladies. Il existe des modalités d'imagerie γ comme la TEP (tomographie par émission de positons ou PET en anglais), qui utilise l'émission simultanée de deux photons lors de l'annihilation d'un positon émis par les radio traceurs injectés au patient (comme le fluor 18) afin de les détecter en coïncidence sur une ligne pour localiser la source. Sa variante la TEMP (tomographie par émission monophotonique ou SPECT en anglais) possède un principe en utilisant des substances émettrices de photons et une collimateur. La sensibilité est plus faible mais on peut utiliser des substances émettant des photons dans une gamme d'énergie plus large. Ces examens sont réalisés en cancérologie, en imagerie cérébrale et en cardiologie majoritairement (Figure 1.16).

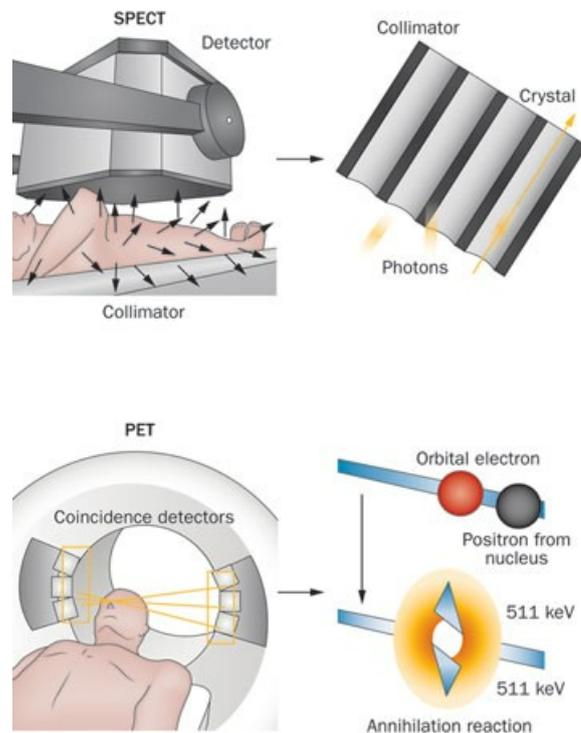


FIGURE 1.16 – Schéma de fonctionnement de l'imagerie SPECT (haut) et PET (bas) [24].

L'imagerie gamma peut également s'appliquer à la recherche afin d'étudier les substances radio-pharmaceutiques. Il est possible d'effectuer des tests sur des petits animaux par exemple afin de tester de nouvelles substances radiopharmaceutiques et des particules diagnostiques. Ces recherches peuvent parfois s'étendre à la recherche sur l'organisme humain, et aboutir dans certains cas à la découverte de substances théranostiques, qui peuvent à la fois servir pour le diagnostic et pour le traitement.

Sécurité et radioprotection

Dans la radioprotection il y a plusieurs situations où il est nécessaire d'identifier et de localiser une source radioactive.

- Contrôler et caractériser les déchets radioactifs, afin d'optimiser leur évacuation, transport et traitement.
- Surveiller les potentielles sources de radioactivité dans l'environnement : l'air, les sols ou un bâtiment industriel (comme une centrale nucléaire).

En ce qui concerne la sécurité la principale application est la vérification des transports et déchets ou la surveillance des foules.

1.2 Description des détecteurs CZT

1.2.1 Caractéristiques du CZT

Les détecteurs semi-conducteurs utilisés pour l'imagerie nucléaire doivent posséder certaines caractéristiques physiques importantes :

- Un numéro atomique Z et une densité élevés pour un pouvoir d'absorption des photons incidents efficace sur des épaisseurs de détecteur limitées (de l'ordre du centimètre).
- Un compromis sur la largeur de la bande interdite, assez large pour une résistivité importante à température ambiante mais limitée pour garder une énergie de création de paires électron-trou E_{paire} assez faible. Ceci permet de minimiser le bruit des fluctuations dues au courant d'obscurité.
- Des propriétés de transport de charges intéressantes (durée de vie τ et mobilité des porteurs μ) afin d'avoir un libre parcours moyen $\mu\tau$ le plus élevé possible.

Matériau	Z	Densité ($\text{g} \cdot \text{cm}^{-3}$)	Absorption à 140 keV (cm^{-1})	Résistivité ($\Omega \cdot \text{cm}$)	Bande interdite (eV)	E_{paire} (eV)	Produit $\mu\tau$ électron ($\text{cm}^2 \cdot \text{V}^{-1}$)
<i>Ge</i>	32	5,32	1,35	50	0,74	2,98	7.2×10^{-2}
<i>Si</i>	14	2,33	0,35	10^6	1,16	3,76	4.2×10^{-2}
<i>CdTe</i>	48/52	6,06	4	$> 10^9$	1,47	4,43	1.1×10^{-3}
<i>CZT</i>	48/30/52	6	3,84	$> 10^{10}$	1,5	4,64	3×10^{-3}
<i>HgI₂</i>	80/53	6,4	8,96	10^{13}	2,13	4,2	10^{-4}
<i>TlBr</i>	81/35	7,56	11,2	$> 10^{10}$	2,68	-	7.2×10^{-3}

TABLE 1.1 – Caractéristiques des principaux semiconducteurs utilisés pour la détection de photons X et γ [25].

D'après le tableau précédent le CZT présente des caractéristiques intéressantes pour une utilisation en imagerie, notamment un numéro atomique Z et une densité élevés, une bande interdite assez fine, des propriétés de transport correctes mais aussi un coût raisonnable et une disponibilité sur le marché.

Les cristaux CZT présentent néanmoins certains inconvénients, comme une croissance du cristal moins maîtrisée que pour le silicium ou le germanium, entraînant la formation de défauts dans la maille cristalline pouvant impacter la création de signal électrique via des inhomogénéités dans le champ électrique.

1.2.2 Physique des semi-conducteurs

Les équations théoriques de la physique liées au fonctionnement des semi-conducteurs permettent de faire corrélérer les résultats expérimentaux avec certaines hypothèses. Après validation on peut transcrire ces équations dans les outils de simulation afin de modéliser les détecteurs et simuler des interactions dans celui-ci.

Transport des porteurs

Les équations de l'électromagnétisme permettent, sous réserves d'hypothèses simplificatrices, de modéliser le trajet des charges et de calculer les charges induites sur les électrodes et le long des lignes de champ. On se place dans le modèle de dérive-diffusion,

qui est une variante du modèle hydrodynamique auquel on ajoute des hypothèses supplémentaires. On suppose en plus que la température ne varie pas, ainsi on peut négliger l'équation de conservation d'énergie [26].

On peut modéliser dans un premier temps le champ électrique à l'intérieur du volume du détecteur grâce à la conservation de la charge et la loi d'Ohm :

$$\vec{\nabla} \cdot \vec{j} + \frac{\partial \rho}{\partial t} = 0 \quad (1.5)$$

Avec $\vec{j} = \sigma \vec{E}$ la densité de courant, ρ la densité volumique de charge, σ la conductivité et \vec{E} le champ électrique. On a également l'équation de Maxwell traduisant la loi de Faraday :

$$\vec{\nabla} \wedge \vec{E} = 0 \quad (1.6)$$

On a donc un champ électrique \vec{E} dérivant d'un potentiel scalaire tel que :

$$\vec{E} = -\vec{\nabla} \phi \quad (1.7)$$

Lorsque le semi-conducteur est à l'équilibre, on peut alors simplifier la loi d'Ohm pour obtenir l'équation de Poisson :

$$\vec{\nabla} (\sigma \cdot \vec{\nabla} \phi) = 0 \quad (1.8)$$

Maintenant que l'on peut modéliser le potentiel électrique dans le détecteur, on peut étudier le comportement des porteurs. Ces derniers (électrons et trous) sont soumis à deux phénomènes principaux : la dérive et la diffusion. La dérive est due à la présence d'une différence de potentiel entre les deux électrodes. Soit μ la mobilité des porteurs de charge, la vitesse due à la dérive est

$$\vec{v} = -\mu \vec{\nabla} \phi \quad (1.9)$$

Pour les électrons cette vitesse est orientée à l'opposé du champ électrique, et dans le même sens pour les trous. Le phénomène de diffusion est quant à lui causé par la présence d'un gradient de concentration des porteurs dans le cristal de détection suivant la loi de Fick :

$$\vec{j}_{n,p} = -D_{n,p} \vec{\nabla} (n, p). \quad (1.10)$$

Ce gradient entraîne un mouvement des porteurs afin d'homogénéiser la répartition de la charge. La relation d'Einstein [27] permet de calculer les coefficients de diffusion pour les électrons et les trous D_n et D_p

$$D_n = \mu_n \frac{kT}{q} \text{ et } D_p = \mu_p \frac{kT}{q} \quad (1.11)$$

Avec μ la mobilité du porteur, k la constante de Boltzmann, T la température absolue et q la valeur absolue de la charge élémentaire des porteurs. En sommant les contributions de ces deux phénomènes on obtient les équations des densités de courant des porteurs

$$\begin{cases} \vec{J}_n(\vec{r}, t) = -q\mu_n n(\vec{r}, t) \vec{\nabla} \phi - qD_n \vec{\nabla} n(\vec{r}, t) \\ \vec{J}_p(\vec{r}, t) = q\mu_p p(\vec{r}, t) \vec{\nabla} \phi - qD_p \vec{\nabla} p(\vec{r}, t) \end{cases} \quad (1.12)$$

Avec n et p la densité respective des électrons et des trous, \vec{r} le vecteur position et t le temps. Ainsi on a la densité de courant totale :

$$\vec{J}(\vec{r}, t) = \vec{J}_n(\vec{r}, t) + \vec{J}_p(\vec{r}, t) \quad (1.13)$$

Il y a cependant un phénomène non négligeable à considérer pour le transport des porteurs : le piégeage-dépiégeage.

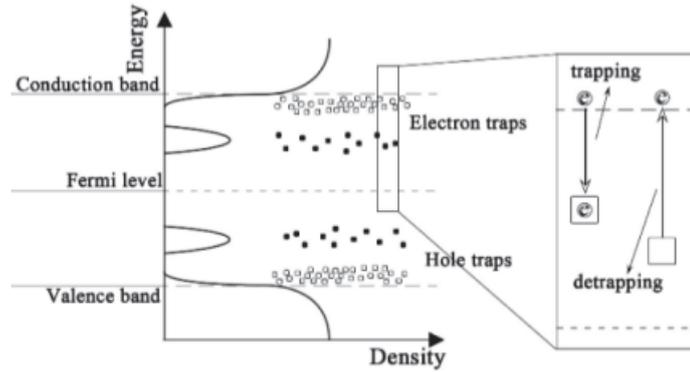


FIGURE 1.17 – Illustrations des phénomènes de piégeage/dépiégeage [28].

Ces processus sont possibles à cause des défauts présents dans le cristal semi-conducteur et dont une description plus détaillée sera faite dans une partie ultérieure. Ces défauts font apparaître des états d'énergie intermédiaires dans la bande interdite. Quand un porteur se retrouve dans un de ces nouveaux états d'énergie, il peut soit se recombiner soit revenir dans son état initial. On note les équations donnant l'évolution temporelle des densités de porteurs piégés

$$\frac{\partial n_t(\vec{r}, t)}{\partial t} = \frac{n(\vec{r}, t)}{\tau_n^c} - \frac{n_t(\vec{r}, t)}{\tau_n^d} \quad \text{et} \quad \frac{\partial p_t(\vec{r}, t)}{\partial t} = \frac{p(\vec{r}, t)}{\tau_p^c} - \frac{p_t(\vec{r}, t)}{\tau_p^d} \quad (1.14)$$

Avec $n_t(\vec{r}, t)$ et $p_t(\vec{r}, t)$ les densités des porteurs piégés à distinguer des densités de porteurs libres. τ^c et τ^d sont respectivement les constantes de temps de piégeage et de dépiégeage.

En assemblant toutes les équations régissant le fonctionnement de chaque procédé physique mis en jeu, on obtient les équations de conservation de la charge pour les deux types de porteurs

$$\begin{cases} \frac{\partial n(\vec{r}, t)}{\partial t} = G_n + \frac{1}{q} \vec{\nabla} \cdot \vec{J}_n - \frac{n(\vec{r}, t)}{\tau_n^c} + \frac{n_t(\vec{r}, t)}{\tau_n^d} \\ \frac{\partial p(\vec{r}, t)}{\partial t} = G_p + \frac{1}{q} \vec{\nabla} \cdot \vec{J}_p - \frac{p(\vec{r}, t)}{\tau_p^c} + \frac{p_t(\vec{r}, t)}{\tau_p^d} \end{cases} \quad (1.15)$$

Où G_n et G_p représentent les termes sources, c'est-à-dire la création des porteurs libres par le dépôt d'énergie d'un photon.

Induction de charges sur les électrodes

Le transport des porteurs de charge est la source de création du signal électrique qui sera collecté sur les électrodes. Dans un semi-conducteur comme le CZT, les différences de propriétés de transport entre les électrons et les trous font que l'exploitation du signal électron est beaucoup plus intéressante du point de vue des performances du détecteur, les

libres parcours moyen de ces derniers étant extrêmement plus faibles. On développe ainsi des structures géométriques, comme les détecteurs à pixel qui vont nous permettre de réaliser la compensation du signal trou. On utilise alors le théorème de Ramo qui permet de calculer le signal induit dans un détecteur à géométrie multiélectrodes [29].

Le théorème de Ramo définit le potentiel de pondération qui sert à calculer le signal de collecte induit par une charge en mouvement [30]. Ce potentiel de pondération n'a pas d'unité physique et représente concrètement l'influence qu'a une électrode sur l'induction de signal en fonction de la distance à laquelle se trouve la particule chargée en mouvement (Figure 1.18). Ce potentiel de pondération w est calculé en résolvant l'équation de conservation du courant de déplacement $\epsilon \frac{\partial \vec{E}}{\partial t}$. Pour l'électrode k on a

$$\vec{\nabla} \cdot (\epsilon \vec{\nabla} w_k) = 0 \quad (1.16)$$

Le courant instantané produit sur une électrode k par une charge q se déplaçant à une vitesse $\vec{v} = \mu \vec{E}$ est égal à

$$i_k = q \vec{v} \cdot \vec{\nabla} w_k = -q \mu \vec{\nabla} \phi \cdot \vec{\nabla} w_k \quad (1.17)$$

Avec w_k le potentiel de pondération sur l'électrode k .

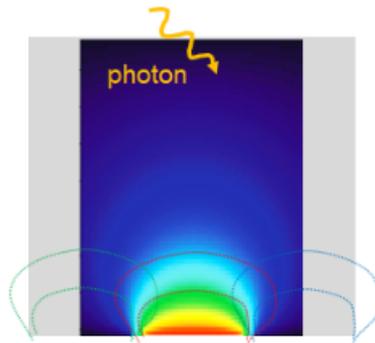


FIGURE 1.18 – Potentiel de pondération sur une anode d'une simulation d'un détecteur pixellisé

Pour un semi-conducteur de forme parallélépipédique à géométrie d'électrodes planes avec une conductivité uniforme, la distribution des deux potentiels est identique ; il s'agit dans les deux cas d'un profil linéaire. La connaissance de la répartition de ces deux potentiels permet d'évaluer la valeur des signaux collectés. La valeur du courant instantané est un produit scalaire des gradients des deux potentiels (appliqué et pondération). De plus son intégrale sur le temps sera proportionnelle à la distance parcourue. Ceci montre que le courant mesuré est fonction de la position initiale des porteurs.

Or l'objectif d'un système de détection spectrométrique est d'assurer un facteur de proportionnalité entre l'information qu'il détecte et le signal recueilli. C'est pour cela que de nombreuses techniques ont été mises en œuvre afin de rendre le signal de courant mesuré indépendant du lieu de création de la charge dans le semi-conducteur (grille de Frisch, ...) (Figure 1.19).

Dans le cas général, les conditions favorables correspondent à un gradient de potentiel électrique important et une distribution du potentiel de pondération avec un fort gradient

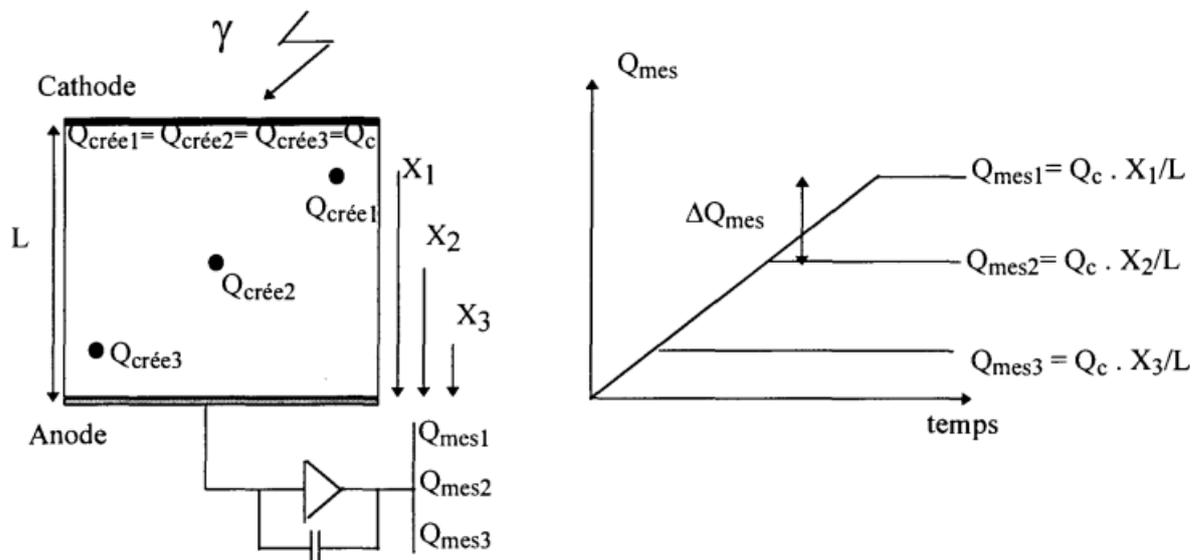


FIGURE 1.19 – Variation de la charge mesurée en fonction de la position d'interaction dans un détecteur plan [29]

du côté de l'électrode collectante (l'anode), pour que le courant induit reste indépendant de la position initiale de la charge.

1.3 Apparition et caractérisation des défauts cristallins du CZT

1.3.1 Structure cristalline du CZT

Le CdTe et le CZT sont deux matériaux à la structure extrêmement semblable. La seule variation correspond à la proportion d'atomes de Zinc qui remplacent de manière aléatoire ceux de Cadmium selon la teneur en Zinc du CZT. La structure cristalline est de type zinc blende, pouvant être décrite comme deux sous-réseaux de type cubique à face centrée décalé l'un de l'autre d'un quart de cellule unitaire de maille [31]. Chacun des sous-réseaux est peuplé par des atomes de Cadmium (ou Zinc selon la teneur) ou de Tellure selon la structure ci dessous (Figure 1.20).

Ces matériaux sont créés artificiellement par croissance pour leur utilisation industrielle. Il existe plusieurs méthodes de croissance des cristaux de CZT avec leurs limites, les défauts apparaissent en parallèle du cristal et vont modifier la structure de la maille cristalline ce qui impacte la qualité du cristal et donc du détecteur. Ceci limite la production de détecteur de surface supérieure à quelques centimètres carrés ce qui représente un des inconvénients du CZT.

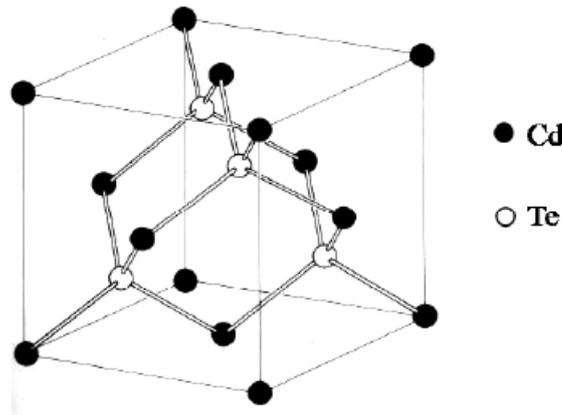


FIGURE 1.20 – Structure de la maille cristalline du CdTe / CZT [31].

1.3.2 Méthodes de croissance usuelles

Méthode Bridgman Verticale (VBM)

La méthode Bridgman est une des méthodes les plus anciennes utilisée pour la croissance de cristaux. Le système peut être orienté de manière verticale ou de manière horizontale, avec un principe relativement similaire.

Cette méthode est basée sur une solidification directionnelle en translatant un matériau fondu vers la partie froide du four utilisé en suivant un gradient de température. La présence d'un échantillon de cristal au bout du creuset assure la croissance d'un cristal monolithique selon une orientation spécifique.

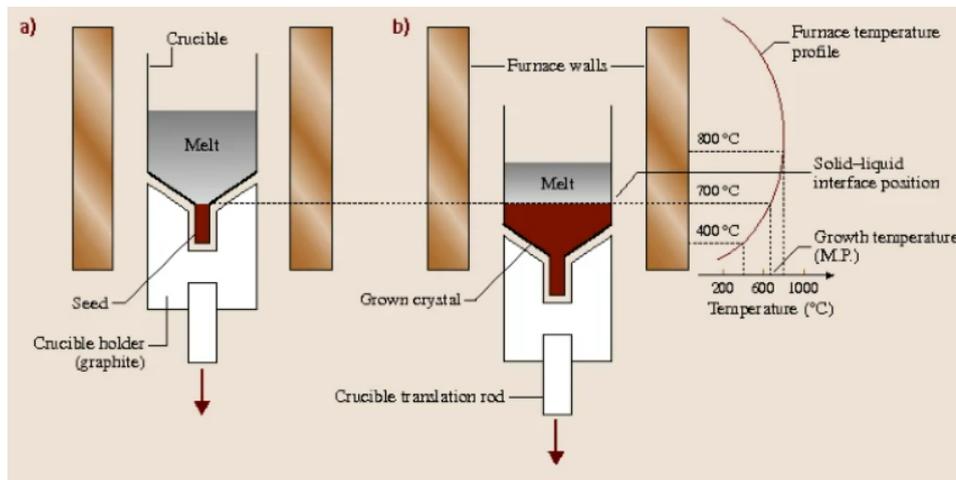


FIGURE 1.21 – Schéma de fonctionnement de la méthode Bridgman verticale : a) A l'état initial et b) avec un cristal partiellement formé [32].

Initialement on place le creuset avec le matériau polycristallin sous forme de poudre et le germe dans la chambre de croissance. On fait le vide dans cette chambre et on la remplit d'un gaz inerte. On fait alors monter la température dans le four jusqu'à fusion du matériau polycristallin et on maîtrise cette dernière précisément grâce à un mécanisme de contrôle. On mélange la zone en fusion de manière mécanique ou convective

pour homogénéiser ce dernier, et on peut alors démarrer la croissance en refroidissant le matériau fondu sur l'extrémité de l'échantillon, par translation du creuset vers la zone plus froide du four (Figure 1.21). On peut contrôler la vitesse de translation afin d'optimiser la qualité du cristal [32].

Méthode de croissance par fusion de zone (THM)

La méthode THM a été mise en place pour obtenir des meilleurs résultats que les techniques « classiques », que ce soit au niveau de la facilité de la mise en place mais également de la qualité du cristal obtenu [33].

Dans un premier temps on place un germe d'un précédent cristal produit dans une ampoule scellée sous vide, avec une solution enrichie en tellure et un polycristal de CZT. On place cette ampoule dans un four annulaire et on applique le profil thermique présenté sur la Figure 1.22. De cette manière on fait fondre la solution enrichie en tellure, et on crée une interface à la fois avec le germe mais aussi avec le polycristal de CZT. On fait ensuite descendre très lentement l'ampoule, permettant à la fois de dissoudre une partie du poly cristal et de le recrystalliser sur le germe, tout en faisant migrer les impuretés avec la zone fondue.

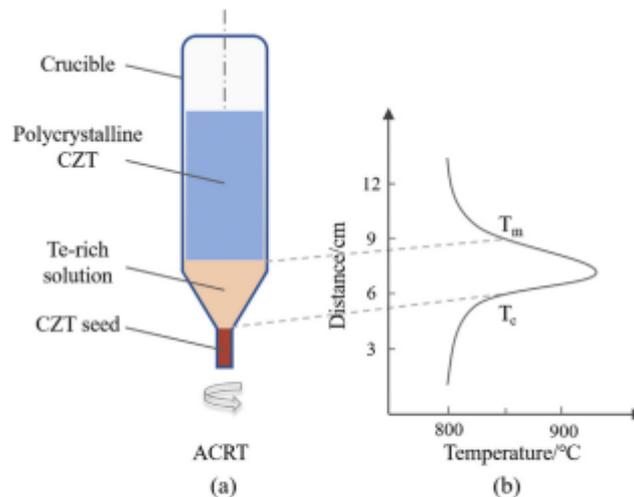


Fig. 1. The sketch of THM-ACRT method (a) the ampoule, (b) thermal field, T_c is the temperature at growth interface, T_m is the temperature at dissolved interface.

FIGURE 1.22 – Schéma de fonctionnement de la méthode THM [34] ;

1.3.3 Caractérisation du détecteur

L'observation des défauts cristallins ainsi que leur caractérisation est une étape clé dans la compréhension de leurs effets sur le cristal. Il existe plusieurs méthodes afin dans un premier temps de les mettre en évidence, puis pour caractériser les propriétés des porteurs de charge en fonction de leur environnement dans la maille.

On recense plusieurs types de méthodes afin de révéler et caractériser les défauts dans un cristal CZT :

- Les méthodes chimiques : le principe est d'attaquer chimiquement la surface du cristal afin de faire apparaître les défauts macroscopiques. Par exemple pour les dislocations, la gravure chimique est plus rapide sur les atomes de la dislocation due à l'énergie de contrainte autour de celle-ci. On obtient alors des motifs, que l'on appelle « *etch pits* », permettant de repérer les dislocations et leur orientation sur la surface du cristal (Figure 1.23). En fonction du cristal étudié les solutions les plus efficaces peuvent être différentes, car les défauts présents sont de natures différentes et sont plus ou moins visibles en fonction des traitements chimiques.
- Microscopie électronique à balayage (MEB) : Cette technique permet de cartographier la surface d'un échantillon en utilisant un faisceau d'électrons soumis à une tension de quelques kilovolts. Ces électrons interagissent avec les atomes de surface et peuvent ioniser ces derniers lorsque l'énergie est assez importante. On a alors émission d'électrons secondaires dont l'analyse permet d'obtenir une image de la surface du cristal, et ainsi observer les éventuels défauts.

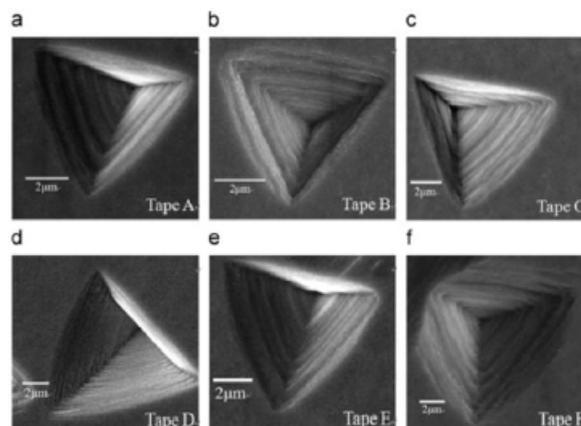


FIGURE 1.23 – Image MEB de six *etchpits* révélés sur une face tellure [35]

- Topographie par faisceau blanc en transmission : c'est une méthode de caractérisation de la structure dans le volume du cristal. On irradie l'échantillon avec un faisceau blanc de rayons X. Ce faisceau polychromatique utilise des rayons de longueur d'onde dont l'ordre de grandeur correspond à celui des distances inter-atomiques. Le faisceau est alors diffracté dans une certaine direction dans le cristal selon la loi de Bragg :

$$2d_{hkl} \sin \theta = \lambda \quad (1.18)$$

avec hkl l'indice du plan diffractant dans la direction θ , d_{hkl} la distance entre 2 plans diffractants et λ la longueur d'onde du faisceau incident. De cette manière chaque longueur d'onde diffracte dans la direction des plans correspondant et on voit apparaître des tâches de diffraction. Les différences locales d'orientation et de taille de maille provoquent des différences d'intensité des tâches et on peut ainsi identifier les défauts (Figure 1.24).

On va s'intéresser à différentes méthodes de caractérisation utilisées afin d'observer différentes caractéristiques des échantillons étudiés.

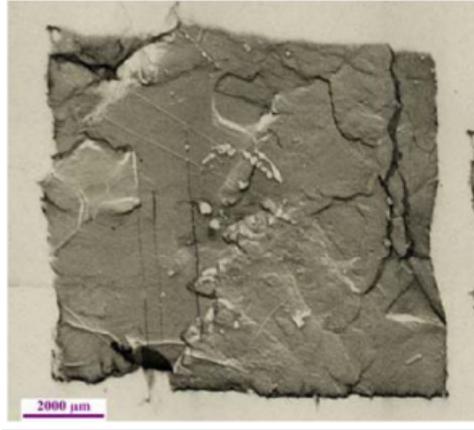


FIGURE 1.24 – Topographie aux énergies entre 4 et 50 keV pour un échantillon de CZT de 2mm d'épaisseur [36].

- La mesure IBIC (Ion Beam Induced Current) consiste à appliquer un faisceau d'ions dont l'ordre de grandeur de l'énergie est le MeV. Cette technique permet de scanner l'échantillon pour mesurer et cartographier les propriétés des matériaux semi-conducteurs [37]. Les avantages de cette méthode sont notamment le ratio signal sur bruit, une collimation facile et la localisation en surface de l'échantillon. L'utilisation d'ions énergétiques fait que toute l'énergie se transmet aux électrons de l'échantillon après de nombreuses interactions. Le dépôt de cette énergie permet dans le cas d'un semi-conducteur de produire un grand nombre de paires électron-trou. Les porteurs de charge, soumis au champ électrique du détecteur vont migrer vers les électrodes. Ainsi chaque pulsation émise permet de générer un histogramme d'amplitude. Dans le cas d'une irradiation mono énergétique, les coups sont distribués autour d'un pic central, dont la position est proportionnelle à la charge mesurée par l'électrode. Ainsi en calculant le rapport entre cette charge et la charge générée par ionisation on détermine l'efficacité de collecte de charge induite, dont le calcul sera détaillé par la suite. On peut également citer les méthodes EBIC (Electron Beam Induced Current) et XBIC (X-ray Beam Induced Current), dont le principe est similaire avec des rayons de nature différente, ce qui implique des distributions en énergie et en profondeur différentes.
- L'analyse par ToF-SIMS (Time-of-Flight Secondary Ion Mass Spectrometry) consiste à bombarder l'échantillon par un faisceau d'ions d'une énergie de quelques keV. Ces ions vont produire des ions secondaires suite à leur interaction avec les atomes de surface du cristal. Ces derniers sont accélérés dans un analyseur à temps de vol avec une certaine énergie cinétique. Ainsi grâce à la relation $E_c = 1/2mv^2$ la durée de transport des ions dans cet analyseur est proportionnelle à la racine carrée de la masse [38]. L'utilisation à la fois du balayage par faisceau et de l'analyseur de temps de vol permet de cartographier la surface de l'échantillon (Figure 1.25).
- On peut citer les méthodes de caractérisation en profondeur comme SCLC (Space Charge Limited Currents), PICTS (Photon Induced Current Transient Spectroscopy), DLTS (Deep-Level Transient Spectroscopy) ou la TSC (Thermally Stimulated Current Spectroscopy). Ces méthodes permettent d'étudier les échantillons en profondeur, en cartographiant les défauts ou en étudiant les niveaux d'énergie

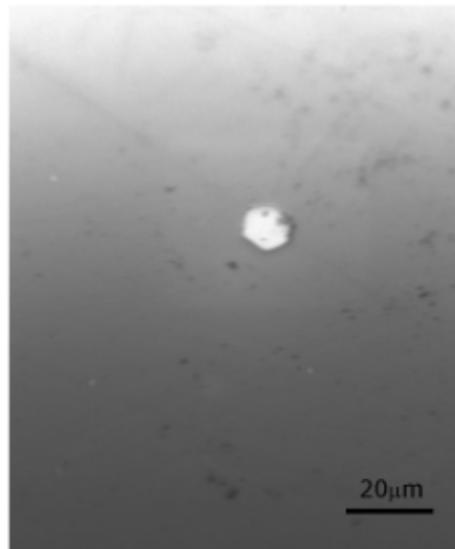


FIGURE 1.25 – Représentation d'une inclusion Te par analyse ToF-SIMS [39].

piège [40], [41], [42].

- Il existe également des méthodes utilisant les infrarouges, telles que FT-IR (Fourier Transform Infrared Spectroscopy), PAS (Photoacoustic Spectroscopy), ou la microscopie infrarouge. Ces méthodes utilisent les propriétés d'absorption et de réflexion du matériau utilisé afin de déterminer différentes caractéristiques [43].

1.3.4 Observation des principaux types de défauts

Comme expliqué précédemment, le CZT peut voir se développer lors de sa croissance des défauts dans sa maille cristalline qui sont de différentes natures. Chaque type de défaut influe sur le transport des charges et le champ électrique du détecteur rendant la mesure du signal fluctuante et non uniforme dans le volume du dispositif. La présence de défauts peut par exemple entraîner la création de nouveaux niveaux d'énergie hors de la bande interdite du CZT où les charges peuvent être piégées. Le piégeage étant plus important la collecte est dégradée. Dans sa thèse, Camille Buis réalise une étude de ces défauts sur des cristaux CdTe, dont la structure est extrêmement similaire au CZT, décrivant leur structure à l'échelle atomique ainsi que leur caractérisation de leurs effets sur les performances macroscopiques du détecteur [44].

Défauts Tellure

Les défauts tellure (principalement les inclusions et les précipités) peuvent être présents en concentration élevée, mais néanmoins uniformément distribués dans un cristal de CZT. La dimension d'une inclusion varie de quelques micromètres à quelques centaines de micromètres et celle d'un précipité est de quelques nanomètres (Figure 1.26). Ils peuvent piéger une fraction des charges du nuage d'électrons de dérive vers l'anode, provoquant ainsi une dégradation de la résolution en énergie et de l'efficacité de détection. Ces défauts proviennent de l'accumulation de Tellure dans le cristal [45].

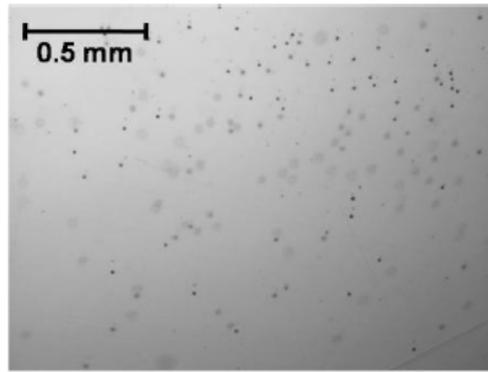


FIGURE 1.26 – Répartition des inclusions tellure en transmission infrarouge dans une zone d'un échantillon de CZT [46].

Plusieurs travaux ont déjà été réalisés pour observer l'impact de ce type de défaut sur le signal de sortie du détecteur. Dans une étude, Bolotnikov et son équipe ont créé un modèle de diffusion du nuage électronique lors d'une interaction d'un photon γ dans un détecteur CZT et y ont inséré des inhomogénéités géométriques correspondant à des défauts tellure afin d'observer leur impact sur le tracé du spectre en énergie (Figure 1.27) [47].

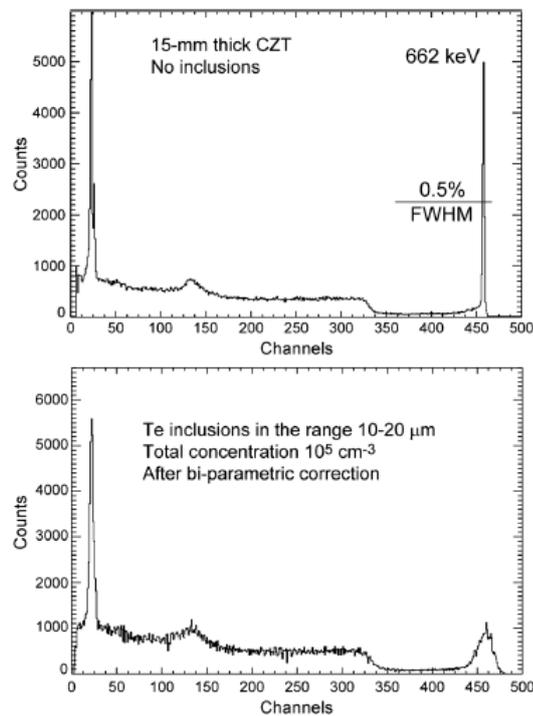


FIGURE 1.27 – Tracé des spectres en énergie pour un détecteur idéal (en haut) et avec des défauts tellure (en bas) [47].

Ce que l'on peut observer sur ces spectres, c'est la non uniformité de collecte avec la présence de défauts, on a une amplitude de pic réduite ainsi qu'un élargissement de celui-ci prouvant que ces défauts Tellure piègent les porteurs de plus haute énergie. De plus la forme des courbes dépend à la fois de la taille et de la densité des défauts considérés, ce

qui ne fait pas de doute quant à leur impact.

Cette même équipe a ensuite étudié les effets liés aux défauts étendus dans le CZT, notamment l'accumulation de défauts Te dans le cristal. Ils ont observé le fait que l'accumulation locale de défauts Te entraîne la formation de défauts en forme d'étoile, dont la présence réduit l'efficacité du système et dégrade les performances du système [48]. Ils concluent en évoquant le fait que ce type de défaut étendu pourrait être responsable du piégeage des trous lors de leur dérive après interaction 1.28.

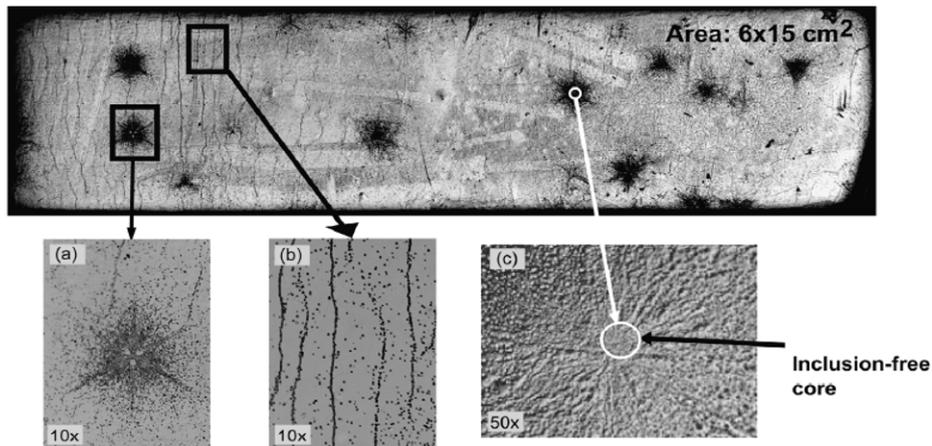


FIGURE 1.28 – Images IR de surfaces de CZT gravées chimiquement. On voit apparaître des lignes de dislocation ainsi que plusieurs défauts en forme d'étoile.

Sous-joints de grains

Les sous-joints de grain, ou mur de dislocations, correspondent à l'arrangement de dislocations organisées de manière parallèle sur un même plan. Ce sont donc des défauts à deux dimensions. La dislocation est un défaut linéaire composé d'un cœur, et sur quelques mailles aux alentours la configuration des liaisons est très différente par rapport à la maille du cristal sans défaut. Les effets des dislocations ont déjà fait l'effet de nombreux travaux de recherche, qui ont permis de mettre en avant les effets de ces défauts sur les détecteurs. Parmi ces derniers l'augmentation de courant de fuite du détecteur [49] ou la diminution de la durée de vie des porteurs de charges [50].

L'étude plus poussée des sous-joints de grains et un peu plus récente, elle fait l'objet d'un article dans lequel il est question de l'observation et de la caractérisation de ce type de défauts notamment dans les cristaux CZT. Différents types de ces défauts sont observés, ce qui aboutit sur un classement en deux catégories principales : des défauts « simples », qui traverse le cristal selon une seule ligne, et les défauts en « réseaux », avec des motifs et des lignes plus complexes [51] (Figure 1.29).

Il apparaît que ces deux variétés de sous-joints de grains impactent différemment le signal mesuré, chose que l'on peut observer sur les spectres en énergie. Les défauts « simples » auront tendance à simplement élargir le pic d'énergie, tandis que les défauts en « réseaux », en plus de l'élargissement du pic vont faire apparaître une augmentation de la traînée, montrant un piégeage plus important des porteurs de charge (Figure 1.30).

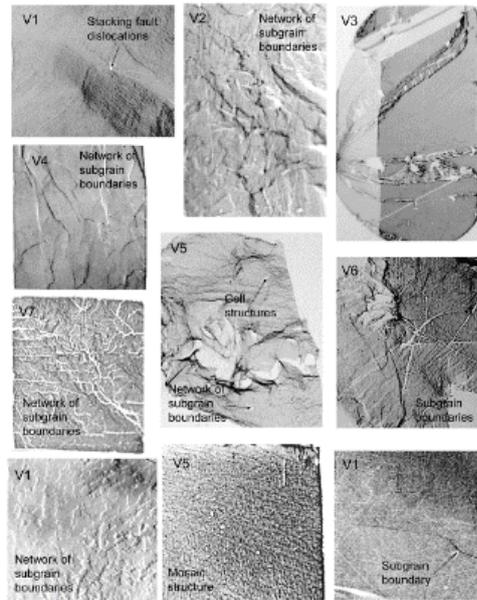


FIGURE 1.29 – Illustration de certains exemples de sous joints de grain observés.

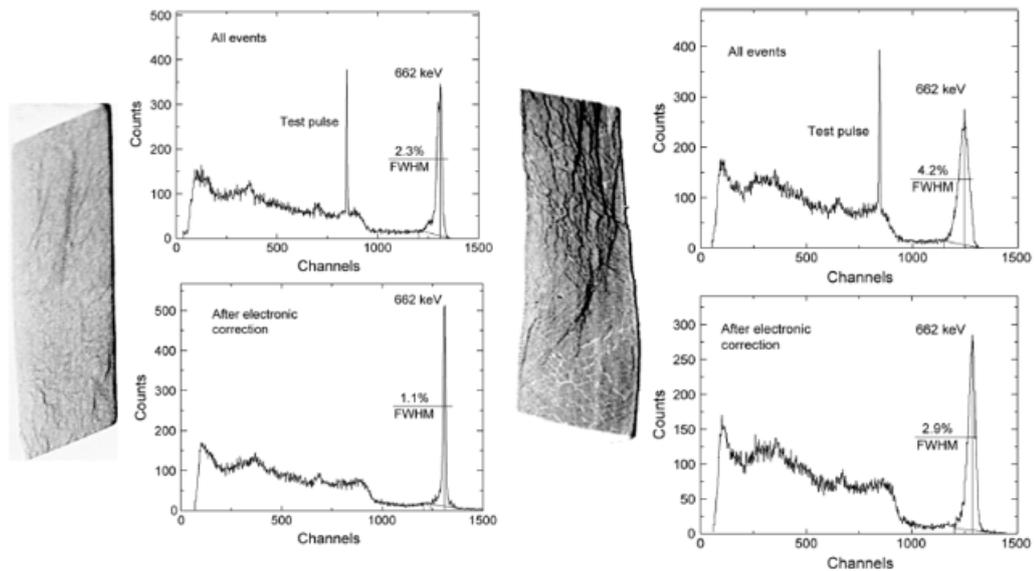


FIGURE 1.30 – Spectres en énergie des signaux impactés par un sous-joint "simple" (gauche) et en "réseau" (droite) avant et après correction.

Défauts profonds

Une autre partie de l'étude des défauts concerne les défauts profonds responsables du piégeage des porteurs. Ils représentent un aspect plus chimique que structurel contrairement aux défauts exposés précédemment. En effet leur apparition proviendrait de réactions entre les différents éléments atomiques du cristal entraînant la formation d'ions ayant un impact sur la mesure du signal. Ces défauts dégradent les performances des détecteurs et une étude plus poussée permettrait d'en connaître plus précisément les origines

pour s'orienter sur des axes d'amélioration concernant les technologies de croissance des cristaux.

Des éléments de réponse sont apportées d'une part par Bale [52], et d'autre par Li et son équipe [53], où l'objectif est de déterminer au niveau atomique quelle sont les types de défauts les moins handicapants pour la détection. Ils utilisent pour cela un détecteur CZT à seize voies et réalisent différents tests sur celui-ci (Figure 1.31). Dans un premier temps la performance en comptage de chaque voie est étudiée. On observe une non uniformité de ce paramètre sur chacune des voies, avec un taux qui diminue avec le numéro du pixel, indiquant une différence de présence des défauts dans différentes régions du détecteur.

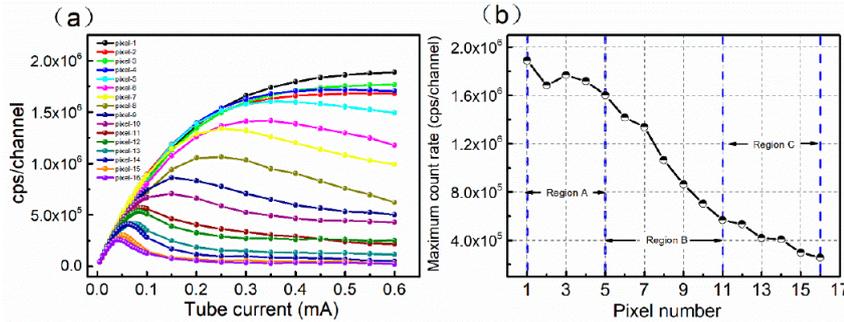


FIGURE 1.31 – (a) Taux de comptage du détecteur CZT à seize pixels ; (b) Courbe des taux de comptage maximum des seize pixels. Les régions A,B et C correspondent la région à haut taux de comptage, la région de performance moyenne et et la région de performance faible [53].

Pour déterminer la raison de ces disparités de comptage différentes mesures sont réalisées sur ce détecteur. Dans un premier temps un spectre de photoluminescence est tracé, concluant sur le fait que ces inhomogénéités de comptage ne proviennent pas de la fluctuation de concentration en zinc, mais de la distribution non-uniforme des défauts dans ce détecteur. Une microscopie de transmission IR est également réalisée et la mesure d'une densité uniforme de dislocations permet de dire que ces dernières n'influent pas sur les mesures de comptage.

Finalement des mesures TSC (Thermally Stimulated Current) sont effectuées pour obtenir des informations sur les pièges dans les seize régions du détecteur. Cette technique permet de déterminer les niveaux d'énergie dans les semi-conducteurs par exemple. Les niveaux d'énergie sont complétés par une stimulation optique ou électrique à une température assez faible, ainsi les porteurs de charge sont émis avec une température plus importante. Une courbe du courant d'émission par la température peut être tracée, et on récupère des informations sur les niveaux d'énergie dans le matériau [54]. Dans notre situation ce type de défauts a assez peu d'impact car l'étude ne porte pas sur le fonctionnement de la dynamique du détecteur. En effet lorsque le détecteur est utilisé et mis sous tension la charge d'espace locale est faible minimisant l'influence des défauts profonds.

Pour synthétiser nous allons dans la suite modéliser les défauts afin de les implémenter dans notre simulation en trois dimensions. Les défauts de type inclusions tellure, de petite taille, peuvent être définis comme des charges ponctuelles distribuées dans le volume du détecteur et qui vont agir sur la physique des porteurs de charge en modifiant localement le champ électrique. Les défauts distribués en surface tels que les sous-joints de grains modifient le champ électrique de manière plus globale sur l'ensemble de leur longueur, ainsi

c'est grâce à un plan chargé que l'on va simuler leur influence. Enfin l'accumulation de défauts peut entraîner la formation de pièges de porteurs qui vont altérer la conductivité locale dans le détecteur et perturber le transport et donc la collecte des charges. On modélise ce type de défaut en insérant des sphères de conductivité plus faible ou plus élevée que celle du détecteur afin d'observer leur influence sur le champ électrique.

1.4 Modélisation des détecteurs

1.4.1 Évolution du développement de la modélisation des détecteurs semi-conducteurs

Modèles analytiques

Dans un premier temps nous allons réaliser un historique du développement des modélisations des détecteurs à semi-conducteurs pour la détection de rayonnement, pour s'intéresser ensuite à chaque point clé dans la démarche. Les premières modélisations utilisaient les équations analytiques afin de comprendre les principaux principes de fonctionnement des semi-conducteurs en se basant sur des résultats expérimentaux. On se focalisait alors sur la collecte de charge, la résolution en énergie et la physique des particules mise en jeu.

Le premier modèle historique analytique est développé par Hecht dans les années 1930. Il donne le lien entre la charge induite sur une électrode, le champ électrique et les différentes constantes physiques liées aux porteurs de charge [55]. La forme de ce modèle dépend des hypothèses que l'on prend en compte, on détaillera une de ces formes dans une partie suivante. Puis le théorème de Ramo-Schockley, présenté dans une partie précédente formulé en 1938 a posé la fondation dans le calcul des charges et du courant induits dans un semiconducteur soumis à une source de rayonnement ionisant. Au fil des années et de l'évolution des matériaux utilisés les modélisations se sont complexifiées permettant de réaliser des calculs pour des situations de plus en plus réalistes.

L'équipe de Herbert Gajewski a commencé à travailler dans les années 1960 sur les équations de modélisation des semi-conducteurs, notamment sur le modèle de dérive [56]. Ils ont notamment travaillé sur l'état stationnaire des porteurs de charge ou les problèmes de valeurs limites dans des cas de champs électrique non linéaires.

Plus récemment, les travaux de Hamel sur la collection de charge ont permis d'introduire le modèle de diffusion-répulsion ou encore les effets de piégeage-dépiégeage [57]. Ces équations analytiques ont été couplées à des simulations numériques afin d'analyser le partage de charge pour des détecteurs CZT (Figure 1.32).

Méthodes numériques

Le besoin d'augmenter le réalisme et la complexité des simulations, l'amélioration des outils informatiques et le développement croissant de l'utilisation des semi-conducteurs dans le développement des imageurs gamma a poussé la communauté scientifique à s'intéresser aux méthodes numériques. Des méthodes de résolution par élément finis à l'aide d'équations différentielles soit ordinaires soit aux dérivées partielles peuvent être utilisées [58]. Dans cette article, il est question de méthodes numériques utilisées afin de résoudre le système d'équations aux dérivées partielles régissant le fonctionnement d'un

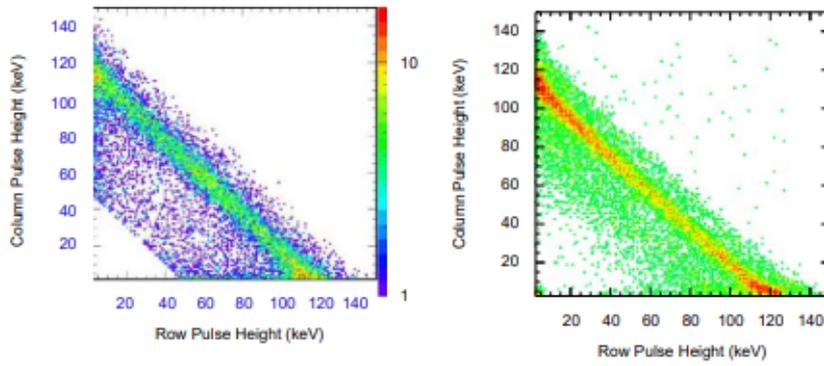


FIGURE 1.32 – Comparaison des histogrammes de partage de charge simulé (à gauche) et expérimental (à droite) pour une source de Cobalt 57 [57].

semi-conducteur. Différentes méthodes comme l'utilisation d'opérateurs non linéaires, le calcul de Jacobien ou la discrétisation des équations sont utilisées, couplées à la méthode des différences finies permettent la résolution de ces équations.

On peut également présenter une méthode qui permet de convertir les équations aux dérivées partielles du semiconducteur en une série d'équations différentielles ordinaires couplées qui pourront ensuite être résolues par un solveur adapté [59]. Dans ce papier on introduit la *method-of-lines* (MOF), qui permet de convertir une équation différentielle aux dérivées partielles de la forme

$$\partial_t u(x, t) = g(x, t, u, \partial_x u, \partial_x^2 u, \dots) \quad (1.19)$$

en un vecteur d'équations différentielles ordinaires grâce à l'approximation des différentes dérivées partielles obtenue à partir de la formule des différences finies pour chaque point d'intérêt discrétisé entre les conditions aux limites du système. Ceci permet d'obtenir une matrice \mathbf{D}_1 qui va être solution de l'équation $\partial_x \mathbf{u} = \mathbf{D}_1 \cdot \mathbf{u}$. On peut aussi calculer cette matrice pour les dérivées d'ordre supérieur grâce à un algorithme récursif. Finalement en appliquant cette méthode avec un solveur et un maillage adaptés on peut résoudre une équation aux dérivées partielles.

Dans le cas du semiconducteur on l'applique aux équations de Poisson et de transport des porteurs de charge, puis après résolution on peut obtenir les courbes de densité des porteurs de charge et du potentiel électrique selon la profondeur du détecteur (Figure 1.33).

On remarque la coïncidence presque parfaite entre les résultats provenant d'une solution analytique exacte et les résultats obtenus par la méthode MOL. Il existe énormément d'autres méthodes utilisant les équations différentielles ordinaires ou aux dérivées partielles pour l'étude du transport, de l'induction et de la collecte des charges, très souvent couplées à des solveurs numériques.

Méthode Monte-Carlo

La méthode Monte-Carlo correspond à un ensemble d'algorithmes permettant de s'approcher d'une valeur numérique en utilisant des procédés aléatoires, comme dans les jeux de hasard au casino de Monte-Carlo. Cette famille d'algorithmes a été conceptualisée en 1947 par Metropolis et mise sous forme mathématique dans un article coécrit avec Ulam

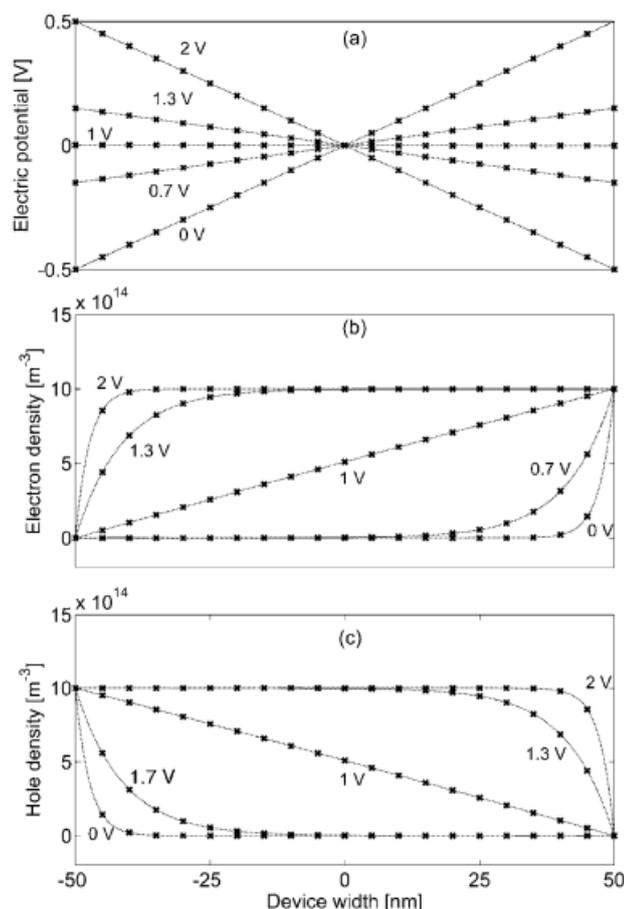


FIGURE 1.33 – Graphique des solutions numériques (lignes en pointillées) et analytiques exactes (croix) pour différents potentiels appliqués de : (a) le potentiel électrique, (b) la densité d'électrons et (c) la densité de trous [59].

en 1949 [60]. Aujourd'hui des codes de simulation tels que Geant4 ou Tripoli sont utilisés pour faire de la modélisation en physique des particules et des rayonnements ionisants.

Concernant l'application aux semi-conducteurs, ces méthodes ont été adaptées notamment pour le transport des porteurs dans le matériau détecteur. Le principe dans ce cas est d'utiliser des nombres aléatoires afin de simuler le caractère aléatoire du mouvement de l'électron. On considère les électrons libres qui interagissent par diffusion dans le matériau. Si on observe les interactions de cet électron pour un certain temps, la distribution des instants de passage de l'électron dans l'espace considéré va reproduire la forme de la fonction de distribution. On applique ce traitement pour un ensemble de porteurs de charge, ce qui permet d'évaluer les propriétés de transport dans une période transitoire pour un échantillon assez important [61]. On peut résumer le fonctionnement de l'algorithme pour un électron comme le suivant :

1. Générer un nombre aléatoire r et déterminer le temps de dérive libre.
2. Enregistrer le temps que la particule passe dans chaque sous espace k du volume durant la dérive.
3. Générer des nombres aléatoires pour déterminer le mécanisme de diffusion de la particule et l'état final.

4. Répéter jusqu'au ce que le nombre d'événements de diffusion soit atteint.
5. Calculer la fonction de distribution, la vitesse de dérive, l'énergie moyenne, . . .

Ces simulations ont permis d'apporter des éclaircissements sur le fonctionnement de certains mécanismes physiques, tel que le processus de dépôt d'énergie dans le matériau détecteur.

Cas du CZT

L'utilisation à large échelle du CZT est relativement récente dans l'histoire des semi-conducteurs. C'est dans les années 1980 qu'on découvre un intérêt dans ce matériau pour la détection de rayonnements X et γ . Les premières modélisations ont permis de déterminer les propriétés de base du matériau, tels que la largeur de bande interdite et les propriétés des transports de charge (temps de vie et mobilité) [62]. Dans un second temps on commence à étudier les modèles de transport de charge relatifs au CZT. Des simulations des mouvements des transports de charge sous l'influence d'un champ électrique sont alors développées, donnant des informations sur l'efficacité de collecte de charge, la résolution spatiale ou les caractéristiques des dépôts d'énergie en considérant les propriétés de transport des charges, le piégeage, la recombinaison et la diffusion [63].

En considérant les mécanismes déjà connus du dépôt d'énergie, de la collecte de charge et du partage de charge entre pixel [64], il est alors possible de simuler le spectre en énergie, la résolution en énergie, la linéarité et la résolution spatiale du détecteur CZT [65] (Figure 1.34).

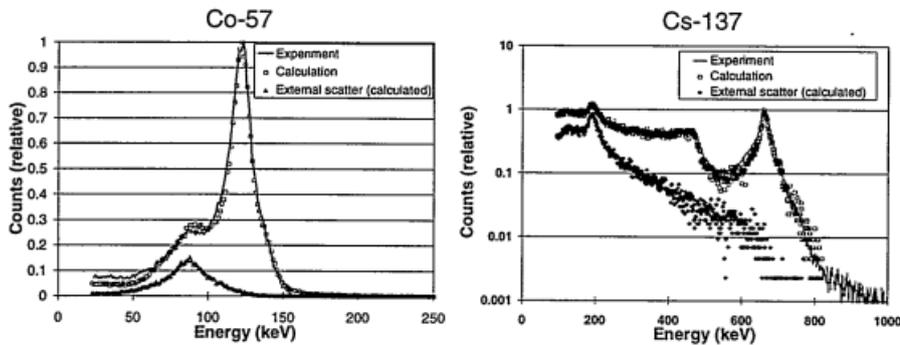


FIGURE 1.34 – Courbes des spectres en énergie du cobalt 57 (gauche) et du césium 137 (droite) simulés et expérimentaux [65]

Les simulations Monte Carlo ont également pu permettre de simuler de manière fiable la physique des interactions ayant lieu dans le cristal CZT, la génération et le transport des charges ainsi que la réponse associée. Toutes ces données ont pu valider les modèles analytiques et donner les limites de ces détecteurs aboutissant sur des solutions d'optimisation [66].

1.4.2 Modélisation et calcul du champ électrique

Il existe différentes approches pour la modélisation du champ électrique. Chacune présente des avantages et des inconvénients, et le choix de tel ou tel style de modélisation

repose principalement sur la complexité du détecteur CZT désiré ainsi que le niveau de précision visé. Les méthodes Monte-Carlo ne sont pas adaptées aux calculs du champ, mais il existe des méthodes analytiques pour les calculs de champ, par exemple en résolvant l'équation de Laplace avec des conditions aux limites. Ces méthodes sont utilisées pour des géométries à trois dimensions simples ou en deux dimensions grâce à l'analyse complexe.

Méthode des différences finies

La méthode des différences finies est une méthode utilisée pour résoudre numériquement des équations aux dérivées partielles en modélisation. Elle consiste à appliquer un maillage de pas fixe sur un volume étudié, puis appliquer les différences finies aux fonctions étudiées à chaque point de la maille. Dans son article, Nagel détaille la méthode d'application de cette méthode pour le calcul du champ électrique en résolvant l'équation de Poisson [67]. On cherche donc à résoudre l'équation $\Delta V = -\frac{\rho}{\epsilon_0}$. On va alors discrétiser cette équation selon la maille que l'on a choisi afin de calculer de manière matricielle la valeur du potentiel en chaque point de la maille. Puis grâce à la relation $\nabla V = -\vec{E}$ on peut déterminer le champ électrique du volume étudié. On se ramène alors à une résolution d'un système linéaire discrétisé selon les éléments de volume.

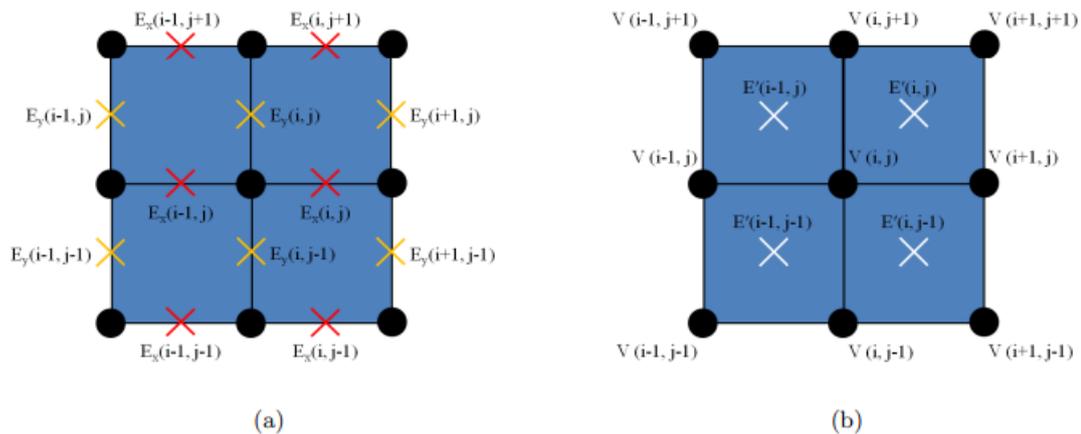


FIGURE 1.35 – Schéma des grilles pour obtenir les échantillons de champ électrique. (a) On calcule les différences centrales de champ électrique selon les axes x et y . (b) On fait la moyenne pour obtenir la valeur du champ électrique sur le pixel [67].

Méthode des éléments finis

La méthode des différences finies est simple à mettre en place mais est limitée à un maillage régulier et un traitement des conditions aux limites plus complexe. La méthode des éléments finis est une autre méthode permettant de résoudre des équations aux dérivées partielles en divisant le volume étudié en élément de volume plus faible que l'on appelle élément fini. Ceci est réalisé en discrétisant les dimensions de l'espace via un maillage. Le principe est alors d'exprimer le champ inconnu (ici le champ électrique) grâce à des fonctions approximatives dans chaque élément de volume de l'espace. Ces fonctions d'approximation, aussi appelées fonctions d'interpolation sont définies à partir des valeurs des variables de champ en certains points spécifiques appelées nœuds.

On peut décrire le fonctionnement de la méthode des éléments finis selon les étapes suivantes [68] :

1. Discrétiser l'espace étudié en sous éléments de volume. Une multitude de formes peuvent être utilisées (Figure 1.36). Le maillage peut être défini manuellement, mais généralement on utilise des outils comme le logiciel GMSH pour avoir une versatilité et une complexité supplémentaire [69].

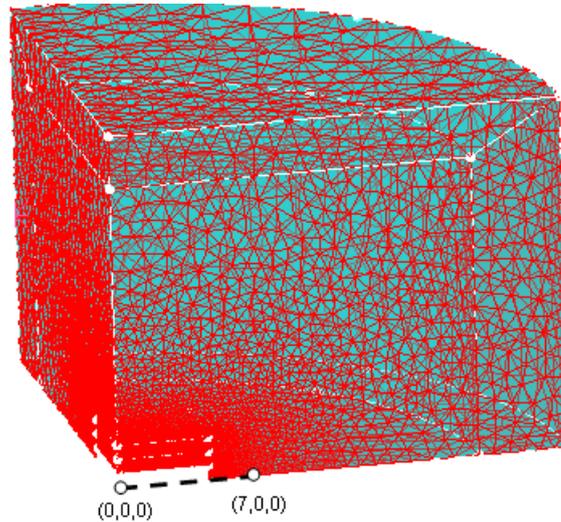


FIGURE 1.36 – Maillage d'un volume d'étude grâce à la méthode des éléments finis [70].

2. Assigner les nœuds à chaque élément et choisir la fonction d'interpolation afin de représenter la variation locale du champ au niveau de l'élément. On choisit généralement des fonctions polynomiales pour leur facilité d'intégration dont le degré va dépendre du nombre de nœuds associés à l'élément.
3. Déterminer la matrice des équations permettant d'exprimer les propriétés de chaque élément de manière individuelle.
4. Assembler les propriétés des éléments pour obtenir les équations globales du système. On combine les matrices de chaque élément pour former la matrice globale.
5. Imposer les conditions aux limites (champ électrique aux extrémités par exemple).
6. Résoudre les équations du système. Suite aux étapes précédentes on obtient un système d'équations à résoudre pour connaître la valeur du champ sur les nœuds où elle est inconnue. On doit parfois opérer de manière itérative afin de calculer des paramètres intermédiaires dans le cas de problème complexe.

L'article de John Barnes présente une application de la méthode des éléments finis à la modélisation de semi-conducteurs en combinant les équations de la physique avec les équations mathématiques de la méthode des éléments finis [71] (Figure 1.37).

Intégrales de frontière et méthodes spectrales

Pour les résolutions de l'équation de champ électrique, on peut se baser sur la connaissance de sa fonction de Green (champ créé par une charge ponctuelle). En l'absence de

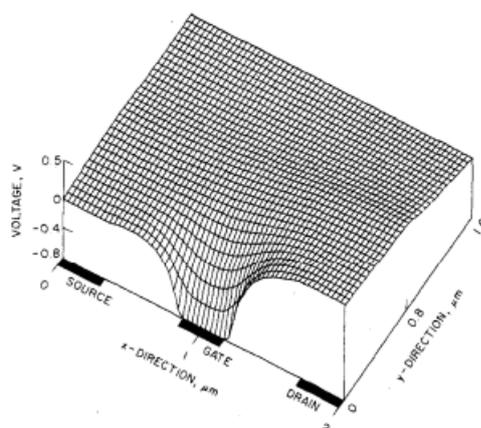


FIGURE 1.37 – Distribution de la tension dans un semi-conducteur GaAs par la méthode des éléments finis [71].

charge à l'intérieur d'un volume, on peut se contenter d'intégrer les fonctions de Green à la frontière du volume.

La convolution de la fonction de Green par la distribution de charge peut avantageusement être réalisé dans le domaine de Fourier. Dans le cas particulier de détecteurs plan sans charges de volume, il a par exemple été proposé par Kavadias une résolution semi-analytique pour déterminer le potentiel dans le volume du détecteur ainsi que les capacités inter-électrodes [72].

Modèles empiriques

Les modèles empiriques sont basés sur des données expérimentales couplées à des modèles théoriques permettant d'obtenir une représentation approchée de la distribution du champ électrique. Généralement la méthode d'application d'un modèle empirique est la suivante :

1. Mesurer de manière expérimentalement le champ électrique dans le détecteur en fonction de la position.
2. Ajuster la distribution du champ électrique obtenue de manière expérimentale en l'approximant pour obtenir une expression paramétrée.
3. Corriger les paramètres de la fonction d'ajustement en utilisant des algorithmes d'optimisation pour minimiser l'écart entre les mesures expérimentales et le modèle. On obtient une approximation empirique de la distribution du champ électrique.
4. Simplifier dans certains cas le modèles afin de simplifier sa représentation en modifiant la géométrie par exemple.
5. Valider le modèle en le comparant avec des données expérimentales provenant d'un autre jeu de données et calculer son efficacité.

L'article de Maura Pavesi [73] présente une méthode de reconstruction du champ électrique en comparant des données expérimentales avec un modèle physique. Les données expérimentales sont obtenues grâce à une technique du courant transitoire induit par laser (TCT ou laser-induced transient-current technique). L'ajustement est réalisé à partir de

la loi du mouvement des électrons selon la profondeur du détecteur couplées à la loi d'induction. Les contraintes sur la tension appliquée et l'épaisseur du détecteur permettent de remonter à la reconstruction du champ électrique. Dans sa thèse, Olivier Alirol allie les simulations et les mesures expérimentales de champ électrique par effet Pockels. Ces mesures permettent dans un premier temps de vérifier son modèle théorique, et de l'affiner dans un second temps [74].

1.4.3 Création et transport des porteurs de charge

Afin d'illustrer cette partie on s'intéresse au travail de thèse de Silvère Lux, qui propose un modèle de calcul des performances des détecteurs CZT afin d'améliorer leur résolution spatiale. Cette étude est basée sur un modèle théorique simplifiée du détecteur et permet l'utilisation d'équations semi-analytiques pour le calcul des performances [75]. Nous allons ici présenter le modèle qu'il a utilisé ainsi que le fonctionnement de l'algorithme associé

Le nombre de paramètres d'entrée du détecteur est limité tout en prenant en compte les approximations que cela entraîne. On va alors considérer six paramètres qui sont la taille des pixels p , l'espace inter-pixel g , l'épaisseur du détecteur d , la tension appliquée au détecteur V , l'énergie des photons E et le bruit électronique σ_B . Les hypothèses d'approximation appliquées dans ce modèle sont les suivantes :

- Le signal induit par une charge au niveau d'un pixel est défini par le potentiel de pondération.
- La distribution spatiale des électrons dans le volume du détecteur suit une distribution gaussienne sphérique dont l'élargissement est dû à la diffusion et à la répulsion coulombienne.
- On considère uniquement le signal induit par les électrons, car leur mauvaises propriétés de transport limite grandement l'influence des trous.
- La trajectoire des électrons est rectiligne, ce qui est vrai au dessus de l'électrode mais ne l'est plus au dessus de l'espace inter-pixel. Toutefois, pour des pixels dont la taille est grande par rapport à l'espace inter-pixel, cela a peu d'influence sur le résultat final.

On définit le fonctionnement de l'algorithme que l'on va utiliser pour le calcul de la résolution spatiale du détecteur modélisé, dont le déroulement est en quatre étapes (Figure 1.38).

1. La première étape consiste à calculer le potentiel de pondération. Il dépend de la taille du pixel, de l'espace inter-pixel et de l'épaisseur du détecteur ; plus exactement il dépend du rapport entre la dimension du pixel et l'épaisseur.
2. L'étape suivante consiste à déterminer la distribution spatiale des électrons, et son évolution temporelle. Celle-ci dépend d'abord de la distribution initiale des électrons due aux interactions rayonnement-matière, et est donc liée à l'énergie du photon. Les électrons ainsi générés transitent vers l'anode et sont soumis à la diffusion thermique et à la répulsion coulombienne. Le temps de transit est lié à l'épaisseur et à la tension appliquée. Pour cette partie les interactions et donc le nuage électronique initial est fait par méthode Monte-Carlo.
3. Connaissant le potentiel de pondération (qui définit le signal en charge induit par un électron) ainsi que la taille du nuage électronique, on effectue la convolution des

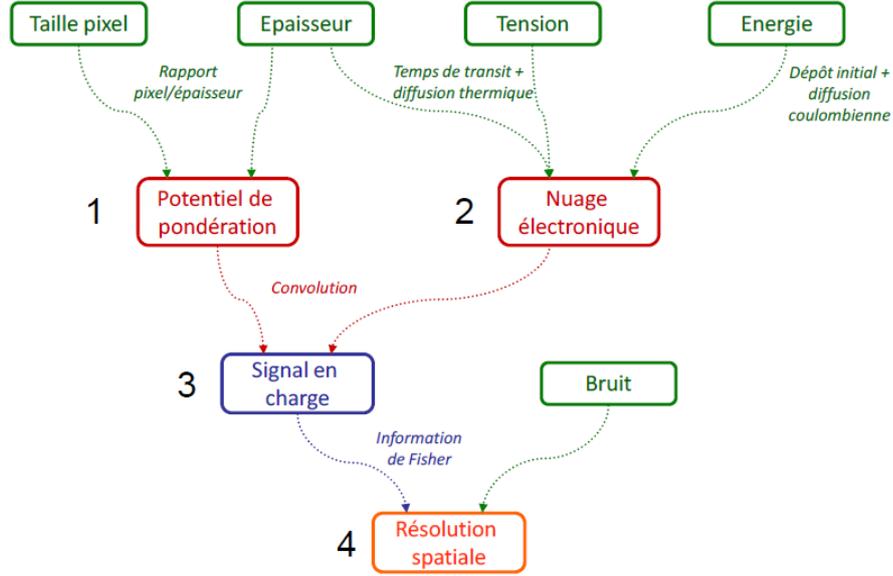


FIGURE 1.38 – Schéma de fonctionnement de l'algorithme, les paramètres d'entrée sont représentés en vert [75].

deux. Ce qui donne un nouveau potentiel de pondération dit effectif qui définit le signal en charge induit par le nuage d'électrons.

4. Finalement, en ayant d'une part le signal et d'autre part le bruit associé à ce signal, on peut calculer la résolution spatiale en s'appuyant sur l'information de Fisher et la borne de Cramér-Rao.

Chaque photon primaire va générer un certain nombre d'électrons secondaires dépendant de son énergie, qui peuvent eux-mêmes générer des photons secondaires et ainsi de suite. On peut alors schématiser l'interaction d'un photon comme étant une suite de dépôts très localisés séparés par une grande distance. Chacune de ces positions de dépôt vont générer un nuage électronique qui va s'étendre sous l'effet de la diffusion thermique et de la répulsion coulombienne.

L'équation décrivant la migration du nuage électronique soumis principalement aux deux phénomènes physiques précédents est :

$$\frac{\partial \rho(r, t)}{\partial t} = D \nabla^2 \rho(r, t) - \mu_e \nabla \cdot (\rho(r, t) E(r, t)) \quad (1.20)$$

Avec ρ la densité de charges, r la distance par rapport au centre du nuage électronique, D le coefficient de diffusion d'Einstein et E le champ électrique défini par rapport au théorème de Gauss. Pour une distribution spatiale sphérique

$$\rho(r, t) = \frac{Ne}{(2\pi\sigma_e^2(t))^{\frac{3}{2}}} \exp\left(-\frac{r^2}{2\sigma_e^2(t)}\right) \quad (1.21)$$

$$E(r, t) = \frac{1}{\epsilon} \int_0^r u^2 \rho(u, t) du \quad (1.22)$$

Avec N le nombre d'électrons dans le nuage, $N = \frac{E}{E_p}$ où E est l'énergie du photon et E_p l'énergie de création de paires du matériau ($E_p=4.63$ eV pour le CZT), ϵ la permittivité du

matériau ($\epsilon \approx 10^{-10} F/m$ pour le CZT). En appliquant l'équation (1.21) dans l'équation (1.20) et en intégrant sur le volume on obtient

$$\frac{\partial \sigma_e^2(t)}{\partial t} = 2 \left(D + \frac{\mu_e N e}{3(2\pi)^{\frac{3}{2}} \epsilon \sigma_e(t)} \right) \quad (1.23)$$

On note D' le nouveau coefficient de diffusion [57]

$$D' = D + \frac{\mu_e N e}{3(2\pi)^{\frac{3}{2}} \epsilon \sigma_e(t)} \quad (1.24)$$

On connaît la valeur initiale ($\sigma_e(t=0)$) et on considère une petite variation temporelle Δt :

$$\sigma_e^2(t + \Delta t) = \sigma_e^2(t) + 2D'(t) \cdot \Delta t \quad (1.25)$$

En résolvant cette équation par un changement de variable on peut alors obtenir l'expansion du nuage électronique selon le temps de dérive.

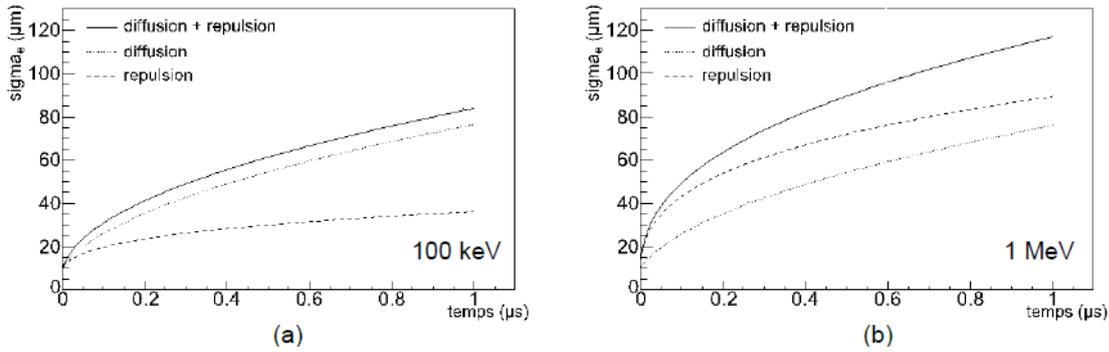


FIGURE 1.39 – Courbes d'évolution temporelle de la taille du nuage électronique en fonction de l'énergie du photon initiale et du type de phénomène physique pris en compte.

1.5 Conclusion

Dans cette partie nous avons pu avoir un aperçu du cadre de cette étude et des différents aspects à approfondir. Le principe de l'imagerie γ et le fonctionnement des détecteurs est un point d'entrée pour rappeler l'objectif qui est de mesurer des interactions dans un détecteur afin d'identifier et de localiser une source émettant des photons. Les détecteurs CZT sont un type de détecteur à semi-conducteur utilisés dans plusieurs domaines pour ses caractéristiques intéressantes. Néanmoins la présence de défauts structuraux impacte le fonctionnement de ces détecteurs et dégrade les signaux permettant la localisation des interactions. L'étude de ces défauts en simulation permettrait à la fois de simuler et étudier le comportement d'un détecteur comportant un ou plusieurs types de défaut de manière réaliste, et d'anticiper les erreurs causées par ces défauts en corrigeant les signaux de manière à passer outre l'effet de ces défauts et augmenter les performances des détecteurs.

Développement de simulations 3D de détecteurs CZT avec défauts

Le cadre de notre étude porte sur l'apprentissage des données d'un détecteur CZT réel via l'utilisation de données simulées. De ce fait il nous faut des données de complexité comparables aux données obtenues de manière expérimentale afin d'obtenir la meilleure précision possible dans le calcul des paramètres recherchés. Cela passe par le développement de simulations de détecteur les plus réalistes possibles, afin notamment de modéliser les défauts et de pouvoir observer leur influence sur les performances du détecteur, comme l'efficacité de collecte. Ces simulations doivent également pouvoir s'exécuter de manière rapide afin d'optimiser l'apprentissage. Dans cette partie nous allons détailler le cheminement du travail réalisé sur ces simulations, de la modélisation du transport à l'implémentation des défauts.

2.1 Modélisation du transport

2.1.1 Équation de Hecht analytique

On peut dans un premier temps s'intéresser à des modélisations très simples et analytiques du fonctionnement du détecteur CZT. L'équation de Hecht permet de décrire l'évolution de la charge induite par la dérive selon la direction de la profondeur d'une chargée générée lors d'une interaction d'un photon dans le semi-conducteur [55]. Cette équation permet également de prendre en charge le piégeage des charges dans le matériau. L'application de cette équation nécessite plusieurs approximations :

- On considère un cas unidimensionnel avec dérive selon la profondeur.
- La distribution spatiale de la charge générée par l'interaction suit une distribution de Dirac.
- Le champ électrique est distribué de manière homogène dans le volume.
- Les porteurs se déplacent selon leur mobilité μ et sont piégées selon leur durée de vie τ .
- On néglige l'influence de la diffusion.

— le champ de pondération est $E_w = \frac{1}{L}$

On part des durées de transit de chaque type de porteur de charge lors d'une interaction à une profondeur z_0 :

$$T_e = \frac{L - z_0}{\mu_e E} \quad (2.1)$$

$$T_h = \frac{z_0}{\mu_h E} \quad (2.2)$$

Avec T_e et T_h les durées de transit moyennes des électrons et des trous, L l'épaisseur du détecteur, μ_e et μ_h les temps de vie moyens des porteurs de charge et E l'intensité du champ électrique. Pour obtenir la charge induite totale sur le temps de transit on a

$$Q = \int_0^T Q(t) u E_w dt \quad (2.3)$$

$$Q = \int_0^T Q(t) \mu E E_w dt \quad (2.4)$$

$$Q_e = \int_0^{T_e} Q_0 e^{-\frac{t}{\tau_e}} \mu_e \frac{E}{L} dt \quad (2.5)$$

$$= \mu_e \frac{Q_0 E}{L} \left[-\tau_e e^{-\frac{t}{\tau_e}} \right]_0^{T_e} \quad (2.6)$$

$$Q_h = \int_0^{T_h} Q_0 e^{-\frac{t}{\tau_h}} \mu_h \frac{E}{L} dt \quad (2.7)$$

$$= \mu_h \frac{Q_0 E}{L} \left[-\tau_h e^{-\frac{t}{\tau_h}} \right]_0^{T_h} \quad (2.8)$$

$$Q = Q_e + Q_h = \frac{Q_0 E}{L} \left[\mu_e \tau_e \left(1 - e^{-\frac{T_e}{\tau_e}} \right) + \mu_h \tau_h \left(1 - e^{-\frac{T_h}{\tau_h}} \right) \right] \quad (2.9)$$

$$Q = \frac{Q_0 E}{L} \left[\mu_e \tau_e \left(1 - e^{-\frac{L - z_{int}}{\mu_e \tau_e E}} \right) + \mu_h \tau_h \left(1 - e^{-\frac{z_{int}}{\mu_h \tau_h E}} \right) \right] \quad (2.10)$$

avec $u = \mu E$ la vitesse de dérive des charges, Q_0 la charge initiale déposée lors de l'interaction et τ le temps de vie moyen des porteurs de charge.

Il existe également une version de cette équation où l'on néglige l'apport des trous dans le calcul de la charge. En effet ceux ci ayant des mauvaises propriétés de transport dans le CZT, avec une mobilité et un temps de vie inférieurs, se font piéger plus facilement. On a alors

$$Q = \frac{Q_0 E}{L} \left[\mu_e \tau_e \left(1 - e^{-\frac{L - z_{int}}{\mu_e \tau_e E}} \right) \right] \quad (2.11)$$

Pour vérifier si l'impact de ces trous est bien négligeable devant celui des électrons, on peut tracer les deux courbes de charge induite selon la profondeur d'interaction pour les comparer. On utilise les valeur numériques suivantes, typique du CZT :

- $Q_0 = q n_0$ avec $q = 1.6 \times 10^{-19}$ C et $n_0 = 2 \times 10^6$
- $E = 1000$ V · cm⁻¹
- $\mu_e = 1050$ cm² · s · V⁻¹ et $\mu_t = 70$ cm² · s · V⁻¹
- $\tau_e = 3 \times 10^{-6}$ s et $\tau_t = 1 \times 10^{-7}$ s

- $L = 0.6$ cm

Le libre parcours moyen pour les électrons est alors de 3.15 cm et pour les trous de seulement 70 μm . On obtient alors les courbes de la figure 2.1.

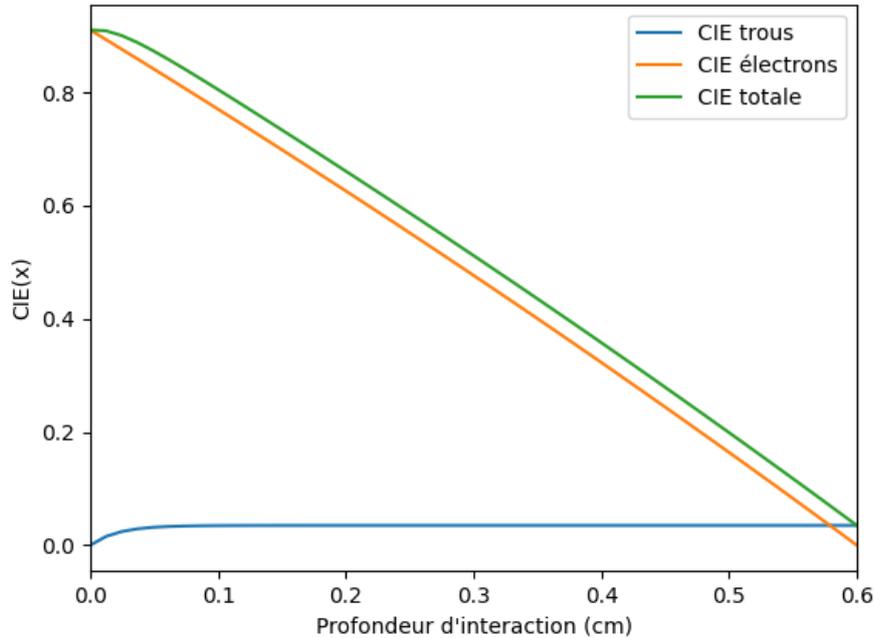


FIGURE 2.1 – Comparaison des courbes d'efficacité d'induction de charge (Q/Q_0) selon la profondeur d'interaction. La cathode est à gauche ($x=0$ cm) et l'anode à droite ($x=0.6$ cm).

On remarque que les deux courbes se rejoignent à partir d'une certaine profondeur d'interaction, ce qui prouve que la contribution des trous dans le calcul de la charge induite est faible, on peut alors considérer n'utiliser que les signaux obtenus sur les anodes pour déterminer la charge induite. Un autre élément nous prouvant ce phénomène est le fait d'obtenir une charge quasiment nulle lorsque l'interaction a lieu près de l'anode, où les électrons induisent immédiatement la charge électrique.

2.1.2 Simulation 1D

En utilisant la formule de Hecht il est alors possible d'obtenir des courbes d'évolution de charge et de position. On s'intéresse à la dérive des charges uniquement selon la profondeur, les électrons vers l'anode et les trous vers la cathode. Au lieu de considérer un champ électrique uniforme on complexifie cette approche simpliste en s'affranchissant de l'hypothèse d'uniformité du champ. Pour illustrer, on choisit ici un champ linéaire augmentant entre la cathode à $x = 0$ et l'anode à $x = L$, selon l'équation

$$E(x) = 100 + \frac{900}{L} \cdot x \quad (2.12)$$

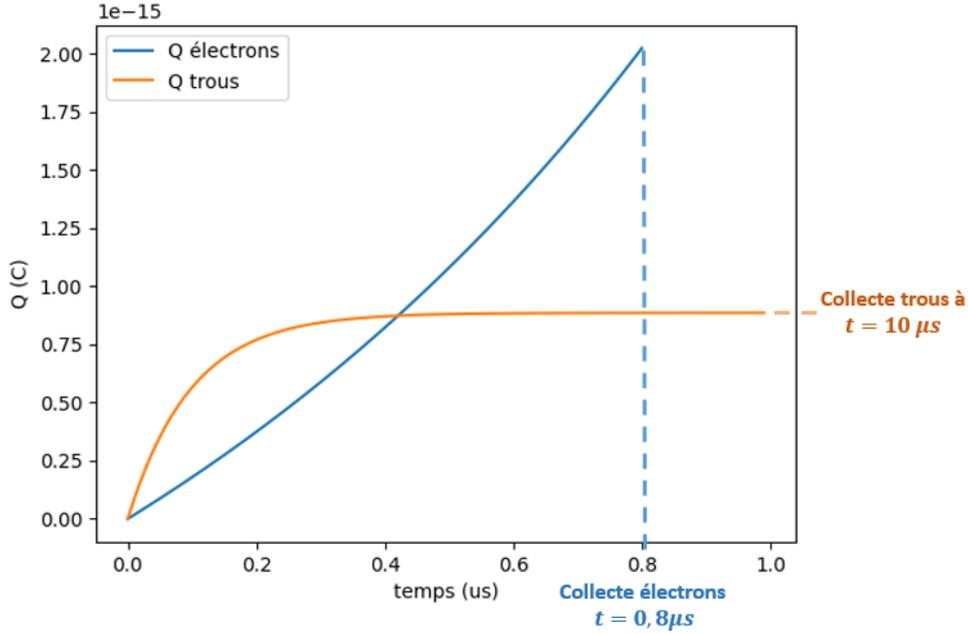


FIGURE 2.2 – Courbes de la charge induite Q dans le temps pour chaque porteur de charge pour une énergie incidente de 660 keV pour une interaction à $z = 0.18$ cm.

On observe sur la figure 2.2 que la majeure partie du total de la charge induite provient des électrons. On a donc un champ électrique variant de 100 à 1000 $V \cdot cm^{-1}$ sur l'épaisseur du détecteur. On décide d'imposer un potentiel de pondération sur le pixel touché afin de calculer le courant induit grâce aux électrons. On applique un potentiel de pondération $V_w = \frac{x}{L}$ pour garder une modélisation simple mais réaliste.

L'équation différentielle du mouvement est résolue numériquement grâce à l'algorithme de Runge-Kutta 4

$$\frac{dx}{dt} = \mu E(x) \quad (2.13)$$

Et on a l'équation d'évolution du piégeage de la charge induite

$$\frac{dQ}{dt} = -\frac{Q}{\tau} \quad (2.14)$$

avec τ la constante de piégeage, la charge s'obtient analytiquement

$$Q(t) = Q_0 e^{-\frac{t}{\tau}} \quad (2.15)$$

On peut alors obtenir le courant d'induction par

$$i(t) = Q(t) \cdot \mu E(x) \cdot E_w \quad (2.16)$$

avec E_w le champ de pondération associé au potentiel V_w .

On peut tracer l'évolution des positions pour chaque type de porteur de charge en fonction de la profondeur d'interaction dans le détecteur (Figure 2.3). De la même manière que précédemment on remarque que le transport des électrons est environ dix fois plus rapide que celui des trous.

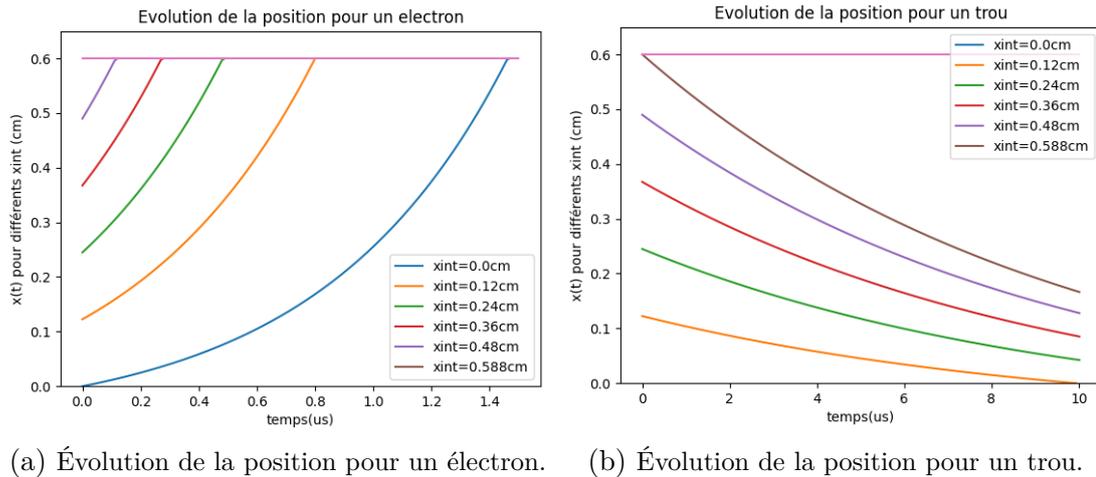
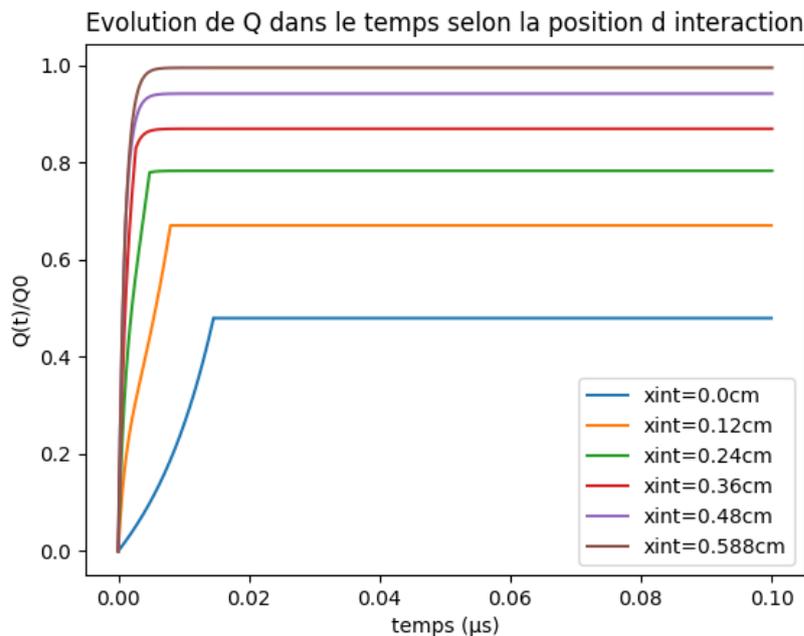


FIGURE 2.3 – Courbes des trajectoires des deux porteurs selon la position d'interaction.

FIGURE 2.4 – Courbes d'évolution temporelles de la charge induite Q sommée selon la position d'interaction

On peut également tracer l'évolution du courant induit normalisé sur les électrodes en fonction du temps et de la position d'interaction (Figure 2.4). On voit apparaître dans un premier temps la contribution des électrons plus rapides avec une pente raide, puis celle des trous avec une convergence vers un plateau, donc l'amplitude dépend de la profondeur d'interaction. Plus le photon interagit loin de l'anode, moins les électrons parviennent à être collecté intégralement et plus la charge collectée est faible devant la charge déposée.

2.1.3 Simulation 3D

Afin de passer à un modèle 3D, on utilise des tableaux 3D de potentiel appliqué et de potentiel de pondération simulés à partir de l'environnement Tasmania [76]. Ce dernier est

un environnement C++ adapté aux modélisations de détecteurs de particules. Il permet l'évaluation du champ électrique par une méthode spectrale utilisant les transformées de Fourier pour l'évaluation du champ électrique, une cartographie des cartes de champ grâce à une intégration l'équation différentielle de la dérive des électrons dans le détecteur, le calcul de la diffusion du nuage d'électrons grâce au Monte-Carlo et la détermination des signaux transitoires. On peut facilement accéder à chaque paramètre de la simulation [72].

La méthode spectrale de calcul du champ électrique par transformée de Fourier est basée sur la résolution de l'équation de Poisson $\Delta V = f$ dans l'espace de Fourier avec $f = -\frac{\rho}{\epsilon_0}$. On a alors

$$k^2 V(k) = f(k) \quad (2.17)$$

$$\Leftrightarrow V(k) = \frac{f(k)}{k^2} \quad (2.18)$$

On peut alors simuler un détecteur CZT, ainsi que les potentiels présents à l'intérieur de celui-ci. Le détecteur est échantillonné en 100 valeurs sur la longueur (x) et la largeur (y), et 52 en épaisseur (z) avec un pas $\Delta = 100\mu m$. Sur la figure 2.5 on peut observer sur les potentiels électrique et de pondération appliqué sur le détecteur simulé. A gauche le champ de potentiel électrique pixellisé, puis au centre et à droite deux coupes du potentiel de pondération appliqué sur l'anode centrale.

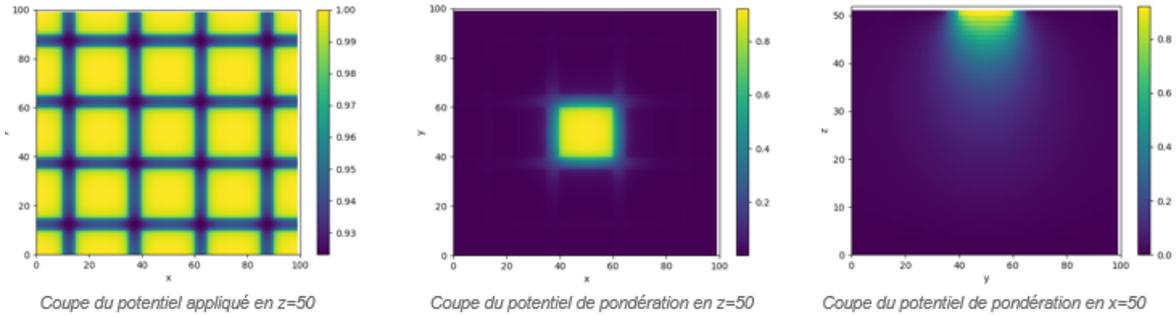


FIGURE 2.5 – Coupes des potentiels appliqués (à gauche) et des potentiels de pondération (au centre et à droite).

Contrairement à la simulation unidimensionnelle précédemment présentée, on ne dispose pas d'une formule analytique du champ électrique nous avons besoin d'interpoler les valeurs du potentiel électrique appliqué afin d'avoir une évaluation de l'intensité du champ électrique en chaque point avec une dérivation par différences finies. On obtient alors trois tableaux pour chaque coordonnée x, y et z de la forme

$$E(x) = \frac{(V_{app}(x + \Delta x) - V_{app}(x - \Delta x))}{2\Delta x} \quad (2.19)$$

avec Δx le pas (ici $100\mu m$). De la même manière que pour la simulation 1D on détermine les trajectoires ainsi que les charges induites pour chacun des types de porteurs de charge (Figure 2.6). On remarque les mêmes comportements que précédemment, avec une légère courbure sur la fin des courbes due à la présence de charges de contact distribuées à proximité des électrodes ainsi qu'à la géométrie pixel.

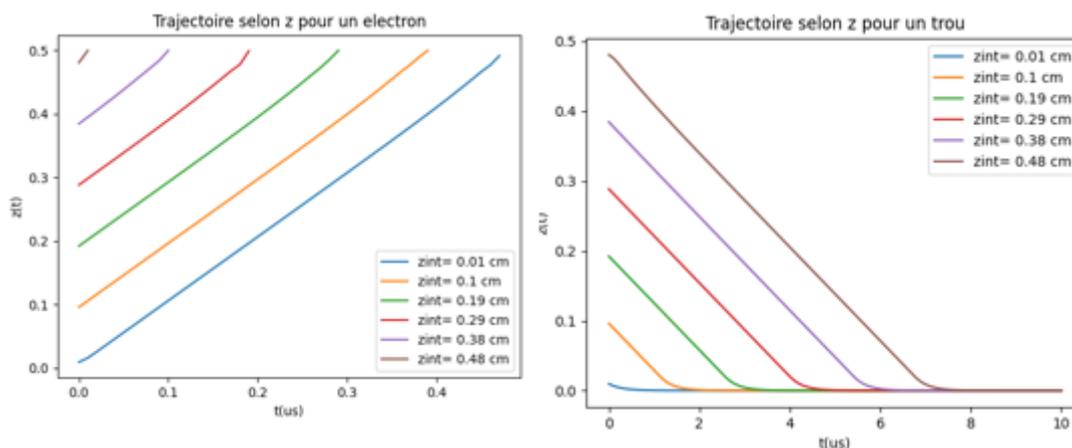


FIGURE 2.6 – Courbe des trajectoires des 2 porteurs selon la position d’interaction en 3D.

Il est également possible de tracer des coupes de champ électrique dans le sens de la hauteur, afin d’observer le comportement dans le volume du détecteur (Figure 2.7).

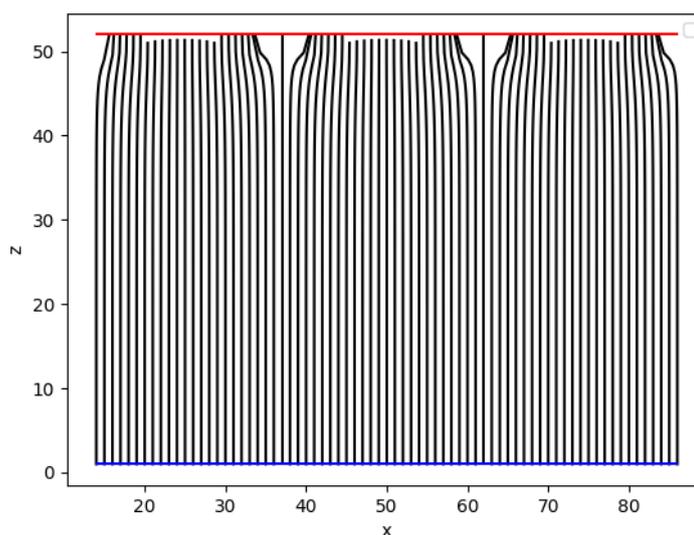


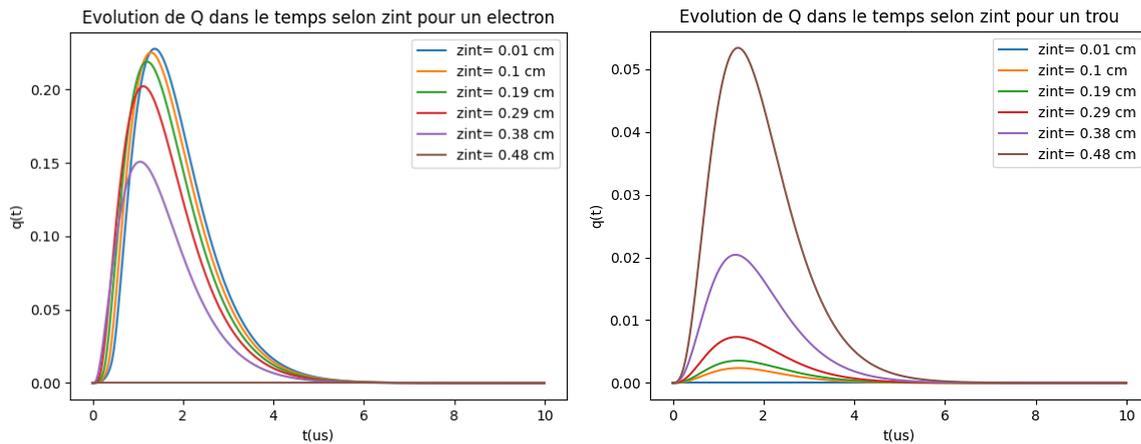
FIGURE 2.7 – Coupe du champ électrique selon x .

La cartographie des lignes de champ permet de bien observer la pixellisation côté anode (en rouge), causée par les espaces inter-pixel qui ne permettent pas la collecte de charge. L’extension du nuage électronique provoque du partage de charge, pour des interactions lointaines et entre deux pixels, où la charge induite finale se retrouve sur deux pixels ou plus.

Pour augmenter le côté réaliste de notre simulation, on décide d’introduire un filtrage sur le signal. En effet lors des mesures expérimentales sur détecteur réel, le signal de sortie est filtré par un filtre de mise en forme analogique $CR - RC^2$ de constante de temps $\tau = 500$ ns qui permet d’optimiser le rapport signal sur bruit [77]. On utilise ce filtre passe-bande facilement réalisable en électronique analogique (Figure 2.8). On peut observer la différence d’amplitude entre la charge induite par un électron et un trou, qui

est supérieure d'un facteur 4 pour l'électron. On remarque également que le maximum d'induction est retardé par rapport à l'instant de collecte, dû à la présence du filtre passe bande.

La figure 2.9 correspond à la superposition des deux courbes précédentes et on peut alors remarquer une similarité importante avec la courbe des charges induites par les électrons, avec une influence faible des trous, mis à part pour les interactions en profondeur. Ceci permet de renforcer le fait de considérer dans le calcul de charge la contribution des électrons.



(a) Évolution temporelle de la charge induite pour un électron. (b) Évolution temporelle de la charge induite pour un trou.

FIGURE 2.8 – Courbes d'évolution de la charge induite pour chaque porteur de charge selon la profondeur d'interaction.

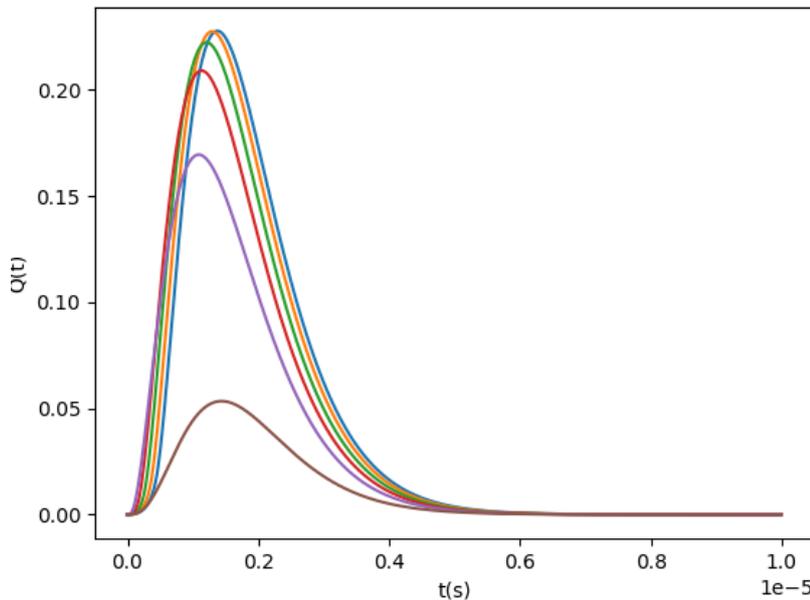


FIGURE 2.9 – Évolution de la somme de la charge induite par les deux types de porteurs.

On peut pointer certaines limites dans ce modèle : la non prise en compte de la diffusion ainsi que du dépiégeage des charges, et l'absence de charges électriques de volume pouvant

simuler les défauts du cristal. Ce dernier point sera étudié de manière plus approfondie dans la suite de l'étude.

2.2 Approche de modélisation avec défauts

2.2.1 Modélisation détecteur

Le couplage entre la modélisation d'un détecteur réaliste avec la présence de défauts et la gestion des interactions rayonnement-matière n'est pas une chose facile à traiter par simulation. Généralement le transport de particules ionisantes est géré par une simulation Monte-Carlo. Puis il faut simuler la génération d'événements dans le volume du détecteur, le transport des porteurs de charge, l'induction de signal et le traitement électronique. On modélise le transport des charges grâce à des équations différentielles ordinaires ou aux dérivées partielles. Néanmoins l'approche temporelle directe est trop complexe car le coût en calcul de ces opérations pour chaque interaction est trop important. De plus, la méthode adjointe utilisant l'équation adjointe du transport des électrons permet de cartographier la CIE (Charge Induction Efficiency), quantifiant l'efficacité du détecteur à mesurer les charges induites lors des interactions de photons, plus rapidement mais possède ses limites [78] : le champ électrique à l'intérieur du détecteur doit être solution de l'équation de Laplace $\nabla \vec{E} = 0$ et la simulation des impulsions transitoires est instable car dans ce cas le nombre de Péclet, défini comme le rapport entre le transport par dérive et le transport par diffusion thermique vaut.

$$Pe = \frac{E\Delta x}{kT}e = \frac{V_{bias}}{V_{th}} \cdot \frac{\Delta x}{L} \sim \frac{500}{50 \times 25 \cdot 10^{-3}} \sim 400 \quad (2.20)$$

Pour être stable, ce nombre doit être inférieur ou égal à 1. Dans notre cas, cela signifie que l'on devrait soit réduire le pas en x à l'ordre du μm soit artificiellement augmenter la diffusion thermique.

Tasmania permet l'accès et la modification de chaque paramètre de la simulation, comme le champ électrique, qui peut être défini comme non-uniforme et réévalué lors de la simulation. La figure 2.10 est une illustration des potentiels électrique et de pondération appliqués sur une géométrie de détecteur classique. On peut retrouver l'anode pixellisée ainsi que le potentiel de pondération sur le pixel central de l'anode. Par exemple les équations de piégeage-dépiégeage peuvent être incluses et utilisées pour mettre à jour le champ électrique $\vec{E} = -\vec{\nabla}V$ de manière itérative comme une fonction d'accumulation de charge, en utilisant l'équation de Poisson. Si p, n désignent la densité de porteurs de charge libres et p_t, n_t la densité de porteurs piégés et s'il y a un unique niveau de piège pour chaque porteur :

$$\frac{\partial p_t}{\partial t} = \frac{p - \theta_h p_t}{\tau_h} \quad (2.21)$$

$$\frac{\partial n_t}{\partial t} = \frac{n - \theta_e n_t}{\tau_e} \quad (2.22)$$

$$\nabla \cdot \vec{E} = \frac{p + p_t - n - n_t}{\epsilon} \quad (2.23)$$

avec $\tau_{e,h}$ les constantes de piégeage des électrons et des trous, $\theta_{e,h}$ les constantes d'équilibre (ratio entre les constantes de piégeage et de dépiégeage) et ϵ la permittivité du matériau.

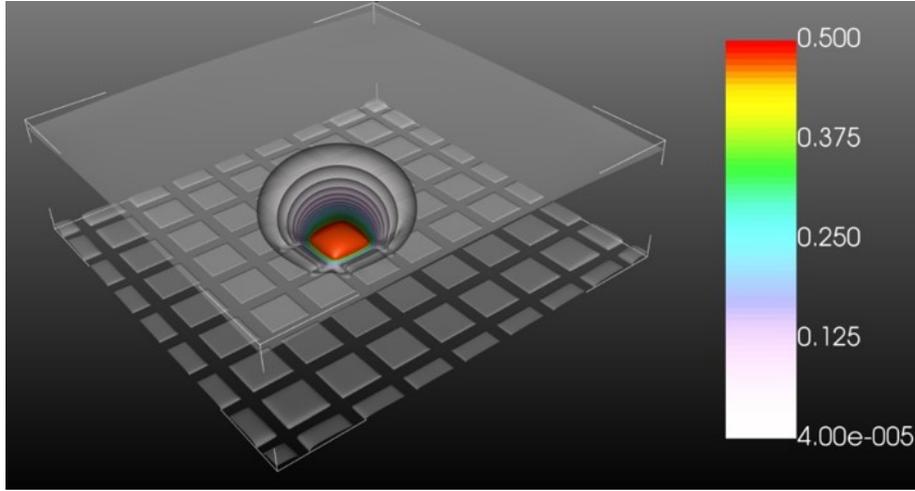


FIGURE 2.10 – Simulation du potentiel de pondération sur le pixel central d'une anode dans un détecteur de 5mm d'épaisseur avec Tasmania. La cathode n'est pas segmentée et l'anode est segmentée en électrodes carrés de 2.5mm de pas.

Pour obtenir le champ tel qui s'établit en régime permanent, on va procéder ainsi : lors de la dérive des charges dans le détecteur, on calcule pour chaque itération la distribution de charges piégée pour mettre à jour le champ électrique. Pour résumer le calcul du champ électrique s'effectue en quatre étapes :

1. Partant d'un champ électrique donné, on simule le transport des charges pour un lot d'interactions.
2. La densité moyenne de porteurs de charge libre est évaluée.
3. En considérant les constantes d'équilibre de piégeage, on calcule la distribution des charges en régime permanent.
4. Le champ électrique est mis à jour en prenant en compte la densité des charges piégées.
5. On itère jusqu'à stabilité du champ électrique.

2.2.2 Modélisation des défauts

On peut utiliser une alternative à la méthode précédente utilisant le calcul des populations de pièges afin de modéliser les défauts. Pour modéliser les inhomogénéités, on combine l'équation de conservation du courant avec la loi d'Ohm. Quand on a une distribution de la conductivité d'un matériau σ , il est possible de résoudre ces équations de manière itérative pour trouver la distribution de charge $p - n$ et le champ résultant :

$$\text{div}(\sigma \vec{E}) = 0 \quad (2.24)$$

$$\vec{\nabla} \sigma \cdot \vec{E} + \sigma \vec{\nabla} \cdot \vec{E} = 0 \quad (2.25)$$

$$-\Delta V = \vec{\nabla} \cdot \vec{E} = \frac{p-n}{\epsilon} = -\vec{E} \cdot \nabla(\ln \sigma) \quad (2.26)$$

Pour simuler le signal de sortie de ce détecteur modélisé, il faut calculer le champ électrique : le potentiel de pondération résulte uniquement de la géométrie tandis que le potentiel appliqué est obtenu en implémentant un ou plusieurs défauts dans le volume du détecteur. La modification de ce dernier entraîne une modification de la distribution de la charge dans le détecteur et ainsi du champ électrique (Figure 2.11).

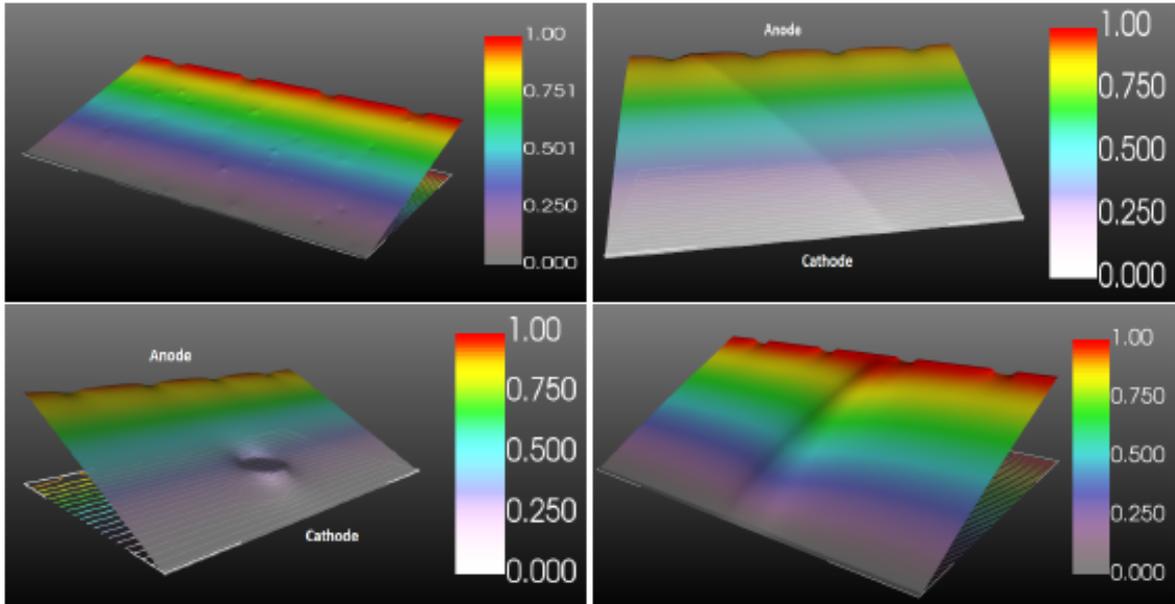


FIGURE 2.11 – Coupes des potentiels appliqués simulés (normalisé à 1) dans le détecteur CZT sous Tasmania. Le détecteur a une épaisseur de 5mm et le pas des anodes est 2.5mm. De gauche à droite et de haut en bas : défauts ponctuels discrets ; charges piégées distribuées sur un plan ; défaut sphérique de haute conductivité ; irradiation de haute intensité par rayons X sur la cathode.

En connaissant le champ électrique à chaque point du volume du détecteur on peut à présent calculer la trajectoire des charges positives et négatives selon l'équation différentielle de la dérive

$$\frac{\partial \vec{r}(x, y, z)}{\partial t} = \pm \mu_{e,h} \vec{E}(x, y, z) \quad (2.27)$$

avec \vec{r} le vecteur position du porteur de charge dans le détecteur, $\mu_{e,h}$ la mobilité du porteur de charge et \vec{E} le champ électrique. Comme le champ électrique est d'une intensité importante, d'après l'expression du nombre de Peclet on peut négliger la diffusion devant la dérive. Pour simplifier et accélérer le temps de calcul, la diffusion thermique et la répulsion de Coulomb sont modélisées analytiquement : l'expansion du nuage électronique est calculée comme une fonction du temps de transit [79]. L'évolution de la forme du nuage due au champ électrique et le partage de charge sont modélisés en divisant la charge du nuage en plusieurs (typiquement de l'ordre de 20 à 50) paquets discrets. Puis on considère ces paquets de manière individuelle en sommant leur contribution sur l'évolution du

champ électrique en fonction de leur localisation dans le détecteur simulé.

Cette équation différentielle est résolue en utilisant la méthode de Runge-Kutta et accélérée par l'unité graphique de l'ordinateur (GPU) en utilisant OpenCL, dont le fonctionnement sera développé dans la partie suivante, en traitant chaque paquet de charge de manière indépendante. L'interaction et le transport de chaque particule étant considérés indépendants, on peut se permettre de paralléliser ces calculs de manière à les optimiser. Cette parallélisation du calcul permet de réduire le temps de calcul par rapport à un usage du processeur uniquement. Dans le cas des charges ponctuelles discrètes et du défaut planaire chargé, le calcul du champ électrique est direct. Cependant la présence d'un défaut de conductivité affecte de manière importante la distribution des charges sur une petite zone, et le champ électrique doit alors être calculé de manière itérative. Enfin dans le cas d'une distribution de charges en transition à cause d'une irradiation sous faisceau X intense, la loi d'Ohm ne suffit plus. Il faut inclure le modèle détaillé de transport de charges, en incluant le piégeage-dépiégeage. On effectue la procédure de calcul décrite dans la partie précédente pour le transport des charges pour réaliser le calcul du champ électrique. Les principales étapes sont résumées sur la figure 2.12.

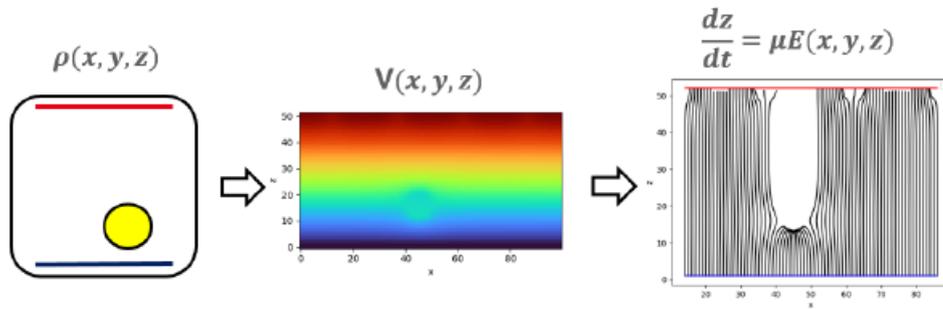


FIGURE 2.12 – Diagramme du processus de calcul du trajectoire des charges avec la présence de défauts. Exemple du défaut de conductivité sphérique.

On réalise ce processus jusqu'à stabilisation du champ électrique dans tout le volume du détecteur. On peut résumer quelques paramètres utilisés dans cette simulation dans le tableau 2.1.

Les défauts simulés de manière assez simple et géométrique ont les caractéristiques suivants, en utilisant des valeurs numériques d'ordre de grandeur de ce qui est mesuré par des méthodes de caractérisations expérimentales :

- Défauts ponctuels : densité des défauts de 100 cm^{-3} et densité de charge de $1.5 \text{ nC} \cdot \text{cm}^{-3}$ répartie de manière aléatoire sur l'ensemble de la distribution des défauts..
- Défaut plan : plan de normale $(1, 1, 1)$ et surface de densité de charge de $250 \text{ pC} \cdot \text{cm}^{-2}$.
- Défaut sphérique : sphère de diamètre 1 mm avec une conductivité 3.3 fois plus élevée que le reste du matériau soit $\rho = 1.5 \times 10^{10} \Omega \cdot \text{cm}$.
- Polarisation : irradiation par rayons X sur la cathode sur un disque de diamètre 1 mm ($10 \text{ nC} \cdot \text{cm}^{-3}$). Le flux est ajusté pour générer une charge nette de 50 pC, formé majoritairement par le piégeage des trous sous la cathode (Figures 2.13 et 2.14).

Cependant même si l'utilisation de ces simulations réalistes est parfaite pour étudier de manière plus approfondie les défauts, leur structure et leur impact sur la physique du

Paramètre	Notation	Valeur
Épaisseur du détecteur	d	5 mm
Pas de l'anode	p	2.5 mm
Potentiel appliqué	V	500 V
Mobilité des électrons	μ_e	$1000 \text{ cm}^2 \cdot \text{V}^{-1} \cdot \text{s}^{-1}$
Mobilité des électrons	μ_h	$100 \text{ cm}^2 \cdot \text{V}^{-1} \cdot \text{s}^{-1}$
Constante de piégeage des électrons	τ_e	5 μs
Constante de piégeage des trous	τ_e	50 ns
Constante d'équilibre des électrons	θ_e	10^{-4}
Constante d'équilibre des trous	θ_h	10^{-4}
Permittivité relative du matériau	ϵ_r	11
Résistivité du matériau	ρ	$5 \times 10^{10} \Omega \cdot \text{cm}$

TABLE 2.1 – Paramètres numériques de la simulation.

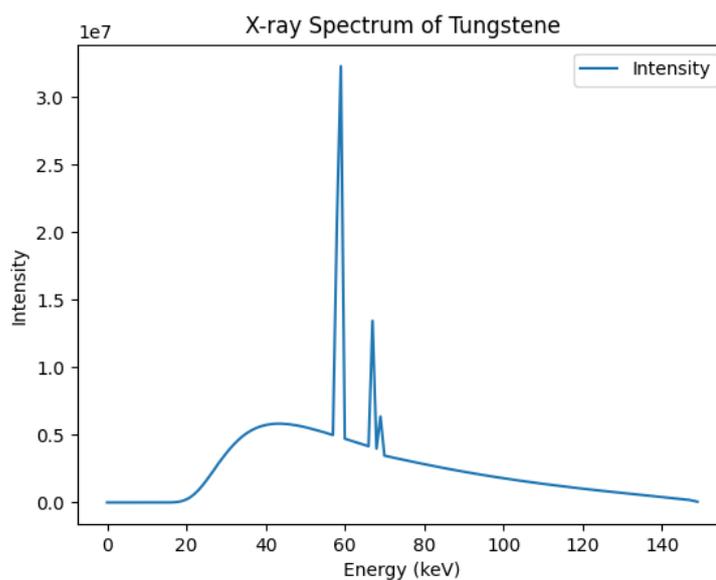
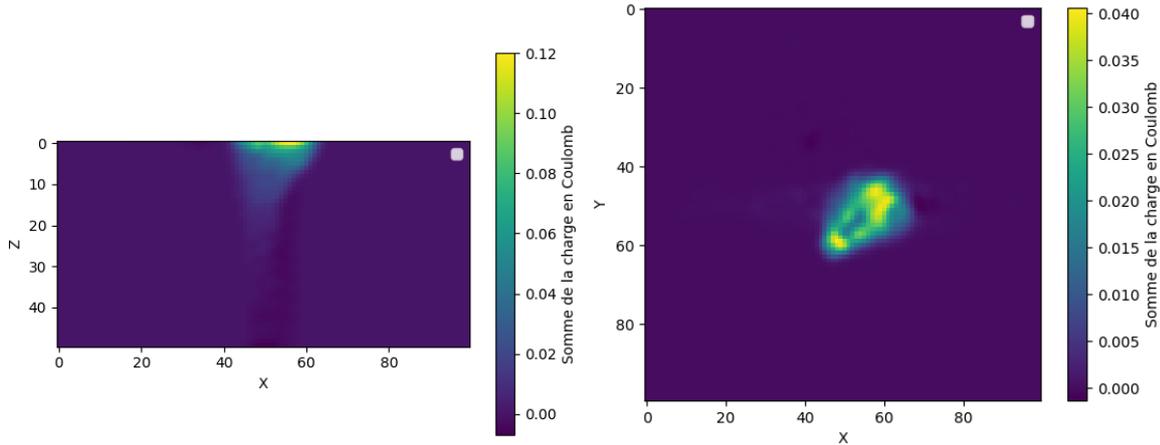


FIGURE 2.13 – Spectre de rayons X émis par une cible de tungstène soumise à un faisceau d'électrons accéléré sous 150 kV.

détecteur (conductivité, champ électrique) et son fonctionnement nous avons besoin de puissance de calcul. En effet ce genre de simulation dans un détecteur modélisé de manière non homogène avec parfois d'importantes perturbations nécessite un temps de calcul plus important, entre 5 et 40 fois supérieur en fonction de la complexité. Une des solutions pour accélérer ce processus est l'utilisation des capacités de calcul de l'unité graphique de l'ordinateur. Nous allons alors transférer les calculs de manière à les paralléliser sur la carte graphique de la machine.



(a) Histogramme 2D de la somme sur la profondeur des charges créées par le faisceau X sur la profondeur. (b) Histogramme 2D de la somme sur la profondeur des charges créées par le faisceau X sur la cathode.

FIGURE 2.14 – Histogrammes 2D de la somme sur la profondeur des charges créées par le faisceau X.

2.3 Utilisation de l'accélération GPU avec OpenCL

Dans les systèmes informatiques, le calcul parallèle permet de traiter plusieurs informations de manière simultanée limitant ainsi le temps d'exécution. Depuis quelques années, le traitement en parallèle des données a été introduit pour différentes applications, notamment pour l'imagerie médicale. Dans le but d'obtenir de hautes performances et de faciliter le développement d'applications massivement parallélisées sous unité de traitement graphique (GPU), l'utilisation du Langage OpenCL (Open Computing Language), standard de Khronos, émerge peu à peu [80] au côté de CUDA, langage propriétaire de NVIDIA.

OpenCL est un langage de programmation permettant d'effectuer des calculs généralistes sur des systèmes hétérogènes. Il combine une API (Application Programming Interface), interface permettant d'utiliser les fonctionnalités, un langage dérivé du C et des extensions. Les applications accélérées grâce à OpenCL peuvent donc, accéder à la puissance de traitement combinée des cœurs GPU (Unité du processeur graphique) (Multicœurs AMD, Intel ou Power, GPU NVIDIA, ATI), CPU (Unité centrale de traitement) ou APU (Unité de calcul accéléré) d'un ordinateur ou d'un serveur sous une même plateforme unifiée.

Un programme OpenCL est divisé en deux parties : une partie s'exécutant sur l'hôte (CPU généralement), et une s'exécutant parallèlement sur le *device* (GPU, CPU multicœurs ou FPGA). Un programme est composé de fonctions déclarées, qui vont s'exécuter sur le *device* : chaque fonction parallélisable est appelée *kernel*. Chaque *kernel* est exécuté sur un ensemble de *work-items* (équivalent d'un thread : la plus petite entité d'exécution) dont le nombre est paramétrable. Chaque *work-item* exécute le même code et l'ensemble peut coopérer au sein d'un *work-group*. Il s'agit d'un parallélisme SIMD (Single Instruction, Multiple Data) : le même code est appliqué à des jeux de données différents.

Lorsqu'un *kernel* est exécuté, on doit spécifier le nombre des *work-items* qui doit être créé en tant qu'une plage de n-dimensions, appelée *NDRange*. Un *NDRange* est un espace

indexé en un, deux ou trois dimensions, qui correspondent souvent aux dimensions des données d'entrée ou de sortie, comme illustré sur la figure 2.15.

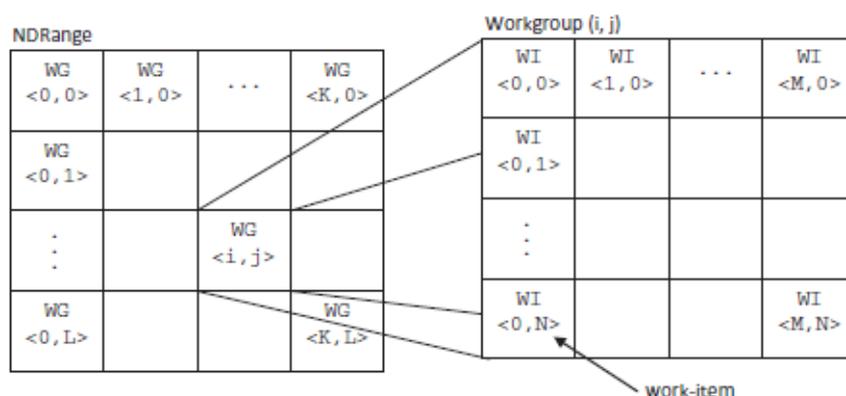


FIGURE 2.15 – Modèle d'exécution d'un programme OpenCL [81].

Le langage OpenCL peut être appelé à partir de différents langages de programmation, appelés hôte, comme C, C++, Java ou Python. L'hôte permet de communiquer entre le CPU et le GPU, et c'est grâce à cet hôte que les programmes sont compilés. L'hôte permet également la création de bloc de données (*buffers* ou image), le transfert des données et le séquençement des kernels. La parallélisation des données effectuée par OpenCL peut être illustrée par l'exemple de la figure 2.16.

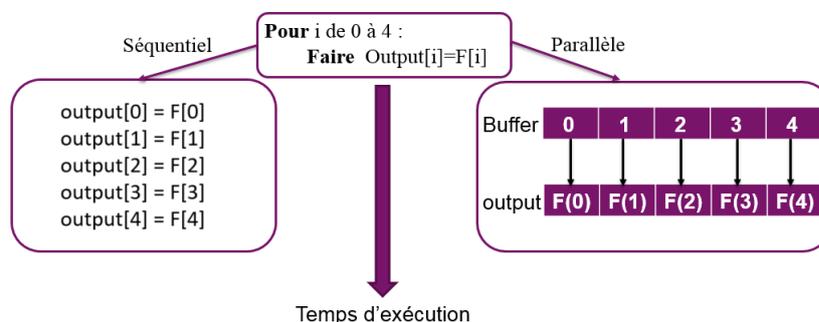


FIGURE 2.16 – Exemple d'exécution d'un programme séquentiel (à droite) et parallèle (à gauche) [82].

Pour simplifier dans un premier temps l'implémentation de la simulation par utilisation de GPU, nous limitons le nombre de paramètres comparé à la méthode séquentielle avec CPU. On décide de se focaliser sur les phénomènes physiques les plus significatifs : la dérive électronique, le piégeage des électrons et la mise en forme du signal du signal d'impulsion. Le transport des trous est négligé car comme précisé précédemment le produit $\mu\tau$ des trous dans le CZT est très faible devant celui des électrons. La diffusion du nuage électronique est prise en compte en convoluant le champ de pondération avec la distribution moyenne du nuage chargé, considérant une distribution isotropique gaussienne de $100\ \mu\text{m}$.

Chaque particule qui interagit dans le volume du détecteur est traitée séparément car considérée indépendante et n'influençant pas le comportement des autres, ainsi il est

possible de paralléliser le transport de manière individuel et ceci permet de diminuer le temps de calcul. On génère plusieurs millions d'interactions chaque seconde (typiquement 10 M/s par TFLOPS de puissance), avec leur énergie et position. Puis on calcule le transport des charges et l'induction du signal. Finalement on peut construire des histogrammes (spectre en énergie et images 2D) avec tous les événements. On peut déjà tracer les lignes de champ électrique d'après le calcul des trajectoires des charges à travers le détecteur, et observer leurs caractéristiques selon le type de défaut (Figure 2.17).

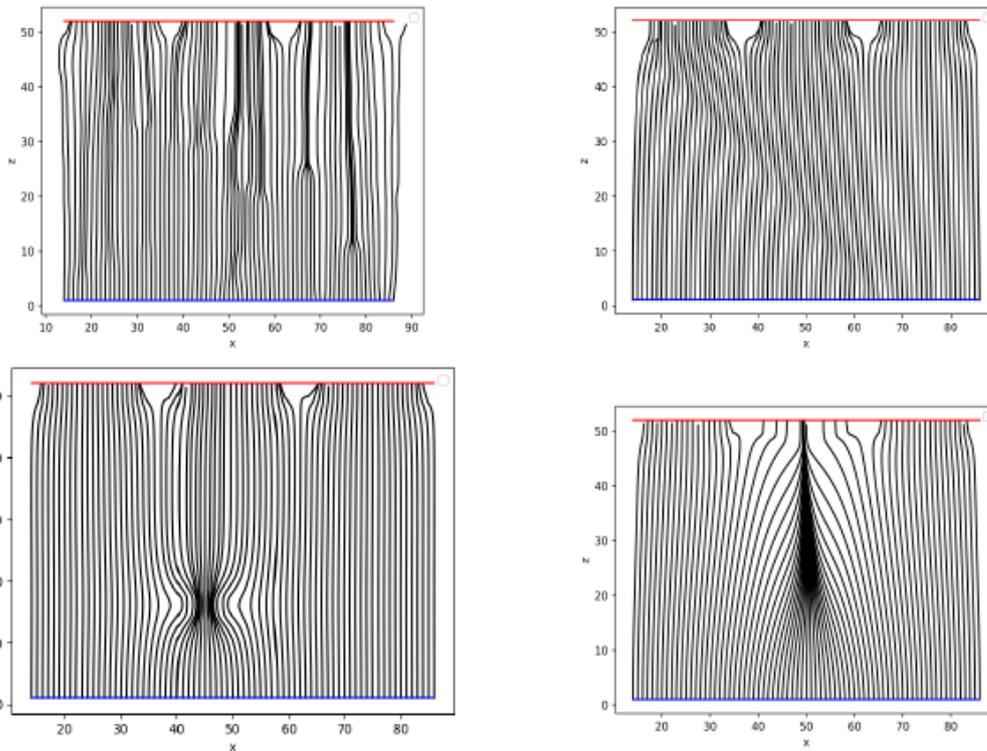


FIGURE 2.17 – Lignes de champ simulées entre les électrodes (anode en rouge, cathode en bleu) pour 4 configurations. De gauche à droite et de haut en bas : défauts ponctuels discrets ; charges distribuées selon un plan incliné à 45° ; défaut sphérique large de 1mm de diamètre de haute conductivité ; irradiation intense par rayons X sur la cathode.

- La présence de défauts ponctuels peut déformer localement les lignes de champ, dégradant la collection des charges sur les électrodes proportionnellement à la taille et profondeur du défaut.
- Un défaut plan a un effet d'attraction sur les lignes du champ, ce qui crée une légère déformation du champ électrique le long du sous-joint de grain, mais impacte de manière plus importante la collection des charges.
- Une sphère de haute conductivité a un impact important sur les lignes de champ, elle les attire et crée une accumulation de charges positives et négatives formant des dipôles qui entraîne un ralentissement de la dérive des charges, ce qui dégrade la collecte des charges.
- Une irradiation locale par rayons X resserre les lignes de champ électrique à proximité de l'anode. Cette irradiation localisée sur la cathode crée une génération importante de

paires électron-trou dans une zone restreinte du détecteur, et les mauvaises propriétés de transport des trous génèrent une zone de charge d'espace positive sous la cathode ce qui serre les lignes de champ entre les électrodes (Figure 2.18).

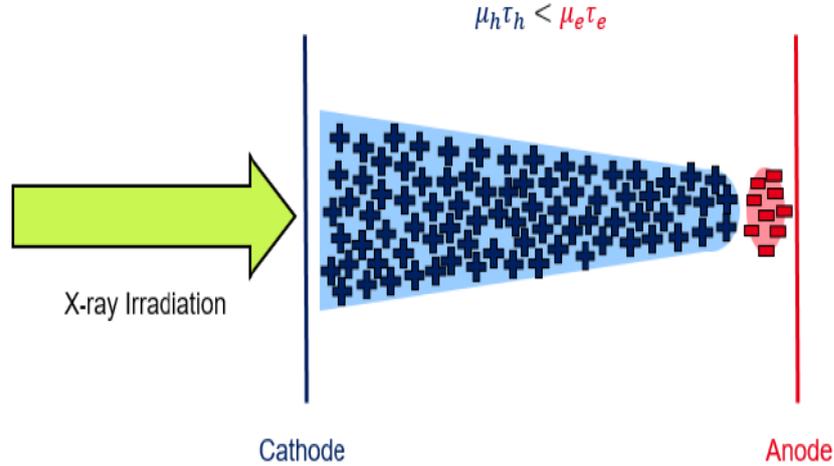


FIGURE 2.18 – Illustration du mécanisme de création de la zone de charge d'espace positive. Les différences de propriétés de transport des porteurs de charge font qu'une charge positive remplit la majorité du détecteur mis à part à proximité de l'anode.

Dans la prochaine section nous allons pouvoir observer les cartes 2D de CIE (Charge Induction Efficiency), ou efficacité de collecte d'induction, de nos simulations afin de les comparer avec la structure des lignes de champ obtenus précédemment.

2.4 Impacts des défauts sur le signal de réponse

Pour observer l'impact des défauts on décide d'utiliser un des facteurs d'évaluation de la collecte des charges qui est la CIE (Charge Induction Efficiency). La CIE représente le rapport normalisé à 1 entre la charge induite intégrée temporellement par un électron sur un pixel et la charge totale déposée par le photon lors de l'interaction.

$$CIE = \frac{Q}{Q_0} = \frac{1}{Q_0} \int_0^T \left[\iiint_{\Omega} q \mu_e n(x, t) E E_w d\Omega \right] \quad (2.28)$$

avec Ω le volume du détecteur, Q_0 la charge induite déposée par le photon lors de l'interaction, Q la charge induite intégrée par un électron sur un pixel, $n(x, t)$ la densité d'électrons à la position x et au temps t , E le champ électrique et E_w le champ de pondération. Le calcul de la trajectoire des charges simulées dans le volume du détecteur est source d'information sur la collecte aux électrodes. Ainsi on peut tracer les cartes 2D de la CIE pour étudier l'impact de la présence de défauts sur les performances du détecteur. On obtient ces cartes CIE en scannant virtuellement notre anode pixellisée avec un faisceau simulé d'une énergie de 122 keV sur une zone de 10 mm x 10 mm avec un pas de 64 μm . De manière expérimentale on peut observer des fluctuations sur les positions des frontières des électrodes et sur les performances de collecte du détecteur. On étudie la collecte sur le pixel central où l'on a appliqué le potentiel de pondération. Dans un premier temps on

trace ces cartes dans le cas d'un détecteur parfait sans défaut afin de pouvoir effectuer une comparaison avec les différents types de défauts étudiés par la suite (Figure 2.19).

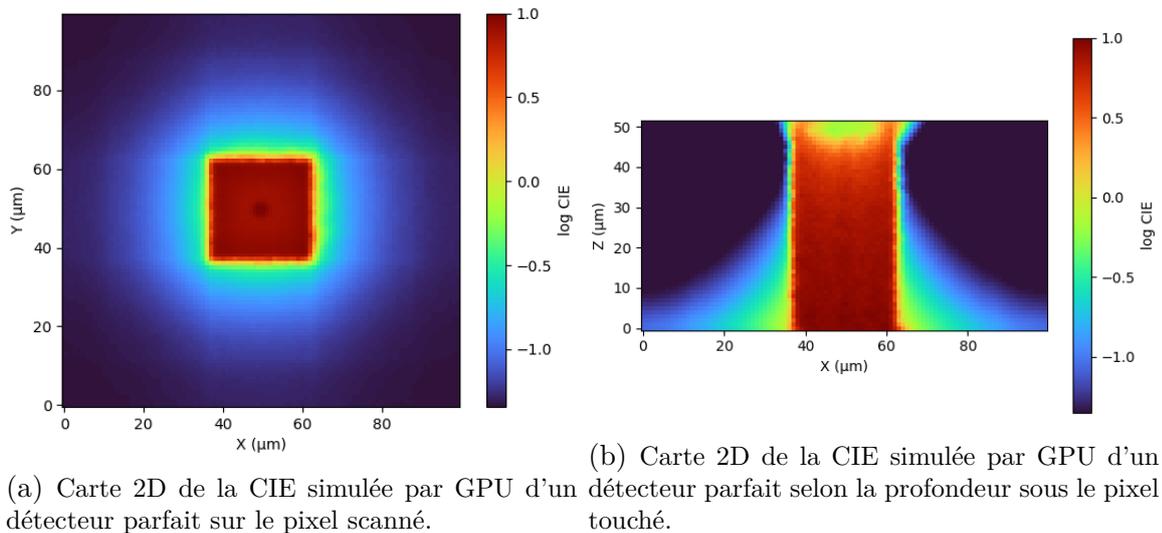


FIGURE 2.19 – Cartes CIE du détecteur parfait.

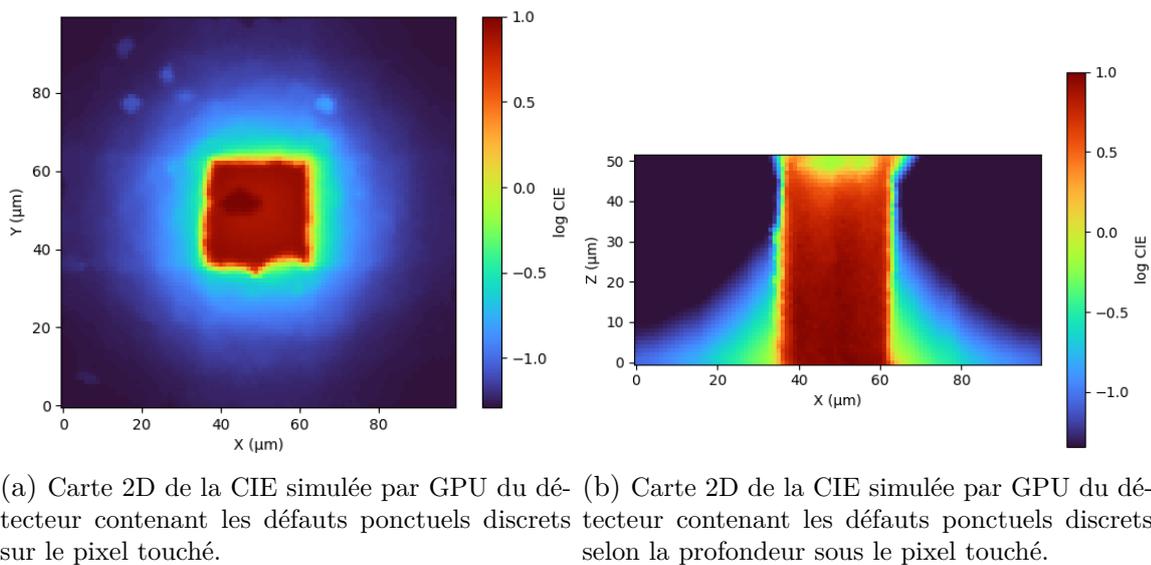
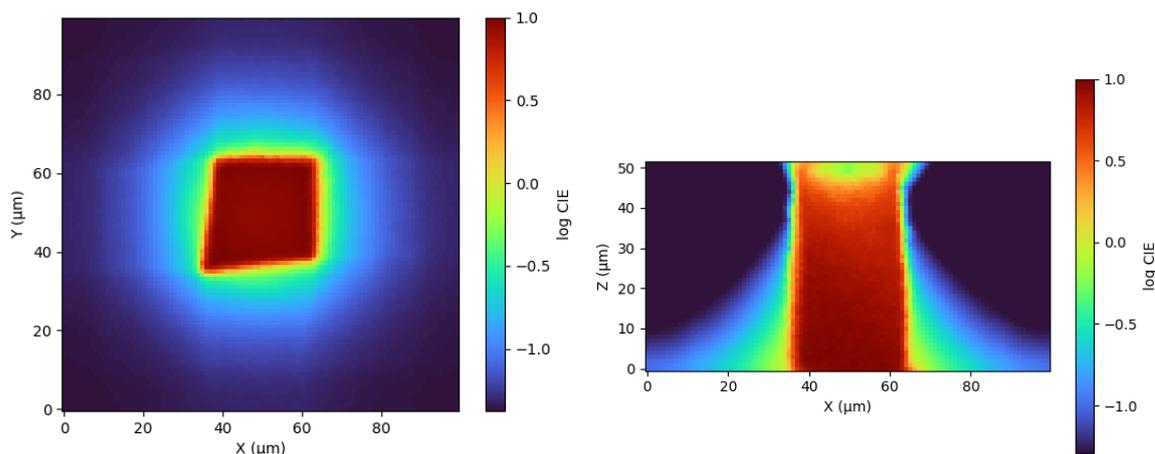


FIGURE 2.20 – Cartes CIE du détecteur avec des défauts ponctuels.

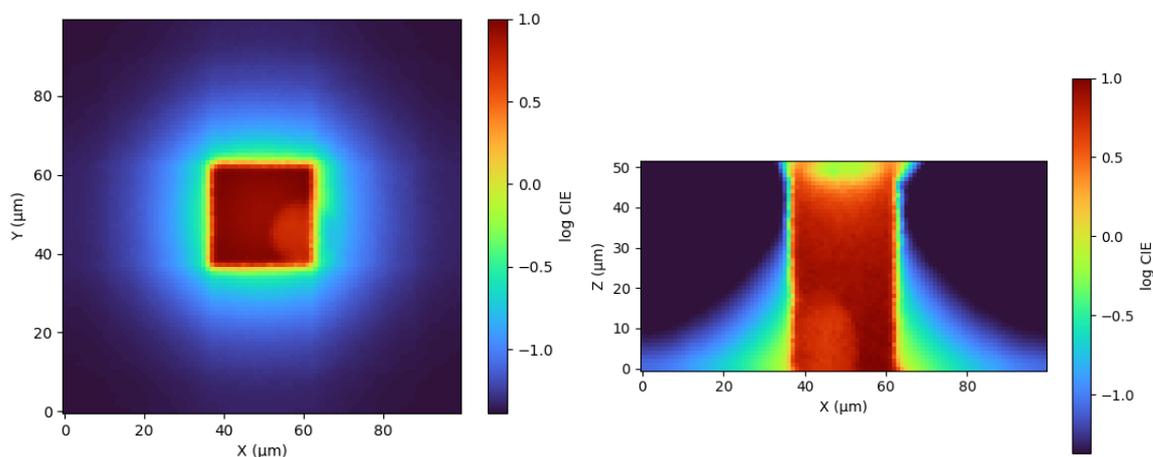
On peut remarquer que la distribution de petits défauts similaires aux inclusions Tellure que l'on peut trouver dans les cristaux CZT ont un impact localisé mais présents partout dans le volume (Figure 2.20). Les positions de collecte ne sont pas modifiées de manière importante mais les valeurs de CIE ne sont pas homogènes et on peut attendre une dégradation de la résolution en énergie et de la résolution spatiale si on utilise un algorithme de positionnement par sous-pixellisation.

Dans le cas du défaut plan, on peut observer une distorsion étendue de la collection de charge dans le coin bas gauche du pixel central due au déplacement des lignes de champ



(a) Carte 2D de la CIE simulée par GPU du détecteur contenant le défaut plan chargé sur le pixel touché. (b) Carte 2D de la CIE simulée par GPU du détecteur contenant le défaut plan chargé selon la profondeur sous le pixel touché.

FIGURE 2.21 – Cartes CIE du détecteur avec le défaut plan.



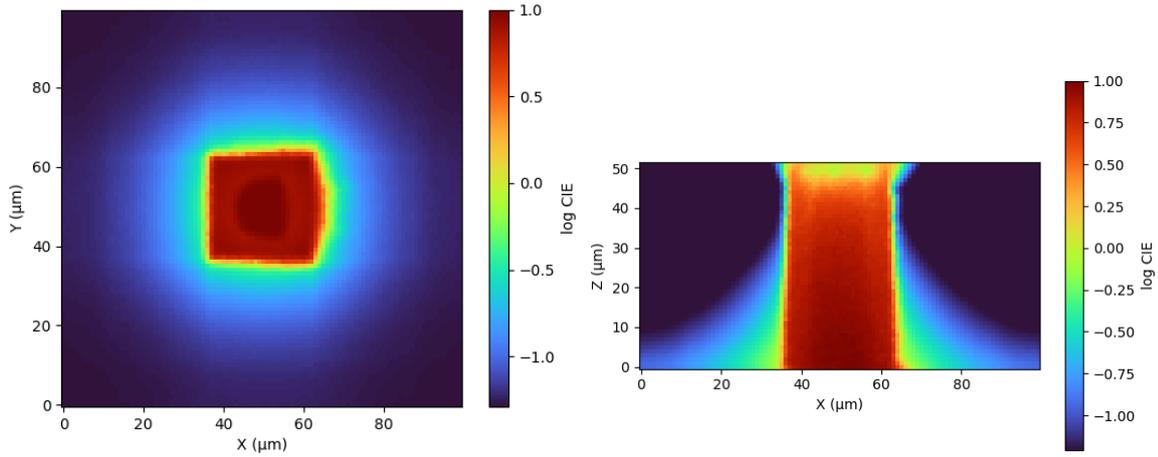
(a) Carte 2D de la CIE simulée par GPU du détecteur contenant le défaut de conductivité sphérique sur le pixel touché. (b) Carte 2D de la CIE simulée par GPU du détecteur contenant le défaut de conductivité sphérique selon la profondeur sous le pixel touché.

FIGURE 2.22 – Cartes CIE du détecteur avec un défaut sphérique.

électrique (Figure 2.21). L'effet sur le spectre sera faible mais l'uniformité de l'efficacité de comptage des photons sera affectée.

On voit clairement l'effet du défaut sphérique de haute conductivité sur la CIE du détecteur modélisé. La sphère est localisée dans la partie basse droite du pixel central proche de la cathode (Figure 2.22). Comme on peut s'y attendre d'après l'observation des lignes de champ, le défaut perturbe la trajectoire des charges ce qui est visible par une déficite de collecte.

Dans le cas de l'irradiation du détecteur par un faisceau X intense, comme cette irradiation se fait du côté de l'anode, il n'y a pas d'impact majeur sur la collecte de charge (Figure 2.23). Cependant comme les lignes de champ sont resserrées l'impact sur le positionnement et sur la qualité de l'image peut être important. De plus une irradiation

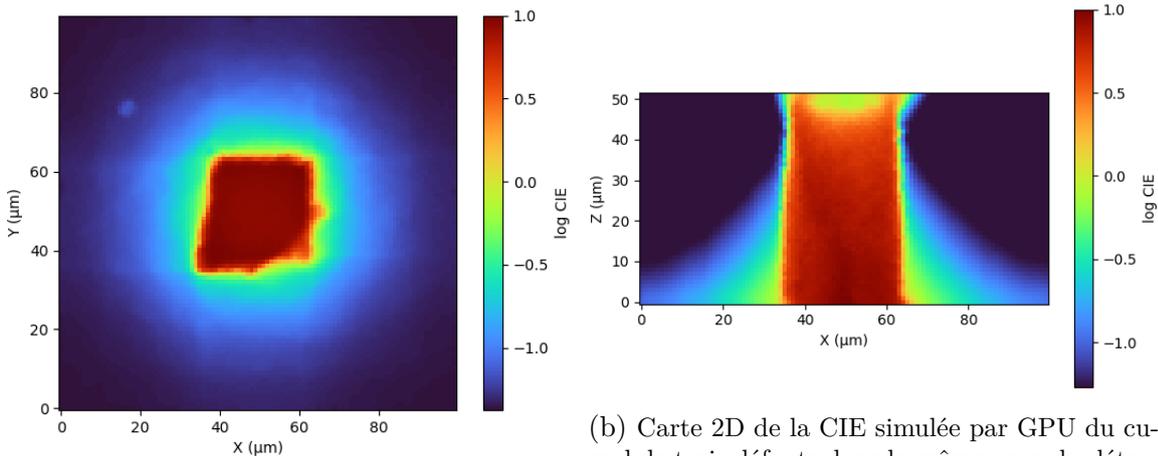


(a) Carte 2D de la CIE simulée par GPU du détecteur irradié par une faisceau X intense sur le pixel touché. (b) Carte 2D de la CIE simulée par GPU du détecteur irradié par une faisceau X intense selon la profondeur sous le pixel touché.

FIGURE 2.23 – Cartes CIE du détecteur soumis à un faisceau X intense.

avec un faisceau très intense peut « pincer » le champ de manière plus intense et créer une zone morte dans le détecteur.

On peut également s’intéresser au cumul des différents types de défauts dans un détecteur, situation que l’on rencontre en réalité car chaque type de défaut peut être présent à tout endroit du détecteur, et se cumuler dans l’espace pour générer encore plus d’incertitudes sur la collecte aux électrodes du détecteur. Par exemple on peut s’intéresser à l’addition du défaut plan, des défauts ponctuels et de la sphère de la conductivité au niveau du pixel central et observer la réponse de la CIE que nous obtenons grâce à la simulation.



(a) Carte 2D de la CIE simulée par GPU du cumul de trois défauts dans la même zone du détecteur : des défauts ponctuels, un défaut plan et une sphère de haute conductivité. (b) Carte 2D de la CIE simulée par GPU du cumul de trois défauts dans la même zone du détecteur : des défauts ponctuels, un défaut plan et une coupe latérale selon la profondeur du détecteur sous le pixel touché.

FIGURE 2.24 – Cartes CIE du détecteur regroupant plusieurs types de défauts.

On observe que le cumul de défauts localisés sur le pixel collectant a un impact important sur la déformation et les pertes de collecte. Ceci permet d'illustrer le fait qu'au delà de l'intensité des défauts c'est probablement leur variété de présence dans certaines zones du détecteur qui dégrade de manière plus intense la collecte du signal sur les électrodes, ce qui amène à des défauts de localisation inévitables (Figure 2.24).

2.5 Conclusion

Le développement de simulations réalistes commence par une compréhension des modèles analytiques que l'on peut appliquer de manière unidimensionnelle ou tridimensionnelle. Néanmoins ces modèles étant limités dans l'intégration de paramètres supplémentaires, il est nécessaire de considérer une approche de modélisation avec résolution numérique grâce à des outils plus performants. L'utilisation de l'outil Tasmania permet l'introduction de différents types de défauts en utilisant des méthodes de résolution basées sur des itérations numériques du transport des charges accélérées par l'utilisation de l'unité graphique de l'ordinateur permettant d'aboutir à son calcul. Le calcul du champ électrique conséquent à cela permet alors de s'intéresser aux résultats de ces simulations comportant des défauts. On évalue de différentes manières l'impact des défauts en observant qualitativement l'effet sur les lignes de champ électrique ou quantitativement en calculant la CIE du voisinage de pixels simulé.

Cette simulation va nous servir par la suite dans les deux autres briques de notre processus : tout d'abord afin d'entraîner un réseau de neurones dense à champ électrique fixe pour faire de l'apprentissage supervisé sur la localisation des interactions, puis dans un second temps dans le but d'apprendre les paramètres de champ d'un détecteur effectif de manière non supervisé. Les simulations vont permettre de faire le relais entre ces deux processus afin d'aboutir sur une utilisation directe de l'apprentissage de position sur données expérimentales.

Apprentissage supervisé de la position par réseau de neurones

Un des objectifs de la thèse est de développer une alternative plus efficace et plus flexible à l'estimation de position et d'énergie des interactions par maximum de vraisemblance. Une limitation technique à laquelle nous devons faire face est que cette nouvelle méthode doit être implémentable électroniquement et doit donc rester assez simple avec un nombre d'opérations limité (de l'ordre de quelques centaines de multiplications par photon détecté typiquement). Les simulations développées dans la partie précédente étant fidèles à la réalité, on peut les utiliser afin d'entraîner un réseau de neurones de manière plus efficace qu'avec des données expérimentales avec des données semblables à la réalité. Dans cette partie on s'intéresse au développement d'un réseau de neurones dense assez simple permettant de calculer la position d'interaction ainsi que l'énergie déposée à partir des signaux induits mesurés en sortie de détecteur. On développera dans ce chapitre les avantages par rapport à la méthode actuelle, la structure du réseau de neurones utilisé ainsi que les résultats obtenus pour un apprentissage sur simulation et un test sur des données expérimentales.

3.1 Localisation par sous-pixellisation

3.1.1 Théorie mathématique

Les détecteurs que nous utilisons possèdent tous une anode segmentée en pixels, chacun d'entre eux comportant une voie de lecture électronique. Le but de la localisation spatiale est de déterminer la position d'interaction d'un photon gamma dans le détecteur CZT avec une précision inférieure à la taille du pixel (Figure 3.1). Dans le cas de l'anode pixellisée on peut utiliser les amplitudes des signaux transitoires du pixel touché et de ses pixels voisins. Une des premières méthodes utilisées se servait d'un calcul de barycentre entre le pixel à droite du pixel touché et celui à gauche pour affiner la position [83]. Il existe plusieurs algorithmes de traitement de ces signaux pour calculer la position d'interaction, La méthode couramment utilisée est basée sur le maximum de vraisemblance.

Pour résumer le principe du maximum de vraisemblance, si on cherche à estimer un

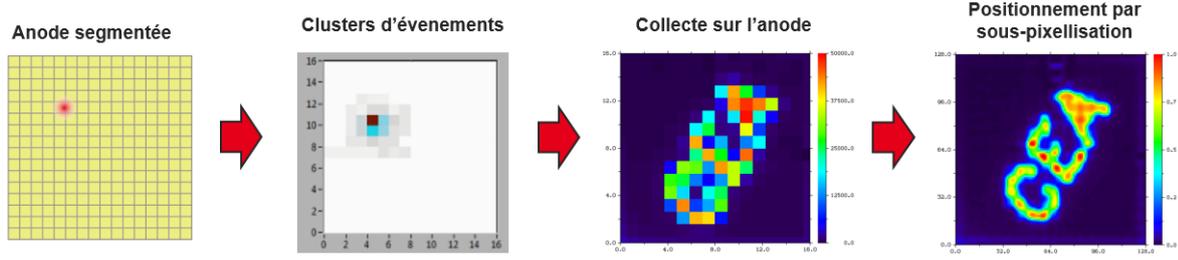


FIGURE 3.1 – Schéma de principe du fonctionnement de la sous-pixellisation pour une anode segmentée de 16x16 pixels et de 2.5mm de largeur : les signaux des différentes anodes sont combinés pour aboutir à une résolution à plus grande précision.

paramètre ou ensemble de paramètres θ , et on a $M = (m_0, \dots, m_n)$ une série de n mesures statistiquement indépendantes dont on connaît la densité de probabilité $f(m_i|\theta)$ sachant la valeur du paramètre θ , alors on définit la vraisemblance telle que

$$L(\theta) = f(m_1|\theta) \times f(m_2|\theta) \times \dots \times f(m_n|\theta) = \prod_{i=1}^n f(m_i|\theta) \quad (3.1)$$

L'estimation de θ , notée $\hat{\theta}$, se fait alors en déterminant la valeur de θ qui maximise la fonction de vraisemblance

$$\hat{\theta} = \arg \max(L(\theta)) \quad (3.2)$$

Dans le cas de la localisation, θ représente soit la variable x , y ou z lorsqu'on estime les coordonnées de manière indépendante, soit le vecteur (x, y, z) . M représente les mesures effectuées sur les n pixels.

$$M = (m_0, \dots, m_n). \quad (3.3)$$

Dans le cas le plus simple, on a besoin de réaliser deux mesures par pixel pour le calcul de la position x (notre paramètre θ) par maximum de vraisemblance : le voisin de droite et le voisin de gauche. Comme $f(m_i|\theta)$ suit une distribution supposées gaussienne et indépendante on peut supposer que $f(\theta|m_i)$ également. L'hypothèse gaussienne provient du fait que la distribution du bruit électronique est effectivement gaussienne, celui-ci étant produit par un grand nombre d'événements indépendants : chaque mouvement élémentaire de porteur de charge dans le détecteur ou le transistor d'entrée contribue au bruit. De plus, les bruits des voies électroniques sont dans une large mesure indépendants les uns des autres, même si une certaine corrélation du bruit de grenaille en haute fréquence est créée par le phénomène de partage d'induction [76]. On va également considérer que la variance V du bruit est égale sur les différentes voies, on suppose en effet que la distribution gaussienne possède les mêmes paramètres sur chaque voie. Si on utilise la modélisation des signaux par les moyennes μ_1 et μ_2 et en considérant donc la variance uniforme, la vraisemblance résultante est :

$$f(\theta|M) = f(\theta|M_1)f(\theta|M_2) = e^{-\frac{1}{2V}[(M_1-\mu_1(\theta))^2+(M_2-\mu_2(\theta))^2]}. \quad (3.4)$$

Si l'on considère le bruit suffisamment faible pour linéariser autour de θ , on obtient la vraisemblance :

$$f(\theta|M) = f(\theta|M_1)f(\theta|M_2) = e^{-\frac{1}{2\sigma_1^2}(\mu_1^{-1}(M_1)-\theta)^2} \cdot e^{-\frac{1}{2\sigma_2^2}(\mu_2^{-1}(M_2)-\theta)^2} \quad (3.5)$$

Or dans le cas à un seul paramètre à estimer :

$$\sigma_i = \sqrt{V} \frac{d\theta}{dm_i} = \frac{\sqrt{V}}{\frac{dm_i}{d\theta}} = \frac{\sqrt{V}}{\mu'_i}. \quad (3.6)$$

D'où

$$f(\theta|M) = e^{-\frac{1}{2V}\mu_1'^2(\mu_1^{-1}(M_1)-\theta)^2} \cdot e^{-\frac{1}{2V}\mu_2'^2(\mu_2^{-1}(M_2)-\theta)^2}. \quad (3.7)$$

Ainsi

$$\frac{df(\theta|M)}{d\theta} = 0 \Leftrightarrow \sum_{i=1}^2 \mu_i'^2 [\mu_i^{-1}(M_i) - X] = 0. \quad (3.8)$$

On obtient finalement l'estimateur de position

$$\hat{\theta} = \frac{\mu_1'^2 \mu_1^{-1}(M_1) + \mu_2'^2 \mu_2^{-1}(M_2)}{\mu_1'^2 + \mu_2'^2}. \quad (3.9)$$

3.1.2 Application à la localisation 3D

La méthode ML (Maximum-Likelihood) permet d'affiner la mesure de position en appliquant un traitement à la CIE mesurée en approchant cette dernière par une fonction liée aux trois coordonnées [84]. Cette technique permet de déterminer la position d'interaction la plus probable d'après les mesures obtenues et leurs calibrations. On utilise un paramétrage en fonction de la distance à la limite du pixel de l'intensité des signaux relatifs des voisins (normalisé avec le signal de l'anode collectante Q_0) que l'on ajuste avec la formule empirique :

$$\frac{Q}{Q_0} = f(X) = e^{-aX^b} \quad (3.10)$$

Durant un stage effectué au laboratoire, Damien Barbes a travaillé sur la mise en place d'une méthode de localisation par maximum de vraisemblance utilisant l'équation (3.10) [85]. Il a notamment calculé les valeurs des coefficients a et b pour chaque pixel du détecteur, et tracé des courbes d'ajustement de l'amplitude des anodes, comme on peut voir sur la Figure 3.2.

La figure 3.3 montre les trois courbes de signaux induits (nommés *center* pour le pixel collectant, *left* pour le pixel à sa gauche et *right* pour celui sur sa droite). On remarque que l'effet des pixels voisins est surtout visible sur les bords du pixel, dû au partage d'induction plus important que l'on observe à ces endroits. Les trois courbes d'amplitude situées vers le bas montrent un exemple de signaux que l'on utilise pour la méthode ML afin d'affiner la localisation de l'interaction selon les axes X et Y à partir des signaux mesurés sur les pixels voisins. Ici on remarque que le signal que l'on estime sur le pixel touché se localise sur le bord droit du pixel, à cause de l'amplitude observée plus importante sur l'électrode voisine située de ce côté.

Cette méthode unidimensionnelle a été étendue à trois dimensions : on approxime la valeur de CIE à un endroit du détecteur comme composées de fonctions dépendantes des coordonnées. Pour effectuer une estimation tridimensionnelle [86], l'approche consiste à

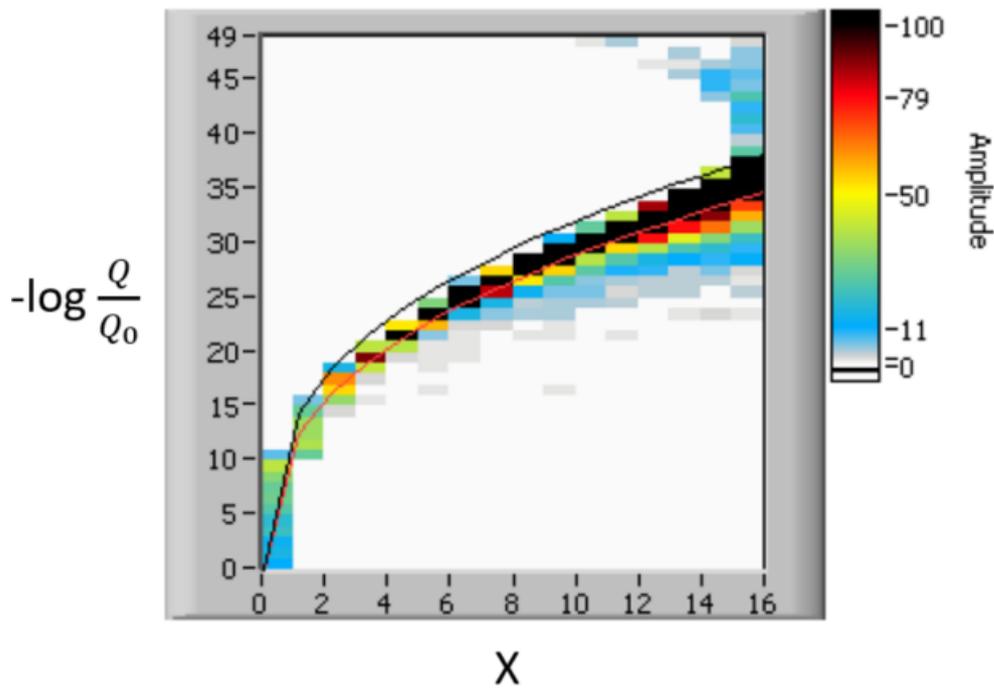


FIGURE 3.2 – Ajustement de l’amplitude relative d’une anodes voisine en fonction de la distance X à l’anode collectante. La courbe rouge représente l’ajustement retenu, la noire est la dernière configuration testée. L’abscisse donne la position X dans le pixel, l’ordonnée donne la valeur de l’atténuation de l’amplitude observée [85]. L’ajustement est représenté par la courbe rouge.

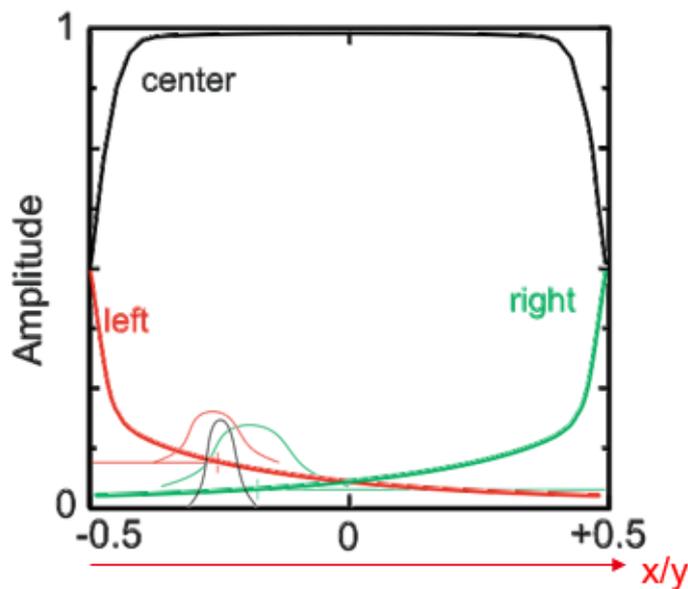


FIGURE 3.3 – Exemples de signaux mesurés sur le pixel central et ses deux voisins de droite et gauche. La méthode ML permet d’affiner la localisation sur le pixel central (courbe noire).

utiliser comme distance X non plus une seule direction latérale mais une décomposition simple en fonction des coordonnées 3D (x, y, z) :

$$X(x, y, z) \approx (g(x) + h(y)) \cdot k(z) \quad (3.11)$$

Cette décomposition est une hypothèse simplificatrice permettant le calcul direct (non itératif) en permettant de mettre le problème sous forme de système linéaire. Les fonctions g et h sont adaptées en fonction de la position de l'anode considérée par rapport à l'anode collectante. La version la plus simple de cette formulation est $g(x) = x$ ou $1 - x$ ou 0 , $h(y) = y$ ou $1 - y$ ou 0 et $k(z) = 1/z$. De cette manière, X est une combinaison linéaire de fonctions de x, y, z dont on peut analytiquement déduire la position. Ces fonctions forment un vecteur paramètre Θ à partir duquel on écrit le vecteur X_i sous forme d'un produit matriciel

$$X_i = \mathbf{M}_i \cdot \Theta$$

$$\frac{Q_i}{Q_0} = CIE(x - x_i, y - y_i, z) = f[\mathbf{M}_i \cdot \Theta]. \quad (3.12)$$

Ensuite, on effectue un maximum de vraisemblance grâce à ce modèle de la charge sur le pixel i :

$$R_i = \frac{Q_i}{Q_0} = CIE(x - x_i, y - y_i, z) + \mathcal{N} \quad (3.13)$$

avec E l'énergie du photon et \mathcal{N} le bruit gaussien de la mesure.

Le fait que le bruit soit gaussien, indépendant et égal sur les différentes voies correspond en maximum de vraisemblance à la minimisation de l'erreur quadratique :

$$\epsilon = \sum_i [R_i - CIE(x - x_i, y - y_i, z)]^2. \quad (3.14)$$

On utilise cette méthode et notamment l'expression de la CIE sous cette forme pour plusieurs raisons. Comme décrit précédemment, le maximum de vraisemblance est un estimateur statistique permettant de déterminer les paramètres optimaux maximisant la fonction de vraisemblance à partir d'une série de mesures. Cette formulation permet d'obtenir une solution de l'équation

$$\nabla \epsilon = 0$$

avec ϵ l'erreur entre estimations et mesures de manière linéaire. Une solution directe est donnée par l'équation vectorielle

$$\sum_i f^{-1}(R_i) f'(R_i)^2 \mathbf{M}_i = \left[\sum_i \mathbf{M}_i^t f'(R_i)^2 \mathbf{M}_i \right] \cdot \Theta$$

Ainsi en séparant les variables de cette manière on peut aboutir à un système linéaire dont les inconnues sont liées aux variables x, y, z . Ce processus est non itératif ce qui rend son exécution plus rapide et moins exigeante en puissance de calcul.

Ce type de méthode de localisation présente néanmoins quelques limites :

- Une correction d'uniformité est nécessaire pour compenser les défauts du détecteur qui varient de pixel en pixel. Il faut d'ailleurs noter que même un estimateur non biaisé produira des résultats non uniformes du fait de la non stationnarité de la variance [75].

- L'hypothèse de bruit faible utilisée par la linéarisation n'est pas nécessairement vérifiée, en particulier à basse énergie.
- Le modèle reste simpliste avec une séparation des trois coordonnées. Le modèle de variation selon la profondeur z multiplicatif et la combinaison additive des variations en x et y ne sont pas exacts [86]. Seules les approches itératives permettent d'utiliser des modèles plus précis.

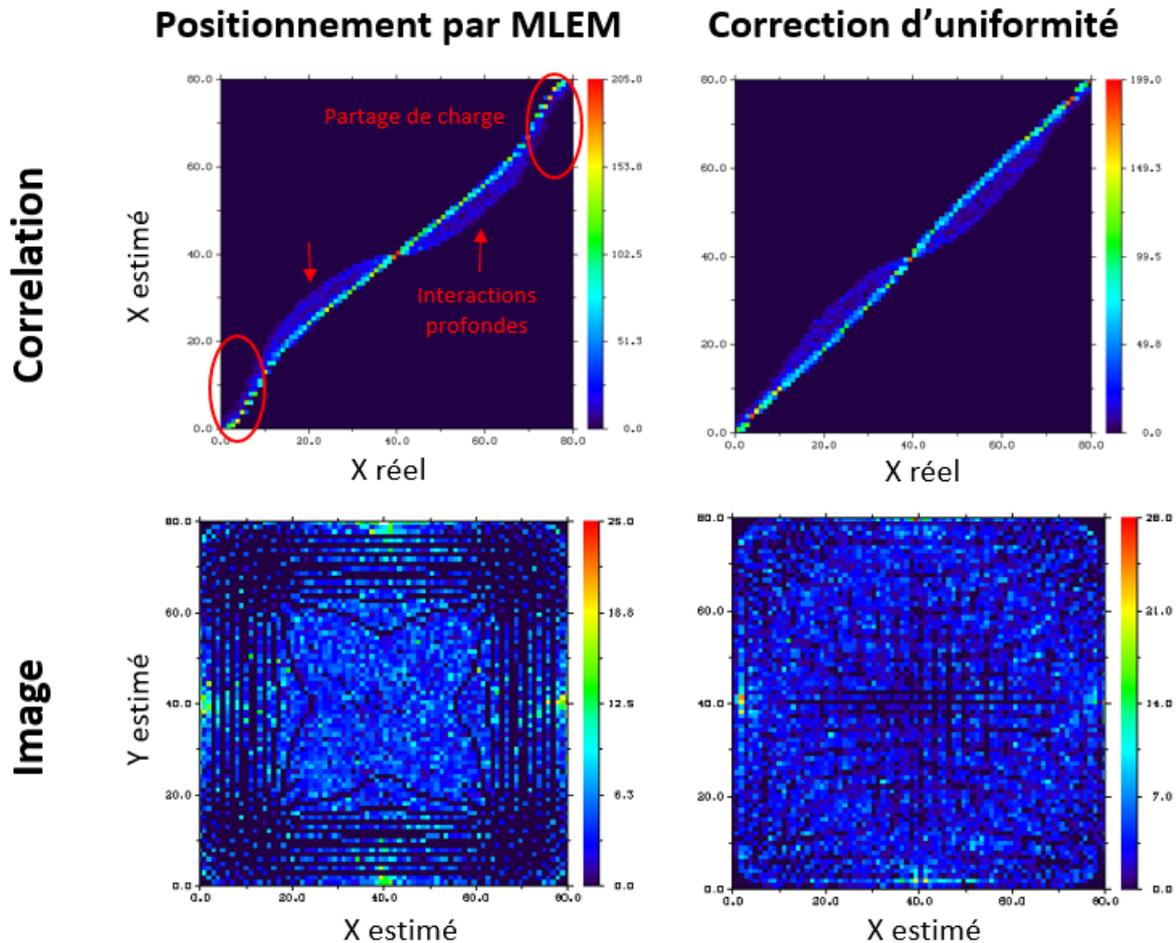


FIGURE 3.4 – Courbe de corrélation (en haut) et image obtenue (en bas) par la méthode du maximum de vraisemblance avant (gauche) et après (correction d'uniformité).

On peut remarquer sur la figure 3.4 la présence de deux phénomènes physiques inhérents au détecteur qui impactent l'estimation de position. Le premier est le partage de charge aux frontières des pixels, qui diminue la précision sur la localisation car une partie de la charge induite sur ce pixel se retrouve dans l'espace inter-pixel voire le pixel voisin. Le second effet est la présence d'interactions profondes. En effet ces dernières induisant peu de signal, lorsqu'elles ont lieu vers le centre de l'électrode le signal induit sur les pixels voisins est quasi nul ce qui limite l'utilisation de la sous-pixelisation. Enfin la présence des non uniformités dans le cristal du détecteur entraîne des difficultés pour estimer correctement la position d'interaction.

3.2 Structure du réseau de neurones

Une des solutions qui permet de s'affranchir de ces limites tout en gardant un estimateur non itératif est l'utilisation d'un réseau de neurones, qui permet d'améliorer la justesse mathématique du modèle, augmenter la résolution sans requérir de résolution itérative et qui est donc facilement implémentable dans l'électronique de traitement. Cette approche a été jusque là très peu étudiée, mais quelques publications récentes ont proposé des solutions de modélisation et d'estimation de position par réseau de neurones.

Dans son article, Banerjee [87] présente un modèle physique de détecteur semi-conducteur dont les paramètres sont déterminés par apprentissage. La technique se base sur une simulation afin d'ajuster ces paramètres qui sont calculés par voxel, comme le champ électrique, les niveaux de pièges ou les propriétés des porteurs de charge. L'usage d'une telle approche pour un usage pratique semble cependant limité : certaines données utilisées pour l'apprentissage supervisé ne sont pas accessibles expérimentalement. Cela n'est donc applicable qu'à des cas simulés.

De son côté, Yang [88] propose une application plus concrète en proposant d'utiliser un réseau de neurones afin de corriger le partage de charge et l'estimation de position d'un détecteur. Il s'agit d'utiliser en apprentissage supervisé la dépendance observée des amplitudes mesurées à énergie et position du photon incident. Il exploite les données expérimentales en considérant en entrée de leur réseau les amplitudes mesurées sur un voisinage de neuf pixels, et en sortie ils classent la perte d'énergie, la position ou les deux simultanément pour effectuer une correction. L'étude porte notamment sur trois architectures différentes de réseau afin de comparer leurs performances (Figure 3.5). Contrairement à notre cas où nous employons la régression, cette approche utilise la classification qui nécessite davantage de neurones de sortie (typiquement un par sous-pixel) et donc plus de ressources *in fine*. Nous pensons que l'apprentissage peut également être plus difficile et moins robuste du fait de l'indépendance entre les neurones de sortie et le fait que davantage de poids sont à déterminer.

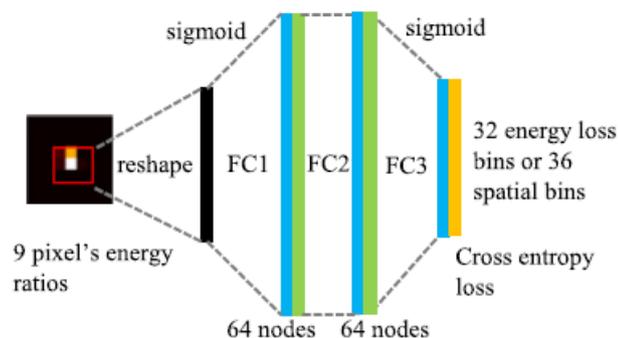


FIGURE 3.5 – Réseau de neurones utilisé pour l'estimation de la perte d'énergie par partage de charge ou le positionnement par sous pixelisation. La sortie correspond au résultat par classification de la perte d'énergie ou du positionnement [88].

3.2.1 Fonctionnement d'un réseau de neurones pour l'apprentissage supervisé

Le domaine du machine learning, ou apprentissage automatique, est actuellement en plein essor car son application dans tous les champs connus pourrait révolutionner la manière de réfléchir et résoudre les problèmes, notamment pour le traitement des données. Il est basé sur plusieurs principes mathématiques tel que la dérivation automatique (ou *automatic differentiation* en anglais). Cette technique permet de calculer la dérivée partielle ou totale d'une fonction définie de manière numérique. Les avantages de cette méthode par rapport à la résolution numérique par différence finie notamment sont sa précision, sa robustesse et sa rapidité d'exécution. D'autres avantages de l'utilisation d'un algorithme d'apprentissage automatique sont par exemple la possibilité de travailler avec des valeurs bornées ce qui va permettre une convergence de manière plus sûre. Il y a également le principe de généralité de ce type de technique, ce qui signifie que l'on peut utiliser des structures similaires voire identiques pour traiter différents types de données, prouvant la robustesse et l'adaptabilité de ce genre d'apprentissage. Le deep learning, ou apprentissage profond, est une branche du machine learning utilisant des réseaux neuronaux avec un grand nombre de couches inspirés du fonctionnement du cerveau humain afin de résoudre des problèmes. On retrouve ces réseaux dans énormément de domaines d'applications, comme les voitures autonomes, le traitement d'images, la reconnaissance faciale ou encore la traduction de langues. [89].

L'apprentissage supervisé est le système d'apprentissage le plus populaire et le plus simple à mettre en place en intelligence artificielle. On utilise un jeu de données dans lequel chaque entrée x est associée à une valeur y . Le but du réseau dans ce système d'apprentissage supervisé est de trouver la fonction f reliant x et y en lui faisant analyser un nombre très important de couples (x, y) et de réaliser un ajustement paramétrique ; on appelle ces paramètres des poids notés w . Après apprentissage des poids le réseau pourra nous donner la valeur y correspondant à une entrée x quelconque.

La structure et le fonctionnement (Figure 3.6) d'un réseau de neurones reste pratiquement toujours les mêmes quel que soit les applications et le type de données traitées :

1. Les données passent dans une couche d'entrée dont la taille correspond à la taille des données $x = u^{(0)}$.
2. Un enchaînement de couches cachées pouvant opérer différents types d'opérations modifient les données. Chaque neurone pondéré de la couche prend en entrée les sorties de la couche précédente et applique une fonction d'activation f et un biais additif b afin d'obtenir une valeur de sortie.

$$u_j^{(k+1)} = f \left(\sum_i w_{ij} u_i^{(k)} + b_j \right) \quad (3.15)$$

3. Un vecteur des valeurs finales est récupéré après la couche de sortie $y = u^{(n)}$.
4. On réalise un calcul d'erreur entre les valeurs estimées par le réseau \hat{y} et les données réelles y auxquelles on souhaite arriver. Si cette erreur est inférieure à la précision désirée l'apprentissage est terminé, sinon on rétropropage le gradient de cette erreur à l'intérieur du réseau. Ainsi les poids correspondant de chaque neurone sont mis à jour et on peut procéder à un nouveau passage dans le réseau, correspondant à

une itération. L'erreur est par exemple une erreur quadratique $\sum_i (y_i - \hat{y}_i)^2$. On peut tracer un graphe de l'évolution de la fonction de perte, correspondant dans notre cas à l'erreur, en fonction de l'avancée de l'apprentissage comme sur la figure. 3.7.

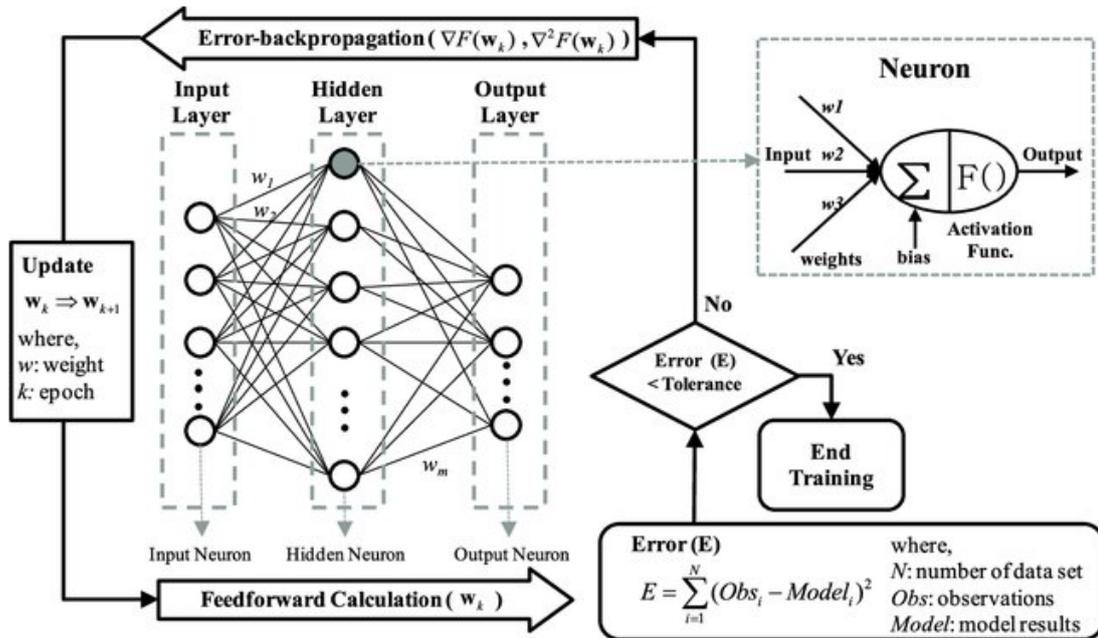


FIGURE 3.6 – Schéma de fonctionnement d'un réseau de neurones [90]

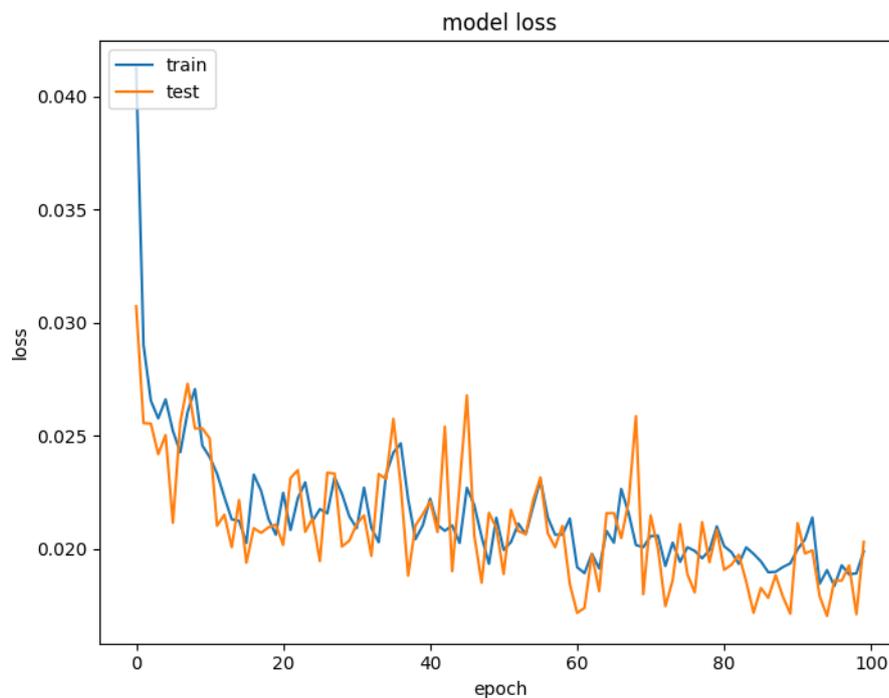


FIGURE 3.7 – Graphique de la valeur de la fonction d'erreur en fonction du nombre d'itérations (*epoch*) pour les données d'entraînement (en bleu) et de test (en orange).

3.2.2 Architecture du réseau utilisé

Dans le cas de notre étude, la structure de nos données est relativement simple. Les données d'entrée correspondent à l'amplitude des signaux induits sur le pixel touché et ses voisins normalisée à 1. Cette normalisation est obtenue, soit en divisant par la valeur maximale ou par la valeur obtenue sur le pixel central, auquel cas cette dernière valeur n'est pas incluse dans les entrées du réseau. Pour chaque pixel nous avons une ou deux valeurs : une valeur transitoire minimale et une valeur transitoire maximale mesurée pour chaque impulsion. D'autre part on choisit d'utiliser les quatre ou huit premiers voisins. Ainsi le vecteur d'entrée pour chaque interaction a une taille de 4 à 18, selon la configuration. Les données de sortie correspondent simplement à la position d'interaction du photon en trois dimensions.

Pour implémenter notre réseau de neurones on utilise l'outil Tensorflow développé pour l'apprentissage automatique, ainsi que l'API¹ Keras qui permet la construction de réseau de neurones de manière assez simple. On procède en quatre étapes distinctes :

1. La préparation de nos données afin qu'elles soient adaptées à la forme du réseau, autrement dit que le vecteur d'entrée et le vecteur de sortie aient les mêmes tailles que les couches d'entrée et de sortie.
2. La construction du réseau en elle même avec l'ajout des différentes couches neuronales.
3. La compilation du modèle en choisissant les fonctions d'optimisation et de perte.
4. L'entraînement du modèle et le test sur une partie des données.

Préparation des données

Que l'on soit dans le cas de données simulées ou expérimentales, la préparation des données sera similaire. Comme décrit précédemment il faut associer à un vecteur de signaux induits la position associée. On va alors séparer en 3 groupes les données de manière aléatoire :

1. 70% des données vont être utilisées comme données d'entraînement. Cela correspond à la partie qui va être utilisée pour entraîner le modèle et mettre à jour les poids du réseau jusqu'à leur valeur correcte.
2. 15% des données seront exploitées pour tester le modèle. Après l'apprentissage on va entrer dans le réseau le vecteur de signaux induits sur les pixels et on va observer si celui-ci arrive à estimer les positions d'interaction associées.
3. 15% des données sont consacrées à la validation du modèle. Cette étape n'est pas indispensable mais permet d'éviter le sur-apprentissage en mettant à jour les hyperparamètres de l'architecture du modèle. Par exemple lorsque l'utilisation du jeu d'entraînement aboutit sur une indécision concernant le choix d'un classificateur, c'est le rôle du jeu de validation de faire ce choix en comparant les performances.

On peut également présenter le concept de validation croisée, qui consiste à diviser le jeu de données plusieurs fois pour obtenir différents jeu d'entraînement, de validation

1. *Application Programming Interface* : interface externe offerte par un logiciel pour s'intégrer à d'autres logiciels.

et de test de telle manière à renforcer la stabilité du réseau de neurones. Pour limiter la présence d'un biais lors de la séparation des données en ces différents groupes on mélange le jeu pour le regrouper ensuite [91].

Construction du réseau

Les données dont nous disposons étant relativement simples, on démarre par une couche d'entrée de taille 4 à 18 selon le nombre de voisins et de données par voisin que l'on considère, on se contente d'un réseau à trois couches cachées denses d'une taille de 16 chacune et la couche de sortie est de taille 3 pour obtenir les valeurs de la position d'interaction associée. Une couche dense (**dense layer** sur Keras) est une couche où chaque neurone reçoit la sortie de tous les neurones de la couche précédente. Contrairement aux réseaux convolutifs on ne se base sur aucune hypothèse de structure ou de symétrie des données. Pour optimiser la taille des couches utilisées on peut effectuer une étude basée sur un fichier de simulation avec les positions d'interaction et l'intensité de la charge induite sur l'électrode et un voisinage de 8 voisins. On peut alors tracer un histogramme des positions pondéré par l'erreur générée par le réseau de neurones pour un même nombre d'itérations afin d'observer l'évolution des erreurs et de leur localisation en fonction de la complexité du réseau utilisé. On garde une taille de 18 en entrée et 3 en sortie pour coïncider avec la forme de nos données (Figure 3.8).

On peut remarquer qu'avec la complexité croissante l'intensité des erreurs diminue sur l'ensemble. Les erreurs plus importantes localisées en profondeur du détecteur et sur les bords du pixel s'estompent petit à petit et de manière moins flagrante les erreurs plus faibles diminuent encore plus avec la complexité. L'architecture 18 – 16 – 16 – 3 est retenue car elle permet d'optimiser les erreurs tout en gardant un temps de calcul très correct, de l'ordre de la microseconde pour une interaction, ce qui est dans les standards de l'algorithme du maximum de vraisemblance implémenté dans l'électronique du détecteur actuellement.

A chaque couche est associée une fonction d'activation. Lors du parcours du réseau pendant une itération, à l'entrée d'un neurone les poids puis la fonction d'activation sont appliqués sur la valeur ce qui donne une valeur de sortie. Dans notre cas on décide d'utiliser la fonction tangente hyperbolique, qui est la version translatée entre -1 et 1 de la sigmoïde utilisée très régulièrement. Ce choix vient du fait que l'on peut centrer nos valeurs en $(0, 0)$ ce qui correspond au centre du pixel (Figure 3.9). La symétrie de cette fonction par rapport à l'origine permet également d'obtenir une réponse symétrique entre un signal négatif et positif de même amplitude.

Compilation du modèle

Cette étape permet comme son nom l'indique de compiler le modèle de réseau de neurones développé précédemment. C'est à ce niveau que l'on décide de l'optimiseur, la fonction de perte et la métrique utilisés.

- L'optimiseur permet de déterminer le poids idéal d'un neurone en fonction de la comparaison entre les valeurs estimées par le réseau et les valeurs réelles. Il permet également d'optimiser la fonction de perte. De nombreuses variétés d'optimiseur existent mais celle que l'on utilise est ADAM (Adaptative Moment Situation). C'est

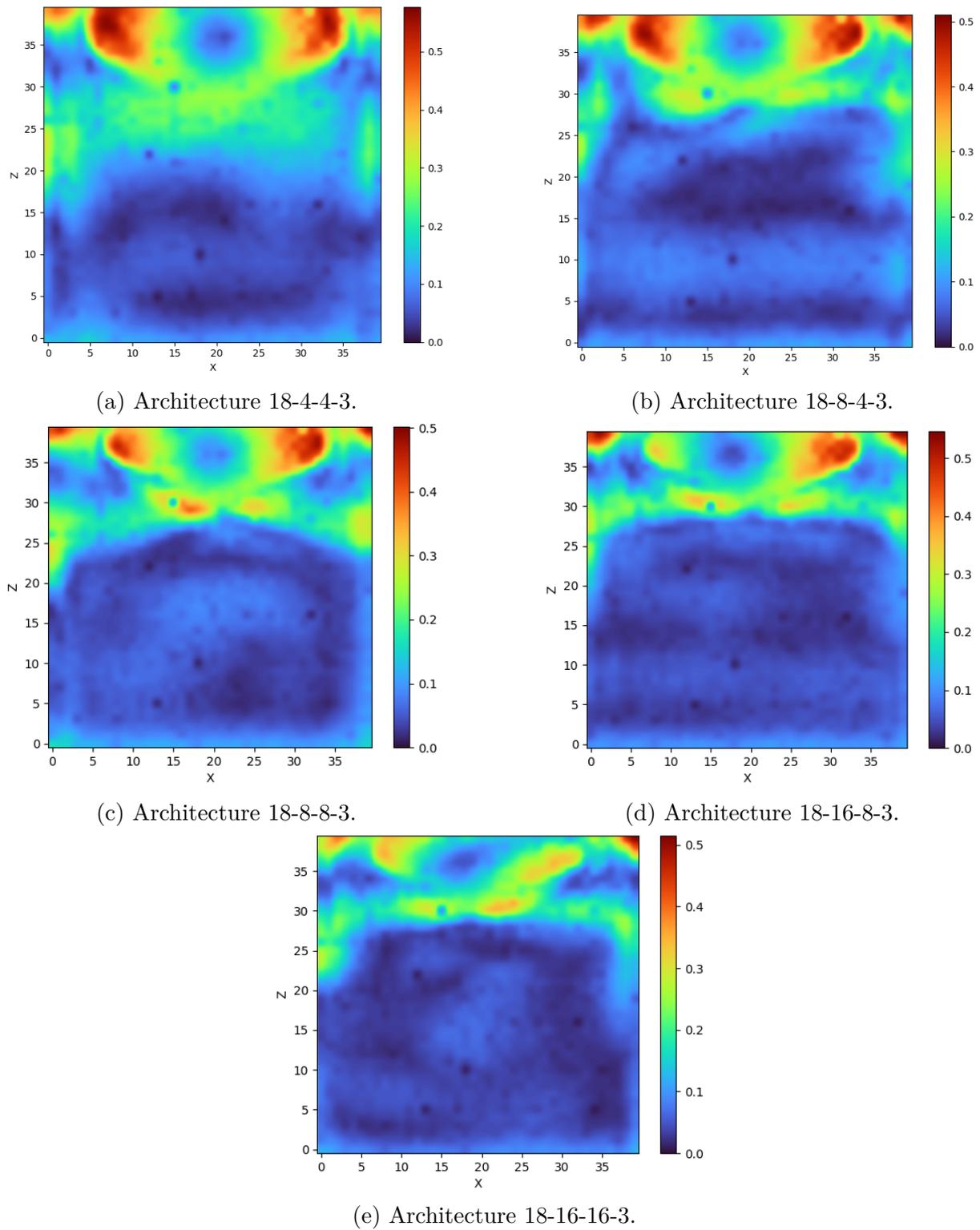


FIGURE 3.8 – Histogrammes 2D des positions pondéré par les erreurs associées en X et Z (coupe en Y au centre du pixel) pour différentes architectures de réseau de neurones.

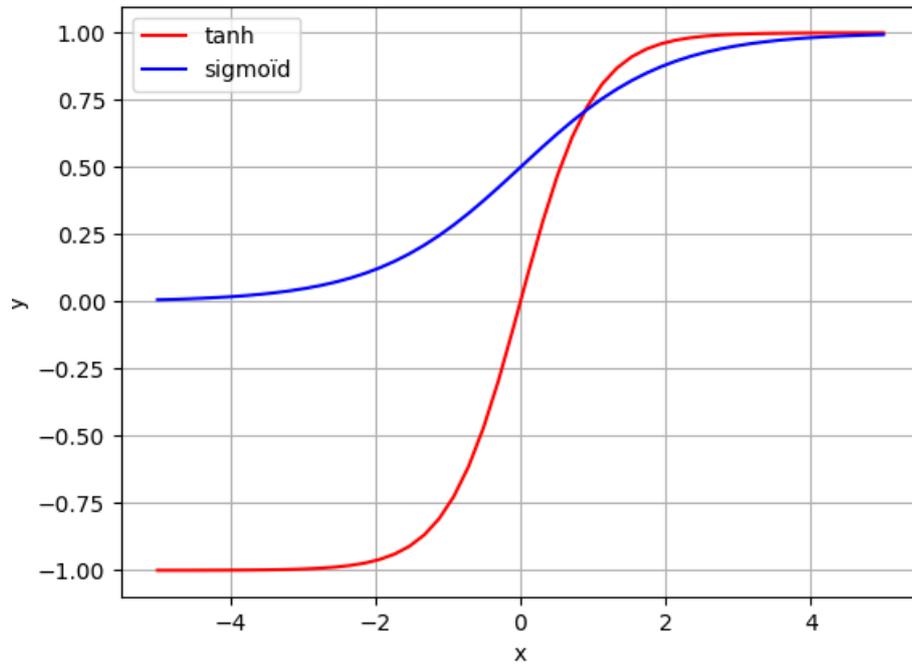
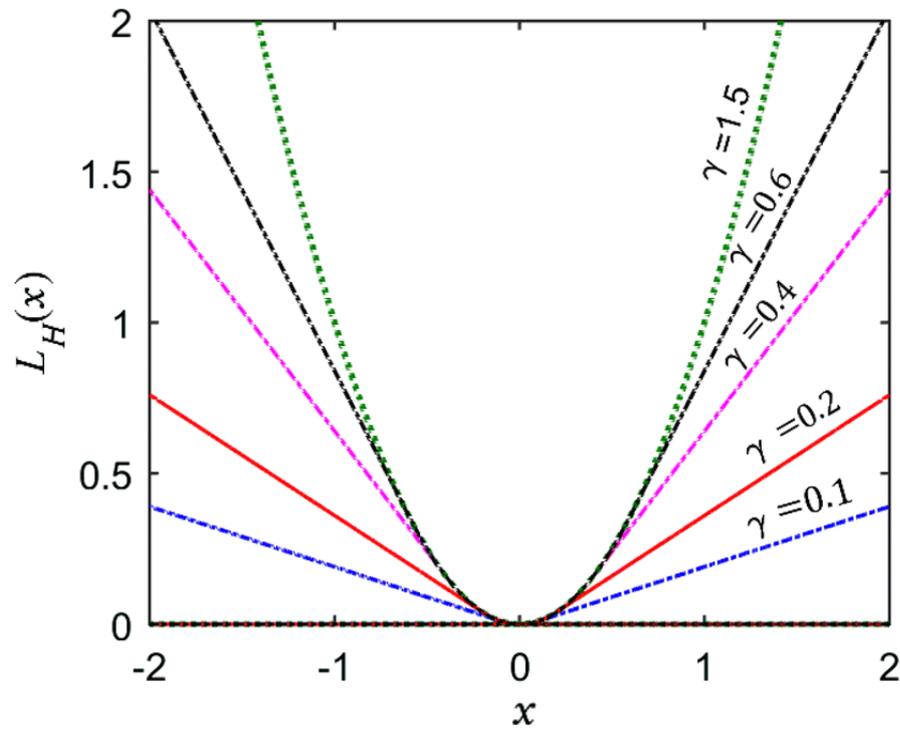


FIGURE 3.9 – Graphes des fonctions tangente hyperbolique et sigmoïde.

FIGURE 3.10 – Graphes de la fonction de Huber en fonction de la valeur de γ [92].

une méthode très populaire qui utilise la méthode de descente de gradient, nécessitant peu de mémoire tout en restant performant [93].

- La fonction de perte permet de déterminer numériquement la quantité que le modèle doit minimiser durant l'entraînement. De même que pour les optimiseurs il existe beaucoup de fonctions de perte disponibles sur Keras. Celle que l'on décide de choisir est la fonction Huber

$$L_{\delta}(y, \hat{y}) = \begin{cases} 0.5 * (y - \hat{y})^2 & \text{si } |y - \hat{y}| \leq \delta \\ \delta * |y - \hat{y}| - 0.5 * \delta^2 & \text{si } |y - \hat{y}| > \delta \end{cases} \quad (3.16)$$

avec y les valeurs réelles, \hat{y} les valeurs estimées et δ la valeur où la fonction de Huber passe de quadratique à linéaire (par défaut $\delta = 1$) (Figure 3.10). L'intérêt de la fonction de Huber est de minimiser l'impact des valeurs estimées aberrantes en les incluant avec une simple norme 1. A l'inverse les valeurs estimées se rapprochant de la réalité ont plus de poids grâce à la norme 2.

Entraînement et test du modèle

Cette partie permet de finaliser notre modèle et d'évaluer ses performances sur le jeu de données dont on dispose. On peut dans un premier temps dire à notre modèle de réaliser un arrêt prématuré (*callback*) s'il s'aperçoit que la valeur de la fonction de perte augmente. Ceci permet d'éviter le sur-apprentissage et de ne pas s'occuper du nombre d'itérations à réaliser (*epochs*). Vient ensuite la méthode d'ajustement, qui est le moteur d'entraînement du modèle pour un nombre d'itérations que l'on décide (qui sera éventuellement réduit par le *callback*). On rentre également les données d'entraînement, de validation et de test puis on compile afin d'obtenir la courbe de la fonction de perte ainsi que les résultats de l'apprentissage.

3.3 Test de l'apprentissage sur simulation

Pour calibrer notre réseau de neurones on s'intéresse tout d'abord aux données simulées, pour lesquelles nous avons une parfaite connaissance de la relation entre les entrées et les sorties. Une fois que les poids du réseau seront ajustés, il sera possible de tester le réseau sur les données expérimentales afin de quantifier le gain de cette méthode de positionnement des interactions.

3.3.1 Etude avec quatre voisins

On s'intéresse d'abord à l'étude avec quatre voisins directs du pixel touché. On ajoute aux données une simulation un bruit électronique afin de se rapprocher des données que l'on pourrait obtenir sur un vrai détecteur (Figure 3.13). Les tableaux de signaux induits et des positions associées sont calculés grâce à Tasmania, pour un détecteur CZT avec une épaisseur de 6mm avec un pas d'anode de 2.5mm, pour une source de Cobalt 57 dont le pic photoélectrique correspond à 122 keV.

Dans un premier temps on réalise un apprentissage de notre réseau de neurones sur un jeu de données simple simulés grâce à Tasmania. On tire aléatoirement un certain

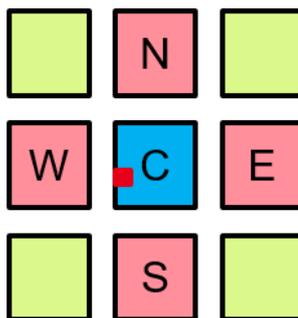


FIGURE 3.11 – Schéma du pixel touché et des quatre pixels voisins étudiés.

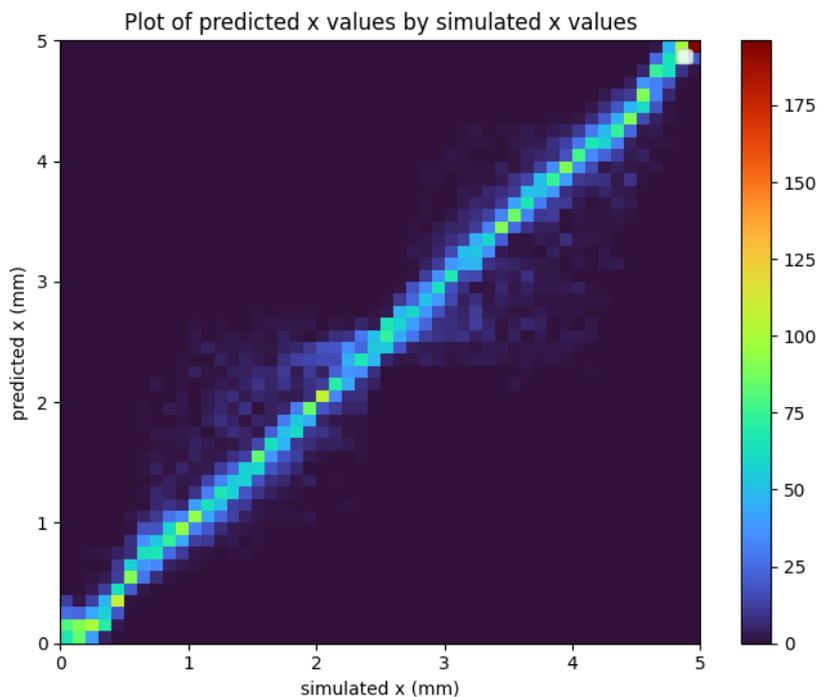
nombre d'événements (500 000 ici) sur un voisinage d'électrodes pour chacune desquelles on récupère une valeur de mesure de position. On s'intéresse à l'électrode collectante et aux quatre électrodes adjacentes telles que présentées sur la figure 3.11. On initialise le réseau en séparant nos données pour l'entraînement, la validation ainsi que le test pour un certain nombre d'itérations qui sera éventuellement réduit pour éviter le sur apprentissage (la valeur de base est 200). Les données sont également normalisées en divisant l'amplitude des voisins par celle du pixel central, et les coordonnées observées ont des valeurs entre -1 et 1. Notre réseau aura donc une couche d'entrée de cinq valeurs correspondant aux cinq amplitudes d'intérêt.

On se place dans le cas d'un détecteur idéal simulé par Tasmania. On simule des interactions de photons de manière aléatoire afin d'obtenir des valeurs de coordonnées qui sont associées à des amplitudes du pixel cible et de ses quatre voisins selon les quatre points cardinaux (Figure 3.11). On peut alors construire notre jeu de données en séparant les entrées (les amplitudes des pixels) et les sorties (les coordonnées du point d'interaction du photon). Comme décrit précédemment on sépare nos données pour les préparer pour le réseau de neurones. Puis, après entraînement, on teste le réseau sur une partie des données d'entrée mise de côté pour observer les performances du modèle. Ainsi on trace sur la figure 3.12 les histogrammes des valeurs des coordonnées prédites par le modèle en fonction des coordonnées simulées afin de voir si elles correspondent.

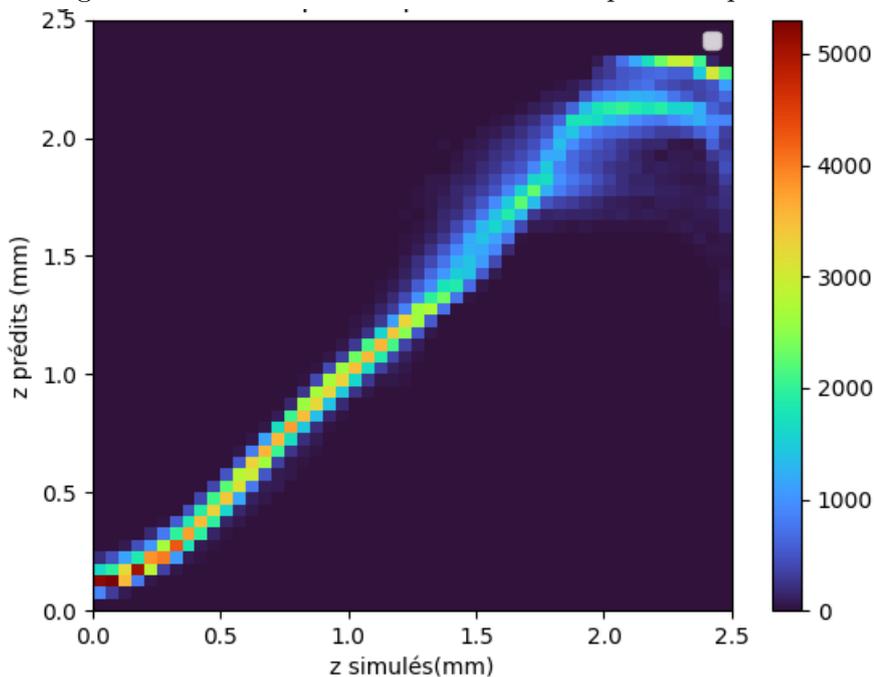
On peut observer que le réseau permet d'estimer de manière efficace les coordonnées d'interaction des photons simulés par Tasmania. On retrouve des légères perturbations au centre du pixel pour la coordonnée x , et une difficulté d'estimation des interactions en profondeur.

L'électronique simulée permet d'ajouter une mise en forme du signal grâce au shaper et un détecteur de crête sur la mesure de position de chaque pixel (Figure 3.13). On utilise également une correction d'uniformité pour obtenir des résultats similaires au calcul de la position avec ML. Pour observer l'effet du bruit électronique on réalise des simulations de positions d'interaction avec différents niveaux de bruit (Figure 3.14). Le bruit électronique est simulé en ajoutant un bruit gaussien aux mesures sur chaque voie d'électronique. Il est exprimé en électrons RMS. Cette valeur est appelée l'ENC (*equivalent noise charge*), correspondant à l'intégration du bruit électronique, dont la densité spectrale de puissance est connue en sortie du préamplificateur dans le filtre de mise en forme (ou *shaper*) utilisé.

On remarque que l'intensité du bruit appliqué sur la simulation dégrade l'estimation des coordonnées des positions d'interaction, ce qui est observable avec l'augmentation de l'épaisseur de la droite de corrélation. On retrouve également les effets physiques dus au



(a) Histogramme d'estimation de la coordonnée x pour chaque interaction.



(b) Histogramme d'estimation de la coordonnée z pour chaque interaction.

FIGURE 3.12 – Histogrammes d'estimation des coordonnées x et z par le réseau de neurones par apprentissage supervisé sur des données de simulation.

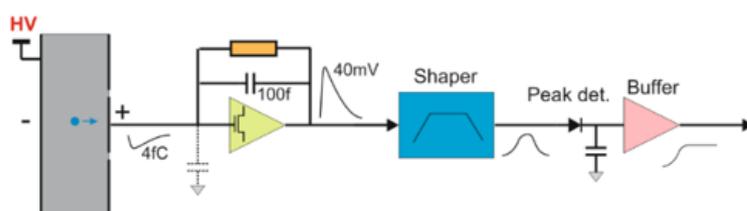


FIGURE 3.13 – Schéma de l'électronique associée au détecteur. Sous l'effet de la haute tension, les électrons (en bleu) migrent dans le détecteur (en gris), induisant un signal de quelque femtocoulombs sur l'électrode. Ce dernier est intégré et converti en tension par le préamplificateur (en vert), mis en forme par un filtre (en bleu). On utilise un détecteur de crête pour déterminer la valeur maximale temporelle au cours du temps.

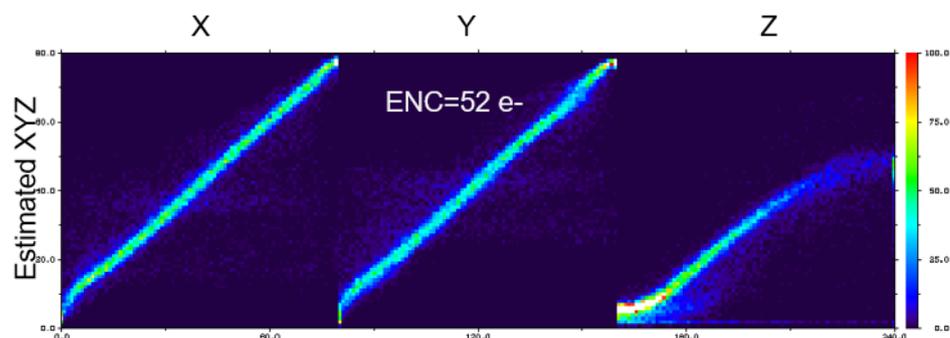
partage de charge et aux interactions en profondeur mais avec un impact beaucoup plus faible qu'avec la localisation par maximum de vraisemblance. Ceci provient notamment de la parfaite connaissance des liens entre position et signaux induits pour données de simulation ainsi que de l'absence de défauts dans le détecteur dans ce premier cas. L'estimation des coordonnées X et Y reste très correcte, l'estimation de la profondeur Z est moins précise et nécessite une correction.

3.3.2 Étude avec huit voisins

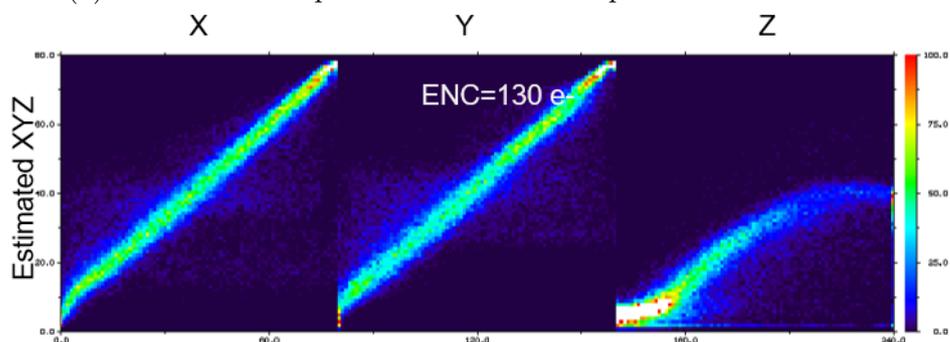
Pour tester notre réseau de neurones en exploitant au mieux les informations disponibles, on utilise les huit voisins du pixel touché et on implémente un second filtre électronique pour obtenir un second signal par pixel (figure 3.15). Dans ce cas on obtient dix-huit valeurs par pixel, ce qui a priori devrait améliorer les performances de localisation du réseau. On doit alors modifier l'électronique de simulation afin d'implémenter ces différences et se rapprocher du type de données que l'on pourrait utiliser pour l'estimation de position plus précise.

Dans le cas où l'on utilise les données sur les huit voisins du pixel collectant on parvient à fortement limiter l'impact des interactions centrées sur le pixel et celles se produisant en profondeur. On peut ainsi conclure qu'il est plus efficace de considérer les valeurs du voisinage des huit voisins plutôt que celui des quatre, la difficulté sera principalement de récupérer assez de données expérimentales afin de pouvoir optimiser la localisation. De plus même si le réseau est assez simple le fait de devoir l'implémenter de manière électronique pourrait impacter la vitesse de localisation. Il faut alors trouver un compromis entre la précision du réseau de neurones et sa rapidité afin de rester dans les références actuelles de puissance accessible et de temps de localisation.

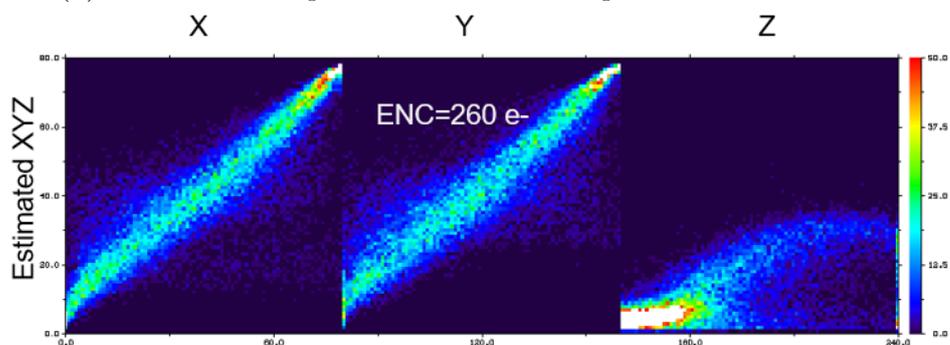
On observe sur la figure 3.17 la comparaison, pour une même valeur de bruit électronique, de l'estimation des coordonnées d'interaction de chaque photon simulé dans le cas où on considère une valeur par pixel et dans le cas où on en utilise deux. On peut déjà remarquer que l'estimation de la profondeur est plus précise qu'en utilisant quatre voisins, et l'estimation des coordonnées X et Y sont très bonnes. L'utilisation de deux valeurs par pixel permet d'améliorer la localisation surtout pour les interactions ayant lieu vers le centre du pixel, en effet on ajoute de l'information de localisation grâce à ce canal supplémentaire ainsi la sous-pixelisation en XY est plus efficace (Figure 3.17).



(a) Estimation des positions d'interaction pour un bruit de 52e-.



(b) Estimation des positions d'interaction pour un bruit de 130e-.



(c) Estimation des positions d'interaction pour un bruit de 260e-.

FIGURE 3.14 – Comparaison des positions estimées par le réseau de neurones avec les positions réelles pour différents niveaux de bruits (utilisation de quatre voisins et une valeur par voisin).

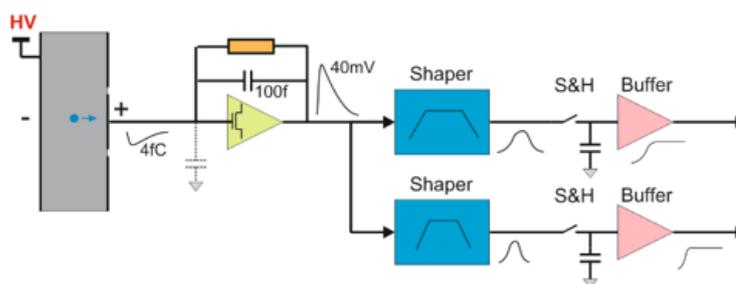


FIGURE 3.15 – Schéma de l'électronique pour deux signaux par canal de sortie : on utilise deux mises en forme parallèles et on mesure leur valeur de manière synchrone grâce à des échantillonneurs-bloqueurs.

Finalement l'ajout d'un canal représente des données en plus à traiter électroniquement et numériquement (multipliées par deux), mais il est très utile afin d'améliorer la localisation des interactions de manière significative, surtout dans le cas de données bruitées.

3.3.3 Comparaison pour d'autres images simulées

On peut élargir l'évaluation du réseau de neurones en recréant des images simulées et comparer les résultats obtenus par une localisation par maximum de vraisemblance.

Irradiation collimatée

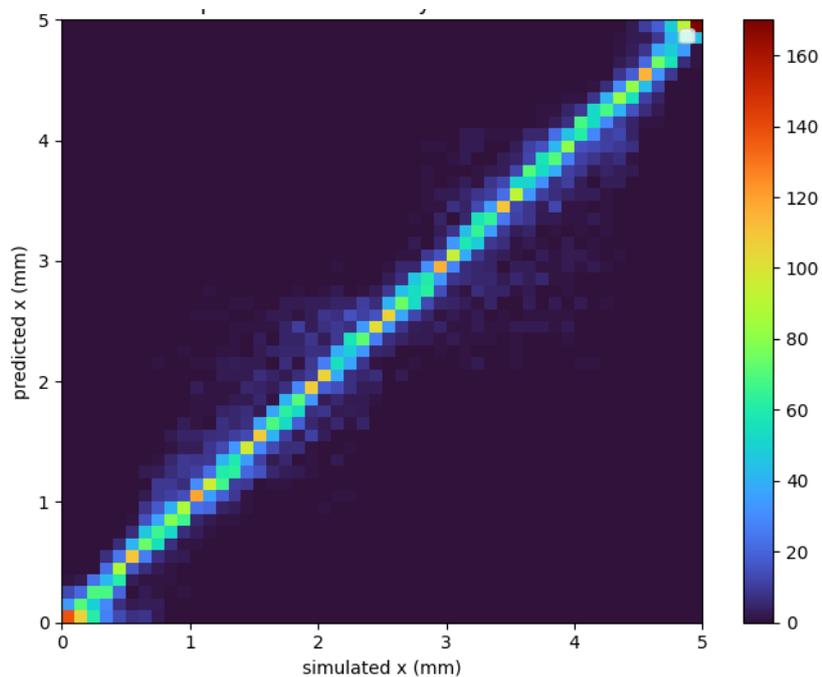
On simule un collimateur circulaire de 1.25 mm de diamètre (Figure 3.18). On compare les images obtenues pour une telle configuration avec les deux types de localisation : par maximum de vraisemblance ou par réseau de neurones.

On remarque que le maximum de vraisemblance donne un résultat moins résolu et plus déformé, en particulier dans le cas où l'on utilise qu'une valeur par anode. Cette déformation peut s'expliquer par le fait que l'algorithme traite chaque variable x et y de manière séparée. En revanche le réseau de neurones, en plus de ne pas faire apparaître cet artefact, possède une localisation plus précise dans les deux cas montrant l'utilité du réseau de neurones dans cette situation (Figure 3.19).

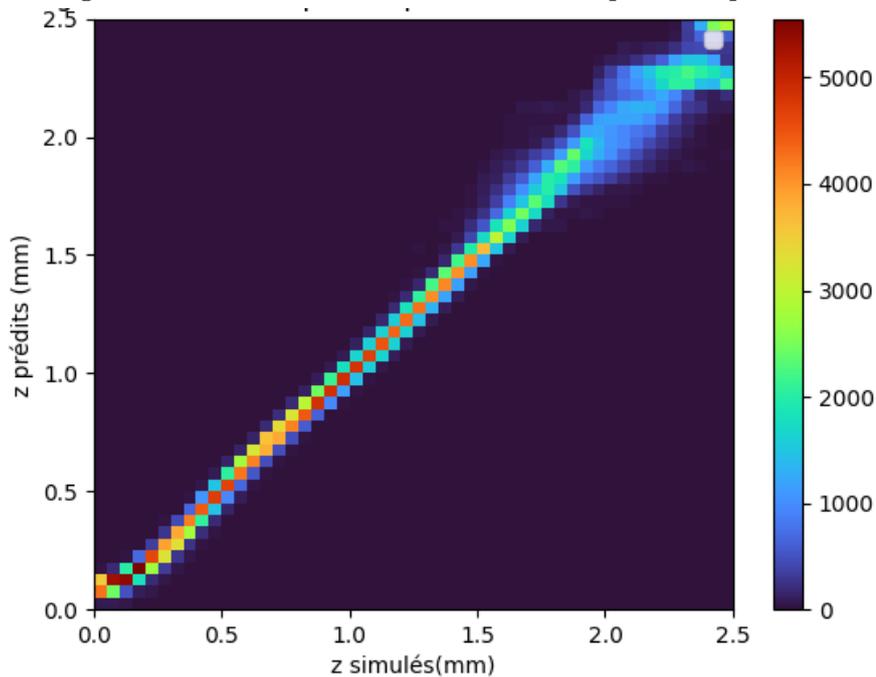
Lame absorbante

De la même manière, on génère une image résultant du masquage d'une partie du pixel par une lame absorbante, typiquement réalisée en tungstène, métal de densité élevée, typiquement $19 \text{ g} \cdot \text{cm}^{-3}$ (Figure 3.20). Comme pour le cas précédent on compare les performances des deux types d'algorithme de localisation (Figure 3.21).

Dans le cas à une valeur par canal, chacune des méthodes a des lacunes pour la localisation. Le maximum de vraisemblance fait apparaître un bruit assez important au bord de l'image rendant difficile la parfaite localisation de la limite. Pour le réseau de neurones certaines valeurs sur le tour du pixel ne sont pas apprises correctement en revanche la limite est plus nette. Par contre quand on considère deux valeurs par canaux le réseau de neurones prend tout son intérêt. Pareillement au cas du faisceau collimaté on retrouve une apparition de pixel avec des effets de bord pour le maximum de vraisemblance, tandis

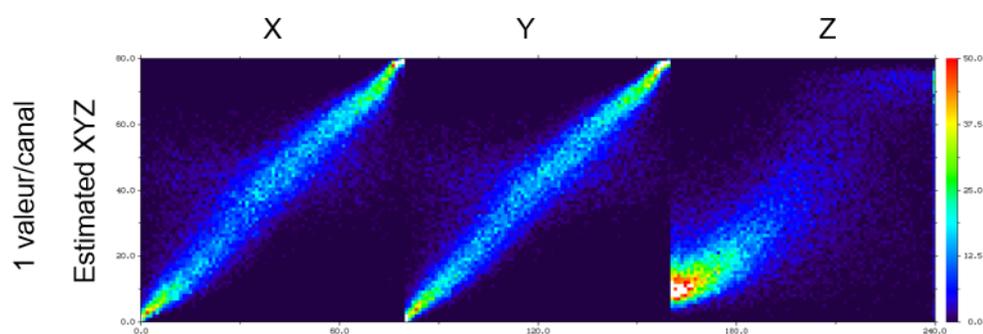


(a) Histogramme d'estimation de la coordonnée x pour chaque interaction.

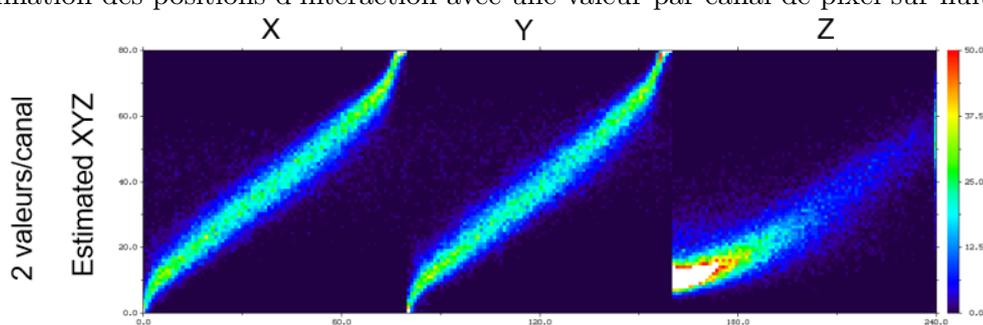


(b) Histogramme d'estimation de la coordonnée z pour chaque interaction.

FIGURE 3.16 – Histogrammes d'estimation des coordonnées x et z par le réseau de neurones par apprentissage supervisé sur des données de simulation.



(a) Estimation des positions d'interaction avec une valeur par canal de pixel sur huit voisins.



(b) Estimation des positions d'interaction avec deux valeurs par canal de pixel sur huit voisins.

FIGURE 3.17 – Comparaison des positions estimées par le réseau de neurones avec les positions simulées pour une ou deux valeurs par canal de pixel.

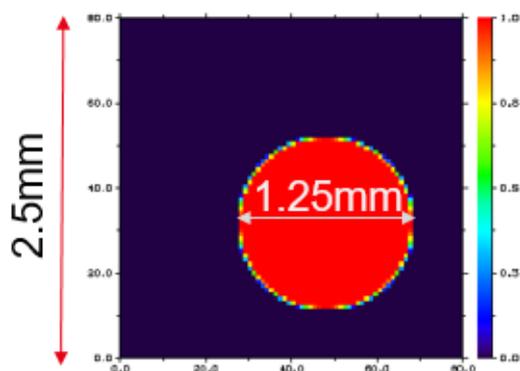


FIGURE 3.18 – Image de la simulation du faisceau collimaté.

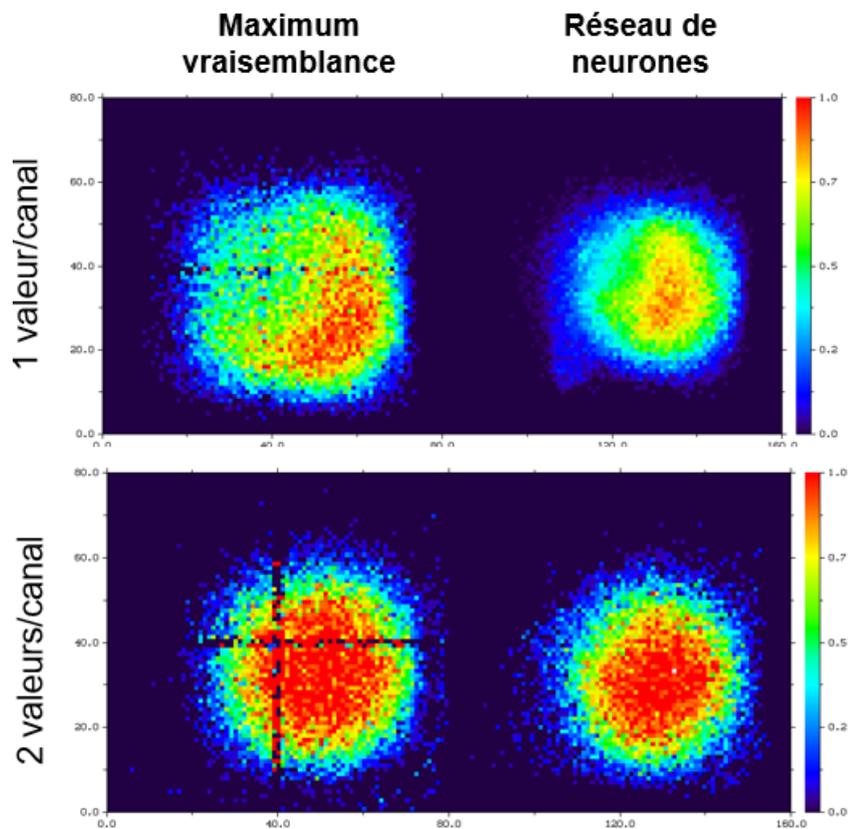


FIGURE 3.19 – Comparaison de la localisation d’un faisceau collimaté simulé de 1.25mm de diamètre par la méthode maximum de vraisemblance (gauche) et par réseau de neurones (droite), en utilisant une (haut) et deux (bas) valeurs par canal de pixel.

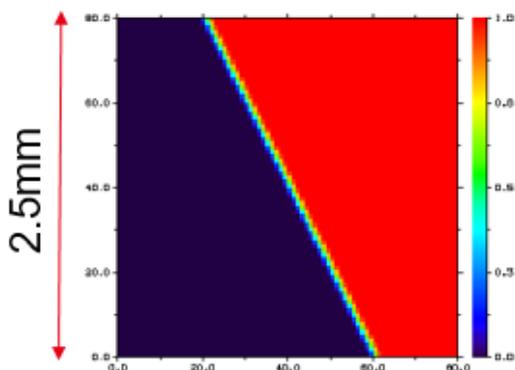


FIGURE 3.20 – Image de la simulation du bord d’un détecteur/pixel.

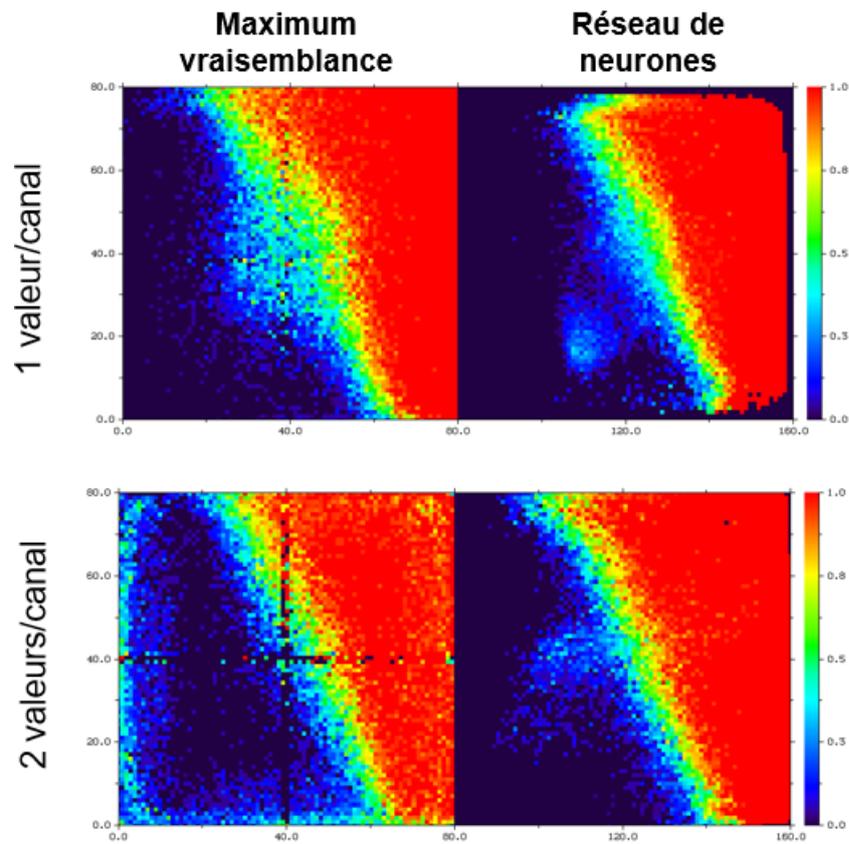


FIGURE 3.21 – Comparaison de la localisation d'un bord de détecteur/pixel simulé par la méthode maximum de vraisemblance (gauche) et par réseau de neurones (droite), en utilisant une (haut) et deux (bas) valeurs par canal de pixel.

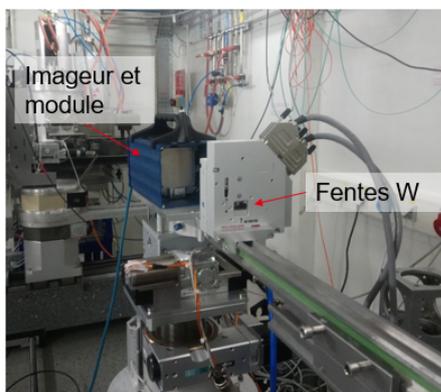
que le bord est plus net avec un léger bruit vers le centre pour la localisation par réseau de neurones. En plus de confirmer les meilleurs résultats du réseau de neurones, on remarque que plus la quantité de données en entrée est importante et plus les écarts de performance se creusent. Ceci est logiquement explicable par le fait que d'une part les variables étant traitées séparément pour le maximum de vraisemblance on ajoute des biais d'estimation, et d'autre part pour le réseau de neurones on nourrit ce dernier avec plus de statistiques ce qui permet d'affiner les poids du réseau.

3.4 Application aux données expérimentales

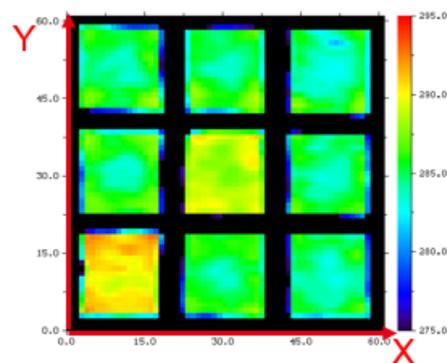
Dans cette partie nous allons nous intéresser aux résultats obtenus pour l'apprentissage de positions sur des résultats expérimentaux. Pour l'instant, le réseau n'est pas calibré pour utiliser des données dont le lien entre entrée et sortie n'est pas parfaitement connu, mais on peut tout de même explorer ses performances sur ces jeux de données imparfaits.

3.4.1 Utilisation des données de l'ESRF

Pour faire les premiers tests de l'algorithme d'apprentissage supervisé par réseau de neurones, on utilise un jeu de données provenant de l'ESRF (European Synchrotron Radiation Facility) à Grenoble. Ces expériences ont été menées quelques années plus tôt dans le cadre d'un autre projet afin de scanner en X et Y un détecteur pour évaluer la résolution spatiale et en énergie de celui-ci (Figure 3.22). L'avantage de ce jeu de données est que le scan effectué est fait avec un pas très faible ($125 \mu\text{m}$), ce qui fait que le lien entre entrées (position d'interaction et énergie déposée) et sorties (signaux induits sur le pixel collectant et ses voisins) est presque connu.



(a) Photo du banc expérimental de mesures à l'ESRF.



(b) Carte CIE 2D du scan réalisé sur un voisinage de 3x3 électrodes pour un détecteur CZT de 6mm d'épaisseur avec un pas d'anode de 2.5mm.

FIGURE 3.22 – Banc expérimental et résultat du scan réalisé à l'ESRF.

L'utilisation de ce jeu de données présente néanmoins certains inconvénients : le flux utilisé est de basse énergie ($\approx 60 \text{ keV}$), le ratio signal sur bruit est plutôt mauvais, on utilise les données du pixel collectant et des quatre voisins adjacents uniquement, on ne

recupère qu'une seule valeur par canal de pixel et un la présence d'un *offset* limite la mesure des signaux de faible amplitude.

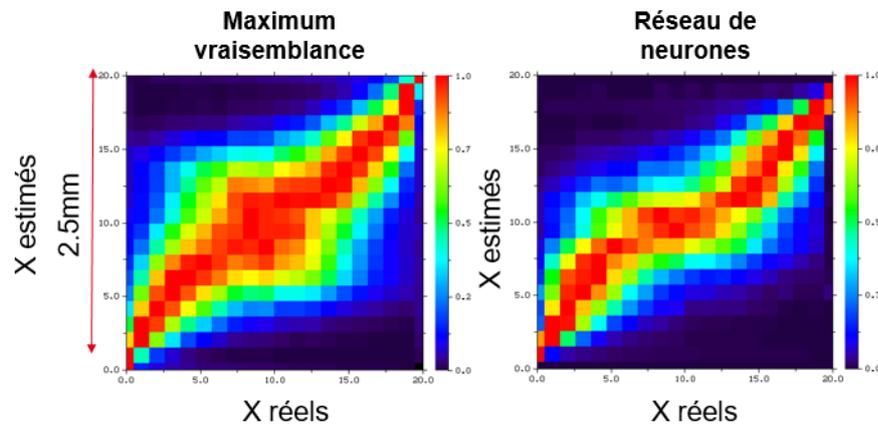


FIGURE 3.23 – Histogramme des coordonnées X d'interaction estimées par la localisation par maximum de vraisemblance (à gauche) et par le réseau de neurones (à droite) avec les X réels (mesurés expérimentalement).

On peut tout de même tester ces données pour réaliser un apprentissage supervisé en les utilisant comme données de référence pour notre réseau de neurones. Même si le pas fin du faisceau de scan aide à l'estimation plus précise de la position, celle-ci n'est pas parfaite et la relation entre données d'entrée et de sortie n'est pas totalement connue. En comparant les résultats de l'estimation de la localisation (par exemple de la coordonnée X) pour nos deux méthodes de calcul on obtient les histogramme de la figure 3.23. On voit clairement la meilleure estimation réalisée par le réseau de neurones, avec un histogramme qui suit plus fidèlement la diagonale $y = x$ qui est l'objectif visé. On notera l'incertitude liée aux interactions au milieu du pixel dans les deux méthodes, même si elle est plus appuyé pour le maximum de vraisemblance.

Grâce à ces données, on a également simulé l'image sur un pixel d'un écran de tungstène (figure 3.24) et d'une irradiation collimatée (figure 3.25). La principale force du réseau de neurone est de mieux se comporter au centre du pixel et de mieux combiner les variations en x et y. En effet, la séparabilité en x et y est une des hypothèses forte de l'implémentation simplifiée du maximum de vraisemblance.

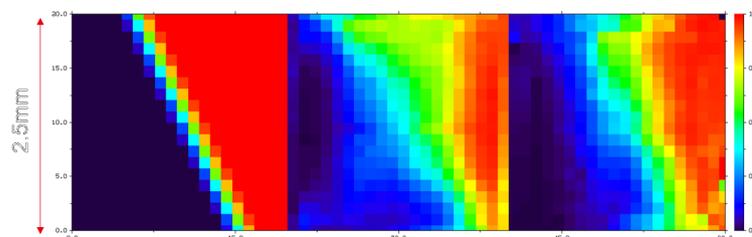


FIGURE 3.24 – Comparaison d'une image synthétique d'un écran de tungstène placé en travers du pixel. Référence (à gauche), maximum de vraisemblance (au centre) et réseau de neurones (à droite).

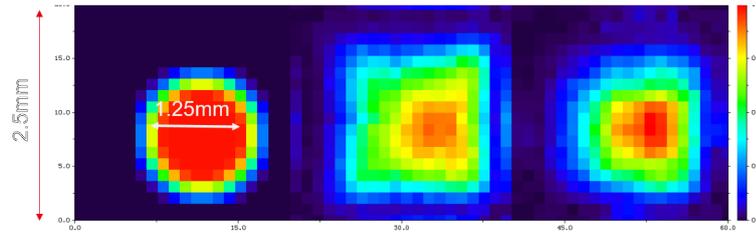


FIGURE 3.25 – Comparaison d’une image synthétique d’un trou de 1.25 mm de diamètre placé devant le pixel. Référence (à gauche), maximum de vraisemblance (au centre) et réseau de neurones (à droite).

3.4.2 Utilisation des données du banc expérimental

Afin de se placer dans des conditions plus réalistes de calibration de détecteur on utilise un banc expérimental afin de réaliser des scans de différents détecteur et obtenir des données de réponse typiques de détecteurs utilisés dans des dispositifs industriels (Figure 3.26). On utilise un banc de translation selon les axes X et Y avec une amplitude de 15 cm, largement suffisant pour notre application. Le détecteur CZT étudié est placé dans un banc de numérisation développé au laboratoire. On place le détecteur dans la partie supérieure et la partie inférieure comporte toute l’électronique de traitement. Les signaux de chacune des 16 voies sont numérisés simultanément, ce qui permet une très grande liberté sur le traitement de signal effectué. On peut placer une source collimatée sur le plateau situé au dessus du système. Le mouvement du banc peut être piloté soit de manière manuelle soit automatisée.

Ce banc nous permet de faire l’acquisition de données correspondant aux signaux de sortie mesurés sur chaque pixel pour chaque interaction. On obtient alors une liste d’événements que l’on peut traiter afin de remonter à la localisation de l’interaction du photon γ dans le détecteur. On applique d’un côté la méthode classique par maximum de vraisemblance utilisée actuellement, et de l’autre l’apprentissage supervisé par réseau de neurones. On obtient alors les histogrammes que l’on peut observer sur les figures 3.27 pour un cristal de 4×4 électrodes.

Les pixels de coins et de bord sont dans une situation défavorable car il leur manque des voisins pour correctement localiser par sous-pixellisation. On observe sur les images un défaut de l’électronique : le pixel supérieur droit qui est déficient (non connecté). On peut remarquer que le réseau de neurones a quelques lacunes d’apprentissage sur les bords des pixels, surtout dans les coins supérieur droit et inférieur gauche. Ceci est peut être dû au manque de statistiques, qui fait que l’on manque d’informations dans ces zones là. En revanche la localisation des interactions sur les bords des pixels est plus performante (visible également sur la figure 3.28). En effet la méthode du maximum de vraisemblance va avoir tendance à manquer de précision au niveau de ces interfaces. Ensuite la méthode par réseau de neurones permet de voir apparaître les pixels voisins au voisinage que l’on étudie, tandis que dans le cas du maximum de vraisemblance on obtient uniquement un contour où toutes les interactions mesurées en dehors du voisinage sont regroupées.

On peut également visualiser la réponse d’une irradiation collimatée à différents endroits du détecteur et de comparer les performances de chacune des méthodes de localisation. On regroupe les images obtenues sur la figure 3.28. On observe sur la première

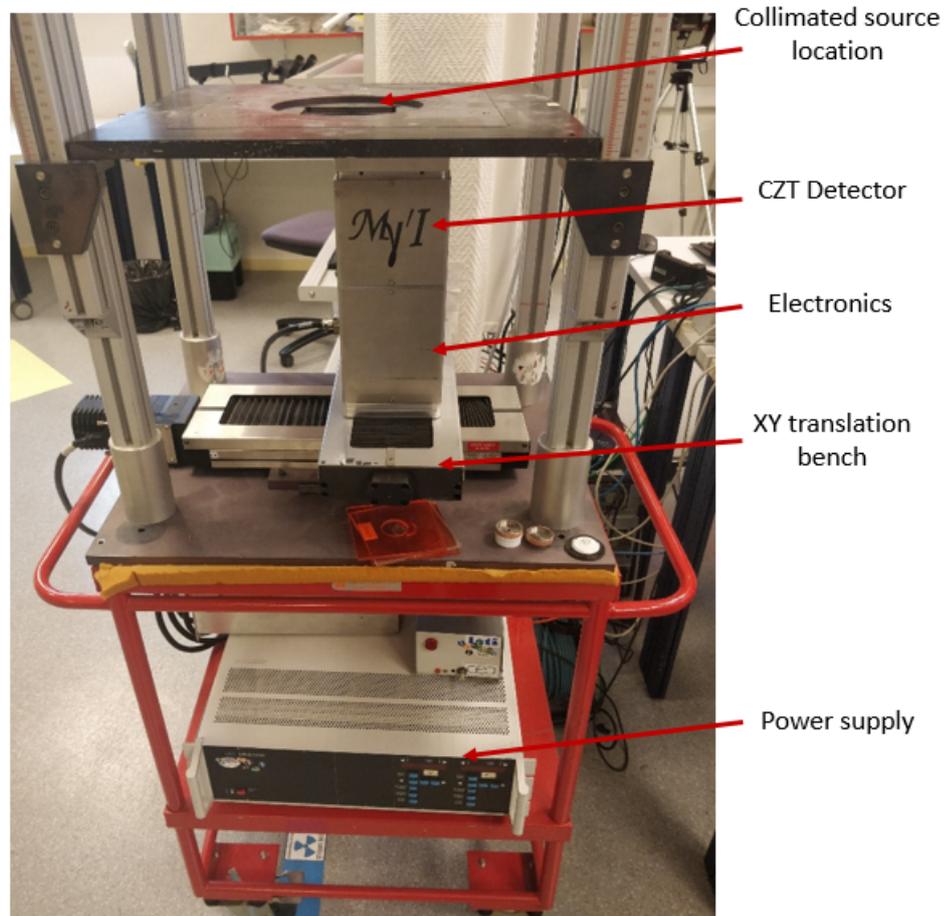


FIGURE 3.26 – Photo du banc expérimental utilisé pour les scans XY de détecteur CZT.

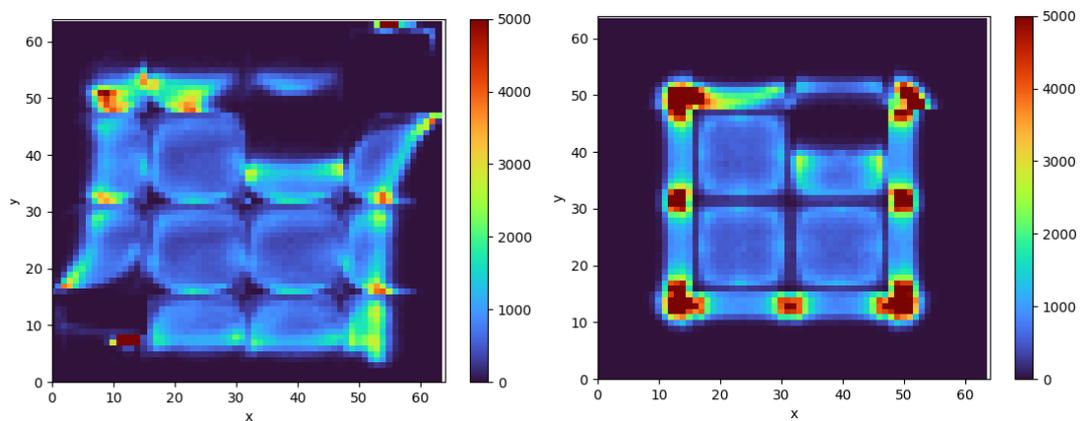


FIGURE 3.27 – Histogramme des localisations des interactions par maximum de vraisemblance (gauche) et réseau de neurones (droite). L'image fait 10×10 mm de dimension et le sous-pixel $156 \mu m$.

image que les performances des deux algorithmes sont similaires avec une reconstitution presque parfaite de la zone d'irradiation. En revanche en déplaçant la source collimatée vers un côté du détecteur on remarque que la technique par maximum de vraisemblance va avoir tendance à tronquer le faisceau sur le bord du pixel et à perdre de l'information, tandis que le réseau de neurones va conserver l'information et la forme du faisceau grâce à sa modélisation du flexible des disparités de réponse entre pixels. On voit sur la dernière image qu'il ne reste presque pas de signal pour le maximum de vraisemblance, tandis que le réseau de neurones arrive à reconstituer la forme de l'irradiation malgré la sortie du pixel. Ceci prouve la meilleure adaptabilité du réseau de neurones, notamment ses performances en bord de détecteur.

3.5 Conclusion

Dans ce chapitre nous avons pu aborder la technique de référence par maximum de vraisemblance, qui est la méthode de localisation utilisée aujourd'hui au laboratoire et intégrée au FPGA afin de déterminer les positions initiales des interactions ayant lieu dans le détecteur. Cette technique présente néanmoins des limites, telles que le besoin d'une correction d'uniformité à cause du partage de charge notamment induisant un biais, ou la non possibilité d'estimer le champ électrique ou les propriétés de transport des porteurs de charge.

Une des solutions envisagées dans cette étude est l'utilisation d'un réseau de neurones dense simple, implémentable dans l'électronique afin de compenser les lacunes de la méthode par maximum de vraisemblance et augmenter la précision de localisation. Ainsi on construit un réseau de neurones à trois couches cachées utilisant en données d'entrée le vecteur des signaux induits du pixel collectant et de ses voisins, et proposant en sortie une estimation des coordonnées de l'interaction ainsi que l'énergie déposée. Les essais de ce réseau d'abord sur des données simulées puis sur des données expérimentales permettent de démontrer sa meilleure efficacité.

Il y a néanmoins quelques limites à utiliser ce type d'apprentissage comme tel : le besoin d'un jeu de données d'apprentissage, difficile à obtenir de manière expérimental dû au fait que le faisceau de scan doit être collimaté finement rendant long le scan XY du détecteur. C'est pour cela que, dans le prochain chapitre, nous allons discuter de l'apprentissage non supervisé des paramètres détecteurs, qui va permettre d'utiliser une simulation optimisée à partir de données expérimentales afin de construire le jeu de données nécessaire à l'utilisation de cette nouvelle méthode de localisation. Le couplage de ces deux outils va nous permettre d'aboutir à une méthode de localisation et de caractérisation du détecteur plus adaptable et plus rapide.

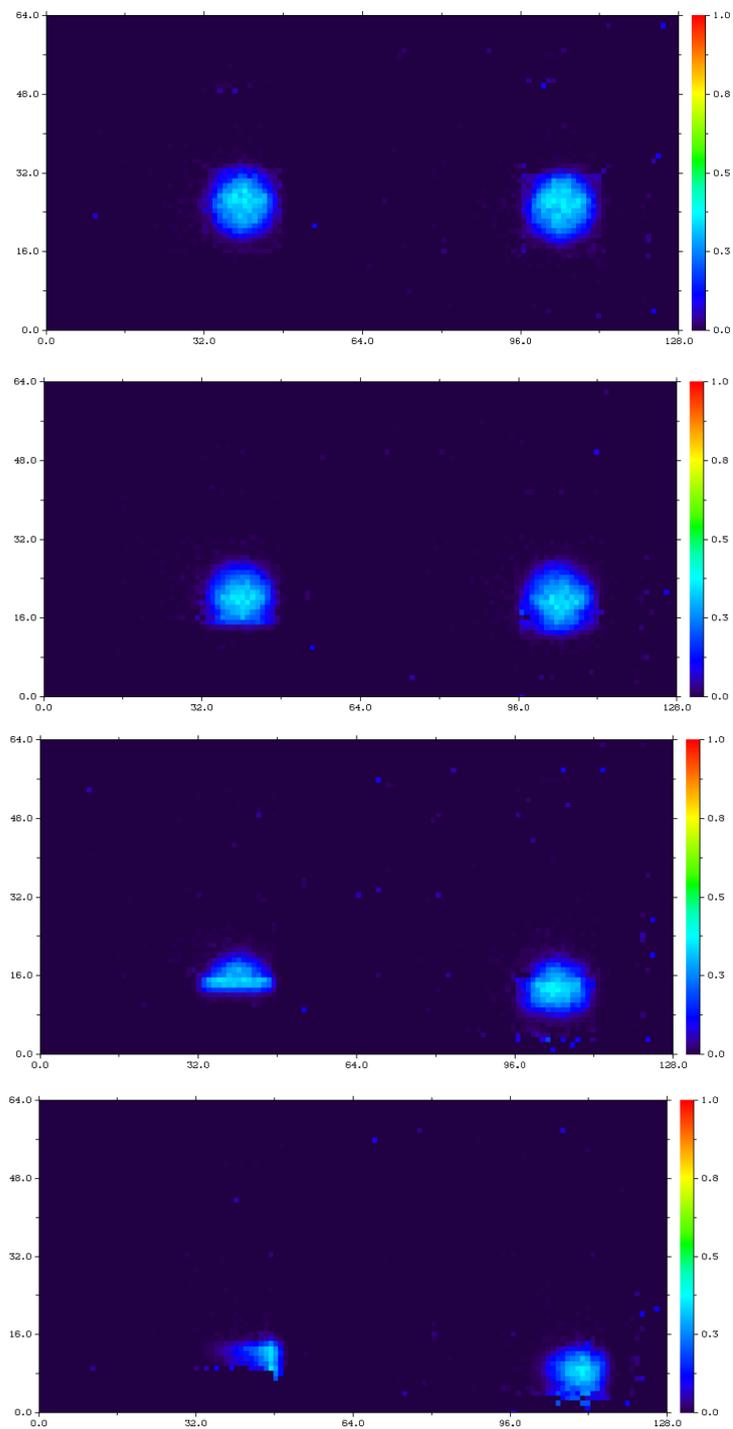


FIGURE 3.28 – Comparaison d’une image d’irradiation collimatée par un trou de 1 mm à différentes positions : maximum de vraisemblance (gauche) et réseau de neurones (droite). L’image fait 10×10 mm de dimension et le sous-pixel $156 \mu m$.

Apprentissage non supervisé des paramètres physiques du détecteur par équation différentielle adjointe

Le réseau de neurones dense développé dans le chapitre précédent présente un certain nombre d'avantages par rapport aux méthodes classiques d'estimation de la position comme la flexibilité et sa bonne performance tout en garantissant une bonne rapidité d'inférence. En revanche on se retrouve devant deux difficultés principales : le besoin d'un jeu de données conséquent afin de pouvoir réaliser l'entraînement de ce réseau de neurones, ainsi que la connaissance parfaite des relations entre les données d'entrée et de sortie afin que le calcul des poids de chaque neurone soit le plus précis possible. Pour résoudre cette situation on décide de développer un second algorithme d'apprentissage plus complet et plus complexe. L'objectif est de simuler la réponse du détecteur CZT étudié en apprenant ses paramètres internes, comme le champ électrique, qui est la principale inconnue¹. De cette manière on pourra utiliser la simulation, beaucoup plus rapide, pour entraîner le réseau de neurones pour la localisation présenté dans la partie précédente avec un détecteur possédant une réponse similaire aux données expérimentales.

4.1 Modèle Neural ODE

4.1.1 Théorie mathématique

En mathématiques, une équation différentielle ordinaire (*ordinary differential equation* ou ODE en anglais) est une équation différentielle dont la ou les fonctions inconnues ne sont dépendantes que d'une seule variable. Elle est opposée aux équations aux dérivées partielles pour lesquelles les fonctions inconnues dépendent de plusieurs variables. On peut décrire ce type d'équation selon sa forme explicite :

$$\frac{dz(t)}{dt} = f(z(t), t) \quad (4.1)$$

1. Un piégeage élevé ou un champ faible se traduisent tous les deux par une perte de charge localisée.

où z est la fonction que l'on cherche à intégrer en fonction de la variable t (ici le temps). La fonction f représente une fonction quelconque dont les variables sont uniquement z et t . Dans le cas d'un problème avec une condition initiale, on connaît l'évaluation de la fonction z pour la valeur initiale de t que l'on note t_0 . On peut également introduire une autre variable θ qui fait office de paramètre dans la fonction f . Pour donner un exemple concret dans un cas simple, dans l'équation $\dot{z} = \sin^2 \theta z$, la fonction f correspond à \sin^2 . Dans un cas multidimensionnel avec des variables vectorielles \mathbf{z} et Θ la structure de nos éléments sera évidemment plus complexe.

On peut alors intégrer notre équation initiale selon une nouvelle forme :

$$z(t) = z(t_0) + \int_{t_0}^t f(z(t), t, \theta) dt. \quad (4.2)$$

Ce type d'intégration peut se réaliser analytiquement dans le cas de fonctions simples, mais généralement nous avons besoin d'intégrer numériquement afin d'obtenir une approximation de la fonction f sur la plage temporelle (t_0, t) avec un certain pas définissant la précision de cette approximation. De nombreuses méthodes ont été développées au fil du temps, comme la méthode d'Euler qui permet d'approcher la fonction le long de sa trajectoire en effectuant des pas faibles et en utilisant la valeur précédente. Pour un pas h on réalise l'approximation

$$z(t+h) = z(t) + hf(z(t), t) \quad (4.3)$$

et en connaissant la valeur initiale on peut tracer une approximation de z .

La technique d'Euler est instable numériquement et peu praticable [94]. Des méthodes plus complexes et performantes comme celle de Runge-Kutta et toutes ses variantes sont apparues au fil du temps [95].

Afin de développer cet algorithme d'apprentissage non supervisé, on s'appuie sur une méthode développée à l'Université de Toronto par l'équipe du professeur R.T. Chen [96]. Il s'agit d'un nouveau type de modèle de réseaux de neurones utilisant les équations différentielles afin de créer des couches continues à la place des couches discrètes des réseaux classiques (Figure 4.1).

Pour résumer la méthode, elle consiste à utiliser une équation différentielle régissant l'évolution de l'état d'un système dépendant de paramètres équivalents aux poids d'un réseau traditionnel. L'intégration de cette équation permet de calculer l'état du système pour tout instant t . Ce temps sera l'équivalent discret d'un numéro de couche. L'inférence consiste alors à intégrer l'équation différentielle pour évoluer d'un état initial en entrée z_0 à un état final z_1 , l'évolution étant régie par les paramètres θ . La difficulté est alors essentiellement d'effectuer la rétropropagation afin d'apprendre θ à partir d'une erreur sur l'observation z_1 . Pour réaliser la mise à jour des « poids » de notre réseau grâce à la fonction d'erreur, on va calculer les gradients de la fonction d'erreur par rapport à chaque composante θ de l'état du système. Pour calculer ces gradients on utilise la méthode des équations adjointes, consistant à effectuer une intégration en inverse de l'équation différentielle. On calcule le gradient d'erreur événement par événement, et on met à jour les paramètres du système de manière globale. L'évolution temporelle du système est liée à une fonction de dynamique f et un vecteur de paramètres θ , tel que dans l'équation (4.2).

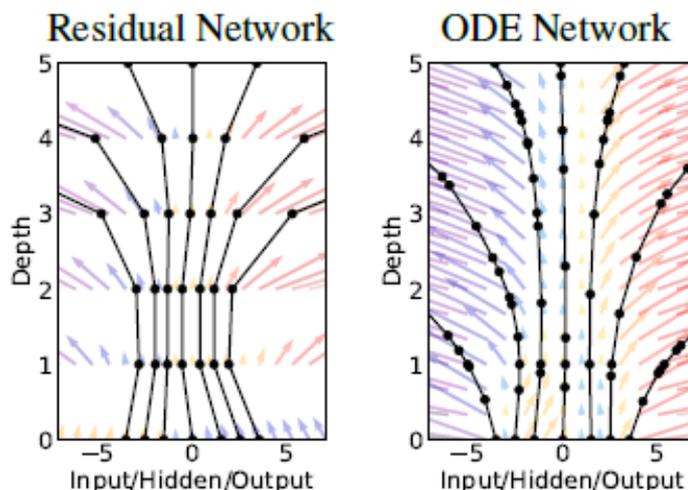


FIGURE 4.1 – Comparaison de la structure d’un réseau de neurones classique (gauche) avec celle de la méthode Neural ODE (droite). Les cercles représentent les endroits d’évaluation.

Cette technique représente plusieurs avantages par rapport à une résolution par un réseau de neurones classique :

- En utilisant la méthode ODE, on n’a pas besoin de garder l’état d’activation dans le réseau. Ainsi on ne stocke pas de données inutiles pour la propagation de l’état du système et le coût mémoire est uniquement dépendant de la profondeur du réseau.
- L’évolution des méthodes de résolution des équations différentielles ordinaires a permis d’aboutir à une maîtrise du niveau d’erreur et une adaptabilité du pas afin d’optimiser la résolution (rapport précision/rapidité).
- La dynamique du système étant paramétrisée par une fonction continue du temps, les paramètres entre couches sont couplés, ce qui facilite l’apprentissage. Leur comportement est plus stable que celui des réseaux neuronaux récurrents discrets (RNN).

4.1.2 Apprentissage par équation adjointe

L’équivalent du formalisme de rétropropagation dans les ODE est l’utilisation de l’équation adjointe : en plus de l’équation directe, on va donc introduire une seconde équation différentielle ordinaire que l’on va résoudre en mode inverse, c’est-à-dire en remontant le temps à partir d’un état final. Cette équation va servir à propager une erreur dans le sens inverse dépendant de l’état du système à l’instant t , ce qui va permettre la mise à jour de l’état initial de manière plus efficace et précise qu’avec une simple propagation avant.

On introduit ϵ la fonction d’erreur lié à l’état $\mathbf{z}(t)$ que l’on cherche à optimiser. Pour un état au temps t_1 , ϵ va correspondre à une mesure de la différence entre l’état mesuré $\mathbf{z}_1 = \mathbf{z}(t_1)$ et l’état estimé de manière calculatoire $\hat{\mathbf{z}}_1 = \hat{\mathbf{z}}(t_1)$. Par exemple :

$$\epsilon = \|\mathbf{z}_1 - \hat{\mathbf{z}}_1\|^2.$$

Pour optimiser cette fonction, il faut dans un premier temps déterminer la dépendance de son gradient par rapport à l’état $\mathbf{z}(t)$ du système à tout instant. On appelle cette

quantité l'adjoint $\mathbf{a}(t) = \frac{\partial \epsilon(\mathbf{z})}{\partial \mathbf{z}(t)}$. Sa dynamique est soumise à une autre équation différentielle déterminée à partir de la composée des dérivées :

$$\frac{d\mathbf{a}(t)}{dt} = -\mathbf{a}(t)^T \frac{\partial f(\mathbf{z}(t), t, \theta)}{\partial \mathbf{z}}. \quad (4.4)$$

Ne connaissant pas l'état initial du système mais l'état final on connaît donc $\mathbf{a}(t_1)$. La résolution de cette équation nécessite la connaissance de $\mathbf{z}(t)$. Pour cela on va reconstruire l'évolution de $\mathbf{z}(t)$ en remontant le temps à l'aide de la valeur de l'adjoint à l'état final. De la même manière concernant le calcul du gradient par rapport au vecteur des paramètres θ on a :

$$\frac{d\epsilon}{d\theta} = \int_{t_1}^{t_0} \mathbf{a}(t)^T \frac{\partial f(\mathbf{z}(t), t, \theta)}{\partial \theta} dt \quad (4.5)$$

Le fonctionnement de l'algorithme de calcul de ces gradients de manière inverse peut être décomposé de la manière suivante :

1. Définir les paramètres d'entrée : θ , temps initial t_0 , temps final t_1 , état final \mathbf{z}_1 , gradient d'erreur $\frac{\partial \epsilon}{\partial \mathbf{z}(t_1)}$.
2. Définir l'état de départ $s_1 = [\mathbf{z}(t_1), \frac{\partial \epsilon}{\partial \mathbf{z}(t_1)}, \mathbf{0}(\theta)]$.
3. Définir les paramètres d'état dynamiques $D = ([\mathbf{z}(t), \mathbf{a}(t)], t, \theta)$.
4. Calculer les dérivées $[f(\mathbf{z}(t), t, \theta), -\mathbf{a}(t)^T \frac{\partial f}{\partial \mathbf{z}}, -\mathbf{a}(t)^T \frac{\partial f}{\partial \theta}]$.
5. Résoudre l'ODE en sens inverse $[\mathbf{z}(t_0), \frac{\partial \epsilon}{\partial \mathbf{z}(t_0)}, \frac{\partial \epsilon}{\partial \theta}] = ODEsolve(s_1, D, t_1, t_0, \theta)$.

Ainsi, on dispose le long de tout le « réseau » des gradients de l'erreur selon θ et \mathbf{z} , de même que dans un réseau de neurones discret on doit calculer les gradients en fonction des poids w et des entrées $x^{(n-1)}$ pour chaque couche.

4.1.3 Application à notre problème

Dans notre problème, les équations de réponse du détecteurs peuvent être simplifiées sous forme d'un système d'équations différentielles. Il s'en suit donc que détecteur en tant que système peut être « appris » grâce à ce type de méthode. Les paramètres comme le champ électrique et le piégeage agissent respectivement sur l'équation d'évolution de la position du nuage de porteurs libres et celle de la quantité de charges libres. Un vecteur paramètre Θ peut donc représenter l'état global du détecteur, en incorporant par exemple le champ électrique et le piégeage. Ensuite, chaque photon détecté va donner lieu à une trajectoire temporelle, d'un état initial (l'interaction) à un état final (la collecte), l'induction se faisant tout au long du trajet.

Par ailleurs, une partie de cet état est observable (le signal induit) et on va le considérer comme notre vecteur des sorties \mathbf{y} , tandis qu'une autre partie (le nuage électronique) est cachée mais déterminé à $t = 0$ par la physique de l'interaction rayonnement-matière. On va partager l'état en deux composantes en écrivant $\mathbf{X} = (\mathbf{x}, \mathbf{y})$ avec \mathbf{x} le vecteur des entrées et \mathbf{y} le vecteur de sortie qui évoluent selon une variable t et paramétrisés par un vecteur Θ . Pour résumer, pour chaque événement on a deux vecteurs dépendant du temps :

- \mathbf{x} est l'état du nuage électronique (charge et position) ;

— \mathbf{y} est l'état du signal induit (en fonction de l'électrode de mesure considéré).

Pour l'intégration, on peut définir des conditions aux limites suivantes :

- la charge initiale peut être connue (si l'énergie du rayonnement incident est connue) ;
- la profondeur finale est la position des anodes, lieu de collecte des électrons ;
- le charge induite initiale est nulle.

Cependant, en général on n'observe que la valeur finale de charge \mathbf{y}_1 ou une certaine intégration temporelle de $\mathbf{y}(t)$ via la réponse impulsionnelle du filtre de mise en forme utilisé.

On peut écrire l'équation différentielle régissant l'évolution temporelle de \mathbf{X} comme une fonction dépendant de \mathbf{X} à un instant donné ainsi que du vecteur des paramètres Θ :

$$\dot{\mathbf{X}} = \mathbf{f}(\mathbf{X}, \Theta) \quad (4.6)$$

Et en version intégrale :

$$\mathbf{X}(t) = \mathbf{X}(t_0) + \int_{t_0}^t \mathbf{f}(\mathbf{X}(t)) dt \quad (4.7)$$

Dans le cas de la propagation des informations dans le réseau, correspondant à la simulation du transport des porteurs de charge dans notre situation, on fixe à l'instant initial t_0 l'état $\mathbf{X}(t_0) = \mathbf{X}_0 = (\mathbf{x}_0, \mathbf{y}_0)$ et on calcule l'état du système à l'instant final t_1 $\mathbf{X}(t_1) = \mathbf{X}_1 = (\mathbf{x}_1, \mathbf{y}_1)$. Pour notre système, l'état à un instant t est défini par deux composantes : la position de l'interaction noté \mathbf{x} et le vecteur des signaux induits associé \mathbf{y} . Nous ne disposons pas d'information sur la position du photon incident et donc ne sommes pas en mesure d'accéder aux valeurs de \mathbf{x} réelles. C'est justement la valeur de \mathbf{x}_0 ainsi que celle de Θ définissant la signature de notre détecteur que nous allons chercher à estimer conjointement afin de mettre à jour correctement le réseau et approcher la réponse du détecteur réel. Nous allons nous baser sur la relation entre \mathbf{y}_1 et $\widehat{\mathbf{y}}_1$, le premier vecteur étant mesuré et le second estimé par propagation selon les équations du système, afin de calculer notre erreur et rétropropager sa dépendance à travers le temps inversé. Connaissant la cinétique d'évolution du système le calcul de \mathbf{X}_1 à partir de \mathbf{X}_0 est relativement simple. En revanche dans le cas où nous avons accès uniquement à l'état final du système \mathbf{X}_1 , il faut trouver un moyen de remonter à l'information sur l'état pour un instant quelconque t antérieur, plus particulièrement à l'état initial \mathbf{X}_0 . Notre objectif est d'implémenter l'accès à ces informations utilisant la méthode neural ODE.

La figure 4.2 permet d'apporter une visualisation plus concrète de la démarche que nous souhaitons mettre en place dans cette partie. L'objectif ici est de s'aider de données de simulation afin de déterminer le vecteur Θ qui est une signature de la forme du champ électrique dans un détecteur réel. On a donc à notre disposition un jeu de mesures correspondant à un état final du système étudié \mathbf{X}_1 . De manière physique le lien entre \mathbf{X}_1 et \mathbf{X}_0 est inconnu, nous supposons néanmoins que le champ électrique et donc le vecteur Θ reste constant. L'outil de simulation présenté précédemment permet de créer les vecteurs $\widehat{\mathbf{X}}_0$ et $\widehat{\mathbf{X}}_1$ qui équivalent à des états estimés pour un certain vecteur $\widehat{\Theta}$. Dans ce cas la relation entre les états initial et final est connue. On va alors procéder en quatre étapes :

1. On réalise une simulation à partir d'un état $\widehat{\mathbf{X}}_0$ quelconque afin d'obtenir un état final $\widehat{\mathbf{X}}_1$.

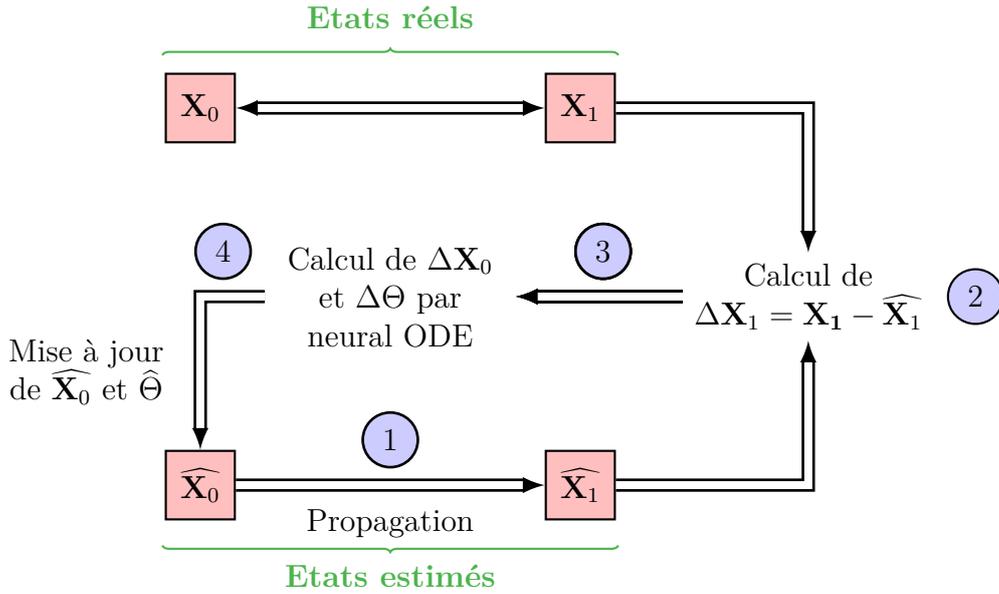


FIGURE 4.2 – Schéma de la démarche de l'apprentissage par neural ODE.

2. On calcule le vecteur d'erreur entre les états estimé et mesuré $\Delta \mathbf{X}_1 = \mathbf{X}_1 - \widehat{\mathbf{X}}_1$.
3. On utilise le calcul par neural ODE pour déterminer les valeurs de $\Delta \mathbf{X}_0$ et $\Delta \Theta$.
4. On met à jour l'état initial estimé $\widehat{\mathbf{X}}_0$ événement par événement et le vecteur $\widehat{\Theta}$ de manière globale.

Pour effectuer la simulation on intègre directement l'équation (4.6). L'inférence est plus compliquée à effectuer. On réalise un calcul de gradient à partir de la dérivation d'une fonction d'erreur noté $\epsilon(\mathbf{X}_1)$. On part du calcul de $\mathbf{X}_1^* = \frac{\partial \epsilon}{\partial \mathbf{X}_1}$ et on rétropropage le gradient pour mettre à jour l'état initial \mathbf{X}_0 d'après l'erreur $\mathbf{X}_0^* = \frac{\partial \epsilon}{\partial \mathbf{X}_0}$. Puis de manière globale pour tous les événements mesurés on calcule l'erreur sur les paramètres $\Theta^* = \frac{\partial \epsilon}{\partial \Theta}$.

Il faut déterminer la fonctionnelle d'erreur entre la valeur simulée par le réseau de neurones et la valeur mesurée. Dans notre cas, le bruit sur les mesures étant gaussien, on va considérer l'erreur quadratique entre les vecteurs des charges induites estimées et mesurées. Afin de calculer l'erreur entre les états initiaux et déterminer la mise à jour à appliquer aux différents paramètres pour minimiser cette erreur on va calculer le gradient de l'erreur par rapport à toutes les variables de l'état du système. Le calcul par équation différentielle adjointe va nous permettre d'effectuer cette différenciation en remontant de manière temporelle la propagation de l'équation d'évolution du système.

$$\epsilon = \|\mathbf{y}_1 - \mathbf{y}_1^{ref}\|^2 \quad (4.8)$$

La figure 4.3 permet d'avoir une idée de l'approche globale de l'estimation de l'état du système que l'on étudie à un moment quelconque $\widehat{\mathbf{X}}_t$ à partir de la propagation de la fonction f le long de la courbe temporelle entre les états estimés initiaux $\widehat{\mathbf{X}}_0$ et finaux $\widehat{\mathbf{X}}_1$. On utilise pour cela le vecteur d'erreur $\Delta \mathbf{X}_1$ entre les états observés et estimés. Le calcul de ce vecteur à partir d'une première estimation de l'état final va permettre à l'aide de la méthode neural ODE de faire remonter cette erreur à travers le temps jusqu'au calcul de l'erreur par rapport à l'état initial. De ce fait on pourra réaliser une nouvelle

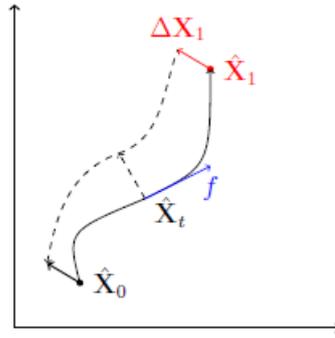


FIGURE 4.3 – Représentation de l'apprentissage non supervisé grâce à l'équation différentielle.

estimation de notre état final en propageant au travers de l'équation différentielle générée par f puis itérer ce processus jusqu'à minimiser l'erreur entre observation et estimation afin d'avoir une estimation réaliste de la fonction f , notamment de ses paramètres Θ comme mentionné précédemment.

4.1.4 Propagation avant

Comme décrit dans la partie précédente notre système est soumis à l'équation différentielle

$$\dot{\mathbf{X}} = \mathbf{f}(\mathbf{X}, \Theta). \quad (4.9)$$

Dans le cas de la propagation avant, c'est à dire déterminer l'état du système de manière directe en réalisant une simple propagation dans le sens d'écoulement du temps on cherche à dériver les équations d'évolution

$$\frac{\partial \dot{\mathbf{X}}}{\partial \mathbf{X}_0} = \frac{\partial \mathbf{f}}{\partial \mathbf{X}_0} \quad (4.10)$$

$$\frac{\partial \dot{\mathbf{X}}}{\partial \Theta} = \frac{\partial \mathbf{f}}{\partial \Theta} \quad (4.11)$$

Les équations (4.6), (4.10) et (4.11) peuvent être intégrées en parallèle dans le sens des temps croissants. A partir d'un état quelconque \mathbf{X}_0 on peut obtenir le résultat \mathbf{X}_1 ainsi que les gradients associés par rapport à \mathbf{X}_0 et Θ . On obtient alors dans un premier temps :

$$\frac{\partial \mathbf{X}}{\partial \mathbf{X}_0} = \int_{t_0}^t \frac{\partial \mathbf{f}}{\partial \mathbf{X}_0} dt + C \quad (4.12)$$

avec C une matrice dont on peut déterminer la valeur grâce à la valeur initiale évaluée en $\mathbf{X} = \mathbf{X}_0$. On arrive alors à

$$\frac{\partial \mathbf{X}_0}{\partial \mathbf{X}_0} = \text{Id} = C \text{ car } \int_{t_0}^{t_0} \frac{\partial \mathbf{f}}{\partial \mathbf{X}_0} dt = 0 \quad (4.13)$$

On aboutit alors à l'expression

$$\frac{\partial \mathbf{X}_1}{\partial \mathbf{X}_0} = \text{Id} + \int_{t_0}^{t_1} \frac{\partial \mathbf{f}}{\partial \mathbf{X}_0} dt \quad (4.14)$$

Cette expression n'est facilement intégrable que dans certains cas particulier, comme le cas unidimensionnel qu'on verra ci-après. On réalise le même raisonnement pour l'équation (4.11) sachant que $\frac{\partial \mathbf{X}_0}{\partial \Theta} = 0$ et on trouve

$$\frac{\partial \mathbf{X}_1}{\partial \Theta} = \int_{t_0}^{t_1} \frac{\partial \mathbf{f}}{\partial \Theta} dt. \quad (4.15)$$

Finalement on peut déterminer la valeur des gradients \mathbf{X}_0^* et Θ^* :

$$\mathbf{X}_0^* = \frac{\partial \epsilon}{\partial \mathbf{X}_0} = \left\langle \frac{\partial \epsilon}{\partial \mathbf{X}_1}, \frac{\partial \mathbf{X}_1}{\partial \mathbf{X}_0} \right\rangle = \left\langle \mathbf{X}_1^*, \frac{\partial \mathbf{X}_1}{\partial \mathbf{X}_0} \right\rangle = \left\langle \mathbf{X}_1^*, \mathbf{Id} + \int_{t_0}^{t_1} \frac{\partial \mathbf{f}}{\partial \mathbf{X}_0} dt \right\rangle \quad (4.16)$$

$$\Theta^* = \frac{\partial \epsilon}{\partial \Theta} = \left\langle \frac{\partial \epsilon}{\partial \mathbf{X}_1}, \frac{\partial \mathbf{X}_1}{\partial \Theta} \right\rangle = \left\langle \mathbf{X}_1^*, \frac{\partial \mathbf{X}_1}{\partial \Theta} \right\rangle = \left\langle \mathbf{X}_1^*, \int_{t_0}^{t_1} \frac{\partial \mathbf{f}}{\partial \Theta} dt \right\rangle. \quad (4.17)$$

L'avantage de cette formulation est qu'elle permet de disposer explicitement de tous les gradients au niveau de \mathbf{X}_1 , incluant leur corrélation. Avoir $\frac{\partial \mathbf{y}_1}{\partial (\mathbf{x}_0, \Theta)}$, qui représente donc la dérivée du vecteur des signaux induits par rapport respectivement à l'état initial du nuage électronique et au vecteur des paramètres d'état du détecteur, permet l'orthogonalisation et la possibilité de calculer la base duale pour une mise à jour découplée des paramètres, mais cela représente davantage de données à traiter et à stocker. En pratique, dans l'espace des mesures \mathbf{y}_1 , on a couramment une métrique euclidienne (bruit blanc gaussien).

Si cette intégration est praticable dans un cas à une dimension, car la trajectoire est figée, elle ne l'est pas dans les cas plus complexes. En effet, le principe de chaînage fait que l'on a besoin à tout instant des gradients des temps ultérieurs afin de remonter le temps. A chaque instant t on a une contribution du gradient en $\mathbf{x}(t + dt)$ et une du gradient en \mathbf{y} à l'instant considéré. De cette manière, pour calculer le gradient $\mathbf{x}(t = 0)$, on a besoin du trajet inverse et donc de la rétropropagation. C'est pour cela que dans la partie suivante on s'intéresse aux calculs en sens inverse afin de s'adapter aux situations les plus complexes.

4.1.5 Rétropropagation

La rétropropagation permet de travailler directement avec le gradient $\nabla \epsilon = (\mathbf{X}^*, \Theta^*)$ de la fonction d'erreur par rapport à l'état du système et des paramètres et non une jacobienne $\frac{\partial \mathbf{X}}{\partial \mathbf{X}_0}$ à propager comme dans le cas de la propagation avant. De manière duale au gradient on a un vecteur d'erreur entre observation et estimation $(\Delta \mathbf{X}, \Delta \Theta)$.

On part de l'erreur observée $\Delta \mathbf{X}_1 = \mathbf{X}_1 - \hat{\mathbf{X}}_1$ pour calculer $\Delta \mathbf{X}_0$ et $\Delta \Theta$ les mises à jour de l'état initial voulues. On cherche donc à obtenir la propagation du gradient entre t_1 et un temps antérieur t_0 . On définit l'erreur $\Delta \epsilon$ de manière analogue aux gradients précédents sauf que dans ce cas on va rétropropager cette erreur grâce à $\Delta \mathbf{X}_1$. Cette erreur se conservant le long de la propagation temporelle on va utiliser cette propriété afin de déterminer l'équation de propagation adjointe :

$$\Delta\epsilon = \left\langle \frac{\partial\epsilon}{\partial\mathbf{X}_1}, \Delta\mathbf{X}_1 \right\rangle = \left\langle \frac{\partial\epsilon}{\partial\mathbf{X}_0}, \Delta\mathbf{X}_0 \right\rangle \quad (4.18)$$

$$\left\langle \mathbf{X}_1^*, \Delta\mathbf{X}_0 + \int_{t_0}^{t_1} \Delta\mathbf{f} dt \right\rangle = \left\langle \mathbf{X}_1^* - \int_{t_0}^{t_1} \dot{\mathbf{X}} dt, \Delta\mathbf{X}_0 \right\rangle \quad (4.19)$$

$$\left\langle \mathbf{X}_1^*, \int_{t_0}^{t_1} \Delta\mathbf{f} dt \right\rangle = \left\langle - \int_{t_0}^{t_1} \dot{\mathbf{X}} dt, \Delta\mathbf{X}_0 \right\rangle \quad (4.20)$$

Si on considère que t_0 et t_1 sont des instants infinitésimalement proches on peut approximer que $\mathbf{X}_0 = \mathbf{X}_1$ et alors

$$\left\langle \mathbf{X}^*, \Delta\mathbf{f} \right\rangle = - \left\langle \dot{\mathbf{X}}^*, \Delta\mathbf{X} \right\rangle \quad (4.21)$$

Et pour un écart $\Delta\mathbf{X} \rightarrow 0$, à un instant t on connaît l'état \mathbf{X} , on peut propager \mathbf{X}^* appliqué aux deux états que l'on appelle \mathbf{X}_u et \mathbf{X}_v avec l'équation suivante :

$$\dot{\mathbf{X}}_u^* = - \sum_v \mathbf{X}_v^* \frac{\partial\mathbf{f}^v}{\partial\mathbf{X}^u} \quad (4.22)$$

On obtient la même chose pour Θ , à la différence près que l'on travaille sur un vecteur de paramètres on va donner noter pour l'indice a de ce vecteur :

$$\dot{\Theta}_a^* = - \sum_v \mathbf{X}_v^* \frac{\partial\mathbf{f}^v}{\partial\Theta^a} \quad (4.23)$$

A partir du gradient $\partial_\mu\epsilon$ plusieurs choix de mise à jour sont possibles. L'objectif est de déterminer le vecteur de modification $\Delta\mathbf{x}$ à appliquer en corrigeant le gradient par la métrique du problème. Il existe différentes techniques [94, 97] plus ou moins complexe et performantes : Newton, Gauss-Newton, Levenberg-Marquardt, etc. Dans notre cas, on décide d'utiliser la méthode du Hessien diagonal. On profite du fait que le bruit soit connu comme gaussien, décorrélé, identique et uniforme en \mathbf{y} pour calculer les termes diagonaux de la métrique.

$$\frac{\partial^2\epsilon}{\partial\mathbf{y}_\mu\partial\mathbf{y}_\nu} = g_{\mu\nu} = \left\langle \frac{\partial\mathbf{y}^\alpha}{\partial\mathbf{x}_\mu} \delta_{\alpha\alpha} \epsilon \frac{\partial\mathbf{y}^\alpha}{\partial\mathbf{x}_\nu} \right\rangle \quad (4.24)$$

avec $g_{\mu\nu} = \partial_{\mu\nu}\epsilon$ la métrique engendrée par la fonction d'erreur ϵ dans l'espace des \mathbf{X} , $\frac{\partial\mathbf{y}^\alpha}{\partial\mathbf{x}_\nu}$ la dérivée du signal induit α par rapport à la sensibilité de l'erreur à l'état du système et $\delta_{\alpha\alpha}\epsilon$ la métrique engendrée par ϵ pour le signal α .

4.2 Cas à une dimension

4.2.1 Équations du problème

Dans cette partie on s'intéresse à un problème en une dimension. C'est à dire que l'on considère uniquement le transport des électrons selon la profondeur du détecteur afin d'avoir une première approche plus simple à implémenter au niveau des différents calculs et de la mise à jour via l'apprentissage.

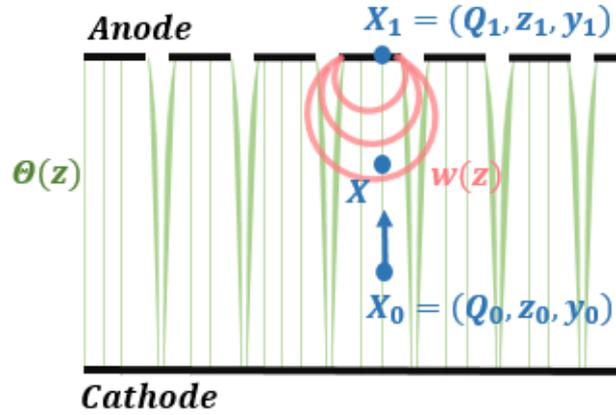


FIGURE 4.4 – Schéma du cas d'un transport unidimensionnel selon la profondeur

Sur la figure 4.4 on peut observer la situation dans laquelle on effectue cette première étude et les différentes variables qui vont être mises en jeu. Pour chaque événement on a \mathbf{X}_0 l'état initial du système, avec $\mathbf{X}_0 = (\mathbf{x}_0, \mathbf{y}_0) = (Q_0, z_0, \mathbf{y}_0)$, Q_0 étant la charge déposée lors de l'interaction, z_0 la profondeur de l'interaction et $\mathbf{y}_0 = 0$ le vecteur des signaux induits, étant nul car aucune induction n'a été réalisée lors de l'interaction. Les porteurs de charge vont évoluer le long des lignes de champ paramétrée par le vecteur Θ définissant le champ électrique du détecteur. A proximité de l'anode les porteurs de charge vont être soumis au champ de pondération $\mathbf{w}(z)$ et vont terminer leur trajet en induisant leur signal sur l'anode nous donnant l'état final du système \mathbf{X}_1 . L'état du nuage électronique pendant le transport est donné par sa charge Q et sa position z . On a à l'état initial

$$\mathbf{x}_0 = \mathbf{x}(t = 0) = (Q_0, z_0) \quad (4.25)$$

L'état final à $t = t_1$ est

$$\mathbf{x}_1 = \mathbf{x}(t = 1) = (Q_1, z_1 = z_{anode}) \quad (4.26)$$

On note \mathbf{y} le vecteur des signaux induits. On considère pour le moment que ces signaux sont des charges induites. A l'état initial on a $\mathbf{y}_0 = 0$ et la valeur finale est \mathbf{y}_1 . On note

- $u(z) = \mu E$ la vitesse selon z ,
- $\mathbf{w}(z)$ le champ de pondération en z ,
- τ la constante de piégeage des électrons.

On paramétrise u avec un paramètre Θ . Les équations d'évolution temporelle du système sont :

$$\frac{d}{dt} \begin{pmatrix} Q \\ z \\ \mathbf{y} \end{pmatrix} = \begin{pmatrix} -Q(t)/\tau \\ u(z) \\ Q(t)u(z)\mathbf{w}(z) \end{pmatrix} \quad (4.27)$$

Cependant on a intérêt à éliminer la variable temporelle pour passer en z , ce qui simplifie le traitement en éliminant une équation, on obtient alors :

$$\frac{d}{dt} \begin{pmatrix} Q \\ \mathbf{y} \end{pmatrix} = \begin{pmatrix} -Q(z) \\ u(z)\tau \\ Q(z)\mathbf{w}(z) \end{pmatrix} \quad (4.28)$$

Les bornes d'intégration vont de z_0 à z_1 sont toutes les deux connues contrairement au temps t_1 , qui est dépendant de z_0 et $u(z)$.

Grâce au changement de variable entre t et z , l'intégration analytique du problème est réalisable, soit de manière directe, soit par rétropropagation (voir Annexe A). Les trois équations de gradient résultantes sont :

$$\frac{\partial \mathbf{y}}{\partial Q} = \frac{\mathbf{y}_1}{Q_0} \quad (4.29)$$

$$\frac{\partial \mathbf{y}}{\partial z_0} = \frac{\mathbf{y}_1}{u_0\tau} - Q_0\mathbf{w}_0 \quad (4.30)$$

$$\frac{\partial \mathbf{y}}{\partial \Theta} = \int_{t_0}^{t_1} \overleftarrow{\mathbf{y}} \frac{\partial u}{u\tau} dt \quad (4.31)$$

où $\overleftarrow{\mathbf{y}}$ correspond à la charge induite au point considéré avec une propagation inverse (partant de l'anode) tel que $\overleftarrow{\mathbf{y}} = \mathbf{y}_1 - \mathbf{y}(t)$.

4.2.2 Mise à jour avec le calcul direct

Calcul des valeurs de mise à jour

Pour chaque événement mesuré \mathbf{y}_1 on effectue la procédure suivante :

1. On initialise (Q_0, z_0) et on effectue une première propagation.
2. On obtient une estimation $\hat{\mathbf{y}}_1$ et on calcule l'écart $\Delta \mathbf{y}_1 = \mathbf{y}_1 - \hat{\mathbf{y}}_1$.
3. On utilise les gradients $\frac{\partial \mathbf{y}_1}{\partial Q_0}$ et $\frac{\partial \mathbf{y}_1}{\partial z_0}$ pour calculer les mises à jour de l'événement :

$$\Delta Q_0 = \frac{\left\langle \Delta \mathbf{y}_1, \frac{\partial \mathbf{y}_1}{\partial Q_0} \right\rangle}{\left\| \frac{\partial \mathbf{y}_1}{\partial Q_0} \right\|^2} \quad (4.32)$$

$$\Delta z_0 = \frac{\left\langle \Delta \mathbf{y}_1, \frac{\partial \mathbf{y}_1}{\partial z_0} \right\rangle}{\left\| \frac{\partial \mathbf{y}_1}{\partial z_0} \right\|^2} \quad (4.33)$$

4. On itère pour obtenir une estimation de l'événement (Q_0, z_0) . Si Q_0 est connue, on la laisse fixe. Si elle est inconnue mais égale pour tous les événements, on la met à jour globalement comme Θ , dans l'étape suivante.
5. On extrait l'écart de Θ donné par l'événement, que l'on moyenne avec les contributions des autres événements.

$$\Delta \Theta = \frac{\left\langle \Delta \mathbf{y}_1, \frac{\partial \mathbf{y}_1}{\partial \Theta} \right\rangle}{\left\| \frac{\partial \mathbf{y}_1}{\partial \Theta} \right\|^2} \quad (4.34)$$

Application à un cas simple

On va appliquer la méthode de mise à jour par gradient direct à un cas simple en une dimension. Pour cela on simule un détecteur plan à partir de différentes constantes : sa longueur L , les constantes de mobilité et de temps de vie des électrons et des trous dans le CZT μ_i et τ_i ainsi que le coefficient d'atténuation linéique du CZT. On définit le champ électrique comme étant linéaire et paramétré par une constante α qui correspond à notre vecteur Θ introduit précédemment. On a alors $E = E_0 + \frac{\alpha}{L}z$ avec z , la profondeur dans le détecteur et E_0 la valeur du champ en $z = 0$. On ne s'intéresse pas ici à l'impact de la charge considérant une charge initiale connue $Q_0 = 1$. On réalise deux tirages aléatoires afin d'obtenir deux listes de profondeurs initiales d'interaction, une mesurée et une estimée avec pour chacune d'elle une valeur de α différente. On introduit deux potentiels de pondération w_1 et w_2 pour les calculs de charge induite sur chacune des électrodes. On calcule la charge induite par électrode avec $y = Q \cdot w$ par résolution numérique grâce à la méthode de Runge Kutta 4.

On peut alors faire l'analogie avec la partie antérieure pour le calcul de la mise à jour des gradients. En reprenant l'équation (4.32) on a $\Delta y_1 = y_1 - \widehat{y}_1$ que l'on calcule avec la différence de charges induites par mesure et par estimation. Les calculs de charge induite se font numériquement. Ensuite on a $\frac{\partial y_1}{\partial z_0} = \int_{z_0}^{z_{anode}} \frac{\partial Q}{\partial z_0} w dz$ avec $\frac{\partial Q}{\partial z_0} = \frac{Q}{u_0 \tau}$, ce qui donne $\frac{\partial y_1}{\partial z_0} = \frac{y_1}{u_0 \tau}$. On arrive donc à une expression simplifiée

$$\Delta z_0 = \frac{\Delta y_1}{\frac{y_1}{u_0 \tau}}. \quad (4.35)$$

On peut alors appliquer cette mise à jour à nos données estimées en itérant plusieurs fois afin de se rapprocher des valeurs de profondeurs d'interaction mesurées de manière expérimentale (ou provenant d'une seconde simulation dans ce cas).

Sur la figure 4.5, on peut remarquer que l'algorithme arrive rapidement à se rapprocher des valeurs mesurées, avec de très bons résultats pour cinq itérations. Après s'être intéressés à la mise à jour de z_0 on va appliquer la mise à jour au vecteur Θ , qui est ici notre constante α . L'équation (4.34) est semblable à celle appliquée pour z_0 , il faut néanmoins calculer une expression de $\frac{\partial y_1}{\partial \alpha}$. D'après l'équation (4.31) On a

$$\int_{z_0}^{z_1} \frac{\overleftarrow{y}}{u^2 \tau} \frac{\partial u}{\partial \Theta} dz \text{ où } \overleftarrow{y} = y_1 - y \quad (4.36)$$

Dans notre cas on a $\frac{\partial u}{\partial \alpha} = z$, et $\overleftarrow{y} = \int_z^{z_{anode}} Q w dz = \int_z^{z_{anode}} \frac{Q_0}{L} e^{-\frac{z}{u\tau}} dz$. Finalement on arrive à l'expression

$$\frac{\partial y_1}{\partial \alpha} = \int_{z_0}^{z_{anode}} \int_z^{z_{anode}} \frac{Q_0}{L u^2 \tau} e^{-\frac{z}{u\tau}} z dz dz' \quad (4.37)$$

On calcule ces intégrales de manière numérique afin d'aboutir au calcul de $\Delta \alpha$. On réalise cette opération pour chaque événement estimé, puis on moyenne la somme afin de déterminer la valeur de mise à jour de α à appliquer pour se rapprocher des valeurs mesurées.

On peut remarquer sur la figure 4.6 que la mise à jour de α par l'algorithme neural ODE permet aux valeurs estimées de coïncider avec les valeurs mesurées de manière très convaincante. Ce graphique est obtenu après uniquement une itération ce qui prouve l'efficacité de cette méthode de mise à jour.

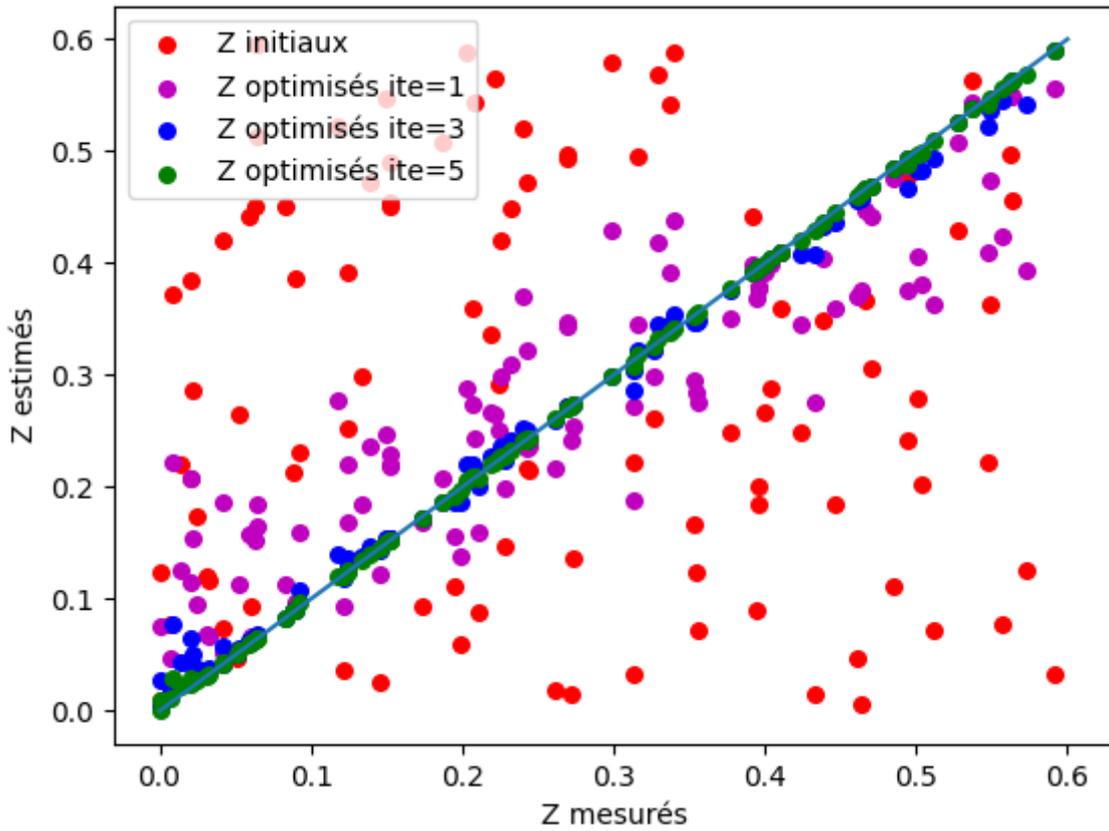


FIGURE 4.5 – Graphique des profondeurs d’interaction mesurées en fonction des profondeurs estimées par l’algorithme neural ODE.

4.3 Cas à trois dimensions

4.3.1 Équations du problème

De la même manière que pour le problème unidimensionnel, on s’intéresse à l’état du nuage électronique durant le transport donné par sa charge Q et sa position $\vec{r} = (x, y, z)$. Initialement on a

$$\mathbf{x}_0 = \mathbf{x}(t = 0) = (Q_0, x_0, y_0, z_0). \quad (4.38)$$

Et l’état final à $t = t_1$ est

$$\mathbf{x}_1 = \mathbf{x}(t = 1) = (Q_1, x_1, y_1, z_1 = z_{anode}). \quad (4.39)$$

On va également introduire des notations avec indices covariants et contravariants pour simplifier les écritures. Les équations de transport temporelles sont

$$\mathbf{f} = \frac{d}{dt} \begin{vmatrix} Q \\ \vec{r} \\ \mathbf{y} \end{vmatrix} = \begin{vmatrix} -\frac{Q}{\tau} \\ \vec{u} \\ Q(\vec{\mathbf{w}} \cdot \vec{u}) \end{vmatrix} = \begin{vmatrix} -\frac{Q}{\tau} \\ \mathbf{u}^i \\ Q\mathbf{w}_i^\alpha \mathbf{u}^i \end{vmatrix} \quad (4.40)$$

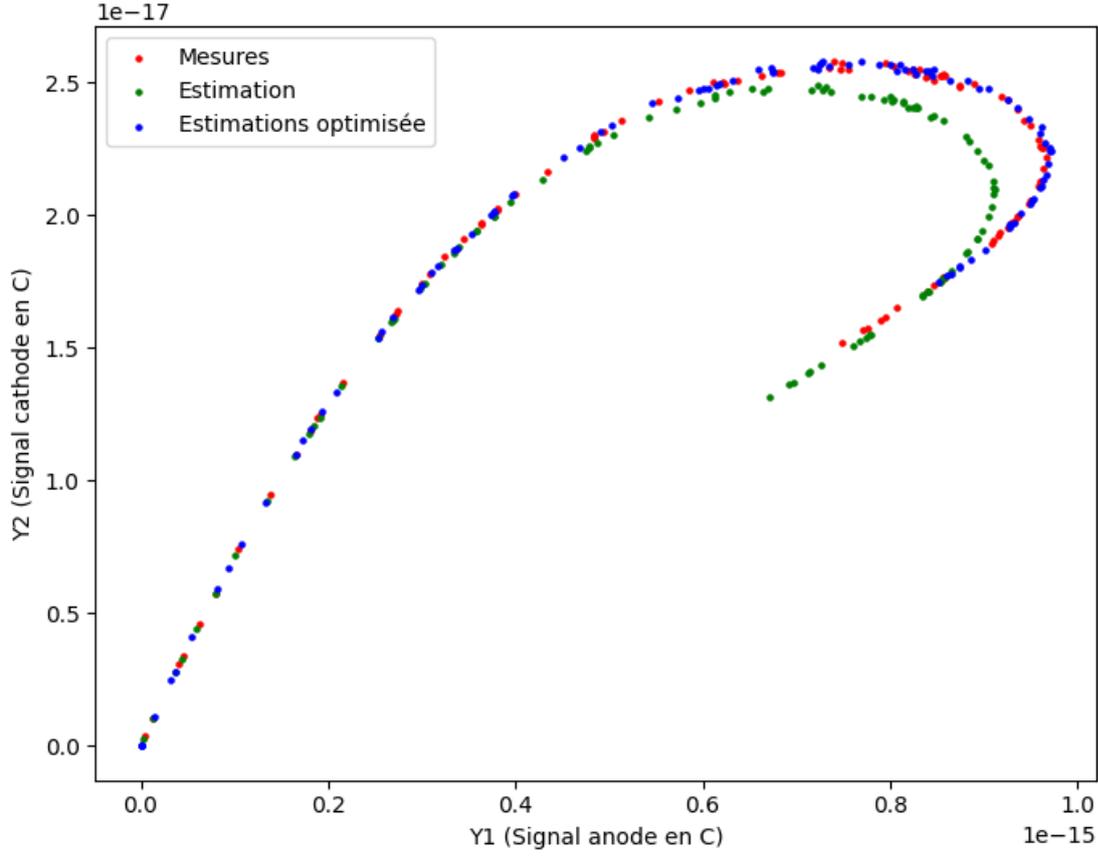


FIGURE 4.6 – Graphique des signaux cathode et anode dans le cas des événements mesurés, estimés et mis à jour grâce à la méthode neural ODE.

$$\frac{\partial \mathbf{f}}{\partial \mathbf{X}} = \begin{bmatrix} -\frac{1}{\tau} & 0 & 0 \\ 0 & \vec{\nabla} \vec{u} & 0 \\ \tilde{\mathbf{w}} \cdot \vec{u} & Q \vec{\nabla}(\tilde{\mathbf{w}} \cdot \vec{u}) & 0 \end{bmatrix} = \begin{bmatrix} -\frac{1}{\tau} & 0 & 0 \\ 0 & \partial_j \mathbf{u}^i & 0 \\ \mathbf{w}_i^\alpha \mathbf{u}^i & Q \partial_j(\mathbf{w}_i^\alpha \mathbf{u}^i) & 0 \end{bmatrix} \quad (4.41)$$

$$\frac{\partial \mathbf{f}}{\partial \Theta} = \begin{vmatrix} 0 \\ \frac{\partial \vec{u}}{\partial \Theta} \\ Q(\tilde{\mathbf{w}} \cdot \frac{\partial \vec{u}}{\partial \Theta}) \end{vmatrix} = \begin{vmatrix} 0 \\ \partial_a \mathbf{u}^i \\ Q \tilde{\mathbf{w}}^\alpha \partial_a \mathbf{u}^i \end{vmatrix} \quad (4.42)$$

En ajoutant le filtrage, permettant d'ajouter l'influence de l'électronique dans notre système d'équations, via sa réponse impulsionnelle F . On obtient

$$\mathbf{f} = \frac{d}{dt} \begin{vmatrix} Q \\ \vec{r} \\ \mathbf{y} \end{vmatrix} = \begin{vmatrix} -\frac{Q}{\tau} \\ \vec{u} \\ F \cdot Q(\tilde{\mathbf{w}} \cdot \vec{u}) \end{vmatrix} \quad (4.43)$$

L'ajout du filtrage reflète le fait que dans un système réel, on mesure l'amplitude d'impulsions mises en forme. Le choix de la mise en forme est dicté par l'optimisation du rapport signal sur bruit, et s'effectue en fonction de la densité spectrale de puissance moyenne du bruit et du signal porteur de l'information qui nous intéresse (énergie, temps, position d'interaction) [75].

4.3.2 Résolution directe

Dans l'hypothèse où l'on est capable d'obtenir une CIE(x,y,z) en direct, la résolution ne demande pas d'intégration numérique. Le gradient spatial de la CIE nous permet de déterminer la position d'interaction. Sa norme carré nous donne la métrique correspondante qui est $\frac{\partial_i C}{\langle \partial_i C, \partial_i C \rangle}$. Enfin, le champ peut être mis à jour le long de la trajectoire. Cependant, il est compliqué dans le cas général d'obtenir des cartes de CIE dans la mesure où :

- Les charges sont piégées, donc le potentiel de pondération seul ne suffit pas et la forme du champ électrique a toute son importance.
- Le champ électrique est quelconque, $\Delta \mathbf{u} \neq 0$ donc les méthodes de calcul adjoint telles que présentées dans [26, 65] ne sont pas applicables.
- De part la mise en forme électronique, les valeurs mesurées ne sont pas égales à la charge collectée. La dépendance est fonction de la forme du signal transitoire, ce qu'on appelle phénomène de déficit balistique.

Cependant, si on sait calculer rapidement la cartographie de CIE, ce procédé est potentiellement utilisable. En 2D, cela peut être potentiellement avantageux. En 3D, on a besoin à peu près d'un million d'événements pour calculer une CIE et celle-ci doit être recalculée à chaque mise à jour de champ électrique ce qui rend le processus trop coûteux en temps de calcul.

4.3.3 Rétropropagation

On écrit l'équation d'évolution des gradients :

$$\frac{d}{dt} \begin{vmatrix} Q^* \\ \mathbf{r}^* \\ \Theta^* \end{vmatrix} = \frac{d}{dt} \begin{vmatrix} \partial_Q \\ \partial_j \epsilon \\ \partial_a \end{vmatrix} = - \begin{bmatrix} -\frac{1}{\tau} & 0 & \mathbf{w}_i^\alpha \mathbf{u}^i \\ 0 & \partial_j \mathbf{u}^i & Q \partial_j (\mathbf{w}_i^\alpha \mathbf{u}^i) \\ 0 & \partial_a \mathbf{u}^i & Q \mathbf{w}_i^\alpha \partial_a \mathbf{u}^i \end{bmatrix} \begin{vmatrix} \partial_Q \\ \partial_i \epsilon \\ \partial_\alpha \end{vmatrix} \quad (4.44)$$

Charge du nuage électronique

On résout l'équation de la même manière que le cas 1D :

$$\dot{Q}^* = -\frac{1}{\tau} Q^* + \mathbf{y}_\alpha^* \mathbf{w}_i^\alpha \mathbf{u}^i \quad (4.45)$$

$$Q^* = \mathbf{y}_\alpha^* e^{t/\tau} \int_t^{t_1} e^{-t'/\tau} \mathbf{w}_i^\alpha \mathbf{u}^i dt' \quad (4.46)$$

$$= \mathbf{y}_\alpha^* \frac{1}{Q} \int_t^{t_1} Q \mathbf{w}_i^\alpha \mathbf{u}^i dt' \quad (4.47)$$

Ce qui pour $t = 0$ donne le résultat déjà vu dans le cas 1D :

$$Q_0^* = \frac{\partial \epsilon}{\partial Q_0} = \frac{\partial \epsilon}{\partial \alpha} \cdot \frac{\mathbf{y}_1^\alpha}{Q_0} \quad (4.48)$$

Cela signifie que la variation de la composante signal α avec la charge du nuage Q en un instant donné est proportionnelle à l'induction sur le trajet jusqu'à l'anode divisée par la charge courante, c'est à dire la CIE correspondant à la position du nuage.

Position du nuage électronique

On reprend toujours la méthode du 1D appliqué aux équations de la 3D

$$-\frac{\partial \dot{\epsilon}}{\partial \mathbf{r}^j} = \frac{\partial \epsilon}{\partial \mathbf{r}^i} \cdot \frac{\partial \mathbf{u}^i}{\partial \mathbf{r}^j} + \frac{\partial \epsilon}{\partial \mathbf{y}^\alpha} \cdot Q \frac{\partial (\mathbf{w}_i^\alpha \mathbf{u}^i)}{\partial \mathbf{r}^j} \quad (4.49)$$

En pratique, et pour faciliter le calcul de la métrique, on va calculer la propagation des gradients des signaux de sortie :

$$-\partial_j \dot{\mathbf{y}}^\alpha = \partial_i \mathbf{y}^\alpha \cdot \partial_j \mathbf{u}^i + Q \partial_j (\mathbf{w}_i^\alpha \mathbf{u}^i) \quad (4.50)$$

Comme la métrique est euclidienne dans l'espace des \mathbf{y} , la métrique en espace est proportionnelle à :

$$g_{ij} = \partial_i \mathbf{y}^\alpha \delta_{\alpha\beta} \partial_j \mathbf{y}^\beta \quad (4.51)$$

Paramètres de champ

Finalement on arrive aux équations concernant les paramètres de champ

$$-\frac{\partial \epsilon}{\partial \Theta^a} = \left(\frac{\partial \epsilon}{\partial \mathbf{r}^i} + \frac{\partial \epsilon}{\partial \mathbf{y}^\alpha} Q \mathbf{w}_i^\alpha \right) \frac{\partial \mathbf{u}^i}{\partial \Theta^a} \quad (4.52)$$

$$-\partial_a \dot{\epsilon} = \partial_i \epsilon \partial_a \mathbf{u}^i + \partial_\alpha \epsilon Q \mathbf{w}_i^\alpha \partial_a \mathbf{u}^i \quad (4.53)$$

$$-\dot{\Theta}_a^* = (\mathbf{r}_i^* + \mathbf{y}_\alpha^* \cdot Q \mathbf{w}_i^\alpha) \partial_a \mathbf{u}^i \quad (4.54)$$

À noter qu'on peut simplifier en prenant comme paramètres intermédiaire la valeur locale du champ

$$-\dot{\mathbf{u}}_i^* = \mathbf{r}_i^* + \mathbf{y}_\alpha^* Q \mathbf{w}_i^\alpha \quad (4.55)$$

Si on s'intéresse au gradient par rapport aux signaux de sortie on obtient

$$-\frac{\partial \dot{\mathbf{y}}^\alpha}{\partial \Theta^a} = \left(\frac{\partial \mathbf{y}^\alpha}{\partial \mathbf{u}^i} + Q \mathbf{w}_i^\alpha \right) \frac{\partial \mathbf{u}^i}{\partial \Theta^a} \quad (4.56)$$

4.4 Développement de l'algorithme

On rappelle que l'objectif de l'algorithme d'apprentissage non supervisé des paramètres détecteur est de servir de relais entre les données expérimentales et l'apprentissage supervisé des positions et de l'énergie. Il permet d'aboutir à une simulation dont la réponse correspond à la réponse du détecteur étudié afin d'entraîner le réseau de neurones implémenté dans l'électronique. Le fonctionnement de cet algorithme est schématisé sur la figure 4.7.

4.4.1 Implémentation et fonctionnement

Nous allons dans un premier temps décrire la construction ainsi que le fonctionnement de cette algorithme. Il se compose de deux briques élémentaires :

- Le parcours direct, permettant de simuler le trajet des porteurs de charge depuis le lieu d'interaction jusqu'à leur collecte (*forward*). Ce premier élément permet de calculer le vecteur des signaux induits \mathbf{y}_1 ainsi que les positions de collecte et la charge finale \mathbf{x}_1 à partir d'une hypothèse \mathbf{x}_0 sur l'interaction.
- Le parcours inverse, qui détermine la trajectoire des porteurs de charge de la collecte jusqu'à leur lieu d'interaction (*backward*). Celui-ci permet d'obtenir les gradients de signaux induits par rapport au lieu d'interaction ainsi que par rapport au champ électrique. On peut alors mettre à jour les paramètres initiaux de l'interaction \mathbf{x}_0 (interaction par interaction) et les paramètres de champ \vec{u} (de manière globale) par gradient stochastique.

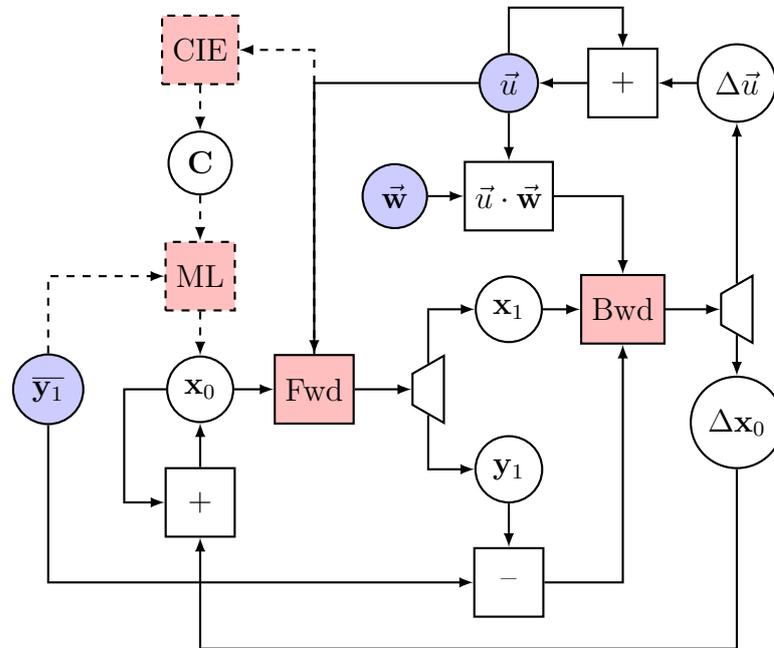


FIGURE 4.7 – Implémentation de l'algorithme d'apprentissage 3D.

Dans un premier temps on s'intéresse à l'initialisation des différents vecteurs. On commence par initialiser les positions initiales d'interaction \mathbf{x}_0 dans un premier temps de manière arbitraire. On utilise ces positions en tant que données d'entrée ainsi que le champ électrique initial \vec{u} dans le programme *Forward*. En sortie on obtient d'un côté les positions finales après collecte \mathbf{x}_1 , et de l'autre le vecteur des signaux induits associés \mathbf{y}_1 . Ce dernier est comparé avec les signaux induits $\bar{\mathbf{y}}_1$ générés de manière expérimentale ou à partir d'une seconde simulation dissociée tenant lieu de référence. La différence dans ces signaux est un paramètre d'entrée du programme *Backward*, ainsi que \mathbf{x}_1 et le produit $\vec{u} \cdot \vec{w}$ calculé au préalable, \vec{w} étant constant. Puis à la sortie de cette simulation inverse on récupère événement par événement la correction sur les positions initiales $\Delta \mathbf{x}_0$, permettant de mettre à jour \mathbf{x}_0 pour chaque interaction, et dans un second temps grâce

à la correction sur le champ $\Delta\vec{u}$ on met à jour ce dernier dans le *Forward* ainsi que pour calculer le nouveau produit $\vec{u} \cdot \vec{w}$.

Au lieu d’initialiser les positions initiales de manière quelconque, on peut utiliser la méthode décrite par les pointillés afin de se rapprocher de données réalistes. Ainsi on utilise le champ \vec{u} ainsi que le vecteur des signaux observés \overline{y}_1 pour initialiser \mathbf{x}_0 . Pour cela on génère une CIE à partir du champ électrique, et le couplage de celle-ci avec \overline{y}_1 et un algorithme de localisation classique (barycentre, maximum de vraisemblance) permet d’obtenir une initialisation plus réaliste des positions de départ.

4.4.2 Résultats sur simulation

La première étape à valider est la partie localisation, c’est à dire le fait de déterminer \mathbf{x}_0 à partir des \overline{y}_1 observés. De manière à limiter la dimension du vecteur \mathbf{y} , on se limite à un voisinage de 2×2 anodes et donc on a quatre signaux observés.

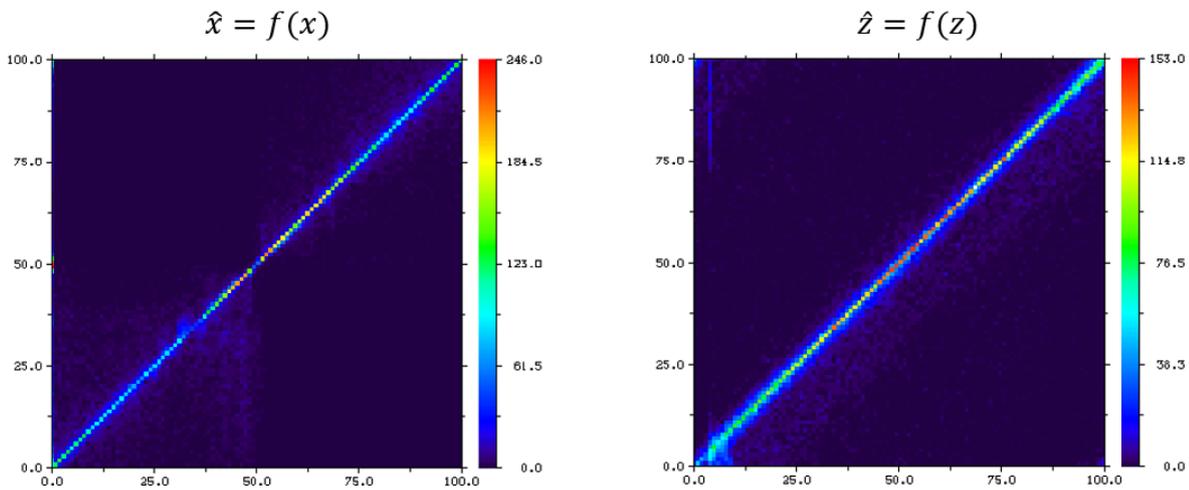


FIGURE 4.8 – Histogramme de la coordonnée X (à gauche) et Z (à droite) estimée par rapport au X et Z réels.

On génère des observations par simulation à partir de réalisations aléatoire de positions $\overline{\mathbf{x}}_0$ et on vérifie que l’algorithme permet bien de retrouver la position d’origine avec une bonne fidélité. En l’absence de bruit (Figure 4.8), la fidélité est très bonne à la fois en latéral (x et y) et en profondeur (z), ce qui paraît donc supérieur aux techniques du chapitre précédent. Il est vrai que l’estimateur est ici basé sur un modèle rigoureusement exact, puisque le modèle de simulation et le modèle utilisé par l’estimateur sont les mêmes et qu’il n’y pas de bruit. Il s’agit donc d’un *inverse crime* qui n’a vocation qu’à valider l’algorithme.

Pour obtenir des données plus réalistes il convient au minimum d’ajouter du bruit électronique sur les données de simulation en introduisant une distribution gaussienne sur les données \overline{y}_1 . On choisit un bruit typique de l’application, à savoir 230 électrons RMS, ce qui fait une résolution énergétique équivalente à 2 % en largeur à mi-hauteur pour des photons incident de 122 keV. On remarque que le bruit détériore l’estimation de profondeur en élargissant la droite de manière homogène, tandis que pour la localisation en X ou Y ce sont les interactions éloignées de l’inter-pixel qui sont impactées de manière importante

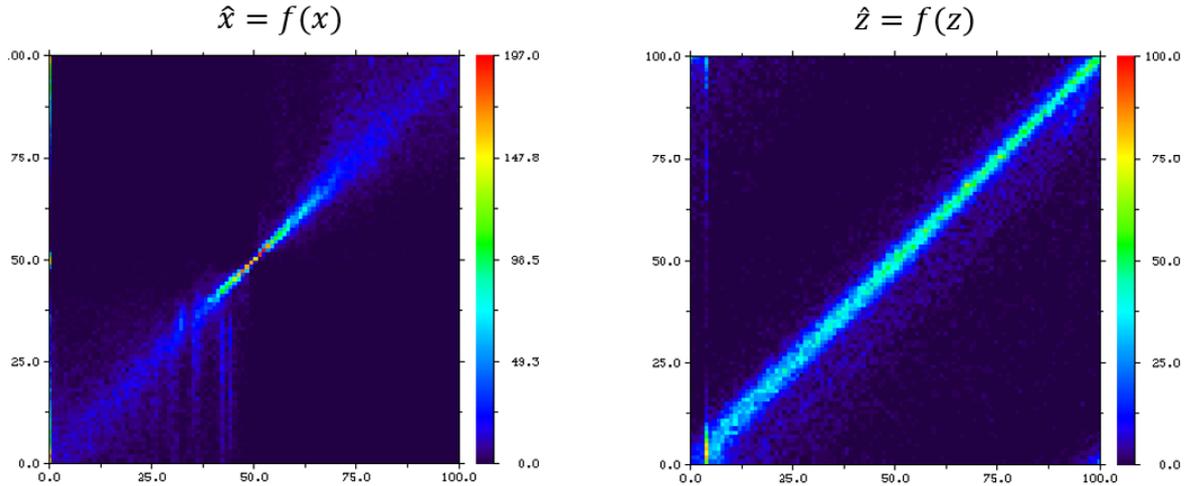


FIGURE 4.9 – Histogramme de la coordonnée X (à gauche) et Z (à droite) estimée par rapport au X et Z réels avec ajout d'un bruit de 122 keV (simulations).

(Figure 4.9). Ce comportement est bien connu puisque plus on s'éloigne de l'inter-pixel, moins le rapport signal sur bruit est favorable à la localisation. Ici on est également limité par le fait qu'on s'est limité à l'exploitation des signaux de quatre électrodes seulement : une électrode collectante ne dispose que de trois voisines donc on ne dispose que d'un voisinage partiel. La localisation en profondeur se comporte en revanche très bien par rapport aux méthodes présentées précédemment, ce qui provient probablement du fait que l'énergie incidente était supposée connue.

4.4.3 Exploitation sur détecteur réel

Pour tester l'algorithme d'apprentissage, on s'est basé sur les données expérimentales ESRF déjà utilisée au chapitre précédent. En effet, sur ces données, on dispose de la position réelle des photons incidents, ce qui nous permet de vérifier si notre estimation fonctionne correctement. Il s'agit donc de données obtenues sur un détecteur au pas de 2.5 mm, une épaisseur de 6 mm et pour une énergie incidente de l'ordre de 58 keV. De manière à limiter la dimension du vecteur \mathbf{y} , on se limite à un voisinage de 2×2 anodes.

Au démarrage du processus d'apprentissage, le champ est initialisé sans charges de volume, donc $\vec{\nabla}u = 0$ à l'intérieur du volume. C'est le cas sans défaut produisant la carte de CIE de la figure 4.10.

On observe que la zone de collecte incomplète est assez étroite au niveau des interpixels traduisant une très haute conductivité de surface et l'absence de charge piégées dans cette zone. En conséquence, les lignes de champ se courbent correctement vers les électrodes, ce qui limite la perte de charge,

La CIE obtenue par le scan expérimental du faisceau est représentée sur la figure 4.11. Le résultat est assez propre, le détecteur étant de bonne qualité mais on observe cependant deux choses : d'une part la perte de charge à l'inter-pixel est notablement plus marquée et d'autre part, les frontières entre pixel sont un peu plus fluctuantes. Pour le reste, la CIE est assez homogène sans gros défaut apparent.

L'algorithme d'apprentissage est testé sur les données $\overline{\mathbf{y}}_1$, sans connaissance de la position d'interaction réelle $\overline{\mathbf{x}}_0$ ni évidemment du champ électrique, Lors des itérations de

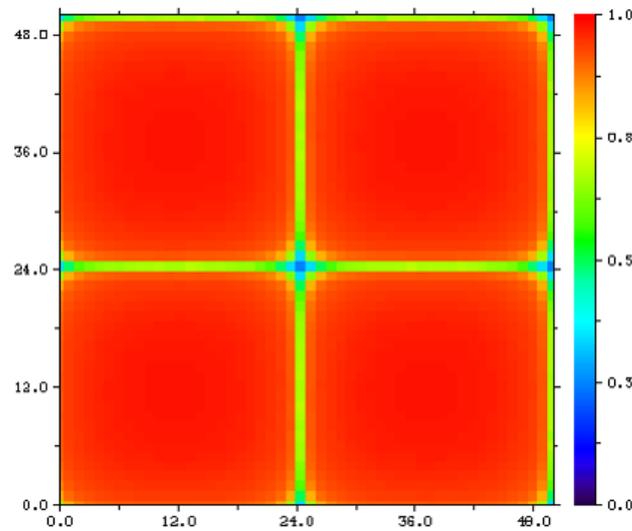


FIGURE 4.10 – Carte bidimensionnelle de CIE dans le voisinage de 2×2 anodes simulé sans défaut.

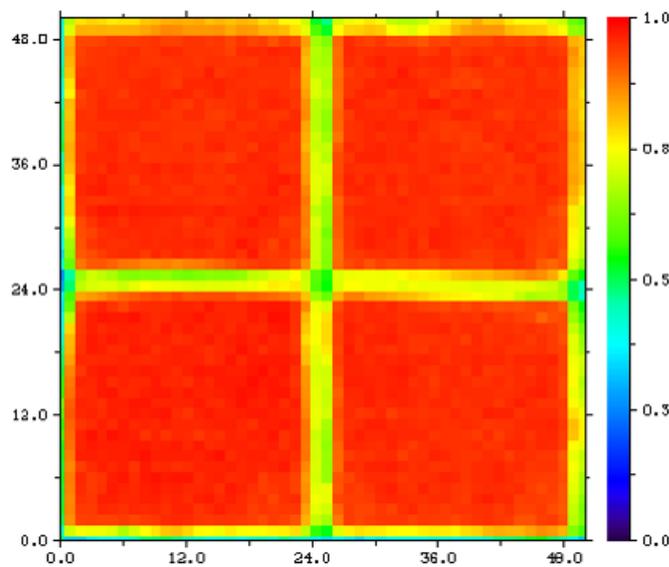


FIGURE 4.11 – CIE obtenue expérimentalement par le scan d'un détecteur CZT sur un voisinage de 2×2 anodes.

l'algorithme, on observe que la zone de perte de charge à l'inter-pixel tend à s'élargir (Figure 4.12) de manière à ajuster le modèle à l'observation. Au niveau de la simulation, cet effet est obtenu par l'ajout de charges électriques dans l'inter-pixel, qui tendent à attirer les lignes de champ et favorise le phénomène de perte de charge. Les légères fluctuations de la frontière entre pixel sont partiellement reproduites mais assez imparfaitement. Il est vrai que leur impact sur le signal observable est plutôt de second ordre. On montre tout de même que l'algorithme d'apprentissage par neural ODE permet d'apprendre une CIE mesurée expérimentalement sur une simulation. Ceci pourra être utilisé afin d'entraîner

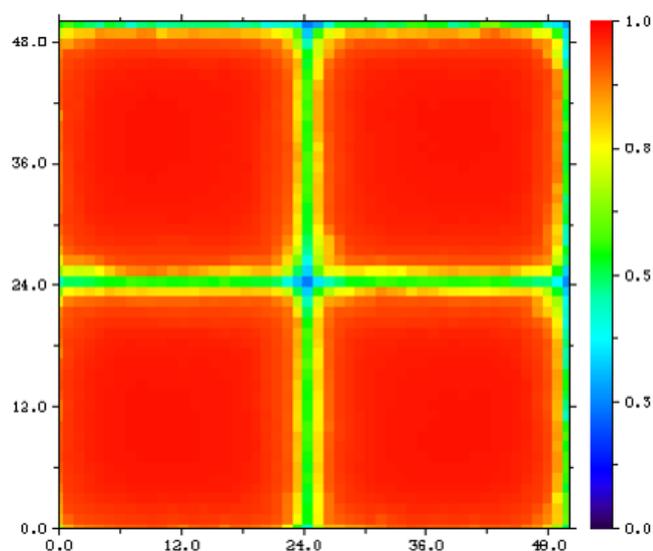


FIGURE 4.12 – CIE obtenue après apprentissage à partir des amplitudes observées expérimentalement sur un voisinage de 2×2 anodes.

le réseau de neurones développé dans la partie précédente pour améliorer la performance de localisation des interactions du détecteur.

4.5 Conclusion

La connaissance des paramètres internes d'un détecteur CZT permet une meilleure compréhension de son comportement physique et des réponses qui y sont liées. Plus particulièrement une cartographie du champ électrique interne permet à la fois de pouvoir reproduire en simulation le fonctionnement de ce détecteur propre et d'appliquer des corrections afin de prendre en compte et minimiser l'impact des inhomogénéités du détecteur. Dans cette partie on met en place une méthode d'apprentissage du champ électrique d'un détecteur quelconque simulé ou réel grâce à un modèle spécifique basé sur une étude d'utilisation d'un réseau de neurones appliqués sur une équation différentielle. Ce modèle permet de remplacer les couches discrètes de neurones par des couches continues où les poids des neurones correspondent aux paramètres que l'on souhaite optimiser. Dans notre cas cela correspond aux valeurs de champ électrique. Ainsi on développe un algorithme qui, pour une interaction dans le détecteur, permet de mettre à jour l'état initial du système et le vecteur des paramètres à partir des données mesurées de manière expérimentale. De cette manière on peut retranscrire le comportement du détecteur étudié. Ceci permet d'adapter notre simulation au détecteur utilisé pour faire les mesures expérimentales. Ainsi on peut entraîner le réseau de neurones pour la localisation développé dans la partie précédente grâce à la simulation qui est beaucoup plus rapide pour obtenir un jeu de données conséquent et optimiser le calcul des poids de ce réseau. Finalement après le réseau de localisation entraîné, on peut utiliser les données obtenues de manière expérimentale afin d'obtenir une localisation des interactions plus flexible et adaptable qu'avec une technique de localisation classique.

La problématique de la présence de défauts structuraux dans les cristaux de CZT utilisés pour la détection gamma est un enjeu important dans le domaine de l'imagerie. Ces défauts entraînent des inhomogénéités dans le comportement du détecteur, en altérant notamment ses paramètres physiques tels que la forme du champ électrique ou les performances de transport des porteurs de charge. Ces effets se répercutent au niveau des performances de localisation des interactions des photons gamma dans le détecteur ou diminuant la précision des mesures de manière importante. L'objectif de cette thèse était de proposer une alternative mathématique aux méthodes traditionnelles de correction de localisation, tout en permettant d'améliorer la compréhension de la structure des détecteurs CZT en simulant leur comportement de manière réaliste, en transposant dans les simulations les effets physiques observables de manière expérimentale pour différents types de défauts.

Le développement de cette méthodologie mathématique se décompose en trois parties se complétant et permettant d'obtenir un outil complet de calibration d'un détecteur CZT de manière rapide et efficace. La première brique concerne le développement de simulations en trois dimensions réalistes de détecteurs CZT. En effet pour pouvoir effectuer un apprentissage il faut être capable de modéliser le phénomène direct. Ces simulations permettent de modéliser et simuler le fonctionnement d'un détecteur CZT en ayant accès à chacun des paramètres le caractérisant, tel que le champ électrique, la distribution spatiale des charges ou encore la conductivité. On peut alors décider de modéliser de manière réaliste des défauts cristallins en modifiant ces paramètres de manière à approcher les comportements observés sur des données réelles. La puissance de calcul nécessaire étant croissante avec la complexité des simulations, on décide d'utiliser les capacités de l'unité graphique de l'ordinateur afin de paralléliser les opérations et optimiser les calculs. Nous avons réussi à simuler et observer le comportement de défauts de différentes natures que l'on peut retrouver en pratique dans un cristal de CZT réel.

La seconde partie est le développement d'un réseau de neurones simple permettant la localisation des interactions des photons gamma dans un détecteur CZT de manière non itérative. On construit un réseau simple, pouvant être implémentable sur FPGA dans l'électronique de mesure des détecteurs et dont le temps de calcul de localisation doit être similaire à l'algorithme analytique utilisé actuellement. Cet algorithme utilise le lien entre amplitudes de sortie par pixel et coordonnées estimées par simulation associées

comme jeu de données afin d'entraîner ce réseau. Après entraînement ce réseau permet d'estimer les coordonnées d'une interaction d'un photon gamma à l'aide des amplitudes des signaux de sortie mesurées pour le pixel collectant et ses voisins. On a pu montrer que les performances du réseau de neurones sont supérieures à celle de l'algorithme de localisation par maximum de vraisemblance utilisé aujourd'hui, en particulier pour traiter des données incomplètes ou atypiques car il est adaptable d'un pixel à l'autre tandis que l'algorithme de référence utilise le même modèle pour tous les pixels.

La troisième brique utilise une méthode d'apprentissage novatrice par équation différentielle, basée sur la méthode mathématique des équations adjointes. L'objectif est de propager une erreur calculée entre les signaux de sortie mesurés et simulés et de rétro-propager le gradient de cette erreur à travers l'équation différentielle décrivant la dynamique du système pour déterminer la mise à jour à appliquer sur l'état d'entrée estimé. Cette méthode permet surtout d'estimer les paramètres physiques d'un détecteur réel, notamment son champ électrique ou la carte de ses performances de collecte à l'aide des simulations développées précédemment. On a notamment pu apprendre les cartographies de CIE (Charge Induction Efficiency) à partir de données expérimentales et les appliquer en simulation. L'étude présentée ici se limitait à un voisinage restreint de deux par deux anodes. Il reste un travail à approfondir pour pouvoir apprendre de manière efficace le champ électrique en trois dimensions d'un détecteur réel d'après les seules mesures d'amplitudes à l'échelle d'une matrice de détection complète.

La combinaison de ces trois outils permet de développer une calibration rapide d'un détecteur CZT basée sur la simulation. On peut faire apprendre à une simulation les paramètres physiques d'un détecteur réel grâce à la méthode neural ODE. Puis cette simulation réaliste va servir à entraîner le réseau de neurones simple pour la localisation des interactions gamma dans le volume du détecteur, la quantité d'informations obtenue en un temps faible permettant d'optimiser le calcul des pondérations du réseau de neurones. Finalement, l'utilisation des données expérimentales d'amplitudes mesurées en entrée de ce réseau de neurones entraîné permet d'estimer de manière rapide et efficace la localisation des interactions dans le détecteur.

Des ajustements restent à faire et à tester, notamment l'implémentation dans le module électronique de traitement du détecteur du laboratoire, afin de pouvoir comparer en conditions réelles les capacités de la méthode de calibration. Également l'observation du champ électrique en trois dimensions d'un détecteur réel après apprentissage par simulation pourrait être mis en place afin de pouvoir observer la structure du cristal CZT utilisé. Ce type de calibration pourrait être par exemple réalisé en amont de l'utilisation d'un cristal afin de déterminer son niveau de performances, que ce soit au niveau de la production du côté industriel ou au niveau de l'exploitation et de la construction de détecteurs.

Au niveau plus global, l'amélioration des performances des détecteurs CZT pour l'imagerie gamma grâce à ces méthodes d'apprentissage par intelligence artificielle peuvent apporter un certain nombre d'avantages pour différentes raisons. Il pourra déjà être possible de réaliser des détecteurs avec des volumes plus importants, intéressant notamment dans les domaines du médical et de l'aérospatial. En effet le fait que les effets des défauts cristallins soient corrigés par l'algorithme, ce qui va minimiser leur impact, peut aboutir à la construction ou à l'assemblage de surfaces de détection plus importantes. D'autre part l'augmentation de la précision des localisations des interactions dans le détecteur permet en amont d'obtenir une meilleure localisation de la source gamma. Cette efficacité

permettrait dans des cas applicatifs de diminuer l'injection de la quantité de radiotraceurs dans l'organisme d'un patient en imagerie médicale, ou de diminuer le temps d'exposition des opérateurs en centrale nucléaire.

Bibliographie

- [1] D. Blanc, *Les rayonnements ionisants*, p. 23. Ed.Masson, 1990.
- [2] W. Leo, *Techniques for Nuclear and Particle Physics Experiments*, p. 5. Springer Berlin, 1994. <https://link.springer.com/book/10.1007/978-3-642-57920-2>.
- [3] S. Nicol, *Etude et construction d'un tomographe TED/TDM pour petits animaux, combinant modules phoswich à scintillateurs et détecteur à pixels hybrides*. PhD thesis, Université de la Méditerranée Aix-Marseille 2, July 2010. http://imxgam.in2p3.fr/doc/SNicol_these.pdf.
- [4] S. Nicol, *Etude et construction dun tomographe TED/TDM pour petits animaux, combinant modules phoswich à scintillateurs et détecteur à pixels hybrides*. PhD thesis, Université de la Méditerranée Aix-Marseille 2, July 2010. http://imxgam.in2p3.fr/doc/SNicol_these.pdf.
- [5] A. Compton, “A Quantum Theory of the Scattering of X rays by Light Elements,” *Physical Review*, vol. 21, no. 5, pp. 483–502, 1923. <https://link.aps.org/doi/10.1103/PhysRev.21.483>.
- [6] C. Foellmi, “Schéma montrant la collision d'un photon avec un électron, pour illustrer la diffusion compton.” <https://commons.wikimedia.org/wiki/File:DiffusionCompton.png>.
- [7] S. Nicol, *Etude et construction dun tomographe TED/TDM pour petits animaux, combinant modules phoswich à scintillateurs et détecteur à pixels hybrides*. PhD thesis, Université de la Méditerranée Aix-Marseille 2, July 2010. http://imxgam.in2p3.fr/doc/SNicol_these.pdf.
- [8] H. Negm, N. K. Ahmed, A. Abbady, and M. M. Reda, “Study of Radionuclides and Radon Exhalation Rate in Soil and Sand Samples from Tiba, Luxor, Governorate,” *World Journal of Nuclear Science and Technology*, vol. 9, no. 2, pp. 84–95, 2019. <http://dx.doi.org/10.4236/wjnst.2019.92006>.
- [9] G. Gilmore, *Interactions of Gamma Radiation with Matter*, pp. 25–38. John Wiley and Sons, 2008. <https://doi.org/10.1002/9780470861981>.
- [10] C. Dujardin and M. Hamel, *Overview on Plastic and Inorganic Scintillators*, pp. 3–33. Springer, 2021. https://doi.org/10.1007/978-3-030-73488-6_1.

- [11] A. Douiri, *Développement à partir du concept CACAO (Caméra A Collimation Assistée par Ordinateur) d'un système tomographique adapté à la mesure 3D de plaies contaminées*. PhD thesis, Université Paris XI, March 2002. <https://www.theses.fr/2002PA112043>.
- [12] H. Anger, "Scintillation Camera with Multichannel Collimators," *Journal of Nuclear Medicine*, vol. 5, pp. 515–531, 1964. <https://pubmed.ncbi.nlm.nih.gov/14216630>.
- [13] P.-C. Gervais, *Études expérimentale et numérique du colmatage de filtre plissé*. PhD thesis, Université de Lorraine, February 2013. <https://www.theses.fr/2013LORR0007>.
- [14] ONSEMI, "Readout Methods for Arrays of Silicon Photomultipliers." <https://www.onsemi.com/pub/Collateral/AND9778-D.PDF>.
- [15] H. Seo and al, "Development of Double-scattering-type Compton Camera with Double-sided Silicon Strip Detectors and NaI(Tl) Scintillation Detector," *Nuclear Instruments and Methods in Physics Research Section A : Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 615, no. 3, pp. 333–339, 2010. <https://doi.org/10.1016/j.nima.2010.02.060>.
- [16] J. Roser and al, "Image Reconstruction for a Multi-layer Compton Telescope : an Analytical Model for three Interaction Events," *Physics in Medicine and Biology*, vol. 65, no. 14, 2020. <https://doi.org/10.1088/1361-6560/ab8cd4>.
- [17] P. Tandon, P. Huggins, R. Maclachlan, A. Dubrawski, K. Nelson, and S. Labov, "Detection of Radioactive Sources in Urban Scenes using Bayesian Aggregation of Data from Mobile Spectrometers," *Information Systems*, vol. 57, pp. 195–206, 2016. <https://doi.org/10.1371/journal.pone.0268556>.
- [18] R. Lecomte, "Novel detector technology for clinical pet," *European Journal of Nuclear Medicine and Molecular Imaging*, vol. 36, pp. 69–85, March 2009. <https://doi.org/10.1007/s00259-008-1054-0>.
- [19] F. Brooks, "Development of Organic Scintillators," *Nuclear Instruments and Methods*, vol. 162, no. 1, pp. 477–505, 1979. [https://doi.org/10.1016/0029-554X\(79\)90729-8](https://doi.org/10.1016/0029-554X(79)90729-8).
- [20] R. Heath, R. Hofstadter, and E. Hughes, "Inorganic Scintillators : A Review of Techniques and Applications," *Nuclear Instruments and Methods*, vol. 162, no. 1, pp. 431–476, 1979. [https://doi.org/10.1016/0029-554X\(79\)90728-6](https://doi.org/10.1016/0029-554X(79)90728-6).
- [21] M. Titov, "Gaseous Detectors : Recent Developments and Applications," 2010. <https://doi.org/10.48550/arXiv.1008.3736>.
- [22] J. Peyré, "Cours détecteurs gazeux," 2019.
- [23] E. Gros d'Aillon, *Etude des performances spectrométriques des détecteurs gamma CdZnTe/CdTe monolithiques*. PhD thesis, Université Joseph Fourier Grenoble 1, November 2005. <https://theses.hal.science/tel-00011119>.
- [24] R. J. Hicks and M. S. Hofman, "Is There Still a Role for SPECT–CT in Oncology in the PET–CT Era?," *Nature Reviews Clinical Oncology*, vol. 9, no. 1, pp. 712–720, 2012. <https://doi.org/10.1038/nrclinonc.2012.188>.
- [25] E. Sakai, "Present Status of Room Temperature Semiconductor Detectors," *Nuclear Instruments and Methods in Physics Research*, vol. 196, no. 1, pp. 121–130, 1982. [https://doi.org/10.1016/0029-554X\(82\)90626-7](https://doi.org/10.1016/0029-554X(82)90626-7).

- [26] M. Picone, *Contribution à la simulation tridimensionnelle de détecteurs semiconducteurs en spectrométrie gamma*. PhD thesis, Université Joseph Fourier, November 2002. https://inis.iaea.org/collection/NCLCollectionStore/_Public/43/035/43035118.pdf.
- [27] E.-J. Cassignol, *Physique et électronique des semi conducteurs*. Bibliothèque Technique Philips, 1965.
- [28] H. Bian, L. Yang, J. Zhang, and Z. Ma, “The Role of Trapping-De-trapping Space Charge in Electrothermal Aging of XLPE,” *IEEE Transactions on Dielectrics and Electrical Insulation*, vol. 27, no. 6, pp. 2039–2047, 2020. <https://doi.org/10.1109/TDEI.2020.008866>.
- [29] M. Rosaz, *Etude et réalisation de nouvelles structures de détection à base de CdTe et CdZnTe pour l'imagerie X et γ* . PhD thesis, Institut National Polytechnique de Grenoble, October 1997.
- [30] S. Ramo, “Currents Induced by Electron Motion,” *Proceedings of the IRE*, vol. 27, no. 9, pp. 584–585, 1939. <https://doi.org/10.1109/JRPROC.1939.228757>.
- [31] Q. Zheng, *Preparation and Development of CdTe and CdZnTe detectors for Gamma Ray Radiations Applications*. PhD thesis, Universidad autonoma de Madrid, October 2012.
- [32] P. S. Dutta, *Bulk Crystal Growth of Ternary III-IV Semiconductors*, pp. 281–325. Springer, 2010. https://link.springer.com/chapter/10.1007/978-3-540-74761-1_10.
- [33] R. Triboulet, “The Traveling Heating Method (THM) for $Hg_{1-x}Cd_xTe$ and Related Materials,” *Progress in Crystal Growth and Characterization of Materials*, vol. 28, no. 1, pp. 85–144, 1994. [https://doi.org/10.1016/0960-8974\(94\)90014-0](https://doi.org/10.1016/0960-8974(94)90014-0).
- [34] B. Zhou and al, “Modification of Growth Interface of CdZnTe Crystals in THM Process by ACRT,” *Journal of Crystal Growth*, vol. 483, pp. 281–284, 2018. <https://doi.org/10.1016/j.jcrysgro.2017.12.003>.
- [35] X. Cui and al, “Characteristics of the Dislocations in CdZnTe Crystals Revealed by Etch Pits,” *Journal of Crystal Growth*, vol. 321, no. 1, pp. 40–44, 2011. <https://doi.org/10.1016/j.jcrysgro.2011.02.009>.
- [36] C. K. Egan, A. Choubey, M. Moore, and R. J. Cernik, “Characterisation of Vapour Grown CdZnTe Crystals using Synchrotron X-ray Topography,” *Journal of Crystal Growth*, vol. 343, no. 1, pp. 1–6, 2012. <http://dx.doi.org/10.1016/j.jcrysgro.2012.01.003>.
- [37] E. Vitonne and al, “Semiconductor Characterization by Scanning Ion Beam Induced Charge (IBIC) Microscopy,” *Nuclear Instruments and Methods in Physics Research Section B : Beam Interactions with Materials and Atoms*, vol. 266, no. 8, pp. 1312–1318, 2008. <https://doi.org/10.1016/j.nimb.2007.12.083>.
- [38] E. Darque-Ceretti, M. Aucouturier, and P. Lehuédé, “Spectrométrie de masse d’ions secondaires : Sims et tof-sims - principes et appareillages,” *Techniques de l’ingénieur Techniques d’analyse*, vol. base documentaire : TIP630WEB, 2014. <https://doi.org/10.51257/a-v4-p2618>.
- [39] G. Yang, “Impurity Gettering Effect of Te Inclusions in CdZnTe Single Crystals.” <https://doi.org/10.1016/j.jcrysgro.2008.09.201>.

- [40] A. Kokil, K. Yang, and J. Kumar, "Techniques for Characterization of Charge Carrier Mobility in Organic Semiconductors," *Journal of Polymer Science Part B : Polymer Physics*, vol. 50, pp. 1130–1144, 2012. <https://doi.org/10.1002/polb.23103>.
- [41] M. Hage-Ali and all, "Microscopic Defect Level Characterization of Semi-Insulating Compound Semiconductors by TSC and PICTS : Application to the Effect of Hydrogen in cdte," *Applied Surface Science*, vol. 50, no. 1, pp. 377–382, 1991. [https://doi.org/10.1016/0169-4332\(91\)90202-U](https://doi.org/10.1016/0169-4332(91)90202-U).
- [42] F. Scholtz, J. Hwang, and D. Schroder, "Low frequency noise and DLTS as Semiconductor Device Characterization tools," *Solid-State Electronics*, vol. 31, no. 2, pp. 205–217, 1988. [https://doi.org/10.1016/0038-1101\(88\)90129-3](https://doi.org/10.1016/0038-1101(88)90129-3).
- [43] F. Arevalo, "Optical Characterization of Semiconductor Materials by using FTIR-PAS," *Journal of Physics : Conference Series*, vol. 134, no. 1, 2008. <https://dx.doi.org/10.1088/1742-6596/134/1/012019>.
- [44] C. Buis, *Etude des corrélations entre les défauts structuraux et les inhomogénéités spatiales des détecteurs de rayons X à base de CdTe pour l'imagerie médicale*. PhD thesis, Université Jean Monnet St Etienne, August 2014. <https://theses.hal.science/tel-01056544>.
- [45] P. Rudolph, A. Engel, I. Schentke, and A. Grochocki, "Distribution and Genesis of Inclusions in CdTe and (Cd,Zn)Te Single Crystals Grown by the Bridgman Method and by the Travelling Heater Method," *Journal of Crystal Growth*, vol. 147, no. 3, pp. 297–304, 1995. [https://doi.org/10.1016/0022-0248\(94\)00617-2](https://doi.org/10.1016/0022-0248(94)00617-2).
- [46] C. Buis, *Etude des corrélations entre les défauts structuraux et les inhomogénéités spatiales des détecteurs de rayons X à base de CdTe pour l'imagerie médicale*. PhD thesis, Université Jean Monnet St Etienne, August 2014. <https://theses.hal.science/tel-01056544>.
- [47] A. Bolotnikov and al, "Cumulative Effects of Te Precipitates in CdZnTe Radiation Detectors," *Nuclear Instruments and Methods in Physics Research Section A : Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 571, no. 3, pp. 687–698, 2007. <https://doi.org/10.1016/j.nima.2006.11.023>.
- [48] A. Hossain and al, "Extended Defects in CdZnTe Crystals : Effects on Device Performance," 2010. <https://doi.org/10.1016/j.jcrysgro.2010.03.005>.
- [49] R. List, "Electrical Effects of Dislocations and other Crystallographic Defects in Hg_{0.78}Cd_{0.22}Te n-on-p Photodiodes," *Journal of Electronic Materials*, vol. 22, no. 1, pp. 1017–1025, 1993. <https://doi.org/10.1007/BF02817519>.
- [50] T. Figielski, "Dislocations as Electrically Active Centers in Semiconductors," *Journal of Physics : Condensed Matter*, vol. 14, no. 48, 2022. <http://dx.doi.org/10.1088/0953-8984/14/48/301>.
- [51] A. E. Bolotnikov and al, "Correlation Between Crystals Defects and Performance of CdZnTe Detectors," *IEEE Transactions on Nuclear Science*, vol. 58, no. 4, pp. 1972–1980, 2011. <https://doi.org/10.1109/TNS.2011.2160283>.
- [52] M. Prokesch, D. S. Bale, and C. Szeles, "Fast High-Flux Response of CdZnTe X-Ray Detectors by Optical Manipulation of Deep Level Defect Occupations," *IEEE Transactions on Nuclear Science*, vol. 57, no. 4, pp. 2397–2399, 2010. <https://doi.org/10.1109/TNS.2010.2049120>.

- [53] Y. Li and al, "Effect of Deep-Level Defects on the Performance of CdZnTe Photon Counting Detectors," *Sensors*, vol. 20, no. 7, 2020. <http://dx.doi.org/10.3390/s20072032>.
- [54] V. M. Gerrish, "Chapter 13 - Characterization and Quantification of Detector Performance," in *Semiconductors and Semimetals*, vol. 43 of *Semiconductors and Semimetals*, pp. 493–530, Elsevier, 1995. [https://doi.org/10.1016/S0080-8784\(08\)62752-0](https://doi.org/10.1016/S0080-8784(08)62752-0).
- [55] K. Hecht, "Zum Mechanismus Des Lichtelektrischen Primärstromes In Isolierenden Kristallen," *Zeitschrift für Physik*, vol. 77, pp. 235–245, 1932. <https://doi.org/10.1007/BF01338917>.
- [56] H. Gajewski and K. Gröger, "On the Basic Equations for Carrier Transport in Semiconductors," *IEEE Transactions on Electron Devices*, vol. 113, no. 1, pp. 12–35, 1986. [https://doi.org/10.1016/0022-247X\(86\)90330-6](https://doi.org/10.1016/0022-247X(86)90330-6).
- [57] M. Benoit and L. Hamel, "Simulation of Charge Collection Processes in Semiconductor CdZnTe γ -ray Detectors," *Nuclear Instruments and Methods in Physics Research Section A : Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 606, no. 3, pp. 508–516, 2009. <https://doi.org/10.1016/j.nima.2009.04.019>.
- [58] R. Bank, D. Rose, and W. Fichtner, "Numerical Methods for Semiconductor Device Simulation," *IEEE Transactions on Electron Devices*, vol. 30, no. 9, pp. 1031–1041, 1983. <https://doi.org/10.1109/T-ED.1983.21257>.
- [59] J. Rogel-Salazar, D. D. C. Bradley, J. R. Cash, and J. C. deMello, "An Efficient Method-of-lines Simulation Procedure for Organic Semiconductor Devices," *Physical Chemistry Chemical Physics*, vol. 11, pp. 1636–1646, 2009. <http://dx.doi.org/10.1039/B813810A>.
- [60] N. Metropolis and S. Ulam, "The Monte Carlo Method," *Journal of the American Statistical Association*, vol. 44, no. 247, pp. 335–341, 1949. <https://dx.doi.org/10.2307/2280232>.
- [61] G. U. Jensen, B. Lund, T. A. Fjeldly, and M. Shur, "Monte Carlo Simulation of Semiconductor Devices," *Computer Physics Communications*, vol. 67, no. 1, pp. 1–61, 1991. [https://doi.org/10.1016/0010-4655\(91\)90220-F](https://doi.org/10.1016/0010-4655(91)90220-F).
- [62] R. James and al, "Material Properties of Large-volume Cadmium Zinc Telluride Crystals and their Relationship to Nuclear Detector Performance," *Journal of Electronic Materials*, vol. 27, pp. 788–799, 1998. <https://doi.org/10.1007/s11664-998-0055-x>.
- [63] L.-A. Hamel and S. Paquet, "Charge Transport and Signal Generation in CdTe Pixel Detectors," *Nuclear Instruments and Methods in Physics Research Section A : Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 380, no. 1, pp. 238–240, 1996. [https://doi.org/10.1016/S0168-9002\(96\)00354-3](https://doi.org/10.1016/S0168-9002(96)00354-3).
- [64] K. Iniewski and all, "Modeling Charge-sharing Effects in Pixellated CZT Detectors," in *2007 IEEE Nuclear Science Symposium Conference Record*, vol. 6, pp. 4608–4611, 2007. <http://dx.doi.org/10.1109/NSSMIC.2007.4437135>.

- [65] T. Prettyman and al, “Physics-based Generation of Gamma-ray Response Functions for CdZnTe Detectors,” *Journal of Radioanalytical and Nuclear Chemistry*, vol. 233, pp. 257–264, 1998. <https://doi.org/10.1007/BF02389681>.
- [66] P. Guerra, A. Santos, and D. Darambara, “Development of a Simplified Simulation Model for Performance Characterization of a Pixellated CdZnTe Multimodality Imaging System,” *Physics in Medicine and Biology*, vol. 53, no. 4, 2008. <https://doi.org/10.1088/0031-9155/53/4/019>.
- [67] J. R. Nagel, “Solving the Generalized Poisson Equation Using the Finite-Difference Method (FDM),” 2011. https://www.researchgate.net/publication/228411289_Solving_the_Generalized_Poisson_Equation_Using_the_Finite-Difference_Method_FDM.
- [68] V. Jagota, “Finite Element Method : An Overview,” *Walailak Journal of Science and Technology*, vol. 10, no. 1, 2013. <https://wjst.wu.ac.th/index.php/wjst/article/view/499>.
- [69] C. Geuzaine and J.-F. Remacle, “GMSH : A Three-dimensional Finite Element Mesh Generator with built-in pre- and post-processing Facilities,” *International Journal for Numerical Methods in Engineering*, vol. 79, no. 11, pp. 1309–1331, 2009. <https://doi.org/10.1002/nme.2579>.
- [70] V. Ardon, *Méthodes numériques et outils logiciels pour la prise en compte des effets capacitifs dans la modélisation CEM de dispositifs d’électronique de puissance*. PhD thesis, Université de Grenoble, June 2010. <https://theses.hal.science/tel-00502938>.
- [71] J. Barnes and R. Lomax, “Finite-element Methods in Semiconductor Device Simulation,” *IEEE Transactions on Electron Devices*, vol. 24, no. 8, pp. 1082–1089, 1977. <https://doi.org/10.1109/T-ED.1977.18880>.
- [72] S. Kavadias, K. Misiakos, and D. Loukas, “Calculation of Pixel Detector Capacitances through Three Dimensional Numerical Solution of the Laplace Equation,” *IEEE Transactions on Nuclear Science*, vol. 41, no. 2, pp. 397–401, 1994. <https://doi.org/10.1109/23.281533>.
- [73] M. Pavesi and al, “Electric Field Reconstruction and Transport Parameter Evaluation in CZT X-Ray Detectors,” *IEEE Transactions on Nuclear Science*, vol. 64, no. 10, pp. 2706–2712, 2017. <https://doi.org/10.1109/TNS.2017.2744327>.
- [74] O. Alirol, *Simulation d’un détecteur CdTe : CI de rayonnement X en mode comptage à fort flux pour la radiothérapie et la tomographie*. PhD thesis, Institut National des Sciences Appliquées de Lyon, September 2010. <https://www.theses.fr/2010ISAL0070>.
- [75] S. Lux, *Amélioration de la résolution spatiale de détecteurs à semi-conducteur, pour l’imagerie gamma et X, par l’exploitation de signaux transitoires*. PhD thesis, Université de Grenoble, 2014.
- [76] G. Montémont and al, “Simulation and Design of Orthogonal Capacitive Strip CdZnTe Detectors,” *IEEE Transactions on Nuclear Science*, vol. 54, no. 4, pp. 854–859, 2007. <https://doi.org/10.1109/TNS.2007.902356>.
- [77] G. F. Knoll, *Pulse Processing and Shaping*, pp. 577–603. Wiley, third ed., 2000.

- [78] F. Mathy and al, “ATthree-dimensional Model of CdZnTe Gamma-ray Detector and its Experimental Validation,” *IEEE Transactions on Nuclear Science*, vol. 51, no. 5, pp. 2419–2426, 2004. <https://doi.org/10.1109/TNS.2004.835906>.
- [79] G. Montémont and al, “Studying Spatial Resolution of CZT Detectors using Sub-pixel Positioning for SPECT,” *IEEE Transactions on Nuclear Science*, vol. 61, no. 5, pp. 2559–2566, 2014. <https://doi.org/10.1109/TNS.2014.2356293>.
- [80] M. Dimmock and al, “An OpenCL Implementation of Pinhole Image Reconstruction,” *IEEE Transactions on Nuclear Science*, vol. 59, no. 4, pp. 1738–1749, 2012. <https://doi.org/10.1109/TNS.2012.2197760>.
- [81] B. Gaster and al, *Heterogeneous Computing with OpenCL*. Morgan Kaufman, 2011. <https://doi.org/10.1016/C2011-0-69669-3>.
- [82] K. Yacoubi, *Validation expérimentale d’un algorithme de reconstruction pour l’imagerie médicale par tomo-scintigraphie*. Master thesis, ENSSAT, 2017.
- [83] W. Warburton, “An Approach to Sub-pixel Spatial Resolution in Room Temperature X-ray Detector Arrays with Good Energy Resolution,” *MRS Online Proceedings Library*, vol. 487, pp. 531–535, 1997. <https://doi.org/10.1557/PROC-487-531>.
- [84] H. Barrett and al, “Maximum-Likelihood Methods for Processing Signals from Gamma-ray Detectors,” *IEEE Transactions on Nuclear Science*, vol. 56, no. 3, pp. 725–735, 2009. <http://dx.doi.org/10.1109/TNS.2009.2015308>.
- [85] D. Barbes, *Evaluation expérimentale d’un module d’imagerie spectrométrique gamma et mise en place de méthode de localisation sous-pixel*. Master thesis, Grenoble INP Phelma, 2013.
- [86] F. Durand, *Méthodes pour l’imagerie spectrométrique à base de détecteurs CdZnTe*. Master thesis, Grenoble INP ENSE3/SICOM, 2014.
- [87] S. Banerjee and al, “A Learning-based Physical Model of Charge Transport in Room-temperature Semiconductor Detectors,” *IEEE Transactions on Nuclear Science*, vol. 69, no. 1, 2022. <https://doi.org/10.1109/TNS.2021.3130486>.
- [88] C. Yang, “Joint Estimation of Interaction Position and Energy Deposition in Semiconductor SPECT Imaging Sensors using Fully Connected Neural Network,” *Physics in Medicine and Biology*, vol. 68, no. 9, 2023. <http://dx.doi.org/10.1088/1361-6560/aca740>.
- [89] W. McCulloch and W. Pitts, “A Logical Calculus of the Ideas Immanent in Nervous Activity,” *The bulletin of mathematical biophysics*, vol. 5, pp. 115–133, 1943. <https://doi.org/10.1007/BF02478259>.
- [90] S. E. Kim and I. W. Seo, “Artificial Neural Network Ensemble Modeling with Conjunctive Data Clustering for Water Quality Prediction in Rivers,” *Journal of Hydro-environment Research*, vol. 9, no. 3, pp. 325–339, 2015. <https://doi.org/10.1016/j.jher.2014.09.006>.
- [91] D. Berrar, “Cross-validation,” in *Encyclopedia of Bioinformatics and Computational Biology*, pp. 542–545, Academic Press, 2019. <http://dx.doi.org/10.1016/B978-0-12-809633-8.20349-X>.
- [92] S. Balasundaram, “Robust Twin Support Vector Regression Based on Huber Loss Function,” *Neural Computing and Applications*, vol. 32, p. 11285–11309, 2020. <https://doi.org/10.1007/s00521-019-04625-8>.

- [93] S. Ruder, “An Overview of Gradient Descent Optimization Algorithms,” *Computing Research Repository*, 2016. <http://arxiv.org/abs/1609.04747>.
- [94] W. Press, S. Teukolsky, W. Vetterling, and B. Flannery, *Numerical Recipes in C, The Art of Scientific Computing*. Bibliothèque Technique Philips, second ed., 1992.
- [95] J. Butcher, “Numerical Methods for Ordinary Differential Equations in the 20th Century,” *Journal of Computational and Applied Mathematics*, vol. 125, no. 1, pp. 1–29, 2000. [https://doi.org/10.1016/S0377-0427\(00\)00455-6](https://doi.org/10.1016/S0377-0427(00)00455-6).
- [96] R. T. Q. Chen, Y. Rubanova, J. Bettencourt, and D. Duvenaud, “Neural Ordinary Differential Equations,” in *Proc. of the 32nd conference on Neural Information Processing Systems*, 2019. <https://doi.org/10.48550/arXiv.1806.07366>.
- [97] Y. LeCun, L. Bottou, G. Orr, and K.-R. Müller, *Efficient BackProp*. Springer, second ed., 2022.

Intégration analytique en une dimension

A.1 Calcul direct des gradients

Pour estimer Θ il va d'abord falloir estimer le $x_0 = (Q_0, z_0)$ de chaque événement mesuré \mathbf{y}_1 puis estimer globalement Θ . On s'intéresse donc au vecteur des dérivées partielles

$$J = \frac{\partial \mathbf{y}_1}{\partial (\mathbf{x}_0, \Theta)} = \begin{pmatrix} \frac{\partial \mathbf{y}_1}{\partial Q_0} \\ \frac{\partial \mathbf{y}_1}{\partial z_0} \\ \frac{\partial \mathbf{y}_1}{\partial \Theta} \end{pmatrix} \quad (\text{A.1})$$

On part de l'équation d'évolution de \mathbf{y} (par rapport à z)

$$\dot{\mathbf{y}} = Q\mathbf{w} \Rightarrow \mathbf{y}_1 = \int_{z_0}^{z_1} Q\mathbf{w}dz \text{ sachant que } \mathbf{y}_0 = 0 \quad (\text{A.2})$$

Le champ de pondération est fixe et connu donc l'évolution des gradients ne dépend que de Q :

$$\frac{\partial \dot{\mathbf{y}}}{\partial (\mathbf{x}_0, \Theta)} = \frac{\partial Q}{\partial (\mathbf{x}_0, \Theta)} \mathbf{w} \Rightarrow \frac{\partial \mathbf{y}_1}{\partial (\mathbf{x}_0, \Theta)} = \int_{z_0}^{z_1} \frac{\partial Q}{\partial (\mathbf{x}_0, \Theta)} \mathbf{w} dz \quad (\text{A.3})$$

En séparant variable par variable on obtient

$$\frac{\partial \dot{\mathbf{y}}}{\partial Q_0} = \frac{\partial Q}{\partial Q_0} \mathbf{w} \quad (\text{A.4})$$

$$\frac{\partial \dot{\mathbf{y}}}{\partial z_0} = \frac{\partial Q}{\partial z_0} \mathbf{w} \quad (\text{A.5})$$

$$\frac{\partial \dot{\mathbf{y}}}{\partial \Theta} = \frac{\partial Q}{\partial \Theta} \mathbf{w} \quad (\text{A.6})$$

Or l'équation d'évolution de la charge est

$$\dot{Q} = -\frac{Q(z)}{u(z)\tau} \quad (\text{A.7})$$

Les deux premières dérivées partielles ne dépendent que de l'état initial mais la troisième est dépendante de l'état du système à l'instant considéré car l'impact de Θ sur Q en un z donné se propage sur la suite de la trajectoire.

$$\frac{\partial Q}{\partial Q_0} = \frac{Q}{Q_0} \text{ car } Q = Q_0 e^{-t/\tau} \quad (\text{A.8})$$

$$\frac{\partial Q}{\partial z_0} = \frac{Q}{u_0 \tau} \quad (\text{A.9})$$

$$\frac{\partial \dot{Q}}{\partial \Theta} = -\frac{\partial}{\partial \Theta} \left(\frac{Q}{u \tau} \right) = -\frac{1}{\tau} \frac{\partial}{\partial \Theta} \left(\frac{Q}{u} \right) \quad (\text{A.10})$$

$$= \frac{1}{u \tau} \cdot \left(\frac{Q}{u} \frac{\partial u}{\partial \Theta} - \frac{\partial Q}{\partial \Theta} \right) \quad (\text{A.11})$$

Lors de l'interaction du photon dans le détecteur, Les valeurs initiales en $z = z_0$ sont :

$$\frac{\partial Q}{\partial \Theta}(z_0) = 0 \quad (\text{A.12})$$

$$\frac{\partial \mathbf{y}}{\partial Q_0}(z_0) = 0 \quad (\text{A.13})$$

$$\frac{\partial \mathbf{y}}{\partial z_0}(z_0) = -Q_0 \mathbf{w}_0 \quad (\text{A.14})$$

$$\frac{\partial \mathbf{y}}{\partial \Theta}(z_0) = 0 \quad (\text{A.15})$$

On peut simplifier le calcul de la dépendance à l'état initial :

$$\dot{Q}_0^* = \frac{\partial \dot{\mathbf{y}}}{\partial Q_0} = \frac{Q}{Q_0} \mathbf{w} \Rightarrow Q_0^* = \frac{1}{Q_0} \int Q \mathbf{w} dz = \frac{\mathbf{y}_1}{Q_0} \quad (\text{A.16})$$

$$\dot{z}_0^* = \frac{\partial \dot{\mathbf{y}}}{\partial z_0} = \frac{Q}{u_0 \tau} \mathbf{w} \Rightarrow z_0^* = \frac{1}{u_0 \tau} \int Q \mathbf{w} dz = \frac{\mathbf{y}_1}{u_0 \tau} - Q_0 \mathbf{w}_0 \quad (\text{A.17})$$

La dépendance aux paramètres Θ est chaînée

$$\frac{\partial \dot{Q}}{\partial \Theta} = \frac{1}{u \tau} \cdot \left(\frac{Q}{u} \frac{\partial u}{\partial \Theta} - \frac{\partial Q}{\partial \Theta} \right) \quad (\text{A.18})$$

$$\frac{\partial \dot{\mathbf{y}}}{\partial \Theta} = \frac{\partial Q}{\partial \Theta} \mathbf{w} \quad (\text{A.19})$$

On peut faire une intégration analytique en exploitant le changement de variable $z \Leftrightarrow t$,

$dz = u dt :$

$$\frac{d}{dt} \frac{\partial \dot{Q}}{\partial \Theta} = \frac{1}{\tau} \cdot \left(\frac{Q}{u} \frac{\partial u}{\partial \Theta} - \frac{\partial Q}{\partial \Theta} \right) \quad (\text{A.20})$$

$$\frac{\partial Q}{\partial \Theta} = e^{-t/\tau} \int_{t_0}^t Q e^{t'/\tau} \frac{1}{u\tau} \frac{\partial u}{\partial \Theta} dt \quad (\text{A.21})$$

$$\frac{\partial Q}{\partial \Theta} = Q_0 e^{-t/\tau} \int_{z_0}^z \frac{1}{u(z')^{2\tau}} \frac{\partial u(z')}{\partial \Theta} dz' \quad (\text{A.22})$$

$$\frac{\partial Q}{\partial \Theta} = Q \int_{z_0}^z \frac{1}{u(z')^{2\tau}} \frac{\partial u(z')}{\partial \Theta} dz' \quad (\text{A.23})$$

$$\frac{\partial \mathbf{y}}{\partial \Theta} = \int_{z_0}^{z_1} Q \mathbf{w} \left(\int_{z_0}^z \frac{1}{u^2 \tau} \frac{\partial u}{\partial \Theta} dz' \right) dz \quad (\text{A.24})$$

$$= \left[\mathbf{y} \int_{z_0}^z \frac{1}{u^2 \tau} \frac{\partial u}{\partial \Theta} dz \right]_{z_0}^{z_1} - \int_{z_0}^{z_1} \mathbf{y} \frac{1}{u^2 \tau} \frac{\partial u}{\partial \Theta} dz \quad (\text{A.25})$$

$$= \int_{z_0}^{z_1} (\mathbf{y}_1 - \mathbf{y}) \frac{1}{u^2 \tau} \frac{\partial u}{\partial \Theta} dz \quad (\text{A.26})$$

$$= \int_{z_0}^{z_1} \overleftarrow{\mathbf{y}} \frac{\partial u}{u^2 \tau} dz \text{ où } \overleftarrow{\mathbf{y}} = \mathbf{y}_1 - \mathbf{y} \quad (\text{A.27})$$

$$= \int_{t_0}^{t_1} \overleftarrow{\mathbf{y}} \frac{\partial u}{u\tau} dt \quad (\text{A.28})$$

A.2 Rétropropagation

On repart de la définition du gradient de l'erreur selon les paramètres d'état du système

$$\dot{\mathbf{X}}^* = \frac{\partial \dot{\epsilon}}{\partial \mathbf{X}} = \left\langle \frac{\partial \epsilon}{\partial \mathbf{X}}, -\frac{\partial \mathbf{f}}{\partial \mathbf{X}} \right\rangle = -\left\langle \mathbf{X}^*, \frac{\partial \mathbf{f}}{\partial \mathbf{X}} \right\rangle \quad (\text{A.29})$$

La variable \mathbf{y}^* est constante car $\frac{\partial \mathbf{f}}{\partial \mathbf{y}} = 0$. Autrement dit $\forall t, y(t) = \mathbf{y}_1^*$. La variable z^* est connue sans avoir à intégrer :

$$z^* = \frac{\partial \epsilon}{\partial z} = -\left\langle (Q^* \mathbf{y}^*), \begin{bmatrix} \frac{dQ}{dz} \\ \frac{d\mathbf{y}}{dz} \end{bmatrix} \right\rangle = -\left\langle (Q^* \mathbf{y}^*), \begin{bmatrix} -Q \\ \frac{1}{u\tau} \\ Q\mathbf{w} \end{bmatrix} \right\rangle. \quad (\text{A.30})$$

Pour Q et Θ , l'intégration est en revanche nécessaire :

$$(\dot{Q}^* \dot{\Theta}^*) = -\left\langle (Q^* \mathbf{y}^*), \begin{bmatrix} -\frac{1}{u\tau} & -\frac{Q}{u^2 \tau} \frac{\partial Q}{\partial \Theta} \\ \mathbf{w} & 0 \end{bmatrix} \right\rangle. \quad (\text{A.31})$$

A.2.1 Charge du nuage électronique

On s'intéresse tout d'abord au calcul de l'erreur sur la charge du nuage électronique Q^* :

$$\dot{Q}^* = \frac{\partial \epsilon}{\partial Q} = Q^* \frac{1}{u\tau} - \langle \mathbf{y}^*, \mathbf{w} \rangle \quad (\text{A.32})$$

L'intégration temporelle de cette équation différentielle donne

$$Q^* = e^{t/\tau} \left\langle \mathbf{y}^*, \int_t^{t_1} e^{-t/\tau} \mathbf{w} u dt \right\rangle \quad (\text{A.33})$$

$$= \frac{1}{Q(t)} \left\langle \mathbf{y}^*, \int_t^{t_1} Q(t) \mathbf{w} u dt \right\rangle \quad (\text{A.34})$$

$$= \frac{\langle \mathbf{y}^*, \mathbf{y}_1^* - \mathbf{y} \rangle}{Q} \quad (\text{A.35})$$

$$= \frac{\langle \mathbf{y}^*, \overleftarrow{\mathbf{y}} \rangle}{Q} \quad (\text{A.36})$$

Ce qui donne pour le calcul de l'erreur sur la charge initiale

$$Q_0^* = \left\langle \mathbf{y}^*, \frac{\partial \mathbf{y}_1}{\partial Q_0} \right\rangle = \left\langle \mathbf{y}^*, \int \frac{Q}{Q_0} \mathbf{w} dz \right\rangle = \frac{\langle \mathbf{y}_1^*, \mathbf{y}_1 \rangle}{Q_0} \quad (\text{A.37})$$

Le sens physique de cette équation est clair : la charge initiale a un effet proportionnel sur l'induction de signal donc la dérivée est égale au ratio des deux valeurs.

A.2.2 Position initiale

A présent on calcule la dépendance de l'erreur sur la profondeur d'interaction z :

$$z^* = Q^* \frac{Q}{u\tau} - \langle \mathbf{y}^*, Q \mathbf{w} \rangle \quad (\text{A.38})$$

$$= \left\langle \mathbf{y}^*, \left(\frac{\overleftarrow{\mathbf{y}}}{u\tau} - Q \mathbf{w} \right) \right\rangle \quad (\text{A.39})$$

$$z_0^* = \left\langle \mathbf{y}^*, \left(\frac{\mathbf{y}_1}{u_0\tau} - Q_0 \mathbf{w}_0 \right) \right\rangle \quad (\text{A.40})$$

On retrouve le résultat obtenu pour la résolution de l'équation en sens direct. Physiquement, on peut expliquer cette équation par le fait qu'éloigner le pont de départ peut allonger le temps de transit et induire un piégeage supplémentaire mais permet également d'augmenter l'induction.

A.2.3 Paramètres de champ

Enfin concernant les paramètres de champ Θ on a

$$\dot{\Theta}^* = - \left\langle \mathbf{X}^*, \frac{\partial \mathbf{f}}{\partial \Theta} \right\rangle \quad (\text{A.41})$$

$$= -Q^* \frac{Q}{u^2\tau} \frac{\partial u}{\partial \Theta} \quad (\text{A.42})$$

$$= \frac{\langle \mathbf{y}^*, \overleftarrow{\mathbf{y}} \rangle}{u^2\tau} \frac{\partial u}{\partial \Theta} \quad (\text{A.43})$$

$$\Theta^* = \left\langle \mathbf{y}^*, \left(\int \frac{\overleftarrow{\mathbf{y}}}{u\tau} \frac{\partial u}{\partial \Theta} dt \right) \right\rangle \quad (\text{A.44})$$

On peut citer la variante temporelle, qui a l'intérêt d'être plus générale applicable au delà du cas unidimensionnel et pour des signaux filtrés.

Abstract :

Since a few years, the wide spread use of CZT-based detectors in gamma imaging drives their performance optimization to stay competitive at the industrial level. However, the presence of structural defect in the CZT crystal deteriorates the output signals quality and holds back the higher volume detectors development.

The purpose of this thesis is the use of optimization and artificial intelligence algorithms using realistic simulations to limit the impact of the defects and improve the localization performances of gamma interactions in the detector. We will develop a mathematical-based method in three steps as an alternative to common characterization and correction methods.

First, we develop 3D CZT detector simulations enabling to implement defects with different natures to observe their impact on output signals. Then we build a simple neural network, which can be introduced in the electronics to localize the gamma interactions in the detector from simulation results. A second network based on a gradient computation method will allow determining the electric field and collection performance of a detector.

The addition of these three steps will be used to learn through simulation the intern parameters of a determined detector such as the electric field. This simulation will serve to train the simple neural network and finally be used on experimental data to improve the localization performance of the detector.

The development of this mathematical approach will help us having a better understanding of the intern structure of a CZT crystal being able to reproduce its behavior in simulation. In addition, the better performance of the detector might be sufficient to decrease the radiotracer dose for medical imaging or limit the exposition time of operators in a nuclear power plant.

Résumé :

L'utilisation généralisée des détecteurs à base de CZT dans l'imagerie gamma depuis certaines années pousse à l'optimisation de leurs performances afin de rester compétitif au niveau industriel. Néanmoins la présence de défauts structurels dans la maille cristalline du CZT dégrade la qualité de signaux de sortie et freine le développement de détecteurs à plus grand volume.

L'objectif de cette thèse est d'utiliser des algorithmes d'optimisation et d'intelligence artificielle en s'aidant de simulations réalistes afin de limiter les effets de ces défauts et d'améliorer les performances de localisation des interactions gamma dans le détecteur. Nous allons nous baser sur une méthode mathématique en 3 étapes alternative aux méthodes de caractérisation et de correction usuelles.

D'abord le développement de simulations de détecteurs CZT en 3 dimensions avec possibilité d'implémenter des défauts de différentes natures pour observer leur impact sur les signaux sortants. Puis nous allons construire un réseau de neurones simple adaptable dans l'électronique afin de localiser les interactions dans le détecteur d'après des résultats de simulation. Un second réseau basé sur une méthode d'optimisation par calcul de gradient va permettre de déterminer le champ électrique et les performances de collecte d'un détecteur.

La concaténation de ces 3 étapes va permettre d'apprendre par simulation les paramètres internes d'un détecteur étudié tel que le champ électrique, simulation qui va servir afin d'entraîner le réseau de neurones simple, pour finalement être utilisé sur les données expérimentales dans le but d'améliorer les performances de localisation du détecteur.

Le développement de cette approche mathématique va permettre d'un côté la meilleure compréhension de la structure interne du cristal CZT et la possibilité de reproduire son comportement en simulation. Également les meilleures performances du détecteur pourraient se retrouver dans l'utilisation de doses de radio traceurs inférieures en imagerie, ou un temps d'exposition réduit pour les opérateurs en centrale nucléaire.