



**HAL**  
open science

# Système pour le diagnostic du cancer du sein exploitant la diffusion RX à petit angle

Vera Feldman

► **To cite this version:**

Vera Feldman. Système pour le diagnostic du cancer du sein exploitant la diffusion RX à petit angle. Biophysique [physics.bio-ph]. Université Grenoble Alpes, 2019. Français. NNT : 2019GREAY057 . tel-02538620

**HAL Id: tel-02538620**

**<https://theses.hal.science/tel-02538620>**

Submitted on 9 Apr 2020

**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 LA COMMUNAUTE UNIVERSITE GRENOBLE ALPES**

Spécialité : **Physique appliquée**

Arrêté ministériel : 25 mai 2016

Présentée par

**Vera FELDMAN**

Thèse dirigée par **Jean-Louis HAZEMANN**, Directeur de recherche, Université Grenoble Alpes, et co-encadrée par **Caroline PAULUS**, Ingénieure de recherche, CEA LETI et **Joachim TABARY**, Ingénieur de recherche, CEA LETI.

préparée au sein du **CEA LETI**  
dans l'**École Doctorale de Physique de Grenoble**

## **Système pour le diagnostic du cancer du sein exploitant la diffusion RX à petit angle**

Thèse soutenue publiquement le **7 novembre 2019**,  
devant le jury composé de :

**Madame Hélène ELLEAUME**

Directrice de recherche, ESRF, Présidente du jury

**Monsieur Dominique BAZIN**

Directeur de recherche, CNRS, Rapporteur

**Madame Paola COAN**

Professeure, Ludwig Maximilians University, Rapporteuse

**Monsieur Philippe DUVAUCHELLE**

Maître de conférences, INSA Lyon, Examineur

**Monsieur Jérôme KIEFFER**

Ingénieur, ESRF, Examineur

**Monsieur Jean-Louis HAZEMANN**

Directeur de recherche, ESRF, Directeur de thèse





# REMERCIEMENTS

---

Mon chemin de thèse touche à sa fin et j'aimerais le conclure en remerciant chaleureusement toutes les personnes qui, de près ou de loin, ont contribué à cette expérience exceptionnelle.

Tout d'abord, je remercie chacun des membres du jury de thèse pour le temps que vous avez consacré à la lecture de mon manuscrit, pour votre intérêt et pour toutes vos questions et remarques. Je voudrais remercier en particulier mes rapporteurs, Dominique Bazin et Paola Coan, ainsi que Hélène Elleaume et Jérôme Kieffer, qui faisaient partie de mon comité de suivi individuel.

Merci aux personnes qui m'ont guidé pendant ces trois années de travail : mon directeur de thèse, Jean-Louis Hazemann, et mes co-encadrants, Caroline Paulus et Joachim Tabary. Merci pour votre sympathie, pour le temps que vous avez consacré à nos réunions, pour vos conseils et votre savoir-faire, que vous avez partagé avec moi.

Merci à tous mes collègues du ex-LDET, du ex-LSIV et du LASP, de m'avoir accueillie dans le village des irréductibles gaulois dans la joie et la bonne humeur. Les échanges matinaux autour du café sur des sujets d'actualité cruciaux me manquent déjà. Un énorme merci à Marie-Claude et Olivier pour votre expertise et votre aide dans la réalisation des manip sur des fantômes exotiques. Merci à Clarisse, c'était un plaisir de partager le bureau avec toi. Merci à Laurence et Nathalie, pour votre assistance dans les tâches administratives et les réponses à mes questions existentielles.

Merci aux thésards que j'ai eu la chance de côtoyer pendant ces trois ans, et particulièrement à Mélanie, pour l'initiation à l'escalade, Cinzia pour ta bonne humeur, et Pierre-Antoine, pour les blagues sur les gnus. Merci à Odran, qui partageait mon enthousiasme pour la rédaction. Basile et Orianne, je vous souhaite bon courage pour la suite.

Je remercie enfin ma famille et mes proches pour vos encouragements et pour tous les moments qui m'ont permis de me changer les idées quand j'en avais besoin. Merci également de ne pas avoir demandé trop souvent si ma thèse se passait bien, ni où j'en étais dans la

---

rédaction. Un grand merci à Émilie pour ton aide lors de mon installation à Grenoble, pour ton support moral et pour la relecture du manuscrit. Et, *last but not least*, merci à toi, Xavier, pour ton support quotidien, ta confiance en moi et ton amour. J'ai beaucoup de chance de t'avoir à mes côtés.



---

# LISTE DES ACRONYMES

---

<b>ADXRD</b>	<i>Angle Dispersive X Ray Diffraction</i> (diffusion cohérente de rayons X à dispersion d'angle)
<b>AGD</b>	<i>Average Glandular Dose</i> (dose glandulaire moyenne)
<b>AUC</b>	<i>Area Under Curve</i> (aire sous la courbe)
<b>BI-RADS</b>	<i>Breast Imaging-Reporting And Data System</i>
<b>CA</b>	<i>Coded Aperture</i> (masque codé)
<b>CACSSI</b>	<i>Coded Aperture Coherent Scatter Spectral Imaging</i>
<b>CAXSI</b>	<i>Coded Aperture X ray Scatter Imaging</i>
<b>CCI</b>	Carcinome Canalaire Infiltrant
<b>CCIS</b>	Carcinome Canalaire in situ
<b>CdTe</b>	tellure de cadmium
<b>CdZnTe</b>	tellure de cadmium dopé au zinc
<b>CEDM</b>	<i>Contrast-Enhanced Digital Mammography</i>
<b>CESM</b>	<i>Contrast Enhanced Spectral Mammography</i>
<b>CHUGA</b>	Centre Hospitalier Universitaire Grenoble Alpes
<b>CLI</b>	Carcinome Lobulaire Infiltrant
<b>CR</b>	Classification par Référence
<b>CSCCT</b>	<i>Coherent Scatter Computed Tomography</i>
<b>DBT</b>	<i>Digital Breast Tomosynthesis</i> (tomosynthèse numérique mammaire)
<b>DQE</b>	<i>Detective Quantum Efficiency</i> (efficacité quantique de détection)
<b>DRM</b>	<i>Detector Response Matrix</i> (matrice de réponse du détecteur)
<b>EDXRD</b>	<i>Energy Dispersive X Ray Diffraction</i> (diffusion cohérente de rayons X à dispersion d'énergie)

<b>FCG</b>	<i>Focal Construct Geometry</i>
<b>FN</b>	<i>False Negative</i> (faux négatif)
<b>FP</b>	<i>False Positive</i> (faux positif)
<b>FWHM</b>	<i>Full Width at Half-Maximum</i> (largeur à mi-hauteur)
<b>HSV</b>	<i>Hue Saturation Value</i> (teinte saturation valeur)
<b>IRIG</b>	Institut de Recherche Interdisciplinaire de Grenoble
<b>IRM</b>	Imagerie par Résonance Magnétique
<b>k-NN</b>	<i>k-Nearest Neighbors</i> (k plus proches voisins)
<b>LDA</b>	<i>Linear Discriminant Analysis</i> (analyse discriminante linéaire)
<b>MLEM</b>	<i>Maximum Likelihood Expectation Maximization</i>
<b>PCA</b>	<i>Principal Component Analysis</i> (analyse en composantes principales)
<b>PixD</b>	<i>Pixellated x ray Diffraction</i>
<b>PMMA</b>	<i>Polymethyl Methacrylate</i>
<b>PS</b>	Polystyrène
<b>RGB</b>	<i>Red Green Blue</i> (Rouge Vert Bleu)
<b>ROC</b>	<i>Receiver Operating Characteristic</i> (caractéristique de fonctionnement du récepteur)
<b>SVM</b>	<i>Support Vector Machine</i> (machine à vecteurs de supports)
<b>TN</b>	<i>True Negative</i> (vrai négatif)
<b>TP</b>	<i>True Positive</i> (vrai positif)
<b>UCL</b>	University College London
<b>XRD</b>	<i>X Ray Diffraction</i> (diffusion cohérente de rayons X)

# LISTE DES SYMBOLES

---

Symbole	Description	Unité / valeur
$\theta$	Angle de diffusion défini comme l'angle entre le rayon incident et le rayon diffusé	rad
$h$	Constante de Planck	$6.626 \cdot 10^{-34}$ J.s
$c$	Célérité de la lumière dans le vide	$2.998 \cdot 10^8$ m.s <sup>-1</sup>
$D_a$	Dose absorbée	mGy
$d$	Distance interréticulaire	m
$E$	Énergie du rayonnement	keV
$I$	Intensité du courant électrique	mA
$p$	Pixel	
$(X Y)$	Produit scalaire de $X$ par $Y$	
$(X \otimes Y)$	Produit terme à terme de $X$ par $Y$	
$U$	Tension électrique	kV
$q$	Transfert de quantité de mouvement ( <i>momentum transfer</i> )	nm <sup>-1</sup>

---



# SOMMAIRE

---

<b>Remerciements</b>	<b>ii</b>
<b>Liste des acronymes</b>	<b>vii</b>
<b>Liste des symboles</b>	<b>ix</b>
<b>Introduction</b>	<b>1</b>
<b>1 Contexte médical et technologique</b>	<b>3</b>
1.1 Outils de diagnostic cliniques . . . . .	3
1.1.1 Mammographie . . . . .	4
1.1.2 Tomosynthèse . . . . .	5
1.1.3 Échographie mammaire . . . . .	6
1.1.4 Imagerie par Résonance Magnétique : IRM . . . . .	6
1.1.5 Analyse anatomo-pathologique . . . . .	7
1.1.6 Conclusion . . . . .	7
1.2 Diffusion cohérente de rayons X pour l'imagerie du sein . . . . .	7
1.2.1 Principes généraux de l'interaction rayonnement - matière . . . . .	8
1.2.2 Diffusion cohérente . . . . .	9
1.2.3 Schéma général des systèmes XRD . . . . .	11
1.2.4 Applications envisagées . . . . .	15
1.2.5 Conclusion . . . . .	16
1.3 Systèmes XRD pour la caractérisation des tissus mammaires . . . . .	16
1.3.1 Systèmes XRD proposés dans la littérature . . . . .	16
1.3.2 Systèmes développés au CEA . . . . .	25
1.3.3 Conclusion . . . . .	34
1.4 Contexte médical et physique : conclusion . . . . .	34
<b>2 Méthode proposée</b>	<b>35</b>
2.1 Géométrie et modèle . . . . .	35
2.1.1 Montage expérimental . . . . .	35

2.1.2	Construction du modèle direct . . . . .	39
2.1.3	Conclusion . . . . .	43
2.2	Estimation de la matrice système . . . . .	43
2.2.1	Matrice de réponse du détecteur . . . . .	43
2.2.2	Estimation du spectre incident atténué . . . . .	45
2.2.3	Calibration de la matrice de réponse géométrique $R_c$ . . . . .	46
2.2.4	Construction de la matrice globale . . . . .	51
2.2.5	Conclusion . . . . .	51
2.3	Reconstruction : inversion du modèle direct . . . . .	52
2.4	Classification : présentation des méthodes . . . . .	53
2.4.1	Classification par référence . . . . .	53
2.4.2	$k$ plus proches voisins . . . . .	54
2.4.3	Machine à vecteur de supports . . . . .	54
2.4.4	Analyse discriminante linéaire . . . . .	55
2.4.5	Analyse en composantes principales . . . . .	55
2.4.6	Conclusion . . . . .	56
2.5	Estimation de la dose absorbée . . . . .	56
2.5.1	Dose absorbée : définitions . . . . .	57
2.5.2	Méthode d'estimation . . . . .	58
2.6	Étude de l'influence de la dose sur la classification . . . . .	60
2.6.1	Métriques . . . . .	60
2.6.2	Simulation . . . . .	61
2.6.3	Classification en fonction de la dose . . . . .	63
2.6.4	Analyse et discussion . . . . .	66
2.7	Système : conclusion . . . . .	66
<b>3</b>	<b>Résultats</b>	<b>69</b>
3.1	Fantôme plastique . . . . .	69
3.1.1	Mesures expérimentales . . . . .	69
3.1.2	Reconstruction . . . . .	71
3.1.3	Classification . . . . .	73
3.1.4	Conclusion . . . . .	74
3.2	Mise au point de fantômes réalistes . . . . .	76
3.2.1	Recherche d'échantillons de tissus mammaires humains . . . . .	76
3.2.2	Glandes mammaires de souris . . . . .	78
3.2.3	Fantômes de viande . . . . .	80
3.3	Fantômes de viande . . . . .	82
3.3.1	Mesures expérimentales . . . . .	84
3.3.2	Reconstruction . . . . .	86
3.3.3	Classification . . . . .	89

---

3.3.4 Conclusion . . . . .	96
3.4 Conclusion . . . . .	97
<b>Conclusions et perspectives</b>	<b>99</b>
<b>A Classification BI-RADS</b>	<b>105</b>
<b>B Actes de conférences</b>	<b>107</b>
<b>Bibliographie</b>	<b>125</b>
<b>Résumés</b>	<b>130</b>



# INTRODUCTION

---

Le cancer du sein est une pathologie très répandue, qui touche environ une femme sur huit en France. C'est pourquoi, son diagnostic et son traitement sont très étudiés dans le domaine de la recherche médicale. En effet, la détection précoce de cette pathologie permet de la soigner au mieux et de minimiser les séquelles. La mammographie est l'outil le plus utilisé, mais elle est souvent complétée par d'autres méthodes d'imagerie, telles que l'échographie ou l'IRM. Cependant, toutes les méthodes ont des limites, c'est pourquoi de nombreuses équipes de recherche travaillent sur des méthodes alternatives.

La diffusion cohérente des rayons X est une des méthodes explorées. Dans le cas des matières cristallines, la diffusion cohérente donne lieu au phénomène de diffraction de rayons X. La découverte de ce phénomène, qui permet de caractériser les cristaux, a valu à Max von Laue le Prix Nobel de Physique en 1914. Par ailleurs, la diffusion cohérente permet également d'étudier des matières amorphes, telles que les tissus biologiques.

Plusieurs équipes de recherche étudient l'utilisation de la diffusion cohérente de rayons X pour le diagnostic du cancer du sein, et leurs résultats attestent de l'intérêt de cette méthode. Toutefois, la plupart de ces équipes travaillent en conditions incompatibles avec les contraintes d'une utilisation clinique : utilisation d'un synchrotron comme source de rayons X, détecteurs non résolus en énergie ou utilisant des systèmes de refroidissement encombrants, temps d'acquisition extrêmement longs, etc.. Au CEA Leti de Grenoble, des détecteurs spectrométriques à tellure de cadmium dopé au zinc (CdZnTe), fonctionnant à température ambiante, ont été développés. Ces détecteurs sont optimaux en terme de compacité, de sensibilité, et de résolution énergétique et spatiale, et offrent de nouvelles possibilités pour la mise au point de systèmes compatibles avec les contraintes de l'imagerie médicale. C'est pourquoi, la faisabilité d'un tel système a été étudiée par Fanny MARTICKE et par Damien BARBES au sein du CEA Leti, pendant leurs travaux de thèse. À l'issue de leurs études en simulation et par expérimentation, ils ont proposé des systèmes permettant de discerner le tissu cancéreux du tissu sain.

Cette thèse s'inscrit dans la continuité de leurs travaux, dans un objectif d'amélioration des per-

formances, afin de s'approcher des exigences de l'imagerie médicale, notamment en terme de dose déposée et de fiabilité. La thèse est organisée en trois chapitres.

Le premier chapitre traite du contexte médical et technologique associé à la diffusion cohérente de rayons X. Les outils de diagnostic de cancer du sein utilisés cliniquement, ainsi que leurs avantages et inconvénients sont présentés. Ensuite, nous décrivons les principes physiques de la diffraction cohérente des rayons X, et expliquons ce que cette méthode peut apporter à l'imagerie du sein. Finalement, nous présentons les systèmes expérimentaux, exploitant cette technique, proposés par les différentes équipes de recherche.

Le deuxième chapitre est dédié au système expérimental et analytique mis au point pendant cette thèse. Il est composé d'un montage expérimental, permettant la mesure des photons diffusés, et d'une chaîne de traitement de données, qui permet de caractériser les tissus à partir de cette mesure. Le montage expérimental a été initié par Damien BARBES pendant ses travaux de thèse. Le traitement utilisé n'avait pas mené aux résultats attendus, c'est pourquoi une autre méthode de traitement de données a été mise en place pendant cette thèse. Le traitement proposé est effectué en plusieurs étapes : tout d'abord un modèle de réponse du système est calculé en utilisant des connaissances *a priori* et des données de calibration. Ensuite, la réponse du système est inversée pendant l'étape de reconstruction afin de remonter aux fonctions de diffusion du tissu, qui caractérisent sa structure moléculaire. Finalement, une classification en tissu sain ou cancéreux est effectuée. Les rayons X étant ionisants, la dose absorbée et son influence sur les performances du système sont également étudiées.

Enfin, dans le troisième chapitre, nous présenterons nos résultats expérimentaux sur différents types de fantômes. Dans un premier temps, nous avons effectué des mesures expérimentales dans des conditions quasi-identiques aux expériences de Damien BARBES, en utilisant un fantôme plastique. Nous avons observé une nette amélioration des résultats, ce qui nous a permis de valider notre méthode de traitement des données. Ensuite, nous avons cherché à réaliser des expériences sur des fantômes plus réalistes. Après plusieurs essais, nous avons mis au point des fantômes à base de tissus bovins. Les résultats obtenus sont globalement encourageants, cependant les performances obtenues ne correspondent pas encore aux exigences de l'imagerie médicale.

En conclusion, nous présenterons une synthèse des résultats obtenus, et proposerons des pistes d'amélioration de la méthode.

# CONTEXTE MÉDICAL ET TECHNOLOGIQUE

---

Le cancer du sein chez la femme est un enjeu majeur de santé publique : en France en 2017 cette pathologie a donné lieu à 59 000 nouveaux cas et a engendré 12 000 décès, d'après l'Institut National du Cancer (INCa) [1]. Le dépistage est essentiel, car un diagnostic précoce augmente les chances de réussite du traitement et diminue les séquelles.

Dans ce chapitre nous décrivons les différents types de cancer et les outils cliniques de diagnostic avec leurs points forts et leurs limites. Ensuite, nous verrons les principes physiques de la diffusion cohérente de rayons X (XRD, *X Ray Diffraction*) et ce qu'elle peut apporter à l'imagerie du sein. Finalement, une revue des systèmes XRD étudiés récents sera proposée.

## 1.1 Outils de diagnostic cliniques

La structure du sein est représentée sur la figure 1.1 : les lobules ont pour fonction la production de lait, qui est ensuite amené jusqu'au mamelon par les canaux. L'ensemble est entouré de tissu adipeux. Globalement, le sein sain est constitué de tissu fibroglandulaire et de tissu adipeux, en proportion variable. La densité du sein augmente avec la proportion de tissu fibro-

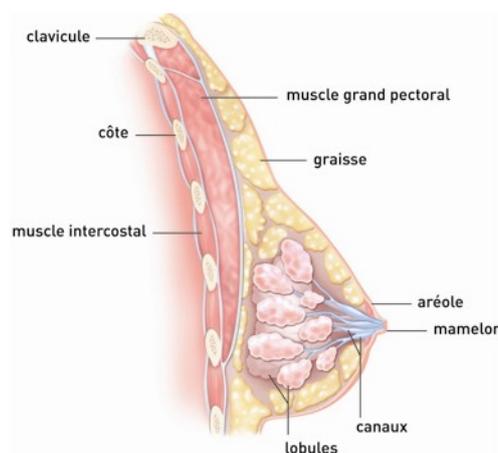


FIGURE 1.1 – Schéma représentant la structure du sein, coupe sagittale [1]

glandulaire.

La plupart des cancers du sein se développent à partir des cellules des lobules ou des canaux : on parle alors respectivement de cancers "*lobulaires*" ou "*canalaires*". Si les cellules cancéreuses restent confinées en leur point d'origine, le cancer est dit "*in situ*", par opposition au cas où les cellules se propagent dans les tissus avoisinants, et où le cancer est dit "*infiltrant*". Les trois types de cancers les plus fréquents par ordre d'incidence sont :

- le Carcinome Canalaire Infiltrant (CCI)
- le Carcinome Canalaire in situ (CCIS)
- le Carcinome Lobulaire Infiltrant (CLI)

Le CCIS est un cancer non-invasif bénin. Cependant, il présente un risque d'évolution vers un cancer invasif, et les femmes ayant eu un CCIS, ont plus de chances de développer un cancer invasif dans le futur [2]. La présence de micro-calcifications, dépôts minuscules de calcium, peut indiquer la présence de ce type de cancer [1, 3]. La tumorectomie est le traitement le plus fréquent pour ce type de cancer [4].

Les carcinomes canalaire et lobulaire infiltrants représentent respectivement 80% et 10% des cancers du sein infiltrants [4]. Ce sont des cancers invasifs, qui affectent surtout les femmes de plus de 50 ans. Ils sont traités de manière chirurgicale. Des traitements complémentaires, comme la radiothérapie ou la chimiothérapie, peuvent être également réalisés.

En fonction de la technique d'imagerie, les cancers peuvent être plus ou moins faciles à détecter en fonction de leur type. De manière générale, la **mammographie** est utilisée pour le dépistage et pour le diagnostic en cas d'apparition de symptômes, tels que la douleur ou la déformation anormale. Ces dernières années, la **tomosynthèse** vient souvent compléter, voire remplacer la mammographie, en tant qu'outil de diagnostic. Si les résultats présentent une incertitude ou si la patiente a des seins particulièrement denses, la mammographie est suivie d'une **échographie** ou d'un **examen par Imagerie par Résonance Magnétique (IRM)**. S'il existe une anomalie, même faiblement suspecte, un prélèvement de tissu sera réalisé par biopsie pour effectuer une **analyse anatomo-pathologique**. Cette analyse est très fiable et est effectuée avant toute décision de traitement.

Il existe d'autres méthodes d'imagerie mammaire, telles que la tomographie par émission de positons. Mais elles sont plutôt utilisées une fois qu'une tumeur est confirmée pour déterminer l'avancement de la maladie ou pour en assurer le suivi. Dans cette thèse, on ne s'intéresse qu'aux techniques de dépistage. C'est pourquoi seules les cinq méthodes mentionnées ci-dessus sont présentées.

### 1.1.1 Mammographie

La mammographie est la méthode la plus répandue dans le diagnostic mammaire. Il s'agit d'un examen radiographique dédié à l'imagerie du sein. C'est une méthode rapide (environ 15 minutes pour l'examen complet), relativement indolore et non invasive. De plus, les images ob-

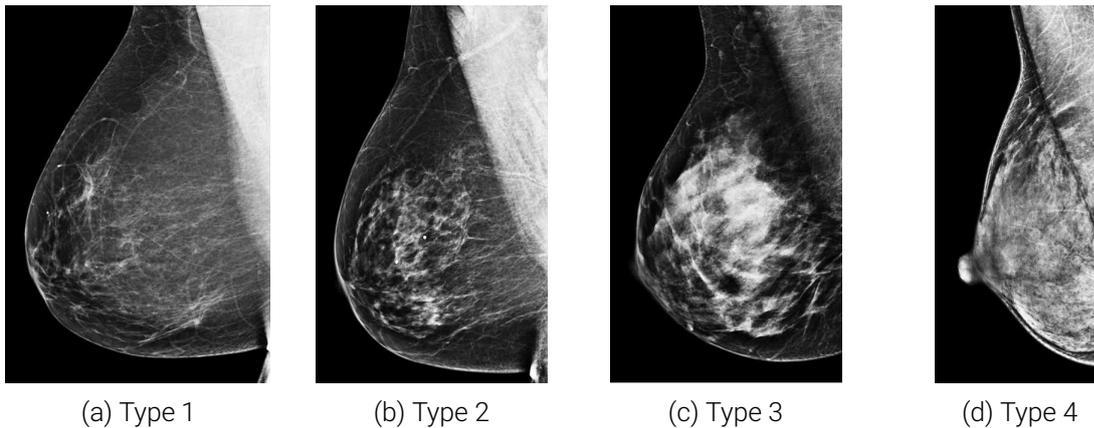


FIGURE 1.2 – Images obtenues par mammographie pour différentes catégories de densité du sein [5], projection oblique médiolatérale

tenues par mammographie ont une très haute résolution spatiale ( $< 100\mu m$ ), ce qui permet de détecter les micro-calcifications.

Pour cette modalité, chaque sein est imagé avec deux prises de vue pour obtenir deux projections, appelées *oblique médiolatérale* et *craniocaudale*. Pour chaque vue, le sein est comprimé jusqu'à une épaisseur de 5-6 cm. Les rayons X de basse énergie sont utilisés (environ 30 keV). En fonction de leur densité, les tissus atténuent plus ou moins le rayonnement incident. Le rayonnement atténué est mesuré par un détecteur X. Dans les systèmes les plus récents, sont utilisés les détecteurs à conversion directe (voir §1.3), qui ont une efficacité quantique de détection (DQE, *Detective Quantum Efficiency*) de 60% [6]. La dose glandulaire moyenne (AGD, *Average Glandular Dose*) varie autour de 1-2 mGy selon les systèmes et les modèles de calcul utilisés [7, 8].

Cependant, la densité globale du sein impacte grandement le résultat. American College of Radiology a proposé en 1998 une classification appelée *Breast Imaging-Reporting And Data System* (BI-RADS). Cette classification distingue, entre autres, quatre types de sein selon leur densité (voir annexe A, table A.1). Un exemple de classification est présenté sur la figure 1.2 [5]. Pour les types 3 et 4, la présence importante de tissu fibroglandulaire peut cacher des masses anormales. Le risque de non-détection est donc plus important dans les seins denses [9]. L'injection d'un produit de contraste (*Contrast-Enhanced Digital Mammography* (CEDM) [10]) permet d'augmenter la sensibilité, notamment dans le cas de seins denses [11].

Les produits de contraste sont également utilisés dans la mammographie spectrale. La technique correspondante porte le nom de *Contrast Enhanced Spectral Mammography* (CESM) [12, 13]. Elle est assez récente, c'est pourquoi elle est encore peu présente dans les hôpitaux.

### 1.1.2 Tomosynthèse

La tomosynthèse, ou tomosynthèse numérique mammaire (DBT, *Digital Breast Tomosynthesis*), tout comme la mammographie, utilise le rayonnement X pour produire du signal. Cepen-

dant, la caméra ou le détecteur se déplace pour acquérir des projections avec des angles différents. Après reconstruction, cela permet d'obtenir des coupes en profondeur, et donc des mesures volumétriques. Certains imageurs permettent de réaliser la DBT en même temps que la mammographie. Selon les systèmes et les modèles de calcul utilisés, la dose varie entre 1 et 4 mGy [14], ce qui est très proche de la mammographie. Aux États-Unis, la DBT est, comme la mammographie, une technique d'examen et de dépistage depuis 2016.

### 1.1.3 Échographie mammaire

Lors d'une échographie, un transducteur en contact direct avec le sein, émet des ondes sonores à haute fréquence, et mesure les ondes réfléchies. Cette méthode a l'avantage de produire des images 2D en temps réel sans ionisation. Le praticien peut adapter la direction d'inspection en temps réel et donc étudier la zone d'intérêt sous différents angles, si nécessaire. Cette technique est souvent utilisée en complément de la mammographie car elle est plus efficace sur les seins denses, et améliore considérablement le diagnostic. Cependant, en tant qu'outil de dépistage chez les patientes asymptomatiques, l'échographie est très peu fiable. C'est pourquoi, pour le dépistage elle est toujours utilisée en complément d'une autre méthode [9].

### 1.1.4 Imagerie par Résonance Magnétique : IRM

L'IRM est une technique 3D qui possède des avantages très intéressants :

- elle n'utilise pas de rayonnement ionisant et ne présente pas d'effets secondaires ;
- elle permet d'imager les deux seins en même temps et sans compression ;
- elle est efficace sur les seins denses ;
- elle est indolore ;

Mais elle a également des limites : sa résolution spatiale est de l'ordre du millimètre, c'est pourquoi elle ne permet pas de détecter de manière sûre les micro-calcifications. Elle a un taux assez élevé de faux positifs et, de manière générale, est moins performante que les autres méthodes pour le diagnostic du CCIS [9]. Un autre inconvénient est plutôt de l'ordre logistique : d'après l'étude 2018 de Cemka, le délai d'attente moyen en France pour un examen prescrit en cancérologie est de 32 jours. À cause de ce facteur, l'IRM n'est pas utilisée de manière systématique pour le diagnostic. Par contre, une fois que la tumeur est confirmée, elle est utilisée pour déterminer le traitement le plus adapté. Elle est également quasi-systématique chez les patientes qui ont des prédispositions génétiques au cancer du sein.

D'après une étude publiée en 2017, l'IRM pourrait être concurrencée par la CESM [15]. Dans cette étude, les résultats des diagnostics obtenus par IRM et CESM ont été comparés sur 66 cas de lésions confirmées. Pour les deux modalités, la sensibilité était de 100% et la résolution spatiale d'environ 4 mm. Par contre, la CESM nécessite moins de temps d'examen et de temps d'accès aux soins.

### **1.1.5 Analyse anatomo-pathologique**

L'analyse anatomo-pathologique est l'ultime étape pour confirmer le diagnostic. Une biopsie est réalisée pour prélever une petite partie de tissu suspect. Ce tissu sera mis sous lame, puis coloré, et analysé au microscope. Ceci permet de déterminer si la lésion est bénigne ou maligne et, le cas échéant, de proposer le traitement le plus adapté. Cette méthode est particulièrement fiable, mais prend du temps, est invasive et très stressante pour les patientes.

En annexe A, la table A.2 indique la démarche à suivre en fonction des résultats d'imagerie. La biopsie, et l'analyse anatomo-pathologique qui suit, sont préconisées pour toute anomalie, même faiblement suspecte. Ceci explique pourquoi les résultats de biopsie sont négatifs la plus part du temps [16].

L'analyse anatomo-pathologique a également une utilisation per-opératoire : après une tumorectomie, elle permet d'effectuer un "contrôle de marges". Il s'agit de vérifier que la tumeur a été enlevée dans sa totalité : pour s'en assurer, l'analyse anatomo-pathologique sera effectuée pour vérifier si le tissu excisé présente des marges d'au moins quelques millimètres de tissu sain.

### **1.1.6 Conclusion**

Dans cette section, nous avons vu les examens cliniques principaux pour le diagnostic du cancer du sein. La biopsie est réalisée systématiquement en présence d'une anomalie suspecte, et dans la plupart des cas, le résultat de la biopsie est négatif. On peut en déduire que les méthodes d'imagerie, utilisées en amont, n'ont pas suffi à infirmer la présence de la tumeur. La biopsie étant invasive et stressante, avec des temps de réponse longs, il est intéressant de réduire le nombre d'examens "inutiles". Pour cela, des techniques d'imagerie, qui permettent de détecter de manière fiable l'absence de tumeur, sont nécessaires. Pour répondre à cette problématique, de nombreuses équipes de recherche travaillent pour améliorer les méthodes existantes ou proposer des méthodes nouvelles. Une des possibilités explorées, qui est étudiée dans cette thèse, est l'utilisation de la diffusion des rayons X, dont le principe physique et les possibilités d'utilisation sont présentées dans la section suivante.

## **1.2 Diffusion cohérente de rayons X pour l'imagerie du sein**

La diffusion cohérente de rayons X (XRD, *X Ray Diffraction*) est une technique de caractérisation des matériaux, qui permet d'évaluer leur structure moléculaire ou cristalline. Dans cette partie, nous verrons les principes physiques sur lesquels repose cette technique, ainsi que le schéma général des systèmes qui l'utilisent. Finalement, nous verrons les applications envisagées.

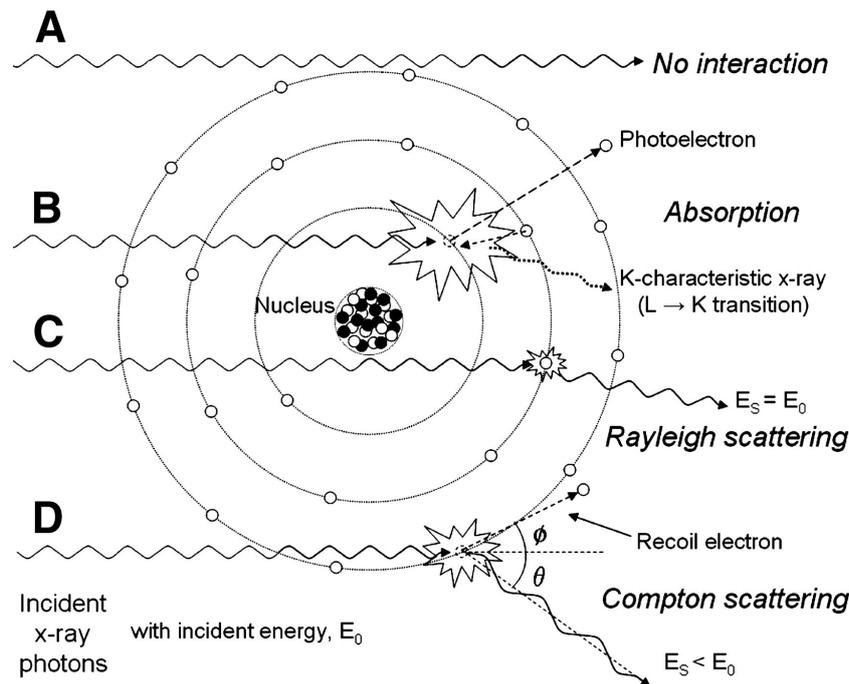


FIGURE 1.3 – Schéma illustrant les différents types d'interaction possibles entre un photon X et un atome [17]

### 1.2.1 Principes généraux de l'interaction rayonnement - matière

Le rayonnement X est un rayonnement électromagnétique, dont les longueurs d'onde sont comprises approximativement entre  $10^{-12}$  et  $10^{-8}$  mètres. L'énergie des photons X est comprise entre une centaine d'eV et quelques MeV. Ci-dessous, nous expliquons comment ces photons interagissent avec la matière au niveau atomique, puis comment ils sont produits et mesurés. Dans le cadre de cette thèse, l'énergie des photons est comprise entre 20 et 160 keV, et pour cette gamme, quatre scénarios d'interaction entre un photon X et un atome sont possibles. Ils sont illustrés sur la figure 1.3 :

- A** Le photon passe sans interagir avec l'atome. La probabilité de ce scénario diminue avec le numéro atomique de l'élément : plus il y a d'électrons dans le cortège électronique et plus il y a de chances que le photon interagisse avec l'un d'entre eux;
- B** Le photon est absorbé par *effet photoélectrique*. Lorsque l'énergie du photon est supérieure à l'énergie de liaison des électrons, le photon X incident peut éjecter un électron de sa couche et exciter l'atome. Un électron de la couche supérieure vient combler le trou ainsi créé. Ce faisant, cet électron perd de l'énergie, qui est émise sous forme de photon X, dont l'énergie est caractéristique de l'atome;
- C** Le photon est dévié de sa trajectoire, mais conserve son énergie : c'est la diffusion cohérente, ou diffusion Rayleigh;

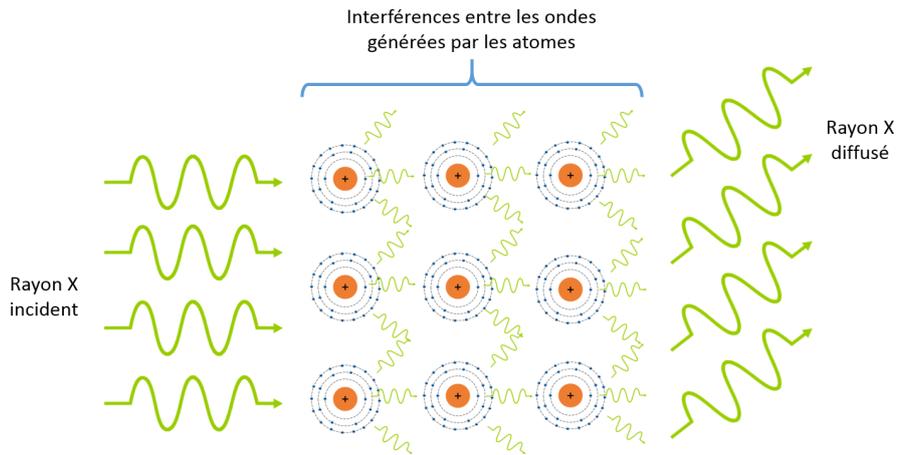


FIGURE 1.4 – Schéma illustrant la diffusion Rayleigh [18]

- D** Le photon est dévié de sa trajectoire et perd de l'énergie : c'est la diffusion incohérente, ou diffusion Compton.

### 1.2.2 Diffusion cohérente

La XRD repose sur la diffusion Rayleigh. Elle peut être décrite de manière ondulatoire : on considère le photon X comme une onde électromagnétique incidente, qui perturbe le cortège électronique d'un atome. Cette perturbation entraîne le déplacement du barycentre des charges négatives de l'atome, et forme ainsi un dipôle élastique qui oscille à la même fréquence, que l'onde incidente. Comme illustré sur la figure 1.4, ce phénomène s'additionne sur l'ensemble des atomes composant un matériau. Les ondes diffusées ainsi produites interfèrent entre-elles de manière plus ou moins constructive, selon la structure du matériau.

Dans le cas des cristaux, le phénomène d'interférences constructives est appelé "diffraction". Les atomes sont parfaitement ordonnés et les interférences constructives ont lieu pour des longueurs d'onde et des directions précises, décrites par la loi de Bragg :

$$2 d \sin(\theta/2) = n \lambda \quad (1.1)$$

où  $d$  est la distance entre les plans réflecteurs,  $\theta$  l'angle entre les rayons incident et diffusé,  $\lambda$  la longueur d'onde et  $n$  l'ordre de diffraction ( $n \in \mathbb{N}$ ). Le *transfert de quantité de mouvement*, noté  $q$ , relie l'énergie du photon X et l'angle de diffusion. Il est défini comme :

$$q = \frac{E}{hc} \sin(\theta/2) \quad (1.2)$$

où  $E$  est l'énergie correspondante à la longueur d'onde  $\lambda$ ,  $h$  la constante de Planck et  $c$  la célérité de la lumière dans le vide. Ainsi, la diffraction a lieu lorsque  $q$  est un multiple entier de  $1/2d$ . Donc, en étudiant l'intensité de diffusion en fonction de  $q$ , on obtient une distribution  $F(q)$  avec des pics caractéristiques d'un cristal donné, ce qui permet son identification.

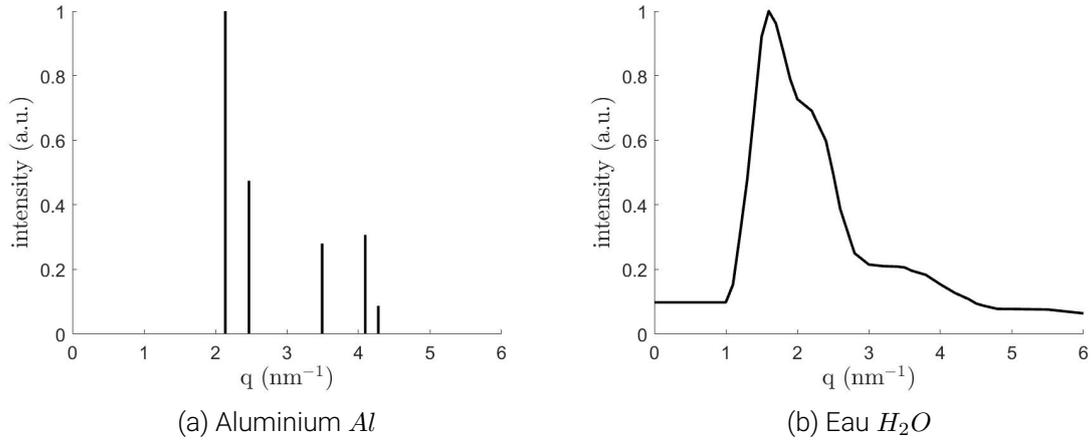


FIGURE 1.5 – Distribution normalisée de l'intensité de diffusion en fonction de  $q$  pour l'aluminium (cristallin) et pour l'eau (amorphe)

Lorsqu'on s'intéresse non plus aux cristallins, mais aux matériaux amorphes, la distribution des atomes est nettement moins régulière. Si on trace  $F(q)$  on n'obtient plus des pics, mais une fonction continue : même si la matière n'est plus parfaitement ordonnée, certaines distances privilégiées existent. Elles sont liées aux tailles caractéristiques des molécules constituant le matériau, et à leur agencement entre-elles. Un exemple de ces distributions pour le sel, qui est un cristal, et l'eau, qui est amorphe, est donné sur la figure 1.5.

Le nombre de photons diffusés dépend de la probabilité de diffusion, donnée par la section efficace différentielle de diffusion cohérente,  $d\sigma_{coh}/d\Omega$ , définie comme :

$$\frac{d\sigma_{coh}(\theta, q)}{d\Omega} = \frac{d\sigma_{Th}(\theta)}{d\Omega} f^2(q) \quad (1.3)$$

où  $d\sigma_{Th}/d\Omega$  et  $f^2$  représentent deux contributions distinctes à la diffusion cohérente [19]. Le premier terme, la diffusion Thomson, est la section efficace différentielle de diffusion cohérente d'un photon sur un électron libre. Elle est définie comme :

$$\frac{d\sigma_{Th}(\theta)}{d\Omega} = \frac{r_e^2}{2} (1 + \cos(\theta)^2) \quad (1.4)$$

où  $r_e$  est le rayon classique d'un électron. Le deuxième terme,  $f^2$ , est le facteur de forme atomique selon le modèle d'atomes indépendants. Il représente la contribution de l'agencement des atomes sur la diffusion cohérente. Dans le cas des amorphes, tels que les tissus biologiques qui font l'objet de cette thèse,  $f^2(q)$  est une fonction décroissante en  $q$  qui tend vers 0 à l'infini [19].

Dans ce travail de thèse, nous cherchons à évaluer la *fonction de diffusion*, notée  $F$ , et définie

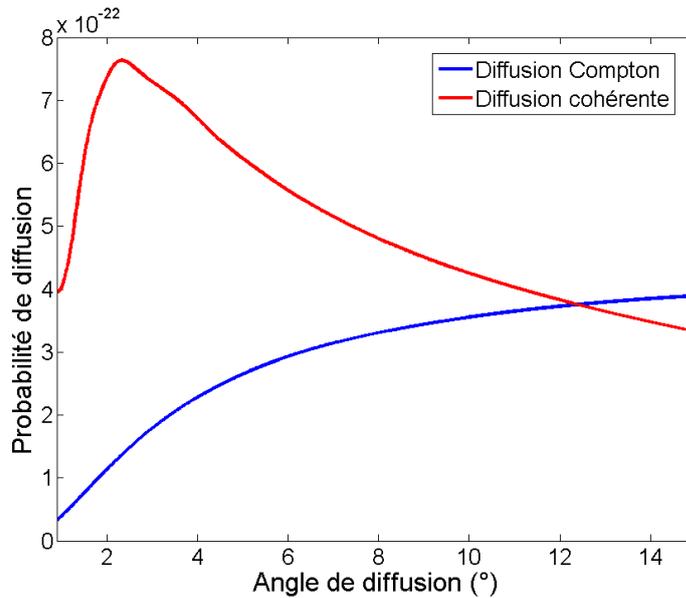


FIGURE 1.6 – Comparaison des probabilités de diffusion pour la diffusion cohérente et la diffusion incohérente (Compton) en fonction de l'angle de diffusion pour le cas de l'eau [18]

comme la section efficace de diffusion cohérente par unité de volume :

$$F(\theta, q) = N_{mol} \frac{d\sigma_{coh}(\theta, q)}{d\Omega} \quad (1.5)$$

où  $N_{mol}$  est le nombre de molécules par unité de volume.  $F$  s'exprime en  $\text{cm}^{-1} \cdot \text{str}^{-1}$ .

Cependant, lorsqu'on mesure les photons diffusés, on ne peut pas savoir s'ils proviennent de la diffusion cohérente ou incohérente. La diffusion incohérente, ou diffusion Compton, va nous gêner dans nos mesures. C'est pourquoi, nous cherchons à travailler dans des conditions qui minimisent sa probabilité. Damien BARBES a montré dans son travail de thèse, que pour des angles de diffusion faibles ( $< 8^\circ$ ), la diffusion Rayleigh prédomine sur la diffusion Compton [18]. Pour illustrer ce propos, la figure 1.6 représente les probabilités de diffusion pour la diffusion cohérente et la diffusion incohérente, sommées sur les énergies comprises entre 1 et 160 keV, pour le cas de l'eau. Les tissus biologiques que nous allons étudier dans cette thèse, sont en grande partie composés d'eau. C'est pourquoi, nous faisons l'hypothèse que l'effet Compton est négligeable, tant que l'angle de diffusion est inférieur à  $8^\circ$ .

### 1.2.3 Schéma général des systèmes XRD

Nous venons de voir la théorie physique qui supporte la XRD. Dans la sous-section qui suit, nous allons voir comment elle est mise en pratique, en présentant un schéma général des systèmes exploitant la XRD.

L'objectif d'un tel système est d'évaluer la fonction de diffusion  $F(q)$ . Cependant, la quantité de transfert de moment  $q$  n'est pas une grandeur mesurable. Comme nous l'avons vu dans

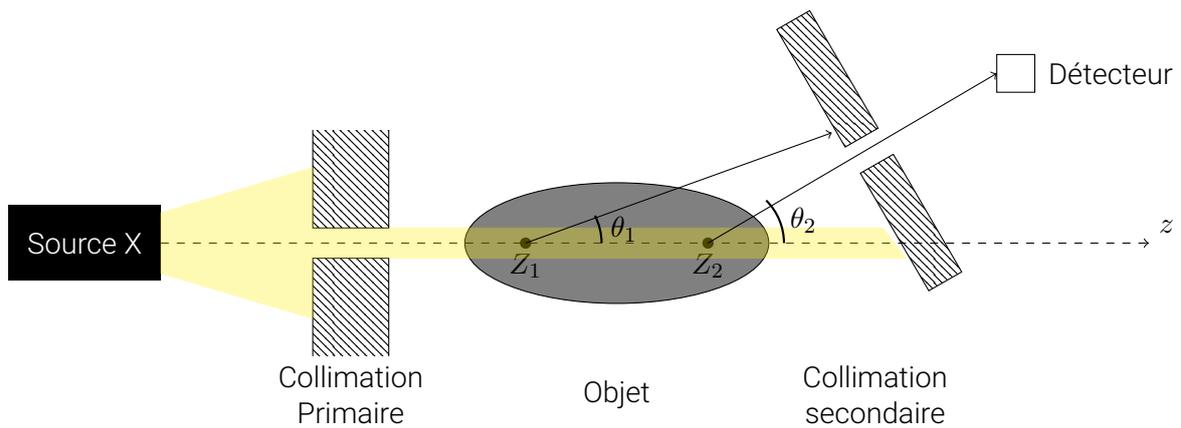


FIGURE 1.7 – Schéma général d'un système XRD

l'équation 1.2,  $q$  dépend de deux paramètres : l'angle de diffusion  $\theta$  et l'énergie des photons  $E$ . Ainsi les systèmes XRD vont mesurer une intensité de diffusion (nombre de photons) pour une énergie et un angle de diffusion donnés. On peut définir deux familles de géométries XRD, où l'un des paramètres varie et l'autre reste fixe :

- la diffusion cohérente de rayons X à dispersion d'angle (ADXRD, *Angle Dispersive X Ray Diffraction*) : les mesures sont effectuées avec une source d'énergie monochromatique pour des angles différents, soit en déplaçant le détecteur, soit en utilisant plusieurs détecteurs simultanément ;
- la diffusion cohérente de rayons X à dispersion d'énergie (EDXRD, *Energy Dispersive X Ray Diffraction*) : la source d'énergie fournit un spectre continu polychromatique et la mesure est effectuée pour un angle de diffusion constant, à l'aide d'un détecteur résolu en énergie.

Il existe également des systèmes qui combinent les deux modalités. Par exemple, certains d'entre eux utilisent une source polychromatique pour réaliser des observations dans différentes directions pour construire des sinogrammes, comme lors d'une tomographie classique. Dans tous les cas, un algorithme de traitement adapté à la géométrie sera utilisé pour estimer la signature du matériau à partir de l'intensité mesurée.

Dans tous les systèmes exploitant la XRD, quelle que soit la modalité choisie, on pourra retrouver les éléments suivants, illustrés sur la figure 1.7 :

### Une source de rayons X

Pour pouvoir mesurer un rayonnement diffusé, il est nécessaire d'éclairer l'objet avec une source de rayonnement X, puis de mesurer les photons diffusés.

La source va produire les photons pour éclairer l'objet. La source est caractérisée par son spectre : la variation de l'intensité, exprimée en nombre de photons par unité de temps, en fonction de l'énergie des photons. Le spectre peut être polychromatique ou monochroma-

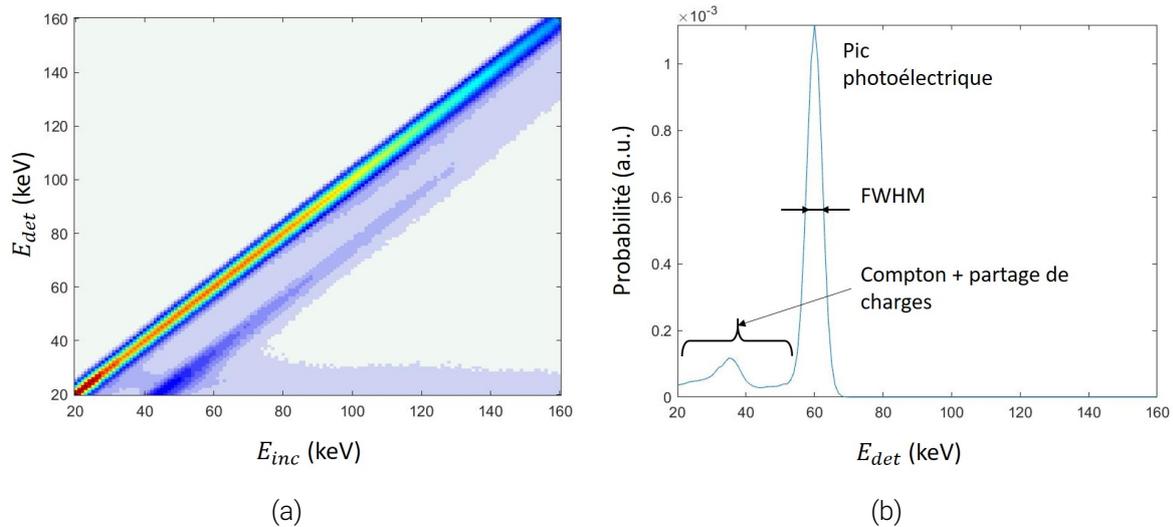


FIGURE 1.8 – (a) DRM du détecteur Imadif, utilisé dans cette thèse. (b) Colonne correspondant à  $E_{inc} = 60$  keV. Chaque colonne représente la la probabilité de détection d'un photon d'énergie  $E_{inc}$  avec l'énergie  $E_{det}$

tique.

Dans le milieu médical, les spectres sont polychromatiques et produits avec des tubes à rayons X. Dans ces tubes, une forte tension est appliquée entre une anode et une cathode afin d'accélérer les photons qui arrivent de la cathode et viennent bombarder une cible métallique fixée sur l'anode, ce qui va produire des rayons X. La forme du spectre va dépendre de la tension appliquée et du métal (souvent en tungstène ou molybdène) utilisé pour la cible.

Un rayonnement monochromatique peut être obtenu en utilisant des filtres après un tube conventionnel ou à partir d'un synchrotron.

### Un collimateur primaire

Le collimateur primaire bloque certaines directions du rayonnement à l'aide de matériaux fortement atténuants, comme le plomb ou le tungstène. Ainsi la collimation primaire détermine la portion de l'objet traversée par le faisceau X, et donc la partie de l'objet qui va donner lieu au phénomène de diffusion. Les collimations les plus fréquemment utilisées sont en pinceau (*pencil beam*), en éventail (*fan beam*) et en cône (*cone beam*).

### Un détecteur

De nombreux types et modèles de détecteurs à rayons X existent, et le choix est orienté par l'application recherchée. Les premiers détecteurs utilisaient les films photosensibles. Les rayons X déclenchaient une réaction chimique, dépendant du nombre de photons reçus, ce qui, à son tour, conduisait à la formation d'une image. Aujourd'hui sont plutôt utilisés les détecteurs numériques, qui convertissent les photons en signal électrique et permettent le comptage des photons et la mesure de leur énergie. Deux types peuvent être distingués : les détecteurs à

conversion directe ou indirecte :

- les détecteurs à conversion indirecte utilisent un intermédiaire, souvent un "scintillateur", pour convertir d'abord le rayonnement X en rayonnement visible, puis ce rayonnement est mesuré avec des photo-détecteurs. Ce type de détecteurs ne permet pas la mesure de l'énergie des photons incidents.
- les détecteurs à conversion directe utilisent le courant généré par l'absorption d'un photon X dans un matériau semi-conducteur. De nombreux matériaux peuvent être utilisés : le silicium, le germanium, le tellure de cadmium (CdTe) ou tellure de cadmium dopé au zinc (CdZnTe), etc. L'avantage des détecteurs CdTe et CdZnTe est d'être utilisable à température ambiante, et donc peu encombrants.

Par ailleurs, les détecteurs peuvent être caractérisés par leur nombre de pixels, la surface de détection, le flux de photons maximal supporté, ainsi que leur matrice de réponse du détecteur (DRM, *Detector Response Matrix*). La figure 1.8 montre la DRM du détecteur à base de CdZnTe "Imadif" [20], utilisé pendant cette thèse. Elle représente les "imperfections" du détecteur. Chaque colonne de la matrice donne la probabilité de détection d'un photon d'énergie  $E_{inc}$  avec une énergie  $E_{det}$ . Pour un détecteur idéal avec une résolution énergétique infinie, cette matrice vaudrait 1 sur la diagonale, où  $E_{inc} = E_{det}$ , et 0 ailleurs. Le comportement d'un détecteur réel est plus complexe, comme on peut l'observer sur la figure 1.8 (b). L'énergie détectée est majoritairement proche de  $E_{inc}$  : la distribution présente un pic centré sur  $E_{inc}$ , dont la largeur définit la résolution en énergie. On observe également des détections avec des énergies inférieures. Ceci est dû à différents phénomènes, comme le partage de charges dans le semi-conducteur, l'induction, etc..

### Une collimation secondaire

Si l'épaisseur de l'objet est négligeable, en une position donnée du détecteur le rayonnement diffusé va arriver d'une direction unique. On peut alors soit utiliser un détecteur spectrométrique pour balayer une gamme d'énergie (EDXRD), soit déplacer le détecteur pour étudier plusieurs directions de diffusion (ADXRD), pour finalement construire la fonction de diffusion en fonction de  $q$ . Sur la figure 1.7 est illustré l'autre cas où l'épaisseur de l'objet n'est pas négligeable. Sans collimation secondaire, le détecteur reçoit le rayonnement diffusé, arrivant de tous les points constituant l'objet et sous des angles différents. Il y a mélange de toutes ces informations et il serait alors impossible de discriminer la provenance des photons mesurés. C'est pourquoi une *collimation secondaire* peut être utilisée, comme montrée sur la figure 1.7. En fonction de sa géométrie, on peut savoir *a priori* quelles gammes de positions et d'angles sont visibles par le détecteur.

Une fois la mesure effectuée, un *algorithme de traitement* adapté à la géométrie sera utilisé pour calculer la fonction de diffusion.

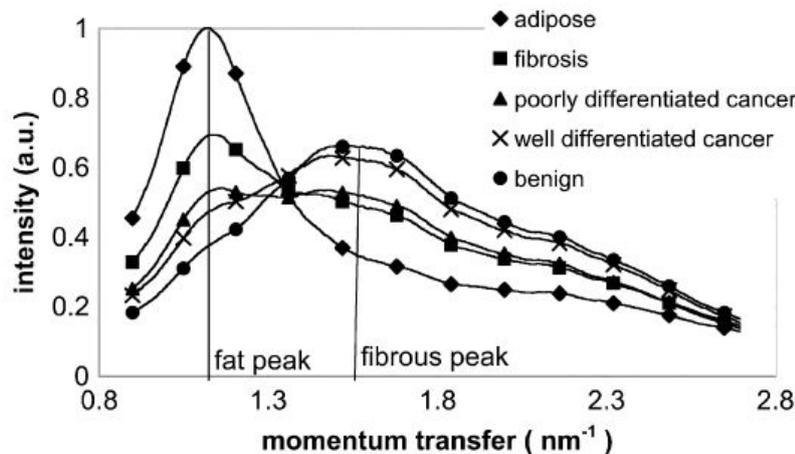


FIGURE 1.9 – Comparaison des fonctions de diffusion normalisées pour les tissus mammaires : adipeux, fibreux, cancer bien et peu différenciés, ainsi que pour les lésions bénignes [21]

#### 1.2.4 Applications envisagées

La XRD permet de différencier les tissus biologiques, et notamment les tissus mammaires. La figure 1.9 présente les fonctions de diffusion des tissus mammaires sains et pathologiques, obtenues par Silvia PANI et ses collègues [21]. Le tissu mammaire sain est composé de tissus adipeux et fibreux, respectivement représentés avec des marqueurs en forme de losange et de carré. Les cellules mammaires saines sont différenciées. Lorsqu'un cancer se développe, les cellules atteintes vont perdre progressivement leur structure et devenir non différenciées. Les fonctions de diffusion de cellules atteintes de cancer bien différenciées et peu différenciées sont représentées avec des marqueurs en forme de croix et de triangle respectivement. Les marqueurs ronds correspondent aux tumeurs bénignes, telles que le CCIS. On remarque que les fonctions de diffusion des tissus sains présentent un maximum bien marqué vers  $1.1 \text{ nm}^{-1}$  ("fat peak" sur la figure). Les fonctions de diffusion des tissus pathologiques sont plus plates, avec un maximum à  $1.6 \text{ nm}^{-1}$  ("fibrous peak" sur la figure). Ainsi, on observe bien une différence dans les fonctions de diffusion des tissus sains et pathologiques. Ci-dessous sont présentés deux champs d'application envisagés, qui exploitent cette différence pour caractériser les tissus mammaires.

##### Application ex vivo

Lorsqu'une tumeur est excisée par intervention chirurgicale, un contrôle de marges est nécessaire pour s'assurer que l'intégralité de la tumeur a bien été enlevée. Pour cela, le tissu prélevé est analysé par analyse anatomo-pathologique, ou à l'aide d'appareils d'imagerie dédiés, tels que BioVision (Faxitron Bioptics, LLC).

L'analyse anatomo-pathologique est très fiable, mais peut prendre du temps (jusqu'à une heure), ce qui allonge considérablement le temps de l'opération. À l'inverse, les appareils d'imagerie

dédiés sont très rapides. Ils utilisent l'atténuation des rayons X pour former l'image et ont une très haute résolution spatiale. Cependant, cette technique ne reflète que la densité des tissus, et non pas la structure moléculaire.

L'intégration de la XRD dans les appareils dédiés à l'imagerie des échantillons excisés peut apporter l'information sur la structure moléculaire, qui manque pour le moment. Pour cette application, le système XRD devra :

- être fiable;
- donner une bonne résolution, de l'ordre du millimètre, pour qu'il soit possible d'effectuer correctement un contrôle de marges;
- être rapide, afin de fournir un diagnostic en maximum quelques minutes;
- être compact, afin de pouvoir être placé dans une salle d'opération.

### **Application *in vivo***

Nous envisageons également une utilisation *in vivo* de la XRD, en deuxième niveau, dans le cas où des examens supplémentaires seraient nécessaires après la mammographie. La caractérisation moléculaire par XRD peut infirmer la présence d'une tumeur maligne et réduire ainsi le nombre de biopsies. Comme pour l'application *ex vivo*, les principales contraintes seraient le temps et la précision, auxquelles vient s'ajouter la contrainte de la dose.

### **1.2.5 Conclusion**

Dans cette section, nous avons vu que la XRD est une technique de caractérisation moléculaire des matériaux et des tissus, qui exploite le rayonnement X diffusé. Nous avons présenté le schéma général des systèmes, qui utilisent cette technique et avons donné deux champs d'application potentiels pour le diagnostic du cancer du sein. Dans la section suivante, nous décrivons les différents systèmes expérimentaux proposés dans la littérature.

## **1.3 Systèmes XRD pour la caractérisation des tissus mammaires**

Un certain nombre de systèmes exploitant la XRD a été proposé récemment dans la littérature. Dans cette partie, nous présentons quelques-uns des systèmes proposés par différentes équipes de recherche. Nous concluons avec une comparaison de ces systèmes.

### **1.3.1 Systèmes XRD proposés dans la littérature**

Dans cette partie nous allons présenter les travaux récents sur l'exploitation de la XRD pour la caractérisation des tissus mammaires. La XRD est également présente dans les domaines de la sécurité et du contrôle non destructif, et les équipes qui s'intéressent à la XRD travaillent souvent en parallèle sur toutes ces applications. Cependant, l'application médicale impose des contraintes, qui lui sont spécifiques :

- la dose absorbée et, par conséquent, le temps d'acquisition suffisamment faibles, afin que les effets néfastes de la radiation ne dépassent les bénéfices d'un diagnostic amélioré;
- la géométrie et la nature des objets à imager : les tissus mammaires sont hétérogènes et ont une épaisseur non négligeable, notamment dans la perspective d'une utilisation clinique *in vivo*;
- la précision du diagnostic en termes de sensibilité et de spécificité : un système moins précis que les modalités d'imagerie existantes aurait peu d'intérêt.

C'est pourquoi, dans cette partie ne sont présentées que les configurations pensées pour l'imagerie mammaire :

- la tomographie par diffusion, dont la géométrie est assez similaire à celle de la tomographie classique;
- la géométrie *Pixellated x ray Diffraction* (PixD) proposée en 2013 par l'équipe de D. O'FLYNN et R. D. SPELLER de l'University College London (UCL) [22];
- la géométrie *Coded Aperture Coherent Scatter Spectral Imaging* (CACSSI), utilisant les détecteurs spectrométriques et les masques codés comme collimation secondaire, proposée par l'équipe de Duke University [23].

### Tomographie par diffusion

Une géométrie de systèmes XRD similaire à la tomographie a été proposée pour la première fois en 1985 par HARDING et KOSANETZKY [24]. Par la suite, d'autres chercheurs ont également travaillé sur des géométries semblables, qu'on trouve dans la littérature sous les noms *Coherent Scatter Computed Tomography* ou *X-ray Diffraction Tomography* (XDT).

La géométrie proposée en 1985 est présentée sur la figure 1.10. La configuration est très similaire à la tomographie en transmission : l'objet est tourné et / ou translaté afin que le faisceau monochromatique incident le traverse avec des angles de vue différents. Cependant, contrairement à la tomographie, l'objet est éclairé avec un faisceau très fin qui ne traverse qu'un petit volume.

Le signal diffusé mesuré avec différentes prises de vue va former un sinogramme, qui peut être reconstruit avec des algorithmes adaptés tels que, par exemple, la rétroprojection filtrée.

En 2010, PANI et ses collègues ont proposé un système CSCT à dispersion d'énergie, qu'ils ont testé avec succès sur des échantillons de tissu mammaire [21]. La nature des tissus présents dans ces échantillons a été déterminée par analyse anatomopathologique. L'image 1.11 montre les résultats obtenus sur un de ces échantillons : les images (a) et (b) montrent les intensités de diffusion obtenues respectivement pour  $q=1.1 \text{ nm}^{-1}$  et  $q=1.6 \text{ nm}^{-1}$ . L'image (c) montre la coupe histologique, prise pour référence. Comme attendu, le tissu adipeux diffuse plus à  $1.1 \text{ nm}^{-1}$ , et la fibrose à  $1.6 \text{ nm}^{-1}$ .

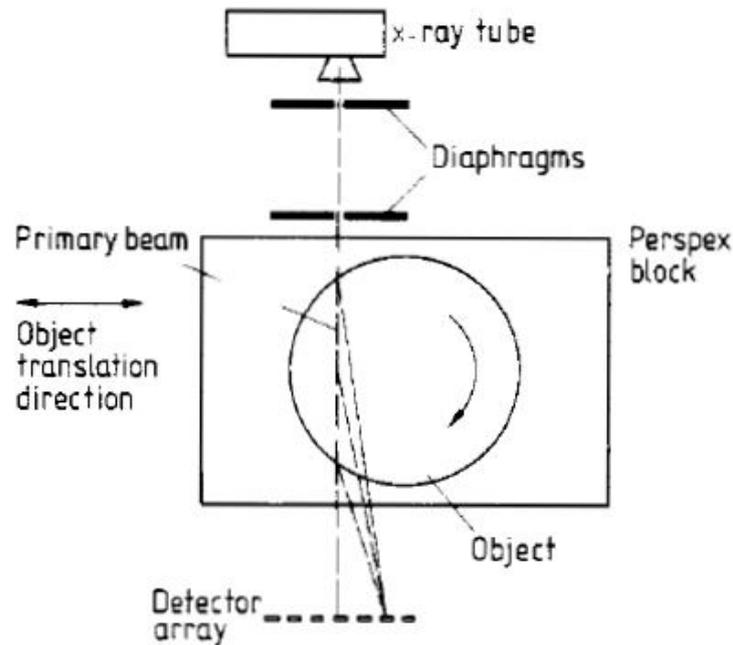


FIGURE 1.10 – Première configuration *Coherent Scatter Computed Tomography* (CSCT) proposée par Harding et Kosanetzky en 1985 [24]

Cependant, le processus est extrêmement long : les échantillons étaient imagés sur  $360^\circ$  par pas angulaires de  $1.8^\circ$  et 14 pas en translation. Le temps d'acquisition par position était de 4 secondes, ce qui correspond à plus de 3 heures d'acquisition au total.

Une équipe de chercheurs de University of Central Florida travaille également sur la tomographie par diffusion. En 2018, ils ont présenté une méthode qui permet de réduire le temps d'acquisition [25]. Pour cela, ils proposent de ne pas étudier tout le fantôme, mais une zone d'intérêt plus restreinte. En effet, l'analyse par diffusion est positionnée en deuxième niveau après une mammographie : en application clinique, il existerait donc une zone suspecte que l'on souhaite étudier plus en détail sans imager tout le sein. La figure 1.12 montre l'échantillon utilisé dans cette étude et les résultats obtenus. L'échantillon est un morceau de  $23 \text{ mm} \times 23 \text{ mm}$  de viande porcine et la zone d'intérêt a un diamètre de 5 mm. La zone d'intérêt est présentée en couleur sur la figure 1.12 (a). À droite de la figure, deux disques représentent la fonction de diffusion  $F(q)$ , correspondants à cette zone, pour  $q = 0.08 \text{ \AA}^{-1}$  et  $q = 0.16 \text{ \AA}^{-1}$ . L'acquisition sur la zone d'intérêt réduite a pris 2.3 heures, alors que, d'après leurs estimations, imager l'objet en entier prendrait 13.4 heures. Il y a bien une réduction considérable de temps et de la dose (83% d'après l'article), mais pour le moment, cela reste excessivement long pour une application clinique.

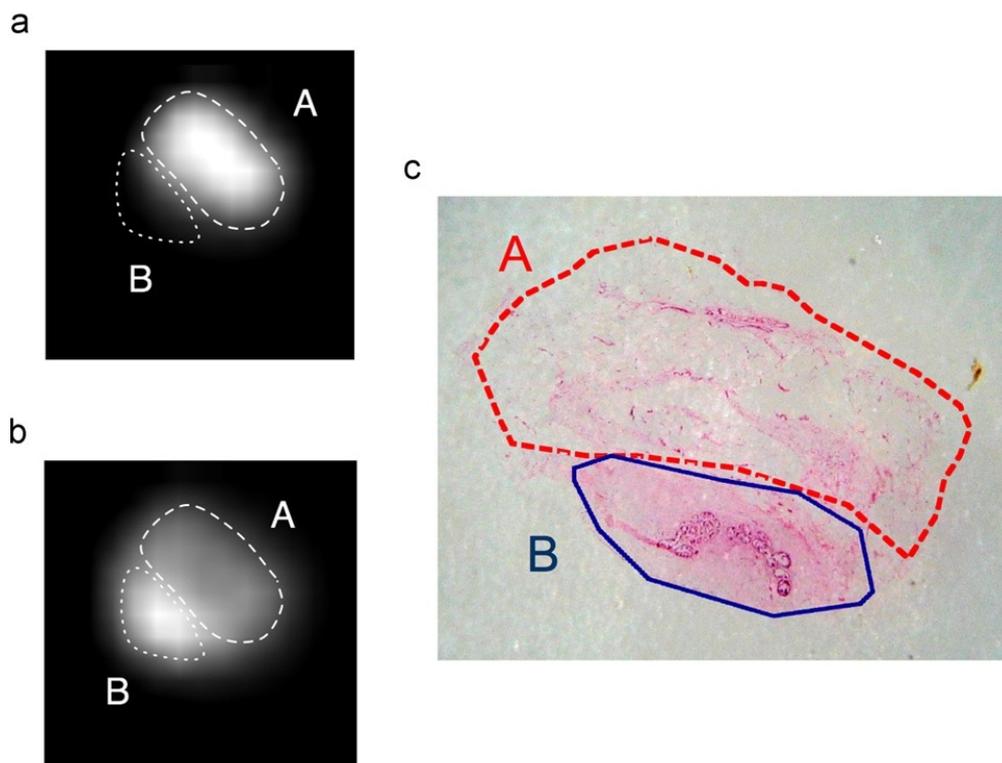


FIGURE 1.11 – Résultats obtenus par PANI et al. en 2010 [21]. Sur la photo (c) est représentée la coupe, ayant subi l'analyse anatomopathologique : le trait pointillé en rouge (A) délimite le tissu adipeux, et le trait continu en bleu (B) délimite la fibrose. Les images (a) et (b) montrent les intensités de diffusion reconstruites, respectivement à  $q=1.1 \text{ nm}^{-1}$  (a) et  $q=1.6 \text{ nm}^{-1}$  (b).

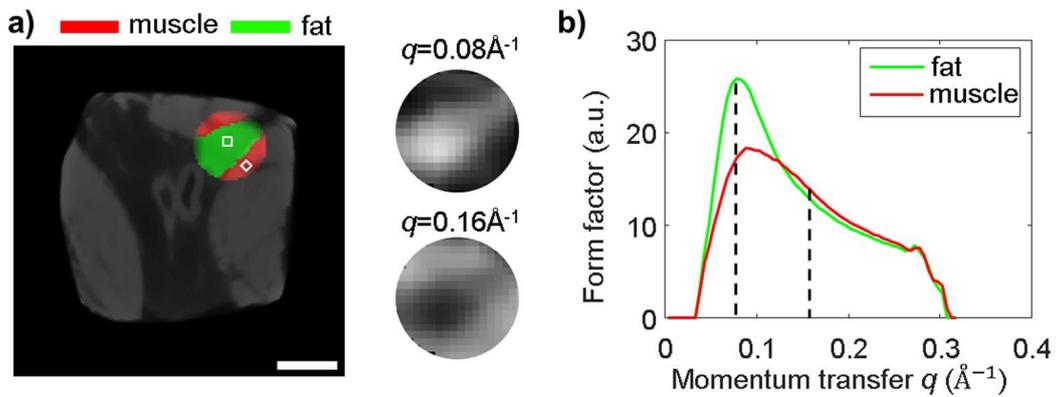


FIGURE 1.12 – (a) Reconstruction en CSCT de la zone en couleur superposée sur l’image de l’échantillon obtenue par tomographie classique. L’échantillon est un morceau de 23 mm × 23 mm de viande porcine. La zone d’intérêt a un diamètre de 5 mm. (b) Facteurs de forme estimés pour le tissu adipeux et le muscle à l’intérieur de la zone d’intérêt [25]

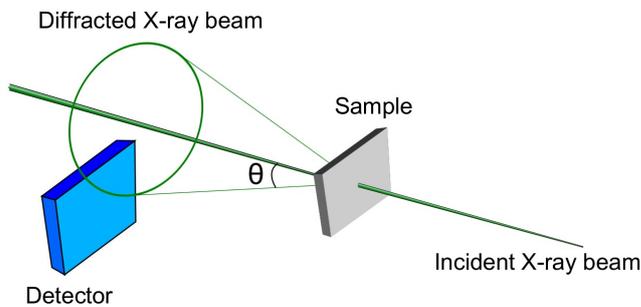


FIGURE 1.13 – Architecture du système PixD [22]

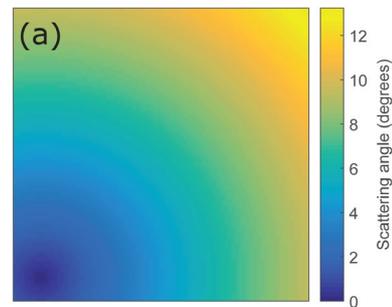


FIGURE 1.14 – Distribution des angles de diffusion sur la surface du détecteur [26].

### Pixellated X-ray diffraction

La géométrie PixD, illustrée sur la figure 1.13, a été initialement développée pour la détection des drogues, des explosifs et des contrefaçons de médicaments [22, 27, 26, 28]. Ensuite, elle a été utilisée en 2017 dans une étude d’imagerie médicale [29].

Dans cette géométrie, l’objet est éclairé par un faisceau incident collimaté en pinceau. Un détecteur spectrométrique plan HEXITEC™ (Quantum Detectors Ltd.) est placé après l’objet dans le plan perpendiculaire au faisceau incident, sans l’intercepter. Le rayonnement diffusé à un angle  $\theta$  constant définit un cône. L’intersection de ce cône avec le détecteur est un arc de cercle centré sur l’axe du faisceau incident : la figure 1.14 illustre la variation de l’angle de diffusion sur la surface du détecteur. Ainsi, tous les pixels se trouvant sur un arc de cercle centré autour du faisceau incident, vont mesurer les photons correspondant à un même angle de diffusion. La somme des mesures des pixels correspondant à un  $\theta$  constant permet de réduire le bruit de mesure.

En 2017 [29], cette géométrie a été utilisée pour mesurer les spectres diffusés sur huit échan-

tillons de tissu mammaire issus de tumorectomie et provenant de 4 individus différents, appelés ci-après "pièces". Les pièces mesurent environ 30 mm × 25 mm et ont une épaisseur allant jusqu'à 5 mm. Après opération, les pièces ont été fixées au formol et conditionnées en pochettes plastiques scellées pour préserver les tissus. Deux des échantillons sont montrés sur les figures 1.15 (a) et (b).

Les mesures XRD ont été obtenues en balayant toute la surface des pièces, comme montré sur les figures 1.15 (c) et (d) : chaque point coloré correspond à une mesure XRD. Le temps d'acquisition par point était de 300 secondes. Une analyse en composantes principales (PCA, *Principal Component Analysis*) [30] a été réalisée sur ces mesures. Elle a montré que les deux premières composantes principales, notées  $PC_1$  et  $PC_2$ , suffisent pour décrire les mesures. Le code couleur est défini dans l'espace teinte saturation valeur (HSV, *Hue Saturation Value*), où la teinte et la saturation sont fonction de  $PC_1$  et  $PC_2$ .

L'analyse anatomopathologique est réalisée sur une coupe fine (4  $\mu\text{m}$ ) de chaque pièce (voir figures 1.15 (e) et (f)) pour obtenir une classification de référence en tissus mammaires normaux, adipeux et cancéreux.

La répartition des mesures dans le plan ( $PC_1, PC_2$ ) est illustrée sur la figure 1.16. Le tiret rouge délimite l'ensemble des mesures qui correspondent aux points classifiés comme cancéreux par l'analyse anatomopathologique. On constate qu'il y a bien corrélation entre les résultats obtenus par analyse anatomopathologiques et la projection des mesures XRD dans le plan ( $PC_1, PC_2$ ) : les points correspondant au tissu cancéreux se retrouvent tous dans un cône de cet espace.

Ces résultats sont très intéressants, car il s'agit de la première fois où la corrélation entre la XRD et l'analyse anatomopathologie a été démontrée sur des pièces mammaires. Malheureusement, le temps d'acquisition (300 secondes par point) reste excessivement long, ce qui ne permettrait pas une utilisation clinique en l'état.

### **Coded Aperture Coherent Scatter Spectral Imaging**

L'équipe de Duke University étudie des systèmes XRD où un masque codé (CA, *Coded Aperture*) [31, 33] est utilisé comme collimateur secondaire, ce qui permet d'estimer la fonction de diffusion de l'objet à différentes positions. De manière générale, un CA est un objet plan comportant des parties opaques et transparentes qui forment un motif. Le principe de son utilisation est illustré sur la figure 1.17. Le masque codé, placé après l'objet, engendre un multiplexage du signal. Le motif du masque est dimensionné pour faciliter la déconvolution du signal mesuré, et un processus de reconstruction adapté permet de remonter au signal démultiplexé.

Une telle collimation permet, d'une part, d'encoder la géométrie de l'objet : l'onde électromagnétique provenant de l'objet va passer à travers le masque codé et engendrer des ombres portées sur le détecteur. Ces ombres reproduisent le motif du masque avec un agrandissement qui dépend de la distance entre l'objet et le masque. Ainsi, connaissant le motif, on peut remonter à

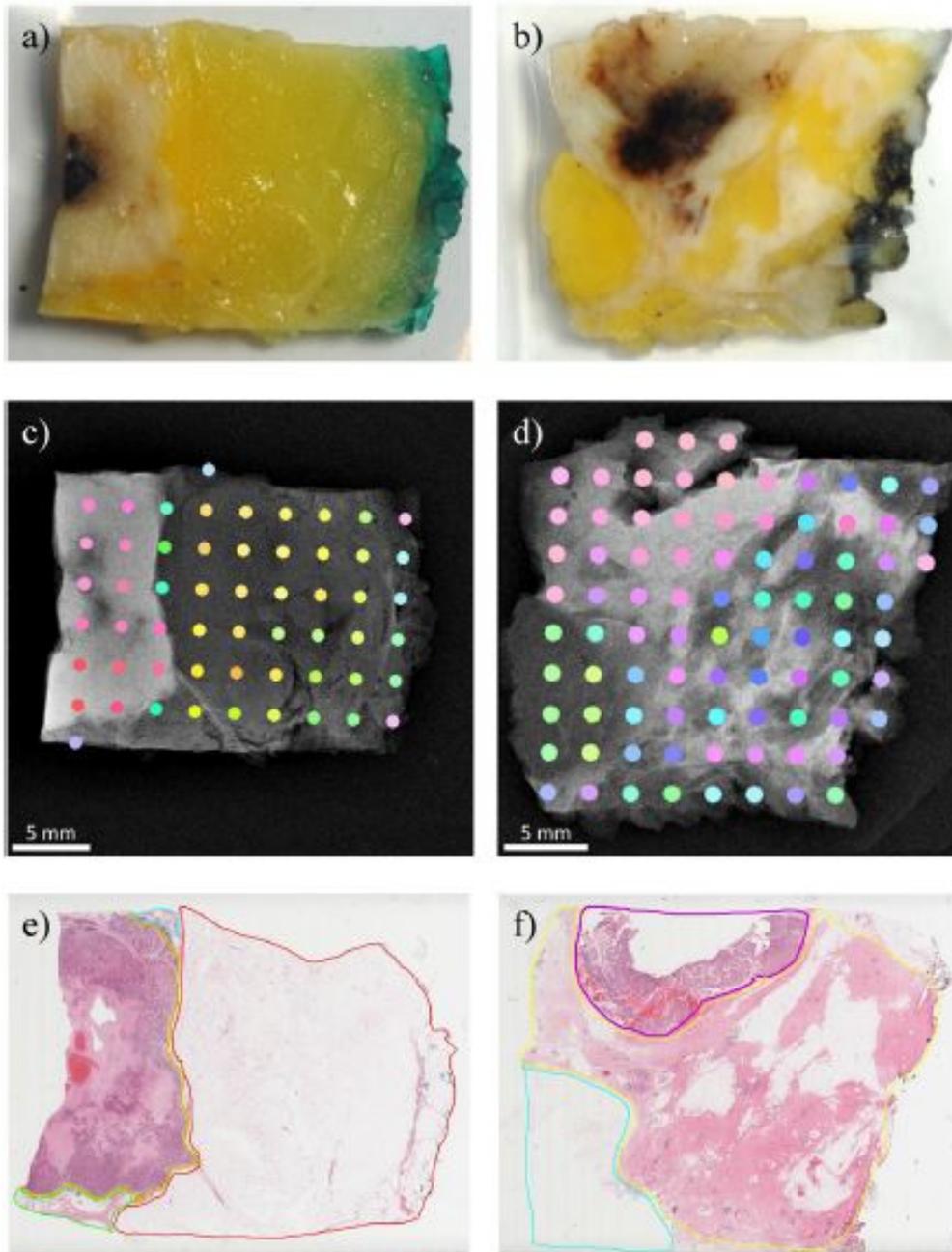


FIGURE 1.15 – Exemple de deux échantillons analysés [29]. (a) et (b) sont les photographies des échantillons préparés. (c) et (d) sont les mammographies haute résolution, les différents points de couleur correspondant aux points de mesure XRD. La couleur dépend des coefficients obtenus par PCA. (e) et (f) montrent les images obtenues par microscopie numérique sur une couche de chaque échantillon. Les limites correspondent, respectivement, au tissu mammaire normal (bleu / vert), au tissu adipeux (jaune) et au tissu cancéreux (rouge).[29]

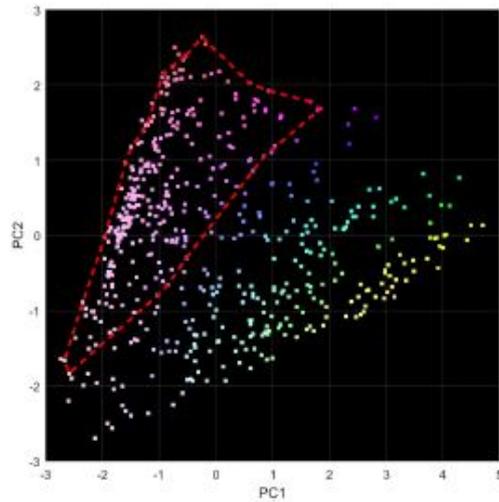


FIGURE 1.16 – Distribution des mesures dans l'espace  $(PC_1, PC_2)$ . Le code couleur des points est fonction de  $(PC_1, PC_2)$ . Le trait en pointillés rouge délimite l'ensemble des points classifiés comme cancéreux par l'analyse anatomopathologique.[29]

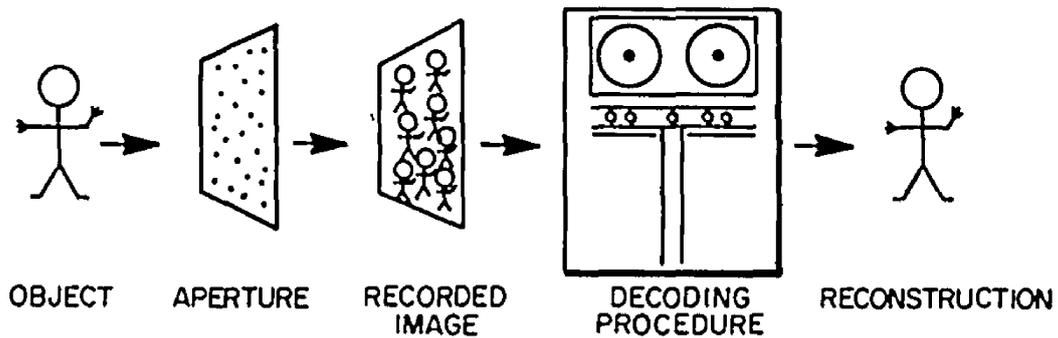


FIGURE 1.17 – Illustration de la chaîne d'acquisition et de traitement du signal à l'aide d'un masque codé (CA, *Coded Aperture*) [31]

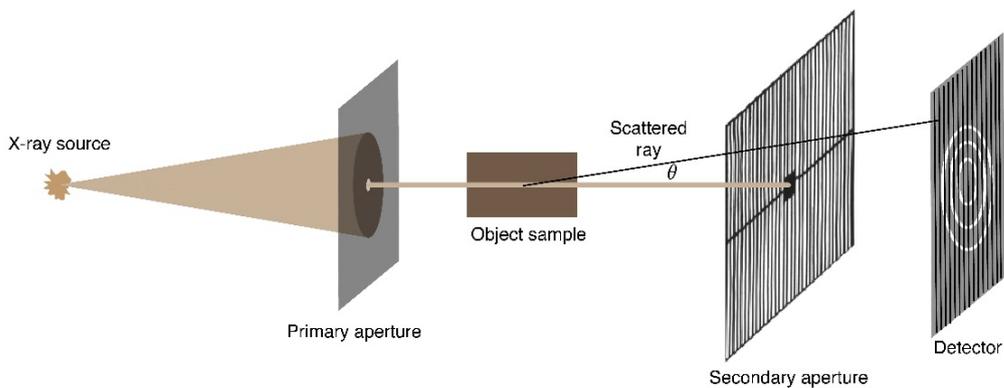


FIGURE 1.18 – Architecture du système *Coded Aperture X ray Scatter Imaging* (CAXSI) [23]

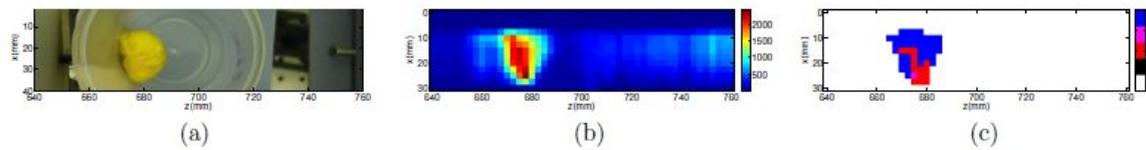


FIGURE 1.19 – Images obtenues à partir d'un échantillon de tissu mammaire humain : (a) photographie de l'échantillon ; (b) une coupe de l'image en intensité de diffusion ; (c) classification selon le type de tissu : normal (rouge), fibroglandulaire (magenta), adipeux (bleu) [32]

la position de l'objet. D'autre part, comparée à une collimation à diaphragme (*pinhole*), elle est plus ouverte et laisse passer plus de photons, et permet donc d'obtenir un meilleur rapport de signal à bruit.

La figure 1.18 représente le dispositif expérimental proposé par l'équipe de Duke University en 2012. L'objet est illuminé avec un faisceau de rayons X, issu d'un générateur conventionnel avec une anode de tungstène, collimaté en pinceau. Le CA, placé après l'objet, a une forme très simple : il s'agit d'une fonction porte périodique, tel un peigne, avec, au milieu du masque, un cache qui arrête le faisceau direct (*beamstop*). L'image produite est mesurée sur un détecteur à scintillation plan. Comme ce type de détecteur ne permet pas de mesurer l'énergie des photons, plusieurs filtres passe-bande ont été utilisés pour restreindre la répartition énergétique du faisceau. Il s'agit donc d'une architecture plutôt de type ADXRD.

Pour la reconstruction, un algorithme itératif a été développé [34]. Il repose sur la méthode d'estimation du maximum a posteriori (MAP) et permet d'estimer la fonction de diffusion de l'échantillon en fonction du transfert de quantité de mouvement et de la profondeur.

En 2015, cette collimation a été validée pour des échantillons de tissu mammaire humains [32]. La géométrie n'a pas changé, à part le détecteur à scintillation plan, qui a été remplacé par un détecteur spectrométrique ME-100 (MultiX™). Ce détecteur se présente sous la forme d'une ligne de 128 pixels. Comme ce détecteur est résolu en énergie, toute la gamme d'énergies émises par le tube peut donc être utilisée.

Une acquisition permet d'obtenir une fonction de diffusion en fonction de  $q$  et  $z$ , position selon la direction du faisceau. Dans l'expérience décrite, l'objet a été déplacé par pas réguliers dans le plan perpendiculaire au faisceau, pour couvrir les deux autres directions de l'espace. Ainsi, la fonction de diffusion est estimée en fonction des trois dimensions de l'espace et de  $q$ . On peut parler de reconstruction tomographique (par coupe), c'est pourquoi la méthode est présentée comme *Coherent Scatter Computed Tomography*.

Dans la suite du processus de traitement, pour chaque voxel, le tissu est classifié en comparant la fonction de diffusion obtenue à une base de données [35]. Les résultats obtenus à partir d'un des échantillons sont représentés sur la figure 1.19. L'échantillon était identifié comme non pathologique avant l'expérience, et on observe bien l'absence de tissu cancéreux, à l'issue de la classification des données obtenues par XRD. Cependant, ces résultats n'ont pas été confrontés aux résultats de l'analyse anatomopathologique : outre l'absence de tissu tu-

moral, la composition exacte de l'échantillon n'est pas présentée. Il est donc difficile d'évaluer la précision de la méthode.

### 1.3.2 Systèmes développés au CEA

Nous venons de voir les principales géométries proposées dans la littérature : la tomographie par diffusion, la géométrie PixD et la géométrie utilisant les masques codés. Chacune de géométries présente des inconvénients : la tomographie nécessite un temps d'exposition extrêmement long, et la géométrie PixD ne permet pas d'analyser des échantillons d'épaisseur non négligeable. Les résultats de Duke University sont les plus intéressants, cependant le détecteur spectrométrique utilisé, le ME-100, est plus adapté aux applications à fort flux de photons, comme l'imagerie en transmission.

Le laboratoire, où s'est déroulée cette thèse, est spécialiste des détecteurs spectrométriques pour des applications en imagerie X et gamma, et du développement de systèmes utilisant ces détecteurs. Notamment, le détecteur ME-100 y a été développé pour MultiX™. Par ailleurs, le laboratoire a également développé des détecteurs spectrométriques pour des applications à faible flux de photons, telles que la diffusion cohérente de rayons X. L'utilisation de ces détecteur dans les systèmes d'imagerie mammaire exploitant la XRD pourrait être très intéressante et donner de meilleurs résultats. Le développement d'un système XRD pour le diagnostic du cancer du sein a fait l'objet de deux thèses :

- "Nouveaux systèmes d'imagerie médicale exploitant la diffraction X en dispersion d'énergie à l'aide de détecteurs spectrométriques CdZnTe", soutenue par Damien BARBES en octobre 2016 [18];
- "Optimization of an X-ray diffraction imaging system for medical and security applications", soutenue par Fanny MARTICKE en juillet 2016 [36].

Chronologiquement, Damien BARBES a proposé une méthode de reconstruction des facteurs de forme à partir des spectres mesurés en utilisation un matériau de référence. Il l'a validée expérimentalement en utilisant une géométrie très simple. Suite aux premiers résultats encourageants, Fanny MARTICKE a réalisé en simulation une étude de dimensionnement afin de proposer une géométrie optimale pour le diagnostic du cancer du sein. La géométrie idéale ainsi obtenue n'était pas fabricable en l'état. Par conséquent, pour la suite des études expérimentales, Damien BARBES a adapté le design du collimateur idéal afin de fabriquer un collimateur réel, optimisé pour le détecteur "Imadif" [20] fabriqué au sein de notre laboratoire. Dans cette partie sont présentés les résultats-clés de ces deux thèses.

#### **Méthode du matériau de référence et validation expérimentale : D. BARBES**

Dans le contexte d'imagerie XRD médicale, Damien BARBES a proposé une méthode de reconstruction dite "par matériau de référence". Il a mis en place une expérience avec une géométrie

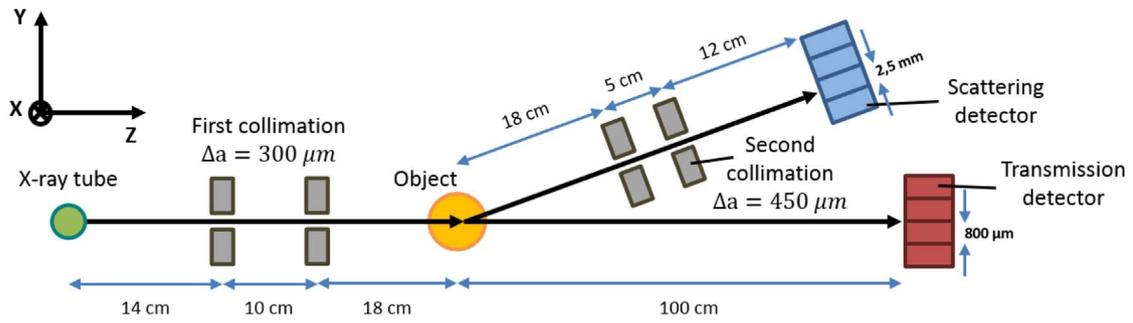
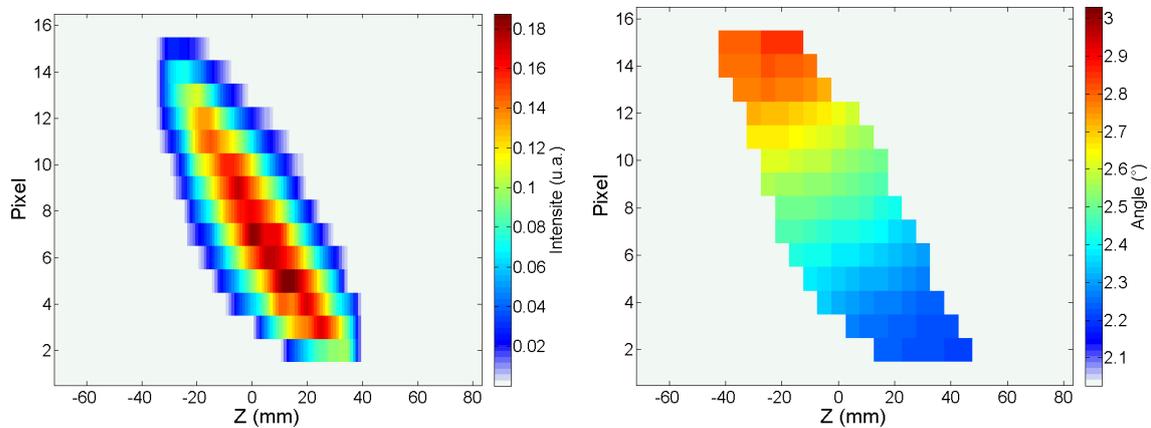


FIGURE 1.20 – Schéma du système utilisé : les photons incidents collimatés en pinceau illuminent l'objet. Le rayonnement diffusé est mesuré avec un angle constant de  $2.5^\circ$  par un détecteur spectrométrique CdZnTe. Le rayonnement atténué est mesuré par un deuxième détecteur.



(a) Distribution de l'intensité de diffusion  $A(p, z)$  (b) Distribution des angles de diffusion  $\Theta(p, z)$

FIGURE 1.21 – Cartes de champ de vue obtenues par calibration expérimentale

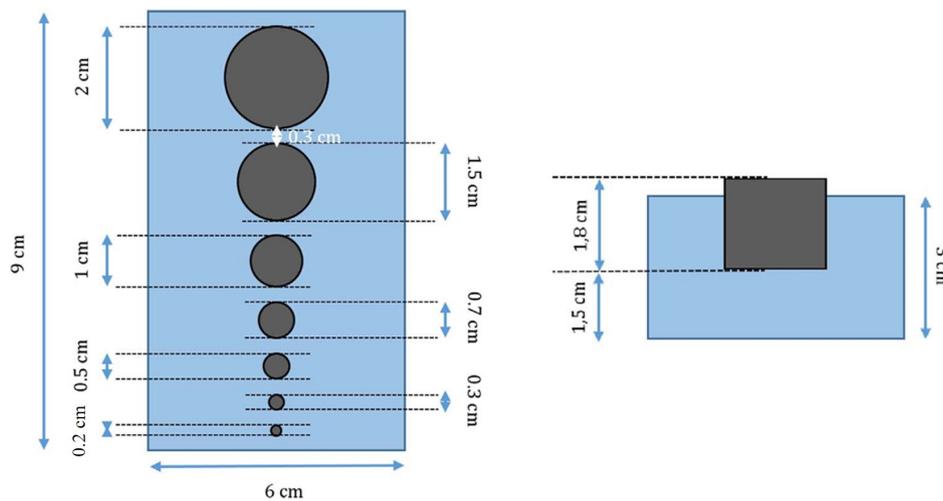


FIGURE 1.22 – Schémas du fantôme dans les plans  $(z, x)$  et  $(z, y)$ , respectivement à gauche et à droite (relativement au repère du schéma 1.20) : le support (en bleu) est fabriqué en Polystyrène (PS) et les inserts (en gris) sont en nylon

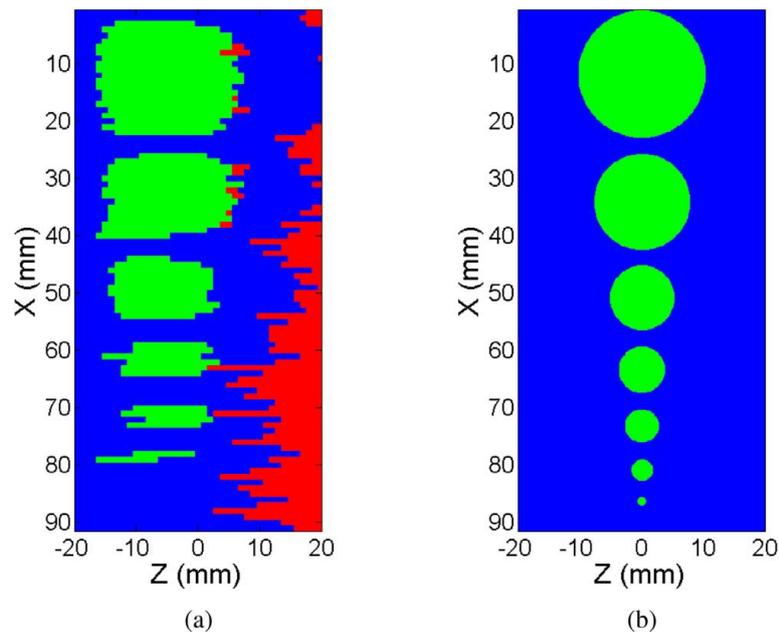


FIGURE 1.23 – Résultats de reconstruction (a) et l'objet réel (b). Les couleurs représentent les matériaux : PS (bleu), nylon (vert) et *Polymethyl Methacrylate* (PMMA) (rouge).

simple et un fantôme constitué de plastiques (Polystyrène (PS) et nylon), choisis pour la ressemblance de leur facteurs de forme à ceux des tissus mammaires [37].

La géométrie utilisée est représentée schématiquement sur la figure 1.20. Un jeu de double fentes est utilisé pour collimater le faisceau de rayons X en pinceau. Un deuxième jeu est utilisé comme collimation secondaire pour ne mesurer que les photons diffusés avec un angle de  $2.5^\circ$ . Deux détecteurs spectrométriques ont été utilisés, respectivement pour le rayonnement direct et le rayonnement diffusé.

Le rayonnement direct est mesuré avec un détecteur spectrométrique CdZnTe composé d'une ligne de 16 pixels avec une taille de  $800 \mu\text{m}$  [38]. Ce détecteur a été développé au laboratoire d'accueil pour des applications à fort flux. Il s'agit d'une version antérieure du détecteur ME-100™ (MultiX, France). Le rayonnement diffusé a été mesuré avec un autre détecteur CdZnTe, adapté au faible flux. Il comporte  $4 \times 4$  pixels sur une surface de détection de  $2.5 \text{ mm} \times 2.5 \text{ mm}$ . La collimation secondaire a été réglée de sorte à ce qu'un seul pixel reçoive les photons diffusés. Ce pixel a été virtuellement divisé en 16 sous-pixels en utilisant la technique de sur-pixellisation [39, 40].

En utilisant ce système expérimental, des mesures de rayonnement diffusé et atténué ont été réalisées avec le fantôme représenté sur la figure 1.22. Il est composé d'un support en PS avec des inserts cylindriques en nylon de différentes tailles, allant de 2 cm à 1 mm. Le système permet d'obtenir seulement une ligne (selon  $z$ ) en une acquisition, donc le fantôme a été déplacé selon  $x$  par pas de 1 mm pour imagier l'ensemble du fantôme dans le plan  $(x, z)$ . Au total, il a eu 90 acquisitions de 5 minutes. Le tube opérait à 160 kV et 11.25 mA.

Afin de traiter les données et reconstruire la composition en matériaux en fonction de la position selon  $z$ , il est nécessaire d'établir une carte de correspondance entre les pixels, les angles de diffusion et les positions le long de l'axe  $z$ . Pour cela, une calibration est effectuée avec une plaque d'aluminium. Ainsi sont obtenues les cartes de champ de vue, qui détaillent l'intensité (par rapport au nombre de photons) et l'angle de diffusion en fonction des pixels  $p$  et des positions  $z$ . Les cartes obtenues sont présentées sur la figure 1.21.

Afin de calculer les fonctions de diffusion à partir des spectres diffusés et atténués mesurés, un matériau de référence, ici le PMMA, a été utilisé. Des acquisitions de spectres diffusés et de spectres atténués sont réalisées avec l'objet de référence, exactement dans les mêmes conditions que le fantôme. Sa fonction de diffusion est connue. Un ratio entre les mesures du fantôme et de l'objet de référence permettent de simplifier la reconstruction et obtenir la fonction de diffusion  $F(z, q)$  du fantôme en fonction de  $q$  et de la position  $z$ .

Finalement, un matériau est attribué à chaque position  $z$  du fantôme, en comparant  $F(z, q)$  aux fonctions de diffusion de référence, obtenues dans la littérature [41, 21, 42]. La méthode des moindres carrés a été utilisée pour trouver une fonction de diffusion de référence qui ressemble le plus à celui qui a été calculé précédemment, et le matériau correspondant est attribué à la position  $z$ .

Les spectres acquis ont été traités et le résultat obtenu est représenté sur la figure 1.23 (a), en comparaison avec la vérité terrain (b). Le PS est représenté en bleu, le nylon - en vert, et le matériau de référence (PMMA) - en rouge. On constate que tous les inserts, sauf le plus petit, sont détectés. Pour la majorité des positions, le bon matériau est attribué. Par contre, il y a une zone entre 10 et 20 mm selon  $z$ , où le PMMA apparaît, alors qu'il n'y en a pas dans l'objet réel. Cet effet de bord était lié au bruit, amplifié par la division par le spectre du matériau de référence.

Globalement, les résultats de cette expérience étaient très encourageants. Plusieurs pistes d'amélioration du système étaient envisagées :

- Amélioration du processus de classification;
- Dans la configuration utilisée un seul pixel du détecteur était utilisé. La conception d'un collimateur secondaire adapté aux dimensions du détecteur permettrait d'utiliser en intégralité la surface de détection.
- Une étude de la dose absorbée est nécessaire pour les applications médicales, et particulièrement, les applications *in vivo*.

### **Géométrie idéale : thèse de F. MARTICKE**

En parallèle de l'étude décrite précédemment, Fanny MARTICKE a mené un travail de dimensionnement d'un collimateur secondaire idéal. Les résultats de ce étude ont fait l'objet d'un article détaillant la géométrie et les résultats obtenus en simulation [43]. L'objectif était de proposer un système pour une utilisation *in vivo* en second niveau après une mammographie,

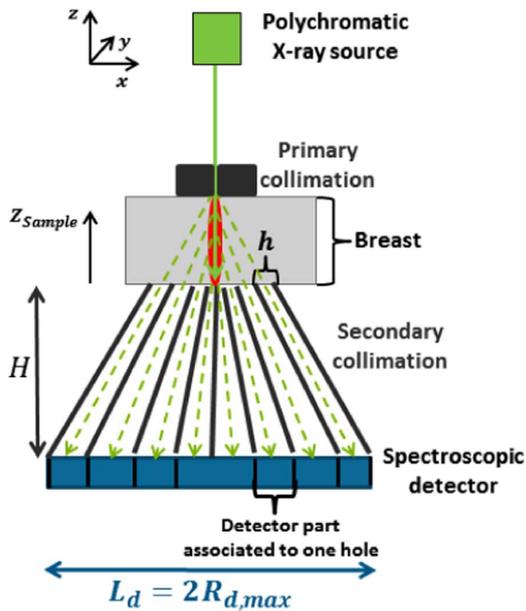


FIGURE 1.24 – Schéma du système proposé : il s'agit d'un système EDXRD qui permet de mesurer la diffusion à différents angles grâce au collimateur secondaire multifocal.

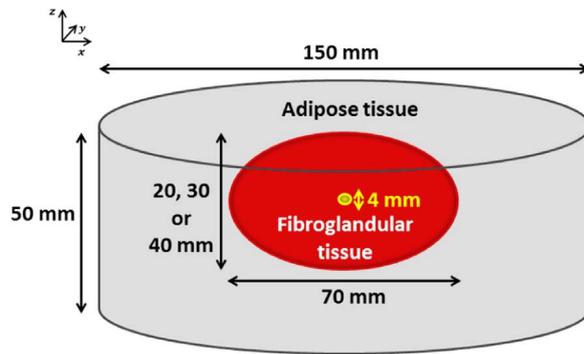


FIGURE 1.25 – Fantôme de sein : à l'intérieur d'une médium composé de tissu adipeux (en gris), le fantôme présente une masse fibroglandulaire (en rouge), au milieu de laquelle il y a une tumeur (en vert).

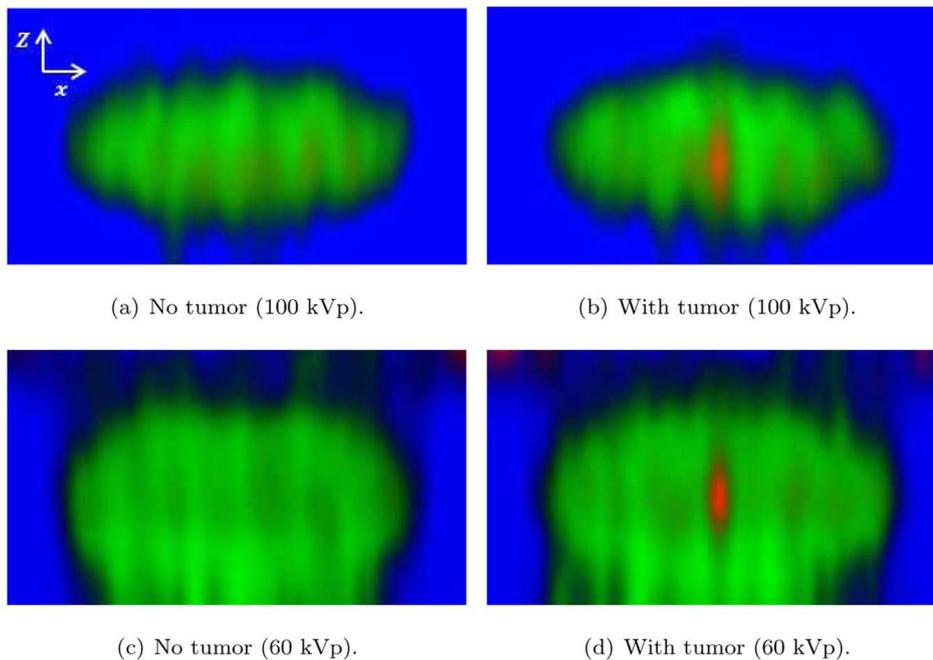


FIGURE 1.26 – Résultats obtenus avec et sans tumeur, respectivement à droite et à gauche, pour deux voltages différents du tube X : 100 kVp en haut et 60 kVp en bas. Les couleurs représentent les différents types de tissu : adipeux (bleu), fibreux (vert) et cancéreux (rouge).

Nombre de cônes	$H$ (mm)	$h$ (mm)	$p$ (mm)	$L_d$ (mm)
7	230	0.8	1.0	100

TABLE 1.1 – Paramètres du système, utilisés sur la figure 1.24

pour enquêter une zone suspecte.

Le système proposé est illustré sur la figure 1.24. Les longueurs utilisées sont détaillées dans la table 1.1. Les photons incidents issus d'une source polychromatique sont collimatés en pinceau qui illumine la zone suspecte de sein (en rouge sur la figure) dans toute son épaisseur. Le sein est comprimé entre le collimateur primaire et le collimateur secondaire. Le collimateur secondaire a une configuration multifocale grâce aux cônes concentriques, qui le composent et qui visent les différentes profondeurs du sein. Pour le détecteur spectrométrique, un semi-conducteur en CdZnTe de 5 mm d'épaisseur a été simulé. La surface de détection était de 10 cm  $\times$  10 cm et les pixels étaient carrées, avec un pas de 2.5 mm.

Le fantôme étudié est illustré plus en détail sur la figure 1.25. Il est constitué d'un cylindre de tissu adipeux, à l'intérieur duquel se trouve un ellipsoïde de tissu fibroglandulaire. Une tumeur de 4 mm est placée à l'intérieur du tissu fibro-glandulaire. Plusieurs épaisseurs de tissu fibroglandulaire ont été testées, comme indiqué sur la figure 1.25.

Les mesures de spectres diffusés ont été simulées avec la méthode Monte-Carlo en utilisant le logiciel PENELOPE [44]. Le but des simulation était de démontrer qu'il est possible de détecter correctement la tumeur par XRD, ainsi que d'estimer la dose absorbée nécessaire pour la différenciation entre les tissus sains et les tissus pathologiques. Comme la dose absorbée dépend, en partie, de la tension maximale du générateur X, plusieurs voltages ont été testés. Les spectres simulés ont été reconstruits en utilisant un algorithme itératif *Maximum Likelihood Expectation Maximization* (MLEM) [45]. La classification par type de tissu est intégrée dans cette reconstruction. Ceci permet de représenter les résultats des reconstructions sous la forme d'images en couleur, dont un exemple est présenté sur la figure 1.26. Y sont représentées les reconstructions pour deux configurations de fantôme (sans et avec tumeur, respectivement pour la première et la deuxième colonne) et pour deux tensions maximales du générateur X (100 kVp et 60 kVp, respectivement pour la première et la deuxième ligne). On constate que dans les deux cas la présence ou l'absence de tissu tumoral est correctement détectée.

Cette étude a montré que la géométrie proposée permet d'identifier correctement une tumeur de petite taille dans un fantôme de sein. La dose absorbée a été estimée pour les différents paramètres de simulation, tels que la tension maximale du tube ou l'épaisseur du tissu fibroglandulaire. Dans toutes ces configurations, la dose nécessaire pour avoir une séparation de  $3\sigma$  entre un spectre XRD sans et avec une tumeur au centre du fantôme est proche de 0.25 mGy. Cette dose est plus basse que celle délivrée pendant une mammographie standard, qui est d'environ 1 à 2 mGy [7]. La fiabilité du système a également été étudiée en fonction de la

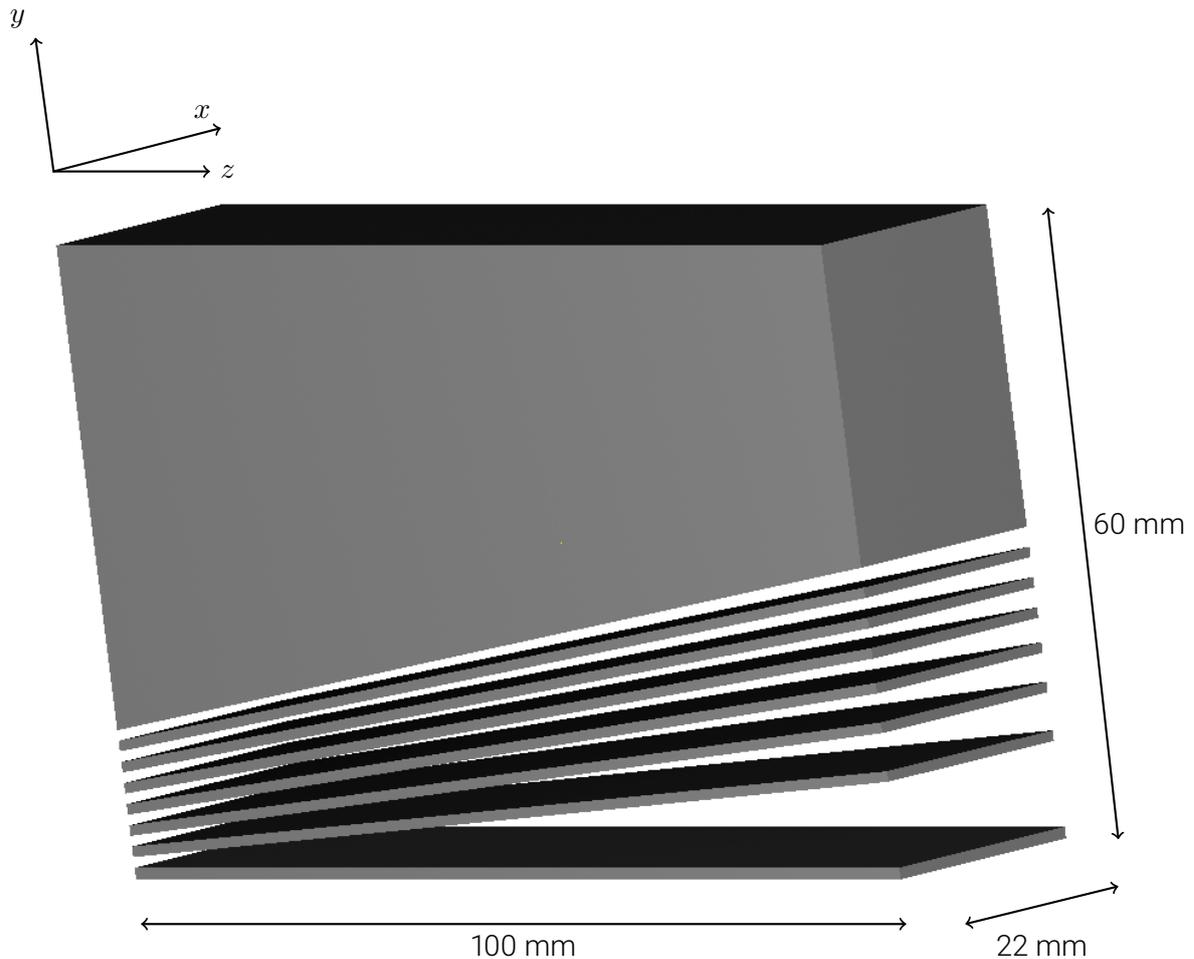


FIGURE 1.27 – Schéma 3D de la collimation hybride.

dose en utilisant les courbes caractéristique de fonctionnement du récepteur (ROC, *Receiver Operating Characteristic*). Pour une séparation de  $3\sigma$  et une dose de 0.25 mGy, l'aire sous la courbe (aire sous la courbe (AUC, *Area Under Curve*)) était supérieure à 0.9.

### **Systeme avec une géométrie optimisée réelle : thèse de D. BARBES**

La géométrie idéale proposée par Fanny MARTICKE n'a pas pu être réalisée à cause des difficultés de fabrication (lames conique avec des épaisseurs variables). Un nouveau collimateur, dont la géométrie a été inspirée par le design idéal, a donc été mis au point en prenant en compte ces limites de fabrication. Ce collimateur, représenté en 3D sur la figure 1.27, a été appelé "hybride". Fabriqué en tungstène, il comporte des lames planes d'épaisseur variable, qui permettent de viser simultanément plusieurs angles de diffusion. Les dimensions ont été optimisées pour le détecteur spectrométrique Imadif [20], développé au sein du laboratoire d'accueil. Ce détecteur CdZnTe comporte  $8 \times 24$  pixels ( $x, y$ ) mesurant  $2.75 \text{ mm} \times 1.25 \text{ mm}$ .

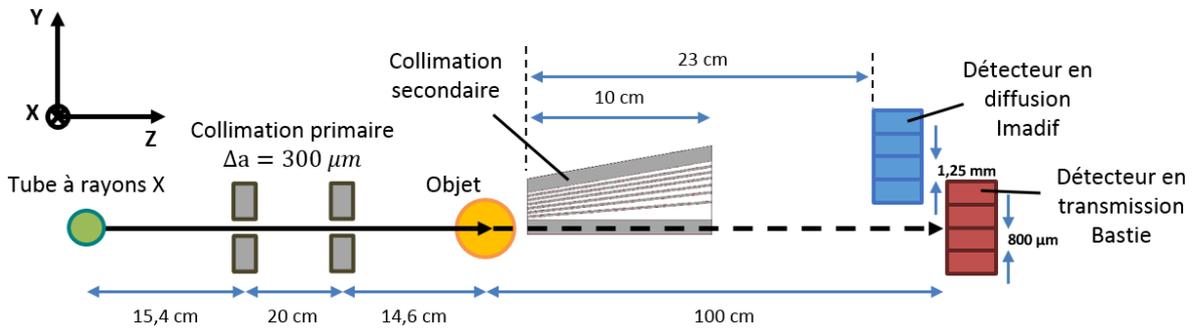
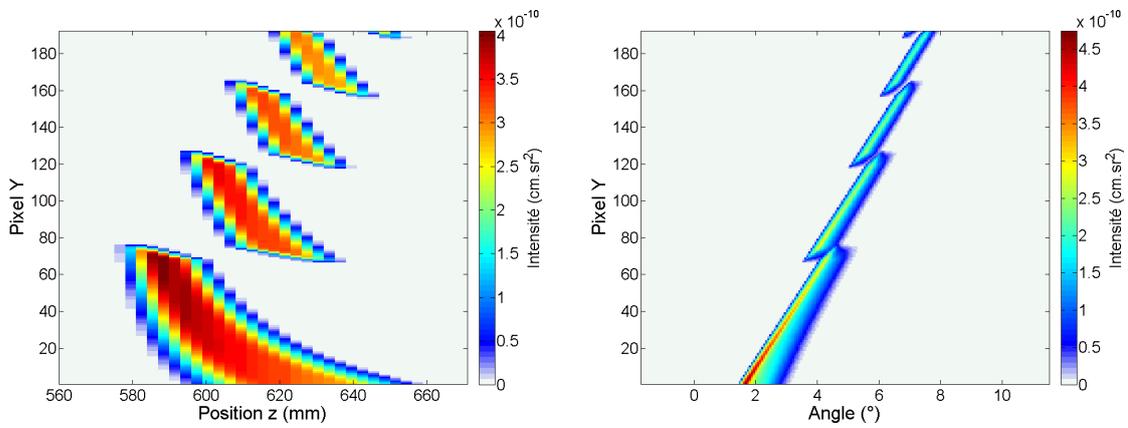


FIGURE 1.28 – Schéma du système expérimental avec la collimation hybride.



(a) Distribution de l'intensité de diffusion  $A(p, z)$

(b) Distribution des angles de diffusion

FIGURE 1.29 – Cartes de champ de vue obtenues en simulation.

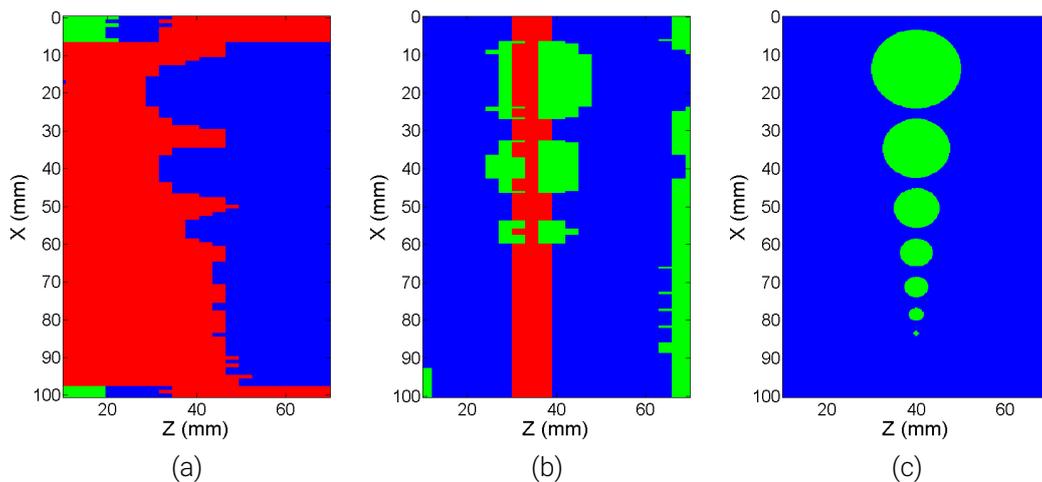


FIGURE 1.30 – Résultats de reconstruction des spectres obtenus avec la collimation hybride, respectivement expérimentalement (a) et par simulation (b). L'image (c) montre l'objet réel. Les couleurs représentent les matériaux : PS (bleu), nylon (vert) et PMMA (rouge).

Géométrie	Résolution spatiale en profondeur	Temps d'acquisition	Fiabilité
CSCT	de l'ordre du mm	30+ min par tranche	À tester
PixD	Aucune	5 min	À tester
CACSSI	de l'ordre du mm	10 s par position $(x, y)$	À tester
<b>CEA</b>			
Simulation	de l'ordre du mm	Non estimé	AUC > 0.9
Expérimental	de l'ordre du mm	5 min	À tester

TABLE 1.2 – Performances des systèmes présentés dans §§1.3.1 et 1.3.2.

Il a une résolution énergétique de 2.4% à 122 keV. Ce collimateur a été testé en simulation et expérimentalement dans la configuration représentée sur la figure 1.28. Les cartes de champ de vue correspondantes, obtenues par simulation, sont représentées sur la figure 1.29. On remarque que, contrairement à la collimation idéale, le multiplexage est bien présent : les  $z$  sont vues simultanément par plusieurs groupes de pixels.

Des acquisitions expérimentales ont été réalisées avec le même fantôme que précédemment. Le générateur de rayons X a opéré à 100 kV et 15 mA. Le temps d'acquisition par position  $x$  était de 5 min. Le résultat de cette expérience, après reconstruction et classification, est représenté sur la figure 1.30a. On constate que les inserts de nylon sont bien détectés, par contre, la classification est très loin de la vérité terrain (figure 1.30c). Pour mieux comprendre ce résultat, les spectres diffusés ont été simulés pour la même configuration expérimentale. Ces spectres ont été traités de la même manière que les spectres expérimentaux. Le résultat de ce traitement est présenté sur la figure 1.30b. La classification est meilleure : les couleurs sont correctes, sauf au bord, où on voit du nylon, et au milieu, où une bande de PMMA apparaît. Tous les inserts ne sont pas détectés.

Ces résultats sont moins bons que ce qui était espéré. Malheureusement, faute de temps, ils n'ont pas été améliorés. Cependant, plusieurs pistes avaient été envisagées :

- Augmenter le temps d'acquisition pour un meilleur rapport de signal à bruit ;
- Comparé à la première expérience, la gamme d'angles couverte est beaucoup plus large. La calibration devient plus difficile, car il est nécessaire de détecter le pic principal d'aluminium dans une gamme d'énergies plus large qu'auparavant. L'amélioration du processus de calibration pourrait être très bénéfique. L'utilisation de plusieurs matériaux de calibration est envisagée.
- La processus de reconstruction peut être amélioré avec une reconstruction MLEM.

### 1.3.3 Conclusion

Différents systèmes d'imagerie par XRD pour le diagnostic du cancer du sein ont été présentés. Dans l'hypothèse d'utilisation de cette méthode dans le milieu clinique, et particulièrement *in vivo*, trois critères de qualité ont été dégagés :

- la résolution spatiale, afin d'imager le sein en profondeur;
- le temps d'acquisition et la dose absorbée;
- la fiabilité du diagnostic, évaluée par la sensibilité et la spécificité.

La table 1.2 récapitule les performances des systèmes présentés dans §§1.3.1 et 1.3.2, en se basant sur ces critères. La dernière ligne "Expérimental" concerne la première validation expérimentale avec la collimation simple. La collimation hybride, proposée par Damien BARBES, était très prometteuse, mais faute de temps, certains problèmes n'ont pas pu être réglés.

## 1.4 Contexte médical et physique : conclusion

Dans ce chapitre, nous avons passé en revue les outils cliniques utilisés pour le diagnostic du cancer du sein, tels que la mammographie, la DBT, les ultrasons, l'IRM et, finalement, l'analyse anatomo-pathologique par biopsie. Ces méthodes d'imagerie ne permettent pas d'infirmier de manière fiable la présence d'un cancer sans faire de biopsie, c'est pourquoi d'autres méthodes de diagnostic de cancer du sein, telles que la diffusion cohérente de rayons X (XRD, *X Ray Diffraction*), sont étudiées par les chercheurs.

La XRD révèle la structure moléculaire des tissus et a le potentiel d'être utilisée *ex vivo* pour le contrôle de marges per-opératoire, ou *in vivo*, en deuxième niveau après la mammographie pour améliorer le diagnostic. Un système utilisant la XRD devra répondre à trois objectifs :

- posséder une résolution spatiale en profondeur de l'ordre du millimètre;
- être rapide, idéalement avec un temps d'exposition en dessous de la minute et une dose absorbée de l'ordre du milligray, pour l'application *in vivo*;
- apporter de l'information pertinente au diagnostic. En particulier, un taux très faible de faux positifs est recherché, ce qui correspond à une spécificité proche de 1.

Nous avons présenté quelques systèmes proéminents développés par les différentes équipes de recherche : le système CSCT, la géométrie PixD et la géométrie utilisant les masques codés CACSSI. Chacun de ces systèmes a des avantages et des inconvénients, mais pour le moment, aucun des systèmes ne répond aux trois objectifs à la fois. C'est pourquoi, au CEA deux thèses ont été menées en parallèle pour mettre au point un système qui répond à tous ces critères. À l'issue de ces deux thèses, un système prometteur avec une collimation multifocale a été proposé. Le travail de cette thèse s'inscrit dans la continuité de ce projet. Dans le chapitre suivant sera présenté en détail le montage expérimental, et en particulier, la méthode de traitement de données mise au point.

# MÉTHODE PROPOSÉE

---

Dans ce chapitre, nous allons passer en revue la méthode utilisée pendant cette thèse. Son architecture globale est illustrée sur la figure 2.1. La finalité est de déterminer  $C(z)$ , la classification en tissu "sain" ou "cancéreux" d'un objet donné en fonction de sa profondeur  $z$ , à partir des mesures de ses spectres diffusés,  $S_{diff}$ , et atténués  $S_{att}$ . Pour cela, nous allons acquérir des mesures et des données de calibration avec un montage expérimental. Ces données de calibration vont contribuer à construire une matrice système. Puis, en utilisant cette matrice, nous allons calculer des fonctions de diffusion  $F(z, q)$  pendant l'étape de reconstruction, pour finalement, les classifier et obtenir  $C(z)$ .

Dans ce chapitre, nous allons commencer par présenter le montage expérimental et ses différents composants, ainsi que le modèle direct qui le décrit. Puis, nous allons détailler la chaîne de traitement : la construction de la matrice système à l'aide de la calibration, la reconstruction et la classification. Finalement, une étude sur la dose absorbée et son influence sur les résultats sera proposée.

## 2.1 Géométrie et modèle

Dans cette section, nous présentons tout d'abord le montage expérimental et ses différents composants. Puis, nous allons présenter leur modèle direct.

### 2.1.1 Montage expérimental

La figure 2.2 montre une photographie du banc expérimental en configuration d'acquisition des spectres diffusés. Ce montage, schématisé sur la figure 2.3, comprend :

- Le tube de rayons X polychromatique, avec une anode de tungstène (W);
- la collimation primaire en pinceau;
- l'objet à évaluer;

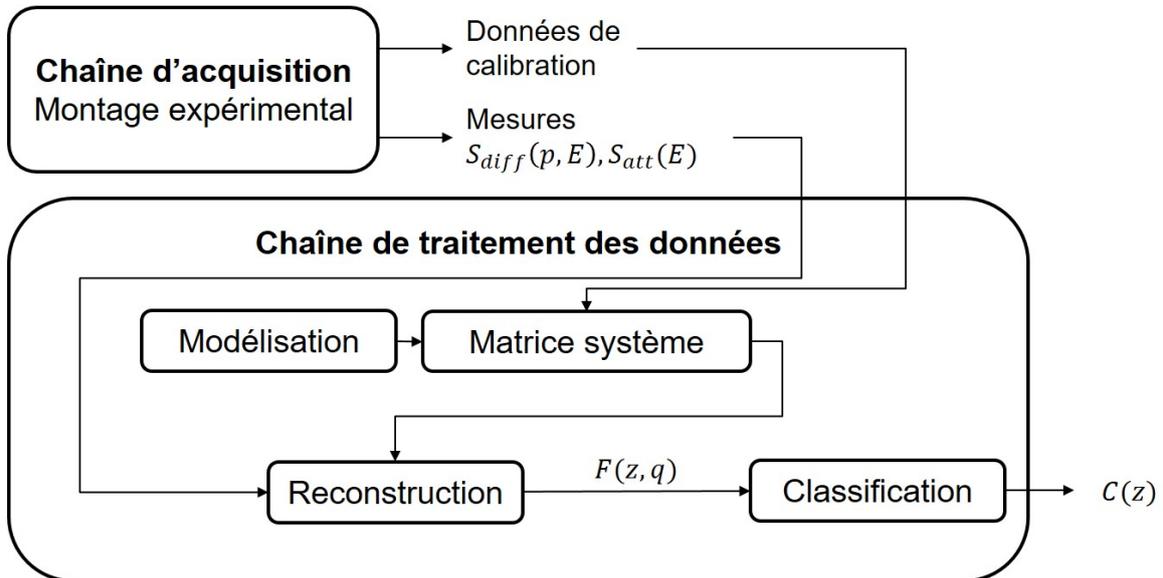


FIGURE 2.1 – Schéma représentant l'architecture globale de la méthode : il comprend une chaîne d'acquisition de données et une chaîne de traitement de données

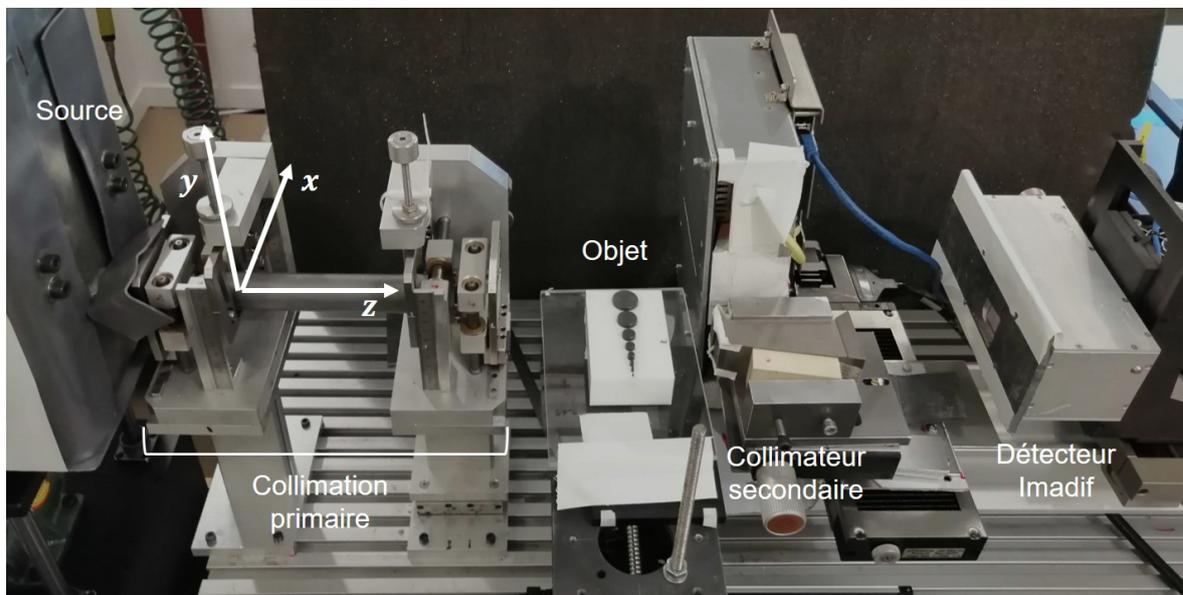


FIGURE 2.2 – Photo du montage expérimental en configuration d'acquisition des spectres diffusés.

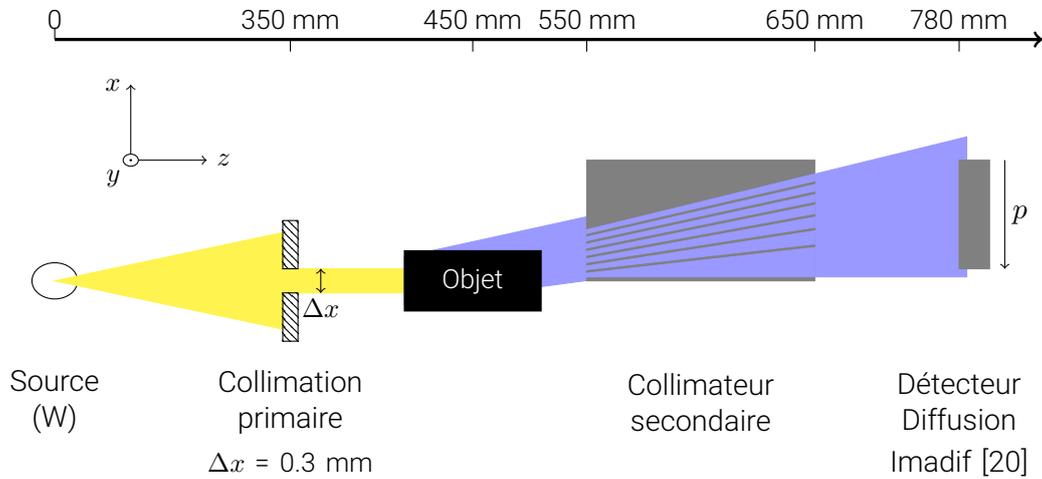
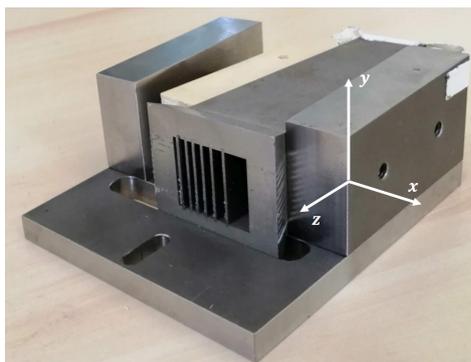
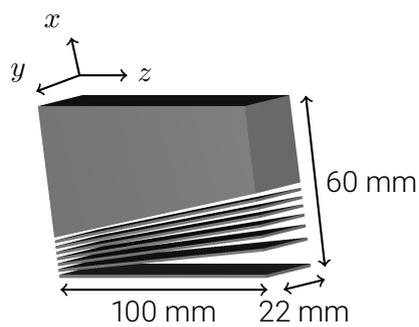


FIGURE 2.3 – Schéma dans le plan  $z, x$  (vue de dessus) du montage expérimental en configuration d’acquisition des spectres diffusés. Le rayonnement direct (avant objet) est représenté en jaune. Le rayonnement diffusé est représenté en bleu. Pour plus de lisibilité, l’échelle n’est pas respectée.



(a)



(b)

FIGURE 2.4 – Collimateur secondaire. Il est fabriqué en tungstène et comporte 7 fentes. (a) photographie du collimateur dans son socle, (b) modèle 3D.

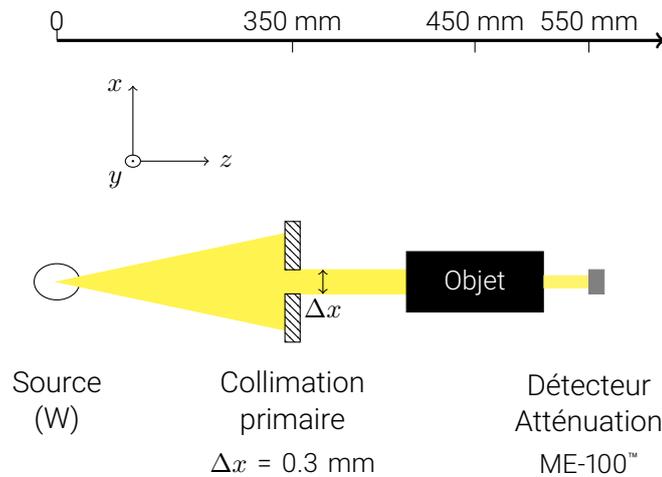


FIGURE 2.5 – Schéma dans le plan  $(z, x)$  du montage expérimental en configuration d'acquisition des spectres transmis. Le rayonnement direct (avant objet) et transmis (après l'objet) sont représentés en jaune. Pour plus de lisibilité, l'échelle n'est pas respectée.

- le collimateur secondaire, représenté plus en détail sur la figure 2.4. Il est fabriqué en tungstène, et il comporte 7 fentes, qui permettent de voir différentes parties de l'objet. Le champ de vue de ce collimateur sera abordé dans la sous-section §2.2.3;
- le détecteur spectrométrique Imadif [20]. Il s'agit d'un détecteur à base de tellure de cadmium dopé au zinc (CdZnTe), composé de  $24 \times 8$  pixels, sur une surface de détection de  $660 \text{ mm}^2$ . La technique de "surpixelisation" [39], mise au point dans notre laboratoire, permet de diviser virtuellement les pixels selon les lignes en 8 sous-pixels chacun. Ainsi, on obtient un détecteur avec  $192 \times 8$  pixels virtuels. La résolution en énergie est de 4.25% à 60 keV.

Comme nous verrons dans la sous-section suivante, notre modèle prend également en compte l'atténuation des rayons X par l'objet. Pour l'estimer, nous mesurons également les spectres en configuration de transmission, comme représenté sur la figure 2.5. Les acquisitions de spectres diffusés et transmis se font séparément : le collimateur secondaire est alors décalé pour laisser place au détecteur spectral ME-100™ (MultiX). Il s'agit d'un détecteur CdZnTe, composé d'une rangée de 128 pixels. Pour nos besoins, un seul pixel suffit pour la mesure du spectre transmis.

En effet, l'agencement de notre banc expérimental ne permet pas actuellement d'effectuer des mesures de spectres diffusés et transmis simultanément, sans déplacer les composants. À terme, dans un système industriel, il faudra évidemment que les mesures puissent se faire en parallèle.

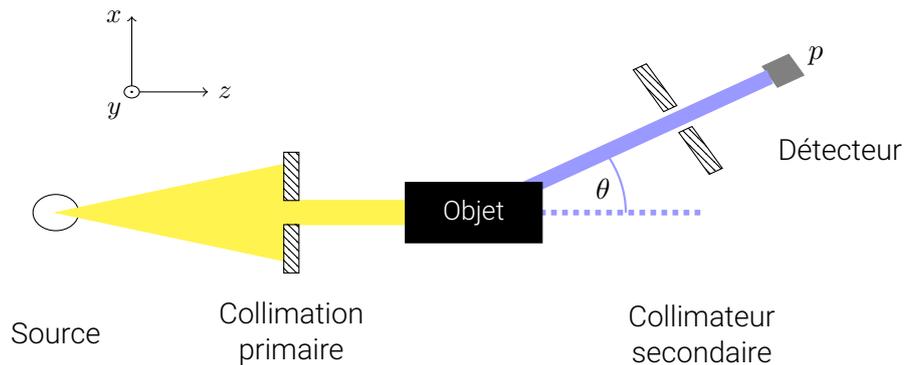


FIGURE 2.6 – Schéma dans le plan  $(z, x)$  d'un système simplifié : ici, l'angle diffusion  $\theta$  est unique, et le détecteur ne comporte qu'un seul pixel  $p$ . Le rayonnement incident et le rayonnement diffusé sont représentés, respectivement, en jaune et en bleu.

### 2.1.2 Construction du modèle direct

Pour calculer la fonction de diffusion de l'objet à partir de son spectre diffusé, nous avons besoin de relier ces deux grandeurs par un modèle. Ce modèle prendra en compte l'influence de chacun des composants du montage expérimental sur le trajet des photons :

- le spectre incident,  $S_{inc}(E)$ , qui a une distribution énergétique non-uniforme;
- l'atténuation du spectre incident par l'objet,  $Att(E)$ ;
- la fonction de diffusion,  $F(q)$ , qui impacte l'angle de diffusion des photons en fonction de leur énergie;
- le collimateur secondaire, qui limite les angles de diffusion, et donc les positions  $z$  visibles par le détecteur;
- la réponse en énergie du détecteur spectral.

Notre montage expérimental étant assez complexe, nous allons commencer par présenter une modélisation pour un montage simplifié ne comportant qu'un seul pixel. Ensuite, nous allons généraliser ce modèle à un système multipixel.

#### Modèle direct simplifié

Le montage simplifié est représenté sur la figure 2.6 : le détecteur ne comporte qu'un seul pixel, et le collimateur secondaire est une fente qui ne laisse passer que les photons, qui ont été diffusés avec un angle  $\theta$ . L'ouverture de la fente influence la résolution spatiale et la résolution angulaire du système, comme illustré sur la figure 2.7.

Nous allons illustrer les transformations, que subit le spectre, en passant par les différents composants avec l'exemple de l'aluminium [18] (voir figure 2.8) :

- a** Le spectre incident, dont la distribution énergétique dépend des paramètres opératoires du générateur X;

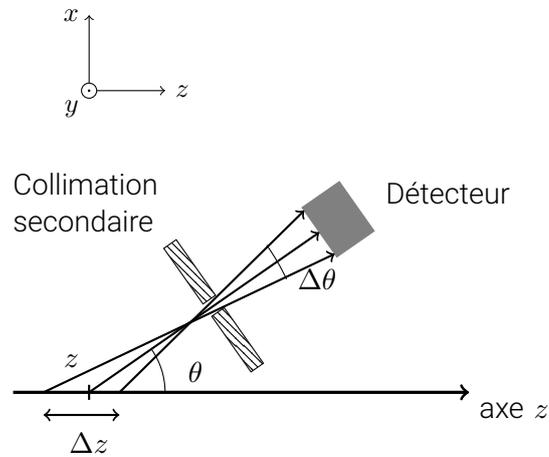


FIGURE 2.7 – Effet de la collimation secondaire sur la résolution spatiale, et la résolution angulaire.

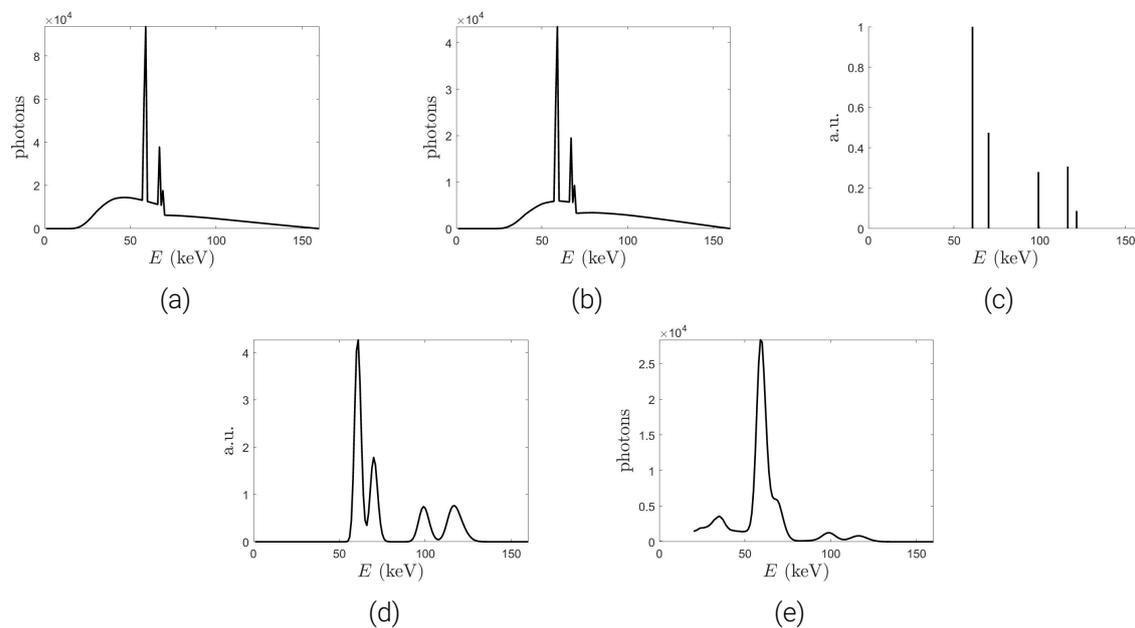


FIGURE 2.8 – Influence des différents composants de la chaîne d'acquisition sur le spectre diffusé par l'aluminium [18] : (a) Spectre incident  $S_{in}(E)$ ; (b) Spectre incident atténué par l'objet  $S_{att}(E)$ ; (c) Spectre diffusé par l'aluminium à l'angle de diffusion moyen; (d) Spectre diffusé à travers la collimation secondaire, effet de flou dû à la résolution angulaire et spatiale; (e) Spectre diffusé, mesuré par le détecteur.

- b** Le spectre incident est atténué par l'objet. Comme nous travaillons avec des matériaux peu atténuants, nous faisons l'hypothèse simplificatrice que tous les photons diffusés mesurés ont subi la même atténuation, quel que soit l'angle de diffusion;
- c** Spectre diffracté par l'aluminium situé en un point  $z$  ponctuel, à l'angle de diffusion moyen  $\theta$ , fixé par la collimation secondaire;
- d** Spectre diffracté par l'aluminium situé dans l'ensemble du volume inspecté  $\Delta z$ , en tenant compte de l'effet de flou dû à la collimation secondaire ( $\Delta z$  et  $\Delta\theta$ , représentés sur la figure 2.7);
- e** Spectre diffracté, mesuré par le détecteur, en tenant compte la réponse du détecteur.

En tenant compte de ces composants et de leur influence sur le spectre, le modèle physique de la chaîne d'acquisition s'écrit :

$$\forall E_{det}, E_{inc}, z, q$$

$$S_{diff}(E_{det}) = \sum_{E_{inc}} \left( R_{det}(E_{det}, E_{inc}) Att(E_{inc}) S_{inc}(E_{inc}) \sum_{z,q} (R_c(E_{inc}, z, q) F(z, q)) \right) \quad (2.1)$$

où :

- $E_{det}$ ,  $E_{inc}$ ,  $z$  et  $q$  représentent, respectivement, l'énergie détectée, l'énergie incidente, la position selon l'axe des rayons X, et le transfert de quantité de mouvement.
- $R_{det}$  est la matrice de réponse du détecteur (DRM, *Detector Response Matrix*).
- $Att$  est l'atténuation de l'objet.
- $S_{inc}$  est le spectre incident.
- $R_c$  est la matrice de réponse géométrique du collimateur secondaire, elle reflète la résolution spatiale et angulaire du système. Elle décrit finalement la relation

$$q = \frac{E}{hc} \sin(\theta/2)$$

pour l'ensemble des points du volume inspecté par le système.

- $F$  est la fonction de diffusion de l'objet, que l'on cherche à déterminer.

### Modèle direct pour le montage expérimental complet

Notre montage expérimental, représenté sur la figure 2.3, est plus complexe à modéliser, car sa collimation secondaire comporte plusieurs fentes avec des ouvertures différentes, ainsi qu'un détecteur spectrométrique composé de  $192 \times 8$  pixels. En plus des phénomènes modélisés dans le cas précédent, nous avons donc besoin d'établir la relation entre les pixels, les angles de diffusion et les positions  $z$ . La figure 2.9 illustre cette relation : chaque couleur correspond à une fente du collimateur secondaire. On peut ainsi observer, pour chaque fente,

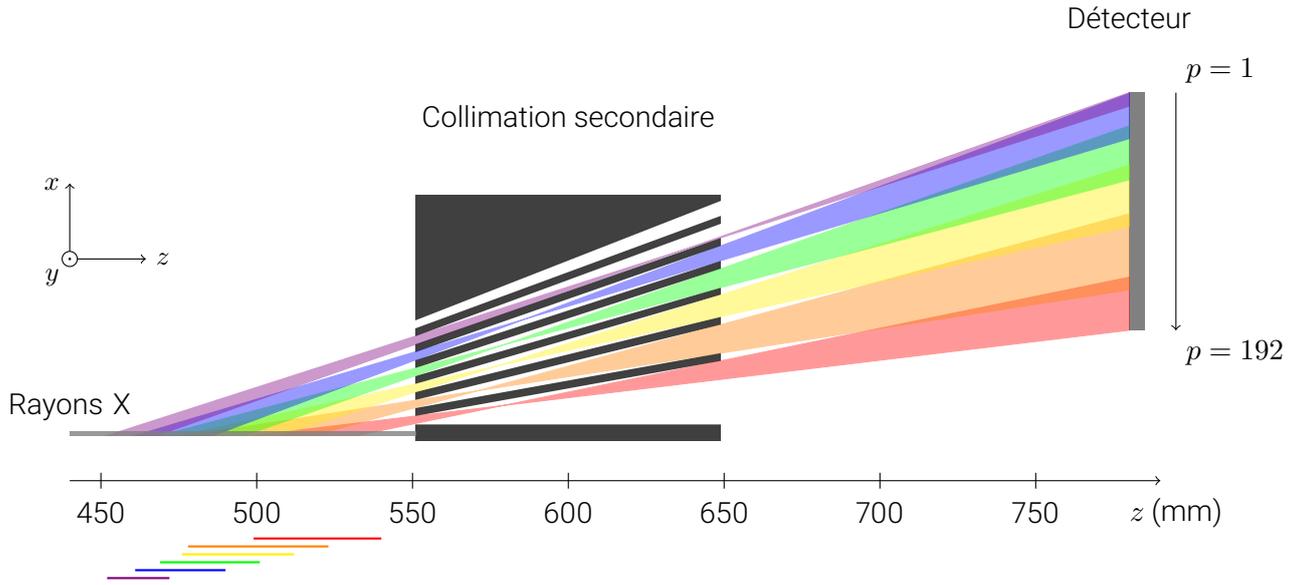


FIGURE 2.9 – Schéma dans le plan  $(z, x)$  du champ de vue du collimateur secondaire. Chaque couleur est associée à une fente unique et illustre la relation entre les positions  $z$  et les pixels du détecteur. Les traits de couleur en dessous de l'axe  $z$  représentent les intervalles  $z$  vus par chaque fente. Le faisceau direct est représenté en gris.

quels intervalles selon  $z$  sont "vus" par les différents groupes de pixels, et sous quels angles. Comme indiqué sur la figure, on considère que l'origine des pixels  $p = 1$  correspond au pixel le plus éloigné de l'axe  $z$ . On observe d'une part, que les angles de diffusion diminuent, et d'autre part, que les intervalles  $z$  se rapprochent du collimateur secondaire, avec les pixels. En dessous de l'axe  $z$ , les traits horizontaux colorés représentent les intervalles vus par chaque fente. On peut remarquer, qu'il y a un léger multiplexage.

Pour modéliser cette collimation, une dimension supplémentaire représentant les pixels, notés  $p$ , est ajoutée. Le spectre diffusé est donc mesuré en fonction des énergies  $E_{det}$  et des pixels  $p$ , et la matrice de résolution géométrique,  $R_c$ , intègre également la dimension des pixels. Le modèle établi précédemment devient :

$$S_{diff}(p, E_{det}) = \sum_{E_{inc}} \left( R_{det}(E_{det}, E_{inc}) \text{Att}(E_{inc}) S_{inc}(E_{inc}) \sum_{z,q} (R_c(p, E_{inc}, z, q) F(z, q)) \right) \quad (2.2)$$

Par ailleurs, le détecteur Imadif est bi-dimensionnel, il possède 192 lignes et 8 colonnes. Dans notre étude, nous n'utilisons que le décalage angulaire selon les lignes. La variation angulaire selon les colonnes étant suffisamment faible, nous sommes donc les spectres mesurés sur les 6 colonnes centrales. Les colonnes situées aux bords du détecteur sont assez bruitées et présentent des artefacts, elles ne seront pas prises en compte. Pour alléger les notations, dans la suite de nos travaux, "pixel" désigne un pixel-ligne, et  $S_{diff}(p, E)$  désigne le spectre diffusé, sommé sur les 6 colonnes centrales du détecteur, en fonction des pixels-lignes et des

énergies.

### 2.1.3 Conclusion

Dans cette section, nous avons établi le modèle direct de notre chaîne d'acquisition, qui exprime l'influence des différentes parties du montage expérimental sur le spectre diffusé mesuré. Pour calculer la fonction de diffusion  $F$  à partir du spectre diffusé mesuré, nous avons besoin tout d'abord d'estimer chaque partie de ce modèle. Ces estimations font l'objet de la section suivante.

## 2.2 Estimation de la matrice système

Dans cette section, nous allons estimer les différentes composantes du modèle 2.2, établi précédemment :

- la réponse du détecteur ;
- le spectre incident, et son atténuation par l'objet ;
- la matrice de réponse géométrique,  $R_c$ .

Puis, nous allons regrouper toutes ces informations dans une seule matrice, notée  $R$ , et appelée "matrice système globale". Ainsi, nous pourrions écrire la relation entre le spectre diffusé et la fonction de diffusion comme :

$$S_{diff}(p, E) = R(p, E, z, q) F(z, q) \quad (2.3)$$

Pour alléger les notations, dorénavant  $E$  représente  $E_{det}$ , l'énergie détectée des photons.

### 2.2.1 Matrice de réponse du détecteur

La DRM est propre à chaque détecteur. Dans le cadre de cette thèse, nous avons utilisé les détecteurs Imadif [20] et ME-100 (MultiX). Développés au sein du laboratoire d'accueil, leurs réponses ont été estimées via un outil interne de modélisation [46]. Cet outil repose sur les simulations des effets électroniques des interactions dans le semi-conducteur et de l'induction des signaux par le partage de charges. Le procédé est validé expérimentalement avec des sources de cobalt et américium. La figure 2.10 illustre la DRM du détecteur Imadif. La sous-figure (b) illustre la répartition des énergies détectées pour une énergie incidente de 60 keV : l'énergie du pic principal correspond bien à l'énergie des photons incidents. La largeur de ce pic définit la résolution énergétique du détecteur pour une énergie donnée. Par ailleurs, on observe un deuxième pic, plus faible, d'énergies inférieures à l'énergie incidente. Ce pic résulte des phénomènes de fluorescence, de partage de charges, etc..

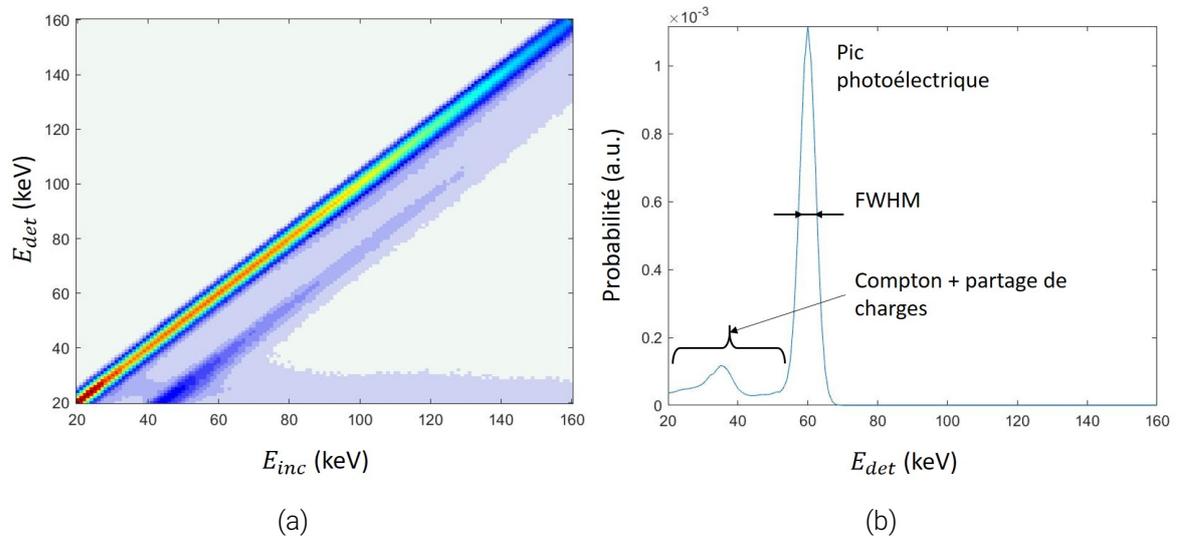


FIGURE 2.10 – (a) Matrice de réponse du détecteur Imadif, (b) Réponse du détecteur Imadif à l'énergie incidente de 60 keV

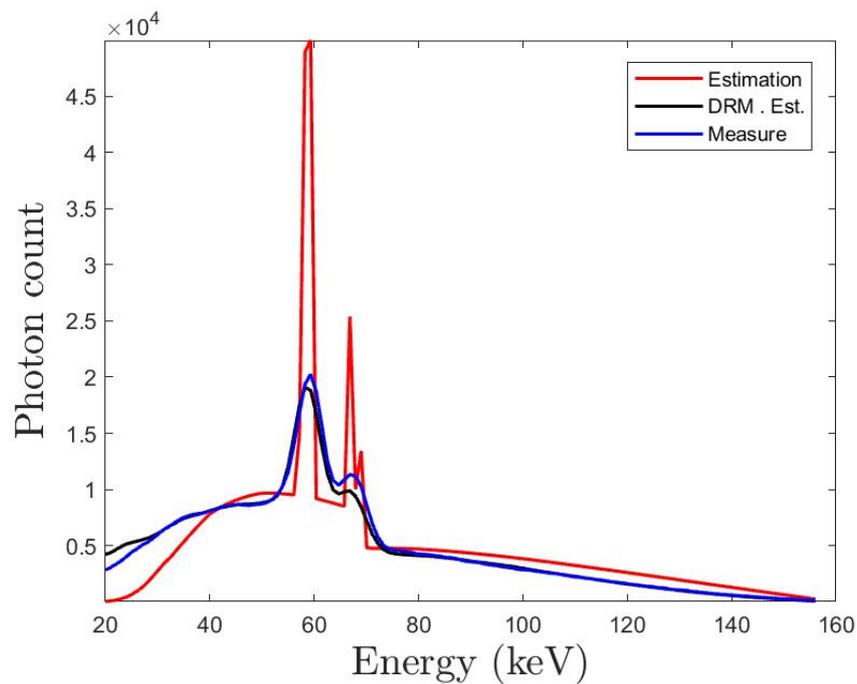


FIGURE 2.11 – Résultat de la régression non-linéaire pour l'estimation du spectre incident atténué par 50 mm d'eau. Les courbes représentent : en bleu, le spectre atténué mesuré,  $S_{att}$ ; en rouge, l'estimation obtenue,  $(S_{inc} \otimes Att)$ ; en noir, le produit de l'estimation par la DRM du détecteur,  $R_{ME-100}(S_{inc} \otimes Att)$

## 2.2.2 Estimation du spectre incident atténué

Nous cherchons à estimer le spectre incident et l'atténuation de l'objet étudié. Pour cela, nous mesurons le spectre atténué en utilisant la configuration illustrée sur la figure 2.5. Cependant, cette mesure correspond au spectre incident atténué par l'objet, multiplié par la DRM du détecteur ME-100, qu'on peut modéliser de la manière suivante :

$$S_{att,M}(E_{det}) = R_{ME-100}(E_{det}, E_{inc}) (S_{inc}(E_{inc}) \otimes Att(E_{inc})) \quad (2.4)$$

où :

- $S_{att}$  est le spectre atténué mesuré ;
- $R_{ME-100}$  est la DRM du détecteur ME-100 ;
- $(S_{inc} \otimes Att)$  est le produit terme à terme du spectre incident par l'atténuation de l'objet.

La matrice  $R_{ME-100}$  est connue de la même façon que pour le détecteur Imadif, comme décrit dans la sous-section précédente. Le spectre incident  $S_{inc}$ , à un facteur d'échelle  $G$  près, a été obtenu par simulation avec un modèle semi-empirique [47]. Nous allons donc déterminer ce facteur  $G$  et l'atténuation de l'objet, à partir de la mesure  $S_{att}$  et des grandeurs connues,  $R_{ME-100}$  et  $S_{inc}$ .

L'atténuation de l'objet est modélisée par la loi de Beer-Lambert :

$$Att(E) = exp\left(-\sum_i (\mu_i(E) l_i)\right) \quad (2.5)$$

où  $\mu_i$  et  $l_i$  sont, respectivement, les coefficients d'atténuation linéique et les épaisseurs des différents matériaux  $i$  constituant l'objet. Or, nous cherchons à identifier un objet dont, justement, nous ne connaissons pas la composition. Nous avons donc utilisé la décomposition d'ALVAREZ-MACOVSKI [48], qui permet d'exprimer l'atténuation d'un matériau inconnu, dont on ne connaît ni le coefficients d'atténuation linéique  $\mu_i(E)$ , ni leurs longueurs respectives  $l_i$ , à l'aide de deux matériaux  $M_1$  et  $M_2$ , dont on connaît les coefficients d'atténuation linéique  $\mu_1$  et  $\mu_2$  :

$$\sum_i (\mu_i(E) l_i) \approx \mu_1(E) l_1 + \mu_2(E) l_2 \quad (2.6)$$

où  $l_1$  et  $l_2$  sont les épaisseurs équivalentes de  $M_1$  et  $M_2$  à déterminer. Ceci est possible car les matériaux qu'on étudie n'ont pas de raies  $K$ , et leurs atténuations ne présentent donc pas de discontinuité en fonction des énergies.

Ainsi, le modèle de la mesure du spectre incident atténué s'écrit :

$$S_{att}(E_{det}) = G R_{ME-100}(E_{det}, E_{inc}) (S_{inc}(E_{inc}) \otimes exp(-\mu_1(E) l_1 - \mu_2(E) l_2)) \quad (2.7)$$

Pour inverser ce modèle, et calculer  $Att$  à partir de la mesure expérimentale  $S_{att}(E)$ , nous utilisons une régression non-linéaire. La figure 2.11 montre un exemple d'estimation obtenu

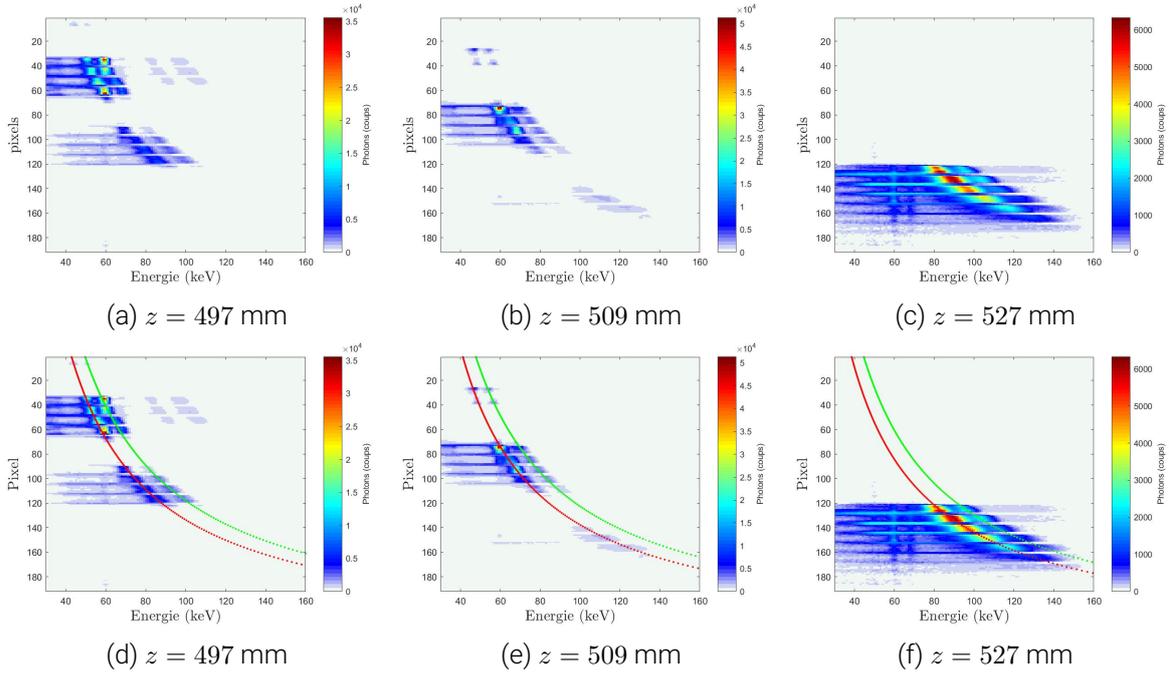


FIGURE 2.12 – Spectres diffusés par la plaque d'aluminium. Les figures (a), (b) et (c) illustrent, respectivement, les spectres diffusés mesurés pour les positions  $z = 497, 509$  et  $527$ mm. Sur les figures (d), (e) et (f), sont tracés les modèles  $E_{111}(p, z)$  et  $E_{200}(p, z)$ .

pour le spectre incident atténué par 50 mm d'eau :

- en bleu, le spectre atténué mesuré, noté  $S_{att}$  ;
- en rouge, représente l'estimation obtenue,  $(S_{inc} \otimes Att)$  ;
- en noir, le produit de l'estimation par la DRM du détecteur,  $R_{ME-100}(S_{inc} \otimes Att)$

Les courbes bleue et noire sont quasi-identiques, on en déduit que l'estimation du spectre atténué, en rouge, est assez fiable.

### 2.2.3 Calibration de la matrice de réponse géométrique $R_c$

La calibration expérimentale a pour objectif de calculer la matrice de réponse géométrique,  $R_c$ , qui exprime l'intensité de diffusion en fonction des pixels, des énergies, des positions  $z$  et du transfert de quantité de mouvement  $q$ . Pour cela, nous allons commencer par déterminer les matrices, dites "de champ de vue", qui expriment l'intensité de diffusion en fonction de  $p, z$  et de l'angle de diffusion. Puis, en utilisant la loi de Bragg, nous en déduisons  $R_c$ .

Dans cette partie, on néglige la DRM, car on considère les pics de diffraction. Ainsi, on peut simplifier les notations :  $E_{inc} = E_{det} = E$ .

#### Calcul des matrices de champ de vue

Pour commencer, nous allons calculer deux matrices de champ de vue :

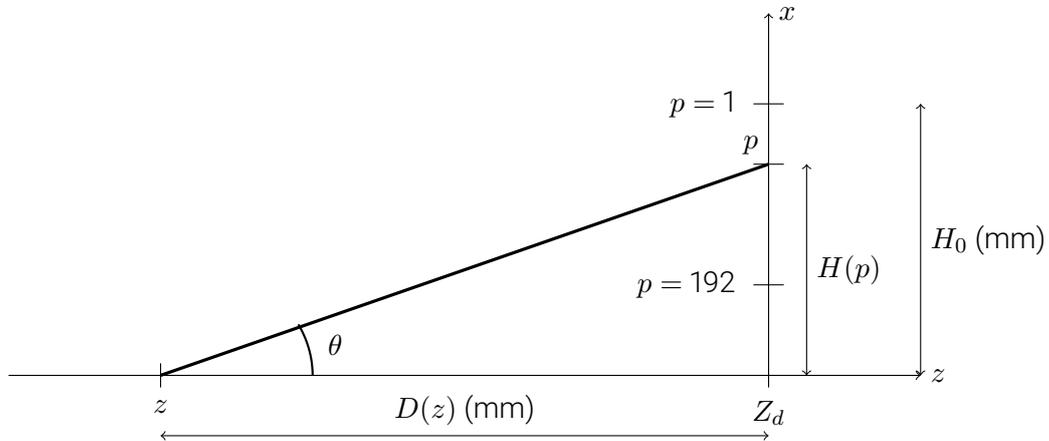


FIGURE 2.13 – Schéma illustrant la relation entre  $p$ ,  $z$  et l'angle de diffusion  $\theta$ .  $Z_d$  est la position du détecteur selon  $z$ , et  $D(z)$  représente la distance selon  $z$  entre le détecteur et un point  $z$  de l'objet.  $H_0$  représente la distance en millimètres, entre le haut du détecteur et l'axe  $z$ .  $H(p)$  correspond à la distance entre l'axe  $z$  et le centre du pixel  $p$ .

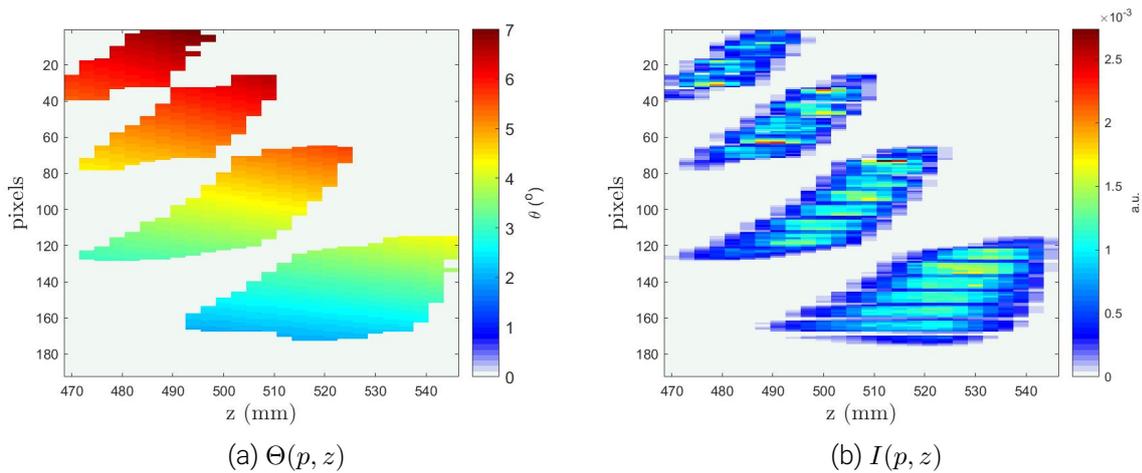


FIGURE 2.14 – Matrices de champ de vue obtenues expérimentalement

$\Theta(p, z)$  qui représente l'angle de diffusion du photon diffusé en  $z$ , détecté par le pixel  $p$ ;

$I(p, z)$  qui représente l'intensité de diffusion à la position  $z$ , perçue par le pixel  $p$ . Cette intensité est définie comme le ratio entre le nombre de photons issus de  $z$ , détectés en  $p$ , et le nombre total de photons détectés sur toute la surface du détecteur.

Pour cela, nous utilisons des spectres diffusés par une plaque d'aluminium, de 3 mm d'épaisseur. Les spectres diffusés sont mesurés en décalant la plaque selon  $z$  par pas de 3 mm. Ainsi, nous pouvons observer quels pixels voient la plaque à chaque position, et déduire la relation entre les pixels et les positions  $z$ . L'aluminium est un matériau cristallin dont les paramètres de maille sont connus. Dans nos expériences, nous pouvons observer ces deux pics de diffraction principaux, qui correspondent aux distances inter-réticulaires des plans (111) et (200) :  $d_{111} = 2.34 \cdot 10^{-10}$  mètres et  $d_{200} = 2.02 \cdot 10^{-10}$  mètres. Les énergies, qui correspondent à ces distances, vont changer en fonction de l'angle de diffusion, qui varie de pixel en pixel. Ainsi, en observant les positions en énergies de ces pics en fonction des pixels, nous pouvons déduire la relation entre les angles de diffusion et les pixels. Donc, pour établir la matrice de réponse géométrique, nous allons tout d'abord relever les énergies correspondantes aux pics de diffraction, en fonction des pixels et des  $z$ .

La première ligne des figures 2.12 montre des exemples de spectres diffusés mesurés pour les positions  $z = 497, 509$  et  $527$  mm. Nous pouvons observer que, en fonction de la position  $z$ , des groupes différents de pixels sont éclairés. De plus, nous remarquons que le nombre de photons varie en fonction des énergies et des pixels : il s'agit des pics de diffraction de l'aluminium qui se décalent vers les hautes énergies en fonction des pixels. Pour un pixel  $p$  donné, le pic de plus basse énergie est le pic principal correspondant à  $d_{111}$ . Cependant, comme on peut l'observer sur les pixels 30 à 60 de la figure 2.12 (a), le pic principal n'est pas toujours celui de plus grande intensité. En effet, lorsque les pics sont proches de 60 keV, ils sont amplifiés par les raies K- $\alpha$  et K- $\beta$  du tungstène. Dans ce cas, si le deuxième pic est amplifié, il paraît alors plus intense que le pic principal. Par ailleurs, comme on peut l'observer sur la figure 2.12c, pour les angles de diffusion importants, le pic s'élargit de plus en plus. Tous ces facteurs rendent la détection de l'énergie du pic principal, en fonction des pixels et des positions  $z$ , relativement compliquée. Les algorithmes simples, comme la détection de maxima par exemple, ne permettent pas d'y répondre. Pour résoudre à ce problème, nous avons décidé de procéder en plusieurs étapes.

Tout d'abord, nous avons sélectionné manuellement, en regardant les spectres diffusés pour chaque position, quelques couples  $(z, p)$ , où le pic principal est plus intense que le deuxième et facilement détectable. Pour ces couples, nous détectons la position en énergie du pic principal,  $E_{111}$ , grâce au maximum d'intensité. Cela nous donne une valeur  $E_{111}$  pour chaque couple  $(z, p)$ . Par exemple, pour le spectre mesuré à  $z = 509$  mm (figure 2.12b), nous avons sélectionné les pixels  $p \in [73-101]$ , puis détecté la position en énergie du maximum d'intensité.

La relation entre l'angle de diffusion,  $z$  et  $p$  est purement géométrique, comme illustré sur la figure 2.13 :

$$\begin{aligned} \tan(\theta) &= \frac{H(p)}{D(z)} \\ &= \frac{Z_d - z}{H_0 - \delta_p * (p - 1)} \end{aligned} \quad (2.8)$$

où  $Z_d$  est la position du détecteur selon  $z$ ,  $H_0$  la hauteur du détecteur selon  $x$  ( $p = 1$ ), et  $\delta_p$  le pas du pixel. Ainsi, on peut exprimer l'énergie du pic principal,  $E_{111}$ , en fonction de  $z$  et  $p$ , en utilisant la loi de Bragg :

$$\begin{aligned} E_{111} &= \frac{hc}{\sin(\theta/2)} \\ &= \frac{hc}{2 d_{111} \sin\left(\frac{1}{2} \arctan\left(\frac{Z_d - z}{H_0 - \delta_p * (p - 1)}\right)\right)} \end{aligned} \quad (2.9)$$

$Z_d$  et  $H_0$  font partie des spécifications du montage expérimental et sont connus à quelques millimètres près. Nous réalisons une régression non-linéaire avec les couples  $(z, p)$  et les valeurs  $E_{111}$  correspondantes, sélectionnés précédemment. Cela permet d'avoir des valeurs précises de  $Z_d$  et  $H_0$ , et de vérifier simultanément qu'il n'y a pas eu d'erreur dans le montage.  $Z_d$  et  $H_0$  étant estimées, nous avons un modèle théorique qui lie l'énergie de diffusion, ainsi que l'angle de diffusion, aux pixels et aux positions  $z$ . Nous avons tracé sur les figures 2.12 (d), (e) et (f) les énergies théoriques, obtenues avec ce modèle, pour le premier et le deuxième pic, respectivement en rouge et en vert. Nous constatons que le modèle correspond très bien aux données.

Nous utilisons ensuite le modèle  $E_{100} = f(p, z)$  pour calculer  $I(p, z)$ . Connaissant la position théorique du pic, nous pouvons réduire considérablement la fenêtre énergétique,  $E_{win}$ , où nous cherchons sa position. Pour s'affranchir localement de la forme du spectre incident atténué, nous calculons le ratio :

$$\forall E \in E_{win}, Ra(p, E) = S_{diff}(p, E) / S_{att}(E)$$

La position mesurée du pic est alors détectée en réalisant un fit gaussien sur  $Ra$ . Finalement, la valeur  $I(p, z)$  est définie comme la somme de  $Ra$  sur la largeur à mi-hauteur (FWHM, *Full Width at Half-Maximum*) du pic détecté. La FWHM du pic donne la résolution en énergie,  $\Delta E(p, z)$ . En différenciant la loi de Bragg, nous en déduisons la résolution angulaire :

$$\begin{aligned} \Delta\theta &= \left| \frac{\delta\theta}{\delta E} \right| \Delta E \\ \Rightarrow \forall(p, z), \Delta\theta(p, z) &= \frac{2hc}{E_{111}(p, z) \sqrt{(2d_{111}E_{111})^2 - (hc)^2}} \Delta E \end{aligned} \quad (2.10)$$

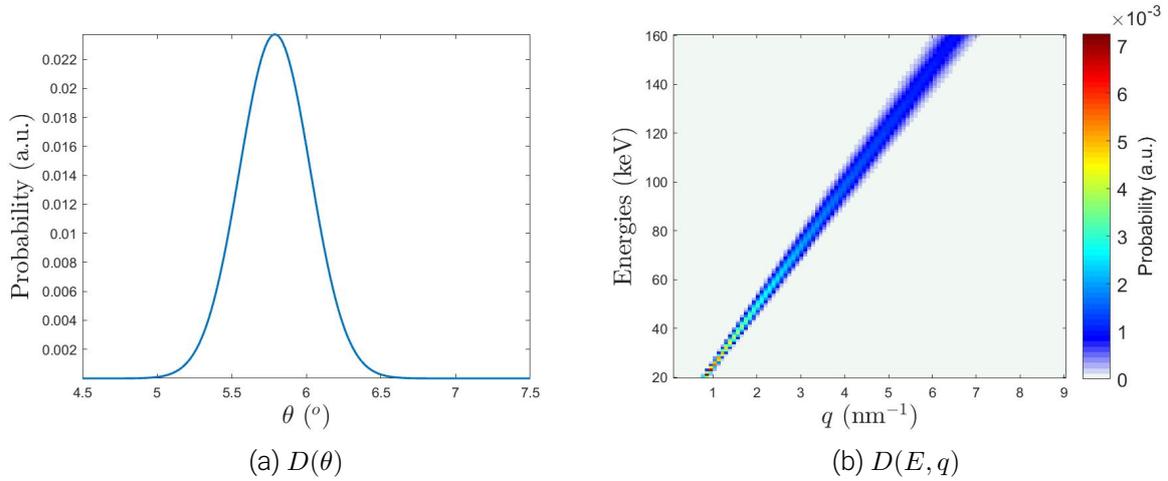


FIGURE 2.15 – Exemple de calcul des distributions pour  $z = 470$  mm et  $p=26$ . (a) représente la distribution angulaire et (b) sa distribution équivalente dans le plan  $(q, E)$ .

Pour tous les couples  $(p, z)$ , où  $I(p, z)$  est non nul, nous pouvons exprimer  $\Theta(p, z)$  et  $\Delta\theta(p, z)$  en utilisant les modèles définis précédemment. La figure 2.14 illustre les deux matrices de champ de vue ainsi obtenues.

### Calcul de la matrice de réponse géométrique

Les matrices  $I(p, z)$ ,  $\Theta(p, z)$  et  $\Delta\theta(p, z)$  nous donnent des informations sur le champ de vue en trois dimensions :  $p$ ,  $z$  et  $\theta$ . Nous allons utiliser la loi de Bragg, pour transformer la dimension des  $\theta$  en deux dimensions  $(E, q)$ , et ainsi construire la matrice de réponse géométrique  $R_c(p, E, z, q)$

1. Pour chaque couple  $(p, z)$  tel que  $I(p, z) \neq 0$ , nous construisons une distribution angulaire gaussienne  $D(\theta)$  de  $\theta$  de moyenne égale à  $\Theta(p, z)$ , et de FWHM égale à  $\Delta\theta(p, z)$  (voir figure 2.15a);
2. En utilisant la loi de Bragg, nous calculons  $D(q, E)$ , la distribution équivalente de la distribution angulaire, dans les plans  $(E, q)$  (voir figure 2.15b).
3. Finalement,

$$\forall(p, E, z, q), R_c(p, E, z, q) = I(z, p) D(E, q)$$

### 2.2.4 Construction de la matrice globale

Une fois que nous avons estimé toutes les composantes du modèle direct (2.2), nous pouvons les utiliser pour construire la matrice globale, afin d'obtenir un modèle plus compact (2.3) :

$$S_{diff}(p, E_{det}) = \sum_{E_{inc}} \left( R_{det}(E_{det}, E_{inc}) \text{Att}(E_{inc}) S_{inc}(E_{inc}) \sum_{z,q} (R_c(p, E_{inc}, z, q) F(z, q)) \right)$$

$$\iff S_{diff}(p, E) = R(p, E, z, q) F(z, q) \quad (2.11)$$

Ainsi, la matrice de réponse globale est obtenue en regroupant tous les termes estimés précédemment :

$$\forall(p, E, z, q), R = R_{det}(E_{det}, E_{inc}) (\text{Att}(E_{inc}) \otimes S_{inc}(E_{inc})) \cdot R_c(p, E, z, q) \quad (2.12)$$

Chacune des quatre dimensions de cette matrice est échantillonnée de la manière suivante :

- les pixels, étant en nombre fini, vont de 1 à 192 par pas de 1;
- le détecteur spectral détecte les énergies entre 20 et 160 keV, c'est pourquoi nous avons choisi d'échantillonner entre 20 et 160 keV par pas de 1 keV;
- comme on peut l'observer sur les figures 2.14, les positions visibles selon  $z$  vont de 470 à 545 mm. De plus, la plaque utilisée pendant la calibration était déplacée par pas de 3 mm, c'est pourquoi, l'axe des  $z$  a été échantillonné également par pas de 3 mm entre 470 et 545 mm;
- la gamme des  $q$  a été choisie de manière à recouvrir toutes les valeurs atteintes par les énergies et les angles. Les angles varient de  $2^\circ$  à  $7^\circ$ , et les énergies vont de 20 à 160 keV. Ainsi, nous avons choisi un échantillonnage régulier sur 100 pas, allant de  $0.14$  à  $9 \text{ nm}^{-1}$ . Cependant, les pics caractérisant les tissus mammaires sains et cancéreux se situent respectivement en  $1.1 \text{ nm}^{-1}$  et  $1.6 \text{ nm}^{-1}$ . C'est pourquoi, dans les représentations, nous ne nous intéressons qu'à l'intervalle  $[0.6 \text{ } 2.1] \text{ nm}^{-1}$ , qui entoure ces deux valeurs particulières.

La matrice  $R$  est une matrice à 4 dimensions, de taille (192, 141, 26, 100). Pour faciliter sa manipulation, nous la "replions" en une matrice à 2 dimensions de taille  $(192 \times 141, 26 \times 100)$ . Ainsi, la matrice  $R$  est similaire à une matrice de passage entre le plan  $(z, q)$  des fonctions de diffusion de l'objet, et le plan des mesures  $(p, E)$ .

### 2.2.5 Conclusion

Dans cette partie, nous avons expliqué en détail le déroulement du procédé de calcul de la matrice système  $R(p \times E, z \times q)$ . Pour cela, nous avons d'abord estimé séparément chacun des termes du modèle direct : la réponse du détecteur, le spectre incident atténué par l'objet et la réponse géométrique du système. Ensuite, nous avons regroupé tous ces termes dans

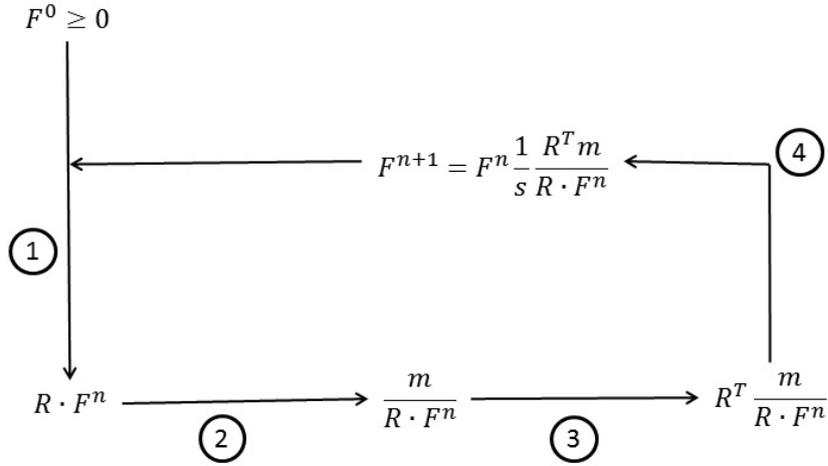


FIGURE 2.16 – Représentation schématique de l'algorithme *Maximum Likelihood Expectation Maximization* (MLEM) : 1) Projection de l'estimation courante, 2) Calcul du ratio entre la mesure et la projection, 3) Rétro-projection pour le calcul d'erreur relative, 4) Mise à jour de l'estimation après normalisation par  $s = \Sigma_{p,E}(R)$

une seule matrice globale qui modélise la chaîne d'acquisition dans son intégralité. Dans la section suivante, nous allons aborder l'algorithme de reconstruction qui utilise cette matrice pour estimer les fonctions de diffusion à partir des spectres diffusés mesurés.

## 2.3 Reconstruction : inversion du modèle direct

L'inversion du modèle 2.3 est un problème mal-posé : la matrice  $R(p \times E, z \times q)$  n'est pas inversible. C'est pourquoi, nous avons utilisé un algorithme itératif MLEM [45] pour calculer une solution  $F(z \times q)$ . Ce type d'algorithme a été utilisé par Fanny MARTICKE pendant ses travaux. Il s'agit d'un algorithme statistique basé sur l'algorithme "Espérance - Maximisation". La figure 2.16 illustre les étapes successives de calcul :

1. Projection de l'estimation précédente  $F^n$  : il s'agit de la mesure qui serait obtenue avec le système, si l'entrée était  $F^n$ .
2. Calcul du ratio entre la mesure  $m$  et de la projection précédente.
3. Rétro-projection du ratio pour obtenir l'erreur relative : il s'agit d'une correction qu'il faudrait apporter à  $F^n$  pour diminuer l'écart entre la mesure et la projection.
4. Mise à jour de l'estimation, après normalisation par  $s = \Sigma_{p,E}(R)$

Finalement, à l'itération  $n + 1$ ,  $F^{n+1}$  s'exprime comme suit :

$$F^{n+1} = F^n \frac{1}{\Sigma_{p,E}(R)} \Sigma_{p,E} \left( \frac{R^T S_{diff}}{\Sigma_{z,q}(R F^n)} \right) \quad (2.13)$$

Pour l'initialisation, nous utilisons une rétro-projection du spectre diffusé :

$$F^0 = R^T S_{diff}$$

L'initialisation étant positive, cela garantit que la solution sera également positive.

Pour notre cas, dix itérations suffisent pour converger à une solution.

## 2.4 Classification : présentation des méthodes

Dans la section précédente, nous avons présenté notre méthode pour obtenir les fonctions de diffusion à partir des spectres diffusés mesurés. Les fonctions de diffusion contiennent des informations sur la structure moléculaire de l'objet, mais elles ne sont pas faciles à interpréter telles quelles. C'est pourquoi, nous avons choisi de mettre un peu en place un procédé de classification des fonctions de diffusion,  $C(z)$ , qui attribue la classe "sain" ou "cancéreux" en fonction de la profondeur  $z$ .

Nous avons envisagé plusieurs modèles de classification. De manière générale, ils peuvent être séparés en deux grands groupes, en fonction de leur type d'apprentissage, non supervisé ou supervisé. Dans les deux cas, le modèle comporte un algorithme qui attribue une classe à chaque réalisation expérimentale. L'apprentissage supervisé nécessite en plus un jeu de données déjà classifiées. La classe à attribuer aux données est alors déduite à partir des données d'apprentissage.

Dans notre cas, nous connaissons les fonctions de diffusion théoriques du tissu sain et du tissu cancéreux. Nous pouvons donc les utiliser comme références, pour mettre au point une méthode de Classification par Référence (CR). Nous avons également décidé de tester plusieurs modèles classiques de classification supervisée :  $k$  plus proches voisins ( $k$ -NN, *k-Nearest Neighbors*), machine à vecteurs de supports (SVM, *Support Vector Machine*) et analyse discriminante linéaire (LDA, *Linear Discriminant Analysis*). De plus, pour ces modèles, nous avons considéré réduire le nombre de dimensions de nos données en utilisant l'analyse en composantes principales (PCA, *Principal Component Analysis*). Nous allons présenter ci-dessous chacun de ces modèles plus en détail.

### 2.4.1 Classification par référence

Pour cette méthode, nous utilisons les fonctions de diffusion de référence, représentées sur la figure 2.17. Nous les notons  $F_a(q)$  et  $F_c(q)$ , respectivement pour le tissu adipeux et le tissu cancéreux. Pour chaque position  $z$ , nous allons comparer  $F(z, q)$  à ces deux fonctions pour attribuer la classe "sain" ou "cancéreux" :

1. En utilisant la méthode des moindres carrés, en chaque position  $z$  nous déterminons les

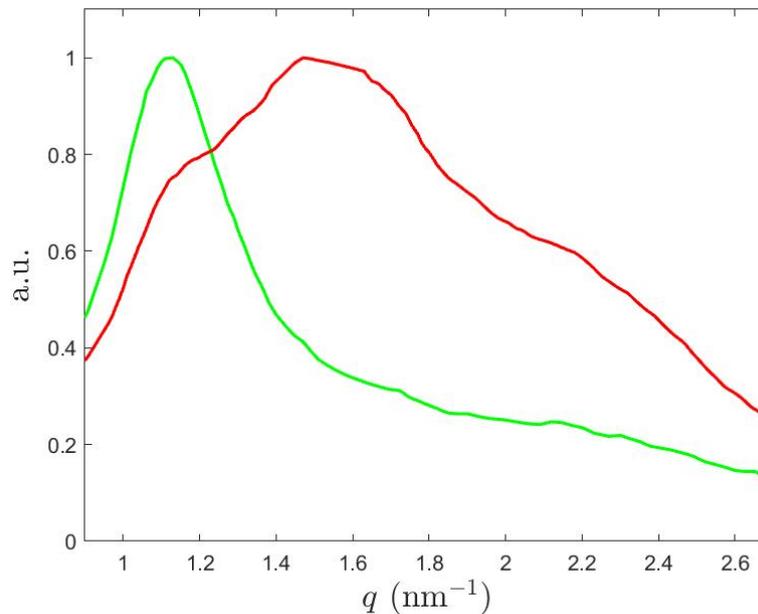


FIGURE 2.17 – Fonctions de diffusion de référence normalisées, pour le tissu adipeux (en vert) et le tissu cancéreux (en rouge).

coefficients  $\alpha$  et  $\beta$  tels que

$$\forall q, F(z, q) \approx \alpha F_a(q) + \beta F_c(q)$$

Nous normalisons ces coefficients par leur somme :  $\alpha + \beta = 1$ .

2. Nous calculons le seuil  $\beta_0$ , en utilisant la méthode d'Otsu [49]. Cette méthode calcule le seuil de séparation optimal entre deux classes, à partir de l'histogramme de  $\beta$ .
3. Pour tous les  $z$ , où  $\beta$  dépasse  $\beta_0$ , on attribue la classe "cancéreux".

### 2.4.2 $k$ plus proches voisins

Le principe de classification  $k$ -NN est illustré sur la figure 2.18a : les carrés bleus et les triangles rouges représentent les données, correspondant à deux classes, fournies en apprentissage. Pour déterminer la classe d'un nouveau point (en vert), l'algorithme va choisir les  $k$  données apprises les plus proches de ce nouveau point. La classe attribuée correspond à la classe majoritaire parmi ces  $k$  plus proches voisins.  $k$  est un paramètre de ce modèle. Dans notre cas, nous avons observé que  $k = 3$  donne de très bon résultats.

### 2.4.3 Machine à vecteur de supports

Le principe de classification SVM est illustré sur la figure 2.18b. Connaissant la distribution de deux classes, l'algorithme va déterminer l'hyperplan, qui permet de séparer au mieux ces deux

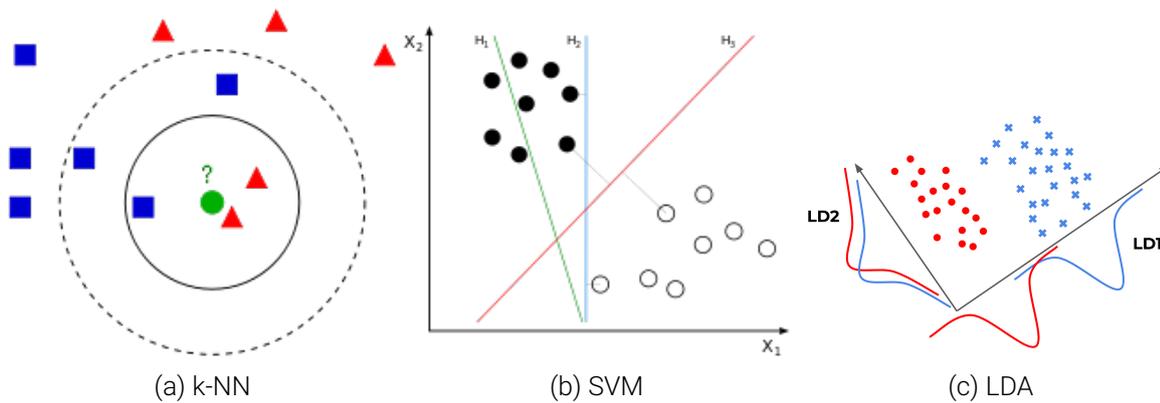


FIGURE 2.18 – Schémas représentant les principales méthodes de classification (a) k-NN [50], (b) SVM [51] et (c) LDA [52]

classes. Il s'agit d'une optimisation où sont maximisées les distances entre la distribution de chaque classe et l'hyperplan. Sur l'exemple présenté sur la figure 2.18b, le plan  $H_3$  est optimal.

#### 2.4.4 Analyse discriminante linéaire

La méthode LDA, illustrée sur la figure 2.18c, est basée sur une projection vectorielle. En connaissant la distribution des données d'apprentissage, avec un nombre total de classes  $N$ , cette méthode détermine un sous-espace de dimension  $N - 1$  optimal, tel que la dispersion des données, projetées sur ce sous-espace, est :

- minimale, pour les données de la même classe;
- maximale, pour les données de classes différentes.

L'algorithme détermine le sous-espace de projection optimal, ainsi que le seuil qui permet de séparer les classes entre elles.

#### 2.4.5 Analyse en composantes principales

Pour toutes les méthodes d'apprentissage supervisé, nous avons envisagé d'appliquer la méthode PCA, avant de classer les données. La PCA est une méthode statistique, utilisées sur les données multi-dimensionnelles. Elle permet de réduire le nombre de dimensions en perdant le moins d'information utile possible.

Dans notre cas, à chaque position physique de l'objet correspond une fonction de diffusion  $F(q)$ . Il s'agit donc d'une observation dont le nombre de dimensions correspond à l'échantillonnage selon  $q$ . Cependant, toutes ces dimensions ne présentent pas la même importance, quant à la discrimination entre le tissu sain et cancéreux. Nous avons remarqué précédemment, que nous pouvons observer la plus grande différence entre ces deux tissus pour les valeurs  $q = 1.1$  et  $q = 1.6 \text{ nm}^{-1}$ . Donc, nous pouvons essayer de classer les données en analysant uniquement les dimensions pertinentes. La PCA permet de les déterminer. En prenant

en entrée un nombre  $n_Z$  de fonctions de diffusion, cette méthode renvoie  $n_Z$  *composantes principales*. Il s'agit de vecteurs, de taille égale à l'échantillonnage en  $q$ , qui sont linéairement indépendants entre eux. Ils sont classés : chaque composante principale a un poids associé. Ce poids correspond au pourcentage de variance des données représenté par la composante principale correspondante.

Dans notre cas, nous avons choisi de garder toutes les composantes principales, qui représentent plus de 1% de variance. En projetant les données sur ces composantes principales choisies, nous obtenons des données de plus faible dimension qui, néanmoins, représentent très bien les données initiales. Parfois, enlever des dimensions superflues permet de simplifier le travail des algorithmes de classification. C'est pourquoi, nous avons choisi de tester si l'application de la PCA avant la classification permettait d'obtenir un résultat plus précis.

#### **2.4.6 Conclusion**

Dans cette section, nous avons passé en revue les différentes méthodes de classification que nous avons sélectionné pour attribuer un type de tissu ("sain" ou "cancéreux") à chaque position de l'objet, en fonction de la fonction de diffusion en ce point. Au total, nous avons testées 7 méthodes :

1. CR
2. k-NN
3. SVM
4. LDA
5. PCA+k-NN
6. PCA+SVM
7. PCA+LDA

Notre objectif est de déterminer la méthode la plus adaptée à notre problématique : détecter de manière fiable un cancer, tout en minimisant la dose délivrée à la patiente. Dans la section suivante, nous allons présenter notre méthode d'estimation de la dose, puis nous présentons une analyse des performances des méthodes de classification en fonction de la dose absorbée.

### **2.5 Estimation de la dose absorbée**

La dose absorbée pendant l'acquisition est un facteur important dans la perspective d'application *in vivo*. En effet, l'augmentation du nombre de photons entraîne une augmentation de la dose absorbée, mais améliore le ratio signal sur bruit. Il est donc intéressant de minimiser la dose absorbée, tout en gardant un nombre de photons suffisant pour obtenir un diagnostic

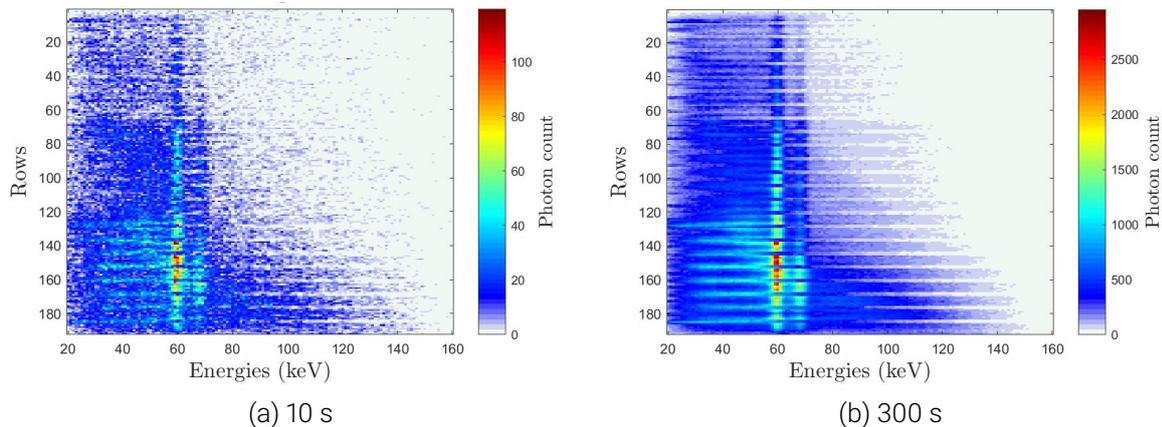


FIGURE 2.19 – Spectres diffusés mesurés. Ils correspondent au même objet, mais le temps d’acquisition varie : les figures (a) et (b) montrent le nombre de photons mesurés en fonction des énergies et des pixels pour un temps d’acquisition de 10 et 300 secondes, respectivement.

fiable. Nous avons donc cherché à évaluer l’efficacité de classification en fonction de la dose. Dans cette section, nous allons présenter quelques définitions concernant la dose absorbée, ensuite nous présenterons notre méthode d’estimation de la dose absorbée par l’objet.

### 2.5.1 Dose absorbée : définitions

Dans le vocabulaire médical, il existe plusieurs types de dose :

- Dose absorbée (ou délivrée) : elle s’exprime en Gray, et est définie comme le rapport entre l’énergie déposée par irradiation, et la masse du volume irradié.
- Dose équivalente : elle s’exprime en Sievert, et est définie comme la dose délivrée, pondérée par un coefficient  $w_R$ . Ce coefficient dépend du type de radiation, et représente les effets de l’exposition sur la santé.
- Dose efficace : elle s’exprime également en Sievert, et correspond à la dose équivalente, pondérée par un coefficient  $w_T$ , qui dépend de l’organe irradiée. En effet, certains organes sont plus sensibles que d’autres, et donc leur exposition présente des risques de santé plus importants.

Dans les cas des rayons X et du sein,  $w_R = 1$  et  $w_T = 0.12$  [53].

Dans les articles scientifiques traitant de l’imagerie médicale, le plus souvent la dose absorbée est utilisée. C’est pourquoi, dans la suite de ce travail, nous utiliserons cette grandeur.

Par définition, la dose absorbée  $D_a$  s’écrit :

$$D_a = \frac{E_D}{m_{irr}} \quad (2.14)$$

où  $E_D$  est l’énergie déposée (en Joules), et  $m_{irr}$  la masse du volume irradiée (en kilogrammes). Pour calculer la dose nous avons donc besoin d’estimer ces deux grandeurs.

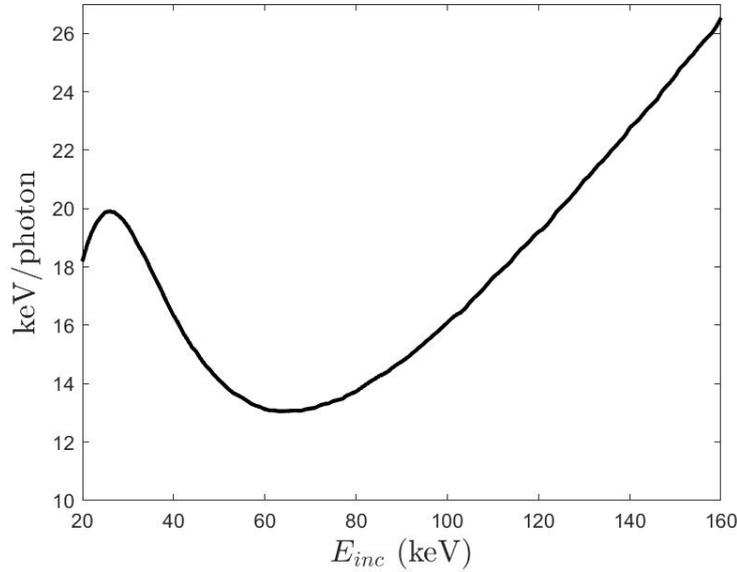


FIGURE 2.20 – Graphique représentant l'énergie moyenne déposée,  $\bar{E}_{ph}$ , (en keV/photon) dans un sein de 6cm d'épaisseur, en fonction de l'énergie du photon

### 2.5.2 Méthode d'estimation

Pour estimer l'énergie déposée dans un objet, nous avons besoin de connaître :

1. L'intensité et la distribution en énergie du spectre incident perçu par l'objet  $S_{inc,obj}$  ;
2. La quantité d'énergie déposée par un photon, en fonction de son énergie et du matériau.  
En effet, en fonction de la densité de l'objet, l'absorption varie en fonction des énergies.

#### Spectre incident perçu par l'objet

$S_{inc,obj}$  représente le spectre incident perçu par l'objet. Il dépend donc de l'angle solide  $\Omega_{obj}$ , avec lequel la source éclaire l'objet.  $\Omega_{obj}$  est défini par l'ouverture de la collimation primaire. Pour  $S_{inc,obj}$ , nous avons normalisé le spectre incident  $S_{inc}$ , calculé dans la section 2.2.2, par l'angle solide. En effet, ce spectre est calculé à partir de la mesure sur la surface d'un pixel du détecteur ME-100. Connaissant la distance entre la source et le détecteur, ainsi que la surface du pixel, nous obtenons ainsi l'angle solide de la mesure.

De plus, nous faisons l'hypothèse que le nombre de photons perçus par l'objet dépend linéairement du temps. En effet, nos mesures sont effectuées avec des temps fixes et il est donc intéressant d'exprimer la dose en fonction du temps de mesure.

Nous notons donc  $S_{inc,norm}$  le spectre incident normalisé par l'angle solide de la mesure et le temps de l'acquisition. Alors le spectre incident perçu par l'objet s'exprime comme :

$$S_{inc,obj} = T \Omega_{obj} S_{inc,norm} \quad (2.15)$$

où  $T$  représente le temps de mesure, et  $\Omega_{obj}$  l'angle solide sous lequel la source éclaire l'objet.

### Estimation de l'énergie déposée

Nous avons estimée l'énergie déposée par photon, en fonction de son énergie et du matériau, en effectuant une simulation Monte-Carlo avec le logiciel Penelope [44]. Nous avons simulés les différentes interactions entre les photons et la matière (effet photo-électrique, diffusion cohérente et incohérente) pour des conditions similaires à un examen mammographie. L'objet simulé avait une épaisseur de 6 cm, et avait la même densité ( $\rho = 0.98 \text{ g.cm}^{-3}$ ) et la même composition chimique moyenne (approx. 50% oxygène, 38% carbone et 12% hydrogène), que le tissu mammaire. Grâce à cette simulation, nous avons obtenu  $\bar{E}_{ph}$ , l'énergie moyenne déposée par un photon, en fonction de son énergie, représentée sur la figure 2.20.

Nous pouvons alors exprimer  $E_D$  comme le produit scalaire entre l'énergie moyenne déposée par photon, et le spectre incident perçu par l'objet :

$$\begin{aligned} E_D &= (\bar{E}_{ph} | S_{inc,obj}) \\ &= T \Omega_{obj} (\bar{E}_{ph} | S_{inc,norm}) \end{aligned} \quad (2.16)$$

### Estimation de la masse irradiée

Pour calculer la masse du volume irradiée, nous utilisons simplement la formule

$$m_{irr} = \rho V_{irr}$$

où  $\rho$  est la masse volumique, et  $V_{irr}$  le volume irradié. Il est défini comme l'intersection entre le volume total de l'objet, et le cône défini par le faisceau de rayons X. Notre objet a une épaisseur de 6 cm et, selon l'axe des rayons X, il est positionné entre  $Z_{min}$  et  $Z_{max}$ . Entre ces deux positions, l'intégralité du faisceau passe par l'objet. Nous sommes donc dans une configuration où le volume irradié correspond au volume du cône, défini par le faisceau, tronqué entre  $Z_{min}$  et  $Z_{max}$ . On peut définir ce cône à partir de l'angle solide, avec lequel la source éclaire l'objet :

$$\begin{aligned} V_{irr} &= \int_{Z_{min}}^{Z_{max}} \Omega_{obj} z^2 dz \\ &= \Omega_{obj} \frac{Z_{max}^3 - Z_{min}^3}{3} \end{aligned} \quad (2.17)$$

où  $\Omega_{obj}$  est constant en tout  $z$ .

### Calcul final de la dose absorbée

Ainsi, grâce aux expressions 2.16 et 2.17, l'équation 2.14 devient :

$$\begin{aligned}
 D_a &= \frac{T I \Omega_{obj} (\bar{E}_{ph} | S_{inc, norm})}{\rho \Omega_{obj} \frac{Z_{max}^3 - Z_{min}^3}{3}} \\
 &= \frac{3 T I (\bar{E}_{ph} | S_{inc, norm})}{\rho (Z_{max}^3 - Z_{min}^3)}
 \end{aligned} \tag{2.18}$$

$\Omega_{obj}$ , indépendant de  $z$ , se simplifie dans le calcul.

Avec les paramètres de notre système, nous estimons ainsi la dose absorbée par un objet similaire à un sein de 6 cm d'épaisseur, à 1.7 mGy par seconde.

## 2.6 Étude de l'influence de la dose sur la classification

Dans les deux sections précédentes, nous avons présenté plusieurs méthodes de classification des fonctions de diffusion, ainsi que notre méthode d'estimation de la dose absorbée. Notre objectif final est de choisir la classification la plus efficiente vis-à-vis de la dose absorbée. Afin de déterminer la méthode de classification la plus performante, nous avons réalisé une étude en simulation sur l'influence de la dose sur les performances des méthodes de classification. Nous allons commencer par décrire les métriques, que nous avons utilisées pour mesurer les performances des classifieurs. Puis, nous présentons les conditions de simulation. Finalement, les performances des classifieurs en fonction de la dose seront évaluées.

### 2.6.1 Métriques

Notre objectif final est d'obtenir le diagnostic le plus fiable possible, tout en gardant la dose absorbée suffisamment faible. C'est pourquoi, nous avons choisi d'évaluer les facteurs de mérite suivants, en fonction de la dose absorbée estimée :

- la sensibilité, définie comme :

$$\frac{TP}{TP + FN}$$

- la spécificité, définie comme :

$$\frac{TN}{TN + FP}$$

où les notations représentent les vrai positif (TP, *True Positive*), les vrai négatif (TN, *True Negative*), les faux positif (FP, *False Positive*) et les faux négatif (FN, *False Negative*). Dans le cas, où le fantôme est composé de tissus adipeux, fibreux et cancéreux, nous considérons les deux premiers comme négatifs, car leur présence dans un sein est normale, et le dernier comme positif.

La sensibilité et la spécificité sont comprises 0 et 1, et mesurent la confiance que l'on peut

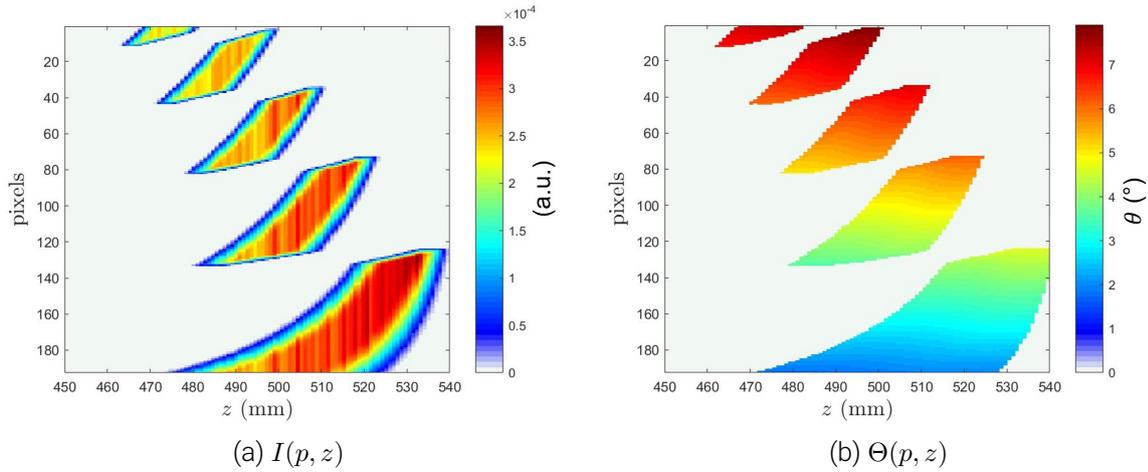


FIGURE 2.21 – Matrices de champ de vue du système simulé

accorder aux résultats de la classification. Si le système est parfaitement sensible (sensibilité = 1), alors la présence d'un cancer entraîne toujours une détection i.e. il n'y a jamais de faux négatif. Si le système est parfaitement spécifique (spécificité = 1), alors une détection positive signifie qu'un cancer est effectivement présent, i.e. il n'y a jamais de faux positif.

## 2.6.2 Simulation

Pour tester les différentes méthodes de classification, nous avons réalisé une simulation complète de notre système. Tout d'abord, nous calculons la matrice globale de réponse du système par simulation. Pour cela, la matrice de réponse géométrique est obtenue grâce à un code de lancer de rayons entre la source, le volume d'inspection et le détecteur. Ce code utilise le logiciel de CAO BRL-CAD © [54] pour dessiner précisément la géométrie du système, et particulièrement le collimateur secondaire, et pour lancer les routines de lancer de rayon. Le reste de la simulation est réalisé analytiquement en s'appuyant sur le modèle direct présenté dans la partie 2.1.2 (p. 39).

Les matrices de champ de vue, obtenues par cette simulation, sont présentées sur la figure 2.21. Ensuite, nous avons créé dix fantômes, de 6 centimètres d'épaisseur, composés en différentes proportions de tissus adipeux, fibreux et cancéreux. Ces fantômes sont représentés sur la figure 2.22. Le déroulement de la simulation, sur l'exemple du fantôme n° 5 et d'un temps d'acquisition de 5 secondes, est présentée sur la figure 2.24 :

1. À partir des fonctions de diffusion de référence, représentées sur la figure 2.23, et de la composition des fantômes, nous construisons les fonctions de diffusion pour chaque fantôme  $F_i(z, q)$ ,  $i \in \{1, \dots, 10\}$  (figure 2.24a)
2. Connaissant les densités et épaisseurs respectives de chaque matériau de chaque fantôme, nous simulons leurs spectres incidents atténués pour des temps d'acquisition, notés  $T_{acq}$ , allant de une à cinquante secondes. La figure 2.24b montre le spectre inci-

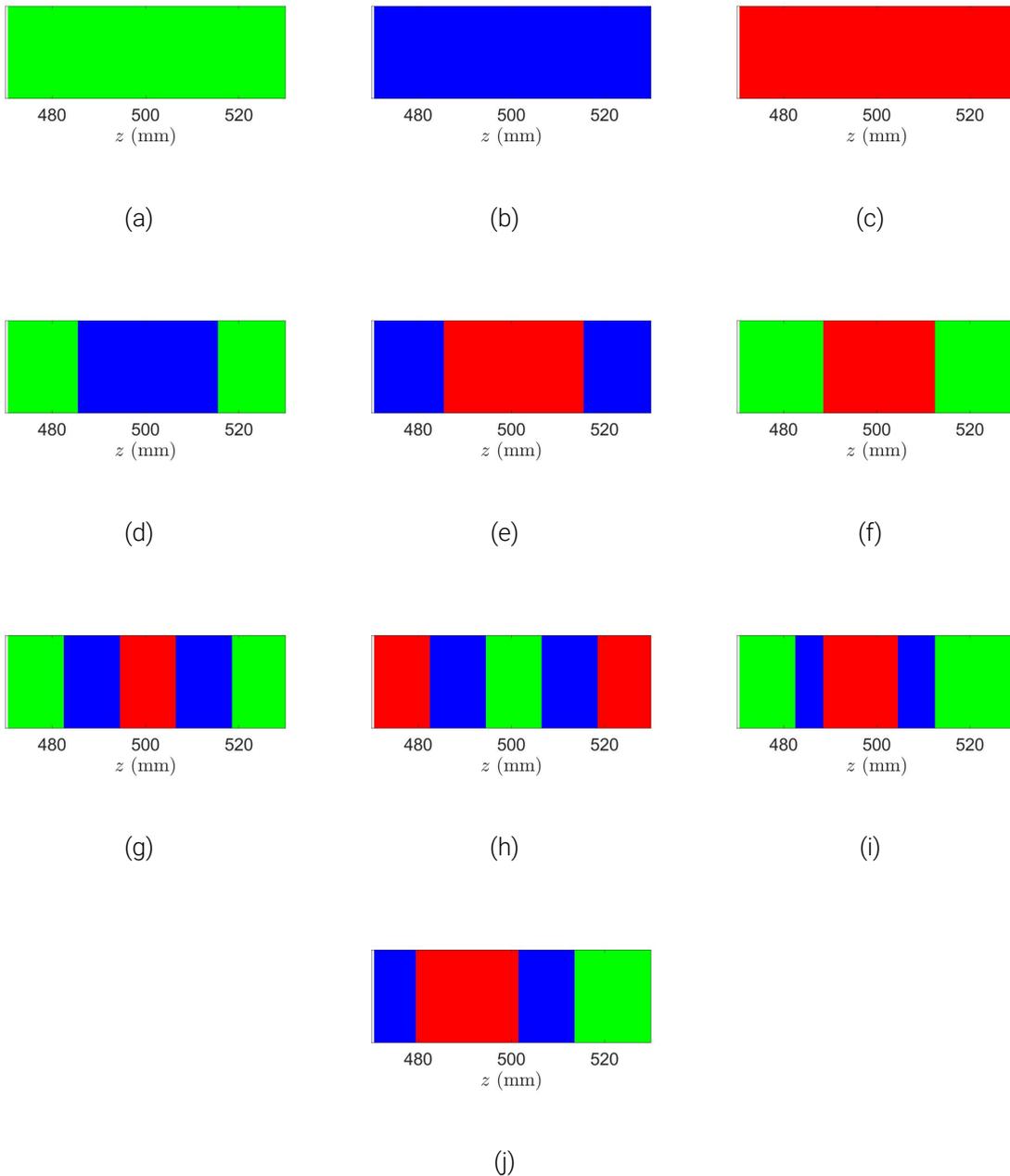


FIGURE 2.22 – Composition des 10 fantômes utilisés pour la simulation. Ils ont tous une épaisseur de 6 cm. Les couleurs représentent le type de tissu : adipeux (vert), fibreux (bleu) ou cancéreux (rouge).

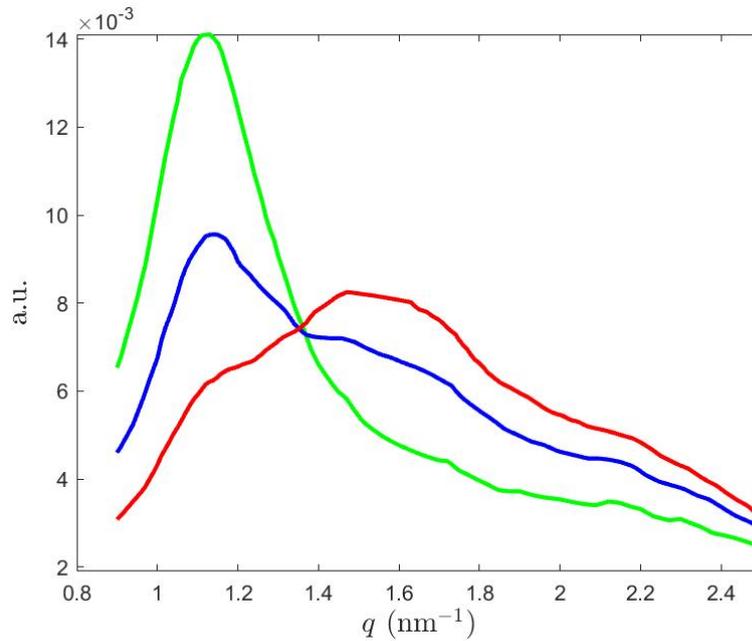


FIGURE 2.23 – Fonctions de diffusion, normalisées par leur intégrale sur l'intervalle représenté, pour les tissus adipeux (vert), fibreux (bleu) et cancéreux (rouge).

dente atténué par le fantôme n° 5 pour  $T_{acq} = 5$  secondes.

3. Puis nous calculons les matrices systèmes,  $R_{t,i}(p, E, z, q)$ , pour chaque fantôme et chaque temps d'acquisition, en utilisant les spectres incidents atténués, la DRM du détecteur Imadif et la matrice de réponse géométrique, calculée précédemment. Ainsi, pour chaque temps d'acquisition et pour chaque fantôme, nous pouvons simuler les spectres diffusés :

$$S_{diff,t,i} = R_{t,i} F_i$$

Nous avons appliqué un bruit de Poisson sur les spectres diffusés ainsi obtenus (figure 2.24c).

4. En ayant ces spectres diffusés simulés, nous les reconstruisons en utilisant l'algorithme MLEM, décrit précédemment dans la partie 2.3 (p. 52). Ainsi, pour chaque fantôme et pour chaque temps d'acquisition, nous avons une fonction de diffusion reconstruite  $F_{i,t,rec}(z, q)$  (figure 2.24d)

### 2.6.3 Classification en fonction de la dose

Nous disposons de fonctions de diffusion  $F_{i,t,rec}(z, q)$  pour dix fantômes, et pour des temps d'acquisition différents. Ces temps ont été convertis en dose absorbée en utilisant la méthode de calcul présentée dans la partie §2.5.

Pour tester les méthodes à apprentissage supervisé, nous avons séparé les données en deux parties, pour l'apprentissage et la validation, respectivement. Nous avons choisi d'utiliser la

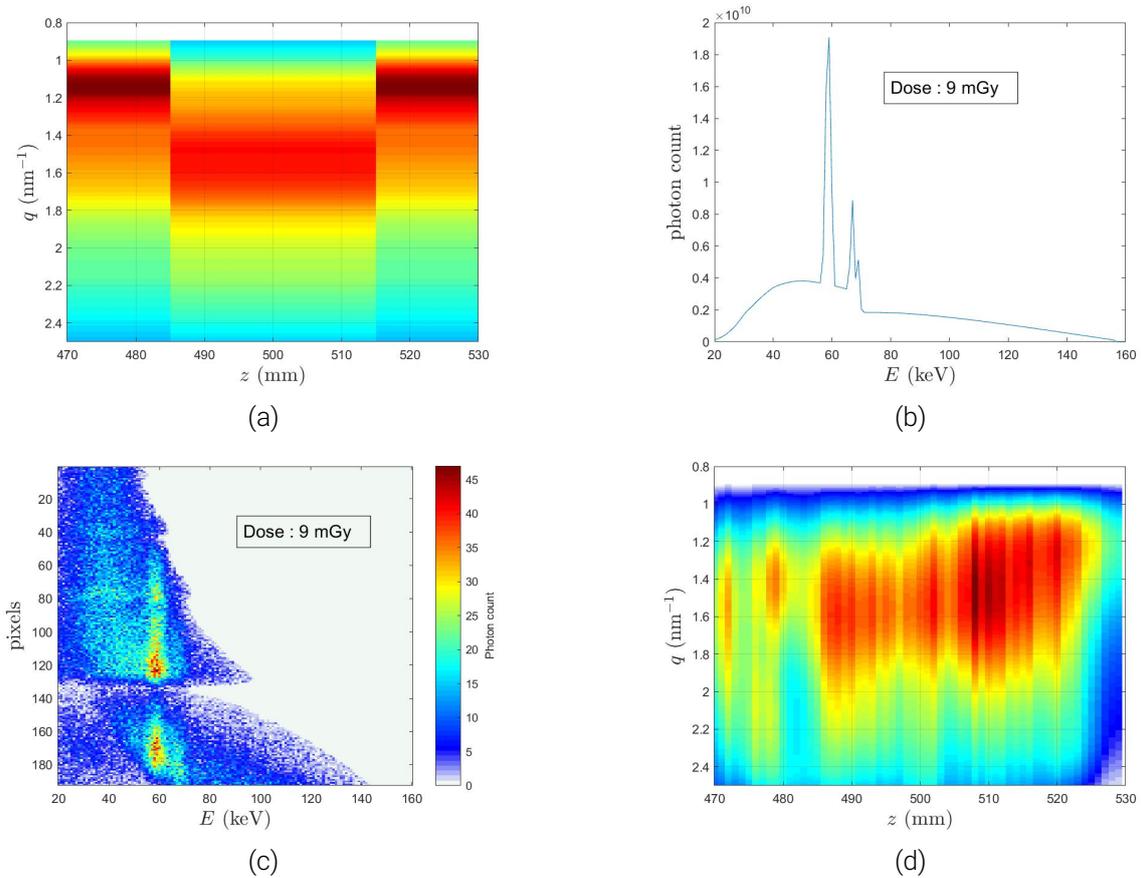


FIGURE 2.24 – Déroulement de la simulation, sur l'exemple du fantôme n° 5. (a) Fonction de diffusion théorique, (b) Spectre incident atténué simulé, (c) Spectre diffusé simulé et (d) Fonction de diffusion reconstruite. Le temps d'acquisition est de 5 secondes, ce qui correspond à une dose de 9 mGy.

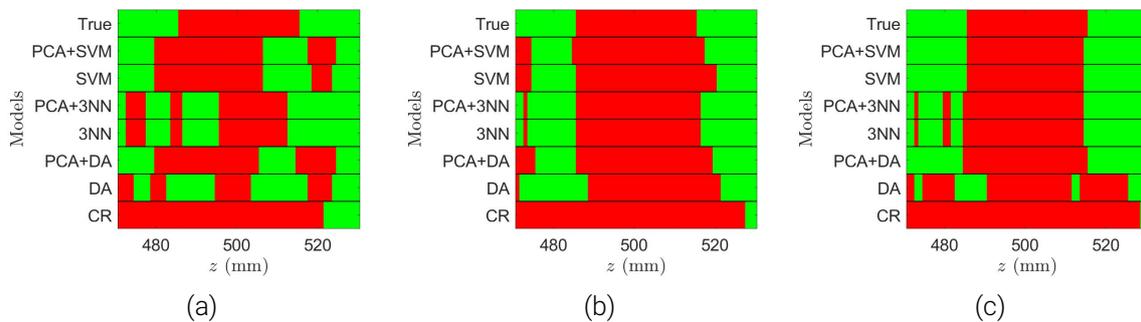


FIGURE 2.25 – Exemple de classifications du fantôme n° 5, obtenues avec les différentes méthodes, pour les doses absorbées de 2 mGy (a), 9 mGy (b) et 34 mGy (c). Les couleurs correspondent aux classes : "sain" (vert) et "cancéreux" (rouge).

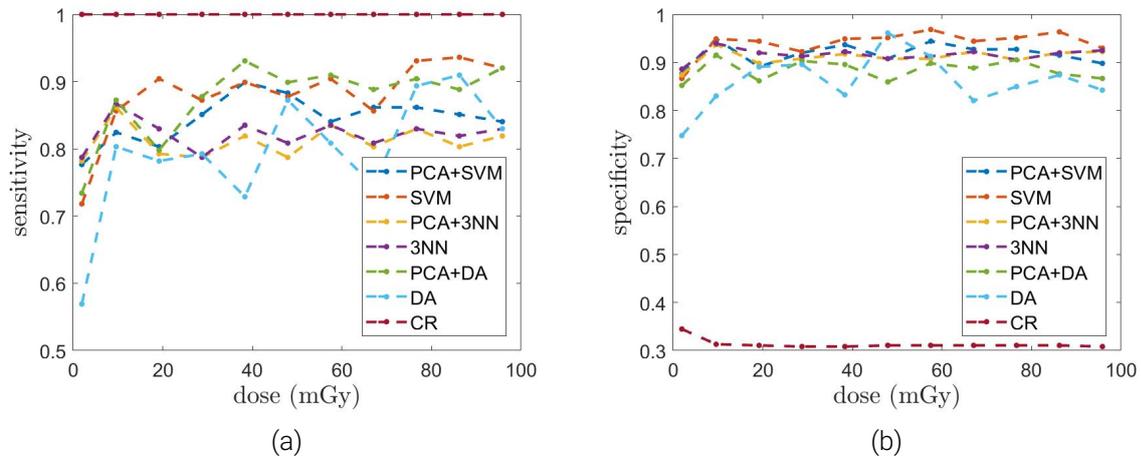


FIGURE 2.26 – Sensibilité (a) et spécificité (b) obtenues pour chaque méthode de classification, en fonction de la dose (en mGy).

méthode de validation croisée : pour classifier les fonctions de diffusion du fantôme  $i$ , nous avons utilisé tous les autres fantômes  $j \in \{1, \dots, 10\}, j \neq i$ , pour l'apprentissage.

Ainsi, pour chaque dose et chaque fantôme, nous avons classifié les données avec les différentes méthodes, et avons comparé les classifications obtenues à la vérité terrain. La figure 2.25 représente les classifications obtenues sur le fantôme n°5 pour 3 doses différentes. Chaque ligne représente le résultat pour une méthode, la vérité terrain étant représentée sur la première ligne.

En comparant l'ensemble des fantômes en chaque  $z$  à la vérité terrain pour chaque dose et chaque méthode, nous avons calculé les TP, les FP, les TN et les FN. Puis, nous en avons déduit la sensibilité et la spécificité de chaque méthode de classification, en fonction du temps. Elles sont représentées sur les figures 2.26.

Nous remarquons que la méthode de classification par référence présente une sensibilité parfaite quelle que soit la dose, cependant elle a une spécificité bien trop faible : cette méthode attribue facilement le label "cancéreux". Ainsi, aucune tumeur n'est manquée, mais au prix d'un grand nombre de faux positifs, ce qui n'est pas très intéressant.

Pour les autres méthodes, nous observons que la sensibilité croît assez rapidement sur l'intervalle de 2 à 10-15 mGy, puis la croissance devient plus faible, voire stagnante, comme dans le cas de la méthode k-NN. En particulier, avec ou sans PCA, la méthode des plus proches voisins stagne à une sensibilité proche de 0.8. Il semblerait que l'utilisation de la PCA n'a pas d'impact significatif sur ce modèle de classification. Dans le cas de la méthode LDA, l'utilisation de la PCA semble intéressante : la méthode "PCA+DA" a une courbe de sensibilité plus lisse, et est globalement supérieure à la méthode "NA", qui n'utilise pas de PCA. À l'inverse, dans le cas de la méthode SVM, l'utilisation de la PCA dégrade la sensibilité et la spécificité.

Globalement, pour toutes les méthodes la spécificité varie moins en fonction de la dose, que la sensibilité. On observe peu d'écart entre les différentes méthodes au niveau de la spécificité,

cependant la méthode SVM présente de bonnes performances à la fois en termes de sensibilité et de spécificité.

On remarque également que à partir d'une dose de 10 mGy, les améliorations apportées par l'augmentation de la dose deviennent faibles. Il n'y a donc aucun intérêt à utiliser des doses élevées.

#### **2.6.4 Analyse et discussion**

Dans cette section, nous avons présenté une étude en simulation sur l'effet de la dose absorbée sur les performances de classification. Nous avons utilisé le taux d'erreur de classification, la sensibilité et la spécificité pour mesurer les performances des méthodes de classification présentées dans la section §2.4. Par simulation, nous avons obtenu des spectres diffusés et atténués pour différents temps d'acquisition, pour dix fantômes composés de différentes proportions de tissu adipeux, fibreux et cancéreux. Puis, nous avons reconstruit les fonctions de diffusion et les avons classifiées. En comparant les différentes méthodes de classification, nous avons constaté que la classification non-supervisée est très peu spécifique. Les différentes méthodes à classification supervisée présentent des résultats proches, toutefois la méthode SVM semble être légèrement meilleure, même à faible dose. C'est donc cette méthode de classification que nous allons privilégier par la suite.

Cependant, dans cette étude en simulation, il est important de noter que, pour chaque fantôme, la simulation des spectres diffusés et leur reconstruction ont été réalisées avec la même matrice. Ceci ne sera pas le cas dans une étude expérimentale, car nous ne pouvons pas estimer exactement cette matrice de réponse système. Ainsi, les résultats obtenus dans cette étude représentent une estimation optimiste des performances que l'on peut espérer expérimentalement.

### **2.7 Système : conclusion**

Dans ce chapitre, nous avons présenté notre système de diagnostic du cancer du sein, exploitant la diffusion cohérente des rayons X. Nous avons décrit le montage expérimental et ses différents composants. Puis, nous avons modélisé la relation entre les spectres diffusés mesurés et les fonctions de diffusion. Nous avons estimé toutes les parties de ce modèle, afin de calculer les fonctions de diffusion à partir des mesures, en utilisant l'algorithme itératif MLEM. Puis, nous avons présenté les différentes méthodes de classification à notre disposition, qui permettent d'attribuer un diagnostic, "sain" ou "cancéreux", en fonction des fonctions de diffusion. Nous avons également étudié la dose absorbée par l'objet pendant l'acquisition et avons présenté une étude en simulation, afin de, d'une part, comparer les performances des différents modèles de classification en fonction de la dose absorbée, et, d'autre part, de déterminer le modèle le plus efficace. D'après les résultats de cette étude, le classifieur SVM donne les meilleurs résultats. C'est donc cette méthode de classification que nous avons choisi.

sie d'utiliser dans nos expériences.

Dans le chapitre suivant, nous allons présenter les résultats obtenus expérimentalement sur différents fantômes, et discuter les performances réelles de notre système.



# RÉSULTATS

---

Dans ce chapitre, nous présentons les résultats obtenus avec le système d'acquisition, de traitement de données et de classification, décrit précédemment. Dans la section 3.1, une comparaison entre la méthode de Damien BARBES et la nôtre est effectuée sur le fantôme plastique, en Polystyrène (PS) et nylon, utilisé par Damien BARBES. Ainsi nous montrons la validité de notre méthode face à celle de Damien BARBES qui donne des résultats très limités.

Puis, nous validons notre méthode sur des fantômes de sein plus réalistes. Dans la section 3.2, nous présentons les différentes étapes, qui nous ont menés à la fabrication des fantômes de viande bovine. Les résultats obtenus sur ces fantômes sont présentés dans la sections 3.3.

### 3.1 Fantôme plastique

Dans cette section, nous allons étudier les résultats obtenus à partir des acquisitions réalisées avec le fantôme plastique représenté sur la figure 3.1. Le bloc principal est composé de PS, avec sept inserts en nylon, dont les diamètres font 2, 1.5, 1, 0.7, 0.5, 0.3 et 0.2 cm. Nous utilisons comme référence les fonctions de diffusion du nylon et du PS obtenues par Brian KING [42]. Elles sont présentées sur la figure 3.2. Ces matériaux ont été choisis car, par la position de leurs maxima, leurs fonctions de diffusion sont similaires aux fonctions de diffusion des tissus mammaires. Ainsi, le PS représente le tissu mammaire sain, et le nylon représente le tissu mammaire cancéreux.

Tout d'abord, nous présenterons les conditions d'acquisition et mesures obtenues, puis les fonctions de forme obtenues après la reconstruction, ainsi que les résultats de la classification.

#### 3.1.1 Mesures expérimentales

Pour se rapprocher des conditions expérimentales correspondant aux acquisitions de Damien BARBES, nous avons utilisé les paramètres décrits dans la table 3.1. Depuis les expériences menées par Damien, des blindages de plomb supplémentaires ont été installés, afin de limiter le

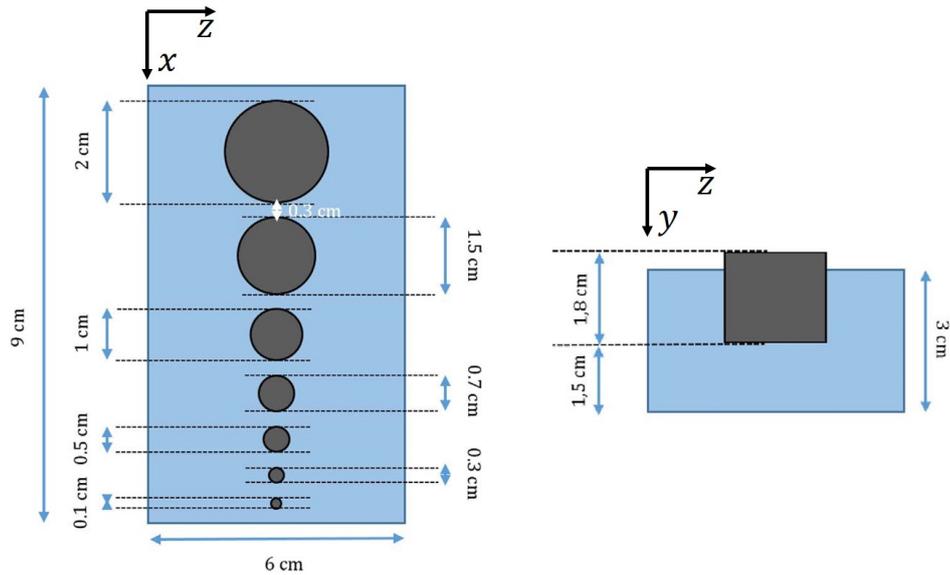


FIGURE 3.1 – Schémas du fantôme dans les plans  $(z, x)$  et  $(z, y)$ , respectivement à gauche et à droite : le support (en bleu) est fabriqué en PS et les inserts (en gris) sont en nylon

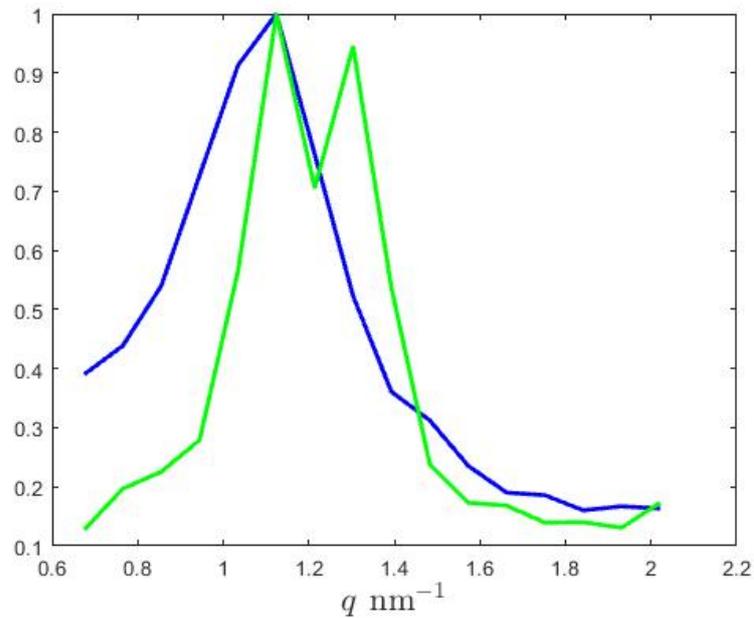


FIGURE 3.2 – Fonctions de diffusion normalisées du nylon et du PS, respectivement en vert et en bleu [42].

Tension du tube X	Intensité du courant	Temps d'acquisition	Pas selon $x$
100 kV	6.4 mA	5 min / position $x$	1 mm

TABLE 3.1 – Paramètres expérimentaux pour l'acquisition des spectres diffusés sur fantôme plastique

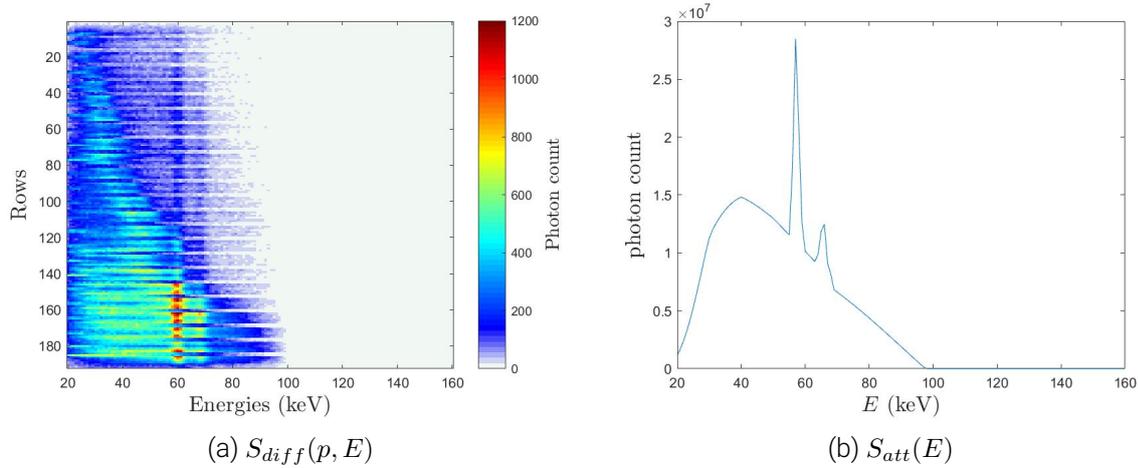


FIGURE 3.3 – Spectre expérimental diffusé (a) et spectre expérimental transmis (b) par le fantôme pour  $x = 15$  mm. Cette position correspond au centre de l’insert de 20 millimètres de diamètre

rayonnement diffusé ambiant.

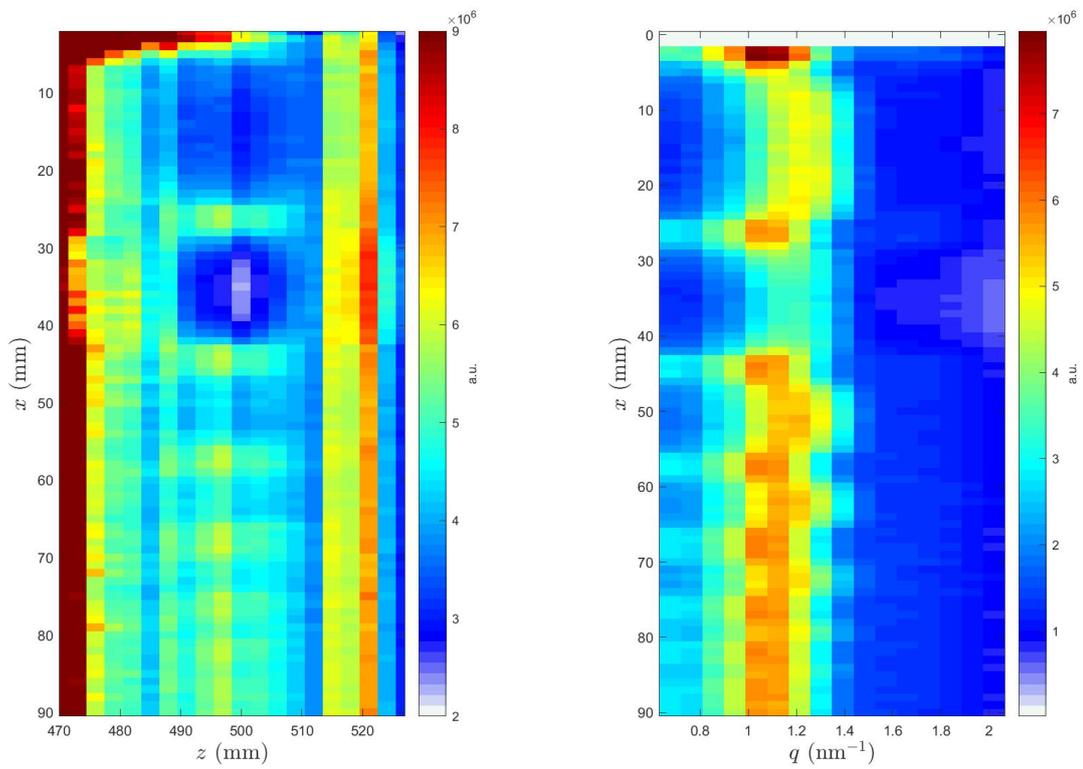
Le générateur de rayons X a été opéré à 100 kV, avec une intensité de 6.4 mA et de 1 mA pour la mesure de spectres diffusés et atténués, respectivement. Les acquisitions de spectres diffusés et transmis ont été réalisées en translatant le fantôme selon  $x$  par pas de 1 millimètre. Pour chaque position  $x$ , le temps d’acquisition des spectres diffusés et atténués était respectivement de 5 minutes et 1 minute. Sur la figure 3.3 sont présentés les spectres diffusés et transmis pour la position  $x = 15$  mm. Elle correspond au centre de l’insert de 20 millimètres de diamètre.

### 3.1.2 Reconstruction

Après la reconstruction avec l’algorithme *Maximum Likelihood Expectation Maximization* (MLEM), nous obtenons des fonctions de diffusion en fonction des positions  $(x, z)$  du fantôme :  $F(x, z, q)$ . Pour faciliter la visualisation de cette matrice tri-dimensionnelle, nous présentons 2 projections sur la figure 3.4 :

- La sous-figure (a) représente la fonction de diffusion en  $q = 1.03 \text{ nm}^{-1}$ , correspondant à la position des maxima du nylon et du PS.
- La sous-figure (b) représente la fonction de diffusion à la position  $z = 497$  mm, qui correspond à la verticale centrale du fantôme.

Sur la figure 3.4a, nous pouvons observer aux positions extrêmes selon  $z$ , des artefacts dûs aux effets de bords. En effet, aux bords, on s’attend d’avantage à une diminution d’intensité. Par ailleurs, aux positions  $(x, z)$  qui correspondent aux inserts de nylon, nous observons une diminution d’intensité. Sur la figure 3.4a, nous observons un déplacement des maxima en fonction de  $x$  : pour les positions  $x$  correspondant au PS, le maximum local correspond à  $q \approx 1.1 \text{ nm}^{-1}$ ;



(a)  $F(x, z, q = 1.03 \text{ nm}^{-1})$

(b)  $F(x, z = 497 \text{ mm}, q)$

FIGURE 3.4 – Projections de la fonction de diffusion  $F(x, z, q)$  sur (a)  $z = 497 \text{ mm}$  et (b)  $q = 1.03 \text{ nm}^{-1}$

pour les positions  $x$  correspondant au nylon, les fonctions de diffusion sont maximales pour  $q \in [1-1.4] \text{ nm}^{-1}$ . Ceci est cohérent avec les fonctions de diffusion théoriques, présentées sur la figure 3.2. La résolution selon  $q$  n'est pas suffisante pour séparer les deux pics du nylon, ils apparaissent donc comme un seul maximum avec une largeur à mi-hauteur plus importante. On remarquera que sur les deux figures 3.4a et 3.4b, les changements d'intensité sont cohérents avec la position des inserts de nylon.

### 3.1.3 Classification

À partir des fonctions de diffusion  $F(x, z, q)$  nous souhaitons obtenir une classe, "positif" pour le nylon et "négatif" pour le PS, en fonction des positions  $(x, z)$ . Pour cela, nous avons utilisé le modèle machine à vecteurs de supports (SVM, *Support Vector Machine*) (voir §2.4). Utilisant l'apprentissage supervisé, ce modèle requiert des données classifiées en entrée. Or, dans cette expérience, nous n'avons qu'un seul fantôme. Nous avons choisi de classer une partie des positions  $(x, z)$ , en utilisant les positions restantes en apprentissage. Cette démarche ne conviendrait pas en conditions cliniques. Néanmoins, elle nous permet de réaliser une première validation expérimentale de notre méthode de traitement des données.

Le fantôme ayant une épaisseur de 6 centimètres selon  $z$ , cela correspond à 20 points de mesure. Au total, nous avons  $N = 91 \times 20 = 1820$  fonctions de diffusion, pour chaque couple  $(x, z)$  du fantôme, à classer. Ainsi, pour étudier statistiquement les performances de la classification, nous avons suivi la démarche suivante :

1. Nous avons réalisé 100 séparations aléatoires des  $N$  fonctions de diffusion en deux parties, notées  $F_1$  et  $F_2$ , de taille égale  $N/2$ .
2. Nous avons classifié  $F_1$ , en utilisant  $F_2$  en apprentissage. Puis, nous avons classifié  $F_2$ , en utilisant en apprentissage  $F_1$ . Ainsi, pour chacune des 100 séparations aléatoires, la totalité des positions du fantôme a été classifiée. La figure 3.5 montre la vérité terrain (fantôme échantillonné pour correspondre à l'échantillonnage en  $x$  et  $z$  des données reconstruites) et une classification du fantôme, obtenue pour une des 100 séparations aléatoires.
3. Pour chaque séparation aléatoire, nous avons calculé la sensibilité et la spécificité, en comparant la classification obtenue à la vérité terrain : à tout point  $(x, z)$ , nous avons attribué le label vrai positif (TP, *True Positive*), vrai négatif (TN, *True Negative*), faux positif (FP, *False Positive*) ou faux négatif (FN, *False Negative*). En sommant sur toutes les positions  $(x, z)$ , nous avons obtenu le nombre total de TP, TN, FP et FN pour chaque séparation aléatoire, puis nous en avons déduit la sensibilité et la spécificité.

Nous avons ainsi obtenu 100 valeurs de sensibilité et spécificité. Leurs moyennes et écart-types sont donnés dans la table 3.2.

Nous pouvons observer sur la figure 3.5, que la classification obtenue est très proche de la vérité terrain. Les valeurs statistiques, présentées dans la table 3.2 le confirment : nous avons

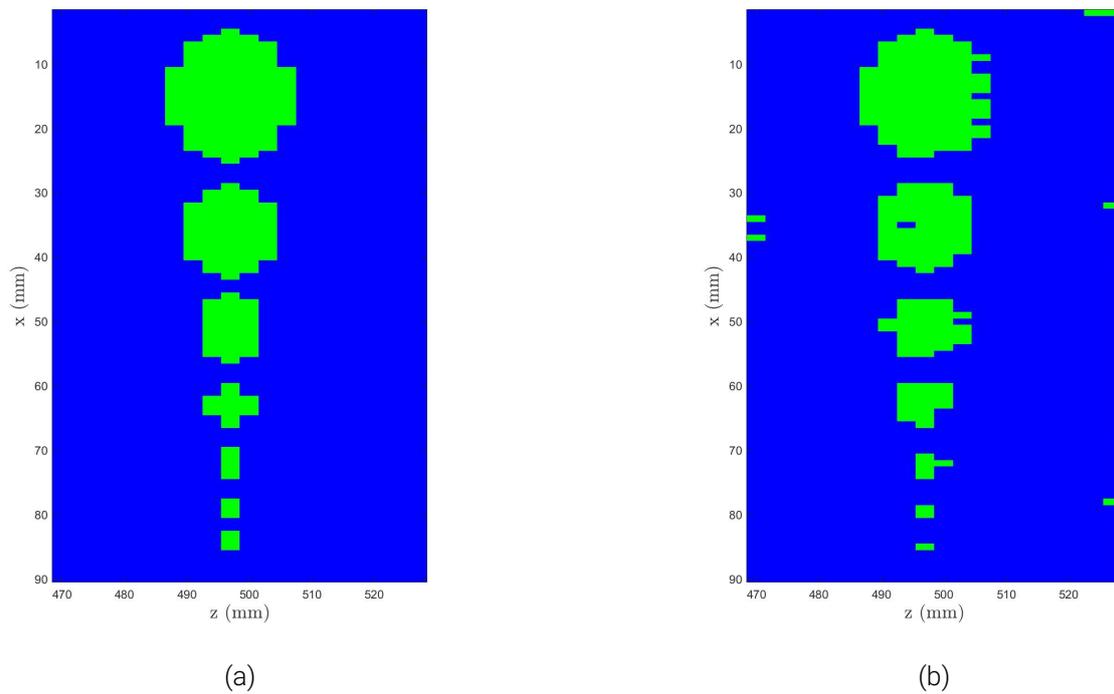


FIGURE 3.5 – Classification du fantôme. (a) représente la vérité terrain, (b) représente le résultat de la classification SVM pour une des séparations aléatoires. Le bleu représente le PS, et le vert représente le nylon

une sensibilité et une spécificité très proches de 1 pour toutes les séparations aléatoires des données.

Dans cette expérience, le temps d'acquisition était de 5 minutes par position, ce qui donne des données très peu bruitées. Par contre, la dose absorbée était excessivement importante et équivaut à environ 500 mGy, en utilisant la méthode de calcul présentée dans la section §2.5.

### 3.1.4 Conclusion

Plusieurs observations peuvent être faites à l'issue de cette expérience :

	Sensibilité	Spécificité
moyenne	0.906	0.984
écart-type	0.019	0.002

TABLE 3.2 – Sensibilité et spécificité obtenues pour la classification du fantôme avec le modèle SVM

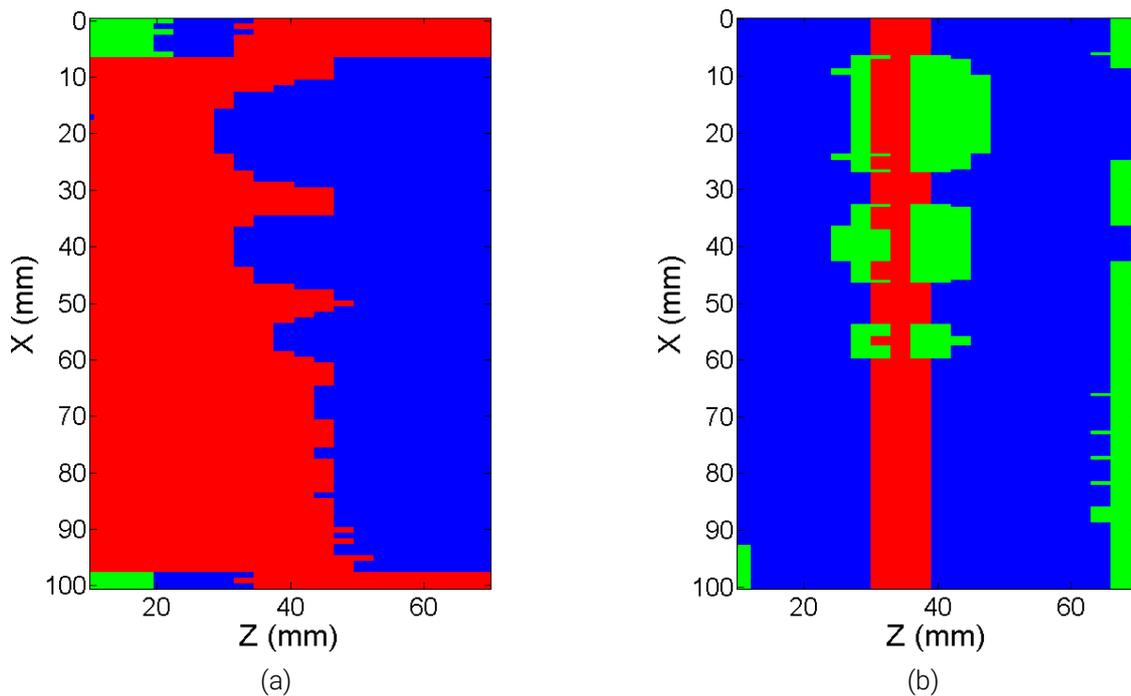


FIGURE 3.6 – Rappel des résultats obtenus par Damien BARBES, respectivement expérimentalement (a) et par simulation (b) [18]. Le bleu représente le PS, le vert représente le nylon, et le rouge représente le *Polymethyl Methacrylate* (PMMA), utilisé comme matériau de référence

1. En comparant la classification obtenue, présentée sur la figure 3.5, et les derniers résultats de Damien BARBES (rappel présenté sur la figure 3.6), nous observons que les classes sont nettement mieux attribuées et que des inserts de plus faible diamètre sont discernables. Nous pouvons donc en déduire, que le traitement des données permet une nette amélioration des résultats. Par ailleurs, l'environnement expérimental est mieux maîtrisé depuis les premières expériences. Notamment, des blindages supplémentaires ont été installés, afin de réduire le rayonnement diffusé ambiant, et donc le bruit. Ces modifications ont sans aucun doute contribué à l'amélioration des résultats.
2. Nous avons choisi de reproduire au mieux les conditions expérimentales de Damien BARBES, notamment avec un temps d'acquisition très long. Les mesures étaient donc très peu bruitées. En diminuant le temps d'acquisition, nous risquons d'obtenir des résultats moins bons. Cependant, l'expérience réalisée nous donne une indication sur les performances maximales du système.
3. Nous avons utilisé une partie des données issues du même fantôme et de la même expérience pour classifier la partie restante des données. Cette démarche semble trop favorable par rapport à une application clinique, où, éventuellement, les résultats d'une patiente seront classifiés en utilisant en apprentissage les examens réalisés sur d'autres personnes.
4. Globalement, les résultats obtenus sur ce fantôme sont très encourageants. Le fantôme

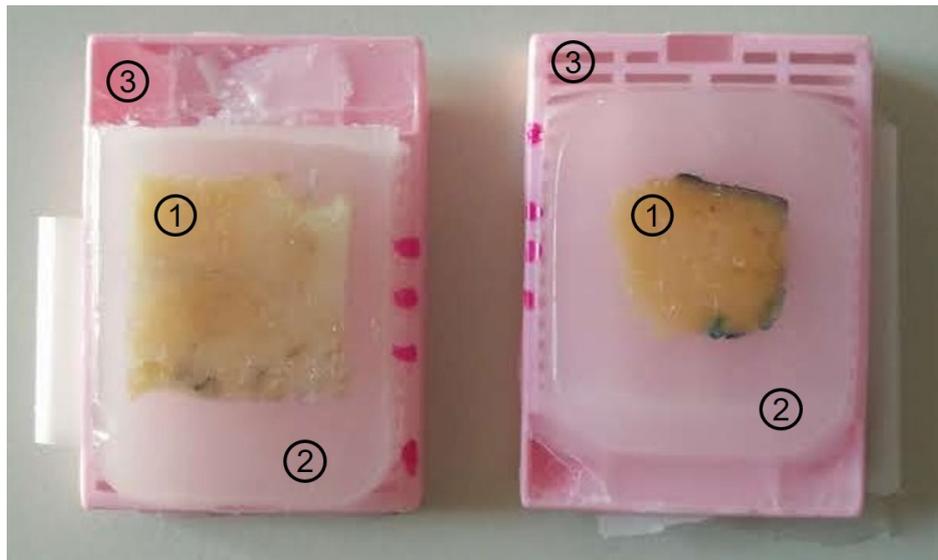


FIGURE 3.7 – Photographie de deux échantillons fournis par le Centre Hospitalier Universitaire Grenoble Alpes (CHUGA) : sain (à gauche) et cancéreux (à droite). Les numéros indiquent les différentes parties : l'échantillon de tissu (1) est entouré de paraffine (2), et fixé sur une cassette plastique (3). Chaque cassette mesure 3 cm × 4 cm.

utilisé est constitué de matériaux plastiques, qui ont été choisis pour imiter au mieux les fonctions de diffusion des tissus mammaires. Cependant, les matériaux plastiques sont bien plus structurés que les tissus biologiques, et nous ne savons pas si les mêmes performances peuvent être atteintes sur un sein réel.

Ces trois derniers points constituent trois limitations de cette expérience : une dose trop importante, un apprentissage trop favorable et des matériaux insuffisamment réalistes. Dans la suite de nos travaux, nous avons donc visé à corriger ces limitations. Notamment, nous avons cherché à obtenir des fantômes plus proches d'un sein réel. Ces recherches sont présentées dans la section suivante.

## 3.2 Mise au point de fantômes réalistes

### 3.2.1 Recherche d'échantillons de tissus mammaires humains

Notre première piste dans la recherche de fantômes réalistes était d'obtenir des échantillons de tissus mammaires humains, sains et cancéreux. Dans cette perspective, nous avons contacté le Centre Hospitalier Universitaire Grenoble Alpes (CHUGA). Nous souhaitions avoir des pièces d'au moins quelques centimètres carrés de surface, où les différents types de tissus étaient préalablement identifiés. Le CHUGA nous a proposé des pièces mammaires issues de tumorectomie.

En effet, quand une tumorectomie est réalisée, la partie excisée est envoyée au laboratoire d'analyse anatomo-pathologique du CHUGA, afin de réaliser un contrôle de marges. Pour cela,

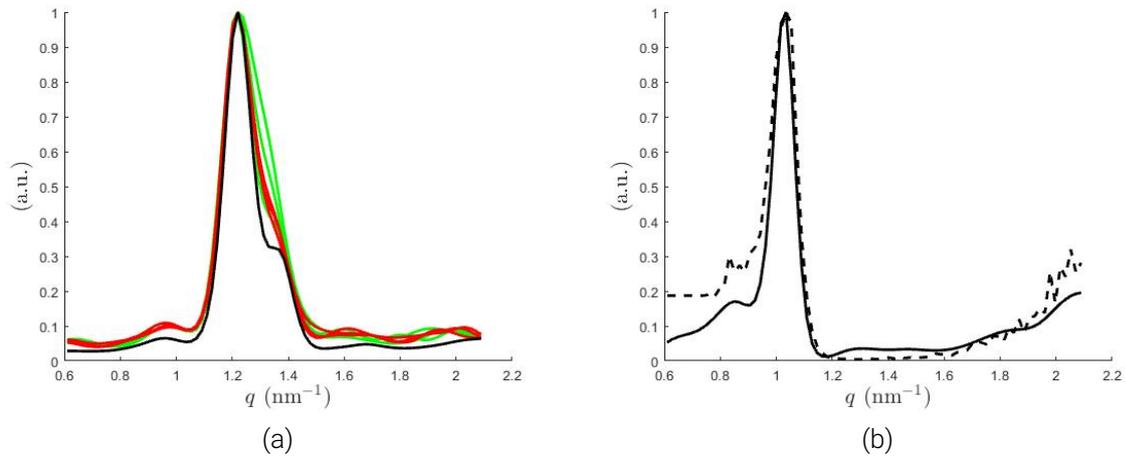


FIGURE 3.8 – Fonctions de diffusion, normalisées par leur maximum : (a) tissu sain, tissu cancéreux et paraffine, respectivement en vert, rouge et noir ; (b) Téflon, reconstruction expérimentale et référence, respectivement en trait plein et en pointillés.

la pièce, prélevée pendant l'opération chirurgicale, subit un traitement :

1. La pièce, dont la plus grande dimension est d'au maximum quelques centimètres, est mise dans plusieurs bains successifs, contenant, notamment, du formaldéhyde et de l'alcool à différents degrés, afin de la préserver de la décomposition.
2. La pièce est ensuite enrobée de paraffine et mise en cassette, pour lui donner de la solidité, et ainsi faciliter la mise sous lame.
3. L'excès de paraffine est enlevé, afin que le tissu soit à la surface.
4. Une fine coupe est réalisée à la surface de l'échantillon. Cette coupe est mise sous lame, colorée, puis analysée au microscope : c'est l'analyse anatomo-pathologique.

À l'issue de cette analyse, la composition exacte de l'échantillon est déterminée. Le CHUGA a alors toutes les informations nécessaires pour le diagnostic, et n'a plus l'utilité des tissus en cassette, et peut nous les envoyer. Le recueil du consentement, la vérification de l'absence d'opposition des patientes concernées, ainsi que l'anonymisation irréversible des échantillons, ont été effectuées par le CHUGA.

La figure 3.7 montre deux échantillons reçus, sain (à gauche) et cancéreux (à droite). Les cassettes font 3 cm × 4 cm, et l'épaisseur des échantillons fixés avec la paraffine mesure 4-5 millimètres.

Comme les échantillons subissent un traitement chimique, nous avons cherché, dans un premier temps, à nous assurer que les fonctions de diffusion n'étaient pas impactées par ce traitement. Nous avons donc réalisé plusieurs acquisitions avec ces échantillons, les rayons X les traversant dans la direction perpendiculaire au plan de la figure 3.7. Dans cette direction, les échantillons ont une très faible épaisseur (4-5 mm). Nous avons donc réalisé les acquisitions avec un montage expérimental similaire à celui présenté sur la figure 2.6. En effet, dans ce

Tension du tube X	Intensité du courant	Temps d'acquisition diffusé	Temps d'acquisition atténué
160 kV	4 mA	10 min	1 min

TABLE 3.3 – Paramètres expérimentaux pour l'acquisition des spectres diffusés et atténués pour les glandes mammaires de souris

cas, nous n'avons pas besoin de résolution spatiale en profondeur, et pouvons donc utiliser une géométrie avec un angle de diffusion unique.

Les spectres diffusés et atténués ont été acquis pour trois points différents pour chaque échantillon, ainsi que pour la paraffine seule et le Téflon. Les fonction de diffusion reconstruites sont présentes sur la figure 3.8. Sur la sous-figure (a), ont été tracées les fonctions de diffusion correspondant aux échantillons sain et cancéreux, respectivement en vert et en rouge, et à la paraffine seule, en noir. Nous constatons que les fonctions de diffusion sont quasi-identiques, et que les maxima attendus à  $q = 1.1 \text{ nm}^{-1}$  et  $q = 1.6 \text{ nm}^{-1}$ , respectivement pour le tissu sain et le tissu cancéreux, ne sont pas observables. Il semblerait que le traitement chimique subi par les échantillons les rend impossible à distinguer de la paraffine. Pour s'assurer que le problème ne vient pas de la reconstruction, nous avons étudié la fonction de diffusion du Téflon. La sous-figure (b) montre la fonction de diffusion obtenue expérimentalement, en trait plein, et la fonction de diffusion théorique, en pointillés. Les deux courbes sont très proches. Nous en déduisons que la reconstruction fonctionne correctement.

Ainsi, d'après les résultats obtenus, nous avons conclu que le traitement chimique, préservant les échantillons, dénature les fonctions de diffusion, et rend inexploitable les pièces mammaires fournies par le CHUGA. Cependant, ce traitement est nécessaire pour réaliser une analyse anatomo-pathologique de l'échantillon et connaître précisément sa composition. Nous avons donc décidé d'abandonner cette piste, et, à défaut de pouvoir travailler avec des tissus humains, nous avons essayé des expériences avec des tissus mammaires de souris.

### 3.2.2 Glandes mammaires de souris

Nous avons cherché à réaliser des acquisitions avec les glandes mammaires saines et tumorales de souris. Dans un premier temps, notre objectif était de confirmer que notre système permet de différencier les glandes saines et tumorales, à partir de leur fonction de diffusion, et également de se comparer aux signatures de la littérature.

Les glandes mammaires saines et cancéreuses nous ont été fournies par l'Institut de Recherche Interdisciplinaire de Grenoble (IRIG). Elles ont été obtenues à partir de femelles de souris "nude", âgées entre 6 et 8 semaines, provenant de Charles River Laboratories. Les souris ont été injectées avec des cellules cancéreuses MDA-MB-231, provenant de *American Type Culture Collection* [55], provoquant le développement de tumeurs mammaires. Les animaux ont été sacrifiés dès que la taille des tumeurs atteignait 10 mm de diamètre. Ensuite, les

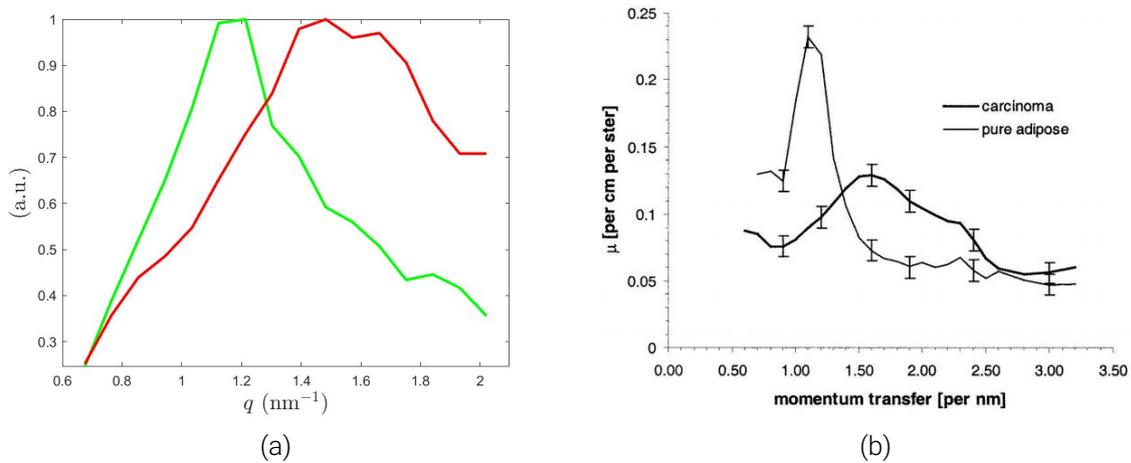


FIGURE 3.9 – Fonctions de diffusion des tissus mammaires. (a) fonctions de diffusion obtenues expérimentalement, normalisées par leur maximum, des glandes mammaires de souris saines et tumorales, respectivement en vert et en rouge. (b) fonctions de diffusion obtenues par Kidane et al. pour le carcinome et le tissu adipeux, respectivement en trait fin et en trait gras [35].

glandes mammaires saines et cancéreuses ont été isolées. Aucun conservateur chimique n'a été utilisé, et nos expériences ont été réalisées dans les 24 heures qui suivaient l'obtention des glandes. Les glandes saines et tumorales ont été conditionnées séparément dans deux tubes de 8 mm de diamètre.

Ces expériences ont été menées dans le respect des lois européennes et françaises sur l'expérimentation animale.

Les paramètres expérimentaux pour les acquisitions des spectres diffusés et atténués sont présentés dans la table 3.3.

Les spectres diffusés et atténués ont été mesurés et reconstruits, afin d'obtenir les fonctions de diffusion présentées sur la figure 3.9. Les fonctions de diffusion correspondent aux glandes saines (en vert) et cancéreuses (en rouge). Nous observons bien des maxima pour  $q = 1.1$  nm<sup>-1</sup> et  $q = 1.6$  nm<sup>-1</sup>, pour les glandes saines et tumorales, respectivement. Les deux fonctions de diffusion sont clairement distinctes et permettent de caractériser les tissus.

Ainsi, nous avons obtenu les résultats attendus. Cependant, les glandes mammaires de souris ont un volume très faible. Il serait très difficile de fabriquer un fantôme d'épaisseur équivalente à un sein, ce qui est nécessaire pour évaluer la résolution spatiale et l'intérêt du système en conditions réelles. C'est pourquoi nous avons choisi d'utiliser comme références les fonctions de diffusion obtenues à partir des glandes mammaires de souris. Ensuite, nous avons effectué des essais avec la viande bovine, plus facilement accessible, pour déterminer si nous pouvons l'utiliser pour imiter les fonctions de diffusion des glandes mammaires de souris, saines et tumorales.

Quelques résultats obtenus sur les glandes mammaires de souris ont été présentés à la confé-

Tension du tube X	Intensité du courant	Temps d'acquisition diffusé	Temps d'acquisition atténué
160 kV	4 mA	10 min	1 min

TABLE 3.4 – Paramètres expérimentaux pour l'acquisition des spectres diffusés et atténués pour les tissus bovins

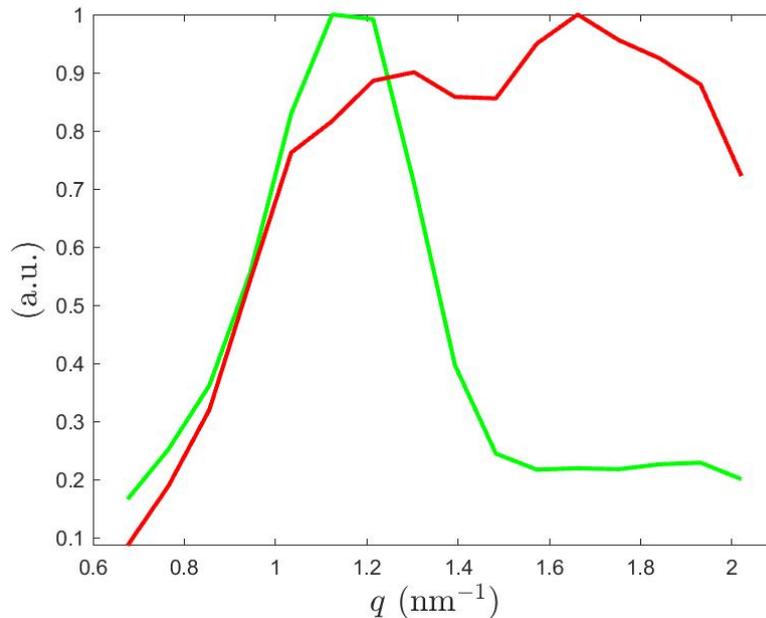


FIGURE 3.10 – Fonctions de diffusion, normalisées par leur maximum, correspondant au tissu bovin adipeux et musculaire, respectivement en vert et en rouge.

rence SPIE Medical Imaging à Houston, Texas en 2018 (voir Annexe B).

### 3.2.3 Fantômes de viande

Nous avons obtenu du gras et de la viande maigre de boeuf, qui correspondent, respectivement, aux tissus adipeux et musculaire. Nous avons réalisé des acquisitions sur ces tissus en utilisant les paramètres expérimentaux décrits dans la table 3.4.

Puis, après reconstruction, nous avons obtenu les fonctions de diffusion pour les tissus adipeux et musculaire. Elles sont représentées sur la figure 3.10. Nous observons des similarités entre les fonctions de diffusion des glandes mammaires de souris et les fonctions de diffusion des tissus bovins. La figure 3.11 permet de comparer les glandes mammaires saines avec le tissu adipeux, et les glandes mammaires tumorales avec le tissu musculaire :

- les fonctions de diffusion du tissu adipeux et des glandes mammaires saines sont maximales en  $q = 1.1-1.2 \text{ nm}^{-1}$ . Par contre, la fonction de diffusion des glandes mammaires saines est plus évasée. Ceci s'explique par le fait que, théoriquement, les glandes saines

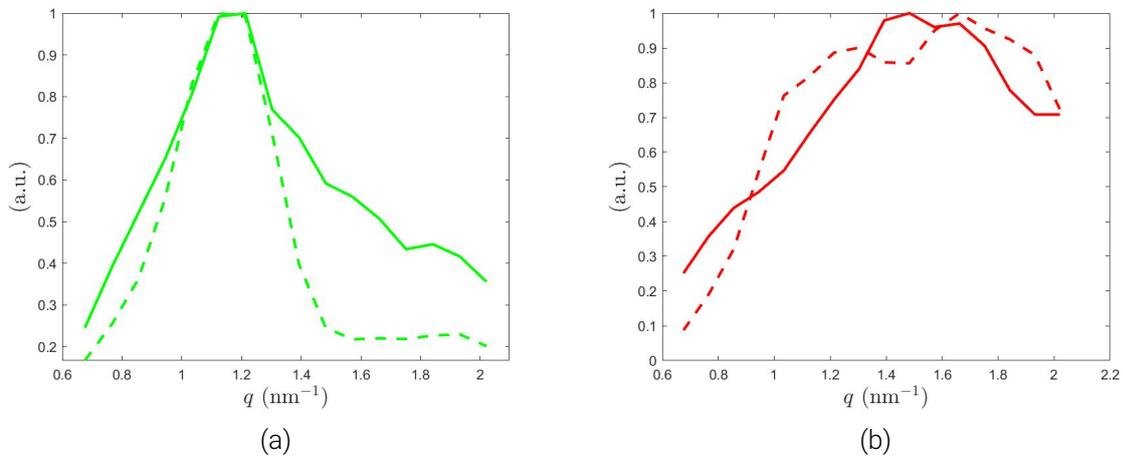


FIGURE 3.11 – Comparaison des fonctions de diffusion, normalisées par leur maximum : (a) glandes mammaires de souris saines et tissu adipeux bovin, respectivement en trait plein et pointillé. (b) glandes mammaires de souris tumorales et tissu musculaire bovin, respectivement en trait plein et pointillé.

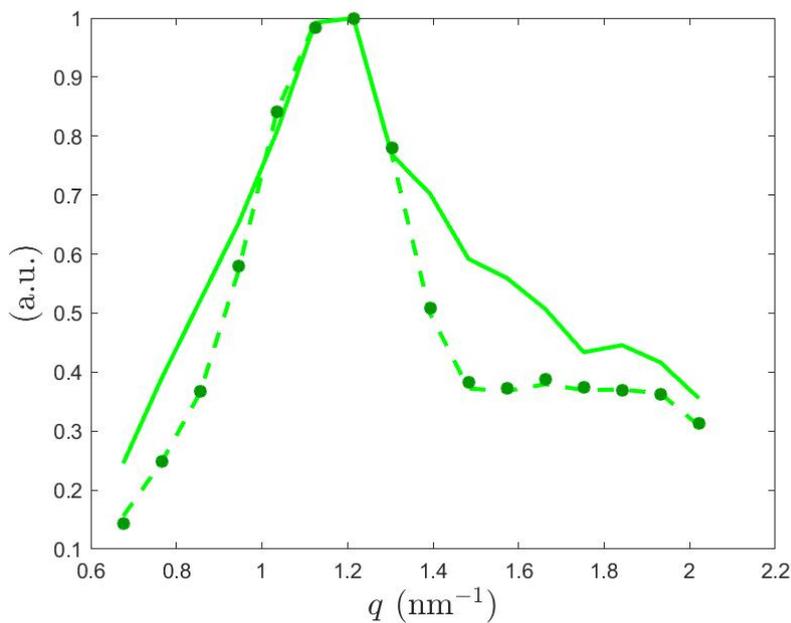


FIGURE 3.12 – Comparaison des fonctions de diffusion, normalisées par leur maximum : en trait plein, les glandes mammaires de souris saines; en pointillés, mélange théorique, composé de 65% de tissu adipeux et 35% de tissu musculaire; les points en vert foncé, la fonction de diffusion de ce mélange, obtenue expérimentalement.

sont composées de tissu adipeux, mais également de tissu fibro-glandulaire.

- les maxima des fonctions de diffusion des glandes mammaires tumorales et du tissu musculaire sont proches de  $q = 1.5-1.6 \text{ nm}^{-1}$ . Les deux fonctions de diffusion n'ont pas exactement la même distribution; cependant, elles sont toutes les deux relativement plates par comparaison avec les glandes mammaires saines ou le tissu adipeux.

Dans le chapitre 1 (§1.1), nous avons vu que le sein sain était constitué de tissus adipeux et fibro-glandulaire. Dans l'hypothèse de présence d'un cancer, le sein comporte également du tissu tumoral.

Pour imiter les glandes mammaires tumorales, nous avons choisi le tissu musculaire bovin. Par contre, le tissu adipeux seul n'est pas assez représentatif des glandes saines. Donc, pour imiter les glandes saines, nous avons choisi de mélanger le tissu adipeux et le tissu musculaire, pour obtenir une fonction de diffusion plus évasée. Pour cela, nous avons utilisé la méthode des moindres carrés, pour trouver les proportions de tissu adipeux et cancéreux :

$$F_{MIX} = \alpha F_{ADP} + \beta F_{MSC} \quad (3.1)$$

où  $F_{MIX}$ ,  $F_{ADP}$  et  $F_{MSC}$  sont les fonctions de diffusion, respectivement, du mélange, du tissu adipeux et du tissu musculaire. Après calcul, nous obtenons  $\alpha = 0.65$  et  $\beta = 0.35$ . Sur la figure 3.12, les fonctions de diffusion des glandes mammaires saines de souris et ce mélange théorique sont représentées, respectivement en trait plein et en pointillés. Ensuite, nous avons réalisé ce mélange,  $\alpha$  et  $\beta$  étant les pourcentages massiques de tissus adipeux et musculaire, respectivement. La fonction de diffusion de ce mélange, obtenue expérimentalement, est représentée avec des points en vert foncé sur la figure 3.12. Nous remarquons que les fonctions de diffusion du mélange, théorique et expérimentale, sont parfaitement confondues. De plus, nous avons bien obtenu une fonction, dont la distribution approche mieux celle des glandes mammaires saines, que la fonction de diffusion du tissu adipeux seul.

Ainsi, nous avons choisi d'utiliser le tissu musculaire pour imiter les glandes mammaires tumorales. Pour imiter le tissu mammaire sain, nous avons utilisé le mélange de tissus adipeux et musculaire décrit précédemment, et le tissu adipeux seul. Dans la section suivante, nous présentons les résultats obtenus à partir de fantômes fabriqués avec ces tissus.

### 3.3 Fantômes de viande

En utilisant les tissus adipeux et musculaire bovins, nous avons fabriqué plusieurs fantômes dans l'objectif de tester le pouvoir de caractérisation de notre système en termes de résolution spatiale et d'identification des tissus. Les fantômes, dont les photographies sont présentées sur la figure 3.13, sont composés de tissus adipeux, musculaire et de leur mélange. Ils sont notés respectivement **ADP**, **MSC** et **MIX** sur les figures pour alléger les notations. Ils sont conditionnés dans des contenants plastiques de taille 100 mm × 50 mm.

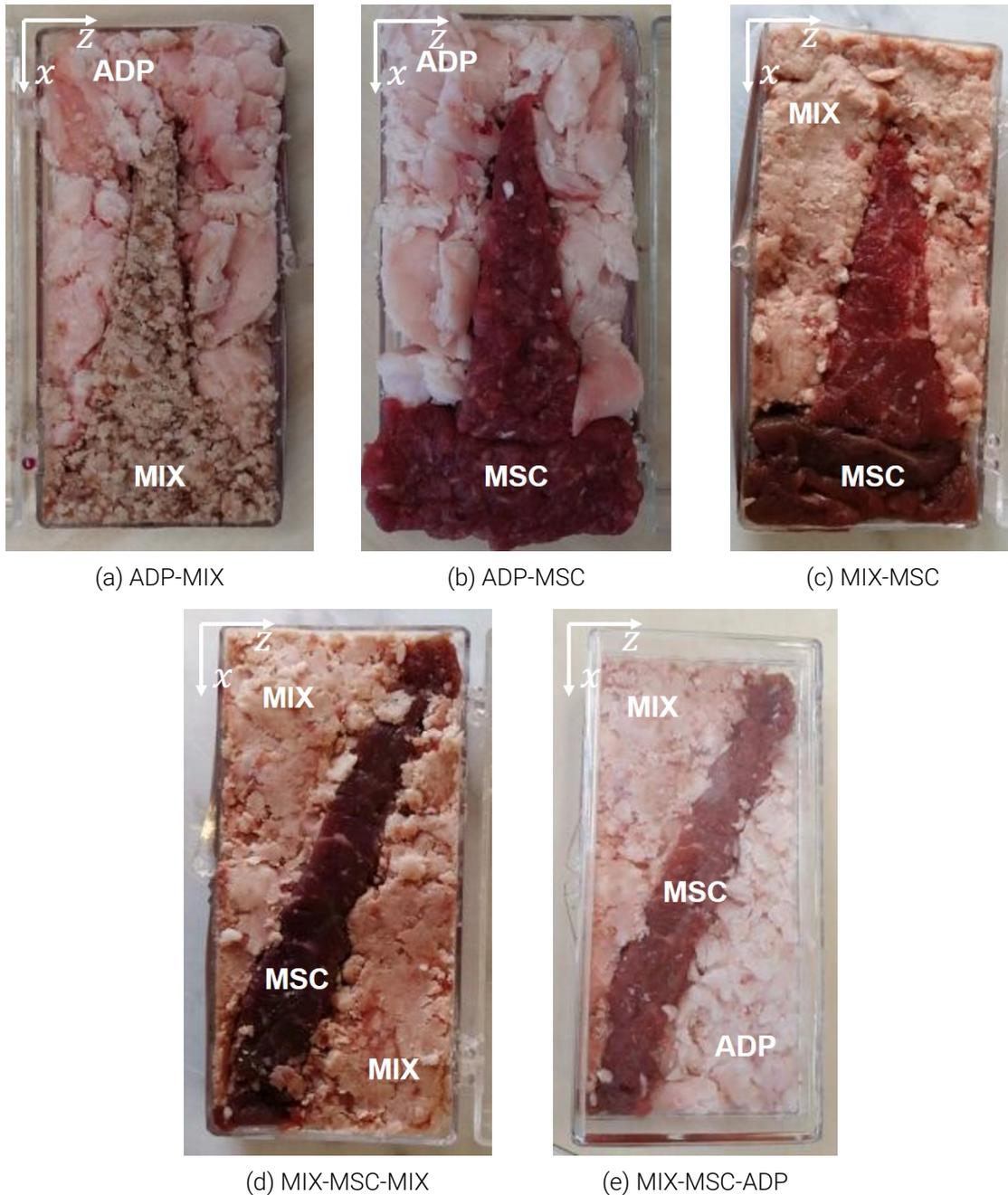


FIGURE 3.13 – Photographies des fantômes, composés de tissu adipeux (ADP), tissu musculaire (MSC) et mélange (MIX) : (a) mélange + tissu adipeux, (b) tissu musculaire + tissu adipeux, (c) tissu musculaire + mélange, (d) mélange + tissu musculaire, (e) tissu adipeux + tissu musculaire + mélange. Les contenants mesurent 100 mm × 50 mm.

Fantôme	Tension du tube X (kV)	Intensité du courant (mA)	Temps d'acquisition par position $x$ (min)	Positions $x$ min : pas : max (mm)
ADP-MIX ADP-MS MIX-MS	160	4	30	-120 :15 :-30
MIX-MS-MIX	160	4	10	-120 :10 :-40
MIX-MS-ADP	160	4	10	-120 :10 :-30

TABLE 3.5 – Paramètres expérimentaux pour l'acquisition des spectres diffusés pour les différents fantômes

Les fantômes ADP-MIX, ADP-MS et MIX-MS, représentés sur les figures (a) à (c), ont été fabriqués de manière à estimer la résolution spatiale du système : chacun de ces fantômes comporte une partie triangulaire, constituée de mélange ou de tissu musculaire. Ainsi, en observant à partir de quelle épaisseur on arrive à détecter cette partie triangulaire, nous avons une estimation de la résolution spatiale. Les fantômes MIX-MS-MIX et MIX-MS-ADP, représentés sur les figures (d) et (e), comportent une bande diagonale de tissu musculaire. La largeur de la bande est relativement constante. Par la géométrie de notre collimateur, la résolution spatiale n'est pas uniforme selon  $z$ . C'est pourquoi la géométrie de ces deux fantômes a été choisie afin d'étudier la qualité de reconstruction et de classification en fonction de la profondeur.

Nous allons présenter les conditions d'acquisition et les mesures obtenues, puis les résultats de reconstruction. Finalement, nous allons présenter les résultats de classification en fonction de la dose déposée.

### 3.3.1 Mesures expérimentales

Les mesures ont été effectuées avec la géométrie présentée dans le chapitre 2 sur la figure 2.3. Pour chaque fantôme, les acquisitions sont réalisées selon la direction  $z$ , en tradant les fantômes selon  $x$ , afin d'obtenir des spectres transmis et diffusés selon plusieurs lignes horizontales. Les paramètres d'acquisitions sont résumés dans la table 3.5. Les positions  $x$  vont de -120 mm à -30 mm, ce qui correspond au déplacement de la table, sur laquelle est positionné le fantôme. Pour chaque position  $x$ , les dimensions de chaque tissu, composant le fantôme, ont été relevées afin de créer les cartes de vérité terrain utilisées pour la classification. Elles sont représentées sur la figure 3.14.

Les acquisitions de spectres diffusés ont été réalisées sur des temps très longs, allant jusqu'à 30 minutes. Nous avons utilisé un mode de lecture spécial, qui enregistre les événements (détectons de photons) en fonction du temps. Ainsi, après la fin de l'acquisition, il est possible

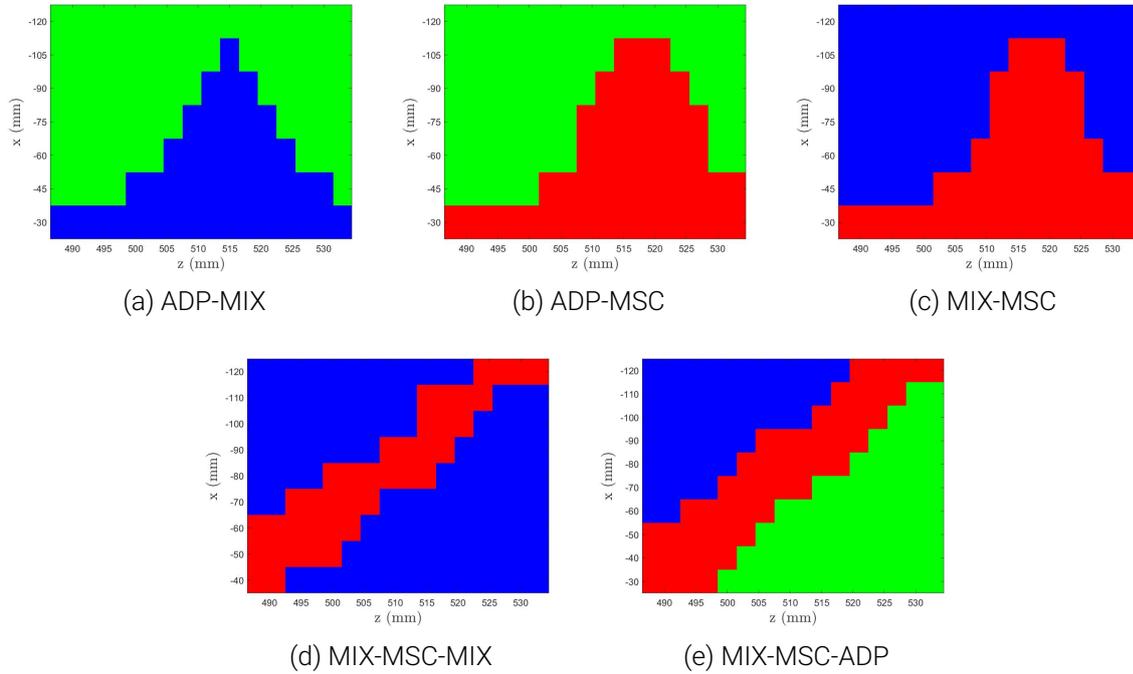


FIGURE 3.14 – Vérité terrain pour chaque fantôme, en fonction des positions ( $x$ ,  $z$ ). Les couleurs représentent le tissu adipeux (en vert), le mélange (en bleu) et le tissu musculaire (en rouge).

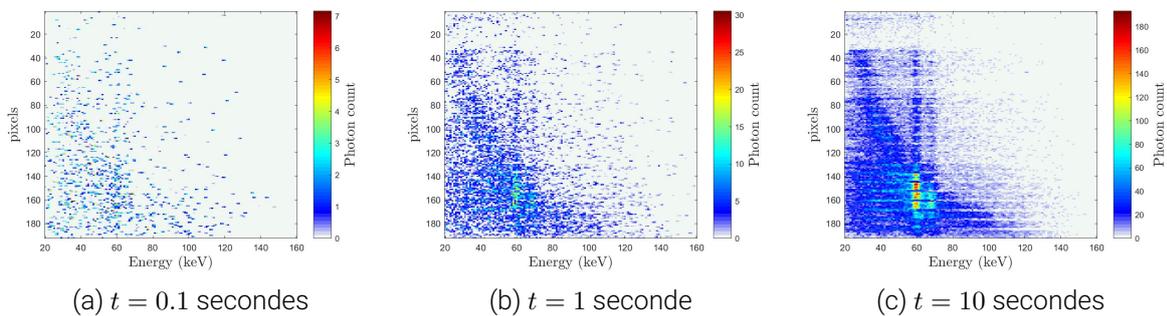


FIGURE 3.15 – Spectres diffusés par le fantôme ADP-MIX, à la position  $x = -90$  mm, pour différents temps d'acquisition  $t$ . Le nombre total de photons mesurés pour chaque figure vaut, respectivement,  $2.0 \cdot 10^3$ ,  $2.1 \cdot 10^4$  et  $2.1 \cdot 10^5$ .

de découper l'enregistrement en temps plus courts. À partir de l'acquisition totale, nous avons sélectionné des temps d'acquisition, notés  $t$ , allant de 0.1 à 50 secondes. Afin d'effectuer par la suite une étude statistique en fonction de la dose déposée, nous avons suivi l'algorithme suivant :

---

Pour chaque fantôme  $\phi$  :

Pour chaque position  $x$  de  $\phi$  :

Pour chaque temps  $t$  :

Pour chaque réalisation  $r \in [1 - 25]$

Soit  $t_0$  un temps aléatoire entre 0 et  $T_{max} - t$ , où  $T_{max}$  est le temps total de l'acquisition;

Extraction des évènements qui ont eu lieu entre  $t_0$  et  $t_0 + t$ .

---

Ainsi, pour chaque fantôme, pour chaque position  $x$  et pour chaque temps  $t$ , nous obtenons 25 mesures de spectres diffusés. La figure 3.15 représente un exemple de spectres diffusés obtenus pour le fantôme ADP-MIX (voire figure 3.14a), à la position  $x = -90$  mm pour trois temps différents.

### 3.3.2 Reconstruction

Après la reconstruction avec l'algorithme MLEM, nous obtenons pour chaque fantôme, pour chaque temps d'acquisition et chaque répétition, une fonction de diffusion en fonction de  $x$  et  $z$ ,  $F(x, z, q)$ . Pour faciliter la visualisation de ces fonctions de diffusion tri-dimensionnelles, nous avons choisi de représenter une grandeur dérivée,  $R_F(x, z)$  :

Soit, pour tous  $x$  et  $z$ ,  $F_{ADP}$  la moyenne de  $F$  par rapport à  $q$  sur l'intervalle  $Q_1 = [1.0 - 1.2]$   $\text{nm}^{-1}$  et  $F_{MSC}$  la moyenne de  $F$  par rapport à  $q$  sur l'intervalle  $Q_2 = [1.5 - 1.7]$   $\text{nm}^{-1}$ . Alors

$$R_F(x, z) = \frac{F_{MSC}(x, z)}{F_{ADP}(x, z)}$$

Comme on peut l'observer sur les figures 3.10 et 3.12, les fonctions de diffusions du tissu adipeux et du mélange sont maximales pour  $q \in Q_1$ , et plus faibles pour  $q \in Q_2$ . À l'inverse, la fonction de diffusion du tissu musculaire est légèrement plus importante sur l'intervalle  $Q_2$ , que sur l'intervalle  $Q_1$ . Ainsi,  $R_F(x, z)$  permet de mettre en exergue les différences propres à la nature des tissus. La figure 3.16 présente des exemples de  $R_F$ , calculés pour les cinq fantômes, avec les fonctions de diffusion, reconstruites à partir des spectres diffusés correspondant à 0.1, 1 et 10 secondes d'acquisition.

Nous observons qu'avec l'augmentation du temps d'acquisition, des structures apparaissent plus clairement sur les figures. Ces structures sont cohérentes avec la vérité terrain de chaque fantôme, présentée sur la figure 3.14. En étudiant les figures de la troisième colonne, nous

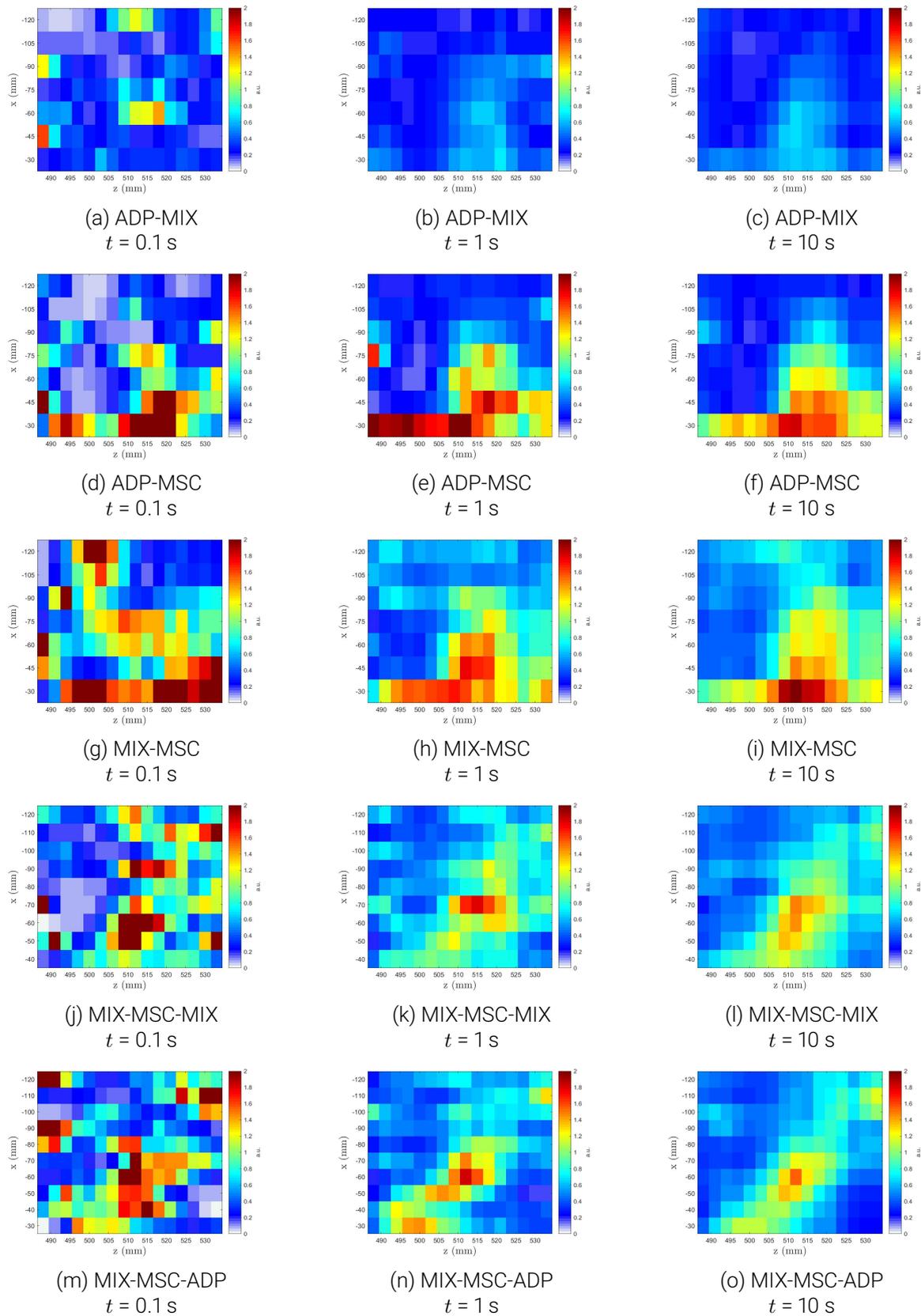


FIGURE 3.16 – Cartes  $R_F$  obtenues pour différents fantômes (selon les lignes) et différents temps d'acquisition (selon les colonnes). L'échelle de couleurs est identique pour toutes les figures.

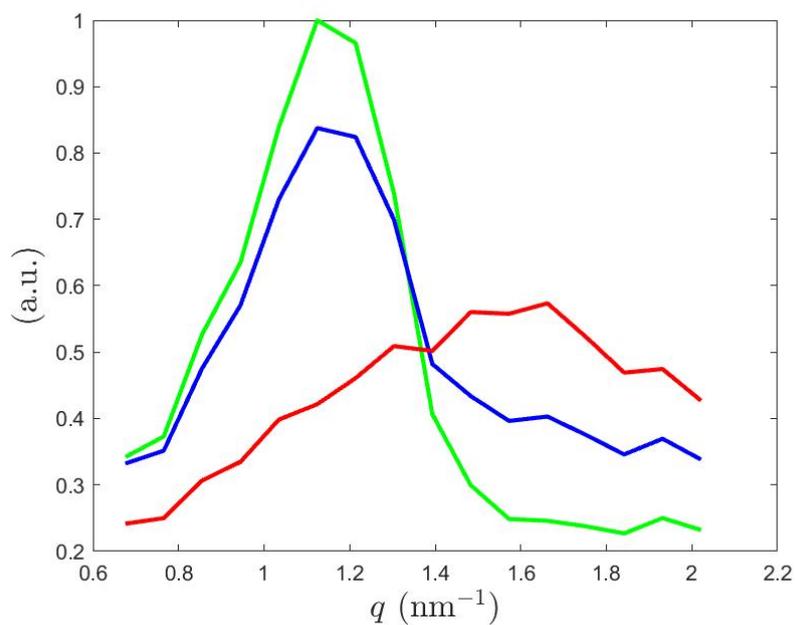


FIGURE 3.17 – Fonctions de diffusion du tissu adipeux, du mélange et du tissu musculaire, respectivement en vert, bleu et rouge. Ces fonctions de diffusion ont été obtenues à partir des acquisitions à  $t = 50$  secondes. Les courbes correspondent aux moyennes sur toutes les réalisations et sur toutes les positions  $z$ , aux positions  $x$  correspondantes aux tissus : le tissu adipeux a été calculé sur les fantômes ADP-MIX et ADP-MS, à la position  $x = -120$  mm, le mélange a été calculé sur le fantôme ADP-MIX à la position  $x = -30$  mm, et le tissu musculaire sur les fantômes ADP-MS et MIX-MS à la position  $x = -30$  mm.

distinguons clairement les positions correspondant au tissu musculaire, par rapport au tissu adipeux et au mélange. On peut remarquer que le contraste est plus élevé entre le tissu musculaire et le tissu adipeux, qu'entre le tissu musculaire et le mélange. Ceci est cohérent avec les fonctions de diffusion présentées sur la figure 3.17 : le mélange est un intermédiaire entre le tissu adipeux et le tissu musculaire. Il est donc logique que le contraste entre le tissu adipeux et le tissu musculaire soit plus élevé qu'entre le mélange et le tissu musculaire.

### 3.3.3 Classification

Les fonctions de diffusion, obtenues précédemment, ont été classifiées en utilisant l'algorithme SVM. Pour chaque temps d'acquisition et chaque répétition, les fonctions de diffusion d'un fantôme, pour tous les  $x$ , ont été classifiées en utilisant en apprentissage les fonctions de diffusion des quatre autres fantômes. La classification se fait sur deux classes, "sain" et "pathologique". Ici, le tissu adipeux et le mélange sont considérés comme "sains", et le tissu musculaire comme "pathologique". La classification est notée  $C(x, z)$ .

Pour mieux comparer la vérité terrain et les résultats de classification, nous avons calculé des cartes  $M(x, z)$ , qui représentent, pour chaque classification les TP, les TN, les FP et les FN correspondants.

Des exemples de classification et de cartes  $M$  correspondantes, pour les différents fantômes et pour des données correspondant à des temps d'acquisition  $t = 0.1, 1$  et  $10$  secondes, sont présentés sur les figures 3.18, 3.19, 3.20, 3.21 et 3.22.

Globalement, sur toutes les reconstructions, nous observons une amélioration de classification avec l'augmentation des temps d'acquisition.

La figure 3.18 représente les résultats de classification du fantôme ADP-MIX. Ce fantôme devrait être entièrement classifié comme "sain", et c'est effectivement ce qu'on observe pour la majorité des positions. À très faible dose, quelques faux positifs sont présents.

La classification obtenue pour le fantôme ADP-MS, représenté sur la figure 3.19, ne comporte pas de faux positifs. Cependant, pour beaucoup de positions, le tissu musculaire n'est pas identifié. Ces résultats sont un peu décevants, d'autant plus que le contraste est assez marqué sur les cartes  $R_F$  correspondantes, comme nous pouvons l'observer sur la figure 3.16. De plus, le tissu musculaire paraît bien plus petit en taille.

Le résultat obtenu pour le fantôme MIX-MS, représenté sur la figure 3.20, est assez similaire au précédent. Quelques faux positifs apparaissent à  $t = 10$  secondes. Cependant, l'estimation de la taille du tissu musculaire est légèrement meilleure. Ce cas est le plus représentatif de notre problématique de détection de tumeur dans un milieu de tissu sain, et on observe un résultat globalement encourageant.

Le fantôme MIX-MS-MIX, dont les résultats de classification sont donnés sur la figure 3.21, est également représentatif de cette problématique. Nous observons que le tissu "pathologique" est grossièrement détecté à partir de  $t = 1$  seconde. Cependant, il semblerait qu'aux bords du fantôme, la détection soit plus difficile. De plus, nous observons un décalage selon

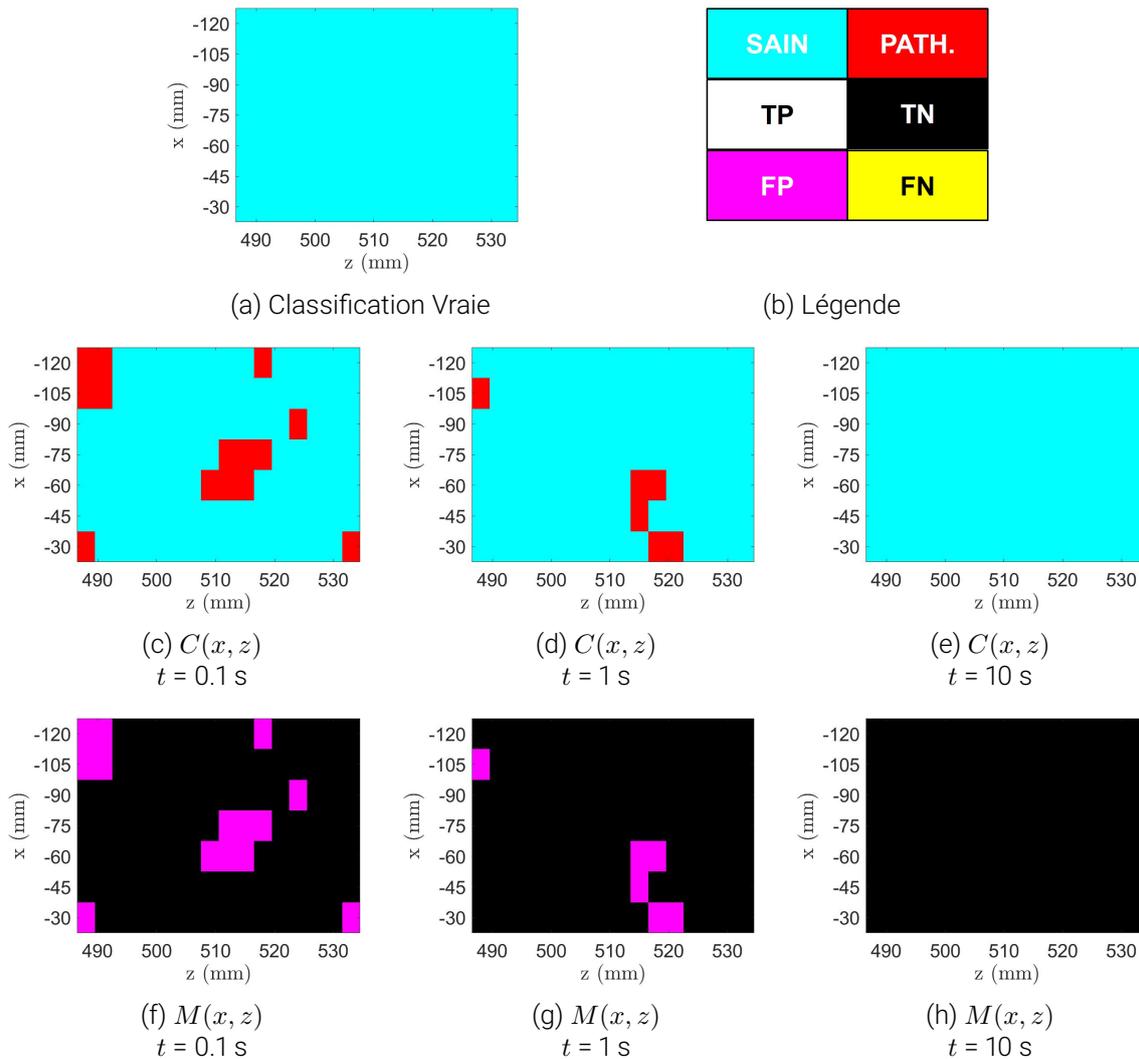


FIGURE 3.18 – Classification du fantôme ADP-MIX. (a) classification, (b) rappel de la légende. (c), (d) et (e) représentent les classifications  $C(x, z)$  obtenues par SVM, à partir de données correspondant aux temps d'acquisition de  $t = 0.1, 1$  et  $10$  secondes. (f), (g) et (h) montrent les cartes  $M(x, z)$  correspondantes.

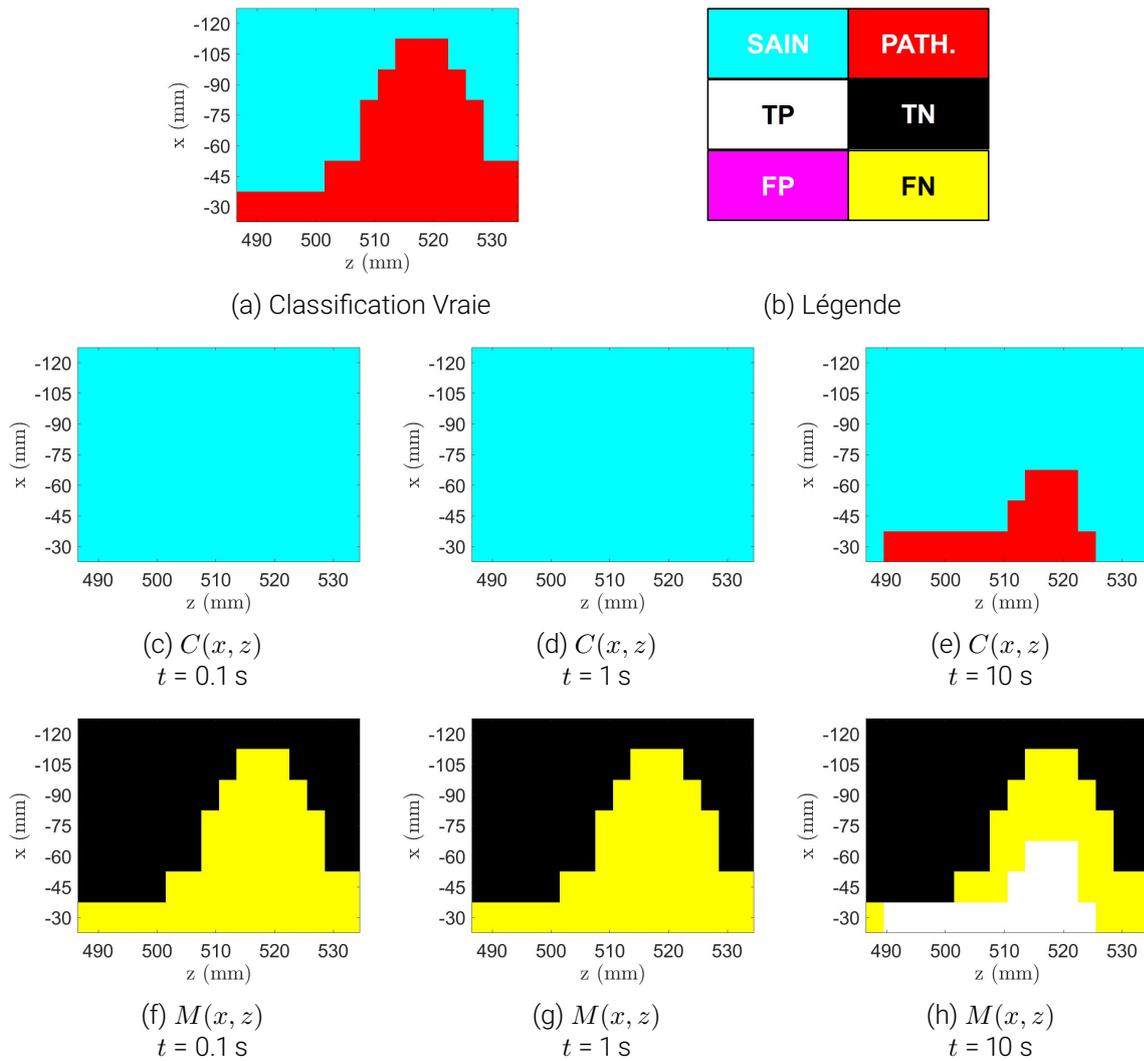


FIGURE 3.19 – Classification du fantôme ADP-MSC. (a) classification, (b) rappel de la légende. (c), (d) et (e) représentent les classifications  $C(x, z)$  obtenues par SVM, à partir de données correspondant aux temps d'acquisition de  $t = 0.1, 1$  et  $10$  secondes. (f), (g) et (h) montrent les cartes  $M(x, z)$  correspondantes.

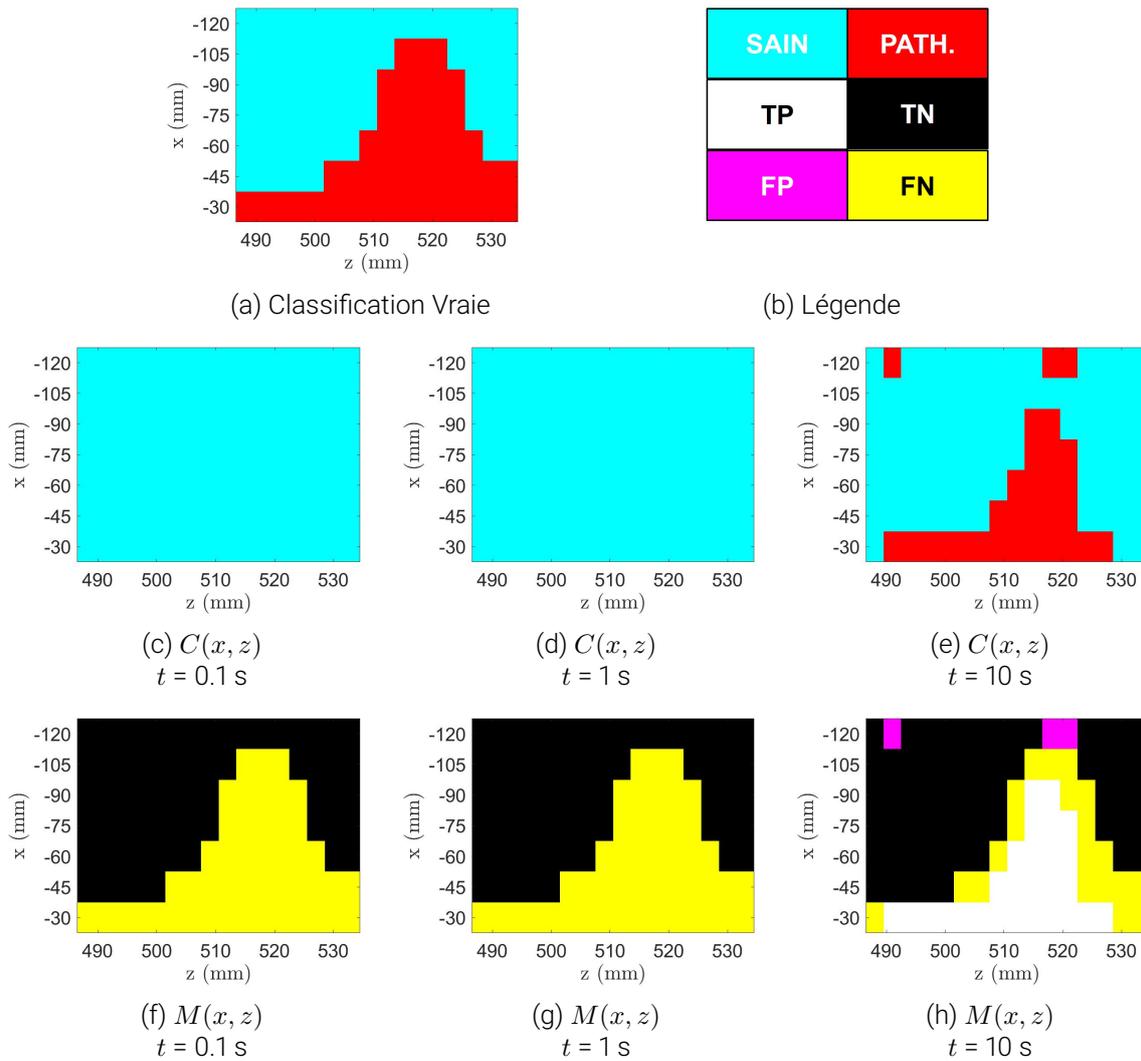


FIGURE 3.20 – Classification du fantôme MIX-MSC. (a) classification, (b) rappel de la légende. (c), (d) et (e) représentent les classifications  $C(x, z)$  obtenues par SVM, à partir de données correspondant aux temps d'acquisition de  $t = 0.1, 1$  et  $10$  secondes. (f), (g) et (h) montrent les cartes  $M(x, z)$  correspondantes.

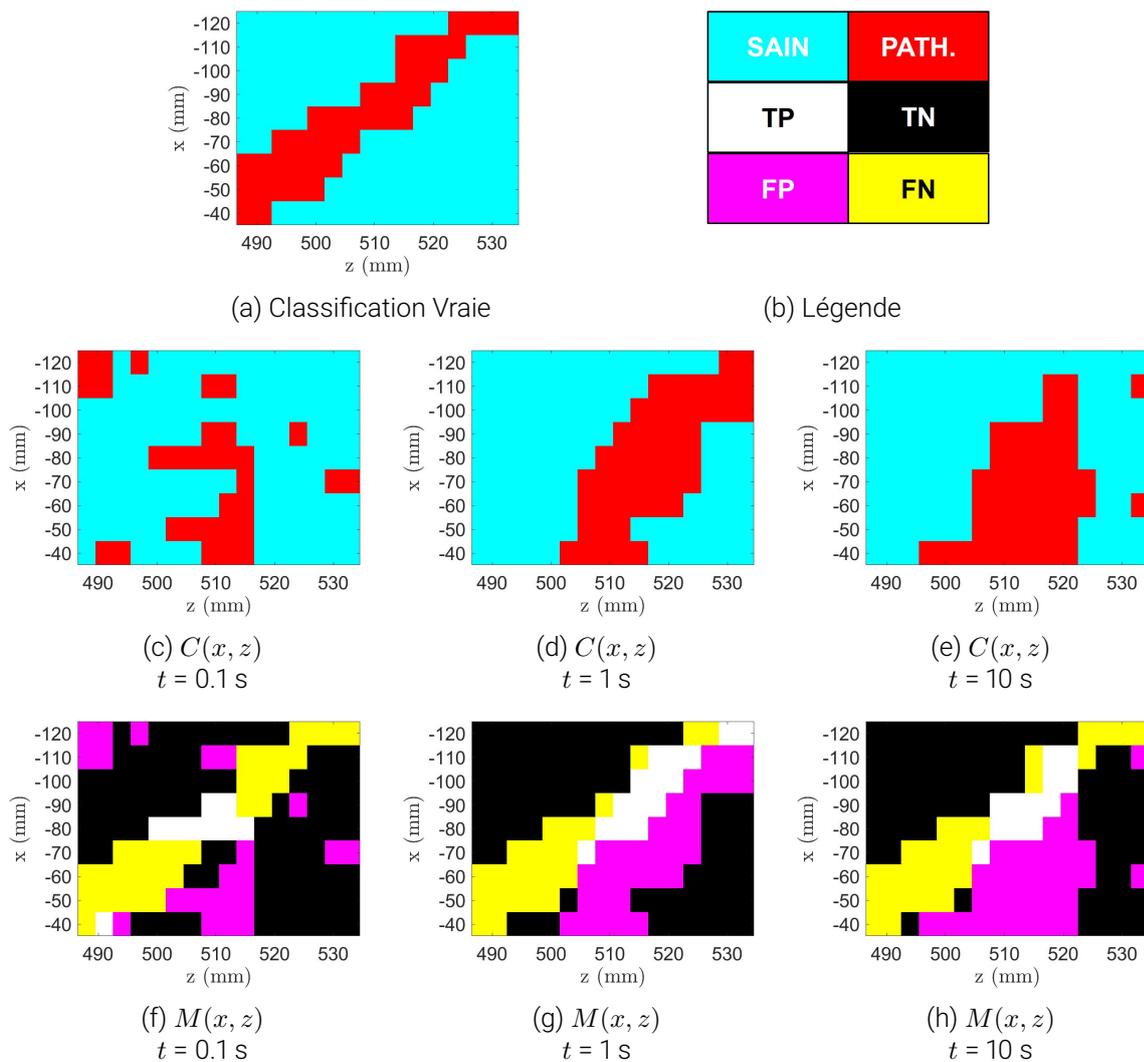


FIGURE 3.21 – Classification du fantôme MIX-MSC-MIX. (a) classification, (b) rappel de la légende. (c), (d) et (e) représentent les classifications  $C(x, z)$  obtenues par SVM, à partir de données correspondant aux temps d'acquisition de  $t = 0.1, 1$  et  $10$  secondes. (f), (g) et (h) montrent les cartes  $M(x, z)$  correspondantes.

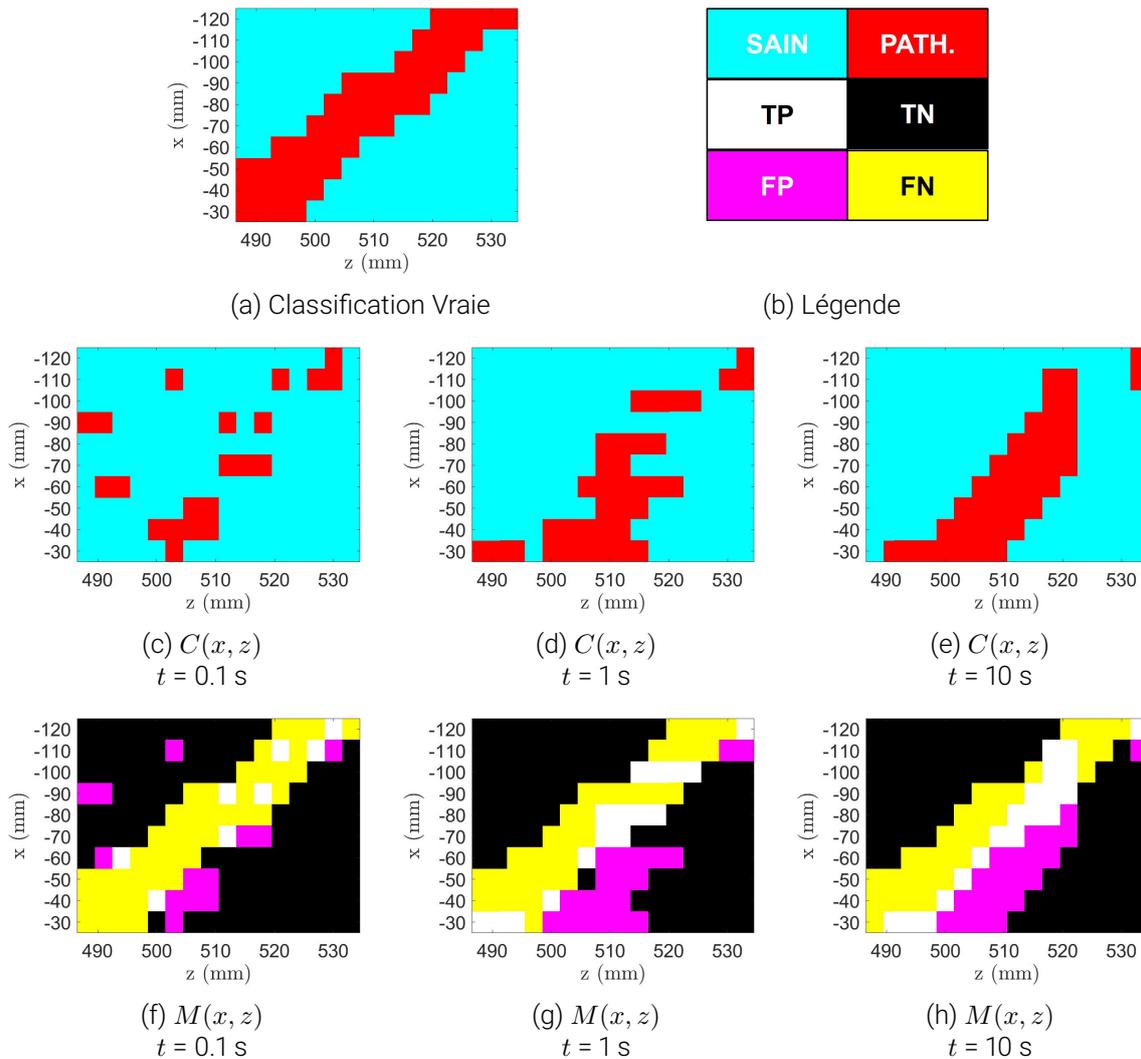


FIGURE 3.22 – Classification du fantôme MIX-MSC-ADP. (a) classification, (b) rappel de la légende. (c), (d) et (e) représentent les classifications  $C(x, z)$  obtenues par SVM, à partir de données correspondant aux temps d'acquisition de  $t = 0.1, 1$  et  $10$  secondes. (f), (g) et (h) montrent les cartes  $M(x, z)$  correspondantes.

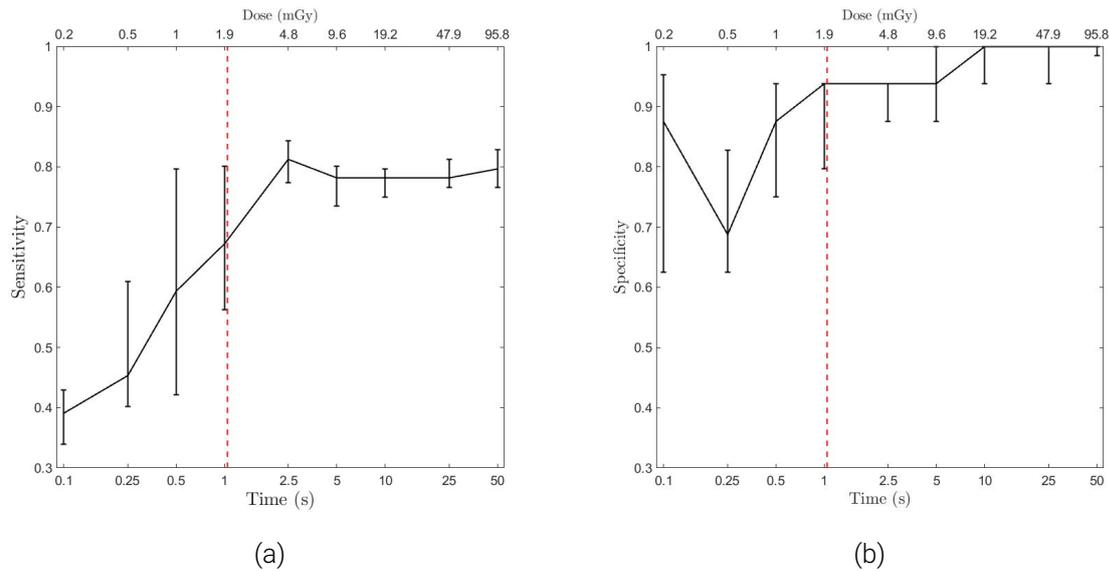


FIGURE 3.23 – Sensibilité (a) et Spécificité (b) en fonction du temps (en bas) et de la dose absorbée correspondante (en haut). Les points affichés correspondent aux médianes sur 25 réalisations, et les barres d'erreur verticales correspondent respectivement au 1<sup>er</sup> et au 3<sup>ème</sup> quartiles. Pour d'avantage de lisibilité, l'échelle logarithmique est utilisée pour l'axe des temps, en bas, et des doses, en haut. Le trait pointillé rouge vertical représente 2 mGy, la dose moyenne d'une mammographie.

$x$  : sur les figures, le tissu musculaire est détecté plus bas qu'il ne l'est en réalité.

Ce décalage est toujours présent sur la figure 3.22, qui illustre la classification du fantôme MIX-MS-ADP. De plus, le bord  $z = 530$  mm est toujours problématique. Par contre, l'épaisseur de la diagonale semble à peu près respectée.

En observant notamment les figures 3.19 et 3.20 pour  $t = 10$  secondes, nous avons constaté que souvent le tissu "pathologique" était détecté, mais ces dimensions paraissent plus faibles qu'elles ne le sont en réalité. Dans la perspective d'utilisation de notre système en tant qu'outil de dépistage, ceci n'est pas problématique : tant qu'on est capable de détecter une tumeur de manière fiable, une erreur sur ces dimensions est acceptable. En effet, si une tumeur est détectée, un autre examen pourra être réalisé pour déterminer précisément sa taille. Par contre, nous avons une exigence sur la fiabilité de diagnostic : nous ne voulons ni manquer des tumeurs, ni en détecter alors que le tissu est sain. Pour refléter cet objectif, nous avons choisi de modifier notre manière de calculer la sensibilité et la spécificité.

En effet, dans le cas du fantôme plastique, nous avons attribué les labels TP, TN, FP ou FN à chaque position  $(x, z)$ . Dans le cas présent, nous constatons que la résolution selon  $z$  est trop grossière. En effet, le temps d'acquisition ayant été fortement réduit, de cinq minutes par position, pour le fantôme plastique, à maximum 10 secondes pour les figures présentées, cela a impacté la résolution spatiale. Pour évaluer la sensibilité et la spécificité, nous avons donc choisi de ne pas tenir compte de la résolution selon l'axe  $z$ .

Ainsi, les TP, les TN, les FP ou les FN ont été calculés pour tous  $x$ , tous les fantômes, toutes

les réalisations, et pour des temps d'acquisition variant de 0.1 à 50 secondes. Nous avons sommé les nombres obtenus sur les  $x$  et les fantômes, afin d'obtenir les TP, les TN, les FP ou les FN pour 25 réalisations, en fonction du temps. Nous en avons déduit 25 réalisations de la sensibilité et de la spécificité, en fonction du temps d'acquisition. En utilisant notre formule de calcul de la dose absorbée en fonction du temps (voir la section §2.5 p.56), nous avons obtenu les doses absorbées correspondantes aux différents temps d'acquisition. Dans cette expérience, les bords de l'objet correspondent à  $Z_{min} = 485$  mm, et  $Z_{max} = 535$  mm. Ainsi, d'après la formule 2.18, nous obtenons une dose absorbée de 1.9 mGy par seconde. La figure 3.23 présente l'évolution de la sensibilité et de la spécificité, en fonction du temps et de la dose absorbée. Les points correspondent aux valeurs médianes sur les 25 réalisations, et les barres d'erreur correspondent au 1<sup>er</sup> et au 3<sup>ème</sup> quartiles.

On observe, comme attendu, une augmentation de la sensibilité et de la spécificité, et une diminution de leurs dispersions respectives, en fonction du temps et de la dose. La sensibilité semble stagner à  $\approx 0.8$ , et la spécificité dépasse 0.9, à partir de  $t = 2.5$  secondes. La dose correspondante à  $t = 2.5$  secondes est estimée à 4.8 mGy, alors qu'une mammographie dépose une dose entre 2 et 4 mGy.

### 3.3.4 Conclusion

Dans cette section, nous avons présenté les résultats expérimentaux obtenus à partir de fantômes composés de tissus adipeux et musculaire bovins, imitant les tissus mammaires. L'épaisseur des fantômes valait 5 centimètres, ce qui correspond à l'épaisseur moyenne de sein pendant une mammographie. Les acquisitions ont été traitées avec l'algorithme MLEM afin d'obtenir des fonctions de diffusion selon les positions  $(x, z)$  des fantômes. Ensuite, ces fonctions de diffusion ont été classifiées en utilisant le modèle SVM. La sensibilité et la spécificité globales ont été évaluées en fonction du temps et de la dose.

À partir des valeurs obtenues, nous avons observé qu'une sensibilité de 0.8 et une spécificité supérieure à 0.9 peuvent être atteintes, à partir d'une acquisition de 2.5 secondes qui correspond à une dose absorbée de 4.8 mGy. Cette dose est proche de la dose moyenne absorbée pendant un examen mammographique standard.

La résolution spatiale selon l'axe  $z$  est moins précise que ce qu'on a obtenu pour le fantôme plastique. En effet, nous avons réduit le temps des acquisitions, donc les mesures sont plus bruitées que dans l'expérience précédente sur le fantôme plastique. De plus, les tissus utilisés ont des fonctions de diffusion très proches, ce qui rend la discrimination entre eux plus difficile. Par ailleurs, la résolution spatiale dépend de la matrice globale du système. Celle-ci a été calculée avec une résolution selon  $z$  de 3 millimètres. Il serait intéressant de réduire ce pas, par interpolation, ou en effectuant une calibration plus fine.

Globalement, les résultats obtenus sont encourageants. Cependant, une sensibilité de 0.8 signifie que 20% des tumeurs ne sont pas détectées, et une spécificité de 0.9 correspond à 10%

de faux positifs. Des améliorations sur ces deux critères sont nécessaires avant que l'utilisation de ce système en conditions cliniques soit envisageable.

Par ailleurs, les cartes  $R_F$  semblent mieux mettre en valeur le contraste entre les tissus, ce qui remet en cause la classification SVM. En effet, le choix de la classification a été effectué en se basant sur des résultats simulés. Une nouvelle étude sur données expérimentales serait intéressante.

### 3.4 Conclusion

Dans ce chapitre, nous avons présenté nos résultats expérimentaux.

Premièrement, nous avons reproduit aussi fidèlement que possible, la dernière expérience de Damien BARBES avec un fantôme composé de polystyrène et de nylon. Nous avons observé une nette amélioration des résultats, ce qui nous a permis de valider les changements effectués sur le processus de traitement des données.

Ensuite, nous avons effectué plusieurs essais, afin d'obtenir des fantômes plus proches des tissus mammaires. Nous avons réalisé des acquisitions sur les échantillons de tissus mammaires paraffinés, sur les glandes mammaires de souris, et, finalement, sur les tissus adipeux et musculaire bovins. Suite à ces essais, nous avons réalisés des fantômes à partir de tissus bovins, imitant les fonctions de diffusion des glandes mammaires de souris.

Finalement, nous avons réalisé des acquisitions avec différents fantômes composés de tissus bovins adipeux et musculaire, et de leur mélange. Nous avons étudié les performances, en termes de sensibilité et de spécificité, en fonction du temps et de la dose. Les résultats obtenus sont globalement positifs : en quelques secondes et pour une dose absorbée similaire à ce qui est déposé en mammographie, nous avons obtenu une sensibilité de 0.8, et une spécificité supérieure à 0.9. Cette performance est très encourageante, même si elle doit encore être améliorée. Par ailleurs, l'objectif initial d'une résolution spatiale selon  $z$  de l'ordre du millimètre n'est pas encore atteint.

Ainsi, ces résultats valident les améliorations apportées pendant cette thèse, mais également mettent en évidence les faiblesses restantes du système. En conclusion, nous proposerons quelques pistes de résolution de ces points problématiques pour la suite de ce projet.



## CONCLUSIONS ET PERSPECTIVES

---

Dans cette thèse, nous nous sommes intéressés à la problématique d'utilisation de la diffusion cohérente des rayons X pour le diagnostic du cancer du sein. Plusieurs équipes de recherche ont démontré que les différents types de tissus mammaires ont des signatures de diffusion différentes, et qu'il est donc possible de distinguer les tissus mammaires sains des tissus cancéreux, en utilisant une mesure de la diffusion cohérente. Un système d'imagerie exploitant cette méthode pourrait donc être mis au point, s'il répond aux exigences de l'imagerie médicale. Nous avons envisagé deux applications : *in vivo*, il servirait à vérifier les résultats de la mammographie en cas de doute, et *ex vivo* il serait utilisé pour le contrôle de marges per-opératoire. Dans les deux cas, le système devra donner un diagnostic fiable rapidement, et avoir une bonne résolution spatiale. De plus, pour une utilisation *in vivo*, la dose absorbée devra être limitée. Dans nos travaux, nous nous sommes intéressés surtout à l'application *in vivo*, qui est la plus contraignante.

Plusieurs équipes travaillent sur la mise au point d'un tel système. En particulier, au CEA Leti, Damien BARBES et Fanny MARTICKE ont réalisé une étude de faisabilité pendant leurs thèses respectives, soutenues en 2016. En utilisant des détecteurs spectrométriques, développés au sein du laboratoire d'accueil, ils ont proposé un premier système, composé d'un montage expérimental et d'un algorithme de traitement des données. Les derniers résultats obtenus ont laissé penser que le traitement des données devait être étudié davantage, afin d'obtenir des résultats plus intéressants. L'amélioration du procédé de traitement des données était donc l'objectif de cette thèse, ainsi que son utilisation sur un système expérimental avec des fantômes réalistes.

Après des études et des essais en simulation, un procédé de traitement des données a été mis en place. Il est composé de trois grandes parties :

1. le traitement des données de calibration, qui permet de fusionner les connaissances *a priori* avec les mesures expérimentales, afin d'établir un modèle représentatif du système;
2. la reconstruction, qui permet d'inverser le modèle, et d'obtenir les fonctions de diffusion

du fantôme étudié;

3. la classification, qui permet d'attribuer un label "sain" ou "cancéreux" à chaque position du fantôme.

Ce traitement a été utilisé pour traiter des mesures expérimentales, obtenues sur un fantôme en plastique, et sur des fantômes plus réalistes, composés de tissus bovins. Les mesures expérimentales ont été effectuées avec un système exploitant la diffusion cohérente à dispersion d'énergie. Ce système comprend, entre autres, un collimateur multi-fentes et un détecteur spectrométrique multi-pixel à tellure de cadmium dopé au zinc (CdZnTe), qui offre un compromis entre la sensibilité et la résolution spatiale.

Les résultats obtenus avec le fantôme plastique étaient très bons, car nous avons obtenu une très bonne résolution spatiale et une classification très proche de la vérité terrain. Cependant, les matériaux plastiques utilisés s'avèrent plus différents entre eux que les tissus mammaires, en termes de fonctions de diffusion. De plus, les temps d'acquisitions étaient très longs, ce qui a donné des mesures très peu bruitées, mais également une dose absorbée trop importante. C'est pourquoi, par la suite, des fantômes plus réalistes ont été utilisés, et l'influence de la dose absorbée sur les performances de classification a été également étudiée. Pendant cette deuxième campagne expérimentale, nous avons obtenu un résultat stable, en terme de classification, pour un temps de mesure raisonnable, et une dose absorbée proche de la dose déposée en mammographie. Cependant, la classification, ainsi que la résolution spatiale, restent encore insuffisamment précises pour une application clinique.

Plusieurs pistes ont été envisagées afin d'améliorer ces résultats :

- Une amélioration du procédé de mesure des données de calibration pourrait avoir un impact positif sur la résolution spatiale du système. En effet, une matrice système moins bruitée pourrait améliorer la stabilité de la reconstruction.
- Nous avons observé une inhomogénéité de résolution spatiale en fonction de la profondeur. Pour l'application *in vivo*, visée dans ces travaux, un compromis entre la dose acceptable et la résolution spatiale souhaitée doit être choisi. Le collimateur secondaire a été conçu dans l'objectif d'offrir une discrimination spatiale suffisante, même à faible dose. Cependant, nos derniers résultats montrent qu'une amélioration du collimateur serait souhaitable. Par exemple, une lame supplémentaire pourrait être ajoutée afin de réduire la zone  $z$  visible à travers chaque fente.  
Par ailleurs, pour une utilisation *ex vivo* dans un contexte de contrôle de marges, on pourrait augmenter le flux de photons et utiliser une collimation plus résolue.
- Un deuxième ensemble, formé d'un détecteur spectrométrique et d'un collimateur secondaire, pourrait être placé de manière symétrique au premier par rapport à l'axe des rayons X. Cela permettrait de mesurer deux fois plus de photons pour un même temps

d'exposition, et donc une même dose absorbée. En augmentant ainsi la quantité du signal, nous pouvons espérer au final un résultat de classification plus fiable. De plus, le deuxième collimateur pourrait avoir une géométrie légèrement différente du premier, afin de pallier à la perte de résolution spatiale à certaines profondeurs.

- Les derniers résultats obtenus avec les fantômes de viande ont mis en exergue les limites de la classification choisie. Une étude plus approfondie, sur des données expérimentales, serait très intéressante. Notamment, des algorithmes de "deep learning" pourraient être envisagés.

Finalement, il serait intéressant de réaliser des expériences sur tissus mammaires humains, qui n'ont pas pu être faites pendant cette thèse. Un éventuel partenariat avec une structure médicale pourrait faciliter les démarches. De plus, cela permettrait de créer un environnement d'échange entre le milieu médical et le milieu technique, et d'avoir accès en temps réel aux besoins du personnel médical vis-à-vis d'un système de diagnostic du cancer du sein.



# Annexes



## CLASSIFICATION BI-RADS

---

Le système de classification *Breast Imaging-Reporting And Data System* (BI-RADS)<sup>®</sup> a été mis au point pour l'utilisation des cliniciens par le American College of Radiology [5]. La table A.1 montre la classification des seins en fonction de leur densité. La table A.2 comprend la classification utilisée pour évaluer les images issues d'examens mammaires afin de déterminer le risque de cancer et la démarche conseillée.

Type	Description
1	Sein composé presque entièrement de tissu adipeux
2	Sein comportant des zones dispersées de tissu fibroglandulaire
3	Sein dense de manière hétérogène, ce qui peut gêner la détection de petites masses
4	Sein extrêmement dense, ce qui diminue la sensibilité de la mammographie

TABLE A.1 – Table des catégories BI-RADS pour la composition mammaire [5]

	Catégorie	Démarche à suivre	Probabilité de cancer ( $P$ )
0	Classification initiale en attente d'examen		N/A
1	Examen normal	Surveillance normale	$P = 0\%$
2	Anomalies bénignes	Surveillance normale	$P = 0\%$
3	Anomalies probablement bénignes	Surveillance à court terme (3 ou 6 mois) conseillée	$0\% < P \leq 2\%$
4	Anomalie		
	a faiblement suspecte	Biopsie	$2\% < P \leq 10\%$
	b modérément suspecte		$10\% < P \leq 50\%$
	c très suspecte		$50\% < P \leq 95\%$
5	Anomalie évocatrice d'un cancer	Biopsie	$P \geq 95\%$
6	Tumeur maligne confirmée par biopsie	Excision chirurgicale, si applicable	N/A

TABLE A.2 – Table des catégories BI-RADS pour l'évaluation d'images d'examen mammaire [1, 5]

## **ACTES DE CONFÉRENCES**

---

Pendant la thèse, des résultats intermédiaires ont été présentés aux conférences SPIE Medical Imaging de 2018, à Houston, et 2019, à San Diego. Les actes de conférence correspondants peuvent être consultés ci-après.

En 2018, nous avons présenté les résultats d'expérience sur des glandes mammaires de souris, en utilisant une géométrie d'acquisition très simple. Ensuite, en 2019, nous avons montré des résultats expérimentaux avec la géométrie complète, telle qu'elle a été présentée dans cette thèse. Les expériences ont été réalisées avec un fantôme composé d'huile d'olive et de tissu musculaire bovin. Ce fantôme n'a pas été présenté dans le manuscrit de thèse, car, par la suite, il a été jugé peu réaliste. Néanmoins, ce fantôme nous a servi de point de départ pour la mise au point des fantômes en tissus adipeux et musculaires bovins, présentés dans la partie §3.3.

# Ex-vivo mice mammary glands characterization using energy-dispersive X-ray diffraction and spatially resolved CdZnTe detectors

Vera Feldman<sup>a,b</sup>, Joachim Tabary<sup>a,b</sup>, Caroline Paulus<sup>a,b</sup>, Jean-Louis Hazemann<sup>c,d</sup>, Odile Filhol<sup>a,e,f</sup>, Isabelle Vilgrain<sup>a,e,f</sup>, and Loïck Verger<sup>a,b</sup>

<sup>a</sup>Univ. Grenoble Alpes, F-38000 Grenoble, France

<sup>b</sup>CEA, LETI, MINATEC Campus, F-38054 Grenoble, France

<sup>c</sup>Univ. Grenoble Alpes, Inst NEEL, F-38042 Grenoble, France

<sup>d</sup>CNRS, Inst NEEL, F-38042 Grenoble, France

<sup>e</sup>Inserm, UMRS1036, Biology of Cancer and Infection, Grenoble, F-38054, France

<sup>f</sup>CEA, Biosciences & Biotechnology Institute of Grenoble, Biology of Cancer and Infection, Grenoble, F-38054, France

## Abstract

False-positive mammography results in breast cancer diagnosis can lead to unnecessary biopsies which are invasive and time-consuming. X-Ray Diffraction (XRD) has the potential of providing diagnosis-relevant information and thus can be used after a mammography to verify its results and possibly avoid needless biopsy. We present an Energy-Dispersive X-Ray Diffraction (EDXRD) system and data analysis method which allowed us to characterize healthy and cancerous mice mammary glands ex-vivo. Our technique showed decent gland localization along the z-axis as well as scatter signatures coherent with ones previously described in the literature. We used an in-house spatially resolved CdZnTe detector, along with a subpixelation technique which enhances spatial resolution. Acquisition time and dose delivered are to be optimized yet, however our results demonstrate the potential of EDXRD systems for depth-resolved breast imaging. Different geometries and processing algorithms are currently being investigated in the development of a future EDXRD clinical system for breast cancer diagnosis.

**Keywords:** Energy-dispersive X-ray diffraction, X-ray scatter, breast cancer diagnosis, tissue characterization, spectral imaging, CdZnTe detectors, subpixelation.

## 1 Introduction

Mammography is the most used imaging tool for screening and breast cancer diagnosis. It is based on the difference of x-ray absorption in healthy and cancerous tissues. However, in some cases this difference is low and the produced contrast is insufficient for diagnosis.<sup>1</sup> Additional tests are then performed and uncertain or positive results are ultimately confirmed by histological analysis, the golden standard in breast cancer diagnosis. It is reliable but requires biopsies, which are invasive, stressful for patients and time consuming. This is why the minimization of false-positive results in upstream tests is needed.

Usually ultrasound or MRI tests are performed after the mammography, but other techniques such as X-Ray Diffraction (XRD)<sup>2-4</sup> are being investigated to complement the existing ones in the future. Coherently diffracted X-rays carry information on the molecular structure which is different for healthy and cancerous tissue so XRD can be efficiently used for breast tissue characterization. The diffraction intensity plotted against the momentum transfer  $\chi$  forms a signature, called **form factor**, which is tissue-specific and allows the differentiation between healthy and cancerous breast tissues. The momentum transfer is a function of the x-ray energy  $E$  and diffraction angle  $\theta$  according to Bragg's law:

$$\chi = \frac{E}{hc} \cdot \sin(\theta/2) \quad (1)$$

---

Further author information: (Send correspondence to V.F.)

V.F.: E-mail: vera.feldman@cea.fr

Medical Imaging 2018: Physics of Medical Imaging, edited by Joseph Y. Lo, Taly Gilat Schmidt, Guang-Hong Chen, Proc. of SPIE Vol. 10573, 105732B  
© 2018 SPIE · CCC code: 1605-7422/18/\$18 · doi: 10.1117/12.2292866

Proc. of SPIE Vol. 10573 105732B-1

where  $h$  is the Planck's constant and  $c$  is the light speed. In Energy-Dispersive X-Ray Diffraction (EDXRD) the intensity is measured at different energies with a constant diffraction angle in order to cover the desired momentum transfer range. We present an EDXRD setup and demonstrate its differentiation potential on ex-vivo healthy and cancerous mice mammary glands. Different spatial separations between the two tubes were tested in order to determine whether our method is capable of accurate spatial resolution in addition to correct samples characterization.

## 2 Methods

### 2.1 Experimental setup

#### 2.1.1 Mammary tissue samples

The MDA-MB-231 mammary carcinoma cell line was obtained from the American Type Culture Collection.<sup>5</sup> Cells were grown in DMEM medium supplemented with 10% fetal bovine serum plus 100 U/ml penicillin, and 100  $\mu\text{g/ml}$  streptomycin. Six- to eight-week-old female Nude mice (15 – 25 g) were obtained from the Charles River Laboratories. Animal maintenance and experiments were performed in accordance with the animal care guidelines of the European Union and French laws. Mice were injected in the mammary fatpad with  $10^6$  early passage MDA-MB-231-luc-2 cells. Tumor growth was assessed morphometrically using electronic calipers, and tumor volumes were calculated according to the formula  $V (\text{mm}^3) = L (\text{major axis}) \times W^2 (\text{minor axis})/2$ .<sup>6</sup> Animals were sacrificed as soon as tumor size reached 1.0 cm in diameter and normal or xenografted mammary glands were isolated.

#### 2.1.2 Acquisition protocol

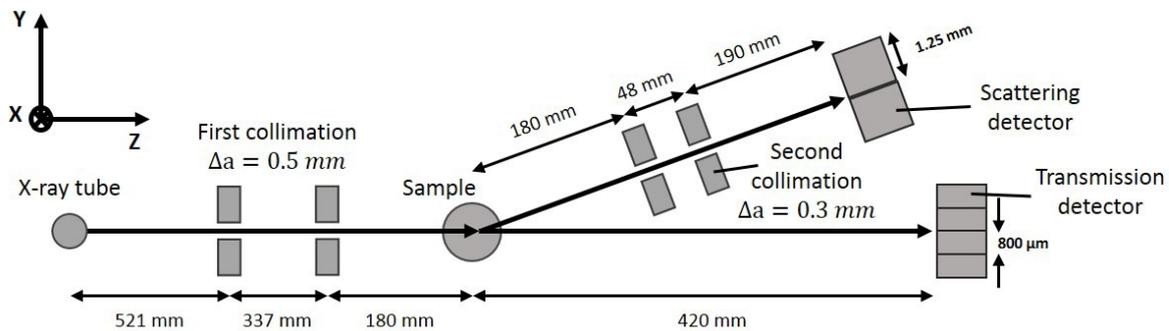


Figure 1: EDXRD and attenuation acquisition setups diagram

The glands were placed in separate 8 mm in diameter plastic tubes with a window cut out along the beam path to eliminate any scattering by the tube itself. The tubes were positioned along the  $z$ -axis on the acquisition bench illustrated on the figures 1 and 2.

Healthy and cancerous glands were positioned in field on view with different spatial spacings between them. Spacings ranging from 0 cm (tubes next to one another) to 5 cm between the tubes by steps of 1 cm were tested to estimate spatial resolution of our system. Diffracted and transmitted spectra were obtained with two separate acquisitions. In both cases the x-ray beam is first collimated into a pencil beam by two slits with a 0.5 mm opening.

For diffraction, a second pair of slits with a 0.3 mm opening was placed between the sample and the detector in order to select an average  $2.5^\circ$  diffraction angle (diffraction angle distribution shown on figure 3b). The in-house energy-resolved "Imadif"<sup>7</sup> detector has an energy resolution around 3% at 122 keV and is composed of 8 columns and 24 rows (X and Y axes on the figure 1 respectively). Shielding plates were placed in order to hide all rows but two, as two rows were sufficient for the purposes of our experiment. Measured spectra were summed along the columns to improve signal over noise ratio. We used subpixelation<sup>8,9</sup> which provides the location of the photon - detector encounter with a spatial resolution below the pixel pitch without compromising sensitivity. With this technique two real pixels with a 1.25 mm pitch are each segmented to eight virtual subpixels with a 0.15 mm pitch.

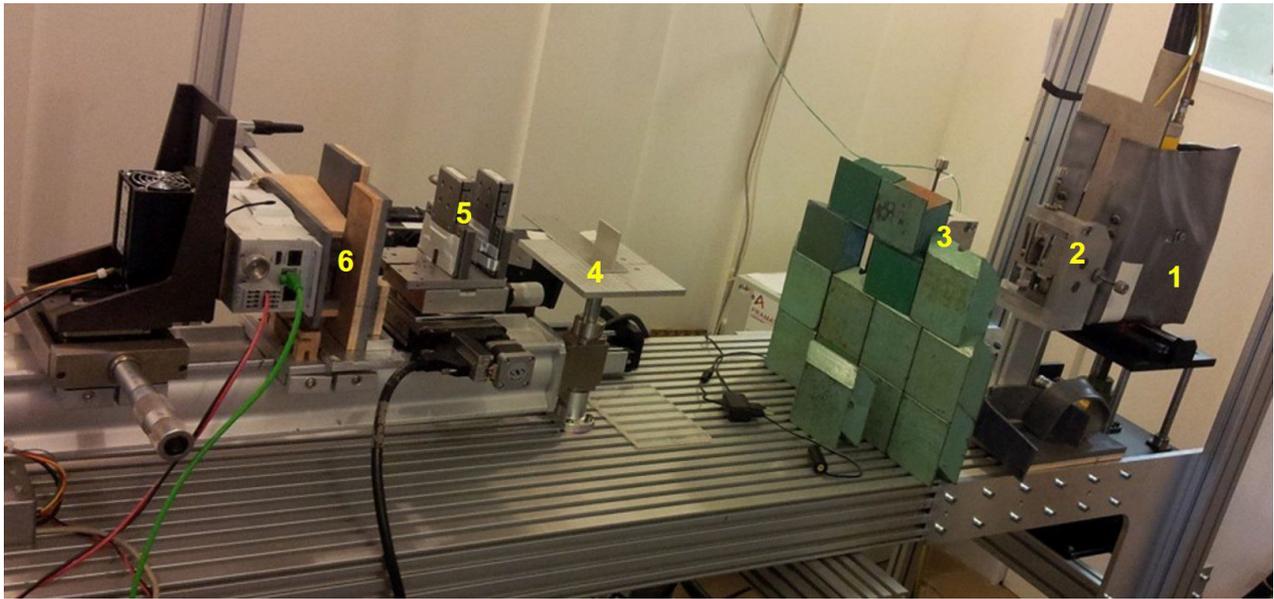
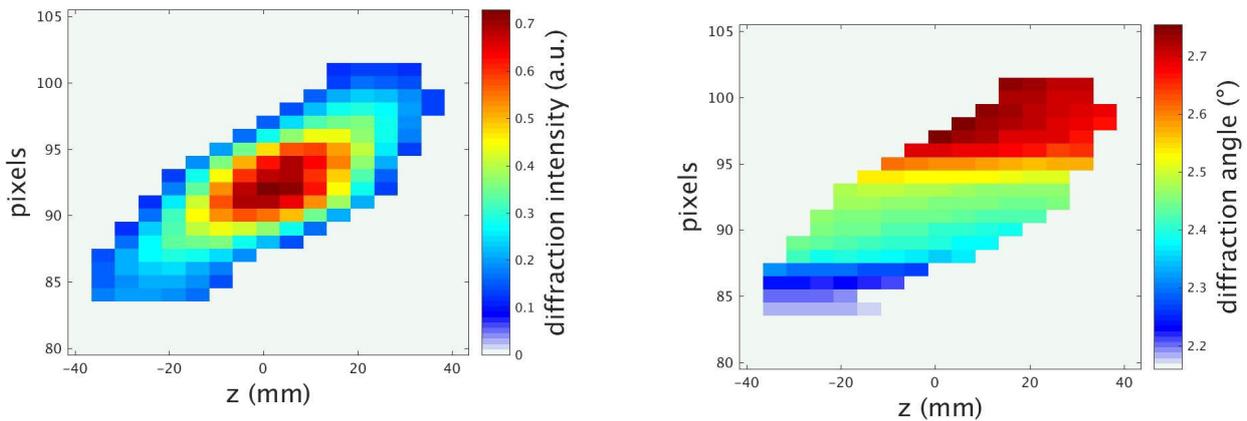


Figure 2: Picture of the experimental EDXRD setup. "1" x-ray tube source, "2" collimation window, "3" first collimation and lead beam shielding, "4" sample plate, "5" second collimation and "6" Imadif detector behind the beam stop



(a) Normalized diffraction intensity as a function of pixels and  $z$  position.

(b) Diffraction angle as a function of pixels and  $z$  position.

Figure 3: Field of view maps. On the left the diffraction intensity perceived by the detector depending on the  $z$  position. On the right the diffraction angle of the detected photons depending on the  $z$  position. 0 on the  $z$  axis is positioned at the center of the field of view which corresponds to the center of the sample on the figure 1.  $z$  axis is oriented in the same direction as the x-ray beam, as illustrated on the figure 1.

Further in the text the subpixels are also referred to as "pixels". The figure 3a shows the normalized diffraction intensity as a function of the pixels and  $z$  position obtained with this geometry. The matrix is diagonal-shaped, each pixel sees a slightly different  $z$  zone, so we can expect a good spatial reconstruction results.

The acquisition time for diffracted spectra was 30 minutes with a tungsten-anode x-ray source operated at 160 kVp and 18.75 mA. An optimization of these settings is to be performed in the near future, and recent results suggest that acquisition time can be greatly reduced. Transmitted spectra were also measured as they are used in the data analysis for attenuation correction. They were acquired over 3 minutes with an older version of the ME-100 Multix<sup>®</sup> detector composed of 16 energy-resolved pixels with a 800 $\mu$ m pitch. The energy resolution of

the detector is about 5.5% at 122 keV. For this part of the acquisition the source was operated at 160 kVp and 1 mA.

## 2.2 Data analysis

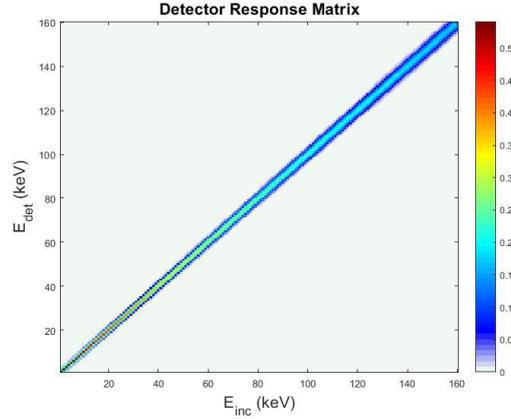


Figure 4: Simulated detector response matrix for Imadif for energies ranging from 0 to 160 keV

Our setup measures the diffracted and attenuated spectra of the samples. Data analysis is needed in order to recover sample's form factor which can then be identified as healthy or cancerous. Our data analysis process has been described in more detail in a previous article.<sup>10</sup>

The detected diffraction spectrum is modeled as follows:

$$S(p, E_{det}) = \sum_{E_{inc}} \left( R(E_{inc}, E_{det}) S_{inc}(E_{inc}) Att(E_{inc}) \sum_{\chi} \left( R_{\theta}(p, \chi, E_{inc}) \sum_z A(p, z) F(z, \chi) \right) \right) \quad (2)$$

where

- $p$  is the pixel,  $E_{det}$  is the detected energy,  $E_{inc}$  is the incident energy,  $\chi$  is the momentum transfer,  $z$  is the position in the object along the incident beam direction;
- $S(p, E_{det})$  is the diffracted spectrum measured by each pixel;
- $F(z, \chi)$  is the form factor;
- $S_{inc}(E_{inc})$  is the incident spectrum emitted by the x-ray tube;
- $Att(E_{inc})$  is the attenuation of the incident beam by the sample. As the diffraction angle is very low, we assumed the attenuation due to passage through the object is the same for transmitted and diffracted spectra;
- $R_{\theta}(p, \chi, E_{inc})$  is the angular resolution of the system;
- $R(E_{inc}, E_{det})$  is the detector response matrix (DRM). It is a 2D distribution which represents the probability that a photon with incident energy  $E_{inc}$  will be detected at a given energy,  $E_{det}$ . It was simulated by interpolating detector responses to americium and cobalt sources (59 and 122 keV respectively) and by assuming a Gaussian distribution (see figure 4);
- $A(p, z)$  is the matrix of intensities detected by the pixels as a function of the position,  $z$ , of the subvolumes viewed in the object. This matrix is shown on the figure 3a. The product  $\sum_z A(p, z) F(z, \chi)$  gives the form factor measured by the pixels  $F(p, \chi)$ .

$R_\theta(p, \chi, E_{inc})$  and  $A(p, z)$  were obtained through system calibration with a 2-mm-thick graphite plate which was translated by 5-mm steps through the field of view along the  $z$  axis. The reconstruction of  $F(z, \chi)$  from  $S(p, E_{det})$  uses the diffraction and transmitted spectra of a known reference material (PMMA here) and is comprised of 4 steps :

1. Transmitted spectra measured by the detector are modeled as a function of the detector response matrix  $R_t(E_{det}, E_{inc})$ , the incident photon flux  $S_{inc}(E_{inc})$  and the attenuation  $Att(E_{inc})$ :

$$S_{att}(E_{det}) = \sum_{E_{inc}} \left( R_t(E_{det}, E_{inc}) S_{inc}(E_{inc}) Att(E_{inc}) \right) \quad (3)$$

Knowing  $R_t$ ,  $S_{inc}$  and  $S_{att}$  we use Alvarez-Makovski property<sup>11</sup> in order to recover the attenuation  $Att(E_{inc})$  from transmitted spectra.

2. A ratio between diffraction and attenuation spectra of the sample and of a known reference material is computed:

$$Ratio(p, E) = \frac{S(p, E)}{S_{ref}(p, E)} \frac{Att_{ref}(E)}{Att(E)} \quad (4)$$

The main purpose of this step is to remove the influence of the incident spectrum, especially the K-edges of the tungsten anode, of the detector response and of the material attenuation.

3. Angular readjustment is used to compute  $Ratio(p, \chi)$  from  $Ratio(p, E)$  using the Bragg's law and mean diffraction angle per pixel (see figure 3b). The obtained ratio is used along with the form factor of the reference material to compute  $F(p, \chi)$  :

$$F(p, \chi) \approx F_{ref}(p, \chi) \cdot Ratio(p, \chi) \quad (5)$$

4. Finally spatial reconstruction is performed to obtain  $F(z, \chi)$  from  $F(p, \chi)$ .

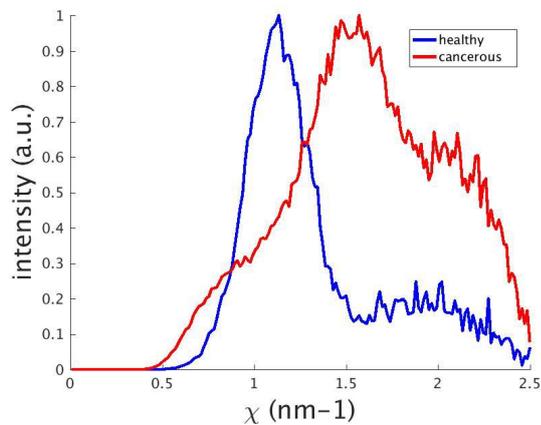
$$F(p, \chi) = A(p, z) F(z, \chi) \quad (6)$$

where  $F(p, \chi)$  et  $A(p, z)$  are known and  $F(z, \chi)$  is to be determined, is an ill-posed problem, as  $A(p, z)$  is not invertible (see figure 3a), so we use Maximum Likelihood Expectation Maximization (MLEM) algorithm to compute  $F(z, \chi)$ .

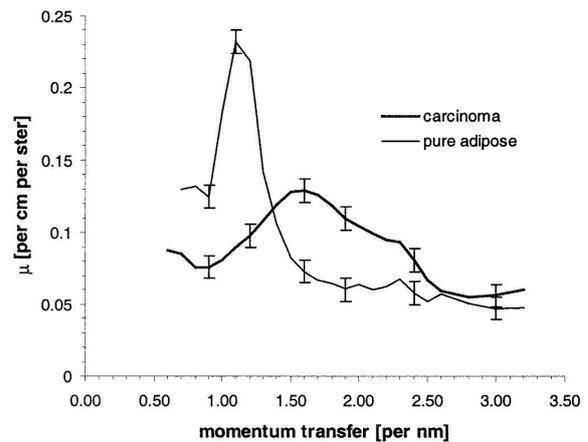
### 3 Results

Form factors for healthy and cancerous glands were obtained by averaging  $F(z, \chi)$  on the  $z$  positions corresponding to healthy and cancerous glands respectively. On the figure 5a are shown the results obtained from the 2 cm separation acquisition. The reference form factors<sup>12</sup> are represented on the figure 5b. Healthy and cancerous glands present a peak at  $1 \text{ nm}^{-1}$  and at  $1.6 \text{ nm}^{-1}$  respectively. They match the peak positions of adipose and pure carcinoma represented on the figure 5b and we can thus conclude that the obtained form factors are consistent with previous studies.

The figure 6 shows scatter signatures obtained with different spatial separations in  $(\chi, z)$  space. The healthy glands are located above the cancerous glands. The dotted lines show the true localization of both tubes. The localization is quite accurate for smaller separations (0 to 2 cm). For wider spacing the localization is much less precise. Indeed, as shown on the figure 3a), most of the detected photons come from the  $z \in [110; 130]$  mm region, while photons coming from the zone closer to the borders of the field of view are much less seen by the detector. The signal-to-noise ratio is thus higher for the center of the field of view and better spatial resolution can be achieved.



(a) Normalized 1D healthy (blue) and cancerous (red) form factors obtained from 2 cm separation acquisition.



(b) Reference 1D form factors for adipose tissue (thin line) and pure carcinoma (bold line).<sup>12</sup>

Figure 5: Form factors of healthy and cancerous mammary glands

#### 4 Conclusion

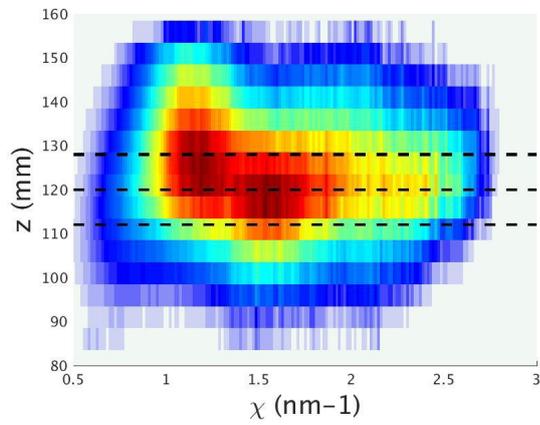
We presented an EDXRD setup and tested its capacity of spatial separation and characterization of healthy and cancerous mice mammary glands as a proof-of-concept of EDXRD potential in breast tissue imaging. The measured form factors of healthy and cancerous glands are consistent with those found in the literature. In-depth spatial resolution requires enhancement, however the results are encouraging considering the geometrical simplicity of our setup. Actually we are exploring more complex collimating geometries and corresponding processing tools, such as coded apertures<sup>3</sup> or multi-focal systems,<sup>13</sup> in order to benefit from the whole detector surface in our future experiments.

Also in the near future a classification algorithm will be integrated into the data analysis process to facilitate interpretation. Furthermore acquisition time and deposited radiation dose will be carefully examined in the perspective of future clinical applications.

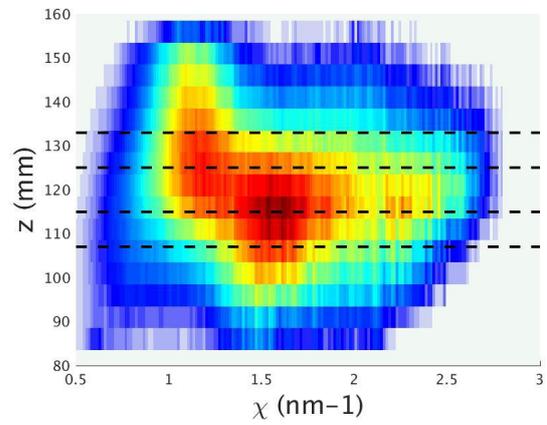
#### REFERENCES

- [1] P. C. Johns and M. J. Yaffe. X-ray characterisation of normal and neoplastic breast tissues. *Physics in Medicine and Biology*, 32(6):675–695, 1987.
- [2] S. Pani. Applications of x-ray diffraction imaging to mammography. *Recent Patents on Medical Imaging*, 2(2):111–119, 2012.
- [3] James R. Spencer, Joshua E. Carter, Crystal K. Leung, Shannon J. McCall, Joel A. Greenberg, and Anuj J. Kapadia. Coded aperture coherent scatter spectral imaging for assessment of breast cancers: an ex-vivo demonstration. In *Progress in Biomedical Optics and Imaging - Proceedings of SPIE*, 2017.
- [4] R. M. Moss, A. S. Amin, C. Crews, C. A. Purdie, L. B. Jordan, F. Iacoviello, A. Evans, R. D. Speller, and S. J. Vinnicombe. Correlation of x-ray diffraction signatures of breast tissue and their histopathological classification. *Scientific Reports*, 7(1), 2017.
- [5] K. J. Chavez, S. V. Garimella, and S. Lipkowitz. Triple negative breast cancer cell lines: One tool in the search for better treatment of triple negative breast cancer. *Breast Disease*, 32(1-2):35–48, 2010.
- [6] G. Carlsson, L. Ekelund, L. Stigsson, and L. Hafstrom. Vascularization and tumour volume estimations of solitary liver tumours in rats. *Annales Chirurgiae et Gynaecologiae*, 72(4):187–191, 1983.
- [7] G. Montémont, D. Kosciesza, O. Monnet, S. Stanchina, J. P. Schlomka, and L. Verger. An autonomous czt module for x-ray diffraction imaging. In *2013 IEEE Nuclear Science Symposium and Medical Imaging Conference (2013 NSS/MIC)*, 2013.
- [8] G. Montémont, S. Lux, O. Monnet, S. Stanchina, and L. Verger. Studying spatial resolution of czt detectors using sub-pixel positioning for spect. *IEEE Transactions on Nuclear Science*, 61(5):2559–2566, 2014.

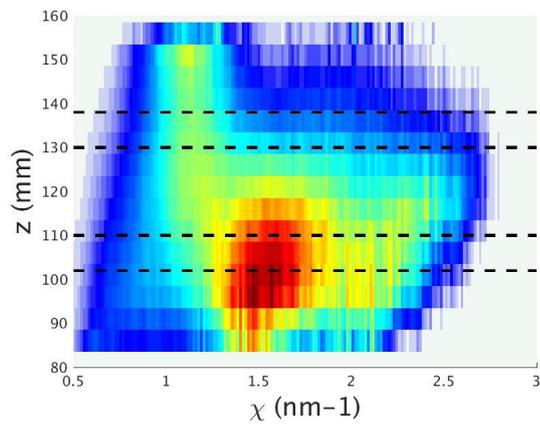
- [9] J. Tabary, C. Paulus, G. Montémont, and L. Verger. Impact of sub-pixelation within cdznte detectors for x-ray diffraction imaging systems. In *Proceedings of SPIE - The International Society for Optical Engineering*, volume 10187, 2017.
- [10] D. Barbes, J. Tabary, C. Paulus, J.-L. Hazemann, and L. Verger. Material-specific imaging system using energy-dispersive x-ray diffraction and spatially resolved cdznte detectors with potential application in breast imaging. *Nuclear Instruments and Methods in Physics Research A*, 848:91–98, 2017.
- [11] R. E. Alvarez and A. Macovski. Energy-selective reconstructions in x-ray computerised tomography. *Physics in Medicine and Biology*, 21(5):733–744, 1976.
- [12] G. Kidane, R. D. Speller, G. J. Royle, and A. M. Hanby. X-ray scatter signatures for normal and neoplastic breast tissues. *Physics in Medicine and Biology*, 44(7):1791–1802, 1999.
- [13] F. Marticke, G. Montémont, C. Paulus, O. Michel, J. I. Mars, and L. Verger. Simulation study of an x-ray diffraction system for breast tumor detection. *Nuclear Instruments and Methods in Physics Research, Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 867:20–31, 2017.



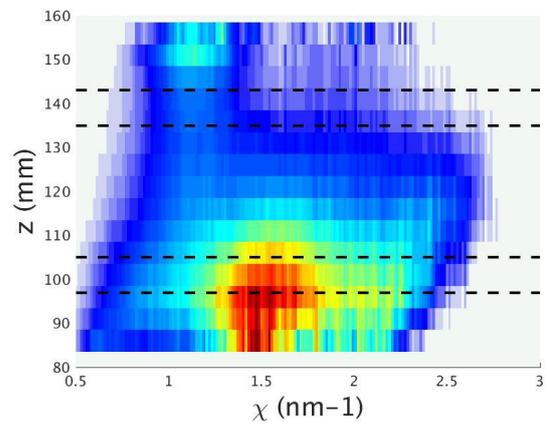
(a) 0 cm separation



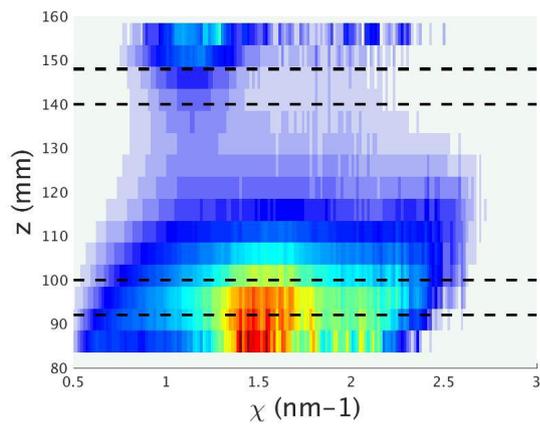
(b) 1 cm separation



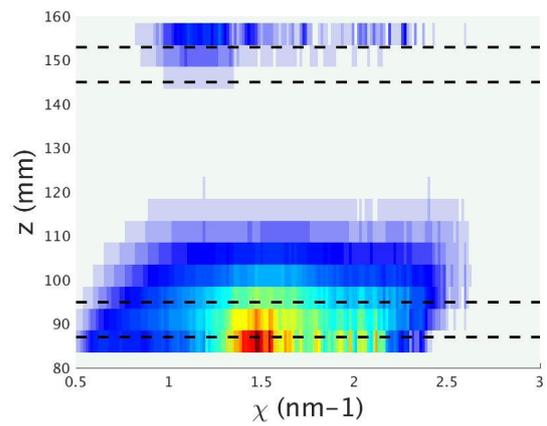
(c) 2 cm separation



(d) 3 cm separation



(e) 4 cm separation



(f) 5 cm separation

Figure 6: Healthy and cancerous mammary form factors in  $z, \chi$  space. On the graphs the healthy gland is located above (higher  $z$ ) the cancerous gland. The dotted lines show the real locations.

# Novel geometry for X-Ray diffraction mammary imaging: experimental validation on a breast phantom

Vera Feldman<sup>a,b</sup>, Joachim Tabary<sup>a,b</sup>, Caroline Paulus<sup>a,b</sup>, and Jean-Louis Hazemann<sup>c,d</sup>

<sup>a</sup>Univ. Grenoble Alpes, F-38000 Grenoble, France

<sup>b</sup>CEA, LETI, MINATEC Campus, F-38054 Grenoble, France

<sup>c</sup>Univ. Grenoble Alpes, Inst NEEL, F-38042 Grenoble, France

<sup>d</sup>CNRS, Inst NEEL, F-38042 Grenoble, France

## ABSTRACT

Mammography is the first tool in breast cancer diagnosis. Its contrast relies on the difference of X-Ray attenuation in healthy and diseased tissues, which is quite limited. This leads to frequent false-positive or inconclusive results and requires further testing. X-Ray Diffraction provides information about molecular structure and can differentiate between healthy and cancerous breast tissues. It can thus be used in synergy with existing imaging methods to provide complementary diagnosis-relevant insight.

We present a novel geometry of such an imaging system and its validation on a breast phantom composed of olive oil and beef muscle, imitating respectively the molecular structure of healthy and cancerous breast tissue. Our system combines energy-dispersive and angle-dispersive X-Ray diffraction by means of an energy-resolved CdZnTe detector and a multi-slit collimation in order to achieve depth-resolved imaging.

The position of the tube with beef muscle inside the oil was varied in this experiment. The obtained results are satisfactory regarding the estimated position of the tube which is very promising for future ex-vivo experiments on human breast tissue samples. Further investigations are carried out on dose reduction and reliable classification algorithms in order to prepare this method for clinical applications.

**Keywords:** Energy-dispersive X-Ray diffraction, Angle-dispersive X-Ray diffraction, X-Ray coherent scatter imaging, breast cancer diagnosis, tissue characterization, spectral imaging, CdZnTe detectors.

## 1. INTRODUCTION

Mammography is the most used tool in breast cancer screening and diagnosis. It is fast (less than 15 minutes for the whole procedure) and has excellent spatial resolution ( $< 100\mu m$ ). Its source of contrast is the difference of X-Ray attenuation coefficients in different types of tissue: micro-calcifications and tumors appear brighter than healthy tissue. However, the difference of X-Ray attenuation in healthy and pathological breast tissue is rather small<sup>1,2</sup> and in some cases, may be insufficient for correct diagnosis.

X-Ray diffraction carries information about molecular structure and provides better contrast between healthy and cancerous breast tissue than attenuation-based techniques.<sup>3</sup> In 2017, Moss et al. demonstrated on breast tissue samples the correlation between the anatomopathological analysis, the golden standard in breast tissue diagnosis, and the data obtained by X-Ray diffraction.<sup>4</sup> It can thus be used as a complementary technique for breast cancer diagnosis or as a tool for surgical margin control in breast surgery.<sup>5-9</sup>

We present the results obtained with a novel geometry which combines energy- and angle-dispersive diffraction. We used a multi-slit secondary collimator for multi-angle selection and a CdZnTe energy-resolved pixelated detector. The diffracted spectra  $S_{diff}$  were measured for different pixels  $p$  and energies  $E$  on a breast phantom made of olive oil and a tube containing beef muscle, imitating healthy and cancerous breast tissue respectively. The measurements were reconstructed in order to assess the *form factors*  $F(z, \chi)$  for different positions  $z$  and momentum transfers  $\chi$ . These form factors reflect the molecular structure and show the difference between oil and muscle. The measurements were realized for different positions of the tube with muscle inside the phantom in order to estimate the system's accuracy regarding spatial resolution.

## 2. METHODS

The goal of our experiment is to estimate the form factor  $F(z, \chi)$  of our phantom,  $z$  being the position along the beam axis, and  $\chi$  the momentum transfer. Its definition is derived from Bragg's law :

$$\chi = \frac{E}{hc} \sin(\theta/2) = \frac{1}{2d} \quad (1)$$

where  $E$  is the energy in keV,  $hc = 1.24 \text{ keV}\cdot\text{nm}$  is the product of the Planck's constant and of the value of light speed in vacuum,  $\theta$  is the angle between incident and diffracted photon directions and  $d$  is the spacing between diffraction planes, used in crystallography. Unlike sharp diffraction peaks observed in crystallography, amorphous materials present a continuous distribution, called *form factor*, noted  $F$  hereafter.

In this section, we detail our acquisition setup and protocol, which yields diffraction spectra  $S_{diff}(p, E)$  depending on detector pixels  $p$  and energy channels  $E$ . We briefly describe the model which links  $S_{diff}(p, E)$  to  $F(z, \chi)$  and the method we use estimate the form factor from diffraction spectra.

### 2.1 Acquisition setup

Our reconstruction process requires both transmission and diffraction measurements. They were acquired separately as the spatial configuration of our experimental bench does not allow simultaneous acquisition, but it will be considered in the future.

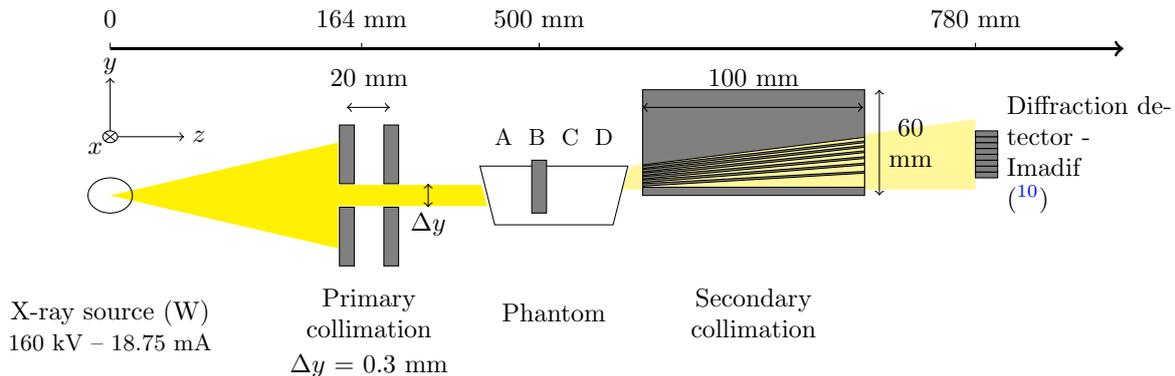


Figure 1: Experimental setup: system geometry (not to scale) for diffraction measurements. The letters above the phantom indicate possible positions of the tube inside the oil container.

#### 2.1.1 Scatter acquisition

The experimental setup is shown schematically on figure 1. The tungsten-anode X-Ray tube was operated at 160 kV and 18.75 mA during 10 minutes. The beam was collimated with two slits into a 2 mm x 0.3 mm ( $x, y$ ) pencil beam before reaching the sample.

The measurements were realized with our in-house energy-resolved detector Imadif.<sup>10</sup> It is composed of 24 rows and 8 columns (along  $y$  and  $x$  directions respectively) on a 660 mm<sup>2</sup> area. For the purposes of our experiment we used only row-wise pixels, the signal received column-wise was summed after angular correction.

Diffracted X-Rays were collimated using a multi-slit collimator designed by our team.<sup>11</sup> It is made of tungsten and has 7 slits with different angular opening aiming at different positions along  $z$ -axis, as can be seen on figure 2. The colored areas show, how the position of the object along  $z$  is translated into pixel position. Each color

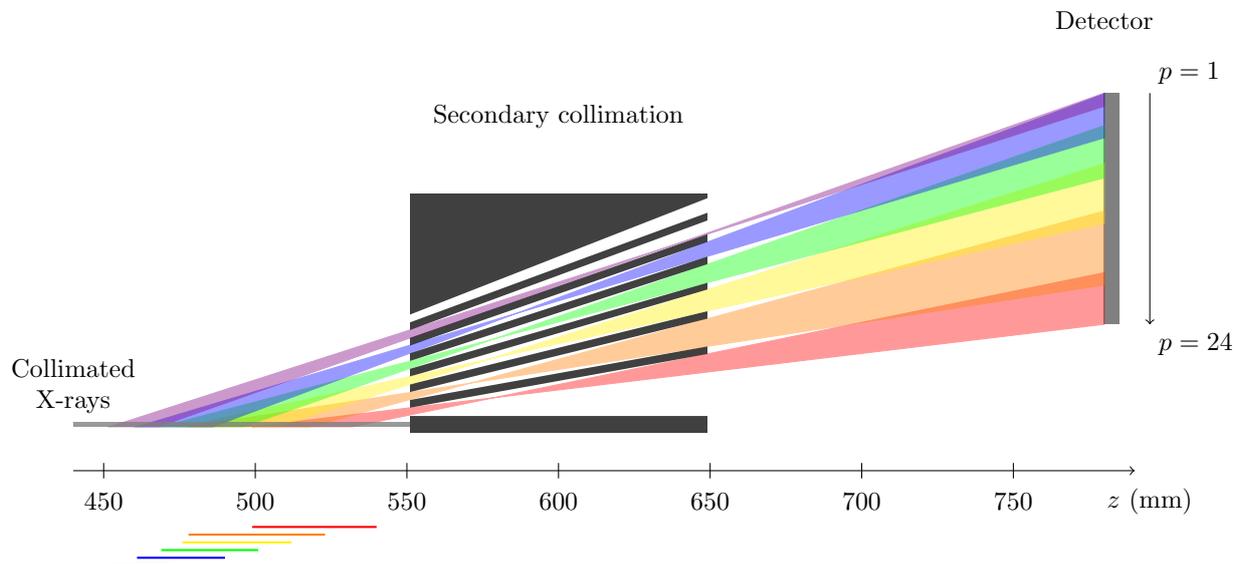


Figure 2: Scaled representation of the  $z$  positions seen by each slit. Below  $z$ -axis are shown the  $z$ -projections of the intersection between the field of view of each slit and the phantom volume illuminated by the X-Rays (gray area).

corresponds to one slit.  $z$  positions "seen" through different splits overlap slightly to cover the whole object area without gaps.

Calibration is required in order to learn the exact relation between  $z$  positions, pixels and diffraction angles. For this purpose we used a 3mm-thick aluminum plate. It was translated along  $z$  by 3 mm steps and diffraction spectra were acquired for each position. We used the aluminum diffraction peak corresponding to  $\chi = 2.14 \text{ nm}^{-1}$ : for each  $z$  position of the calibration plate, we note the pixels, that "see" the aluminum plate at this position  $z$ , and corresponding energies. From this, we infer a matrix  $M(p, z, \theta)$  representing the ratio between the number of photons seen by the pixel  $p$  coming from a position  $z$  at a diffraction angle  $\theta$ , and the total number of detected photons. The 2-dimensional projections of this matrix are shown on the figure 3 : the subfigure (a) shows the relative diffraction intensity by pixel and  $z$  position and the subfigure (b) shows the diffraction angle depending on pixels and  $z$  positions.

These maps show that all the detector pixels are used to image a zone along the  $z$  axis between 480 and 540 mm. The zone between 500 and 540 mm will yield most photons, as this area corresponds to the first slit (lowest on the figure 2), which has the biggest opening. Each pixel "sees" a different zone, even though slight multiplexing is present. However, the diffraction angle varies rapidly between pixels, so we can expect good spatial resolution after reconstruction.

### 2.1.2 Transmission acquisition

The geometry used was exactly the same, except for the secondary collimation, which was removed, and a MultiX ME-100<sup>TM</sup> energy-resolved detector was used instead of the Imadif detector. The X-Ray tube was operated at 160 kV and 1 mA. The transmission acquisition is used to compute the attenuation of the sample, which is used in the reconstruction process to build the forward matrix described in the section 2.4.

## 2.2 Phantom

The phantom is represented schematically in  $(z, y)$  plane on figure 1 and a more detailed top view is shown figure 6. It is made of a thin plastic container filled with olive oil. A plate with cutout holes has been placed on top

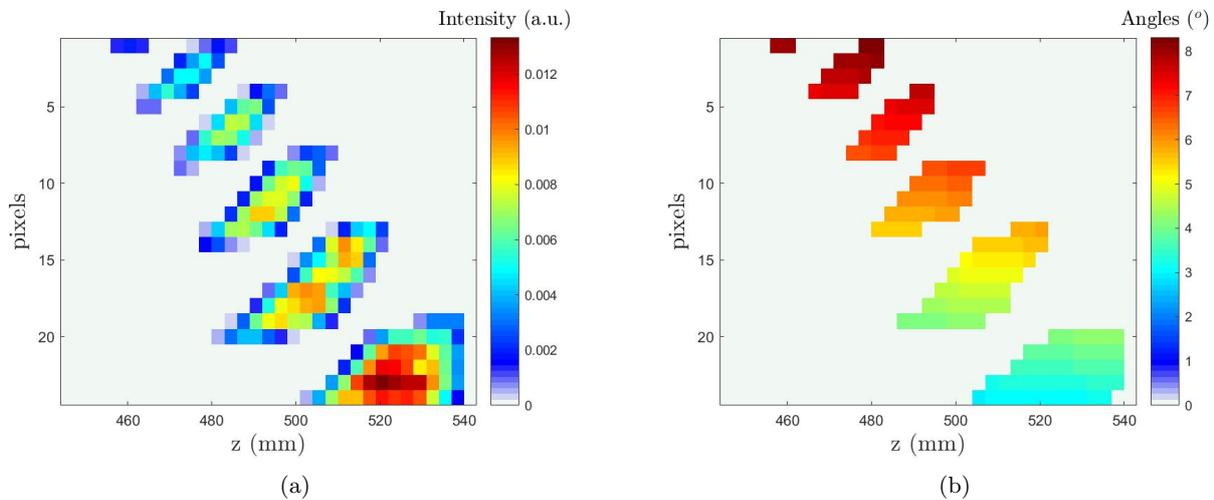


Figure 3: Field of view obtained by calibration : diffraction intensity in a. u. (a) and diffraction angle in degrees (b) per  $z$  position and per pixel  $p$

to hold one plastic tube containing beef muscle tissue at a chosen position along the beam direction. This way, oil occupies all the field of view along  $z$ , except one spot where the tube is placed. These positions are noted with letters going from "A" (closest to the source) to "D" (closest to the detector). On the figure 1 the tube is represented in the position "B". The figure 6 shows  $z$ -coordinates of each hole. During the acquisitions only one of the holes is occupied by the tube and oil occupies all the other  $z$ -positions.

Separation between the centers of consecutive holes is 15 mm. Tube's inner diameter measures 8 mm and its position along  $x$  axis is centered on the X-Ray beam.

Normalized reference form factors for adipose, well-differentiated cancer<sup>12</sup> and beef muscle<sup>13</sup> are shown on figure 4. Olive oil has been chosen because its form factor should be close to adipose tissue and present a peak at  $\chi = 1.1 \text{ nm}^{-1}$  as does healthy breast tissue.<sup>3</sup> Beef muscle form factor has a fibrous peak<sup>3</sup> ( $\chi = 1.6 \text{ nm}^{-1}$ ) and imitates cancerous breast tissue's form factor.

### 2.3 Experimental protocol

The oil container remained at the same position through all the experiment so that the hole "B" was centered in  $z = 500 \text{ mm}$ . 4 diffraction acquisitions were realized for different locations of the tube with beef muscle inside the oil container. An acquisition of olive oil without muscle has also been performed for reference. All diffraction acquisitions lasted 10 minutes.

### 2.4 Data analysis

The relation between form factor  $F(z, \chi)$  and measured diffracted spectrum  $S_{diff}(p, E)$  is modeled as follows:

$$S_{diff}(p, E) = R(p, E, z, \chi) \cdot F(z, \chi) \quad (2)$$

where  $p$  is the pixel number,  $E$  the detected energy,  $z$  the position along the beam axis and  $\chi$  the momentum transfer. The forward matrix  $R(p, E, z, \chi)$  models our whole experimental setup. It converts the form factor expressed in  $z, \chi$  space into a diffracted spectrum, which is measured in  $p, E$  space. This matrix regroups the geometry, described by the matrix  $M(p, z, \theta)$ , the detector response matrix and the incident X-Ray spectrum attenuated by the sample. The geometry matrix  $M(p, z, \theta)$  and the detector response matrix are both obtained by specific calibrations. The attenuated X-Ray beam is estimated using transmission acquisitions.

Inverting the model (2) is an ill-posed problem. In order to reconstruct  $F$  from  $S_{diff}$  we use a Maximum Likelihood Expectation Maximization (MLEM) algorithm.<sup>14</sup>

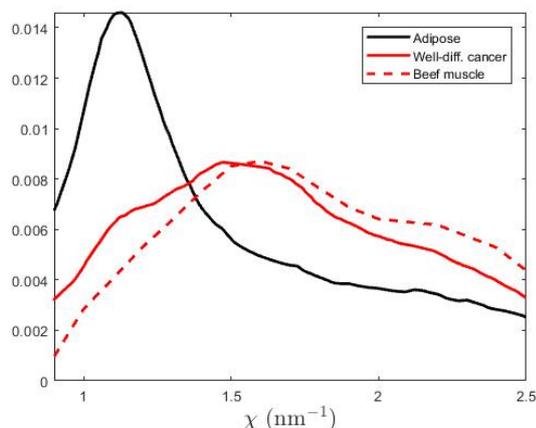


Figure 4: Normalized reference form factors of well-differentiated cancer (red line), beef muscle (red, dashed) and adipose (black)<sup>12,13</sup>

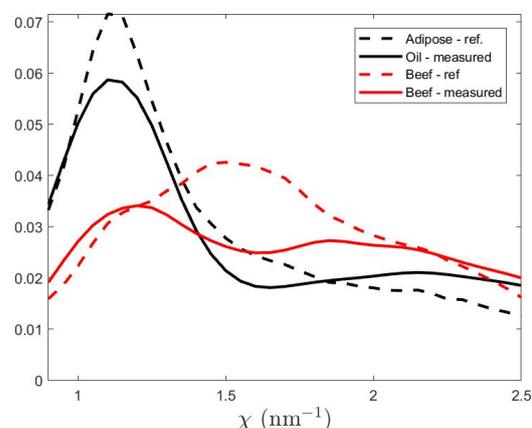


Figure 5: Measured form factors : oil (black) and beef muscle (red); Reference form factors: adipose (black, dashed) and beef muscle (red, dashed)<sup>12,13</sup>

### 3. RESULTS

Two-dimensional form factors  $F(z, \chi)$  were reconstructed from  $S_{diff}(p, E)$  and one-dimensional normalized form factors  $F(\chi)$  have been obtained by averaging the 2D reconstructions on the respective locations of beef muscle and oil. They are presented on figure 5 along with reference form factors.<sup>12,13</sup> As expected oil has a distinct peak around  $1.1 \text{ nm}^{-1}$  and is very similar to the reference. For beef muscle measured and reference form factors are both rather flat, compared to oil and adipose, but not very similar, which is probably due to the oil signature smearing over the position of the tube. Still the difference between oil and beef muscle is clearly visible.

Figure 7 shows the reconstructed 2D form factors for tube positions from "A" to "D". For better visualization the form factors were corrected along  $z$  direction with the form factor of a homogeneous material (oil in this case) in order to remove some irregularities coming from the reconstruction process.

Adipose and beef muscle peak positions at, respectively,  $\chi = 1.1 \text{ nm}^{-1}$  and  $\chi = 1.6 \text{ nm}^{-1}$  are shown by vertical dashed lines (blue and red for adipose and beef respectively). Oil has a well-defined peak around this position. Black dashed lines indicate a 8mm-wide zone where the tube with beef muscle was positioned. We observe a zone with a flat form factor shifting depending on the tube position. This zone and the theoretical position of the tube are very similar which demonstrates the spatial accuracy of our system. Olive oil and beef muscle imitate respectively the form factors of healthy and cancerous breast tissue and these results are very encouraging for future acquisitions and diagnosis on human breast tissue.

### 4. CONCLUSION

We have presented a novel diffraction geometry, which combines angle- and energy-dispersive methods, and tested it on a 6cm-thick breast phantom. Reconstructed form factors  $F(z, \chi)$  show accurate location of the 8mm-wide tube inside the oil container for each acquisition. This demonstrates spatial resolution within centimeter range. This is promising for future tests on human breast tissue samples.

A classification algorithm is going to be implemented in the near future for easier interpretation of the results. Short acquisition times and deposited dose are also two important criteria for clinical imaging. An experimental study is ongoing in order to determine the minimum dose and time, required to obtain insightful analysis.

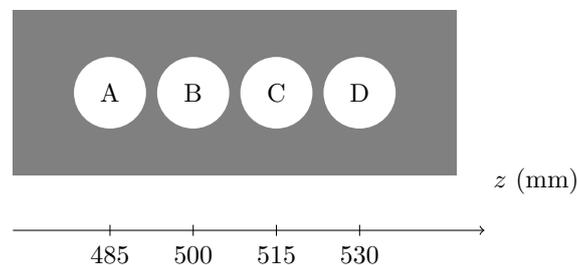


Figure 6: Top view ( $x - z$  plane) of the phantom. In each separate acquisition the tube with beef muscle is located in one of the lettered holes and oil fills all the other positions.

## REFERENCES

- [1] Johns, P. C. and Yaffe, M. J., "X-ray characterisation of normal and neoplastic breast tissues," *Physics in Medicine and Biology* **32**(6), 675–695 (1987).
- [2] Chen, R., Longo, R., Rigon, L., Zanconati, F., De Pellegrin, A., Arfelli, F., Dreossi, D., Menk, R.-H., Vallazza, E., Xiao, T., and Castelli, E., "Measurement of the linear attenuation coefficients of breast tissues by synchrotron radiation computed tomography," *Physics in Medicine and Biology* **55**(17) (2010).
- [3] Kidane, G., Speller, R. D., Royle, G. J., and Hanby, A. M., "X-ray scatter signatures for normal and neoplastic breast tissues," *Physics in Medicine and Biology* **44**(7), 1791–1802 (1999).
- [4] Moss, R. M., Amin, A. S., Crews, C., Purdie, C. A., Jordan, L. B., Iacoviello, F., Evans, A., Speller, R. D., and Vinnicombe, S. J., "Correlation of x-ray diffraction signatures of breast tissue and their histopathological classification," *Scientific Reports* **7**(1) (2017).
- [5] Pani, S., "Applications of x-ray diffraction imaging to mammography," *Recent Patents on Medical Imaging* **2**(2), 111–119 (2012).
- [6] Barbes, D., Tabary, J., Paulus, C., Hazemann, J.-L., and Verger, L., "Material-specific imaging system using energy-dispersive x-ray diffraction and spatially resolved cdznte detectors with potential application in breast imaging," *Nuclear Instruments and Methods in Physics Research A* **848**, 91–98 (2017).
- [7] Marticke, F., Montémont, G., Paulus, C., Michel, O., Mars, J. I., and Verger, L., "Simulation study of an x-ray diffraction system for breast tumor detection," *Nuclear Instruments and Methods in Physics Research, Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **867**, 20–31 (2017).
- [8] Feldman, V., Tabary, J., Paulus, C., Hazemann, J.-L., Filhol, O., Vilgrain, I., and Verger, L., "Ex-vivo mice mammary glands characterization using energy-dispersive x-ray diffraction and spacially resolved CdZnTe detectors," *Proc.SPIE* **10573** (2018).
- [9] Spencer, J. R., Carter, J. E., Leung, C. K., McCall, S. J., Greenberg, J. A., and Kapadia, A. J., "Coded aperture coherent scatter spectral imaging for assessment of breast cancers: an ex-vivo demonstration," in [*Progress in Biomedical Optics and Imaging - Proceedings of SPIE*], (2017).
- [10] Montémont, G., Kosciuszka, D., Monnet, O., Stanchina, S., Schlomka, J. P., and Verger, L., "An autonomous czt module for x-ray diffraction imaging," in [*2013 IEEE Nuclear Science Symposium and Medical Imaging Conference (2013 NSS/MIC)*], (2013).
- [11] Barbes, D., "Nouveaux systèmes d'imagerie médicale exploitant la diffraction x en dispersion d'énergie à l'aide de détecteurs spectrométriques CdZnTe," (2016).
- [12] Pani, S., Cook, E. J., Horrocks, J. A., Jones, J. L., and Speller, R. D., "Characterization of breast tissue using energy-dispersive x-ray diffraction computed tomography," *Applied Radiation and Isotopes* **68**(10), 1980–1987 (2010).

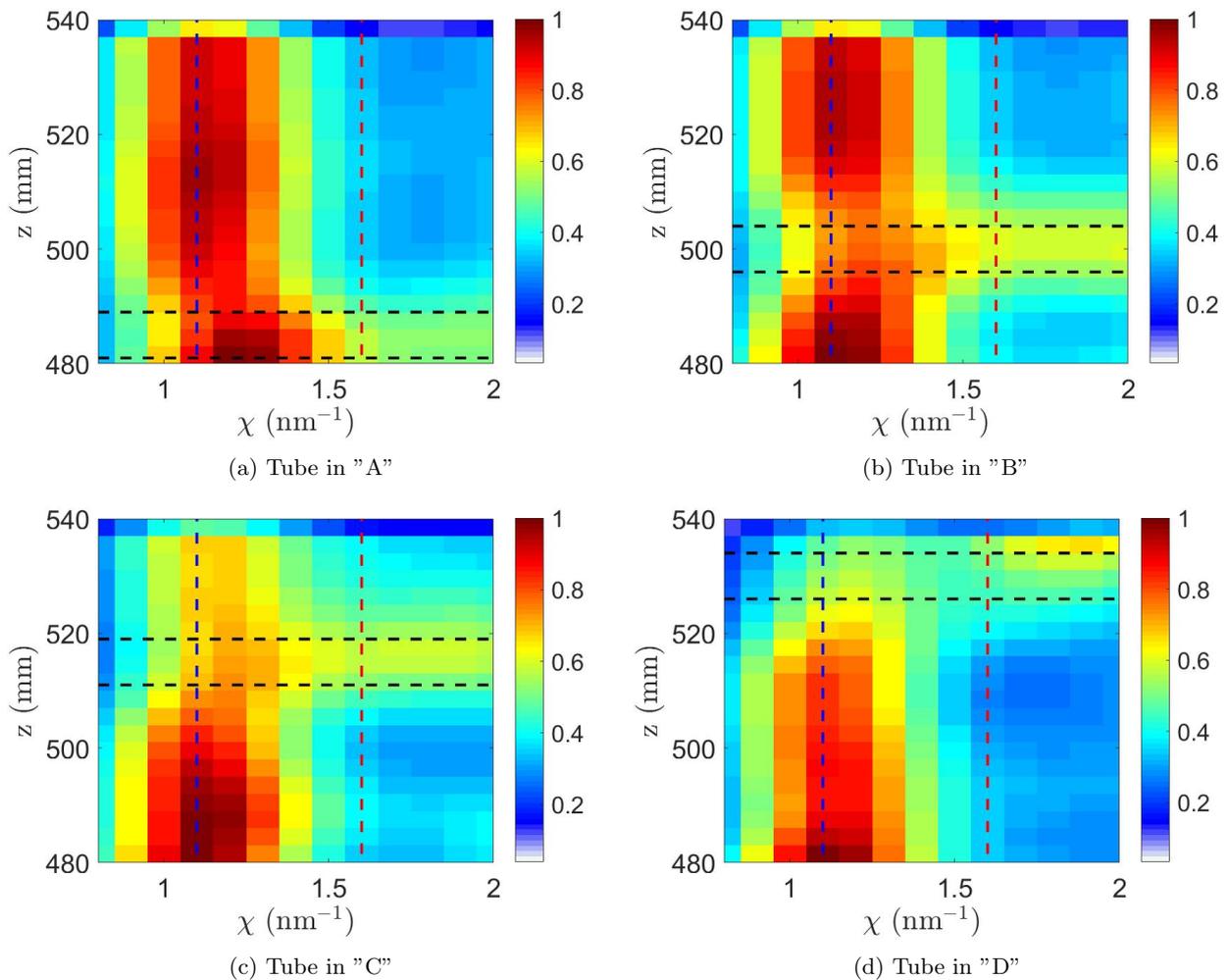


Figure 7: Reconstruction results for tube positions "A" (a) to "D" (d). Vertical dashed lines show the theoretical peak positions of adipose (blue) and beef muscle (red) positions. Horizontal dashed lines delimit the real position of the tube containing beef muscle tissue.

- [13] Peplow, D. and Verghese, K., "Measured molecular coherent scattering form factors of animal tissues, plastics and human breast tissue," *Physics in Medicine and Biology* **43**(9), 2431 (1998).
- [14] Shepp, L. A. and Vardi, Y., "Maximum likelihood reconstruction for emission tomography," *IEEE Transactions on Medical Imaging* **1**(2), 113–122 (1982).



## BIBLIOGRAPHIE

---

- [1] Institut National du Cancer. Les cancers en France. [www.e-cancer.fr](http://www.e-cancer.fr), 2017.
- [2] C. I. Li, K. E. Malone, B. S. Saltzman, and J. R. Daling. Risk of invasive breast carcinoma among women diagnosed with ductal carcinoma in situ and lobular carcinoma in situ, 1988-2001. *Cancer*, 106(10) :2104–2112, 2006.
- [3] L. Wilkinson, V. Thomas, and N. Sharma. Microcalcification on mammography : approaches to interpretation and biopsy. *The British journal of radiology*, 90(1069), 2016.
- [4] Breastcancer.org. Types of breast cancer. <https://www.breastcancer.org/symptoms/types>, 2018.
- [5] E.A. Sickles, C.J. D’Orsi, and L.W. Bassett. ACR BI-RADS® Mammography. In *ACR BI-RADS® Atlas, Breast Imaging Reporting and Data System*. Reston, VA, 2013.
- [6] J. Stines, A. Noel, and P. Heid. La mammographie numérique : paramètres utiles à l’évaluation des systèmes et offre industrielle en France. *Journal de Radiologie*, 88(7) :933–942, 2007.
- [7] D.R. Dance and I. Sechopoulos. Dosimetry in x-ray-based breast imaging. *Physics in Medicine & Biology*, 61(19) :R271–R304, 2016.
- [8] M.E. Suleiman, P.C. Brennan, and M.F. McEntee. Mean glandular dose in digital mammography : a dose calculation method comparison. *Journal of Medical Imaging*, 4(1), 2017.
- [9] S. V. Sree, E. Y.-K. Ng, R. U. Acharya, and O. Faust. Breast imaging : A survey. *World Journal of Clinical Oncology*, pages 171–178, 2011.
- [10] C. Dromain, C. Balleyguier, G. Adler, J.R. Garbay, and S. Delaloge. Contrast-enhanced digital mammography. *European Journal of Radiology*, 69(1) :34–42, 2009.
- [11] F. Diekmann, M. Freyer, S. Diekmann, E. M. Fallenberg, T. Fischer, U. Bick, and A. Pöllinger. Evaluation of contrast-enhanced digital mammography. *European Journal of Radiology*, 78(1) :112–121, 2011.
- [12] Y. Pavia. *Détecteurs spectrométriques pour la mammographie et traitements associés*. PhD thesis, Université de Lyon, 2017.

- 
- [13] B.K. Patel, M.B.I. Lobbes, and J. Lewin. Contrast enhanced spectral mammography : A review. *Seminars in Ultrasound, CT and MRI*, 39(1) :70—79, Feb 2018.
- [14] R.W. Bouwman, R.E. van Engen, K.C. Young, G.J. den Heeten, M.J.M. Broeders, S. Schopphoven, C.R.L.P.N. Jeukens, W.J.H. Veldkamp, and D.R. Dance. Average glandular dose in digital mammography and digital breast tomosynthesis : comparison of phantom and patient data. *Physics in Medicine & Biology*, 60(20) :7893—7907, 2015.
- [15] L. Li, R. Roth, P. Germaine, S. Ren, M. Lee, K. Hunter, E. Tinney, and L. Liao. Contrast-enhanced spectral mammography (cesm) versus breast magnetic resonance imaging (mri) : A retrospective comparison in 66 breast lesions. *Diagnostic and Interventional Imaging*, 98(2) :113—123, 2017.
- [16] American Cancer Society. Breast biopsy. <https://www.cancer.org/cancer/breast-cancer/screening-tests-and-early-detection/breast-biopsy.html>.
- [17] J.A. Seibert and J.M. Boone. X-ray imaging physics for nuclear medicine technologists. part 2 : X-ray interactions and image formation. *Journal of Nuclear Medicine Technology*, 33(1) :3—18, 2005.
- [18] D. Barbes. *Nouveaux systèmes d'imagerie médicale exploitant la diffraction X en dispersion d'énergie à l'aide de détecteurs spectrométriques CdZnTe*. PhD thesis, Université Grenoble Alpes, 2016.
- [19] A. Tartari, A. Taibi, C. Bonifazzi, and C. Baraldi. Updating of form factor tabulations for coherent scattering of photons in tissues. *Physics in Medicine and Biology*, 47(1) :163—175, 2002.
- [20] G. Montémont, D. Kosciesza, O. Monnet, S. Stanchina, J. P. Schlomka, and L. Verger. An autonomous czts module for x-ray diffraction imaging. In *2013 IEEE Nuclear Science Symposium and Medical Imaging Conference (2013 NSS/MIC)*, 2013.
- [21] S. Pani, E. J. Cook, J. A. Horrocks, J. L. Jones, and R. D. Speller. Characterization of breast tissue using energy-dispersive x-ray diffraction computed tomography. *Applied Radiation and Isotopes*, 68(10) :1980—1987, 2010.
- [22] D. O'Flynn, C. B. Reid, C. Christodoulou, M. D. Wilson, M. C. Veale, P. Seller, D. Hills, H. Desai, B. Wong, and R. Speller. Explosive detection using pixellated x-ray diffraction (pixd). *Journal of Instrumentation*, 8(3), 2013.
- [23] K. MacCabe, K. Krishnamurthy, A. Chawla, D. Marks, E. Samei, and D. Brady. Pencil beam coded aperture x-ray scatter imaging. *Optics Express*, 20(15) :16310—16320, 2012.
- [24] G. Harding and J. Kosanetzky. Elastic scatter computed tomography. *Physics in Medicine & Biology*, 30(2) :183, 1985.
- [25] Zheyuan Zhu, Alexander Katsevich, Anuj J. Kapadia, Joel A. Greenberg, and Shuo Pang. X-ray diffraction tomography with limited projection information. *Scientific Reports*, 8(1) :522, 2018.

- 
- [26] D. O'Flynn, C. Crews, I. Drakos, C. Christodoulou, M. D. Wilson, M. C. Veale, P. Seller, and R. D. Speller. Materials identification using a small-scale pixellated x-ray diffraction system. *Journal of Physics D : Applied Physics*, 49(17), 2016.
- [27] C. C. E. Crews, D. O'Flynn, A. Sidebottom, and R. D. Speller. Quantitative energy-dispersive x-ray diffraction for identification of counterfeit medicines : A preliminary study. In *Proceedings of SPIE*, volume 9482, 2015.
- [28] I. Drakos, P. Kenny, T. Fearn, and R. Speller. Multivariate analysis of energy dispersive x-ray diffraction data for the detection of illicit drugs in border control. *Crime Science*, 6(1), 2017.
- [29] R. M. Moss, A. S. Amin, C. Crews, C. A. Purdie, L. B. Jordan, F. Iacoviello, A. Evans, R. D. Speller, and S. J. Vinnicombe. Correlation of x-ray diffraction signatures of breast tissue and their histopathological classification. *Scientific Reports*, 7(1), 2017.
- [30] K. Pearson. On lines and planes of closest fit to systems of points in space. *Philosophical Magazine*, 2 :559–572, 1901.
- [31] E. E. Fenimore and T. M. Cannon. Coded aperture imaging with uniformly redundant arrays. *Appl. Opt.*, 17(3) :337–347, 1978.
- [32] M. N. Lakshmanan, J. A. Greenberg, E. Samei, and A. J. Kapadia. Experimental implementation of coded aperture coherent scatter spectral imaging of cancerous and healthy breast tissue samples. In *Proceedings of SPIE*, volume 9412, 2015.
- [33] T. M. Cannon and E. E. Fenimore. Coded aperture imaging : many holes make light work. *Optical Engineering*, 19(3) :283–289, 1980.
- [34] J. A. Greenberg, K. Krishnamurthy, and D. Brady. Snapshot molecular imaging using coded energy-sensitive detection. *Optics Express*, 21(21) :25480–25491, 2013.
- [35] G. Kidane, R. D. Speller, G. J. Royle, and A. M. Hanby. X-ray scatter signatures for normal and neoplastic breast tissues. *Physics in Medicine & Biology*, 44(7) :1791–1802, 1999.
- [36] F. Marticke. *Optimization of an X-ray diffraction imaging system for medical and security applications*. PhD thesis, Université Grenoble Alpes, 2016.
- [37] D. Barbes, J. Tabary, C. Paulus, J.-L. Hazemann, and L. Verger. Material-specific imaging system using energy-dispersive x-ray diffraction and spatially resolved cdznte detectors with potential application in breast imaging. *Nuclear Instruments and Methods in Physics Research A*, 848 :91–98, 2017.
- [38] A. Brambilla, P. Ouvrier-Buffer, J. Rinkel, G. Gonon, C. Boudou, and L. Verger. Cdte linear pixel x-ray detector with enhanced spectrometric performance for high flux x-ray imaging. In *2011 IEEE Nuclear Science Symposium Conference Record*, pages 4825–4828, 2011.
- [39] G. Montémont, S. Lux, O. Monnet, S. Stanchina, and L. Verger. Studying spatial resolution of cztn detectors using sub-pixel positioning for spect. *IEEE Transactions on Nuclear Science*, 61(5) :2559–2566, 2014.

- 
- [40] J. Tabary, C. Paulus, G. Montémont, and L. Verger. Impact of sub-pixelation within cdznte detectors for x-ray diffraction imaging systems. In *Proceedings of SPIE*, volume 10187, 2017.
- [41] D.E. Peplow and K. Verghese. Measured molecular coherent scattering form factors of animal tissues, plastics and human breast tissue. *Physics in Medicine & Biology*, 43(9) :2431, 1998.
- [42] B.W. King, K.A. Landheer, and P.C. Johns. X-ray coherent scattering form factors of tissues, water and plastics using energy dispersion. *Physics in Medicine & Biology*, 56(14) :4377–4397, 2011.
- [43] F. Marticke, G. Montémont, C. Paulus, O. Michel, J. I. Mars, and L. Verger. Simulation study of an x-ray diffraction system for breast tumor detection. *Nuclear Instruments and Methods in Physics Research, Section A : Accelerators, Spectrometers, Detectors and Associated Equipment*, 867 :20–31, 2017.
- [44] F. Salvat, J. Fernández-Varea, and J. Sempau. Penelope2008 : A code system for monte carlo simulation of electron and photon transport. In *Workshop Proceedings, Data Bank NEA/NSC/DOC(2011)5, Barcelona*, 2006.
- [45] L.A. Shepp and Y. Vardi. Maximum likelihood reconstruction for emission tomography. *IEEE Transactions on Medical Imaging*, 1(2) :113–122, 1982.
- [46] G. Montemont, M.-C. Gentet, O. Monnet, J. Rustique, and L. Verger. Simulation and design of orthogonal capacitive strip cdznte detectors. *Nuclear Science, IEEE Transactions on*, 54 :854 – 859, 09 2007.
- [47] R. Birch and M. Marshall. Computation of bremsstrahlung x-ray spectra and comparison with spectra measured with a ge(li) detector. *Physics in Medicine and Biology*, 24(3) :505–517, 1979.
- [48] R. E. Alvarez and A. Macovski. Energy-selective reconstructions in x-ray computerized tomography. *Physics in Medicine & Biology*, 21(2) :733–744, 1976.
- [49] N. Otsu. A threshold selection method from gray-level histograms. *IEEE Transactions on Systems, Man, and Cybernetics*, 9(1) :62–66, 1979.
- [50] Wikipedia contributors. K-nearest neighbors algorithm — Wikipedia, the free encyclopedia. [https://en.wikipedia.org/w/index.php?title=K-nearest\\_neighbors\\_algorithm&oldid=899016278](https://en.wikipedia.org/w/index.php?title=K-nearest_neighbors_algorithm&oldid=899016278), 2019.
- [51] Wikipedia contributors. Support-vector machine — Wikipedia, the free encyclopedia. [https://en.wikipedia.org/w/index.php?title=Support-vector\\_machine&oldid=904250490](https://en.wikipedia.org/w/index.php?title=Support-vector_machine&oldid=904250490), 2019.
- [52] G. Ho. Linear discriminant analysis for starters. <https://eigenfoo.xyz/lda/>, 2017.
- [53] J. Valentin. The 2007 recommendations of the international commission on radiological protection. *Annals of ICRP*, 103 :64–65, 2007.

- [54] M. Muuss. BRL-CAD. <https://brlcad.org/>, 1979–2019.
- [55] K. J. Chavez, S. V. Garimella, and S. Lipkowitz. Triple negative breast cancer cell lines : One tool in the search for better treatment of triple negative breast cancer. *Breast Disease*, 32(1-2) :35–48, 2010.

## Résumé

Les méthodes actuelles d'imagerie, telles que la mammographie ou les ultrasons donnent parfois des résultats peu concluants et nécessitent des biopsies pour confirmer l'absence de pathologie. Cela mène à un grand nombre de biopsies, qui sont invasives, stressantes, et prennent du temps. Pour répondre à ce problème, de nouvelles méthodes de diagnostic, dont la diffusion cohérente de rayons X, sont explorées. La diffusion cohérente de rayons X reflète la structure moléculaire des tissus, rendant cette technique particulièrement intéressante dans le contexte d'amélioration de diagnostic.

Cette thèse étudie l'exploitation de la diffusion cohérente de rayons X pour affiner la caractérisation *in vivo* des tissus mammaires, en vue d'une future utilisation en complément d'une mammographie suspecte. Un montage expérimental d'acquisition de spectres diffusés et une méthode de traitement de données associée allant jusqu'à la classification sont proposés. La dose déposée est également étudiée, dans la perspective d'une utilisation *in vivo*. La méthode complète fait l'objet de validations expérimentales sur fantômes plastiques et biologiques. Les résultats obtenus sont très encourageants, cependant des améliorations seront nécessaires pour atteindre les exigences cliniques. Quelques pistes pour la suite de ce projet sont également proposées.

**Mots-clés :** application médicale, diffusion cohérente de rayons X, traitement des données

## Abstract

Current breast imaging methods, such as mammography or ultrasounds, sometimes yield inconclusive results and require a biopsy to confirm the absence of pathology. This leads to a large number of biopsies, which are invasive, stressful and time-consuming. To address this problem, new diagnostic methods, including coherent X ray scattering, are explored. Coherent X ray scattering is a promising method for improving diagnosis, as it yields information about the molecular structure of the breast tissues.

This thesis explores the use of coherent X ray scattering to refine *in vivo* breast tissue characterization, as a supplementary technique, in addition to suspicious mammography. An experimental set-up for coherent X ray scattering spectra acquisition and an associated data processing methods, including classification, are proposed. The delivered dose is also studied in view of an *in vivo* application. The complete method is experimentally validated on plastic and biological phantoms. Obtained results are very encouraging, however improvement is still needed in order to meet clinical requirements. A few ways of improvements are suggested for the continuation of this project.

**Key words:** medical application, coherent X ray scattering, data processing