HAL
open science

Novel approaches for photoacoustic imaging of blood vessels : 3D quantitative fluctuation imaging and deep learning assisted reconstruction

Guillaume Godefroy

▶ **To cite this version:**

Guillaume Godefroy. Novel approaches for photoacoustic imaging of blood vessels : 3D quantitative fluctuation imaging and deep learning assisted reconstruction. Physique [physics]. Université Grenoble alpes, 2021. Français. NNT : . tel-03563766

HAL Id: tel-03563766

https://theses.hal.science/tel-03563766v1

Submitted on 9 Feb 2022

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

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

THÈSE

Pour obtenir le grade de

DOCTEUR DE LA COMMUNAUTÉ UNIVERSITÉ GRENOBLE ALPES

Spécialité : **Physique Appliquée**

Arrêtée ministériel : 25 mai 2016

Présentée par

Guillaume GODEFROY

Thèse dirigée par **Emmanuel BOSSY, UGA**
et codirigée par **Bastien ARNAL, CNRS**

préparée au sein du **Laboratoire Interdisciplinaire de Physique (LIPhy)**
dans **l'École Doctorale de Physique**

Nouvelles approches pour l'imagerie photoacoustique des vaisseaux sanguins : imagerie quantitative 3D de fluctuations et reconstruction assistée par apprentissage profond

Thèse soutenue publiquement le **12 octobre 2021**,
devant le jury composé de :

Monsieur Antoine DELON
Professeur des universités, Université Grenoble Alpes, Président du jury

Monsieur Damien GARCIA
Chargé de recherche, INSERM, Rapporteur

Monsieur Mickael TANTER
Directeur de recherche, INSERM, Rapporteur

Madame Lori BRIDAL
Directrice de recherche, CNRS, Examinatrice

Monsieur Christophe MOSER
Professeur associé, EPFL, Examinateur

Résumé

L'imagerie de la vascularisation présente un intérêt considérable pour de multiples applications de santé, incluant le diagnostic de pathologies, le guidage d'intervention chirurgicale, le suivi de traitement ou encore la recherche biomédicale. Si plusieurs modalités permettent aujourd'hui d'imager la vascularisation, l'utilisation de dispositifs bimodaux combinant l'imagerie photoacoustique (PA) et l'imagerie ultrasonore (US) constitue une approche prometteuse. La sensibilité de l'imagerie PA à l'hémoglobine et l'apparition de méthodes sensibles au flux en imagerie US permettent de cartographier avec précision le réseau vasculaire au sein des tissus biologiques à des profondeurs de l'ordre de plusieurs cm et des résolutions allant jusqu'à la dizaine de µm. L'excellente résolution temporelle de l'imagerie US permet d'obtenir une multitude d'informations quantitatives sur la dynamique du flux sanguin. Au travers d'approches spectroscopiques, l'imagerie PA fournit quant à elle de précieuses données sur les molécules imagées telles que leurs concentrations, et des indicateurs biologiques d'intérêt tels que la saturation en oxygène peuvent être cartographiés. Les travaux menés au cours de cette thèse consistent en la mise en place d'un dispositif permettant de réaliser simultanément des images 3D avec ces deux modalités, et le développement de nouvelles méthodes de reconstruction et de traitement en imagerie PA. Ces méthodes ont pour but de corriger sur les images PA les artefacts dits de visibilité et ceux liés à l'utilisation de réseaux de capteurs parcimonieux. Dans une première approche, l'analyse statistique de la fluctuation du signal PA liée au flux sanguin permet de corriger ces artefacts et d'obtenir une image proportionnelle à l'énergie absorbée. L'approche est appliquée in vivo sur des embryons de poulet et une mesure quantitative de la saturation en oxygène est obtenue. Dans une seconde approche s'appuyant sur l'apprentissage profond (deep learning), un réseau de neurone a été implémenté, permettant de résoudre ces artefacts à partir d'une unique acquisition une fois le réseau entrainé. L'architecture utilisée permet d'inclure une estimation de l'incertitude de la prédiction, fournissant la localisation des erreurs dites d'hallucination. D'abord démontrée en 2D sur des échantillons tests, l'approche est ensuite appliquée en 3D et de premiers résultats expérimentaux sur l'embryon de poulet sont présentés.

Abstract

Vascular imaging is of considerable interest for many health applications, including pathology diagnosis, surgical guidance, medical monitoring and biomedical research. While several modalities are currently used to image blood vessels, bimodal devices combining photoacoustic imaging (PA) and ultrasound imaging (US) is a promising approach. The sensitivity of PA imaging to hemoglobin and the emergence of flow-sensitive methods in US imaging allow accurate mapping of the vascular network within biological tissues at depths of several cm and resolutions down to the tens of µm. The excellent temporal resolution of US imaging provides several quantitative information on blood flow dynamics. Through spectroscopic approaches, PA imaging provides valuable data on the imaged molecules such as their concentrations, and biological indicators of interest such as oxygen saturation can be mapped. The work carried out during this thesis involves the implementation of a device allowing simultaneous acquisition of 3D images with these two modalities, and the development of new reconstruction and processing methods for PA imaging. These methods aim to correct the so-called visibility artifacts and those linked to the use of sparse sensor networks on PA images. In a first approach, the statistical analysis of the fluctuation of the PA signal related to the blood flow allows to correct these artefacts and to obtain an image proportional to the absorbed energy. The approach is applied in vivo on chicken embryos and a quantitative measurement of the oxygen saturation is obtained. In a second approach relying on deep learning, a neural network is used to resolve these artifacts from a single acquisition once trained. The used architecture allows for the estimation of the prediction uncertainty, providing the localization of so-called hallucination errors. Firstly demonstrated in 2D on test samples, the approach is then applied in 3D and the first experimental results on chicken embryos are presented.

Remerciements

Je tiens tout d'abord à remercier Lori Bridal, Antoine Delon et Christophe Moser pour leur participation au jury de thèse, ainsi que Damien Garcia et Mickael Tanter qui ont rapporté ce travail.

J'aimerai ensuite remercier chaleureusement Emmanuel Bossy, mon directeur de thèse, qui m'a donné la chance de rejoindre un projet passionnant, et qui m'a permis de le réaliser dans les meilleurs conditions en me soutenant dans mes choix de pistes à explorer. La porte était toujours grande ouverte pour obtenir des explications fournies avec une très grande pédagogie, et pour discuter des avancements ou des difficultés rencontrées. Ses observations claires et pertinentes m'ont permis d'avancer sereinement et rapidement, sans jamais me sentir seul, dans des projets variés avec une grande autonomie.

Ce fut une énorme chance d'être co-encadré par Bastien Arnal. Il s'est fortement impliqué dans les différents travaux à tous les niveaux, contribuant considérablement à la réussite de la thèse : sans sa disponibilité et son expertise scientifique, la qualité des résultats obtenus aurait été bien moindre. Sa présence et ses qualités humaines m'ont permis de progresser énormément dans de nombreux domaines, à la fois sur des aspects théoriques et expérimentaux. Un grand merci.

Merci à Sandrine Fraboulet, Maxime Henry et Véronique Josserand de l'IAB pour nous avoir accompagné lors de nos expérimentations in vivo et de nous avoir donné l'idée d'utiliser les embryons de poulet comme modèle.

La réussite d'une thèse passe aussi par la qualité des membres de l'équipe dans laquelle elle s'intègre. Ce ne fut ainsi que de belles rencontres au cours de ces trois années, aussi bien du côté des membres permanents de l'équipe OPTIMA : Sylvie Costrel, Irène Wang et Phillipe Moreau, que des post-docorants et doctorants : Sergey, Antonio, Sylvain, Bathilde, Adrien, Camille, Dorian et Ivana. L'ambiance dans le bureau, lors des pauses-café et au cours des repas va très certainement me manquer ! Je vous adresse donc un grand merci pour les moments à l'intérieur comme en dehors du labo et les bons moments partagés autour d'une bière, d'un jeu de société, ou encore à ski, en raquettes et à pied en montagne, en espérant que plein d'autres sont encore à venir.

Merci à tous mes proches qui m'ont encouragé de près ou de loin lors de cette belle expérience (petite dédicace aux huîtres), et en particulier ma famille pour son affection et son soutien.

Enfin, à la personne la plus importante, merci Marie, mon amour, du soutien indéfectible et de tes mots qui m'ont permis de rester confiant tout au long de cette thèse, malgré les doutes et incertitudes. Cette thèse est aussi la tienne.

Table des matières

1 Introduction et Etat de l'art **10**
1.1 L'imagerie de vascularisation . . . 11
1.2 Propagation des ondes lumineuse et ultrasonore dans les tissus biologiques . . . 15
1.2.1 Interaction des ondes avec la matière dans les tissus biologiques . . . 15
1.2.2 Rétrodiffusion en imagerie ultrasonore . . . 16
1.2.3 Effet photoacoustique . . . 17
1.3 Reconstruction de l'image photoacoustique et ultrasonore . . . 19
1.3.1 Les sondes échographiques et leurs caractéristiques . . . 19
1.3.2 Réponse impulsionnelle d'un transducteur . . . 20
1.3.3 Reconstruction par *delay-and-sum* . . . 22
1.3.4 Représentation complexe des signaux radiofréquences . . . 24
1.3.5 Résolution spatiale et fonction d'étalement du point . . . 25
1.3.6 Autres méthodes de reconstruction . . . 26
1.4 Techniques en imagerie photoacoustique et ultrasonore . . . 27
1.4.1 Échographie . . . 27
1.4.2 Imagerie Photoacoustique . . . 29
1.4.3 Principales différences entre l'imagerie photoacoustique et ultrasonore . . . 31
1.4.4 Spécificité de l'imagerie 3D . . . 32
1.5 Les artefacts de visibilité . . . 36
1.5.1 Ouverture numérique limitée et directionnalité des transducteurs . . . 36
1.5.2 Réponse fréquentielle limitée . . . 36
1.5.3 Etat de l'art des méthode de correction . . . 38
1.6 Objectifs de la thèse . . . 39

2 Configuration expérimentale et sonde d'imagerie **41**
2.1 Configuration expérimentale . . . 42
2.1.1 Source de lumière . . . 42
2.1.2 Système d'acquisition des ondes ultrasonores . . . 42
2.1.3 Bruit de mesure . . . 44
2.1.4 Synchronisation en imagerie photoacoustique . . . 45
2.1.5 Schéma du dispositif expérimental utilisant le laser Spitelight . . . 45
2.1.6 Spécificité du dispositif expérimental pour l'imagerie 3D en multi-longueur d'onde 46
2.2 Description des sondes utilisées dans ces travaux de thèse . . . 50
2.2.1 Sonde d'imagerie 2D . . . 50
2.2.2 Sonde d'imagerie 3D . . . 52
2.3 Expériences *in vivo* : méthodes . . . 59
2.3.1 Séquence d'acquisition et interfaçage . . . 59
2.3.2 Préparation des embryons de poulet . . . 61
2.3.3 Organisation d'une expérience souris . . . 62
2.4 Conclusion . . . 63

3 Imagerie photoacoustique de fluctuation **65**
3.1 Contexte . . . 66
3.2 Théorie : formation de l'image photoacoustique de fluctuation . . . 68
3.2.1 Définition du phénomène imagé . . . 68
3.2.2 Définition de l'image PA . . . 70
3.2.3 Définition de l'image moyenne . . . 70
3.2.4 Définition de l'image de fluctuation . . . 71
3.2.5 Conséquences sur les artefacts de visibilité . . . 73
3.2.6 L'image de fluctuation en conditions idéales de simulation . . . 73
3.3 Imagerie photoacoustique de fluctuation en conditions expérimentales *in vitro* . . . 75
3.3.1 Nécessité d'un filtrage spatio-temporel en conditions expérimentales . . . 75
3.3.2 Imagerie phototoacoustique de fluctuation *in vitro* 2D . . . 76
3.3.3 Imagerie phototoacoustique de fluctuation *in vitro* 3D . . . 77
3.4 Imagerie photoacoustique de fluctuation et Power Doppler 3D *in vivo* . . . 79
3.4.1 Imagerie photoacoustique de fluctuation 3D de l'embryon de poulet . . . 79
3.4.2 Imagerie Power Doppler ultrasonore 3D de l'embryon de poulet . . . 80
3.4.3 Imagerie photoacoustique de fluctuation et imagerie Power Doppler ultra-rapide simultanées de l'embryon de poulet . . . 81
3.4.4 Imagerie photoacoustique de fluctuation d'une souris : résultats préliminaires . . . 83
3.5 Impact du bruit électronique sur la qualité de l'image de fluctuation . . . 85
3.5.1 Définition de l'image de fluctuation en présence de bruit . . . 85
3.5.2 Conséquences de la présence de bruit sur l'image reconstruite en simulation . . . 85
3.5.3 Conséquences de la présence de bruit sur l'image reconstruite lors des expériences *in vitro* et *in vivo* . . . 87
3.6 Impact d'une variation tir à tir de la fluence du laser . . . 89
3.6.1 Définition de l'image de fluctuation en présence d'une variation tir à tir de la fluence du laser . . . 89
3.6.2 Amplitude de la fluctuation liée au mouvement des globules rouges vis à vis du signal non fluctuant . . . 90
3.6.3 Mécanismes du filtrage spatio-temporel par décomposition en valeurs singulières 91
3.6.4 Influence du choix des bornes de la décomposition en valeurs singulières sur la qualité de l'image . . . 93
3.7 Conclusion . . . 95

4 Imagerie quantitative de fluctuation **98**
4.1 Introduction . . . 99
4.1.1 Imagerie photoacoustique quantitative conventionnelle . . . 99
4.1.2 Contexte et objectifs de ce chapitre . . . 99
4.2 Théorie : l'imagerie photoacoustique quantitative de fluctuation . . . 101
4.2.1 L'imagerie de fluctuation quantitative en conditions idéales . . . 101
4.2.2 Impact du bruit électronique et méthode de correction . . . 103
4.2.3 Impact de la fluctuation tir à tir de la fluence du laser (PEF) et méthode de correction 104
4.2.4 Estimation de la saturation en oxygène en présence de bruit et de PEF . . . 107
4.3 Imagerie photoacoustique quantitative 3D de fluctuation : imagerie d'oxygénation du sang à visibilité totale . . . 108
4.3.1 Méthode de l'imagerie photoacoustique quantitative de fluctuation en conditions expérimentales . . . 108
4.3.2 Imagerie photoacoustique quantitative de fluctuation *in vitro* . . . 110
4.3.3 Imagerie photoacoustique quantitative de fluctuation de l'embryon de poulet . . . 111
4.4 Imagerie photoacoustique quantitative de fluctuation et Doppler ultrasonore 3D simultanées 113
4.4.1 Méthode de l'imagerie Doppler ultrasonore quantitative . . . 113

4.4.2 Images Directional Doppler et Color Doppler 3D de l'embryon de poulet 113
4.4.3 Analyse conjointe de la SO_2 et du flux 115
4.5 Étude sur simulation de l'impact du bruit électronique et du PEF sur l'estimation de SO_2 117
4.5.1 Influence du bruit électronique 117
4.5.2 Etude : impact du filtrage par SVD sur l'estimation de SO_2, simulation 119
4.6 Estimation automatique de la borne inférieure du filtrage par SVD à partir du rapport contraste à bruit sur l'image de fluctuation 123
4.6.1 Impact de la borne inférieure sur l'estimation de la SO_2 sur les données expérimentales 123
4.6.2 Méthode de sélection semi-automatique de la borne inférieure a 124
4.7 Conclusion 126

5 Imagerie photoacoustique assistée par apprentissage profond 128
5.1 Contexte 130
5.2 *Deep learning* : généralités 133
5.2.1 Réseaux de neurones 133
5.2.2 Réseaux convolutifs 134
5.2.3 Sur et sous-apprentissage 136
5.2.4 Hyper-paramètres et leur utilisation dans les réseaux de neurones 137
5.3 Description du réseau de neurones utilisé et de la construction des jeux de données . . . 137
5.3.1 Constitution du jeu de données expérimentales 137
5.3.2 Constitution du jeu de données de simulation 139
5.3.3 Métriques utilisées pour évaluer les performances du réseau de neurones 140
5.3.4 *Deep learning framework* 140
5.3.5 Méthode de reconstructions alternatives 141
5.4 Correction des artefacts de visibilité par une approche *deep learning* : résultats 142
5.4.1 Entraînement du réseau et évaluation sur données de simulation 142
5.4.2 Entraînement du réseau et évaluation sur données d'expérimentales 142
5.4.3 Pré-entraînement du réseau sur des données de simulation 143
5.4.4 Discussion, perspectives et conclusions 145
5.5 Quantification de l'incertitude en *deep learning* 147
5.5.1 Sources de l'incertitude 147
5.5.2 Méthodes d'estimation existantes 148
5.5.3 Prédiction de l'incertitude aléatoire 149
5.5.4 Application des méthodes d'estimation de l'incertitude à nos données : résultats et discussion 149
5.5.5 Conclusion 151
5.6 Études complémentaires 152
5.6.1 Influence du type de données d'entrée sur la qualité de la prédiction 152
5.6.2 Propriétés de généralisation de notre réseau de neurones 152
5.6.3 Influence de la taille du réseau 154
5.6.4 Influence de l'architecture du réseau de neurones 154
5.6.5 Data augmentation : utilisation de GAN 156
5.7 Conclusion 157

6 Imagerie photoacoustique quantitative 3D assistée par apprentissage profond 159
6.1 Contexte 160
6.2 Etudes sur des données de simulations 162
6.2.1 Construction des jeux de données de simulation et présentation de l'architecture du réseau de neurones utilisée 162
6.2.2 Performances du réseau de neurones 166

6.2.3 Présence de bruit 168
6.3 Expériences 170
6.3.1 Méthode 170
6.3.2 Résultats 171
6.4 Conclusion 173

Conclusion générale **174**

Bibliographie **I**

Production scientifique **XI**

A Apodisation lors de l'opération de *delay-and-sum* **XIII**

B Simulation de signaux radiofréquences **XV**

C Aspects quantitatifs des milieux aléatoires **XVIII**

D Impact de la décomposition en valeurs singulières sur l'estimation de la saturation en oxygène en simulation sur une seconde géométrie : la boucle **XXV**

1

Introduction et Etat de l'art

Sommaire

1.1 L'imagerie de vascularisation 11
1.2 Propagation des ondes lumineuse et ultrasonore dans les tissus biologiques 15
1.2.1 Interaction des ondes avec la matière dans les tissus biologiques 15
1.2.2 Rétrodiffusion en imagerie ultrasonore 16
1.2.3 Effet photoacoustique 17
1.3 Reconstruction de l'image photoacoustique et ultrasonore 19
1.3.1 Les sondes échographiques et leurs caractéristiques 19
1.3.2 Réponse impulsionnelle d'un transducteur 20
1.3.3 Reconstruction par *delay-and-sum* 22
1.3.4 Représentation complexe des signaux radiofréquences 24
1.3.5 Résolution spatiale et fonction d'étalement du point 25
1.3.6 Autres méthodes de reconstruction 26
1.4 Techniques en imagerie photoacoustique et ultrasonore 27
1.4.1 Échographie 27
1.4.2 Imagerie Photoacoustique 29
1.4.3 Principales différences entre l'imagerie photoacoustique et ultrasonore 31
1.4.4 Spécificité de l'imagerie 3D 32
1.5 Les artefacts de visibilité 36
1.5.1 Ouverture numérique limitée et directionnalité des transducteurs 36
1.5.2 Réponse fréquentielle limitée 36
1.5.3 Etat de l'art des méthode de correction 38
1.6 Objectifs de la thèse 39

1.1 L'imagerie de vascularisation

Les vaisseaux sanguins, un tissu aux multiples fonctions

L'imagerie des vaisseaux sanguins est un domaine clé de la santé et de la biologie : détection de pathologies, guidage en chirurgie, développement de nouveaux traitements, compréhension des mécanismes biologiques... Ces vaisseaux ont pour fonction principale d'acheminer le sang propulsé par la pompe cardiaque à l'ensemble des cellules de l'organisme. Ce liquide est typiquement composé en volume de 55% de plasma et 45 % de cellules en suspension. Le plasma contient principalement de l'eau dans lequel baignent de nombreuses protéines dont les nutriments nécessaires au bon fonctionnement de tous les processus biologiques à l'oeuvre dans l'organisme. On distingue trois grandes catégories de cellules sanguines : les globules blancs, les plaquettes et les globules rouges. Ces derniers constituent la quasi-totalité des cellules en suspension dans le sang. Composés majoritairement d'hémoglobines, une protéine riche en fer qui donne par ailleurs sa coloration au sang, les globules rouges sont les principaux vecteurs de transport du dioxygène (O_2) des poumons vers les cellules, et participent à l'évacuation du dioxyde de carbone (CO_2) des cellules vers les poumons. La majorité des composés nécessaires au fonctionnement de l'organisme transite donc à un moment ou un autre par le sang. L'analyse de son contenu peut ainsi permettre de détecter une éventuelle défaillance de l'un de nos organes.

La circulation du sang est assurée par le système cardiovasculaire, dont le dysfonctionnement peut entraîner de graves complications puisque l'ensemble de nos organes se nourrissent des éléments transportés par le sang. Il est capital de pouvoir détecter de tels dysfonctionnements. Cette détection peut par exemple être réalisée par l'analyse de divers paramètres hémodynamiques tels que le débit et le flux sanguin, ou morphologiques tels que le rétrécissement d'une artère. Le réseau des plus fins vaisseaux, les capillaires, évolue constamment en fonction des besoins des tissus environnants : sa densité est ainsi un indice de l'activité métabolique et son évolution peut être associée à diverses pathologies telles que l'apparition de tumeurs. L'évolution du réseau vasculaire peut également servir de marqueur biologique dans le cadre d'études plus fondamentales telles que le suivi du développement de formes de vie embryonnaire ou encore d'organes bioartificiels.

Imager les vaisseaux sanguins, une multitude d'approches

L'imagerie des vaisseaux sanguins vise d'abord à obtenir des informations sur la morphologie de ces vaisseaux pour y détecter d'éventuelles anomalies : une bonne résolution spatiale est alors souhaitable, ainsi qu'un large champ de vue. Lorsque la résolution temporelle est suffisante, il est également possible d'obtenir des informations sur la dynamique du flux sanguin en temps réel. Enfin, certaines approches permettent d'identifier et de quantifier certains composants du sang. L'intérêt d'obtenir ces informations via des méthodes d'imagerie est multiple : l'information est localisée, peut être obtenue en temps réel et *in vivo*, de manière non invasive.

Une multitude de méthodes impliquant la majorité des modalités d'imagerie existantes ont ainsi été développées, chacune répondant à des objectifs et exigences particulières.

Les premières méthodes d'imagerie des vaisseaux sanguins sont basées sur l'émission de rayons X et de la mesure de l'énergie absorbée lors de leur traversée d'un tissu. La différence d'absorption de ces rayonnements entre les vaisseaux sanguins et les tissus environnants étant faible, des produits exogènes sont utilisés pour améliorer significativement le contraste. L'utilisation d'agent de contraste combinée avec les techniques de computed tomography (CT) permet d'atteindre une résolution de l'ordre de la cinquantaine de µm [125]. La durée d'acquisition, de l'ordre de la seconde à plusieurs minutes, dépend du nombre de projections, c'est à dire le nombre d'acquisitions réalisées en changeant l'angle d'incidence des rayons. La principale problématique des approches d'imagerie par rayon X concerne le caractère io-

nisant de ces rayons, et donc leur nocivité pour le corps humain.

En imagerie par résonance magnétique (IRM), le contraste est produit par la dynamique de la relaxation d'un signal électromagnétique, qui varie en fonction de la composition chimique des molécules et de leur environnement. L'IRM est principalement utilisée des deux manières suivantes en imagerie des vaisseaux sanguins. En MR-angiographie, un agent de contraste est injecté dans les vaisseaux. En IRM dite fonctionnelle (en anglais *Blood oxygenation level dependent*, BOLD-MRI), l'IRM produit une mesure de la variation du flux sanguin, obtenue en étudiant la perturbation du signal émis par la molécule d'hémoglobine désoxygénée (hémoglobine sans molécule d'oxygène fixée dessus) au cours du temps lors par exemple d'un processus cognitif [42]. Pour ces deux méthodes, le champ de vue est large, de plusieurs cm. Une résolution de 50 µm peut être atteinte au prix d'une faible résolution temporelle en MR-angiographie, de l'ordre de plusieurs dizaines de minutes [120]. A l'inverse, une résolution temporelle de l'ordre de la seconde peut être obtenue en BOLD-MRI, pour une résolution spatiale dégradée de l'ordre du mm [140].

En échographie, une sonde émet des ultrasons dans le milieu à imager puis capte les échos provoqués par l'interaction de ces ondes avec les tissus. La faible échogénicité des globules rouges par rapport aux tissus environnants limite la possibilité d'obtenir directement une image des vaisseaux. De telles images peuvent cependant être obtenues avec l'imagerie Doppler [110] en exploitant la dynamique temporelle des signaux provenant des globules rouges en mouvement. Récemment, l'imagerie ultra-rapide [88] ainsi que l'apparition de nouvelles méthodes de traitement de signaux [28] ont permis d'augmenter la sensibilité de l'imagerie Doppler. L'imagerie Doppler se caractérise par une résolution de l'ordre de la centaine de µm sur un champ de vue de plusieurs cm [29] et d'une résolution temporelle inférieure à la seconde. Le contraste peut également être rehaussé par l'injection d'agents de contraste (microbulles d'air par exemple). Couplé à une approche de super-résolution (localisation acoustique), une résolution spatiale de l'ordre de la dizaine de µm est atteinte, au prix d'une résolution temporelle de l'ordre de la minute [20].

Les approches purement optiques (microscopie conventionnelle, confocale, OCT, 2-photons) permettent d'atteindre des résolutions spatiales de l'ordre du µm et une résolution temporelle très élevée, largement inférieure à la seconde. Cependant, la résolution spatiale diminue fortement avec la profondeur d'exploration de l'ordre au maximum de quelques centaines de µm en raison de la forte diffusion de la lumière dans les milieux biologiques [95]. On peut néanmoins citer l'OCT angiographique [141], qui permet d'imager jusqu'à environ 1 mm de profondeur en milieu turbide, en exploitant la lumière cohérente.

Dans le but d'obtenir une information de nature optique en profondeur, l'imagerie photoacoustique (PA), consistant en la détection d'ultrasons générés par l'absorption d'énergie lumineuse, se développe depuis deux décennies. On distingue l'imagerie PA à résolution optique (ORPAM) et l'imagerie PA à résolution acoustique (ARPAM). En ORPAM, un faisceau est focalisé en un point de l'échantillon à imager, et le signal acoustique produit par une source située en ce point est enregistré. Ce processus se répète pour chaque point du milieu à imager, limitant fortement la résolution temporelle. La résolution spatiale, de l'ordre du µm est définie par la taille du spot lumineux. La profondeur d'exploration reste limitée (centaine de µm) comme pour les méthodes purement optiques. En ARPAM, la totalité du milieu est illuminé en une unique impulsion laser, provoquant la génération d'ondes acoustiques simultanément par l'ensemble des absorbeurs composant le milieu. Ces ondes sont enregistrées par une sonde ultrasonore. La résolution spatiale est fournie par l'onde acoustique, de l'ordre de la centaine de µm. L'ARPAM génère donc une image des propriétés optiques d'absorption des tissus à une profondeur bien supérieure à l'ORPAM, de l'ordre du cm, et à une fréquence d'imagerie très élevée, largement inférieure à la seconde [144]. Comme pour les méthodes d'imagerie purement optiques, un compromis entre profondeur d'exploration et résolution spatiale est nécessaire. Dans la suite de ce manuscrit, le terme imagerie PA référera uniquement à l'approche ARPAM.

Le tableau 1.1 résume de manière non exhaustive les principales approches permettant d'imager la vascularisation. Les caractéristiques de résolution spatiale, temporelle et de profondeur dépendent fortement des applications visées et des compromis réalisés, et sont décrites ici à titre indicatif.

	Angiographie CT	**Angiographie IRM**	**IRM fonctionnelle**	**Doppler ultra-rapide**	**Localisation acoustique**	**OCT Angiographie**	**Photoacoustique (ARPAM)**
Nature du contraste	Absorption des rayons X	Temps de relaxation	Désoxy-hémoglobine	Impédance +dynamique	Impédance +dynamique	Diffusion +dynamique	Absorption optique
Approche dynamique	Non	Oui	Non	Oui	Oui	Oui	Non
Résolution spatiale	≈ 50 µm	≈ 50 µm	≈ mm	≈ 100 µm	≈ 10 µm	≈ µm	≈ 100 µm
Profondeur	> 10 cm	> 10 cm	> 10 cm	≈ 2-5 cm	≈ 2-5 cm	< mm	≈ 1-5 cm
Résolution temporelle	≈ min	≈ 10 min	≈ s	< s	≈ min	« s	« s
Agent de contraste	oui/rayon ionisant	oui	non	non	oui	non	non
Portabilité	Faible	Faible	Faible	Haute	Haute	Haute	Haute
Coût	Très Haut	Très Haut	Haut	Très Bas	Très Bas	Bas	Bas

Tableau 1.1 – Comparaison des différentes modalités d'imagerie des vaisseaux sanguins.

La figure 1.1 est à caractère illustratif et présente une image pour chacune de ces modalités.

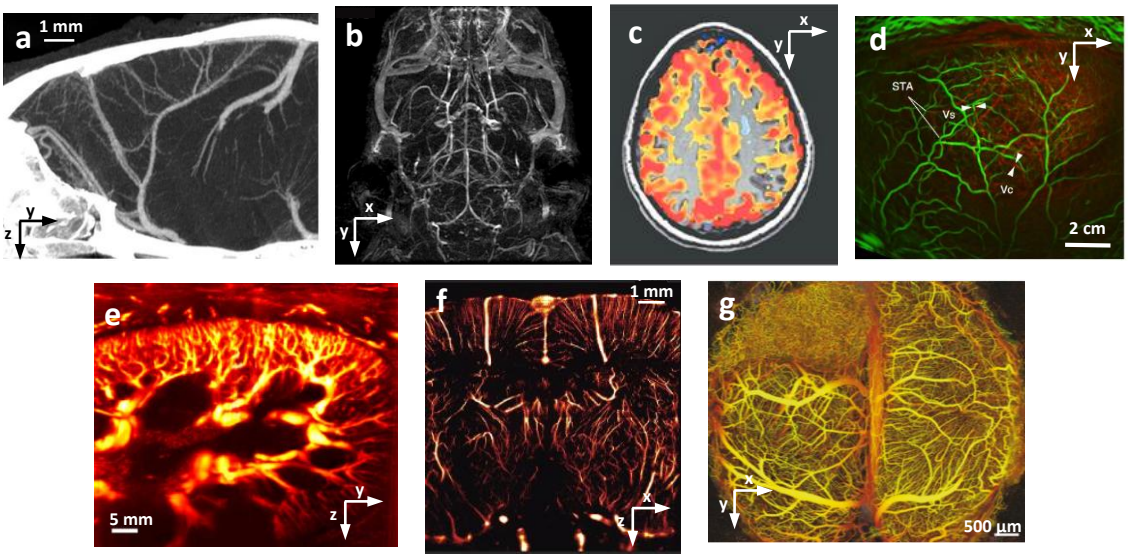

Figure 1.1 – (a) : Image d'angiographie scanner du cerveau d'une souris [125], projection d'une image 3D, vue latérale. (b) : Image du cerveau d'une souris en MR-angiographie [60], projection d'une image 3D, vue du dessus. (c) : Image d'un cerveau humain en IRM fonctionnelle, vue du dessus [54]. (d) : Image photoacoustique d'un cerveau humain après trépanation [90], vue de dessus, projection d'une image 3D. (e) : Image Doppler ultra-rapide d'un rein humain transplanté [28], coupe sagittale. (f) : Image de localisation acoustique d'un cerveau de rat après amincissement du crâne, coupe frontale. [36]. (g) : Image des 2 premiers mm d'un cerveau de souris souffrant d'un glioblastome [130] en optical coherence tomography angiography, projection d'une image 3D, vue du dessus.

Positionnement de l'imagerie PA

Vis à vis des approches basées sur les rayons X ou l'IRM, l'imagerie PA dispose d'une très bonne résolution temporelle et d'une plus grande portabilité (et également d'un coût largement inférieur). Ses performances sont comparables à celles de l'imagerie Doppler ultra-rapide, à une profondeur d'exploration maximale cependant légèrement inférieure. C'est par ailleurs l'approche la plus favorable à l'obtention d'un contraste optique en profondeur à une bonne résolution. On peut noter ici l'existence d'une méthode appelée tomographie optique diffuse, capable de fournir une telle information à ces profondeurs mais avec une résolution spatiale de l'ordre du cm [33]. L'intérêt pour l'imagerie PA s'est également renforcé lors de l'apparition de laser accordable en longueur d'onde et modifiable à haute fréquence. Cela a ouvert la voie à des approches spectroscopiques, permettant notamment d'estimer la saturation en oxygène et de faire émerger l'imagerie PA fonctionnelle [31]. L'imagerie PA permet ainsi d'obtenir simultanément une information similaire à celles fournies par l'angiographie IRM (morphologique) et l'IRM fonctionnelle (saturation en oxygène).

Objectif de ce chapitre

La majeure partie des travaux réalisés au cours de cette thèse se concentre sur le développement d'une nouvelle approche d'imagerie PA (l'imagerie photoacoustique de fluctuation) et de nouvelles méthodes de traitement d'image (principalement basées sur l'apprentissage profond), dans le but de corriger les artefacts dit de visibilité, décrits à la section 1.5 de ce chapitre. En parallèle, nous avons mis en place une plateforme d'imagerie permettant de réaliser simultanément l'imagerie PA et l'imagerie Doppler ultra-rapide et ainsi d'exploiter conjointement les informations fournies par ces deux modalités. Les grands principes de l'imagerie PA ainsi que ceux de l'imagerie US sont présentés dans la suite de ce chapitre. Nous chercherons à mettre en évidence leurs points communs, mais également leurs différences. L'utilisation de ces deux modalités dans le cadre de l'imagerie des vaisseaux sanguins est décrite ainsi que certaines de leurs limitations, notamment celles liées aux artefacts de visibilité. Un schéma résumant les principes de ces deux modalités est disponible en figure 1.2. Une partie des travaux présentés dans ce manuscrit relève de l'imagerie 3D : les spécificités liées à ce type d'imagerie sont donc également abordées dans ce chapitre.

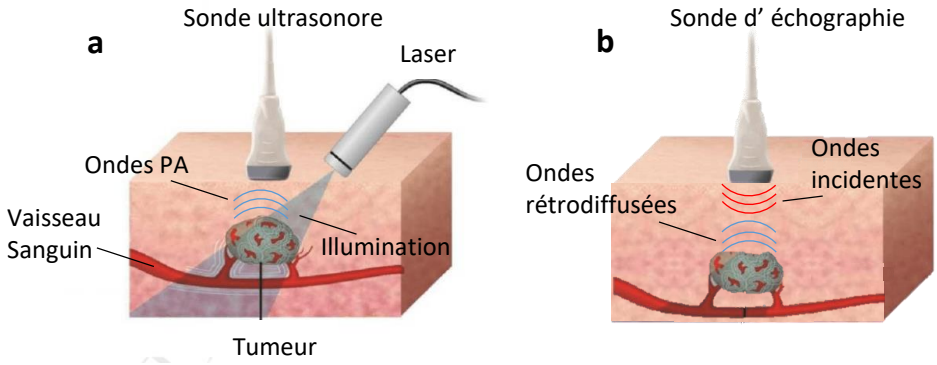

Figure 1.2 – (a) : Représentation de l'imagerie photoacoustique. Une impulsion laser illumine la totalité du champ de vue, provoquant la génération simultanée d'ondes de pression par les absorbeurs composant le milieu. Ces ondes sont enregistrées par une sonde ultrasonore. (b) : Représentation de l'imagerie ultrasonore. Une onde de pression dite incidente est générée par une sonde d'échographie et transmise au milieu à imager. Cette onde est rétrodiffusée par les diffuseurs composant le milieu. Ces échos sont enregistrés par cette même sonde. Modifié à partir de [101].

1.2 Propagation des ondes lumineuse et ultrasonore dans les tissus biologiques

1.2.1 Interaction des ondes avec la matière dans les tissus biologiques

Les interactions entre les ondes et la matière sont à l'origine du contraste en imagerie PA et US, et du compromis entre résolution et profondeur d'imagerie. Ces interactions sont décrites dans cette section pour les ondes lumineuses et acoustiques. L'onde lumineuse est en effet le vecteur de transport de l'énergie vers le milieu en imagerie PA, et l'onde ultrasonore est à la fois le vecteur de transport d'énergie vers le milieu en imagerie US, et le vecteur de transport d'énergie du milieu vers le capteur en imagerie PA et US. Ces interactions sont décrites dans les gammes de longueurs d'onde / fréquences d'intérêt. Le spectre électromagnétique de l'imagerie PA recouvre principalement le visible et le très proche infrarouge, c'est à dire des longueurs d'onde de 350 à 1500 nm (fréquence de l'ordre de la centaine de THz). Les ondes acoustiques d'intérêt en imagerie PA et US ont une fréquence comprise entre 1 et 50 MHz (longueur d'onde de l'ordre de quelques centaines de µm dans les milieux biologiques). Dans ce manuscrit, λ désigne la longueur d'onde de l'onde lumineuse. Le symbole λ_{ac} est utilisé lorsque la longueur d'onde d'une onde ultrasonore est considérée.

Les tissus biologiques sont des milieux complexes et hétérogènes composés d'une grande quantité de particules de différentes tailles. Ils sont néanmoins statistiquement homogènes et peuvent donc être décrits de manière effective. Les interactions entre une onde et un tel milieu peuvent se classer en deux grandes catégories : la diffusion et l'absorption. La diminution de la quantité d'énergie transportée par l'onde au cours de sa propagation dans le milieu résulte de la combinaison de ces deux phénomènes.

L'absorption se définit comme la perte d'une partie de l'énergie transportée par l'onde, absorbée par les atomes et molécules composant le milieu traversé. Cette énergie est principalement convertie sous forme de chaleur ou réémise sous d'autres formes d'ondes ou de rayonnements. Ce phénomène peut être décrit par le coefficient d'absorption effectif μ_a, qui caractérise globalement les propriétés d'absorption d'un milieu en tenant compte de ses fluctuations locales. Il dépend de la densité des molécules/particules dans le milieu et de leurs propriétés d'absorption. L'évolution de l'intensité de l'onde lors de la traversée d'un milieu purement absorbant est décrite par l'expression :

$$I(z) = I_0 e^{-\mu_a z} \tag{1.1}$$

avec z la longueur de propagation et I_0 l'intensité initiale.

La diffusion se définit comme la réémission d'au moins une partie de l'énergie transportée par l'onde dans le milieu environnant sous la forme d'une autre onde de même nature, en raison d'inhomogénéités de l'indice de réfraction du milieu. L'évolution de l'intensité de l'onde lors de la traversée d'un milieu purement diffusant est décrite par l'expression :

$$I(z) = I_0 e^{-\mu_s z} \tag{1.2}$$

avec μ_s le coefficient de diffusion qui décrit la diminution de l'intensité balistique du faisceau. Si ce coefficient μ_s permet d'estimer correctement l'atténuation globale de l'énergie liée à la diffusion simple, on utilise le coefficient de diffusion réduit $\mu_s' = \mu_s(1-g)$, avec g facteur d'anisotropie du milieu lorsque l'on souhaite tenir compte de la diffusion multiple, notamment dans le cas des ondes lumineuses. En effet, à ces longueurs d'onde, le nombre d'évènements de diffusion augmente de manière significative avec la distance de propagation : les rayons lumineux subissent de nombreuses déviations au cours de leur propagation. Une représentation de la diffusion multiple et de la diffusion simple est proposée en figure 1.3. Alors que μ_s décrit l'atténuation provoquée par la diffusion simple, le coefficient de diffusion réduit μ_s' tient compte de la diffusion multiple. Or, c'est bien l'onde lumineuse diffusée qui génère le signal photoacoustique en profondeur.

On peut définir la distance au-delà de laquelle le faisceau sort du régime balistique par $l^* = \frac{1}{\mu_s'}$ le libre parcours moyen de transport, en cm. Il est inférieur au mm dans les tissus biologiques : la lumière a alors perdu toute mémoire de sa direction initiale d'incidence, limitant la localisation des sources par imagerie au-delà de cette distance. C'est la principale limitation des modalités d'imagerie purement optiques. Le tableau 1.2 présente μ_a et μ_s' pour quelques tissus biologiques dans le cas des ondes lumineuses [112].

Tissu	λ (nm)	μ_a (cm^{-1})	μ_s' (cm^{-1})
Peau	630	0.05-1.11	2.26 - 20.95
Cerveau	630	0.02-0.50	3.72 - 21.97
Coeur	630	0.03-1.55	17.56 - 75.06
Poitrine	660	0.037-0.110	11.4 - 13.5

Tableau 1.2 – Propriétés optiques de différents tissus humain

A titre de comparaison, le coefficient d'absorption en imagerie ultrasonore dans le cerveau bovin à 7 MHz est d'environ 0.23 Np/cm^{-1}, et celui de diffusion de 0.41 Np/cm^{-1} [117]. Les écarts entre les coefficients des ondes ultrasonores et lumineuses, très conséquents notamment pour la diffusion, s'expliquent par la différence de longueur d'onde. Alors que la longueur d'onde acoustique est supérieure à la taille des particules, la longueur d'onde lumineuse est du même ordre de grandeur. Les contrastes d'impédance acoustique rencontrés dans les tissus biologiques mous sont faibles, et on peut négliger les phénomènes de diffusion multiple, contrairement au cas de la propagation lumineuse. Ainsi, l'onde acoustique est principalement affectée par des phénomènes de diffusion simple et d'absorption. L'atténuation globale résultante est largement inférieure à celle de l'onde lumineuse, permettant d'atteindre des profondeurs d'exploration de l'ordre de la dizaine de cm.

Jusqu'ici, le cas d'un milieu complexe et hétérogène à l'échelle microscopique mais homogène à l'échelle macroscopique était considéré. Dans le cas de deux milieux complexes dont les caractéristiques globales sont différentes, l'onde peut également être réfléchie au niveau d'une interface (cas de la rupture d'impédance pour l'onde US, et cas de la différence d'indice de réfraction pour l'onde lumineuse).

La figure 1.3 schématise les différents modes de propagation. La flèche représente la direction de propagation de l'onde incidente, et sa largeur l'intensité.

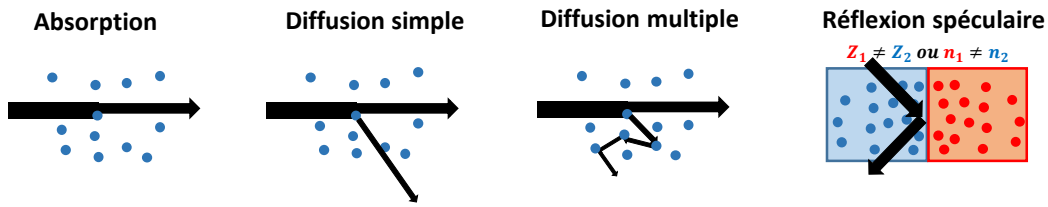

Figure 1.3 – Les différents modes de propagation des ondes dans les milieux biologiques. Z représente l'impédance acoustique du milieu et n l'indice de réfraction optique. La flèche représente la direction de propagation de l'onde incidente, et sa largeur l'intensité.

1.2.2 Rétrodiffusion en imagerie ultrasonore

En imagerie US, une sonde d'échographie émet une onde ultrasonore qui est transmise au milieu à imager. Cette sonde enregistre les ondes ultrasonores rétrodiffusées par les particules composant ce milieu. Cette rétrodiffusion est produite par les fluctuations spatiales de l'impédance entre ces particules.

L'équation de propagation de l'onde de pression peut tout d'abord être définie dans un milieu homogène, en négligeant les phénomènes de dissipation (pas d'absorption) :

$$\Delta p(\mathbf{r},t) - \frac{1}{c_0^2}\frac{\partial^2 p}{\partial t^2}(\mathbf{r},t) = 0 \tag{1.3}$$

avec c_0 la célérité du son et $p(\mathbf{r},t)$ le champ de pression.

Considérons maintenant le cas d'un milieu hétérogène pour lequel la célérité du son c varie faiblement selon $\mathbf{r}$, de valeur moyenne c_0. On nomme l'onde rétrodiffusée $p_r(\mathbf{r},t)$ et l'onde incidente $p_i(\mathbf{r},t)$. On se place dans le cas d'une variation spatiale de l'impédance suffisamment faible pour que le régime de diffusion simple soit valide. Le terme source ne dépend alors pas de $p_r(\mathbf{r},t)$, et on peut écrire :

$$\Delta p_r(\mathbf{r},t) - \frac{1}{c_0^2}\frac{\partial^2 p_r}{\partial t^2}(\mathbf{r},t) \approx \left(\frac{1}{c(\mathbf{r})^2} - \frac{1}{c_0^2}\right)\frac{\partial^2 p_i}{\partial t^2}(\mathbf{r},t) \tag{1.4}$$

On peut approximer que l'onde rétrodiffusée est solution de l'équation d'onde en milieu homogène dont le terme source dépend donc à la fois de l'onde incidente et des propriétés du milieu, dont on considère ici que seule la célérité du son fluctue localement.

1.2.3 Effet photoacoustique

En imagerie photoacoustique, la génération de l'onde acoustique a pour origine l'absorption par des chromophores de l'énergie transportée par un faisceau lumineux, le plus souvent généré par un laser pulsé. Cette énergie est convertie sous forme de chaleur provoquant une hausse insignifiante de température (moins de 0.1 °C en conditions usuelles d'imagerie) laquelle provoque à son tour une hausse de pression. Celle-ci se relaxe sous la forme d'une onde acoustique qui se propage dans le milieu.

Pour définir l'équation de la génération et de la propagation de l'onde photoacoustique, la viscosité du milieu est négligée et les tissus sont alors modélisés par de l'eau pour ce qui concerne les propriétés acoustiques et thermiques. On se place systématiquement dans le cas du confinement thermique c'est à dire lorsque la conduction de chaleur dans l'objet illuminé est négligeable au cours de l'impulsion laser. Les phases d'absorption de lumière et de diffusion de la chaleur sont alors découplées. On peut définir le temps pendant lequel la chaleur se diffuse pour un absorbeur de taille caractéristique D_a par la durée $\tau_{th} = \frac{D_a^2}{\chi}$ avec $\chi \approx 1.410^{-7}$ m^2.s^{-1} (dans l'eau à 20°) le coefficient de diffusion thermique. Le confinement thermique est atteint lorsque $\tau_p << \tau_{th}$ avec τ_p la durée de l'impulsion laser. Pour $\tau_p = 5$ ns, durée typique en imagerie PA, l'hypothèse est vérifiée pour un absorbeur de diamètre grand devant quelques dizaines de nanomètre, ce qui est toujours le cas en pratique. L'équation de la génération et de la propagation de pression des ondes acoustiques $p(\mathbf{r},t)$ est alors :

$$\Delta p(\mathbf{r},t) - \frac{1}{c_0^2}\frac{\partial^2 p}{\partial t^2}(\mathbf{r},t) = -\Gamma\mu_a(\mathbf{r})\frac{\partial \Phi_r(\mathbf{r},t)}{\partial t} \tag{1.5}$$

Le champ de pression est ainsi directement relié à la puissance absorbée par le milieu, définie par $\Gamma\mu_a(\mathbf{r})\frac{\partial \Phi_r(\mathbf{r},t)}{\partial t}$. Ce terme dépend de la quantité d'énergie déposée par le laser, le taux de fluence $\Phi_r(\mathbf{r},t)$ (W.cm^{-2}), du coefficient d'absorption du milieu $\mu_a(\mathbf{r})$ (cm^{-1}), et de Γ le coefficient de Grüneisen, qui quantifie l'efficacité de la conversion de l'énergie absorbée sous la forme de l'onde de pression.

Lorsque le temps de propagation des contraintes induites par l'absorption de l'énergie est long devant la durée des impulsions, on considère le cas du régime de confinement de contrainte [136]. Ce régime est atteint lorsque la durée de l'impulsion laser τ_p est courte devant le temps durant lequel l'onde de pression se propage au sein d'un absorbeur $\tau_{ac} = \frac{D_a}{c}$, ce qui est vérifié pour un absorbeur de diamètre

supérieur à 7.5 µm pour $\tau_p = 5$ ns. Lorsque les deux régimes de confinement sont atteints, l'absorption de lumière est considérée comme instantanée, la durée de l'impulsion lumineuse n'apparaît plus dans le problème : l'absorption de lumière et la propagation de l'onde acoustique sont découplées. L'équation de propagation et de génération de l'onde acoustique peut finalement s'écrire en remplaçant le terme source par des conditions initiales :

$$\begin{aligned} &\Delta p(\mathbf{r},t) - \frac{1}{c_0^2}\frac{\partial^2 p}{\partial t^2}(\mathbf{r},t) = 0 \\ &p(\mathbf{r},t=0) = p_0(\mathbf{r}) = \Gamma\mu_a(\mathbf{r})\Phi(\mathbf{r}) \\ &\frac{\partial p}{\partial t}(\mathbf{r},t=0) = 0 \end{aligned} \tag{1.6}$$

avec $\Phi(r) = \int \Phi_r(r,t)\mathrm{d}t$ la fluence. Dans ce cas, la durée du signal photoacoustique et son contenu fréquentiel dépend alors uniquement de la taille de l'absorbeur. Le signal photoacoustique généré par une boule absorbante de diamètre D_a largement supérieur 7.5 µm est représenté en figure 1.4 (a,b). Dans le cas contraire des sources de dimension trop petite, pour lequel le confinement de contrainte n'est pas atteint, on peut alors considérer le modèle du point source : la durée du signal photoacoustique ainsi que son contenu fréquentiel sont définis par la durée de l'impulsion laser (figure 1.4 c,d).

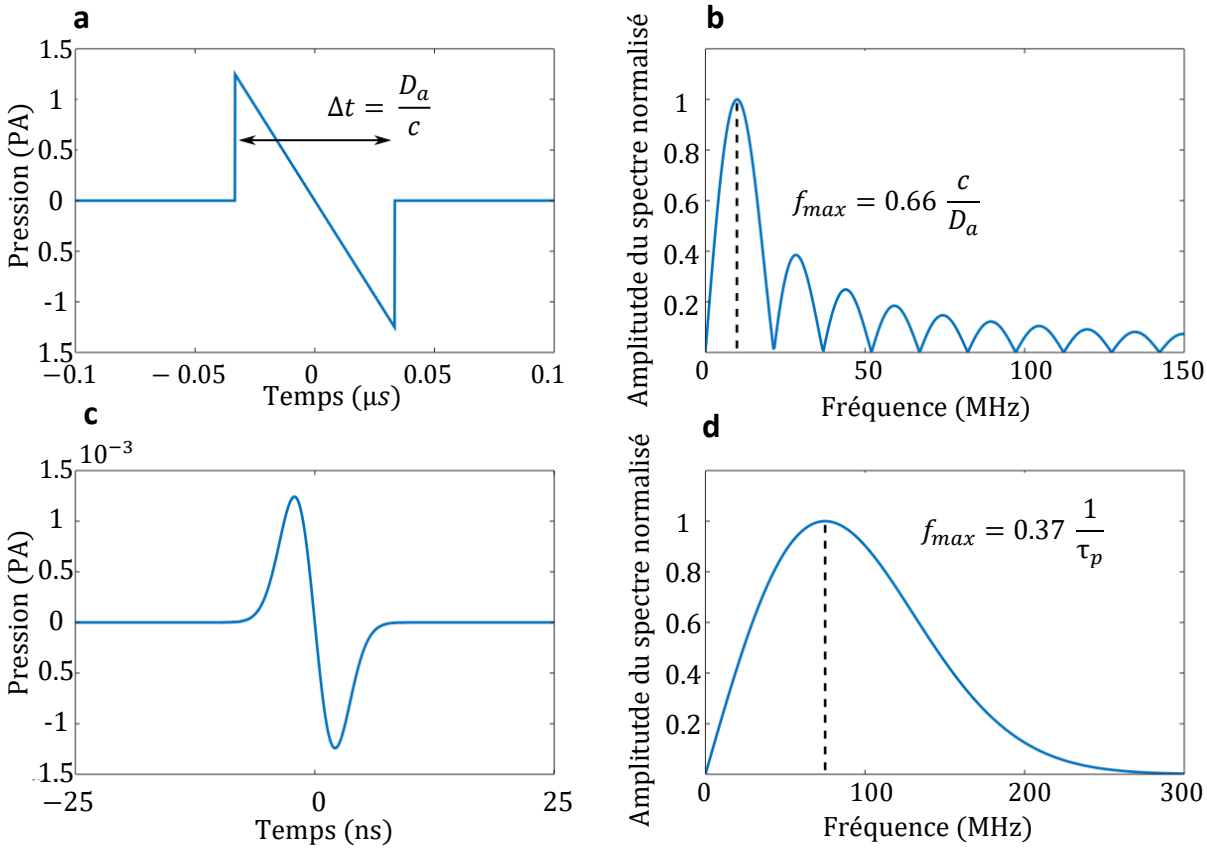

Figure 1.4 – (a) et (b) : Signal photoacoustique généré par une sphère de diamètre $D_a = 100$ µm et son spectre pour une durée d'impulsion laser $\tau_p = 5$ ns et $c = 1500$ m.s^{-1} : le confinement de contrainte est vérifié, la durée du signal et sa fréquence maximale sont définies par D_a. (c) et (d) : signal photoacoustique généré par une sphère de diamètre $D_a = 5$ µm et spectre avec $\tau_p = 5$ ns : seul le confinement thermique est vérifié, la durée du signal et sa fréquence maximale sont définies par τ_p.

Dans le cadre de cette thèse, nous considérons pour la formation des images le cas où le régime de confinement de contrainte est valide. Bien que le diamètre d'un globule rouge soit d'environ 7.5 µm, nous nous placerons systématiquement à l'échelle du vaisseau sanguin pour lequel les globules rouges ne sont pas considérés individuellement mais par 'paquet' au sein d'une *point spread function* acoustique (section 1.3.5), échelle à laquelle le confinement de contrainte est valide. Pour les simulations impliquant la génération photoacoustique à l'échelle d'un globule unique, nous utiliserons par contre le régime du point source.

1.3 Reconstruction de l'image photoacoustique et ultrasonore

En imagerie PA comme en imagerie US, l'image est formée à partir de l'enregistrement d'ondes acoustiques, générées par l'effet PA ou par la rétrodiffusion d'une onde ultrasonore transmise au milieu. Nous décrivons dans cette section le fonctionnement des capteurs conventionnellement utilisés pour enregistrer ces signaux (également utilisés pour l'émission de l'onde de pression en imagerie US) et quelques-unes de leurs propriétés. Nous présenterons également la méthode de reconstruction par *delay-and-sum* aussi appelée *beamforming* puis quelques approches alternatives de reconstruction.

1.3.1 Les sondes échographiques et leurs caractéristiques

L'onde d'intérêt en imagerie PA ou US est une onde de pression générée par un absorbeur ou un diffuseur. Elle se propage dans le milieu jusqu'au capteur qui convertit cette onde en une tension électrique proportionnelle à son amplitude, laquelle pourra ensuite être quantifiée et échantillonnée par un système d'acquisition. Si plusieurs approches permettant de réaliser la conversion amplitude de pression/tension électrique existe, l'effet piézoélectrique est le plus souvent utilisé, le capteur est alors nommé "transducteur". Sous l'application d'une impulsion électrique entre ses deux faces, le matériau qui compose ce type de transducteur se compresse ou se dilate, générant une onde ultrasonore qui peut alors se propager dans le tissu. L'onde incidente en imagerie US est générée ainsi. A l'inverse, lorsque une onde de pression atteint ce matériau, l'effet inverse se produit : il se déforme, générant une tension électrique en fonction du temps proportionnelle à l'amplitude de l'onde. Cette tension électrique est appelée, après quantification et échantillonnage, signal radiofréquence (RF). Ce signal constitue la mesure de base fournie par un échographe.

Les sondes ultrasonores sont généralement constituées d'un réseau de ces transducteurs (aussi appelés 'éléments'). En imagerie 2D, ces sondes se présentent souvent sous la forme d'une barrette linéaire contenant typiquement 128 ou 256 éléments. On distinguera dans ce manuscrit le transducteur, qui correspond à un élément, de la sonde d'imagerie, constituée de plusieurs éléments. On définit h_{el} la hauteur de l'élément et w_{el} sa largeur. Une représentation est proposée en figure 1.5.

Les signaux enregistrés en imagerie PA ou US par une sonde multiéléments peuvent être représentés sous la forme d'une matrice 2D appelée ici $S(t, j)$, pour laquelle l'axe des ordonnées représente le temps et l'axe des abscisses correspond à l'indice de l'élément. La couleur encode la variation en amplitude de l'onde de pression. On parle souvent de B-scan pour décrire cette représentation. Dans la figure 1.5 sont proposées une représentation de l'enregistrement d'une onde de pression générée par une unique particule, et la matrice $S(t, j)$ résultante. Un signal $s_j(t)$ provenant de cette matrice est également observable (on parle souvent de A-scan pour décrire un tel signal unique).

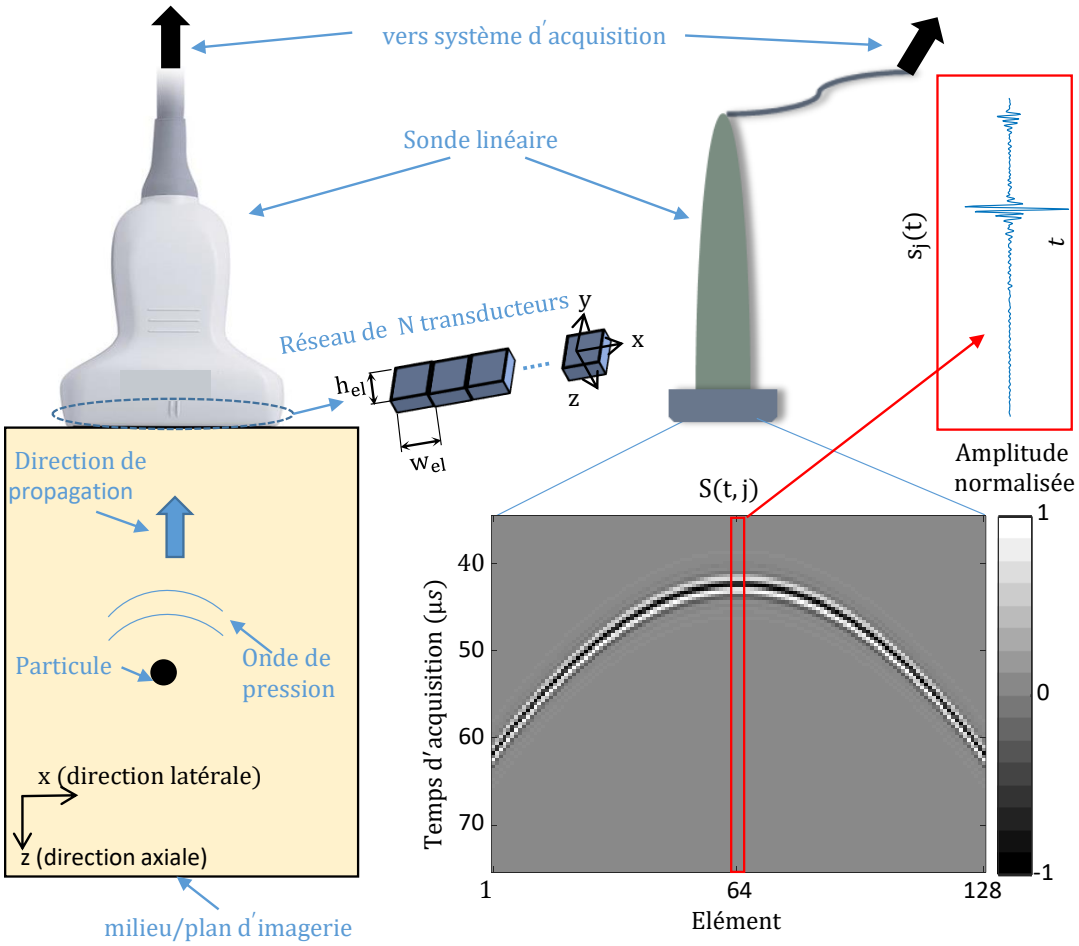

Figure 1.5 – Enregistrement par une sonde d'imagerie constituée de $N = 128$ éléments d'une onde de pression générée par une unique particule, et matrice 2D contenant les enregistrements. La grandeur h_{el} définie la hauteur du transducteur et w_{el} sa largeur. La matrice 2D S(t,j) est composée de l'ensemble des $N = 128$ signaux enregistrés par chacun des transducteurs composant la sonde d'imagerie, en fonction du temps. Les valeurs composant cette matrice sont proportionnelles à l'amplitude de l'onde de pression. L'un de ces signaux est également représenté en fonction du temps dans l'encart rouge.

Par convention, le t_0 du temps d'acquisition correspond au temps de l'émission de l'onde acoustique dans le milieu en US et à l'émission du faisceau lumineux en imagerie PA (donc à la génération de l'onde PA dans le milieu, puisque l'illumination est considérée comme instantanée). Nous verrons dans le chapitre 2 que le t_0 fourni par les systèmes d'acquisition, notamment en imagerie PA, peut être erroné lorsque la synchronisation entre l'illumination et le début de l'acquisition de l'onde est imprécise.

1.3.2 Réponse impulsionnelle d'un transducteur

En réalité, le signal $s_j(t)$ présenté en figure 1.5 mesuré par l'élément j correspond à l'onde de pression, ici générée par un effet PA $s_0(t)$ (ce qui suit est également valable pour l'onde US rétrodiffusée) filtrée par une réponse impulsionnelle. On néglige par la suite l'atténuation de l'onde de pression (géométrique, absorption, diffusion) lors de sa propagation dans le milieu ainsi que le filtrage par ce milieu. La réponse impulsionnelle dépend alors uniquement de la fonction de transfert du transducteur j, appelée $h_j(t, \mathbf{r})$ avec $\mathbf{r}$ la position de la source. On considère dans un premier temps le modèle d'un transducteur ponctuel de position $\mathbf{r}_j$:La fonction de transfert du transducteur traduit alors la conversion de l'onde

de pression en tension et un déphasage qui dépend de la distance entre r et r_j. Le signal mesuré par ce transducteur est alors $s_j(t) = s_0(t) * h_j(t, \mathbf{r})$. Le transducteur est considéré comme omnidirectionnel : il est alors capable de recevoir ou d'émettre de manière similaire dans toutes les directions.

Le modèle du transducteur ponctuel est valide uniquement lorsque la taille caractéristique du transducteur est suffisamment petite devant la longueur d'onde acoustique de l'onde incidente λ_{onde}. Sinon, les effets de diffraction doivent être considérés : la fonction de transfert du transducteur dépend alors de la position absolue de la source. On considère que la fonction de transfert du transducteur est identique en tout point $\mathbf{r}_j$ de sa surface, au déphasage près de l'onde incidente induit par la variation de la distance $(\mathbf{r}, \mathbf{r}_j)$ entre chaque point de la surface $\mathbf{r}_j$. Le signal enregistré par la sonde correspond alors à l'intégrale sur sa surface de l'onde incidente convoluée avec cette fonction de transfert $\tilde{h}_j(t, \mathbf{r})$ variant en fonction de $\mathbf{r}_j$, selon l'expression $s_j(t) = \int_{\mathbf{r}_j} s_0(t) * \tilde{h}_j(t, \mathbf{r}) dr_j$. Dans ce cas, le transducteur n'est plus omnidirectionnel, et le signal mesuré dépend notamment de l'angle θ entre la surface du transducteur et la direction de propagation de l'onde incidente. Le phénomène est schématisé en figure 1.6 dans le cas d'une onde plane arrivant sur un transducteur, générée par exemple par une interface en imagerie US ou une structure allongée dans une direction en imagerie PA. On peut se représenter le transducteur comme une somme de sous-éléments omnidirectionnels (b,c). Lorsque la direction de propagation de l'onde est perpendiculaire à la surface du transducteur, $\theta = \pi/2$, chaque sous élément reçoit l'onde au même t, en phase (b). Lorsque l'angle d'incidence n'est plus perpendiculaire à cette surface ($\theta \neq \pi/2$), les sous-éléments reçoivent l'onde avec un décalage de phase Δt qui augmente avec l'angle θ (c). L'onde ultrasonore étant bipolaire, les contributions positives et négatives se compensent, diminuant de fait l'amplitude effective de la variation de pression appliquée au transducteur, jusque s'annuler complètement lorsque l'angle d'incidence est trop grand : aucun signal n'est alors capté par le transducteur.

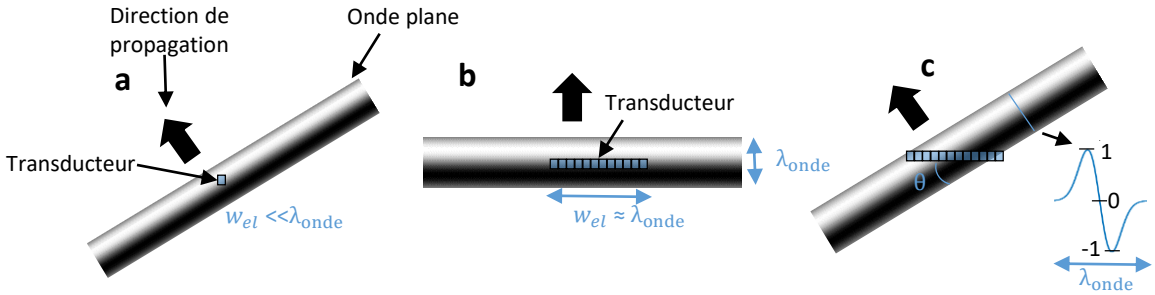

Figure 1.6 – Illustration de la directionnalité des transducteurs. Le cas d'une onde plane bipolaire arrivant sur un transducteur est présenté. (a) : la largeur w_{el} de l'élément est faible devant la longueur d'onde acoustique de l'onde λ_{onde}, le modèle du point est valide. (b) et (c) : la largeur du transducteur est de l'ordre de grandeur de λ_{onde} : l'amplitude du signal enregistré dépend de la direction de l'onde incidente. Le transducteur est alors considéré comme une collection de sous éléments omnidirectionnels. En (b), l'angle entre la direction de l'onde incidente et la surface du transducteur est de $\pi/2$: lorsque l'onde atteint le transducteur, les sous-éléments sont en phase. En (c), l'onde arrive de manière déphasée sur la surface du transducteur. Des contributions positives (partie "blanche" de l'onde sur le schéma) et négatives (partie "noire" de l'onde sur le schéma) sont intégrées "en même temps" à la surface du transducteur. Ces contributions interfèrent négativement, diminuant l'amplitude du signal généré.

Cette directionalité des transducteurs est l'une des causes de l'artefact de visibilité dit de *limited view* : les ondes arrivant avec un angle trop important ne sont pas détectées.

La fonction $h_j(t)$ d'un transducteur peut être estimée en enregistrant l'onde générée par un point source dans un milieu homogène et non absorbant, par exemple, une bille de quelques µm de diamètre située au point focal d'un transducteur dans de l'eau déionisée et dégazée. C'est une manière conventionnelle de définir certaines propriétés de ce transducteur. On a alors $s_j(t) = h_j(t) * \delta(t) = h_j(t)$. La transformée de Fourier de $s_j(t)$ permet notamment de mesurer la gamme de fréquence dans laquelle le transducteur peut émettre et recevoir. Elle est définie par une fréquence centrale f_c et une bande passante Δf. On considère souvent la longueur d'onde acoustique de la sonde d'imagerie définie selon $\lambda_{ac} = \frac{c_0}{f_0}$ pour caractériser une sonde ultrasonore. La réponse impulsionnelle d'une sonde de fréquence centrale 8 MHz et de bande passante 5.2 MHz et sa transformée de Fourier sont présentées en figure 1.7.

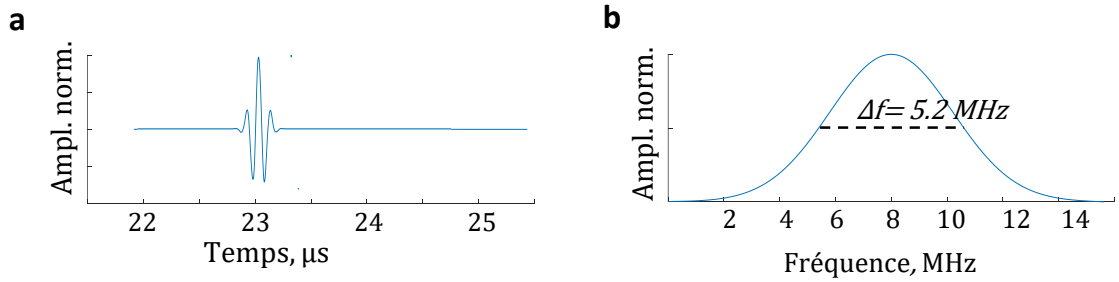

Figure 1.7 – Réponse impulsionnelle d'une sonde de fréquence centrale 8 MHz et de bande passante 5.2 MHz en (a) et sa transformée de Fourier en (b).

La résolution axiale du système dépend en partie de la fréquence centrale et de la largeur de la bande passante (section 1.3.5). Cette réponse fréquentielle limitée est également à l'origine de l'artefact de *limited bandwidth*, qui sera décrit dans la section 1.5. La bande passante est résultante de la conception du capteur et des matériaux utilisés. Des technologies autres que les transducteurs basés sur l'effet piézoélectrique se développent, notamment les CMUTs (*capacitive micromachined ultrasonic transducer*) pour lesquels la transduction de l'énergie est due à la mise en mouvement par force électrostatique d'une membrane de silicone située entre deux électrodes. Les CMUTs ont pour avantage de disposer d'une bande passante plus large.

1.3.3 Reconstruction par *delay-and-sum*

Les méthodes de reconstruction basées sur le principe de *delay-and-sum*, aussi appelé *beamforming*, sont les plus communément utilisées en imagerie ultrasonore et photoacoustique. Dans sa version la plus simple, l'approche se limite à compenser la propagation en supposant une densité et une vitesse du son homogène dans le milieu. On néglige donc tous les phénomènes d'atténuation de l'onde lors de sa propagation, et on considère le modèle du transducteur ponctuel, donc omnidirectionnel. On peut alors émettre l'hypothèse que chaque portion du signal reçu à un temps donné a pu être émis d'une source localisée sur une sphère dont le rayon correspond au temps de vol aller pour la PA ($r = c \times t$) et aller-retour pour l'US (expression de r en fonction de t dépendant du type d'émission). Une illustration pour l'imagerie PA en 2D est présentée en figure 1.8. Un milieu composé de deux absorbeurs est imagé. Les signaux de deux des transducteurs constituant la sonde d'imagerie (en vert el_5 et en orange el_{13}) sont observables. La particule p_1 est située sur le cercle de rayon $c \times t_1$ ayant pour centre el_5, un signal est donc enregistré par el_5 au temps t_1. Aucune particule n'est située sur le cercle de même rayon de centre el_{13}, aucun signal n'est donc enregistré à ce même temps t_1. On observe la même situation au temps t_2 pour la seconde particule. A t_3, les deux particules p_1 et p_2 sont situées sur le cercle de rayon $c \times t_3$ de centre el_{13} : un signal de plus grand amplitude, correspondant à la somme des contributions des deux particules, est visible.

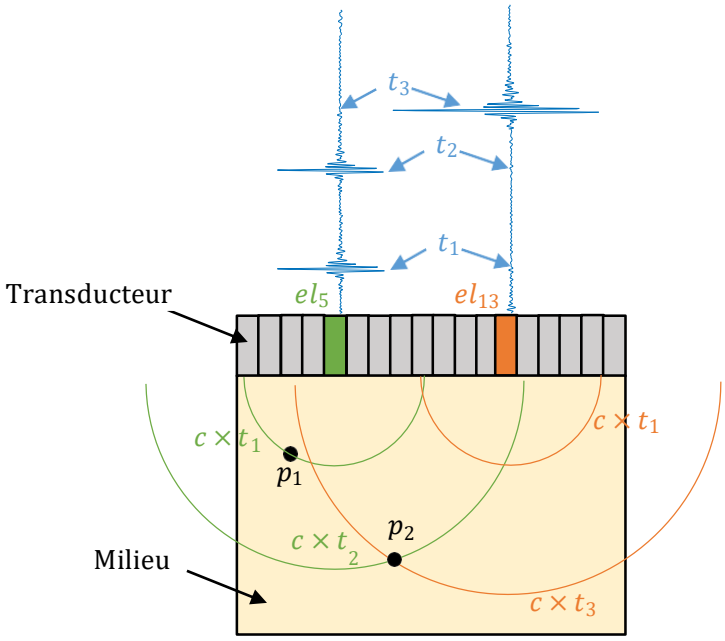

Figure 1.8 – Représentation sphérique du beamforming en imagerie photoacoustique dans le cas de l'imagerie 2D. Un milieu composé de deux absorbeurs est imagé, et les signaux de deux des transducteurs constituant la sonde d'imagerie (en vert, el_5 et en orange, el_{13}) sont présentés. La particule p_1 est située sur le cercle de rayon $c \times t_1$ ayant pour centre el_5, un signal est donc enregistré par el_5 au temps t_1. Aucune particule n'est située sur le cercle de même rayon de centre el_{13}, aucun signal n'est donc enregistré à ce même temps t_1. Même situation pour t_2 avec la seconde particule. A t_3, les deux particules sont situées sur le cercle de rayon $c \times t_3$ de centre el_{13} : un signal de plus grande amplitude, correspondant à la somme des contributions de chacune de des particules, est visible.

C'est l'utilisation d'un réseau de capteurs qui va permettre de localiser les sources par la superposition des différentes sphères projetées depuis chaque transducteur. On considère donc qu'une image est obtenue à partir de l'ensemble des signaux bruts contenus dans $S(t, j)$. Pour chaque pixel constituant l'image, les temps de vol entre le pixel et tous les transducteurs sont calculés. Les signaux de chaque transducteur sont décalés de ces temps de vol afin de les sommer entre eux de manière cohérente. On peut ainsi définir l'équation 1.7 qui définit la valeur d'un pixel de l'image $A(x,z)$ dans le cas 2D :

$$A(x,z) = \sum_{j=1}^{N_{el}} S(t_{propa}(j,x,z), j) \tag{1.7}$$

avec S la matrice contenant les signaux radiofréquences et t_{propa} la matrice de taille $N_{el} \times N_x \times N_z$ avec N_{el} le nombre de transducteurs et N_x et N_z le nombre de pixels dans chacune de ces dimensions. En imagerie photoacoustique, cette matrice stocke les temps de propagation correspondant à la distance entre chaque pixel constituant l'image et chaque transducteur. Dans le cas d'un milieu homogène et isotrope, on peut définir t_{propa} selon :

$$t_{propa}(j,x,z) = \frac{1}{c_0}\sqrt{(x_j - x)^2 + (z_j - z)^2} \tag{1.8}$$

avec x_j et z_j la position du transducteur j. En imagerie US, il faut tenir compte du temps de propagation de l'onde incidente. Ce terme dépend du type d'onde générée (sphérique, plane, angulée ...). On défini $t_{propa,i}$ la matrice t_{propa} lorsque une onde sphérique est émise par l'élément i :

$$t_{propa,US}(j,x,z) = \frac{1}{c_0}\sqrt{(x_i - x)^2 + (z_i - z)^2} + \frac{1}{c_0}\sqrt{(x_j - x)^2 + (z_j - z)^2} \tag{1.9}$$

avec x_i et z_i la position du transducteur i. Dans la figure 1.9, les temps de vol entre deux pixels et chaque transducteur sont représentés en traits verts et rouges pour deux pixels dans le cas de l'imagerie PA.

L'opération de décalage des signaux et de leur sommation cohérente sont représentées pour le pixel vert dont la position coïncide avec celle d'un absorbeur.

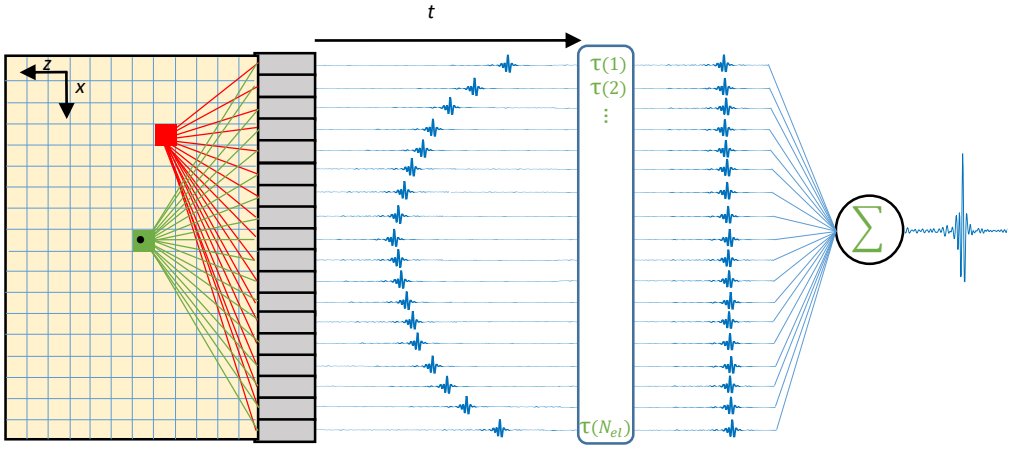

Figure 1.9 – Principe du delay-and-sum en imagerie photoacoustique. La grille de reconstruction qui quadrille le milieu imagé est visible en bleu. Un absorbeur (point noir) est présent dans le milieu. Les signaux enregistrés par chaque transducteur correspondant à l'acquisition en imagerie photoacoustique de ce milieu sont représentés. Les temps de vol entre le pixel rouge ou le pixel vert et chaque transducteur correspondent aux distances représentées par les traits rouges et verts multipliées par c_0. L'opération de delay-and-sum est représentée pour le pixel vert, dont la position coïncide avec celle de l'absorbeur : les signaux sont décalés de la loi de retard $\tau(j)$ correspondant au temps de vol, puis sommés entre eux.

Une matrice de coefficients d'apodisation $W_j(r)$ peut être insérée dans l'équation du beamforming, permettant de prendre en considération la directionnalité des éléments par exemple en réduisant la contribution des signaux provenant de points éloignés de l'axe du transducteur, ce qui permet aussi de diminuer l'impact des lobes secondaires. L'équation du beamforming devient alors :

$$A(x,z) = \sum_{j=1}^{N_{el}} W_j(r) S(t_{propa}(j,x,z), j) \tag{1.10}$$

1.3.4 Représentation complexe des signaux radiofréquences

Les signaux enregistrés par une sonde ultrasonore sont des successions de valeurs réelles : une variation de pression est traduite en une tension électrique par les transducteurs. Des représentations de ces signaux sont disponibles dans les figures 1.7 et 1.8. En pratique, ces signaux sont souvent exploités sous leurs formes en quadrature, obtenues par leurs transformées de Hilbert. On considère alors les signaux $\bar{s}_j(t) = s_j(t) + i.Hilbert(s_j(t)) = s_j(t) + i.q_j(t))$. Un exemple de signal réel et de sa partie imaginaire est proposé en figure 1.10.

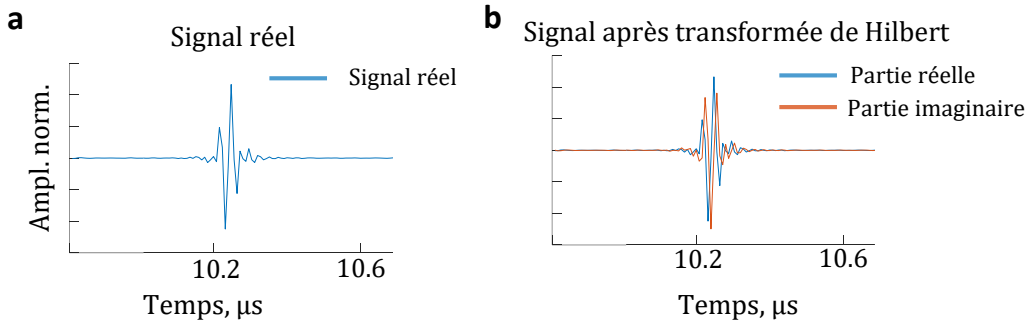

Figure 1.10 – (a) : Signal réel généré par un point source enregistré par un élément de la sonde. (b) : signal sous sa forme en quadrature.

L'image peut être reconstruite par *delay-and-sum* à partir du jeux de signaux $S(t, j)$ ou à partir de sa forme complexe $\bar{S}(t, j)$. Dans le second cas, l'image est également de forme complexe. On parle de démodulation lorsque le module de cette image est calculé, permettant notamment de supprimer les oscillations induites par la réponse fréquentielle de la sonde, et ainsi produire une reconstruction plus conforme à l'objet. C'est l'image après cette opération qui est généralement présentée comme l'image conventionnelle en imagerie PA ou US. Une représentation est visible en figure 1.11 dans la section suivante.

1.3.5 Résolution spatiale et fonction d'étalement du point

En imagerie PA, pour une sonde ultrasonore de diamètre D défini selon $D = w_{el} \times N_{el}$, de fréquence centrale f_c et de bande passante Δf, la résolution latérale (direction perpendiculaire à l'axe de la sonde) à la profondeur R est définie selon :

$$\Delta X \approx \frac{c_0}{f_c} \frac{R}{D} \tag{1.11}$$

Elle dépend donc de la réponse fréquentielle de la sonde et de la géométrie du capteur. La résolution axiale (dans l'axe de la sonde) est définie selon :

$$\Delta Z \approx c_0 \tau_p \approx \frac{c_0}{\Delta f} \tag{1.12}$$

Elle dépend uniquement de la réponse fréquentielle de la sonde. La résolution n'est donc pas invariante spatialement, puisqu'elle dépend de la profondeur R et de la réponse fréquentielle qui varie en fonction de l'angle d'incidence de l'onde générée par la source (donc de la position latérale de la source).

Les valeurs de ΔZ et ΔX peuvent également être évaluées à partir de la fonction d'étalement du point, ou *point spread function* (PSF), qui décrit la réponse d'un système à une source ponctuelle dans un milieu donné. Nous distinguerons le terme réponse impulsionnelle que l'on réservera à la fonction de transfert $h_j(t)$ spécifique à chaque transducteur et défini dans le domaine temporel, du terme PSF qui sera définie comme l'image reconstruite par beamforming du point source. La largeur à mi-hauteur dans chaque direction est une définition possible de la résolution spatiale. La PSF PA au point focal d'une sonde d'imagerie linéaire est présentée en figure 1.11.

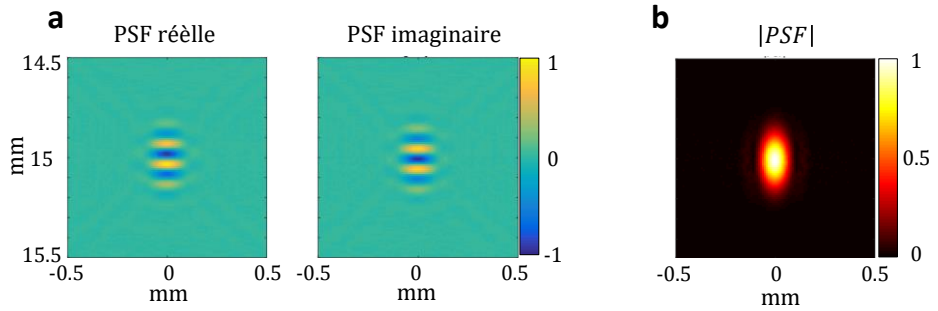

Figure 1.11 – (a) : partie réelle et partie imaginaire de la point spread function (PSF) au point focal du plan d'imagerie. (b) : module de la PSF, les oscillations liées à la réponse de la sonde ne sont plus visibles. Cette image est appelée tache de diffraction, et peut permettre de définir la résolution du système en mesurant la largeur à mi-hauteur dans les directions axiales et latérales.

La résolution dans la direction ΔY (direction transverse au plan d'imagerie en 2D) dépend du mode d'imagerie. Pour une sonde linéaire, on a :

$$\Delta Y = \frac{c_0 R}{f_c h_{el}} \tag{1.13}$$

avec h_{el} la hauteur du transducteur. En imagerie 3D dans le cas d'un réseau de capteur 2D, nous verrons que ΔY est similaire à ΔX par symétrie de révolution. En imagerie US, la résolution axiale est augmentée d'un facteur 2 en raison de l'effet aller-retour et d'un facteur allant de 1 à 2 pour la résolution latérale en fonction de la modalité d'émission, grâce à l'éventuelle focalisation en émission.

1.3.6 Autres méthodes de reconstruction

On distingue deux grands types de méthodes de reconstruction en imagerie US et PA : les méthodes directes de type *delay-and-sum* et les méthodes d'optimisation.

Les méthodes de type *delay-and-sum* peuvent être raffinées en filtrant les signaux dans le domaine temporel ou fréquentiel avant de réaliser l'opération de *delay-and-sum*. Par exemple, des filtres passe-haut ou passe-bas peuvent être appliqués pour rehausser les structures les plus fines ou corriger un problème d'aliasing. Ce filtrage peut également constituer en une déconvolution par une estimation de la fonction de transfert de la source jusqu'à la sortie du transducteur. Cette fonction contient les informations sur le système d'imagerie et sur la propagation de l'onde dans le milieu. Elle peut être estimée par simulation en utilisant la théorie de la propagation. Les propriétés du milieu et du système d'imagerie doivent alors être connues. Cette fonction pourrait également être mesurée expérimentalement, ce qui nécessiterait cependant de mesurer indépendamment les signaux générés par un point source en chaque pixel constituant le milieu, peu envisageable dans les tissus biologiques. L'opération de filtrage peut également être réalisée sur l'image reconstruite avec la PSF après l'opération de *delay-and-sum* : on parle alors de déconvolution 2D.

La validité des méthodes de reconstructions basées sur le *delay-and-sum* repose sur l'hypothèse d'un grand nombre de récepteurs omnidirectionnels répartis tout autour de l'objet en respectant spatialement le critère de Nyquist. C'est rarement le cas en pratique : le problème de la reconstruction d'images PA et US est alors considéré comme *mal posé* : l'information contenue dans les données n'est pas suffisante pour reconstruire correctement l'objet. Elle est également, en plus de l'invalidité de certaines hypothèses réalisées (par exemple une vitesse de son constante dans le milieu), parasitée par le bruit sur les mesures. Pour compenser ces problématiques, des méthodes d'optimisation ont été proposées. L'équation de projection est discrétisée, fournissant un système d'équations linéaires : le problème de reconstruction de

l'image est alors décrit sous sa forme matricielle $Y = HX + b$ avec Y les données mesurées par la sonde, X la distribution initiale de pression représentant l'objet et b le bruit. La matrice H contient les réponses impulsionnelles $h_{r,j}(t)$. La résolution passe par la minimisation d'une équation de type $||Y - HX||$ [97], généralement en présence de termes de régularisation ayant pour but de limiter l'impact du bruit, mais nécessitant en contrepartie de faire certaines hypothèses sur l'objet à reconstruire.

Dans ces travaux de thèse, la version la plus simple du *delay-and-sum* est utilisée pour la reconstruction des images dont la qualité est par la suite améliorée par les méthodes présentées dans ce manuscrit. Dans le chapitre 5, une image reconstruite par une méthode d'optimisation est proposée pour illustrer l'incapacité de telles approches à corriger les artefacts de visibilité. En annexe A, nous montrons sur un exemple que l'utilisation d'une matrice de coefficients d'apodisation $W_j(\mathbf{r})$ impacte peu la qualité de l'image photoacoustique.

1.4 Techniques en imagerie photoacoustique et ultrasonore

1.4.1 Échographie

Généralités

L'échogénicité, l'information présente sur l'image US, traduit une différence de propriétés entre deux milieux. La majorité des ondes mesurées sont produites par la diffusion simple, qui advient lors de l'interaction entre l'onde incidente et des structures d'impédance légèrement différentes du milieu environnant et de petite taille par rapport à la longueur d'onde de l'onde émise par la sonde d'imagerie. Ces ondes contribuent à l'image US sous la forme d'une figure d'interférence appelée tavelure ou speckle. De manière générale, le speckle est formé par les interférences d'ondes cohérentes de phases distribuées aléatoirement, au moins en apparence. Cela correspond bien au phénomène de diffusion dans les tissus biologiques lors de la propagation d'une onde acoustique : les fluctuations d'impédance sont en effet considérées comme distribuées de manière aléatoire et possèdent donc une phase aléatoire, et sont de plus soumises à une excitation avec un délai de phase [52]. Les figures d'interférences dépendent de la concentration des particules de la nature du milieu. Le motif de speckle sur l'image dépendra également des propriétés de la sonde. Si sa présence induit un changement du contraste de speckle dans l'image US, elle permet de différencier deux milieux d'impédance moyenne proche mais de nature différente. Lorsque la différence d'impédance entre deux milieux est conséquente, par exemple à l'interface entre deux tissus (peau/os par exemple) on considère le cas de la réflexion spéculaire. La frontière entre ces modes de diffusion est parfois fine. La figure 1.12 présente ces deux types de contraste sur des images US conventionnelles.

Imagerie US des vaisseaux sanguins

L'échographie semble à première vue peu adaptée à l'imagerie des vaisseaux sanguins, puisque les impédances de ce tissu et de ceux l'enveloppant sont très similaires, fournissant des images peu contrastées. Cependant, la très haute résolution temporelle atteignable a permis de développer les méthodes dites Doppler qui se basent sur l'hypothèse d'un mouvement rapide du sang vis à vis d'un mouvement faible ou inexistant des tissus environnants. Ce mouvement provoque une décorrélation du signal US reçu d'une image à l'autre. Il est alors possible d'extraire les signaux provenant du sang par des méthodes de filtrage et ainsi d'augmenter fortement le contraste. Cela permet également de mesurer les vitesses de flux. L'une des principales limitations des méthodes d'imagerie Doppler est l'hypothèse de stationnarité des milieux environnants vis à vis du flux sanguin, laquelle n'est pas toujours vérifiée, par exemple dans le cadre de l'imagerie de la micro-vascularisation pour laquelle les vitesses sont faibles ou encore lorsque les tissus

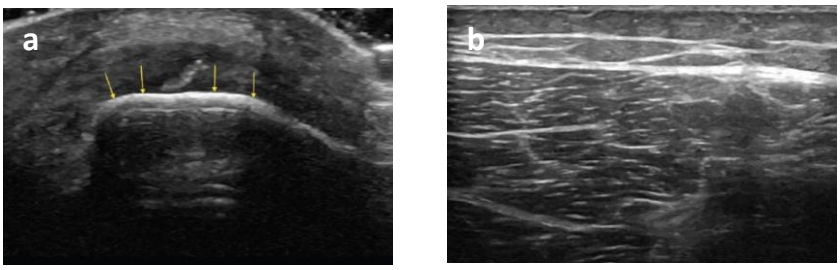

Figure 1.12 – Sonogramme montrant l'apparence d'une réflexion spéculaire en (a), flèche jaune, interface os - tissu, et une réflexion plus diffuse en (b) dans le muscle. Un signal très brillant est visible à l'interface os-tissu, alors qu'une plus grande variété de niveau de gris est visible dans le muscle. Source : https ://radiologykey.com/physics-of-ultrasound-2/

environnants se déplacent rapidement comme dans l'imagerie cardiaque ou abdominale. Le terme *clutter* définit le signal résultant des ondes diffusées par les tissus enveloppant les flux sanguins. De l'ordre de 40 à 100 dB plus grand que celui du sang [13], son impact sur la qualité de l'image est conséquent lorsque les méthodes de filtrage ne permettent pas de le supprimer.

L'imagerie ultra-rapide à ondes planes angulées, de par son excellent compromis entre champ de vue et résolution temporelle, a révolutionné l'imagerie Doppler (entre autres), et son utilisation est aujourd'hui largement répandue dans la communauté. Elle permet notamment de limiter l'impact du mouvement des tissus sur la qualité de l'image US car celle-ci est formée beaucoup plus rapidement, diminuant de fait l'amplitude de ce mouvement à l'échelle de la formation d'une image.
En imagerie US dite conventionnelle, des faisceaux focalisés sont émis, et le milieu est balayé ligne par ligne pour former l'image. La cadence d'imagerie est alors fortement limitée par le nombre de ligne, et donc par la taille du champ de vue, à environ 100 Hz. En imagerie ultra-rapide, la totalité de la zone d'imagerie est balayée en une unique impulsion acoustique, le plus souvent par une onde plane [113] : tous les échos sont donc mesurés en même temps. La résolution de l'image est obtenue en sommant de manière cohérente les signaux enregistrés par les récepteurs pour chaque pixel de l'image reconstruite : c'est l'approche *delay-and-sum*, aussi appelée focalisation en réception. Le gain en résolution temporelle est énorme (jusque 10 000 images par seconde [81]), au prix d'un contraste et d'une résolution diminués (l'intensité en chaque point du milieu est plus faible, superposition des ondes sur les signaux mesurés par la sonde). Pour compenser ces pertes, l'imagerie par combinaison d'ondes planes angulées [88] fut ensuite proposée. Plusieurs ondes planes angulées sont successivement émises dans le milieu. Pour chacune de ces ondes, une image est formée à partir des signaux reçus et recalés puis ces images sont sommées de manière cohérente. Cette méthode permet de créer une focalisation synthétique en émission [68] : un contraste et une résolution similaire à ceux obtenus en imagerie conventionnelle peuvent être atteints pour quelques dizaines d'ondes planes angulées. Un schéma représentant ces trois modes d'imagerie US est disponible en figure 1.13.

La possibilité d'obtenir une image US sur un large champ de vue avec un faible nombre d'émissions et donc une cadence très rapide a permis le développement du Doppler ultra-rapide, permettant une estimation de la vitesse du flux en tout point d'une image, contre une ligne en Doppler conventionnelle pour une estimation de qualité comparable [11], et d'améliorer le contraste et la sensibilité par un facteur 30 en imagerie Power Doppler [83]. L'ensemble de l'image étant obtenue en un temps court, des méthodes de filtrage temporel ont été utilisées pour discriminer le sang du tissu. Cependant, ce filtrage ne peut notamment pas distinguer un flux lent (capillaire sanguin) d'un mouvement physiologique de vitesse similaire. Récemment, une méthode par décomposition en valeur singulière (*Singular Value Decomposition*, SVD) [28] appliquée à une série d'images US acquises à haute cadence a ainsi été proposée : les

cohérences spatiales et temporelles sont utilisées pour discriminer bien plus efficacement le flux sanguin des tissus environnants.

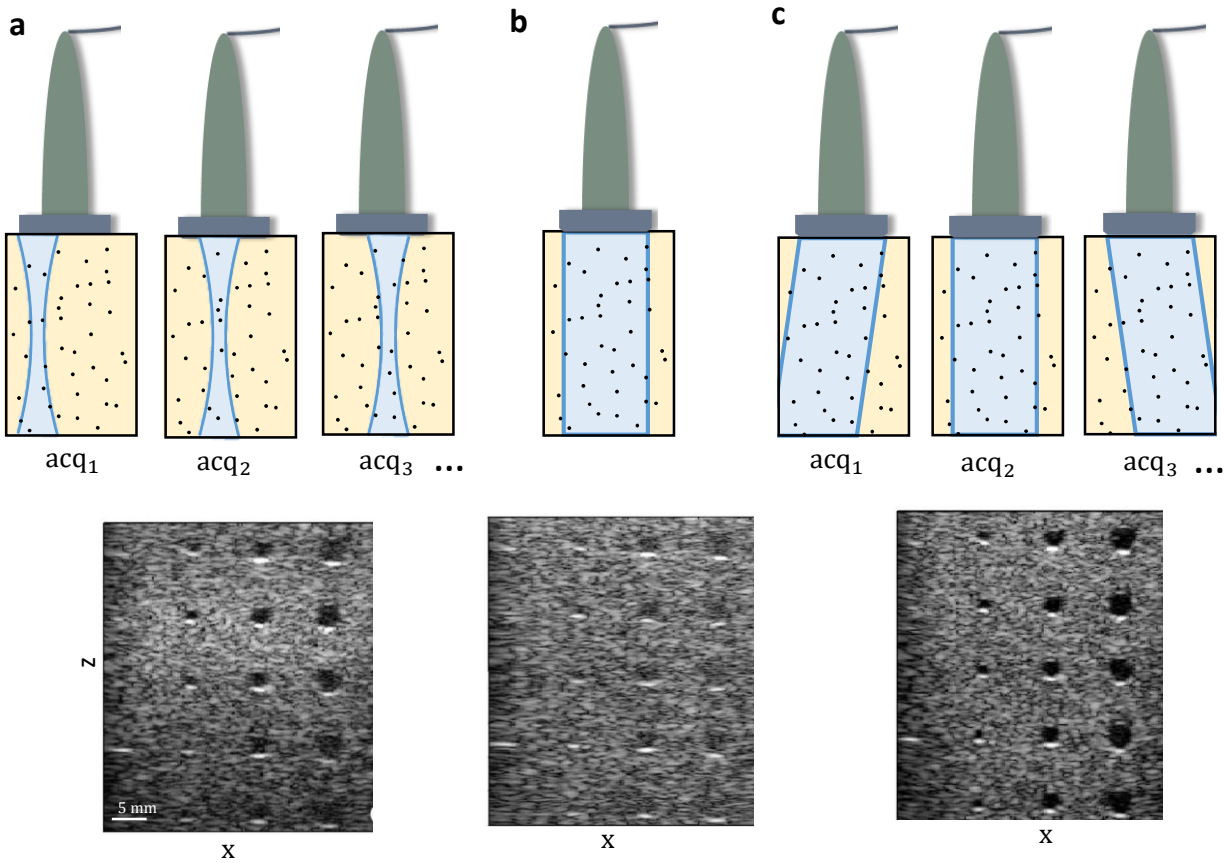

Figure 1.13 – Schéma représentant trois méthodes d'imagerie ultrasonore et images finales correspondantes. (a) : Imagerie US conventionnelle monofocale, plusieurs acquisitions du milieu sont réalisées en décalant latéralement le faisceau incident de $\lambda_{ac}/2$. (b) : imagerie avec une unique onde plane, la cadence d'imagerie est potentiellement très élevée, mais le contraste est faible. (c) : imagerie ultra-rapide à ondes planes angulées, plusieurs ondes planes avec des angles différents sont successivement transmises au milieu. Le contraste est similaire à celui obtenu en imagerie US conventionnelle pour seulement 71 angles. Modifié de [88].

1.4.2 Imagerie Photoacoustique

Généralités

L'équation 1.5 montre que l'amplitude de l'onde PA dépend directement de l'absorption optique μ_a, laquelle est fonction de la composition moléculaire de l'espèce chimique imagée. L'image PA est donc bien une image de l'absorption optique, et le contraste est apporté par la variation de ce coefficient entre les différents tissus. Cette propriété rend l'imagerie PA parfaitement adaptée à l'imagerie des vaisseaux sanguins, car l'absorption optique de l'hémoglobine contenue dans les globules rouges est forte devant celle des tissus environnants. La figure 1.14 présente la signature spectrale de quelques-uns des principaux chromophores rencontrés. Les coefficients d'absorption de l'hémoglobine oxygénée (HbO_2) et désoxygénée (Hbb), dans la gamme de 400 à 1000 nm, sont de 2 à 10 ordres de grandeur supérieurs à ceux des autres chromophores. Il n'est donc pas nécessaire d'utiliser des agents de contraste comme c'est le cas pour certaines méthodes d'imagerie optique telles que l'imagerie de fluorescence.

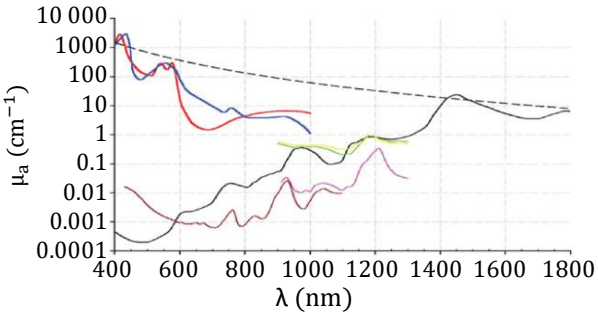

Figure 1.14 – Coefficients d'absorption de différents chromophores. En rouge l'hémoglobine, en bleu la désoxyhémoglobine, l'eau en noir, les lipides en marron et rose, le collagène en vert et l'élastine en jaune. La ligne en pointillé représente le μ_a global de la peau [7].

A l'inverse de l'imagerie US, la résolution spatiale n'est généralement pas limitée par le phénomène à l'origine de la génération de l'onde de pression, ici l'effet PA (section 1.2.3). L'utilisation d'impulsion laser nanoseconde résulte en la génération d'ondes de pression très large bande, de la dizaine à la centaine de MHz. Cependant, la dépendance spectrale de l'atténuation de l'onde ultrasonore limite en réalité la fréquence maximale de l'onde PA qui peut être détectée. Comme en imagerie US, la résolution de l'image dépend finalement de la profondeur d'exploration souhaitée qui déterminera le choix du capteur.

Approche multispectrale

Comme le montre la figure 1.14, la dépendance spectrale des principaux chromophores rencontrés dans les tissus biologiques est conséquente. L'effet PA peut ainsi être exploité par des approches multispectrales. En illuminant successivement un échantillon avec plusieurs longueurs d'onde, il est possible de discriminer ou quantifier les chromophores présents dans le milieu imagé de manière analogue à la spectroscopie par transmission optique conventionnelle. L'une des principales applications de ces approches est la quantification de la saturation en oxygène, un indicateur essentiel pour la détection de certaines pathologies cardiovasculaires, ou encore dans le cadre de l'imagerie fonctionnelle. Par une analyse spectroscopique d'images acquises à plusieurs longueurs d'onde et en utilisant notre connaissance des différences spectrales de HbO_2 et Hbb, on peut déterminer leurs concentrations et ainsi obtenir une mesure de la saturation en oxygène. Un exemple d'image ainsi obtenue est présentée en figure 1.15.

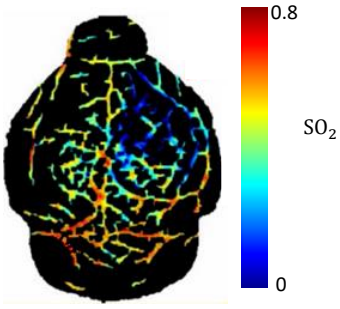

Figure 1.15 – Image de la saturation en oxygène dans un cerveau de souris inoculé par une tumeur. La tumeur, dans la partie droite supérieure du cerveau, présente une hypoxie. Modifié de [78].

Néanmoins, réussir à quantifier de manière absolue ou même relative la concentration d'un chromophore reste un défi. Comme le montre l'équation 1.5, l'énergie absorbée en un point dépend non seulement du coefficient d'absorption optique du chromophore mais également de $\Phi(\mathbf{r})$, la fluence en ce point. Cette quantité dépend de la fluence incidente du laser et de l'atténuation subie par le faisceau lors de sa propagation dans les tissus, donc de μ_a et μ_s dans toute la zone parcourue par la lumière. La forte dépendance spectrale de ces coefficients dans les tissus biologiques soulève une difficulté, puisque le

faisceau est atténué différemment en fonction de la longueur d'onde d'illumination : c'est l'effet de coloration spectrale. Pour estimer $\Phi_\lambda(\mathbf{r})$ la fluence en chaque point de la zone d'imagerie à la longueur d'onde λ, il est alors nécessaire de quantifier cette atténuation, ce qui revient à estimer les propriétés optiques du milieu. En pratique, la fluence est le plus souvent mesurée en sortie du laser, et des facteurs de correction et de calibration calculés à partir des caractéristiques des tissus sont appliqués pour obtenir la distribution de fluence dans les tissus [137]. Ces approches sont limitées par la grande diversité des tissus rencontrés, de leur variation inter-patient et de leur évolution dans le temps. Des méthodes basées sur des modèles d'inversion non linéaire en modélisant le transport de la lumière dans les tissus ont également été proposées [22].

1.4.3 Principales différences entre l'imagerie photoacoustique et ultrasonore

La formation de l'image PA et US, les facteurs affectant la qualité de l'image et sa résolution sont globalement similaires. Cependant, les phénomènes à l'origine de la génération de l'onde de pression (partie 1.2.2 et 1.2.3), donc la source du contraste, sont fondamentalement différents. L'imagerie US fournit une représentation de la différence d'impédance entre des tissus, traduisant une différence de leurs propriétés mécaniques et élastiques. L'image PA représente une distribution de pression initiale qui dépend de la quantité d'énergie lumineuse absorbée, traduisant essentiellement une différence d'absorption optique. L'imagerie PA fournit donc généralement une meilleure spécificité et différenciation que l'imagerie US, de par la plus grande variété de coefficient d'absorption et la faible différence d'impédance rencontrées dans les tissus biologiques. C'est particulièrement vrai en imagerie des vaisseaux sanguins. La forte dépendance spectrale du coefficient d'absorption optique confère un intérêt considérable à l'imagerie PA multispectrale et aux approches spectroscopiques qui en découlent. La mise en place de ces approches a été facilitée par l'apparition de laser dont la longueur d'onde d'émission est rapidement modifiable, jusqu'à pouvoir être changée à chaque impulsion laser à des cadences de l'ordre de la centaine de Hz.

En imagerie US, l'onde excitatrice est une onde ultrasonore, laquelle est moins atténuée par les tissus que l'onde lumineuse utilisée en imagerie PA. Une profondeur d'exploration supérieure peut ainsi être atteinte en imagerie US. La dépendance à l'énergie transportée par l'onde excitatrice est ainsi plus forte en imagerie PA. Par ailleurs, l'onde ultrasonore est résolue temporellement et garde sa forme au cours de sa propagation et peut donc être conçue en amont. Cela a notamment permis le développement de l'imagerie ultra-rapide provoquant à son tour l'émergence et/ou l'amélioration de méthodes d'imagerie US, tel que le Doppler ultra-rapide décrit précédemment qui a permis de fortement améliorer le contraste en imagerie US de la vascularisation.

Un autre point intéressant concerne le speckle. Il est parfois considéré dans la littérature que l'imagerie PA est moins concernée par la formation de speckle que l'imagerie US. Pourtant, ce dernier est bien présent sur les images PA. Le speckle apparaît lorsque l'impédance (en imagerie US) ou le coefficient d'absorption (en imagerie PA) dans un milieu fluctue localement : une telle situation est quasi systématique dans les tissus biologiques. Le speckle est en réalité simplement moins visible sur l'image PA car les différences de contraste entre les tissus et donc la dynamique de l'image est généralement bien plus grande. C'est notamment le cas de l'imagerie de vaisseaux sanguins pour laquelle le contraste entre les coefficients d'absorption du sang et des tissus environnants est conséquent. Ce contraste optique est ainsi largement supérieur aux variations de l'impédance entre les tissus mous. Cependant, on peut retrouver des situations analogues en imagerie US, par exemple lorsqu'un milieu incluant de l'os est imagé : le contraste d'impédance devient important, la dynamique de l'image augmente et le speckle est moins visible sur l'image US. Les forts contrastes du coefficient d'absorption optique sont en partie à l'origine de la présence des artefacts de visibilité sur les images PA. Dans le chapitre 3, nous montrons tout l'intérêt de mettre en évidence le speckle photoacoustique, révélé par l'imagerie de fluctuation, permettant de corriger ces artefacts de visibilité. Il est également intéressant de considérer le cas d'un milieu absorbant et parfaitement homogène (sans fluctuations locales d'impédance ou du coefficient d'absorption). Aucune

onde ultrasonore n'est alors enregistrée par le détecteur à la fois en imagerie US et en imagerie PA du fait de raisons cependant bien distinctes. En imagerie US, l'onde émise par la sonde traverse le milieu sans être rétrodiffusée : aucune onde n'est générée. En imagerie PA en revanche, une onde PA est générée en chaque point du milieu, mais ces ondes interfèrent destructivement entre elles.

1.4.4 Spécificité de l'imagerie 3D

Les tissus biologiques sont des structures fondamentalement tridimensionnelles : les imager en 2D limite donc leur interprétation en terme d'anatomie et d'orientation. L'imagerie 3D trouve ainsi de très nombreuses applications dans l'étude de phénomènes dynamiques 3D tels que l'étude de l'activité cardiaque, l'administration ciblée de médicaments ou encore en chirurgie assistée par ordinateur, laquelle repose de plus en plus sur des méthodes d'imagerie. Quelques exemples d'images 3D en imagerie US et PA sont présentés en figure 1.16.

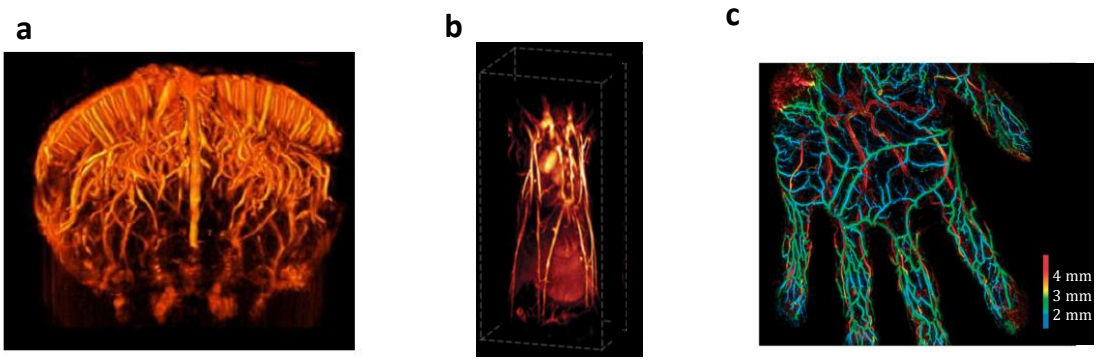

Figure 1.16 – (a) : Image Doppler ultra-rapide d'un cerveau de souris [29]. (b) : Image photoacoustique corps entier d'une souris [27]. (c) : Projection de l'intensité maximale dans le plan XY de l'image photoacoustique 3D d'une main humaine, la couleur encode la profondeur [86].

On distingue deux grandes familles de méthode d'imagerie 3D, l'une basée sur un balayage du volume avec une sonde linéaire, et l'autre basée sur l'utilisation d'un réseau de transducteurs répartis sur une surface permettant d'acquérir toute les données nécessaires à la reconstruction du volume en une unique acquisition.

Imagerie 3D par balayage

En imagerie PA et US, la manière historique pour obtenir une image 3D est de réaliser l'acquisition d'une série de plan 2D qui seront compilées pour reconstruire le volume [67]. La majorité des sondes échographiques 2D sont dotées d'un réseau linéaire de transducteurs, l'organisation la plus flexible quant à l'accessibilité des tissus. Pour certaines applications, telles que l'imagerie du petit animal, le réseau de transducteurs peut être cylindrique, encerclant la zone à imager, ou encore en arc de cercle. Une image 3D peut être obtenue par translation [138, 152] ou rotation [16]. Ce scanning mécanique limite évidemment la résolution temporelle et n'est pas simple à mettre en place, notamment en imagerie PA puisque le faisceau laser doit être déplacé simultanément avec la sonde. Deux approches de balayage sont présentées en figure 1.17.

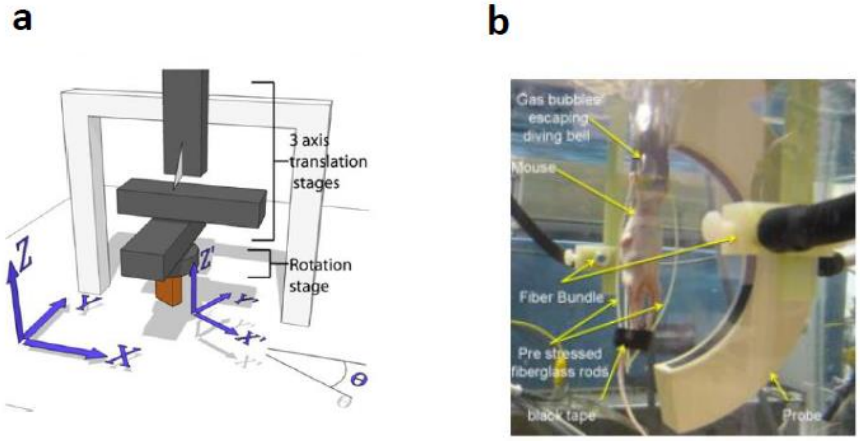

Figure 1.17 – (a) : système permettant un balayage dans les directions x et y, et une rotation autour de l'axe z [29]. (b) : sonde en forme d'arc de cercle qui tourne autour de l'animal [16].

Imagerie 3D avec une sonde matricielle

Les transducteurs peuvent également être repartis sous forme de matrice sur une surface. Avec une sonde matricielle, la totalité du volume peut être imagé en une acquisition. La principale limite de cette approche réside dans le défi technologique associé à l'utilisation d'un grand nombre de détecteurs, nécessaire pour limiter le sous-échantillonnage spatial. Dans l'idéal, la PSF a la forme d'un spot unique, mais en pratique, des lobes secondaires apparaissent dans les dimensions x et y, en raison de la taille finie du détecteur. Une organisation périodique de ces détecteurs, que l'on retrouve par exemple dans les sondes linéaires, peut conduire à des lobes de réseau si le critère de Nyquist n'est pas vérifié, donc si la fréquence d'échantillonnage spatiale ou encore l'espace entre les éléments est supérieur à $\lambda_{ac}/2$ [94] [7]. En plus des lobes de réseau, un sous-échantillonnage spatial conduit nécessairement à des phénomènes d'aliasing et donc des erreurs de quantification et plus globalement de reconstruction sur l'image finale.

Si le critère de Nyquist peut être vérifié en imagerie 2D lorsque que la fréquence centrale n'est pas trop élevée, ce n'est pas le cas en imagerie 3D, à moins de fournir un champ de vue très réduit ou alors de pouvoir disposer d'un nombre très important de transducteurs. Or ce nombre est limité technologiquement (pour une détection classique basée sur des céramiques piézoélectriques) par la difficulté de les câbler en parallèle tout en évitant les interférences. Considérons une sonde linéaire de N_{el} éléments permettant d'imager sur un champ de vue de dimension latérale $L_x = N_{el} \times w_{el}$ et respectant le critère de Nyquist, alors $N_{el} \times N_{el}$ éléments sont nécessaires pour obtenir une image 3D de qualité similaire sur un champ de vue de dimensions latérales $L_x \times L_x$. Pour une sonde linéaire de 128 éléments, son équivalent matriciel serait donc constitué de 16384 éléments. En réalité, les systèmes d'acquisition actuels disposent en clinique en général d'au plus 256 voies, quand la combinaison de plusieurs dispositifs électroniques peut permettre en laboratoire d'atteindre jusqu'à 1024 voies [99].

La surface de la sonde peut être plane [102, 103, 138] ou hémisphérique [91, 92]. Cette seconde géométrie présente l'avantage vis à vis des surfaces planes de mieux couvrir l'objet, au prix d'un champ de vue plus limité. Cette architecture est intéressante en imagerie PA, fortement affectée par l'artefact d'ouverture numérique limitée (section 1.5). Son intérêt est moindre en imagerie US notamment en imagerie ultra-rapide car le design d'ondes planes angulées en est complexifié. Si une telle surface est tout de même utilisée, on pourra alors générer successivement à partir de différents transducteurs une onde divergente, au prix de lobes secondaires et d'une intensité diminuée. La figure 1.18 présente deux exemples de tels systèmes d'imagerie.

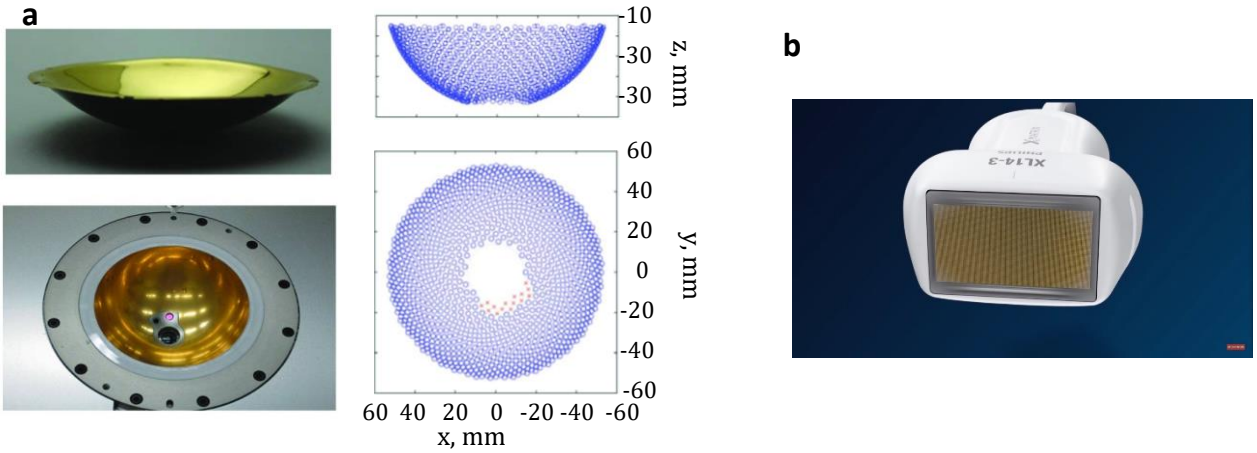

Figure 1.18 – (a) : sonde hémisphérique composée de 1024 éléments [91]. (b) : sonde 2D de type surface plane constituée de 32×32 *éléments (XL14-3 xMATRIX Linear Array transducer Philips Healthcare)*

Plusieurs approches ont été explorées pour augmenter le champ de vue sans augmenter le nombre de voies/éléments. L'utilisation de multiplexeurs a été proposée pour réduire le nombre de voies électroniques. Ils permettent de passer électroniquement d'un ensemble de détecteurs à un autre et donc de conserver un nombre élevé de transducteurs [105] [55]. Le coût est alors temporel, puisqu'une nouvelle acquisition est nécessaire pour chaque nouvel ensemble de transducteurs, et la complexité du système est majorée.

Pour limiter le nombre de détecteurs, un remplissage parcimonieux de la surface plutôt qu'un remplissage dense est possible. L'organisation des détecteurs est alors apériodique : cela présente l'intérêt majeur de limiter les lobes de réseau malgré le non respect du critère de Nyquist. De nombreux travaux visant à optimiser le motif d'échantillonnage ont ainsi été réalisés [19, 94]. En imagerie US, la capacité à former des ondes plane angulées doit être pris en compte lors du choix du motif. Les éléments peuvent à nouveau être répartis sur une surface plane [109] ou une surface hémisphérique [14, 35, 41, 145]. L'utilisation des sondes à réseaux parcimonieux s'est accompagnée du développement de nouvelles méthodes de reconstruction, notamment itératives [51], et plus récemment basées sur l'apprentissage profond (deep learning) [2] [80], dans le but de combler le manque d'information collectée, inhérent à ces réseaux.

Les sondes par adressage ligne-colonne, mises en oeuvre récemment *in vivo* en imagerie US ultrarapide [114] reposent sur la superposition orthogonale de deux réseaux constitués chacun de N électrodes, l'un représentant les lignes et l'autre les colonnes d'une matrice de taille $N \times N$. Un réseau de $N \times N$ éléments est alors virtuellement formé, dont les positions sont obtenues et l'adressage réalisé par le choix d'une paire d'électrodes appartenant chacune à l'un des deux réseaux. Une représentation est proposée en figure 1.19. Le nombre de détecteurs est ainsi diminué par un facteur de $N/2$. La présence d'importants lobes de réseau a néanmoins été rapportée. A noter que cette piste n'est pas adaptée à l'imagerie photoacoustique, car elle repose sur le contrôle de l'émission.

On peut enfin citer les détecteurs basés sur l'utilisation de l'interférométrie optique de type Fabry-Perot [153] qui ont été proposés en imagerie PA. Ces systèmes appartiennent à une catégorie à part, car la détection est réalisée optiquement et par balayage d'un faisceau optique. L'épaisseur du capteur, un film fin en polymère, varie lors du passage de l'onde PA. Ces variations sont mesurées en focalisant un faisceau optique en un point du capteur et en enregistrant la modulation de l'intensité du faisceau réfléchie. En balayant point par point le faisceau focalisé sur la surface du capteur, on synthétise un capteur multiélément permettant alors un reconstruction tridimensionnelle de l'échantillon. L'approche est schématisée en figure 1.20. L'intérêt principal de cette approche est lié au caractère large-bande de la mesure optique. Cette approche est également très efficace pour limiter le sous-échantillonnage spatial,

puisque le signal PA peut être mesuré en autant de points que nécessaire avec une résolution optique. La limitation est évidemment temporelle car le signal en chaque point est mesuré indépendamment, bien qu'elle puisse être améliorée en utilisant plusieurs faisceaux d'interrogation ou en optimisant le balayage de la surface. [61].

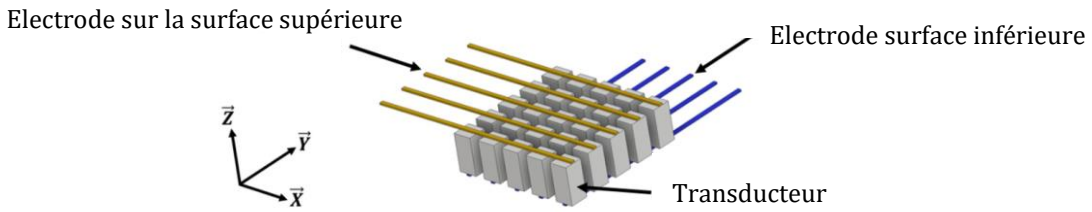

Figure 1.19 – Représentation de la sonde par adressage ligne colonne. Les interconnections entre les N électrodes sur la surface supérieure et les N électrodes de la surface inférieure permettent de former $N \times N$ transducteurs virtuels. Modifié de [114].

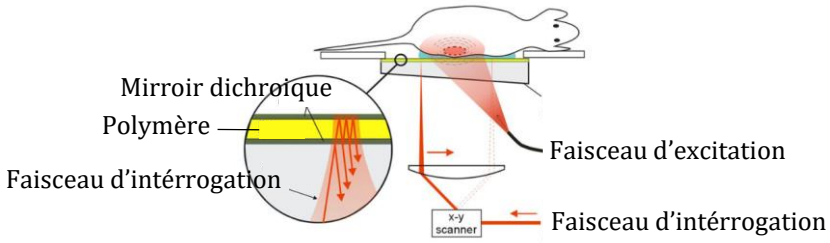

Figure 1.20 – L'épaisseur du capteur, un film fin en polymère, varie lors du passage de l'onde photoacoustique générée par le faisceau d'excitation. Ces variations sont mesurées en focalisant un faisceau d'interrogation en un point du capteur et en enregistrant la modulation de l'intensité du faisceau réfléchie. En balayant point par point le faisceau focalisé sur la surface du capteur, on synthétise un capteur multiélément. Modifié de [50].

Système d'imagerie bimodaux

La similarité des méthodes d'acquisition et de traitement du signal en imagerie PA et US a permis de concevoir rapidement des systèmes bimodaux 2D [93] malgré la différence fondamentale des modes d'excitation. La géométrie de la sonde doit néanmoins être choisie avec précaution : les sondes linéaires sont bien adaptées à l'imagerie US notamment pour l'imagerie ultra-rapide, et moins à l'imagerie PA, de par la faible couverture de l'objet. A l'inverse, les sondes concaves sont mieux adaptées à l'imagerie PA, et moins à l'imagerie US car le design des ondes émises dans le milieu est plus complexe [87]. Les sondes d'imagerie PA gagneraient également à disposer d'une sensibilité fréquentielle plus large.

Les systèmes bimodaux se sont très récemment développés en imagerie 3D, notamment dans le cas de sondes parcimonieuses [106] [39], mais également lorsque le volume est acquis par balayage mécanique [77]. La complémentarité des informations collectées par ces deux modalités est prometteuse pour de nombreuses applications de l'imagerie de vascularisation tel que la détection du cancer du sein [96] ou la caractérisation du ganglion lymphatique sentinelle [39]. Une partie des travaux réalisés au cours de cette thèse consiste ainsi en l'élaboration d'un système d'imagerie 3D multimodal permettant de réaliser simultanément l'imagerie Doppler ultra-rapide et l'imagerie photoacoustique de fluctuation.

1.5 Les artefacts de visibilité

La qualité de l'image PA est affectée par de nombreux artefacts, partagés pour certains avec l'imagerie US. Une mauvaise estimation de la fonction de transfert du milieu induite par la méconnaissance de ses propriétés, telle que la vitesse du son est une source importante d'artefacts. Le bruit, inhérent à toute mesure réelle, limite le ratio signal à bruit. Une fréquence d'échantillonnage inférieure à deux fois la fréquence moyenne de la sonde provoque de l'aliasing, de même que le sous échantillonnage spatial des signaux induits par la distribution parcimonieuse des détecteurs. Enfin, l'imagerie PA est particulièrement affectée par les artefacts dits de visibilité, produits par la directionalité des transducteurs, leur réponse fréquentielle limitée, et la couverture insuffisante de l'objet qu'offrent la plupart des systèmes d'imagerie actuels. Si les mêmes phénomènes à l'origine de ces artefacts existent en imagerie US, ils sont quasi-inexistant en imagerie de vaisseaux sanguins car les contrastes d'impédance sont faibles, à l'exception des gros vaisseaux pour lesquels la paroi est épaisse. Les méthodes qui ont été explorées dans cette thèse visent essentiellement à corriger les artefacts de visibilité en imagerie PA et à réduire les effets du sous-échantillonnage spatial en imagerie 3D. Ces derniers ayant été détaillés dans la dernière section de la partie précédente 1.4.4, nous nous concentrerons dans cette section sur les artefacts de visibilité, de deux types : l'artefact lié à l'ouverture numérique limitée et à la directionnalité des transducteurs (*limited view*) et l'artefact lié à la réponse fréquentielle limitée de la sonde d'imagerie (*limited bandwidth*).

1.5.1 Ouverture numérique limitée et directionnalité des transducteurs

En imagerie PA, l'illumination est considérée comme instantanée. Les ondes PA sont générées simultanément et sont donc fortement cohérentes lorsque les particules ne sont plus réparties aléatoirement mais de manière dense et organisée dans les milieux : de fortes interférences constructives et destructives ont alors lieu.

Lorsque les absorbeurs constituent des structures avec une orientation particulière, telles que des structures allongées dans une direction comme les vaisseaux sanguins, ces interférences produisent une onde se propageant dans une direction privilégiée, perpendiculaire à celle du vaisseau. Si le champ de vue de la sonde d'imagerie ne couvre pas tous les angles possibles, ces ondes peuvent ne pas être détectées, et en conséquence, l'objet n'est pas reconstruit [147]. Ce champ de vue dépend principalement de la géométrie de la sonde et du nombre de directions qu'elle recouvre. Ainsi, une sonde d'imagerie idéale devrait être constituée de transducteurs disposés tout autour de l'objet, sur un cercle complet en 2D, ou sur une sphère en 3D. En pratique, ce sont principalement des sondes linéaires qui sont utilisées en imagerie PA, or ces sondes sont très sensibles à cet artefact car elles ne couvrent qu'un faible nombre de directions. Les transducteurs d'une sonde idéale sont également omnidirectionnels. Or en général, la taille des transducteurs composant les sondes d'imagerie sont supérieures à la longueur d'onde de l'onde acoustique, ils ne disposent donc que d'une gamme limitée d'angles d'incidence pour lesquels l'onde est détectable (section 1.3.2), générant également des artefacts de visibilité.

Ce phénomène peut se retrouver en imagerie US en cas de réflexion spéculaire, donc de forte différence d'impédance entre deux tissus. L'onde est alors réfléchie dans une direction privilégiée, laquelle dépend également de l'angle d'incidence de l'onde d'excitation. Si cet artefact peut apparaître par exemple lorsque l'onde rencontre de l'os, en imagerie des vaisseaux sanguins les fluctuations d'impédance acoustique entre les vaisseaux et les tissus environnants sont trop faibles pour entraîner une réelle directionnalité de l'onde rétrodiffusée.

1.5.2 Réponse fréquentielle limitée

Les interférences se produisent lorsqu'un objet absorbant est large devant la durée de l'impulsion laser : les ondes PA interfèrent alors constructivement aux frontières de l'objet et destructivement à l'intérieur : un signal globalement "basse fréquence" qui dépend de la taille de l'objet se propage alors dans

les tissus. Si l'objet est trop large devant la longueur d'onde acoustique de la sonde d'imagerie, l'onde générée n'est pas enregistrée par la sonde, dont la réponse fréquentielle est limitée dans les basses fréquences (exemple d'une réponse fréquentielle en figure 1.7). Cet artefact se présente sur l'image PA par la quasi-absence de signal au centre des gros vaisseaux, masqué par le fort signal produit par les interfaces. En effet, si du speckle est bien généré entre ces interfaces (à condition que le coefficient d'absorption fluctue localement), son amplitude est généralement bien trop faible pour qu'il soit observable.

Cet artefact est également peu limitant en imagerie US des vaisseaux sanguins, de part le faible écart d'impédance entre les vaisseaux et le milieu environnant, et de l'importance de la fluctuation locale de l'impédance au sein des vaisseaux, suffisante pour faire apparaître le speckle sur l'image. La figure 1.21 illustrent les deux artefacts de visibilité de manière schématique et sur des images *in vitro* et *in vivo*.

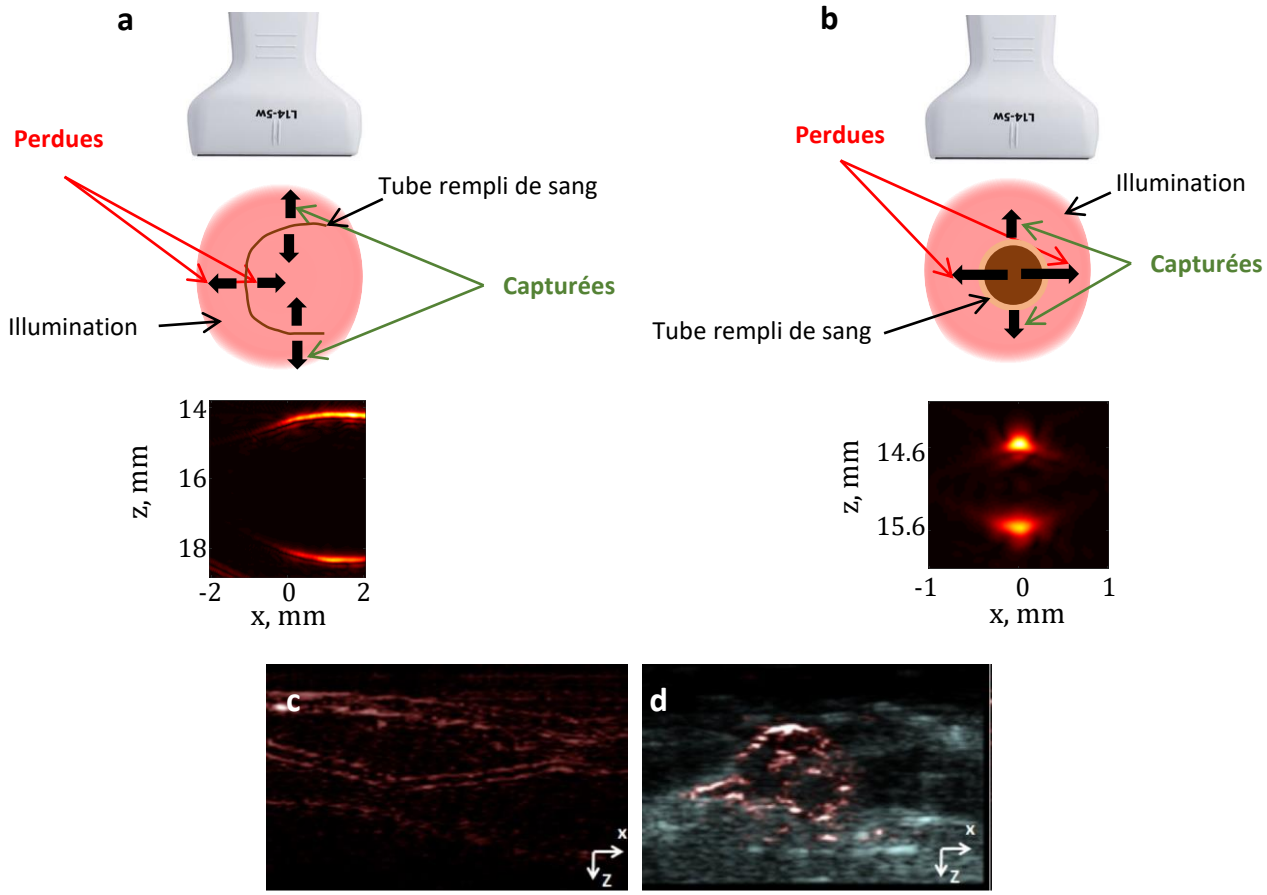

Figure 1.21 – Artefacts de visibilité.
(a,b) : fréquence centrale $f_c = 15$ MHz soit une longueur d'onde acoustique de $\lambda_{ac} = 180$ µm.
(a) : artefact d'ouverture numérique limitée : un tube en forme de boucle rempli de sang est imagé. Seules les parties perpendiculaires à l'axe de la sonde sont reconstruites. (b) : artefact de réponse fréquentielle limitée : un tube d'environ 1 mm de diamètre, supérieur à λ_{ac} est imagé. L'intérieur du tube n'est pas reconstruit.
(c,d) : artefacts d'ouverture numérique limitée et de réponse fréquentielle limitée in vivo (vaisseau en (c) et tumeur en (d) sur un système commercial (AcousticX, hhtps ://www.cyberdyne.jp).

1.5.3 Etat de l'art des méthode de correction

Les problèmes de visibilité sont donc une importante source d'artefacts, et de nombreuses approches ont été proposées pour les réduire. La manière la plus intuitive de résoudre le problème de visibilité est d'augmenter la couverture de l'objet : plutôt que d'organiser les détecteurs de manière linéaire, ceux-ci peuvent être disposés en arc de cercle, ou encore sur un cercle complet [143] [38]. Une sonde conventionnelle peut également être déplacée autour de l'objet [150], ou l'objet peut tourner sur lui même [74], au prix d'une résolution temporelle réduite (voir figure 1.17,b). Enfin, l'utilisation d'une cavité réverbérante dans laquelle sont placés l'objet et la sonde a été proposée [21, 34]. Ces approches sont fondamentalement limitées par la possibilité de pouvoir placer un détecteur sur un nombre important de directions, et ne seront donc pas implémentables pour une majorité d'applications. D'autres méthodes ne requérant pas de détecter des ondes dans toutes les directions ont donc été développées. Les problèmes de visibilité ont lieu lorsque les absorbeurs sont densément distribués. Si ces sources PA sont moins concentrées, ces artefacts disparaissent : l'injection d'absorbeurs exogènes en faible quantité a donc été proposée. Les ondes émises par ces absorbeurs n'interagissent pas ou peu entre elles et peuvent donc être détectées, par des méthodes dites de localisation [26] ou par des combinaisons non linéaires de reconstructions tomographiques représentant la distribution des particules en mouvement [25]. Cette méthode n'est cependant pas directement applicable à des absorbeurs endogènes tels que les globules rouges, trop concentrés dans le sang : l'utilisation d'un agent de contraste est indispensable. Pour obtenir une distribution parcimonieuse de sources PA, une autre approche consiste à générer artificiellement une source en chauffant localement un tissu à l'aide d'un faisceau ultrasonore focalisé [134]. Cette méthode est en pratique limitée par les seuils de sécurité rapidement atteints pour l'onde US focalisée, et requiert, comme pour les méthodes impliquant l'utilisation d'agents exogènes, un temps d'acquisition conséquent.
L'une des approches utilisées dans ces travaux de thèse, dite imagerie de fluctuation, se base sur l'utilisation d'une modulation spatiale du signal PA et de sa variation dans le temps lorsque la répartition des absorbeurs change entre chaque acquisition. Elle est présentée en détail dans le chapitre 3. La seconde méthode implique l'utilisation d'un algorithme de *deep learning* et est décrite dans le chapitre 5.

1.6 Objectifs de la thèse

Les travaux de thèse présentés dans ce manuscrit ont pour premier objectif de développer des méthodes de correction des artefacts de visibilité. Deux approches ont été explorées : une première exploitant la fluctuation du signal PA dans les vaisseaux sanguins, et une seconde s'appuyant sur de l'apprentissage supervisé (*deep learning*). Ces dernières années, l'intelligence artificielle s'est en effet massivement répandue en imagerie médicale et a permis d'atteindre des performances inégalées dans plusieurs tâches, et notamment en traitement d'image. La possibilité de combiner les deux approches pour pallier à leurs faiblesses respectives a ensuite été étudiée. Un dispositif d'imagerie 3D de la vascularisation PA et US a également été développé, sur lequel la méthode d'imagerie par fluctuation a été implémentée ainsi qu'une approche Doppler ultra-rapide. Ce dispositif a été mis en oeuvre et évalué *in vivo*. Ce dispositif a pour particularité d'utiliser une sonde matricielle parcimonieuse à 256 éléments, permettant d'imager en 3D en temps réel en évitant l'utilisation d'un système d'acquisition coûteux et complexe. Dans ce manuscrit, *in vivo*, réfère principalement à l'imagerie d'embryon de poulet, bien que quelques résultats préliminaires sont également présentés dans le cas de l'imagerie du petit animal (souris).

Les dispositifs expérimentaux sont décrits dans le second chapitre, ainsi que les protocoles de préparation de l'embryon de poulet et de la souris, les deux modèles étudiés. En particulier, la caractérisation du dispositif d'imagerie 3D PA/US et son utilisation sont détaillées.

Dans le troisième chapitre, la méthode de correction des artefacts de visibilité appelée imagerie photoacoustique de fluctuation sera présentée de manière approfondie, de la théorie à l'imagerie *in vitro* et *in vivo* et par une étude de l'impact de différents paramètres sur la qualité de l'image. Les résultats *in vivo* sont obtenus à partir du dispositif d'imagerie 3D et seront accompagnés de l'image Power Doppler associée.

Le quatrième chapitre est consacré à l'extension de cette méthode à l'imagerie quantitative. Nous montrons qu'il est possible d'obtenir une image 3D de la saturation en oxygène pour des vaisseaux invisibles sur l'image conventionnelle, mais visible sur l'image de fluctuation. A nouveau, la méthode est implémentée *in vivo* et mise en parallèle avec des images US de la vitesse et de la direction du flux, reconstruites à partir de données acquises simultanément.

Une méthode basée sur un algorithme de *deep learning* est proposée dans le chapitre 5 pour démontrer la capacité de ce type d'approches à résoudre les problèmes de visibilité. La méthode implémentée permet également de localiser des erreurs de prédiction, améliorant la fiabilité des images obtenues. Dans ce chapitre, les grands concepts de l'apprentissage profond seront au préalable présentés.

Le chapitre 6 conclura ce manuscrit par la combinaison de l'approche *deep learning* et imagerie de fluctuation, permettant d'améliorer la résolution temporelle et de constituer un jeu de données quantitatif en milieu opaque. Il sera constitué de résultats de simulation et de quelques résultats expérimentaux préliminaires.

2

Configuration expérimentale et sonde d'imagerie

Sommaire

2.1 Configuration expérimentale 42
- 2.1.1 Source de lumière 42
- 2.1.2 Système d'acquisition des ondes ultrasonores 42
- 2.1.3 Bruit de mesure 44
- 2.1.4 Synchronisation en imagerie photoacoustique 45
- 2.1.5 Schéma du dispositif expérimental utilisant le laser Spitelight 45
- 2.1.6 Spécificité du dispositif expérimental pour l'imagerie 3D en multi-longueur d'onde 46

2.2 Description des sondes utilisées dans ces travaux de thèse 50
- 2.2.1 Sonde d'imagerie 2D 50
- 2.2.2 Sonde d'imagerie 3D 52

2.3 Expériences *in vivo* : méthodes 59
- 2.3.1 Séquence d'acquisition et interfaçage 59
- 2.3.2 Préparation des embryons de poulet 61
- 2.3.3 Organisation d'une expérience souris 62

2.4 Conclusion 63

2.1 Configuration expérimentale

Les résultats expérimentaux présentés dans cette thèse ont été obtenus à partir de différentes configurations, variant en fonction de la nature de l'objet (*in vitro* / *in vivo*), de la dimensionnalité des images (2D/3D) et de l'utilisation ou non de plusieurs longueurs d'onde. Les démarches sont similaires pour toutes ces configurations, ce qui permet de les réunir dans cette partie. De manière générale, une acquisition de signaux PA est réalisée de la manière suivante : une impulsion laser (Surrelite, Continuum, USA ou Spitlight DPSS 250, Innolas laser, Germany) illumine un échantillon couplé à la sonde d'imagerie (de fréquence centrale $f_0 = 15MHz$ et focale à 15 mm pour l'imagerie 2D, et de fréquence centrale $f_0 = 8MHz$ et focale à 35 mm pour l'imagerie 3D). Pour chaque tir de laser, les ondes générées sont captées par la sonde et les signaux enregistrés par le système d'acquisition ultrasonore aussi appelé échographe (High Frequency Verasonics Vantage system), puis les données sont transférées à l'ordinateur. La fluence du laser peut éventuellement être enregistrée par une photodiode. Pour toutes les acquisitions US, la même sonde est utilisée en transmission/réception.

2.1.1 Source de lumière

Deux lasers différents ont été utilisés au cours de cette thèse, le SpitLight d'Innolas et le Surrelite de Continuum. Ces deux lasers sont des lasers impulsionnels, d'une durée d'impulsion τ_p de 5 ns pour l'Innolas et de 6 ns pour le Surrelite. Dans les deux cas, le faisceau lumineux est émis par un laser à cristal Nd-YAG pompé par diode pour le SpitLight et par une lampe flash pour le Surrelite, émettant un faisceau à 1064 nm. Ils disposent tous deux d'un générateur de seconde harmonique laissant la possibilité d'émettre également à 532 nm. Ces lasers sont dits *Q-Switched*. Lorsque la cellule de Pockels, le dispositif qui gère l'émission du faisceau, est inactif, la cavité est fermée et le laser ne 'lase' pas. Pour que le faisceau soit émis, une impulsion de haute tension active brièvement la cellule de Pockels ce qui ouvre la cavité. Le délai entre la décharge de la lampe flash / diode peut être ajusté pour maximiser ou diminuer l'énergie par impulsion. La fréquence de répétition maximale (PRF) est de 100 Hz pour le SpitLight et de 10 Hz pour le Surrelite. Le Spitlight d'Innolas dispose d'un oscillateur paramétrique optique (OPO) permettant de générer des faisceaux de longueur d'onde allant de 680 nm à 980 nm, et de changer cette longueur d'onde à la PRF du laser. Des photographies de ces lasers sont disponibles en figure 2.1.

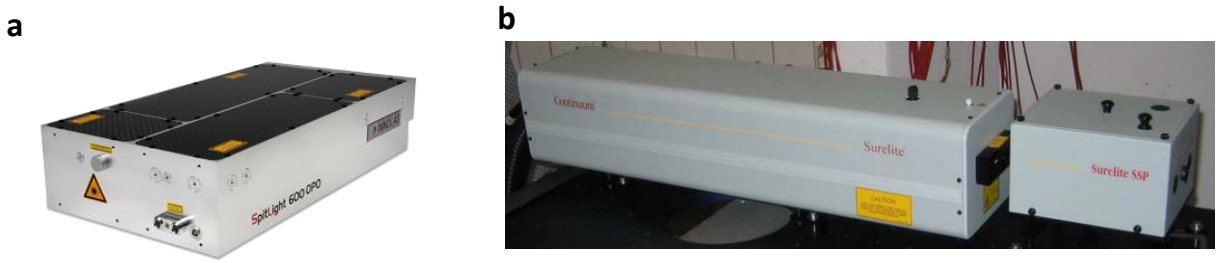

Figure 2.1 – Photographie du laser Innolas en (a) et du laser Surrelite en (b).

Le choix d'utiliser deux lasers différents n'a été dicté que par la volonté de disposer de dispositifs indépendants pour les différentes expériences. Le Surrelite de Continuum a été utilisé pour acquérir le jeu de données utilisés dans le chapitre 4, et le SpitLight d'Innolas a été utilisé pour l'ensemble des autres expériences.

2.1.2 Système d'acquisition des ondes ultrasonores

La détection des ondes de pression et leurs émissions en imagerie US sont assurées par une sonde d'imagerie connectée à un échographe (High Frequency Vantage 256 system, Verasonics, Kirkland, Wa, USA) pouvant gérer jusque 256 voies indépendantes, en émission comme en réception. La connexion

entre les voies de la sonde d'imagerie et l'échographe est réalisée au travers de deux "Cannon-HDI-Format ZIF connecteurs" de 128 canaux chacun. Des photographies ainsi qu'un schéma de l'échographe est disponible en figure 2.2, sur lesquels les différents composants du système d'acquisition sont visibles, notamment :

— *ScanHead Interface* : composant qui contient les connecteurs à la sonde d'imagerie.

— Modules d'acquisition : cartes électroniques permettant la transmission et la réception des signaux pour plusieurs transducteurs, assurant notamment la digitalisation, ainsi qu'un éventuel filtrage et le conditionnement des signaux. Ils incluent également un espace de stockage utilisé avant le transfert des données vers l'ordinateur.

— *Back Plane Module* : ce module contient le séquenceur hardware qui contrôle les opérations des modules d'acquisition.

— *Transmit Power Controller* : ce module fourni l'alimentation électrique à plusieurs modules du système d'acquisition, et notamment l'alimentation haute tension nécessaire aux transducteurs lorsqu'ils sont utilisés en émission dans le cadre de l'imagerie US.

— *I/O Panel* : ce module permet la connexion d'un câble PCI-express qui relie l'échographe à l'ordinateur. Il fournit également deux entrées "Trig in" et une sortie "Trig out" utilisables pour émettre et recevoir des signaux de déclenchement. Ce module fournit également une connexion à l'horloge interne pour éventuellement synchroniser des dispositifs externes.

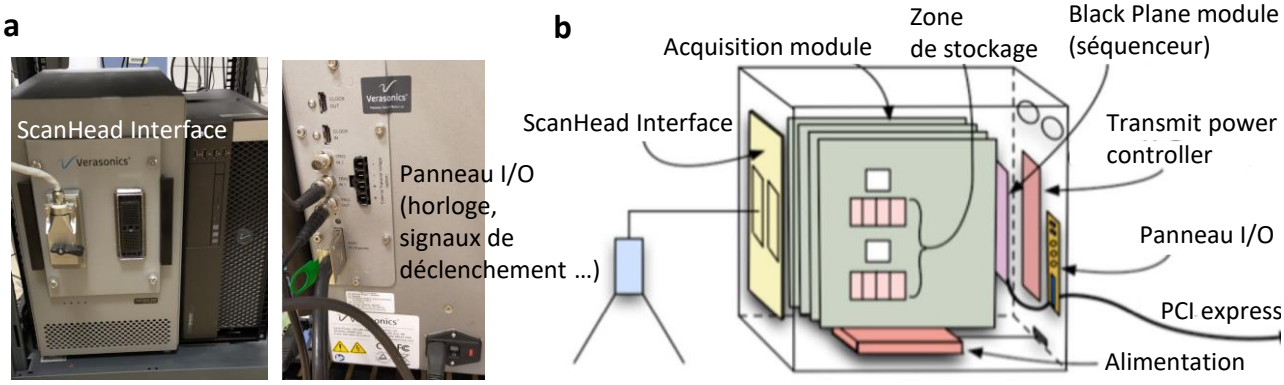

Figure 2.2 – Photographies de l'échographe "High Frequency Vantage 256 system, Verasonics" en (a) et schéma correspondant en (b). Sur la photographie de la Scan Head Interface, l'un des deux connecteurs est relié à une sonde d'imagerie.

La fréquence de l'horloge interne du système d'acquisition est de 250 MHz et la fréquence d'échantillonnage maximale f_e est de 62.5 MHz. La quantification est réalisée en 15 bits signés. Le taux de transfert vers le PC est théoriquement de 27 Mbytes/s par voie lorsque les 2×128 voies sont actives.

La partie software fait le lien entre la partie hardware et l'environnement Matlab à partir duquel le contrôle de l'acquisition et la visualisation des données sont réalisés. Une grande liberté est laissée à l'utilisateur : la majorité des paramètres d'acquisition sont modifiables, notamment :

— Les caractéristiques de la sonde (position des éléments, fréquence centrale f_c etc ...).

— Le gain, dont la valeur dépendra de l'amplitude du signal, l'objectif étant d'optimiser la quantification des signaux tout en évitant la saturation.

— La fréquence d'échantillonnage f_e, laquelle sera d'au moins $2 \times f_{max}$ (pour éviter l'aliasing, avec f_{max} la fréquence maximale à 100% de la bande passante), ainsi que le choix des coefficients de filtres analogiques et digitaux passe-bas et passe-bande.

— Les paramètres de reconstruction tels que la vitesse du son le pas d'échantillonnage spatial ou encore la taille du champ de vue ainsi qu'un éventuel décalage du t_0, la valeur correspondant à l'origine du temps de l'acquisition, laquelle est ajustée en fonction du degré de synchronisation entre l'acquisition et l'excitation (illumination ou émission ultrasonore).

— En imagerie US, le contrôle de la forme des ondes émises de par le choix du nombre de cycles, de leurs phases, le choix des éléments utilisés ainsi que de leurs lois de retard permettant par exemple d'anguler une onde plane.

Une grande flexibilité est également permise dans la conception de la séquence d'imagerie et de l'interface de contrôle, ce qui est illustrée dans la section 2.3.1.

Les signaux enregistrés par une sonde ultrasonore, que l'on appelle signaux radiofréquence (RF), sont fondamentalement réels : une variation de pression est traduite en une tension électrique par les détecteurs. Comme présenté dans le chapitre 1, ces signaux sont souvent exploités sous leurs formes en quadrature. Dans nos expériences, l'échographe nous fournit des signaux RF sous forme réelle. Si nécessaire, nous utilisons une transformée de Hilbert pour les convertir sous leur forme en quadrature.

2.1.3 Bruit de mesure

La présence de bruit électronique sur les données acquises diminue le contraste et influence leurs analyses. Sur notre dispositif expérimental, deux sources ont été identifiées. L'électronique du laser produit un bruit électromagnétique qui est capté simultanément par tous les transducteurs, apparaissant sous la forme de bandes temporelles sur les signaux RFs. Ce bruit peut être réduit en soustrayant la moyenne du signal sur tous les transducteurs pour chaque ligne (une ligne correspondant à un indice de temps). La position de ces bandes variant entre chaque acquisition, le moyennage de jeux de données RFs correspondant à plusieurs acquisitions peut également permettre de diminuer leurs amplitudes.

L'électronique de l'échographe et de la sonde d'imagerie produit également un bruit blanc, caractérisable par son écart type dont la valeur dépendra essentiellement du gain appliqué aux signaux. Le moyennage de données provenant de plusieurs acquisitions permet également de diminuer l'impact de ce bruit, dont l'écart type diminuera en racine carré du nombre d'images moyennées. Lorsque la gain est maximal, un bruit électronique de l'ordre de 30-50 niveaux de quantification (sur 16384) est observé. Ces bruits sont visibles en figure 2.3.

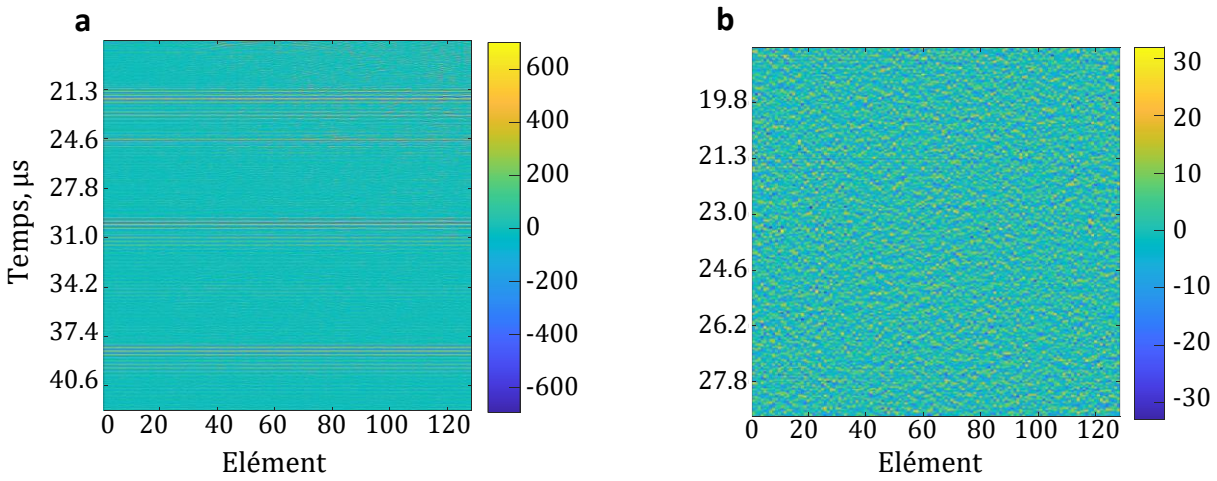

Figure 2.3 – Bruit de mesure sur les signaux radiofréquences expérimentaux.
(a) : bruit électromagnétique généré par l'électronique du laser, apparaissant sous forme de bandes temporelles. (b) : bruit électronique.

2.1.4 Synchronisation en imagerie photoacoustique

La synchronisation entre l'illumination et le début de l'acquisition des signaux est essentielle pour pouvoir reconstruire correctement les images. Cela permet de connaître avec précision le t_0, soit le temps auquel les signaux sont générés dans le milieu vis à vis du début de l'enregistrement par l'échographe. Une mauvaise estimation de ce t_0 conduit à des erreurs lors de l'opération de *delay and sum*, les temps de vol étant calculés en supposant ce t_0 connu. Cette synchronisation est difficile à obtenir en pratique, car les échographes comme les lasers ne sont généralement pas conçus pour être contrôlés avec une précision nanoseconde. On définit le *jitter* comme l'écart type de la durée entre l'instant de l'émission de la lumière et l'origine des temps de l'échographe. Plus ce *jitter* est faible, plus l'estimation de cette durée et donc du t_0 sera précise. On utilisera le *jitter* comme une mesure de la qualité de la synchronisation.

On exclut la possibilité d'utiliser l'échographe comme "maître" pour que laser puisse fonctionner en l'absence de l'échographe, notamment dans le but de le préchauffer en amont d'une expérience afin d'atteindre un état stationnaire à l'intérieur de la cavité. Considérons deux schémas de synchronisation.
1) Le laser est "maître", un signal de déclenchement en provenance de la source de pompage ("Trig FL") du laser est transmis à l'échographe. Celui-ci attend pendant un temps supposé fixe et connu pour que l'énergie dans la cavité atteigne son maximum et que la lumière soit émise (235 µm pour le laser Innolas et µm pour le Spitelight), avant de débuter l'enregistrement des signaux. Cette solution n'est pas idéale, car l'échographe prend en compte les déclenchements externes sur une horloge interne cadencée à la fréquence centrale de la sonde f_c, lente à l'échelle de l'impulsion laser ce qui augmente considérablement le *jitter*, de l'ordre pour ce schéma de 1 µs.
2) Le laser est toujours "maître", mais c'est cette fois l'échographe qui déclenche l'illumination. Comme pour le schéma 1, un signal de déclenchement en provenance de la source de pompage du laser est transmis à l'échographe. Celui-ci marque la même pause que dans le schéma 1, mais transmet cette fois un signal à la cellule de Pockels du laser ce qui déclenche l'illumination. L'horloge interne du laser est ainsi respectée car la source de pompage reste 'maître'. L'illumination est déclenchée sur l'horloge interne de l'échographe : la source d'incertitude liée à la lenteur de cette horloge n'influe donc plus sur la synchronisation. Un *jitter* de 5 ns est obtenu : ce schéma de synchronisation est donc utilisé lors de nos expériences

2.1.5 Schéma du dispositif expérimental utilisant le laser Spitelight

Le schéma du dispositif incluant le Spitelight d'Innolas utilisé dans les expériences mono longueur d'onde est présenté en figure 2.4, ici pour l'acquisition de la *point spread function* (PSF) de la sonde d'imagerie 2D (section 2.2.1). La compatibilité entre la sortie "FL trig" du laser et l'entrée "Trig in" de l'échographe est assurée par un diviseur de tension (Div V). Un générateur amplifie et inverse la polarité de la sortie "Trig out" de l'échographe pour produire le signal d'entrée "Pockels Trig" qui déclenche la cellule de Pockels et l'illumination.

Pour réduire davantage le *jitter*, la synchronisation entre le signal "Pockels trig" et l'ouverture de la cellule de Pockels, donc de l'émission du faisceau a également été augmentée par l'utilisation d'un circuit bypass, permettant de passer d'un jitter de ± 5 ns à ± 1 ns.

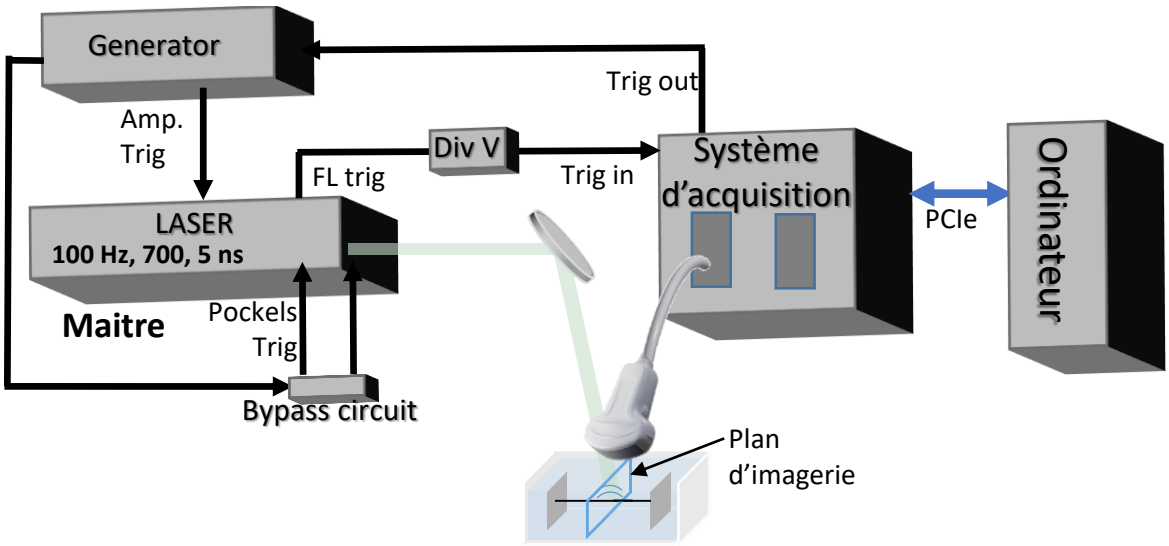

Figure 2.4 – Acquisition de type mono-longueur d'onde 2D.
Le laser émet un signal de déclenchement ("FL trig") qui passe par le diviseur de tension (Div V) pour être converti en un signal de déclenchement de l'électronique ("Trig in") de la Verasonics. Après une pause de 235 µm, un signal de déclenchement ("Trig out") est transmis au générateur lequel à son tour transmet un signal de déclenchement ("Pockels trig") qui traverse le circuit bypass avant d'atteindre la cellule de Pockels du laser, ce qui provoque l'émission du faisceau. L'acquisition des ondes photoacoustiques débute à la réception du signal "Pockels trig".
L'échographe et l'ordinateur sont reliés par une interface PCI express. L'acquisition de la point spread function de la sonde d'imagerie 2D est ici représentée : un fil est placé perpendiculairement au plan d'imagerie de la sonde. Le faisceau lumineux est conduit jusqu'à l'échantillon par une série de miroirs (un seul représenté sur le schéma). La sonde d'imagerie 2D ne dispose que d'un seul connecteur.

Le dispositif expérimental incluant le laser Spitelight est très similaire à celui incluant le laser Innolas et n'est donc pas décrit.

2.1.6 Spécificité du dispositif expérimental pour l'imagerie 3D en multi-longueur d'onde

Couplage acoustique entre la sonde et l'échantillon et optique entre la sonde et le laser

Un bon couplage entre la sonde d'imagerie et l'échantillon est nécessaire pour éviter une réflexion dans le milieu de l'onde de pression générée par les absorbeurs ou diffuseurs, réflexion provoquée par la grande différence d'impédance acoustique entre les tissus biologiques et l'air. Pour les sondes d'imagerie conventionnelles, les transducteurs sont directement au contact des tissus. Un gel de contact est généralement placé à l'interface entre la surface de la sonde et les tissus, permettant d'éviter cette rupture d'impédance. Dans le cas de notre sonde d'imagerie 3D (section 2.2.2), un espace important doit être comblé pour faire la jonction entre les transducteurs et les tissus. Un cône de couplage conçu et fabriqué dans notre laboratoire est alors utilisé. Ce cône est rempli d'eau pour assurer la continuité d'impédance, et le contact avec l'échantillon est assuré par une membrane de latex. La pression subie par cette membrane, contrôlée par le volume d'eau injecté dans le cône, permet de modifier de quelques mm la profondeur de la surface de l'échantillon, sans pour autant perdre le contact acoustique.

Une fibre optique (Ceramoptec, Germany) est couplée à la sonde, permettant d'illuminer le centre du champ de vue et de pouvoir déplacer la sonde sans réajuster la trajectoire du faisceau comme c'est le cas avec la sonde d'imagerie 2D. Une photographie de la sonde (prise lors de la calibration de la photodiode, section 2.1.6) est présentée en figure 2.5.

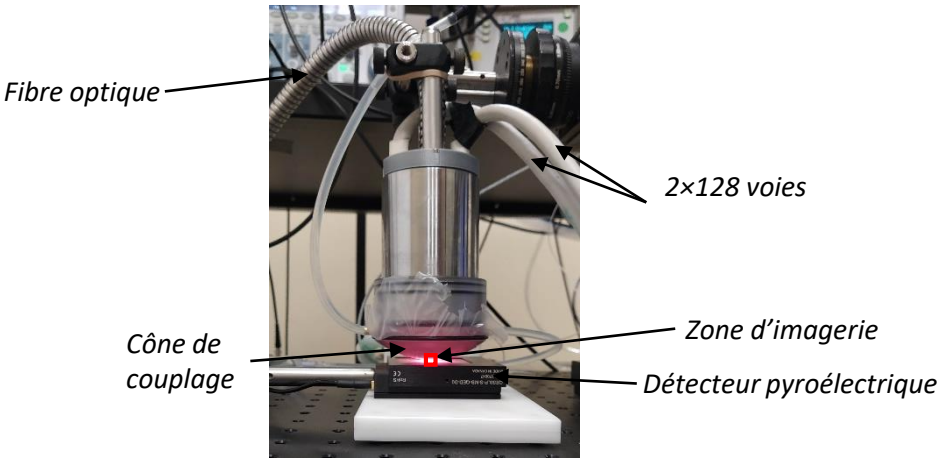

Figure 2.5 – Sonde d'imagerie 3D. La photographie est prise lors de la calibration de la photodiode.

Mesure de la fluence en sortie du laser

Dans le cadre des expériences quantitatives, du chapitre 4, il est nécessaire de connaître la fluence du laser à la surface de l'échantillon pour extraire des images PA une grandeur proportionnelle à l'absorption optique. Le dispositif expérimental pour l'imagerie 3D incluant la mesure de fluence est présenté en figure 2.6 dans le cadre de l'acquisition de la *point spread function* de la sonde. Le diviseur de tension et le circuit bypass sont bien utilisés, mais n'apparaissent pas sur le schéma par soucis de lisibilité. Un signal qui dépend de la fluence en sortie du laser est mesurée en continu par une photodiode rapide (Thorlabs, DET10A). A chaque impulsion, le signal de la photodiode est intégré par un intégrateur à porte (Boxcar, Standford Research System, SR 25) sur une fenêtre temporelle centrée sur l'impulsion dont la taille a été déterminée à l'aide d'un oscilloscope. Le signal résultant est digitalisé par le DAQ, une carte d'acquisition USB (TiePie, HS6). L'étape d'intégration analogique du signal photodiode est nécessaire pour bien quantifier l'énergie par impulsion car la durée de ce signal est d'environ 20 ns (figure 2.7.a) : il est donc difficile à échantillonner précisément. Dans notre cas, la fréquence d'échantillonnage de notre carte d'acquisition n'était pas suffisante.

L'intégration est déclenchée par un signal ("Trig B") fourni par un générateur à chaque illumination. Le signal intégré est visible en figure 2.7.b pour une dizaine d'impulsions. On peut y observer la variation de l'énergie produite par le laser entre chaque impulsion.

Calibration de la fluence mesurée en sortie du laser

Ce dispositif nous permet de mesurer une grandeur qui dépend de la fluence en sortie du laser. Or la valeur d'intérêt est la fluence au niveau de la zone d'imagerie. Entre notre mesure et le moment où le faisceau atteint cette zone, le faisceau transite par la fibre optique et la membrane en latex, il est donc atténué. La fibre et la membrane possèdent par ailleurs leurs propres réponses spectrales, tout comme la photodiode. La fluence mesurée n'est donc pas réellement celle qui atteindra l'échantillon. Plus problématique encore dans le cas d'une acquisition multi-longueurs d'onde, la fluence mesurée n'est pas proportionnelle entre longueur d'onde à celle qui atteint réellement l'échantillon. Il est alors impossible de corriger l'image PA proportionnelle pour obtenir l'absorption optique.

Pour corriger ces effets, une calibration est réalisée avant chaque série de mesures, dont le principe est décrit en figure 2.8. Un détecteur pyroélectrique qui permet de quantifier l'énergie de chaque impulsion (Gentec QE50, France) dont la réponse spectrale est constante, est placé au niveau de la zone d'imagerie (Fig. 2.5). Une série de plusieurs tirs de laser pour différentes longueurs d'onde est alors réalisée, et la fluence est simultanément enregistrée par la photodiode et ce détecteur pyroélectrique (Fig. 2.8.a). Un

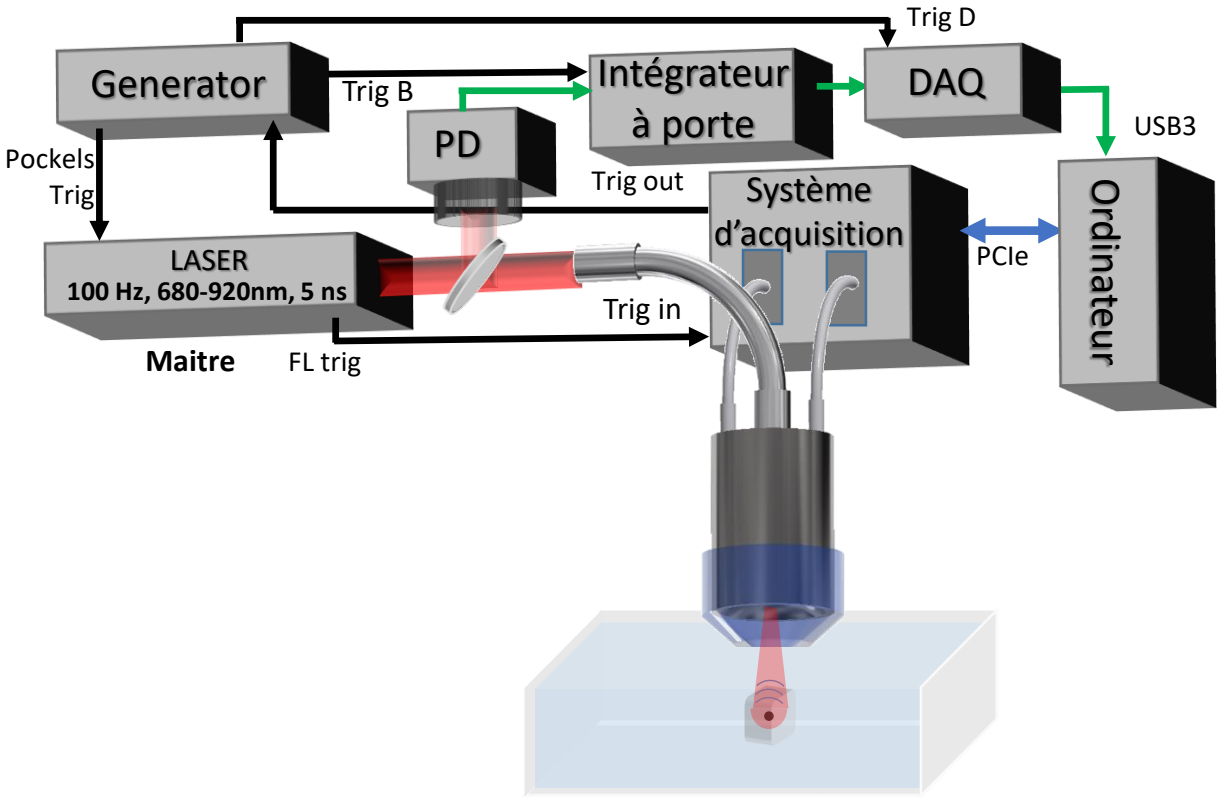

Figure 2.6 – Dispositif expérimental pour l'imagerie 3D avec mesure de fluence. Les signaux de déclenchement ("Trig") sont représentés en noir et le chemin du signal contenant l'information de fluence en vert.
La photodiode enregistre l'énergie de chaque impulsion laser. Ce signal est intégré par l'intégrateur à porte à chaque réception d'un signal de déclenchement fourni par le générateur ("Trig B"). Ce signal de déclenchement ("Trig B") est envoyé à la réception du signal de déclenchement fourni par l'échographe ("Trig out") qui conditionne également le début de l'illumination au travers du "Pockel Trig". Le signal de la photodiode intégré est transmis à la carte d'acquisition (DAQ), lequel l'échantillonne et le convertit numériquement avant de le transmettre au PC.
L'acquisition de la point spread function de la sonde d'imagerie D est ici représentée, pour un faisceau d'illumination de longueur d'onde 700 nm.

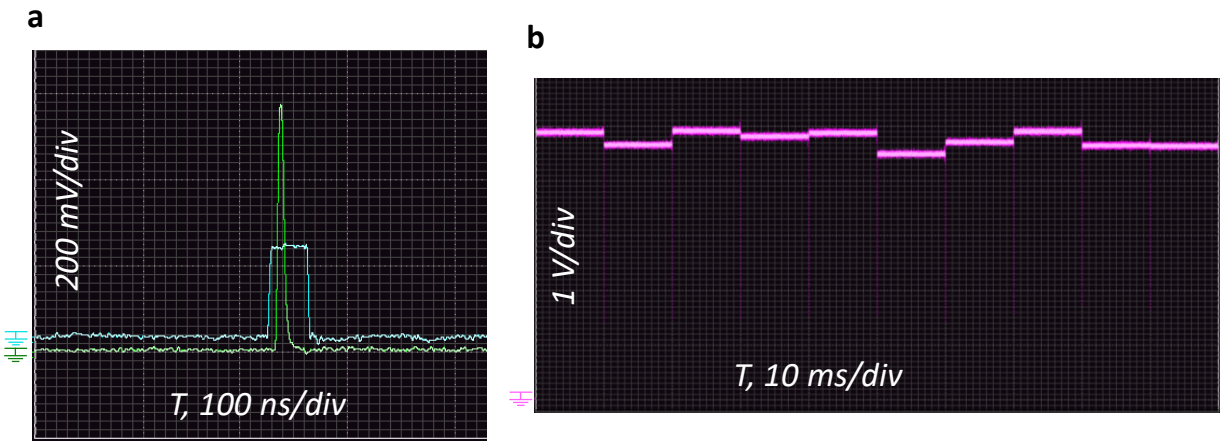

Figure 2.7 – (a) : Signal de la photodiode pour une impulsion en vert et porte d'intégration en bleu. (b) : Résultat de l'intégration pour une dizaine d'impulsions (chaque palier correspond à une impulsion). la variation de la fluence entre chaque tir est visible. C'est ce signal qui est ensuite converti par la carte d'acquisition.

coefficient de calibration est ensuite obtenu pour chaque longueur d'onde λ (Fig. 2.8.c) en réalisant une régression linéaire entre les mesures des deux appareils (exemple en fig. 2.8.b pour $\lambda = 700nm$). Ce coefficient nous permet donc de déduire, à partir de la mesure photodiode, la fluence qui atteint réellement l'échantillon. Une validation de cette calibration est proposée dans le chapitre 4 section 4.3.1 dans le cadre des expériences quantitatives.

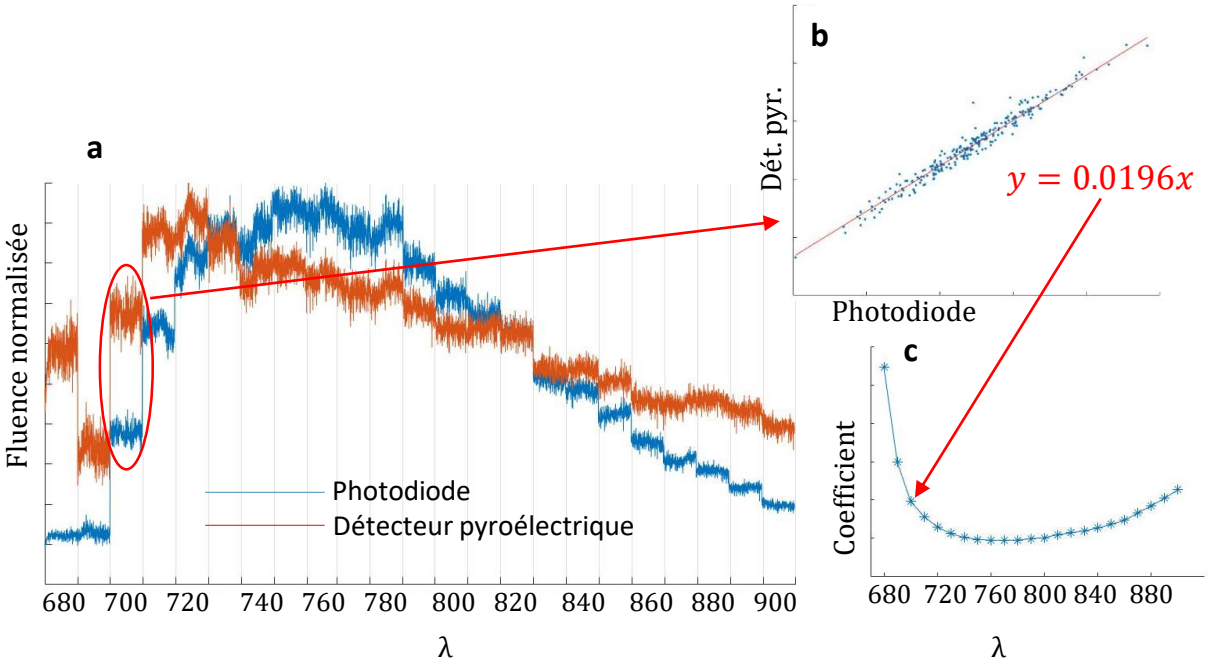

Figure 2.8 – Calibration du signal de la photodiode.
(a) : Le signal de la photodiode (bleu) et celui du détecteur pyroélectrique (orange) sont simultanément enregistrés, ici pour des λ variant de 680 à 900 nm par pas de 10 nm et 300 illuminations par λ. (b) : pour chaque longueur d'onde (ici le cas de $\lambda = 700$ nm est présenté), le coefficient de la régression linéaire (courbe rouge) entre les deux mesures est obtenu. (c) : Représentation des coefficients de calibration de l'ensemble des longueurs d'onde.

2.2 Description des sondes utilisées dans ces travaux de thèse

2.2.1 Sonde d'imagerie 2D

Une sonde de type L228-v (CMUT, Verasonics) de fréquence centrale 15 MHz est utilisée pour les expériences 2D, notamment pour les expériences du chapitre 5. Une photographie de la sonde est proposée en figure 2.9. Bien qu'utilisée au cours de cette thèse en imagerie US et PA, seuls des résultats en imagerie PA sont présentés dans ce manuscrit. La sonde est donc décrite dans cette section dans le cas de l'imagerie PA. Elle dispose au total de 256 éléments mais d'un unique connecteur ne permettant un

Figure 2.9 – Photographie de la sonde d'imagerie 2D L228-v (CMUT, Verasonics)

contrôle simultané de seulement 128 de ces éléments. De largeur 100 µm, ils sont organisés en barrette et disposent d'une focalisation en élévation de 15 mm. La taille des éléments en élévation est de 1.5 mm.

La *point spread function* (PSF) de cette sonde en imagerie PA a été estimée en illuminant un fil d'une épaisseur de 20 µm placé perpendiculairement au plan d'imagerie de la sonde (Fig. 2.4) . La longueur d'onde acoustique λ_{ac} de la sonde dans l'eau étant de 100 µm, ce fil peut être considéré comme un point source dans les directions axiales et latérales. La PSF est présentée dans la figure 2.10.d, ainsi que ses profils axiaux (axe des ordonnées, z) et latéraux (axe des abscisses, x). Elle est obtenue par l'opération de *delay-and sum* appliquée sur les signaux radiofréquences (RF) réels. La mesure de la largeur à mi-hauteur (FWHM) dans chacune de ces directions est une manière de définir la résolution spatiale. Cette résolution est de 121 µm dans l'axe x, et de 132 µm dans l'axe z. La mesure dans l'axe z est réalisée sur la PSF reconstruite à partir des signaux RF sous leur forme en quadrature pour ne pas tenir compte des oscillations liées aux réponses fréquentielles des transducteurs (seul le profil en z de cette PSF est représentée en figure 2.10.d, en orange).

Cette figure présente également les signaux RFs générés par ce fil ainsi que l'allure du signal mesuré par le transducteur central, et sa transformée de Fourier rapide (FFT) laquelle permet d'étudier la réponse fréquentielle de ce transducteur, et de déterminer sa bande passante.

Cette PSF sera notamment utilisée pour générer les données de simulation qui serviront à pré-entraîner le réseau de neurones dans le chapitre 5.

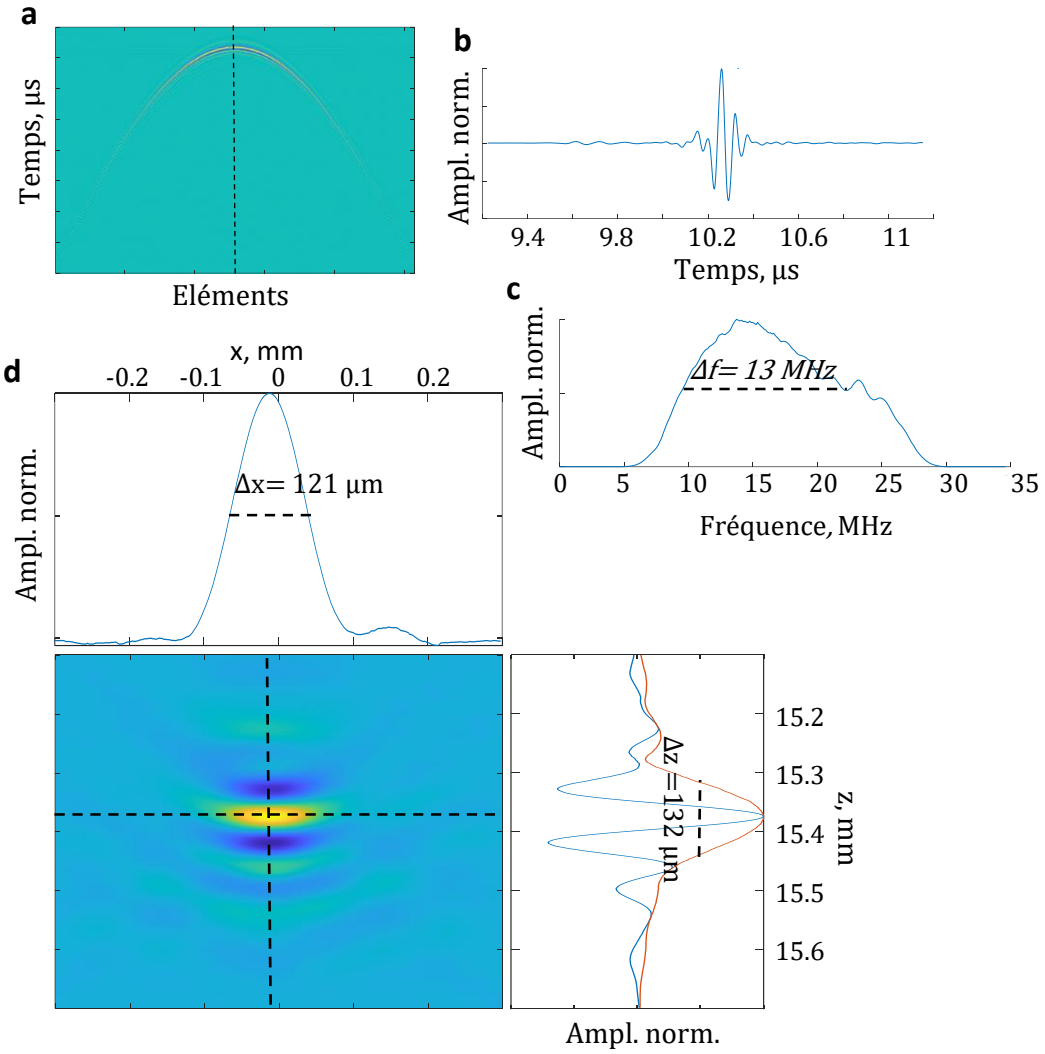

Figure 2.10 – Point spread function (PSF) expérimentale de la sonde d'imagerie 2D en imagerie photoacoustique.
(a) : signaux radiofréquences (RF). (b) : signal mesuré par le transducteur central, repéré par la ligne pointillée sur les signaux RF. (c) : Transformée de Fourier de ce signal. Le Δf représente la bande passante à -6 dB du transducteur en réception. (d) : PSF réelle et profils en X et Z (courbes bleues). La courbe en orange dans la direction z correspond au profil du module de la PSF reconstruite à partir des signaux sous leur forme en quadrature.

2.2.2 Sonde d'imagerie 3D

La sonde d'imagerie 3D utilisée au cours de cette thèse a été conçue et fabriquée au cours de la thèse de Sergey Vilov (2016-2019) réalisée dans notre laboratoire, mais n'avait jamais été mise en oeuvre. Les travaux présentés dans cette thèse constituent les premières expériences en imagerie 3D de l'équipe. Dans cette section, la géométrie de la sonde sera détaillée, puis ses caractéristiques seront étudiées à partir de sa *point spread function* d'abord obtenue en simulation dans le cadre de l'imagerie photoacoustique, puis expérimentalement à la fois en imagerie photoacoustique et en imagerie ultrasonore.

Géométrie et acquisition en imagerie ultrasonore

Cette sonde d'imagerie 3D, fabriquée par Imasonic (France, Voray-sur-l'ognon) dispose d'une fréquence centrale de 8 MHz, et est composée de 256 éléments. Chaque élément a été obtenu par un formage et une découpe sur la surface sphérique. Il en résulte des disques de diamètre 2 mm légèrement concaves et donc focalisés. Les éléments sont organisés sur la surface sphérique sous forme d'une spirale de Fermat. Une telle organisation produit un réseau pseudo-aléatoire qui limite la formation de lobes de réseau [79]. La zone d'imagerie, centrée sur le centre de la sphère de rayon 35 mm, est d'environ $8 \times 8 \times 8$ mm. Un trou de 8 mm de diamètre a été aménagé pour insérer une fibre optique et ainsi amener la lumière au niveau de la zone de l'image. Deux connecteurs permettent d'utiliser simultanément les 256 éléments. Une représentation schématique est présentée en figure 2.11. Une photographie est disponible en figure 2.5.

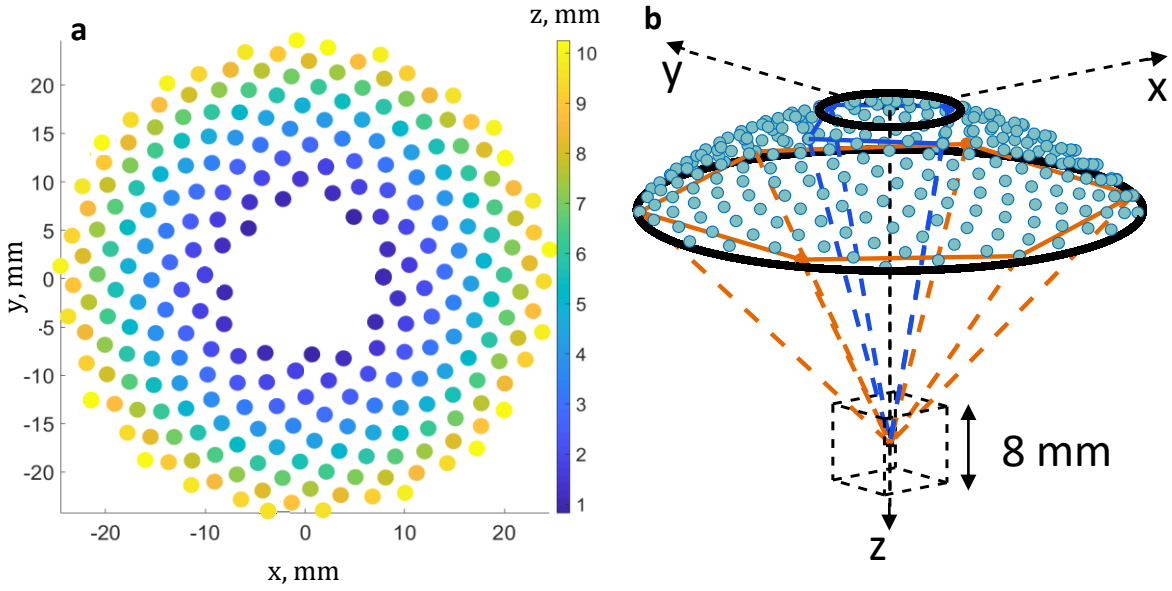

Figure 2.11 – Représentation schématique de la sonde d'imagerie 3D.
(a). Répartition spatiale des éléments avec leur position z encodée par l'échelle de couleur.
(b). Représentation 3D de l'organisation des éléments et du volume d'exploration. Tous les éléments sont dirigés vers le centre du volume d'exploration, comme représenté pour une dizaine d'entre eux par les traits oranges et bleus.

Lors de nos expériences d'imagerie ultrasonore, l'onde excitatrice est générée par un seul des 256 éléments. En effet, l'architecture de notre sonde d'imagerie 3D ne permet pas de facilement conçevoir des fronts d'onde à partir de plusieurs éléments à la manière des sondes linéaires utilisées en imagerie 2D, ou encore des sondes matricielles planes utilisables en imagerie 3D (chapitre 1, section 1.4.4). La géométrie des éléments qui composent notre sonde provoque la génération d'une onde divergente ce qui

permet de couvrir un volume conséquent. Par ailleurs, nous effectuons une approche de type *compounding* : une série d'acquisitions est réalisée en changeant l'élément émetteur. Le positionnement de nos éléments sur une calotte sphérique permet de générer naturellement des ondes excitant le volume avec une direction d'incidence différente, de manière analogue aux ondes planes angulées de l'imagerie ultrarapide "conventionnelle" (section 1.4.1). Dans ce manuscrit, le terme "sous-acquisition" US est utilisé pour décrire l'acquisition du jeu de signaux radiofréquences correspondant à une unique émission. Le terme "sous-image" US est utilisé pour décrire l'image US reconstruite par *delay-and-sum* à partir de ces données. Pour chaque sous-acquisition, une sous-image est donc reconstruite puis ces sous-images sont sommées entre elles pour former ce qu'on appelle dans ce manuscrit l'image US.

Le nombre d'émissions a été choisi de manière empirique et varie entre 10 et 13 en fonction de l'expérience. Le choix des éléments qui émettent est déterminé en fonction de leurs positions de manière à recouvrir au maximum la surface de la sonde pour maximiser le nombre de directions couvertes par les ondes excitatrices. Un exemple de choix d'éléments émetteurs est présenté dans la figure 2.17 lorsque 10 sous-acquisitions par image US sont réalisées. Le signal d'émission correspond à un cycle et demi de la fréquence centrale de la sonde (8 MHz), soit une bande passante de 75%, dans le but d'y maximiser l'énergie.

Point spread function **simulée en imagerie photoacoustique**

Les caractéristiques de la sonde en imagerie PA sont d'abord étudiées à partir de la PSF calculée en simulation à l'aide du logiciel FIELD II [65] [64]. Cet outil repose sur la dérivation discrète de l'équation d'onde en supposant le régime linéaire et une diffusion simple. Pour modéliser les éléments de la sonde, une surface concave représentant la surface totale de la sonde est définie (rayon focal = 35 mm, ouverture de 25 mm). cette surface est subdivisée en sous éléments de taille 0.05 mm^2. Pour un transducteur physique (de rayon 1 mm), le paquet de sous-éléments plans dont la position coïncide avec celle du transducteur physique (environ 124) est conservé. Ce paquet de sous-éléments peut alors être utilisé pour obtenir la réponse impulsionnelle $h_j(t)$ de ce transducteur. Pour cela on simule l'émission d'une onde acoustique, en utilisant une excitation analytique conçue de manière à ressembler à la forme expérimentale de l'onde PA. Le champ de pression résultant de cette émission en fonction du temps est enregistré au point focal de la sonde ($x = 0$ mm, $y = 0$ mm, $z = 35$ mm). D'après le principe de réciprocité de l'opérateur de propagation, le signal émis par ce transducteur et enregistré à la position r, est équivalent au signal émis par une source à la position r et enregistré par ce transducteur. Le signal simulé correspond donc à celui reçu par le transducteur pour une onde PA générée par un absorbeur situé au point focal.

Cette opération est répétée pour chaque transducteur, simulant une acquisition avec la sonde d'imagerie 3D pour une source située au point focal. Ces signaux sont représentés en figure 2.12, a et b. L'image PA de cette source, qui correspond à la PSF de la sonde, peut ensuite être reconstruite par *delay-and-sum*. Comme pour la sonde d'imagerie 2D, la PSF est ici reconstruite à partir des signaux réels. Les courbes bleues correspondent aux profils de cette PSF. Le courbe en orange représente le profil en z du module de la PSF reconstruite à partir des signaux sous leur forme en quadrature.

Contrairement à la sonde linéaire, les signaux RFs sont naturellement en phase : en effet, si le point source est placé exactement au point focal, les distances qui le séparent de tous les transducteurs sont identiques. Comme dans le cas de l'imagerie 2D, les résolutions et la bande passante, ici théoriques, peuvent être mesurées. Cette PSF sera utilisée pour réaliser différents jeux de données de simulation.

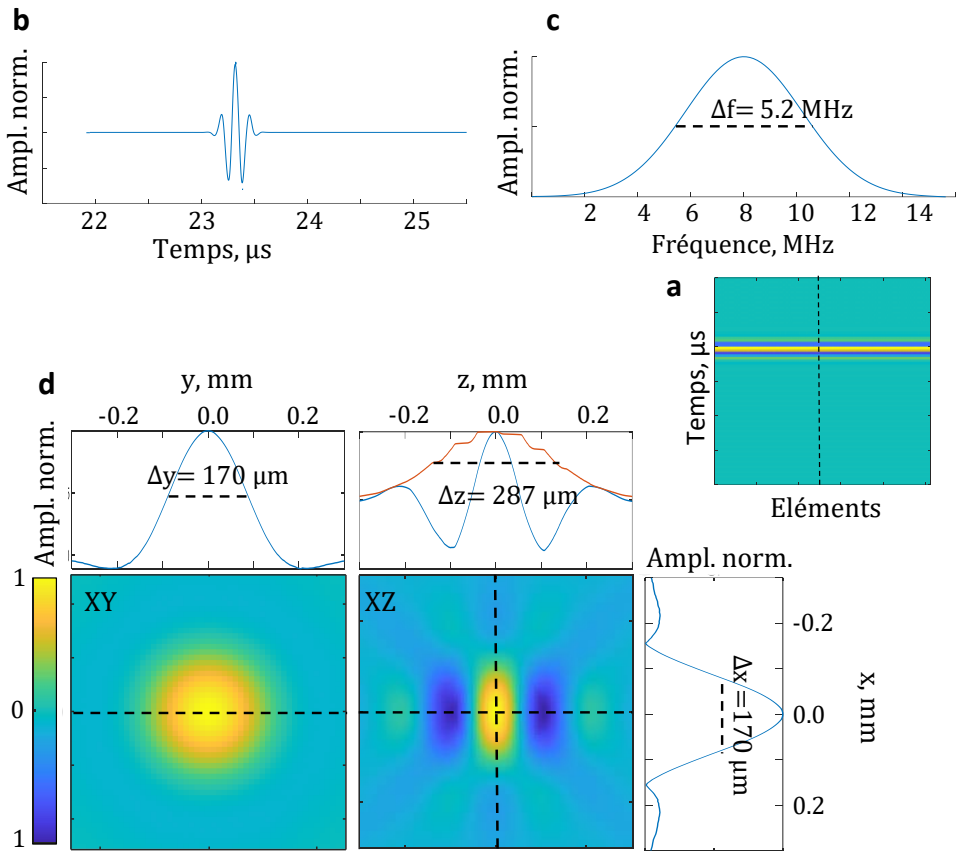

Figure 2.12 – Point spread function simulée de la sonde d'imagerie 3D en imagerie photoacoustique. (a) : signaux radiofréquences (RF). (b) : signal mesuré par le transducteur central, repéré par la ligne pointillée sur les signaux RF. (c) : Transformée de Fourier de ce signal. Le Δf représente la bande passante à -6 dB du transducteur en réception. (d) : Coupe XY et XZ au centre du volume et profils de la PSF réelle en X et Z (courbes bleues). La courbe en orange dans la direction z correspond au profil du module de la PSF reconstruite à partir des signaux sous leur forme en quadrature.

Point spread function expérimentale en imagerie photoacoustique

La PSF expérimentale PA a également été obtenue. L'approximation du fil comme point source, utilisée pour l'imagerie 2D ne fonctionnant pas en 3D, c'est une microsphère noire de polyéthylène de diamètre 50 µm (Cospheric) qui joue le rôle de point source. Pour l'isoler, une suspension très diluée de ces microsphères est versée dans une boite de pétri et observée à l'aide d'un macroscope. Une micropipette dont la pointe a été affinée par thermoformage à l'aide d'un *pipet-puller* est fixée sur un système trois axes et amenée manuellement le plus près possible d'une des sphères. Cette sphère est alors aspirée à l'aide d'une poire fixée à la pipette et vient se fixer sur l'embout. Sans relâcher la dépression, la bille est extraite de la solution puis un cube d'agarose creusé en son centre est placé sous la pipette. La pression est alors relâchée pour déposer la bille dans le puit. Ce puit est ensuite refermé en le remplissant de la solution d'agarose. Des photographies obtenues avec un microscope sont présentées en figure 2.13, où l'on peut voir la sphère fixée à la pipette, puis déposée à la surface du gel d'agarose.

Une acquisition PA de cette bille dans le gel d'agarose est réalisée selon le dispositif expérimental présenté en figure 2.6. L'image résultante correspond à la PSF expérimentale de la sonde. Cette PSF, présentée en figure 2.14 est construite à partir des signaux réels, les courbes bleues correspondent aux profils de cette PSF, et la courbe orange correspond au profil dans la direction z du module de la PSF calculée à partir des signaux sous leur forme en quadrature.

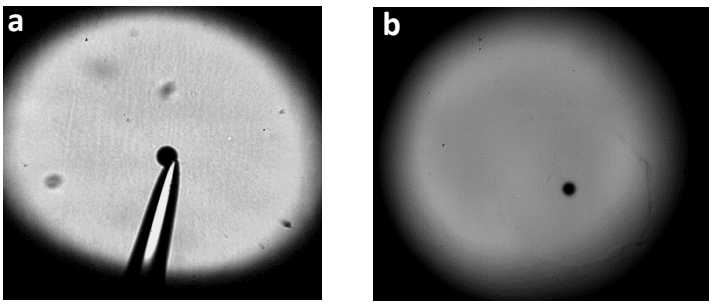

Figure 2.13 – Isolation d'une microsphère de 50 µm.
(a) : la bille est coincée sur l'embout de la pipette, grâce à la différence de pression. (b) : La bille a été relâchée sur le gel d'agarose.

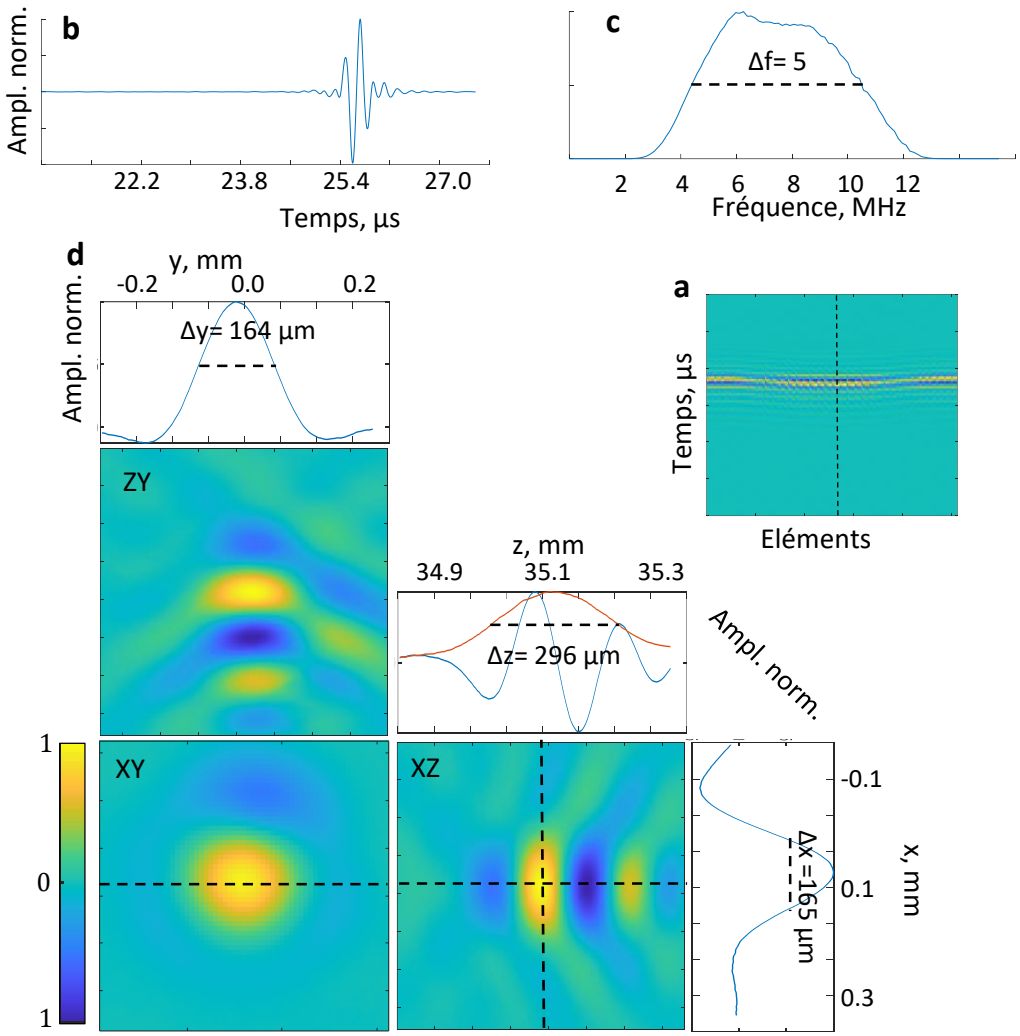

Figure 2.14 – Point spread function (PSF) expérimentale de la sonde d'imagerie 3D en imagerie photoacoustique.
(a) : signaux radiofréquences (RF). (b) : signal mesuré par le transducteur central, repéré par la ligne pointillée sur les signaux RF. (c) : Transformée de Fourier de ce signal. Le Δf représente la bande passante à -6dB du transducteur en réception. (d) : PSF reconstruite à partir des signaux réels, coupe XY, YZ et XZ au centre de la tache de diffraction et profils selon les trois directions. La courbe en orange dans la direction z correspond au profil du module de la PSF reconstruite à partir des signaux sous leur forme en quadrature.

La microsphère n'est pas parfaitement située au point focal (décalage d'environ 80 µm dans la direction x et z), expliquant le léger déphasage des signaux RF. Les résolutions mesurées sont similaires à celles obtenues sur les données de simulation.

Point spread function expérimentale en imagerie ultrasonore

En imagerie ultrasonore, la PSF dépend des conditions d'excitation. Lors de nos expériences d'imagerie, nous effectuons une approche de type *compound* : une série d'acquisition est réalisée en changeant l'élément émetteur, pour chaque acquisition des "sous-images" sont reconstruites par *delay-and-sum* puis ces images sommées entre elles pour former l'image US finale. La PSF US expérimentale est obtenue selon ce principe. On réalise une série d'acquisition de la microsphère précédemment décrite en utilisant en émission les mêmes transducteurs que ceux utilisés lors des expériences d'imagerie. Les "sous-PSF" sont reconstruites pour chaque émission puis ces images sont moyennées entre elles pour former la PSF US finale visible en figure 2.15. Comme pour la PSF de l'imagerie PA, la résolution de la sonde peut être calculée dans les trois directions.

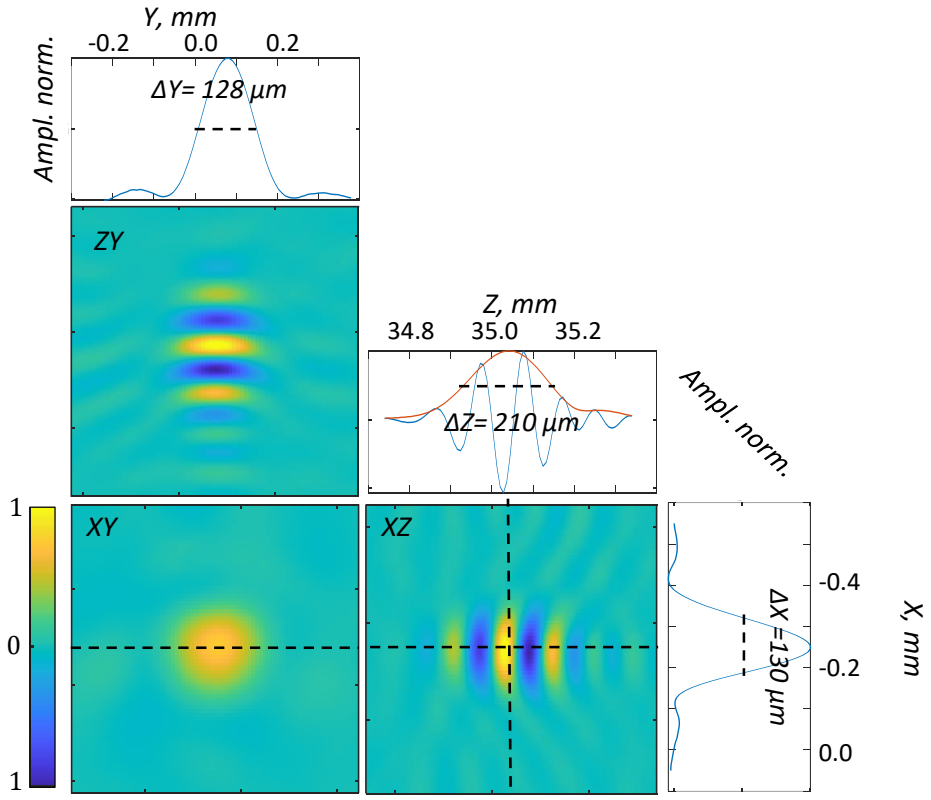

Figure 2.15 – Point spread function expérimentale de la sonde d'imagerie 3D en imagerie ultrasonore, reconstruite à partir d'une série de jeux de signaux radiofréquences réels correspondant à des sous-acquisitions pour lesquelles l'élément émetteur change. Coupe XY, YZ et XZ au centre de la tache de diffraction et profils selon les trois directions. La courbe en orange dans la direction z correspond au profil du module de la PSF reconstruite à partir des signaux sous leur forme en quadrature.

Champs de vue en imagerie photoacoustique 3D

A l'aide du logiciel Field II, on simule l'acquisition PA par l'ensemble des éléments des signaux émis par une série de points sources couvrant un volume de taille 10 mm × 10 mm × 10 mm autour du point focal par pas de 0.2 mm, soit 70 × 70 × 70 points. Pour chaque point source, on répète les étapes décrites pour simuler la PSF. On obtient donc, pour chaque transducteur, un jeu de 70 × 70 × 70 signaux temporels.

La sensibilité spatiale d'un transducteur peut être évaluée en sélectionnant les maxima des enveloppes de ces signaux temporels. Cette carte de maxima est appelée champ de vue pour cet élément. La sensibilité spatiale de la sonde d'imagerie est obtenue en sommant entre elles les cartes de maxima de l'ensemble des transducteurs. Les champs de vue de deux transducteurs sont représentés en 3D et selon une coupe XZ (y=0) en figure 2.16, (a) et (b). Le champ de vue 3D de la sonde et des coupes XZ (y=0 mm) et XY (z = 35 mm) sont représentés en (c).

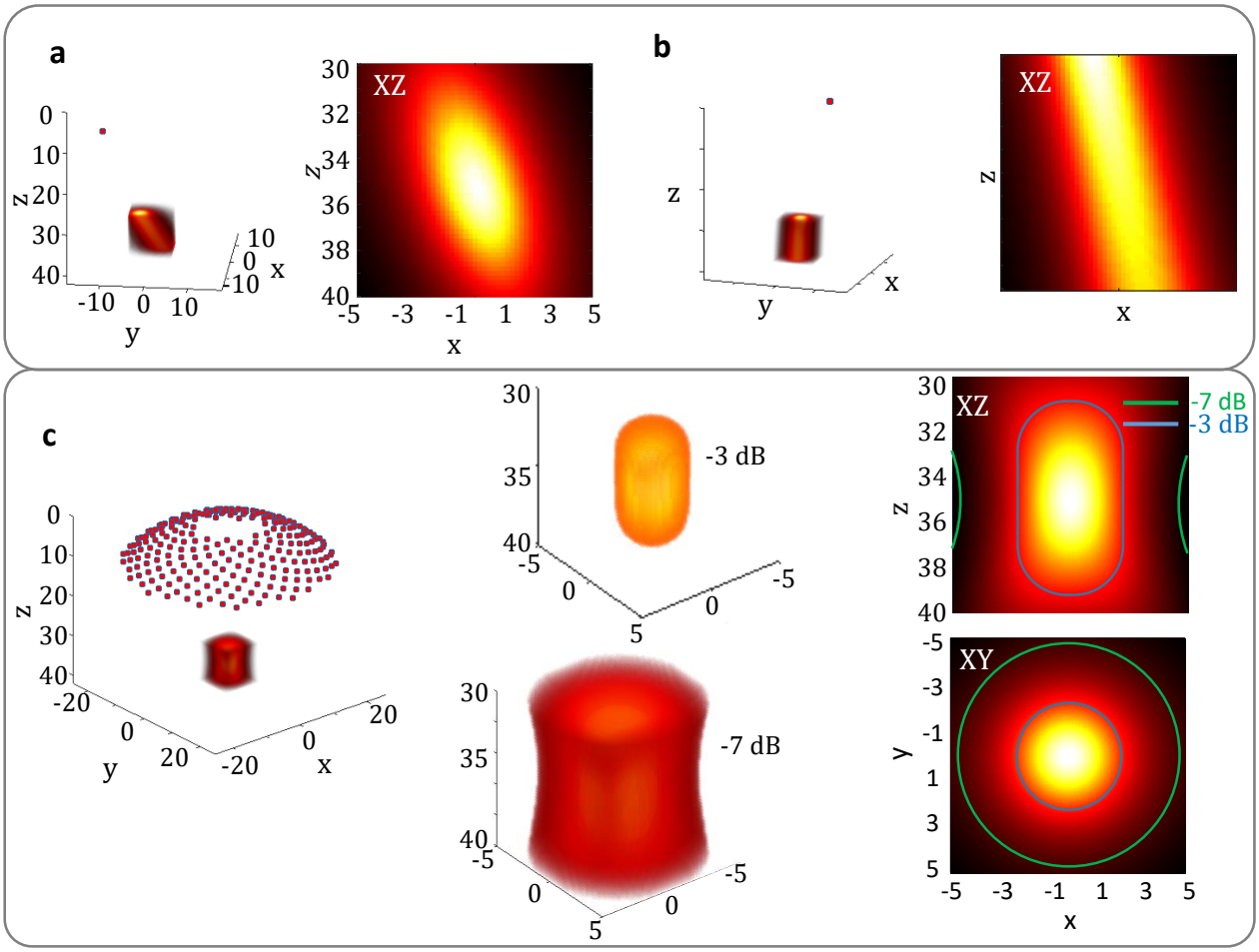

Figure 2.16 – Évaluation du champ de vue de la sonde d'imagerie 3D en imagerie photoacoustique, tous les axes en mm.
(a,b) : champ de vue et profils XZ (y=0 mm) correspondants de deux transducteurs représentés par la pastille rouge. (c) : Champ de vue de la sonde obtenu en sommant les champs de vue des 256 éléments, non seuillé, seuillé à -3 et - 7 dB, et profil XY (z=35 mm) et XZ (y = 0 mm) correspondant.

Les volumes a -3dB et -7dB estimés à partir des données de la figure 2.16, sont respectivement d'environ $4.8 \times 4.8 \times 8$ mm^3 et $9.1 \times 9.1 \times 10$ mm^3 .

Champs de vue en imagerie ultrasonore 3D

Toujours à l'aide du logiciel Field II, on simule l'acquisition US des signaux rétrodiffusés par la série de point définie dans la section précédente. On réalise cette simulation pour différents transducteurs en émission.

Pour obtenir ces signaux rétrodiffusés, on considère que le signal $s_{i,j}(t)$, le signal en provenance de l'élément i rétrodiffusé par la source r et enregistré par l'élément j, peut être défini comme la convolution entre le signal reçu en r pour une émission de l'élément i, et le signal reçu en ce même point pour une émission de l'élément j. On a donc $s_{i,j}(t) = s_i(t) * s_j(t)$. On simule ces signaux pour toutes les positions r. Les maxima d'enveloppe de ces signaux, permettent de définir le champ de vue du couple d'éléments (i, j). En calculant ces champs de vue pour l'ensemble des N_{el} éléments en réception (indice j), on obtient le champ de vue de la sonde pour une émission avec le i^{me} élément.

Ces champs de vue sont simulés pour les 10 éléments utilisés en émission dans nos expériences d'imagerie (les éléments sont représentés en bleu en figure 2.16). On peut finalement déterminer la sensibilité spatiale de notre sonde pour cette séquence d'imagerie en sommant entre eux ces 10 champs.

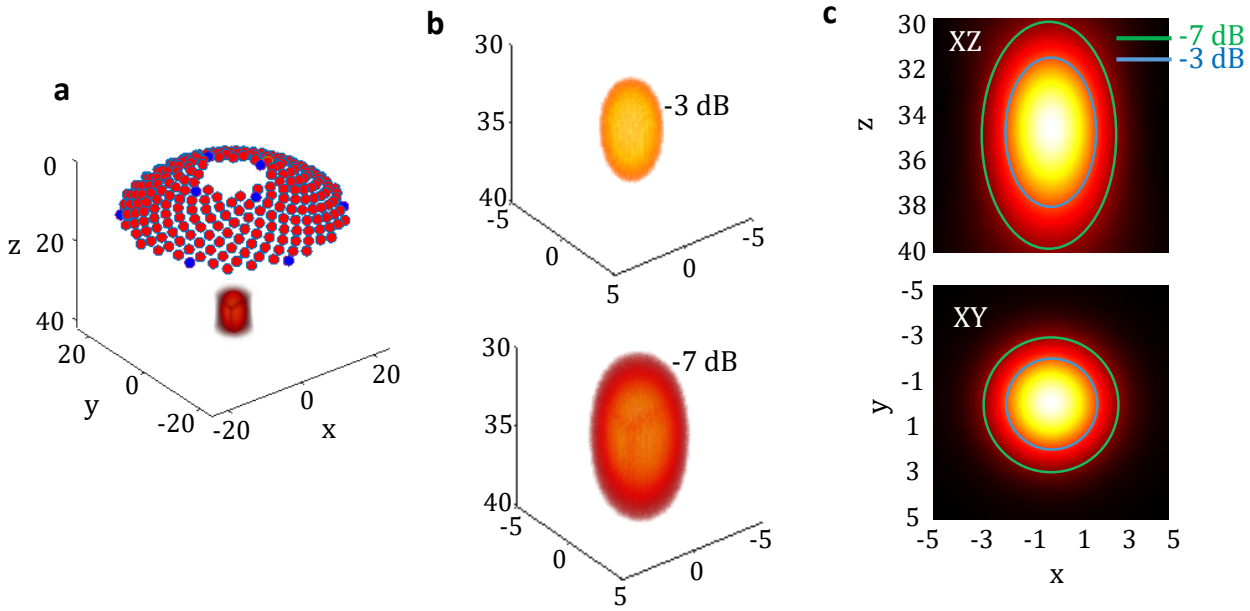

Figure 2.17 – Évaluation du champ de vue de la sonde d'imagerie 3D en imagerie ultrasonore, tous les axes en mm.
(a) : Représentation 3D du champ de vue de la sonde pour une séquence de 10 émissions par des transducteurs différents, représentés par les pastilles bleues. Ce champ de vue correspond à la somme des champs de vue pour ces 10 émissions. (b) : Ce champ de vue est seuillé à -3 et - 7 dB. (c) : profil XY et XZ correspondant au champ de vue en (a).

Les volumes à -3dB et -7dB sont estimés respectivement à environ $4 \times 4 \times 6$ mm^3 et $6 \times 6 \times 8.8$ mm^3. On observe donc que le champ de vue est plus restreint en imagerie US, en raison de la focalisation en émission.

2.3 Expériences *in vivo* : méthodes

2.3.1 Séquence d'acquisition et interfaçage

Lors des expériences *in vivo*, sur l'embryon ou la souris, les données PA et US sont acquises simultanément en intercalant les acquisitions US entre deux acquisitions PA. En fonction de la fréquence d'acquisition souhaitée en imagerie ultrasonore (on appelle cette fréquence PRF_{US}), plusieurs images US peuvent être obtenues entre deux tirs du laser. La PRF_{PA} est fixée par celle du laser, de 100 Hz. Chaque image US est formée à partir de la combinaison de nTX sous-images reconstruites à partir de jeux de signaux provenant de nTX tirs successifs, chacun réalisé avec un élément différent. La durée nécessaire à l'obtention de ces signaux ($c \times d_{max} \times 2 \times nTX$) avec d_{max} la distance maximale entre un élément et un point du volume, déterminera fondamentalement la PRF_{US} maximale. L'acquisition PA peut être multispectrale : M_λ longueurs d'onde sont acquises à la suite. Cette séquence est répétée autant que nécessaire pour obtenir le nombre d'images N souhaitées pour chaque longueur d'onde. Un schéma de la séquence est présenté en figure 2.18 pour une PRF_{US} de 200Hz, 10 tirs US par images et 10 longueurs d'onde.

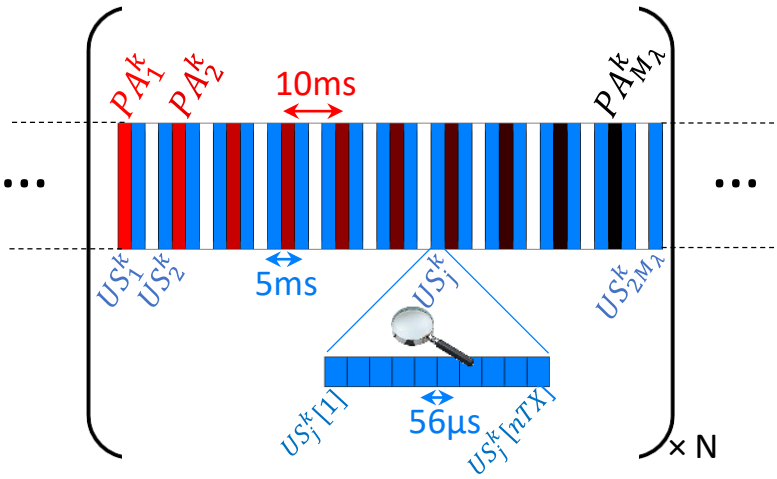

Figure 2.18 – Séquence d'imagerie ultrasonore (US) et photoacoustique (PA) combinés avec $M_\lambda = 10$ longueurs d'onde différentes en imagerie PA, une fréquence d'acquisition en imagerie US de 200 Hz et nTX =10 soit 10 sous-acquisitions en changeant l'élément émetteur pour former une image US.

L'acquisition et le transfert de ces données vers le PC sont réalisés de manière asynchrone par l'échographe. Le *buffer* est un espace dans lequel sont stockées les données radiofréquences avant d'être transmises au PC. A l'origine, un unique *buffer* est utilisé par l'échographe : son temps de transfert vers le PC doit donc être inférieur à la durée espaçant deux acquisitions PA ou US ce qui limite considérablement sa taille. Un système à deux *buffers* a donc été mis en place. Le premier est rempli par les données en train d'être acquises pendant que le second est transféré vers le PC. Le temps de transfert du *buffer* doit alors être inférieur au temps nécéssaire pour le remplir, ce qui limitera en pratique la PRF_{US} et/ou le nombre de tir par image nTX. Dans nos expériences, un *buffer* est typiquement constitué de 50 acquisitions PA et 100 acquisitions US (soit au total nTX$\times$100 "sous-acquisition"), il est donc rempli en environ 500 ms.

En réalité, au début de chaque nouveau *buffer* (soit 2 fois par seconde pour la taille de *buffer* décrite dans cette section), les sous-acquisitions nécessaires à la formation d'une image US dite 'haute résolution' (USHR) sont réalisées. Dans une séquence d'imagerie typique, nous réalisons 20 sous-acquisitions (avec 20 éléments différents en émission) pour former l'image USHR. C'est cette image, de meilleure

qualité que celle obtenue à partir des 10 sous-acquisitions, qui est affichée par défaut lors de la navigation temps réel. La composition d'un *buffer* typique de nos expériences est présenté en figure 2.19.

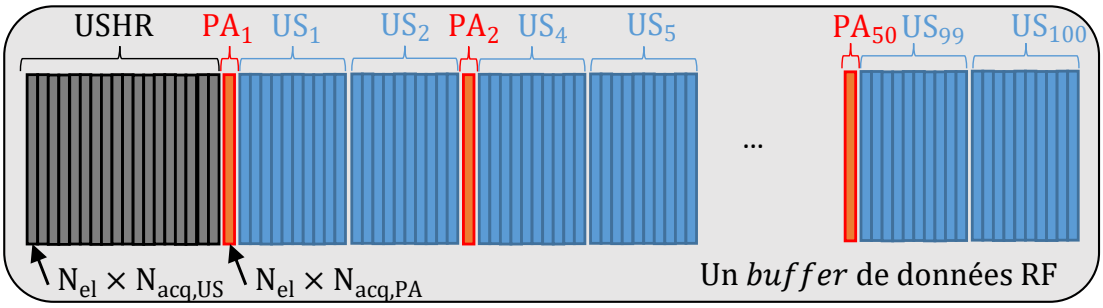

Figure 2.19 – Représentation d'un buffer de données radiofréquences (RF) typique des expériences d'imagerie photoacoustique (PA) et ultrasonore (US) combinées. Un buffer est constitué des données nécessaires à la reconstruction d'une image US "haute résolution" (USHR, 20 sous-acquisitions) utilisée pour la navigation en temps réel, 50 images PA et 100 images US (10 sous-acquisitions). N_{acq} correspond au nombre de pas temporels enregistrés par signal, qui dépend de la fréquence d'échantillonnage et de la profondeur d'imagerie.

L'interface du logiciel d'acquisition de l'échographe a également été adaptée pour facilement naviguer dans l'échantillon. Différentes fonctions y ont été ajoutées :

- affichage de la projection des intensités maximales des images 3D US et PA dans les trois dimensions.
- choix et affichage de coupes de ces images 3D dans les trois dimensions.
- Choix du t_0 pour ajuster la reconstruction, modifiable indépendamment pour l'image US et l'image PA.
- Modification du *Time Gain Compensation* indépendamment pour les données US et PA.
- Affichage des données radiofréquences.
- Passage du mode sauvegarde au mode temps réel sans redémarrage du code d'acquisition.

Deux modes de fonctionnement ont été définis : un mode navigation pour lequel certaines images 3D sont reconstruites, environ deux par seconde ce qui permet un rendu quasi temps réel, et un mode sauvegarde pour lequel les reconstructions et autres affichages sont totalement désactivés. La figure 2.20 permet de visualiser cette interface et les images visibles en temps réel lors de la navigation au cours d'une expérience souris. La boite crânienne est bien visible sur les images ultrasonores (ligne du haut), et des vaisseaux sont observables sur les images photoacoustiques (ligne du bas).

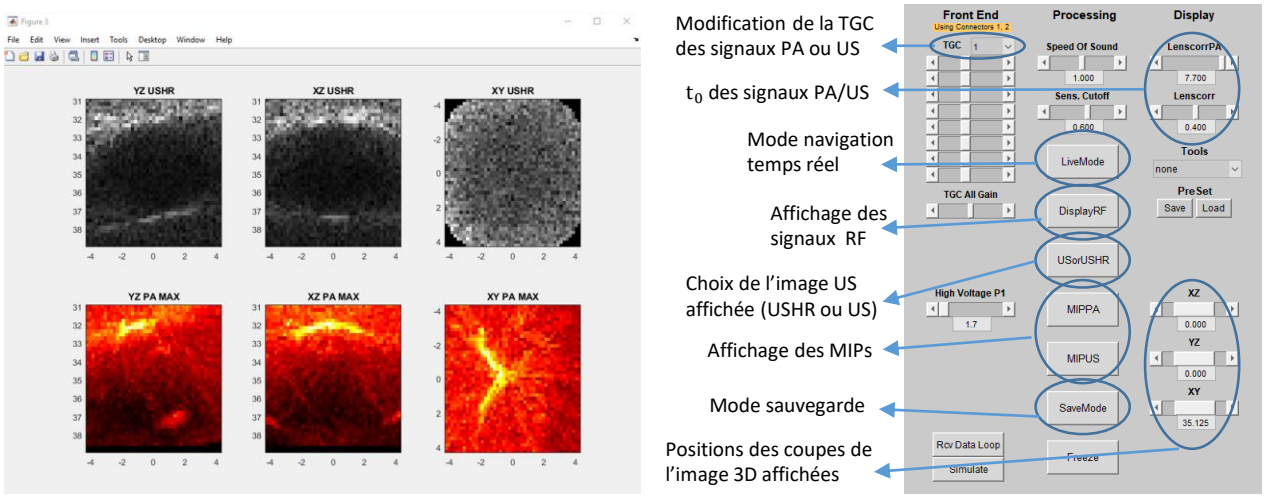

Figure 2.20 – Interface du logiciel de contrôle de l'échographe permettant de modifier différents paramètres en temps réel pour faciliter la navigation, et capture d'écran des images temps réel lors d'une expérience souris (coupes dans les trois plans YZ, XZ et XY en imagerie ultrasonore, US, intensités maximales de projection en imagerie photoacoustique, PA). TGC : Time Gain Compensation. RF : radiofréquence. USHR : imageries ultrasonore haute résolution, chaque image est reconstruite à partir de 20 sous acquisitions.

2.3.2 Préparation des embryons de poulet

Au cours de cette thèse, de nombreuses expériences ont été réalisées sur des embryons de poulet, et plus particulièrement sur leur membrane chorioallantoique, composés de petits vaisseaux sanguins (fig. 2.21) dont le rôle est de capturer de l'oxygène à travers la coquille puis de le transmettre à l'embryon, au centre de l'oeuf.

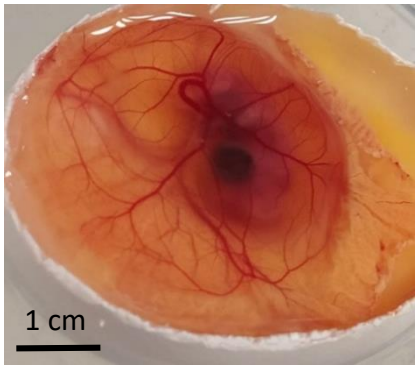

Figure 2.21 – Embryon de poulet à 9 jours de développement, quelques minutes avant le début de son imagerie.

Les vaisseaux sont donc étendus sur une large surface et présentent des orientations tridimensionnelles, rendant l'imagerie d'une telle structure très intéressante dans le but de tester des méthodes d'imagerie 3D, notamment quand elles requièrent du sang et de l'écoulement. Des oeufs de poule fécondés classés sans pathogènes spécifiques (SPF) sont récupérés dans un couvoir spécialisé (Cerveloup) situé à quelques km du laboratoire. Ils sont placés dans un incubateur à 37°C avec un taux d'humidité de 65%. Au bout de 6 à 10 jours de développement, l'embryon est imagé. L'oeuf est tourné sur lui-même à plusieurs reprises pour que la membrane chorioallantoique de l'embryon se décroche de la coquille, avant d'insérer une aiguille pour prélever à l'aide d'une seringue environ 1.5 ml de "blanc d'oeuf". Une fenêtre est délicatement découpée à l'aide de ciseaux à pointes courbées pour accéder à l'embryon.

Une solution tampon (buffer phosphate salin, PBS) est ajoutée dans l'oeuf maintenant ouvert pour faciliter le contact acoustique avec la membrane en latex du cône de couplage de la sonde d'imagerie 3D. L'oeuf est maintenu sur un film alimentaire à la surface d'un bol rempli d'eau thermostaté à 37°C à l'aide d'une plaque chauffante (fig. 2.22).

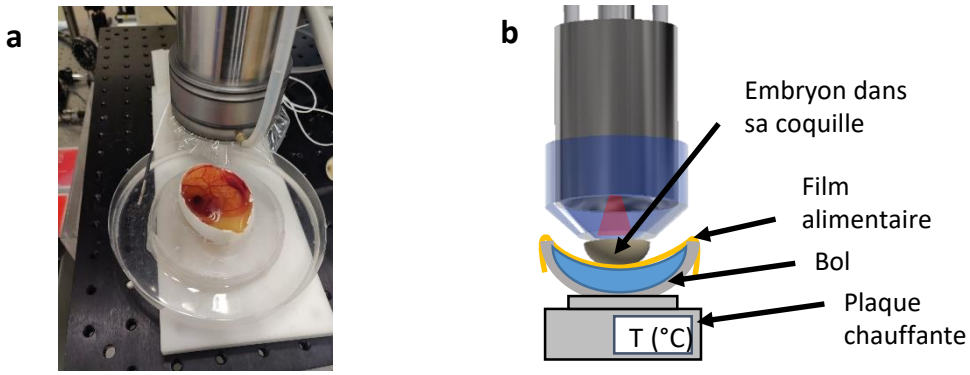

Figure 2.22 – Imagerie de l'embryon de poulet. (a) : Photographie d'une expérience. (b) : Schéma correspondant.

2.3.3 Organisation d'une expérience souris

Dans le chapitre 3, des premières expériences sont présentées sur des souris de type NMRI nude. Ces expériences sont réalisées en collaboration avec l'*Institute for Advanced Biosciences* (IAB) à Grenoble qui dispose d'une animalerie (agréement C38 516 10001), sous le nom de projet APAFIS19706-2019030114369271 qui a reçu l'approbation des autorités compétentes. Toutes les personnes impliquées dans ces expériences ont été formées à l'expérimentation animale.

La souris est amenée de l'animalerie au laboratoire puis placée dans une chambre à induction sous flux d'isofluorane, un gaz anesthésiant. Une fois anesthésiée, la souris est sortie de la chambre et déposée sur un plateau chauffant, son museau inséré dans un tube relié au gaz, et un thermomètre rectal est utilisé pour suivre sa température. Un casque stéréotaxique est utilisé pour limiter le mouvement lorsque le cerveau est imagé. Du gel d'échographie est déposé sur la zone à imager, puis le contact avec la sonde est réalisé. Dans le cas de l'imagerie de la queue, celle-ci est posée sur un pad de gel (Aquaflex) pour limiter les signaux parasites.

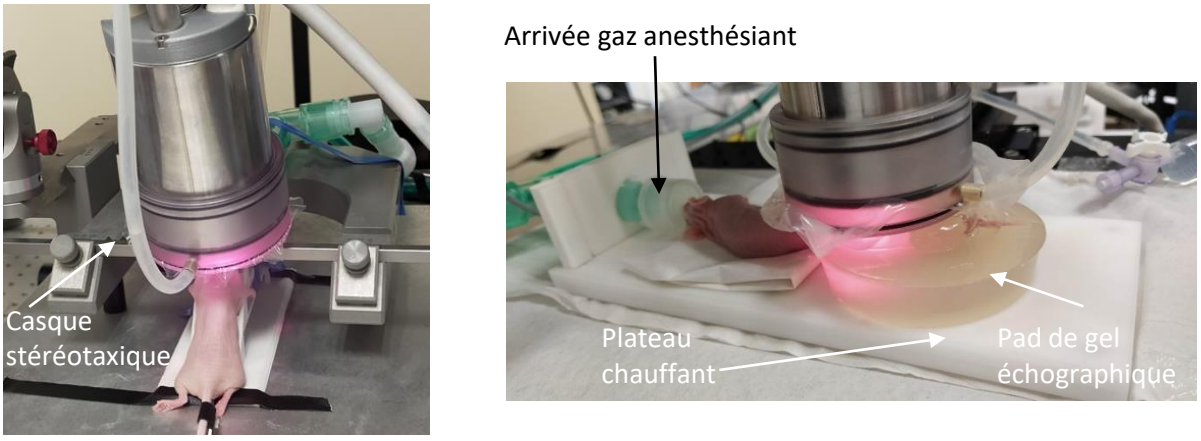

Figure 2.23 – Photographie de deux expériences souris, imagerie du cerveau à gauche et de la queue à droite.

2.4 Conclusion

Nous avons détaillé dans ce chapitre les principaux composants des dispositifs expérimentaux utilisés dans cette thèse (laser, échographe, sonde d'imagerie ...) , en appuyant sur l'importance de la synchronisation entre l'illumination et le début de l'acquisition des ondes de pression en imagerie photoacoustique. Nous avons tout particulièrement insisté sur les aspect d'imagerie 3D et les expériences *in vivo*, et abordé l'aspect quantitatif des mesures photoacoustiques, trois thématiques nouvelles pour notre équipe. Il a notamment été réalisé au cours de cette thèse les éléments suivants, décrits en détail dans ce chapitre :

1. la mise en place d'un dispositif expérimental d'imagerie 3D permettant notamment de mesurer la fluence du laser.
2. la caractérisation d'une sonde parcimonieuse hémisphérique d'imagerie 3D dans le cadre de l'imagerie ultrasonore et photoacoustique (résolution spatiale au travers des *point spread function*, et champ de vue).
3. la conception d'une séquence d'imagerie originale permettant l'imagerie ultra-rapide ultrasonore et photoacoustique en simultanée, et la navigation dans un échantillon avec un affichage quasi temps-réel du volume d'exploration.
4. des expériences *in vivo* sur l'embryon de poulet et la souris.

3

Imagerie photoacoustique de fluctuation

Sommaire

3.1 Contexte **66**

3.2 Théorie : formation de l'image photoacoustique de fluctuation **68**

3.2.1 Définition du phénomène imagé 68

3.2.2 Définition de l'image PA 70

3.2.3 Définition de l'image moyenne 70

3.2.4 Définition de l'image de fluctuation 71

3.2.5 Conséquences sur les artefacts de visibilité 73

3.2.6 L'image de fluctuation en conditions idéales de simulation 73

3.3 Imagerie photoacoustique de fluctuation en conditions expérimentales *in vitro* . . **75**

3.3.1 Nécessité d'un filtrage spatio-temporel en conditions expérimentales 75

3.3.2 Imagerie phototoacoustique de fluctuation *in vitro* 2D 76

3.3.3 Imagerie phototoacoustique de fluctuation *in vitro* 3D 77

3.4 Imagerie photoacoustique de fluctuation et Power Doppler 3D *in vivo* **79**

3.4.1 Imagerie photoacoustique de fluctuation 3D de l'embryon de poulet 79

3.4.2 Imagerie Power Doppler ultrasonore 3D de l'embryon de poulet 80

3.4.3 Imagerie photoacoustique de fluctuation et imagerie Power Doppler ultra-rapide simultanées de l'embryon de poulet 81

3.4.4 Imagerie photoacoustique de fluctuation d'une souris : résultats préliminaires . 83

3.5 Impact du bruit électronique sur la qualité de l'image de fluctuation **85**

3.5.1 Définition de l'image de fluctuation en présence de bruit 85

3.5.2 Conséquences de la présence de bruit sur l'image reconstruite en simulation . . 85

3.5.3 Conséquences de la présence de bruit sur l'image reconstruite lors des expériences *in vitro* et *in vivo* 87

3.6 Impact d'une variation tir à tir de la fluence du laser **89**

3.6.1 Définition de l'image de fluctuation en présence d'une variation tir à tir de la fluence du laser 89

3.6.2 Amplitude de la fluctuation liée au mouvement des globules rouges vis à vis du signal non fluctuant 90

3.6.3 Mécanismes du filtrage spatio-temporel par décomposition en valeurs singulières 91

3.6.4 Influence du choix des bornes de la décomposition en valeurs singulières sur la qualité de l'image 93

3.7 Conclusion **95**

3.1 Contexte

L'imagerie de fluctuation, une démarche commune à plusieurs modalités d'imagerie

L'imagerie de fluctuation exploite une fluctuation temporelle d'un signal : l'image finale est obtenue par une analyse statistique des fluctuations dans une série d'images. En imagerie photoacoustique, l'imagerie de fluctuation a été introduite par notre équipe. Notre approche est similaire dans la démarche à certaines méthodes bien connues en imagerie optique telles que l'imagerie de fluctuation optique super-résolue (en anglais *super-resolution optical fluctuation imaging*, SOFI [30]), et en imagerie ultrasonore avec les méthodes Power Doppler [110]. L'approche SOFI exploite les fluctuations aléatoires et indépendantes de molécules fluorescentes. Une série d'images est acquise avec un microscope conventionnel, leur résolution est donc limitée par la diffraction. Le calcul en chaque pixel des cumulants d'ordre supérieur génère la super-résolution. En effet, si l'image conventionnelle peut être considérée comme le résultat de la convolution de l'objet avec la *point spread function* (PSF) d'ordre 1 du système d'imagerie, l'image de cumulant d'ordre n peut être considérée comme le résultat de la convolution de cet objet avec la PSF d'ordre n. Cette PSF d'ordre n'est plus fine que la PSF d'ordre 1, ce qui génère la super-résolution. L'imagerie Power Doppler ultrasonore exploite les fluctuations de l'amplitude de l'onde rétrodiffusée par les globules rouges en mouvement dans les vaisseaux sanguins. Une série d'images est acquise et filtrée temporellement pour discriminer les signaux provenant de structures en mouvement (les globules rouges) des signaux provenant de structures statiques (les tissus environnants). L'image finale, résultant de l'intégration temporelle de la série d'images filtrées, dispose d'un contraste entre les vaisseaux sanguins et les tissus environnants largement supérieur à l'image ultrasonore conventionnelle. Une image Power Doppler provenant de la littérature est visible dans le chapitre 1 (figure 1.1.e).

En imagerie photoacoustique, l'imagerie de fluctuation a été introduite pour l'imagerie super-résolue [18] [58] [89] et pour la correction des artefacts de visibilité [40] [132]. Ces artefacts, présentés dans le premier chapitre, sont induits par la forte cohérence des ondes de pression générées lors de l'illumination d'un objet dense et absorbant, lesquelles peuvent interférer destructivement entre elles. Lorsque le signal PA fluctue d'une acquisition à une autre, l'homogénéité de l'objet est brisée : en calculant la variance d'une telle série d'images, les structures initialement invisibles apparaissent. Le paragraphe suivant présente les deux catégories de fluctuation utilisées en imagerie photoacoustique de fluctuation.

Sources de fluctuation en imagerie photoacoustique de fluctuation

La fluctuation du signal photoacoustique peut provenir d'une fluctuation de l'illumination. A partir d'une illumination par motif de *speckle* fluctuant à une échelle subdiffraction acoustique, des images photoacoustiques super-résolues [58] [89] ont été obtenues et les artefacts de visibilité corrigés *in vitro* [40] (exemple de résultats en figure 3.1, partie de gauche a-c). Dans l'objectif d'utiliser l'imagerie photoacoustique de fluctuation *in vivo*, l'illumination par motif de speckle aléatoires est limitée par la dimension des grains de speckle dans les tissus, de l'ordre de la longueur d'onde optique (1 µm), [100] soit plusieurs ordres de grandeur en dessous de la longueur d'onde acoustique (200 µm pour une sonde à 8 MHz) . La fluctuation du signal photoacoustique dans la zone résolue (de la taille de la PSF acoustique) est alors extrêmement faible, car résultant de la sommation d'un grand nombre de contributions aléatoires, et donc difficilement détectable dans des conditions normales de *signal to noise ratio* (SNR).

La fluctuation du signal photoacoustique peut également être provoquée par une fluctuation de la distribution d'absorbeurs dans le milieu imagé. En imagerie des vaisseaux sanguins, les absorbeurs sont en suspension dans le sang et leur distribution change continuellement en raison du flux généré par le battement cardiaque. Ces absorbeurs peuvent être exogènes ou endogènes. L'injection de particules absorbantes dans les vaisseaux sanguins a ainsi été proposée pour améliorer la visibilité [25]. De telles injections résultent en une distribution parcimonieuse des particules et le temps d'acquisition nécessaire

pour obtenir une représentation complète de l'objet imagé peut être long. De plus, l'injection de particules constitue une opération invasive. La fluctuation du signal photoacoustique généré par les globules rouges, absorbeurs optiques naturellement présents dans les vaisseaux sanguins, peut également être considérée. Cette approche présente l'avantage, en plus d'être non invasive, de conserver un signal de nature endogène. Son intérêt a d'abord été démontré in vitro pour la super-résolution PA [18] (exemple de résultats en figure 3.1, partie droite d-g). De premiers travaux préliminaires ont également été réalisés lors de la thèse de Sergey Vilov [131], qui a appliqué cette approche *in silico* et *in vitro* avec une sonde linéaire pour corriger les artefacts de visibilité. Les résultats obtenus au cours de ma thèse et présentés dans ce chapitre font suite à ces travaux. Ils consistent notamment au passage de l'imagerie de fluctuation 2D à l'imagerie de fluctuation 3D avec une sonde d'imagerie parcimonieuse et à son implémentation en conditions *in vivo*. Une théorie permettant d'expliciter les mécanismes d'amélioration de la visibilité et de prendre en compte le bruit électronique et la fluctuation de la fluence du laser entre chaque acquisition a également été développée. Enfin, l'imagerie Power Doppler 3D *in vivo* est réalisée simultanément à l'imagerie photoacoustique de fluctuation.

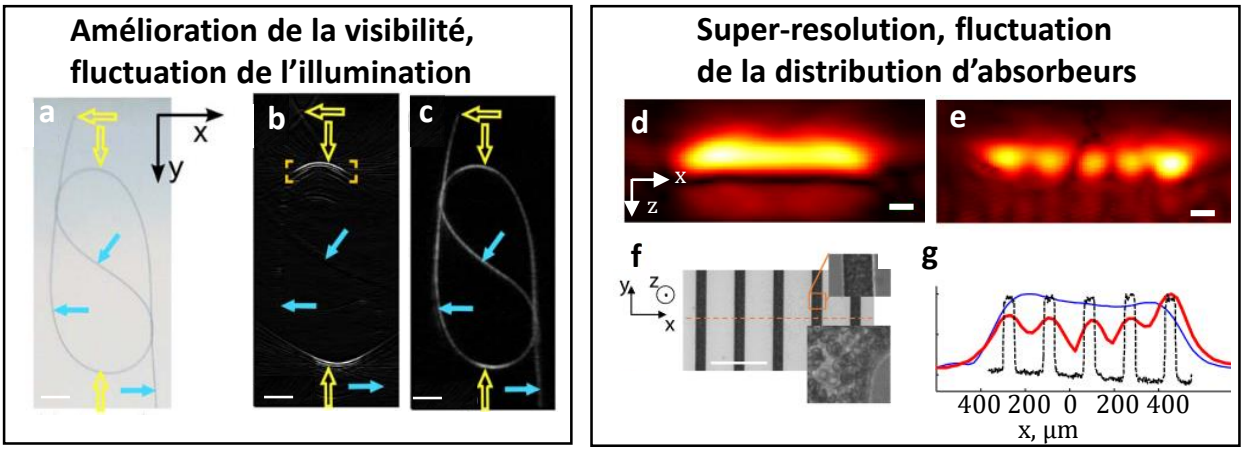

Figure 3.1 – (a-c) : imagerie photoacoustique (PA) de fluctuation pour la correction des artefacts de visibilité par fluctuation de l'illumination [40]. Une série d'images PA 2D d'un phantom (photographie en a) est obtenue. Entre chaque image, le motif de speckle subdiffraction acoustique composant l'illumination change de manière aléatoire. En (b), la moyenne de 50 images : les artefacts de visibilité (flèches bleues) sont présents. En (c), image de fluctuation : les artefacts sont corrigés.
(d-g) : imagerie photoacoustique de fluctuation pour la super-résolution par fluctuation de la distribution d'absorbeurs [18]. On réalise une série d'images PA 2D d'une coupe transversale de 5 canaux microfluidiques (image de microscope du phantom en f) parcourus par du sang humain à flux constant. En (d), la moyenne d'une série d'images PA 2D : les canaux sont indiscernables. En (e), l'image de fluctuation : les canaux sont résolus. Les profils de ces images sont visibles en (g) (bleu, image PA moyenne, rouge, image de fluctuation, noir, image de microscope).

Objectif du chapitre

Si l'imagerie photoacoustique de fluctuation présente également un intérêt pour l'aspect super-résolution, nous nous intéressons ici à ses applications quant à l'amélioration de la visibilité des structures. Dans un premier temps, une théorie de l'imagerie de fluctuation est développée. L'approche est ensuite appliquée *in vitro* avec les sondes d'imagerie 2D et 3D, puis *in vivo* sur l'embryon de poulet et la souris en 3D. Les résultats des acquisitions bimodales sont également présentés : les images de fluctuation PA seront accompagnées d'images Power Doppler US acquises simultanément. L'impact du bruit électronique puis de la présence d'autres sources de fluctuation, telles que la fluctuation tir à tir de l'énergie émise par le laser, sera enfin étudié sur des données de simulation et expérimentales. Une partie des résultats présentés

dans ce chapitre a été publiée dans [132].

3.2 Théorie : formation de l'image photoacoustique de fluctuation

Les développements théoriques qui vont suivre sont illustrés pour l'imagerie photoacoustique de fluctuation induite par le mouvement de particules, bien qu'ils s'appliquent également pour l'illumination par speckles aléatoires. On considère dans cette partie un milieu biphasique composé de phases absorbantes et de phases non absorbantes, dont une série d'images photoacoustiques est obtenue. La distribution des phases change de manière aléatoire entre chaque acquisition k, c'est à dire entre chaque nouvelle illumination. Nous considérons le cas de l'imagerie vasculaire. Les phases absorbantes correspondent alors aux globules rouges, et les phases non absorbantes à des zones du milieu sans globules rouges. Dans cette partie, un éventuel bruit électronique lié au processus d'acquisition est négligé, et la seule source de fluctuation entre deux acquisitions provient du changement de distribution de ces phases.

3.2.1 Définition du phénomène imagé

On considère donc un milieu composé de vaisseaux sanguins et de globules rouges. Ce milieu est décrit par deux fonctions binaires $g_k(\mathbf{r})$ et $f(\mathbf{r})$. La fonction $f(\mathbf{r})$ décrit la structure absorbante imagée : elle vaut 1 dans le vaisseau et 0 à l'extérieur. La fonction $g_k(\mathbf{r})$ définit la position des sources photoacoustiques composant la structure absorbante décrite par $f(\mathbf{r})$, en pratique les globules rouges. La fonction $g_k(\mathbf{r})$ vaut 1 si une source se trouve à la position $\mathbf{r}$ et 0 sinon. La fonction $f(\mathbf{r})$ est identique pour toutes les acquisitions : la structure du vaisseau sanguin n'évolue pas au cours du temps. A l'inverse, la fonction $g_k(\mathbf{r})$ change entre chaque acquisition k. Cette fonction décrit en effet la répartition spatiale des globules rouges, répartition qui change à chaque instant en raison du flux sanguin. Une illustration est proposée en figure 3.2. La structure décrit par $f(\mathbf{r})$ en blanc est identique aux acquisitions k et $k+1$. A l'inverse, la répartition des sources photoacoustiques décrite par $g_k(\mathbf{r})$ en rouge change entre les acquisitions k et $k+1$.

L'information contenue dans l'image PA est la densité d'énergie lumineuse absorbée (chapitre 1, section 1.2.3). Dans cette section, la fluence du laser Φ_0 est considérée constante temporellement et spatialement. Le bruit électronique est nul : la fluctuation du signal PA a pour unique source le caractère aléatoire de $g_k(\mathbf{r})$. L'absorption μ_0 est supposée homogène pour tous les absorbeurs. La distribution d'énergie absorbée peut alors être décrite par l'expression :

$$\alpha_k(\mathbf{r}) = \mu_0 \Phi_0 [g_k(\mathbf{r}) \times f(\mathbf{r})] \tag{3.1}$$

On suppose maintenant que $g_k(\mathbf{r})$ est issu d'un processus aléatoire isotrope et spatialement stationnaire. On peut alors définir ses propriétés statistiques de premier ordre, sa moyenne $\eta =< g_k(\mathbf{r}) >_k$ et sa variance $\sigma_g^2 =< (g_k(\mathbf{r}) - \eta)^2 >_k$ ainsi que celle de second ordre, l'autocovariance :

$$C(\mathbf{r}_1, \mathbf{r}_2) = C\mathbf{r}_1 - \mathbf{r}_2 =< (g_k(\mathbf{r}_1) - \eta)(g_k(\mathbf{r}_2) - \eta) >_k \tag{3.2}$$

L'opération $< . >$ représente le calcul d'une moyenne ici selon k, les différentes acquisitions.

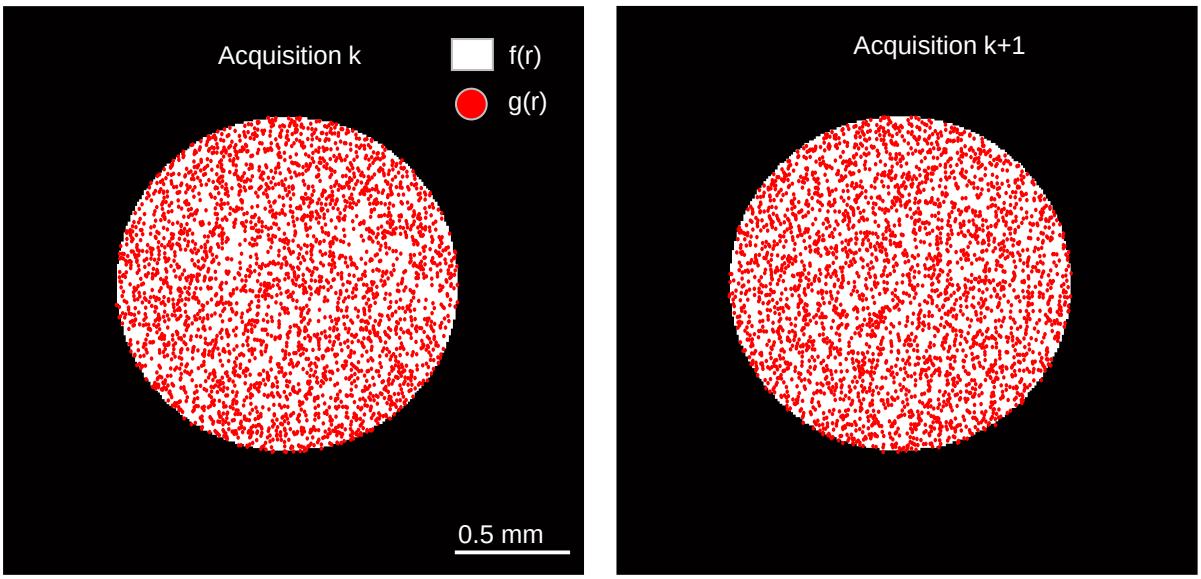

Figure 3.2 – Représentation des fonctions $f(\mathbf{r})$ et $g_k(\mathbf{r})$ en imagerie photoacoustique de fluctuation. La structure absorbante, en pratique un vaisseau sanguin, est décrite par $f(\mathbf{r})$, qui vaut 1 à l'intérieur du vaisseau et 0 à l'extérieur. La fonction $g_k(\mathbf{r})$ décrit la répartition des sources photoacoustiques, ici les globules rouges. Le vaisseau décrit par $f(\mathbf{r})$ en blanc est identique à l'acquisition k et $k+1$. A l'inverse, la répartition des sources photoacoustiques décrite par $g_k(\mathbf{r})$ en rouge change entre les acquisitions k et $k+1$.

On remarque que $C(\mathbf{0}) = \sigma_g^2$. Dans le cadre d'une distribution aléatoire de particules, on peut faire l'hypothèse que la fonction adimensionnée $C(\mathbf{r}_1, \mathbf{r}_2)/C(\mathbf{0})$ est une fonction dont l'intégrale donne un volume fini que nous noterons V_g, de taille caractéristique D_g (avec $D_g^n = V_g$, n=2 ou n=3 selon la dimension du problème). Dans notre contexte, V_g est de l'ordre de grandeur du volume de la particule considérée. La relation entre V_g et V_p le volume de la particule est présentée en annexe. L'intégrale de l'autocovariance I_{cov} est donc :

$$I_{cov} = \int_{\mathbb{R}^n} C(\mathbf{r}) d\mathbf{r} = \sigma_g^2 V_g = \sigma_g^2 D_g^n \tag{3.3}$$

Pour faciliter les démonstrations et la lecture, on néglige l'effet du nombre fini de mesures sur l'estimation des grandeurs statistiques dans les expressions. Par exemple, dans le disque de la figure 3.2, la grandeur η en chaque point (la moyenne selon k) fluctue spatialement car non parfaitement estimée, le nombre de mesure de g_k étant fini. On néglige cependant cette dépendance en $\mathbf{r}$ de η. L'effet du caractère fini du nombre de mesures est néanmoins important notamment lorsque l'on considérera la présence de bruit électronique, et sera donc discuté au long du chapitre.

3.2.2 Définition de l'image PA

On suppose ensuite que chaque image PA reconstruite peut être exprimée par une convolution entre $\alpha_k(\mathbf{r})$ (éq. 3.1) et la *point spread function* (PSF) du système d'imagerie, que l'on appelle $h(\mathbf{r})$:

$$A_k(\mathbf{r}) = \Gamma \cdot \alpha_k(\mathbf{r}) * h(\mathbf{r}), \tag{3.4}$$

avec Γ le paramètre de Grüneisen. L'écriture précédente sous la forme d'une convolution suppose ici que la PSF est spatialement invariante, une hypothèse valide pour un champ de vue suffisamment petit. La démonstration qui suit est valable pour une PSF réelle ou complexe (produite par exemple à partir de la transformée de Hilbert, voir chapitre 1 section 1.3.4). Une représentation des signaux PA et des images reconstruites est proposée en figure 3.3.

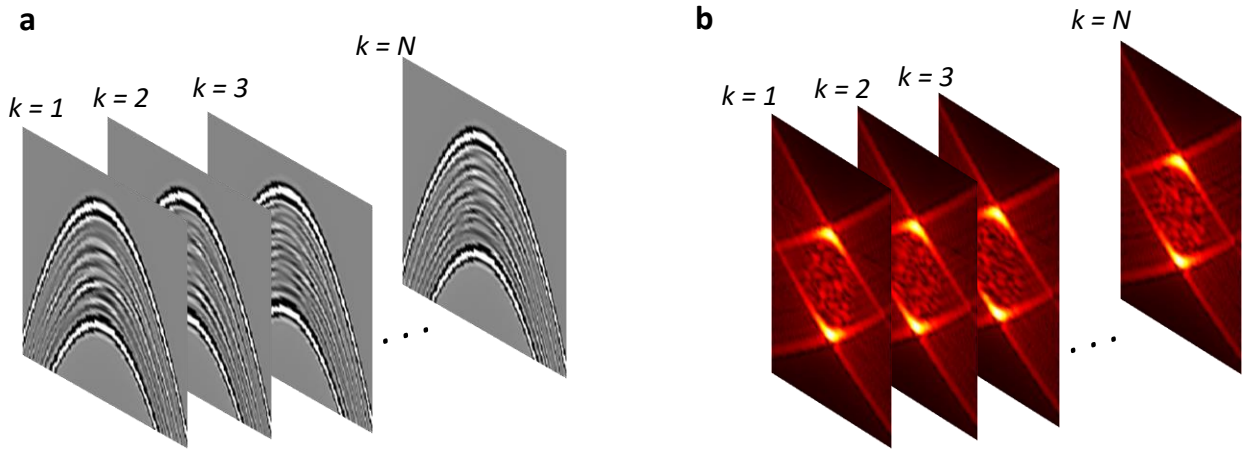

Figure 3.3 – Signaux radiofréquences photoacoustiques (a) acquis pour un milieu de type $g_k(\mathbf{r}) \times f(\mathbf{r})$ avec $f(\mathbf{r})$ un disque de diamètre 1 mm, et images correspondantes en (b). Les barres de couleur sont fortement saturées pour observer la fluctuation du signal à l'intérieur de l'objet.

3.2.3 Définition de l'image moyenne

L'expression de l'image moyenne est donnée par :

$$m(r) = < A_k(\mathbf{r}) >_k = \Gamma \Phi_0 \mu_0 \eta [f(\mathbf{r}) * h(\mathbf{r})] \tag{3.5}$$

avec η toujours égal à $< g_k(\mathbf{r}) >_k$, correspondant dans notre contexte à la fraction du volume de l'objet à imager occupée par les particules, soit la fraction volumique des particules.

L'artefact produit par la réponse fréquentielle limitée de la sonde d'imagerie peut être explicité à l'aide de la formule 3.5. Les PSFs des systèmes d'imagerie classiques étant bipolaires et à moyenne nulle, l'intégrale spatiale de h(r) est égal à 0. Si on s'intéresse à un objet large devant la PSF, on a alors à l'intérieur de l'objet l'égalité $f(\mathbf{r}) = 1$ pour tout $\mathbf{r}$; il en résulte la relation :

$$m[A](r) \propto f(\mathbf{r}) * h(\mathbf{r}) \propto \sqrt{\int_{\mathbf{r}} h(\mathbf{r}) dr} = 0 \tag{3.6}$$

Pour illustrer ce résultat, une représentation 1D de la convolution entre un objet de taille caractéristique 2 mm composé d'absorbeurs ponctuels, et la PSF 1D d'une sonde d'imagerie de fréquence centrale 8 MHz (soit une taille caractéristique de 0.185 µm) est représentée en figure 3.4. Les signaux produits par chaque absorbeur interfèrent entre eux (b). Sur le signal enregistré par la sonde (c), on note l'absence de signal aux temps correspondant aux positions du centre de l'objet. Ces observations s'appliquent également à l'image PA conventionnelle sans moyennage.

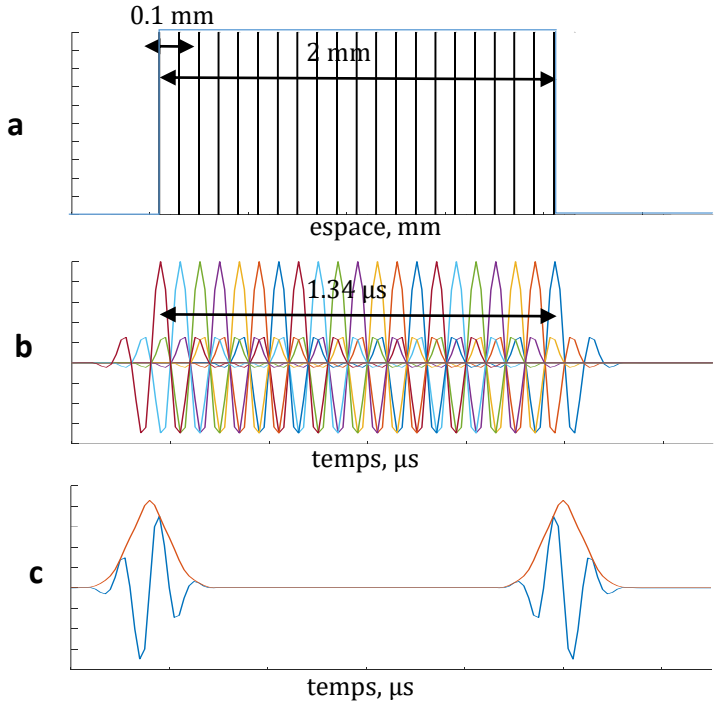

Figure 3.4 – Convolution 1D entre un objet de taille 2 mm et une point spread function (PSF) de longueur caractéristique 0.185 µm. (a) : représentation de l'objet en bleu, position des particules en noir. (b) : les signaux générés par chaque particule sont représentés en différentes couleurs. (c) : somme de ces signaux, correspondant à l'opération de convolution entre la PSF et l'objet, en bleu avec la PSF réelle et en orange avec la PSF complexe, le module du résultat de l'opération de convolution est dans ce deuxième cas représenté.

3.2.4 Définition de l'image de fluctuation

L'image photoacoustique de fluctuation $\sigma[A](\mathbf{r})$ est définie comme la racine carrée de l'image de variance $\sigma^2[A](\mathbf{r})$:

$$\sigma^2[A](\mathbf{r}) =< [A_k(\mathbf{r}) - m[A](\mathbf{r})]^2 >_k =< [A_k(\mathbf{r}) - m[A](\mathbf{r})] \times [A_k(\mathbf{r}) - m[A](\mathbf{r})]^* >_k \tag{3.7}$$

$\sigma^2[A](\mathbf{r})$ peut être exprimée à partir de la fonction d'autocovariance en remplaçant $A_k(\mathbf{r})$ et $m[A](\mathbf{r})$ par leurs expressions respectives, et les convolutions par des intégrales :

$$\begin{aligned}
\sigma^2[A](\mathbf{r}) &= \Gamma^2\mu_0^2F_0^2 < \int_{\mathbf{r}_1}(g_k(\mathbf{r}_1)-\eta)f(\mathbf{r}_1)h(\mathbf{r}-\mathbf{r}_1)d\mathbf{r}_1 \times \\
&\int_{\mathbf{r}_2}(g_k(\mathbf{r}_2)-\eta)f(\mathbf{r}_2)h^*(\mathbf{r}-\mathbf{r}_2)d\mathbf{r}_2 >_k = \\
&= \Gamma^2\mu_0^2F_0^2\int_{\mathbf{r}_1}\int_{\mathbf{r}_2}C(\mathbf{r}_1,\mathbf{r}_2)f(\mathbf{r}_1)f(\mathbf{r}_2)h(\mathbf{r}-\mathbf{r}_1)h^*(\mathbf{r}-\mathbf{r}_2)d\mathbf{r}_1d\mathbf{r}_2
\end{aligned}$$

Pour obtenir une expression simplifiée de l'expression ci-dessus, nous introduisons à présent l'hypothèse suivante : la longueur caractéristique D_g, de laquelle dépend la valeur de l'intégrale de la covariance, est supposée beaucoup plus petite que la longueur caractéristique de $h(\mathbf{r})$ (dimension de la PSF) et $f(\mathbf{r})$ (diamètre des vaisseaux). Dans le contexte de l'imagerie des vaisseaux sanguins, D_g est proche de la taille de la particule absorbante, de l'ordre de 10 µm. La taille de $h(r)$ dépend de la taille de la PSF de la sonde d'imagerie. Pour la majorité des sondes utilisées en imagerie PA, incluant les sondes présentées dans le chapitre précédent, cette grandeur est de l'ordre de la centaine de µm. Notre hypothèse est donc vérifiée dès lors que l'on considère des vaisseaux dont le diamètre est suffisamment supérieur à la taille

d'un globule rouge, c'est à dire quelques dizaines de µm.

On peut alors considérer $C(\mathbf{r}_1,\mathbf{r}_2)$ comme proportionnelle à $\delta(\mathbf{r}_1-\mathbf{r}_2)$ dans l'intégrale, et remplacer ce terme par $[\int_{\mathbb{R}^p} C(\mathbf{r})d\mathbf{r}]\times\delta(\mathbf{r}_1-\mathbf{r}_2)=\sigma_g^2 V_g\delta(\mathbf{r}_1-\mathbf{r}_2)$ dans l'équation 3.2.4 :

$$\sigma^2[A](\mathbf{r})=\Gamma^2\mu_0^2F_0^2\times \iint \sigma_g^2V_g\delta(\mathbf{r}_1-\mathbf{r}_2)f(\mathbf{r}_1)f(\mathbf{r}_2)h(\mathbf{r}-\mathbf{r}_1)h^*(\mathbf{r}-\mathbf{r}_2)d\mathbf{r}_1d\mathbf{r}_2$$

et finalement exprimer $\sigma[A](\mathbf{r})$ sous la forme :

$$\sigma[A](\mathbf{r})=\Gamma\mu_0\Phi_0\sigma_g\sqrt{V_g}\sqrt{f^2(\mathbf{r})*|h|^2(\mathbf{r})} \tag{3.8}$$

L'équation 3.8 relie explicitement l'amplitude de l'image de fluctuation aux propriétés statistiques du processus aléatoire dont elle est issue, au travers de sa moyenne η, son écart type σ_g et le volume V_g de l'autocovariance normalisée . On note également que l'image photoacoustique de fluctuation est directement proportionnelle à μ et peut donc être utilisée pour mener des analyses quantitatives spectroscopiques : cette relation sera approfondie dans le chapitre suivant.

Si $\sigma[A](\mathbf{r})$ est bien l'image dite de fluctuation de par sa proportionnalité à μ, $\sigma^2[A](\mathbf{r})$, l'image de variance, sera régulièrement reprise dans ce chapitre :

$$\sigma^2[A](\mathbf{r})=\Gamma^2\mu_0^2\Phi_0^2\sigma_g^2V_gf^2(\mathbf{r})*|h|^2(\mathbf{r}) \tag{3.9}$$

La variance est en effet une grandeur additive pour des fluctuations indépendantes et s'avère plus pratique à manipuler que l'écart type en présence de bruit électronique.

Des simulations présentées en annexe C nous ont permis de valider la relation de proportionnalité de l'image de fluctuation aux propriétés statistiques V_g et η de g_k, d'étudier influence de ces grandeurs et de relier V_g à V_p. Dans cette annexe, nous relierons également ces grandeurs au facteur de confinement W ou *packing factor*, utilisé par la communauté ultrasonore pour étudier la non linéarité de la dépendance de l'échogénicité à η [6, 12, 115, 121, 128, 129] selon :

$$V_g(\eta)=\frac{W(\eta)}{1-\eta}V_p \tag{3.10}$$

L'image de fluctuation peut alors s'écrire :

$$\sigma[A](\mathbf{r})=\Gamma\mu_0\Phi_0\sqrt{\eta W(\eta)V_p}\sqrt{f^2(\mathbf{r})*|h|^2(\mathbf{r})} \tag{3.11}$$

Grâce à cette relation, il semble possible de relier le coefficient de rétrodiffusion de l'imagerie US avec l'imagerie de fluctuation PA, et des travaux sont actuellement en cours dans le but de définir et d'exploiter cette relation.

Dans la suite de ce manuscrit, on attribue à l'image PA donnée par l'équation 3.4 le nom A_k^{ideal}. Le terme A_k représente toujours l'image photoacoustique mais sa définition dépendra du contexte (présence de bruit électronique, d'autres sources de fluctuation etc ...). On appelle la fluctuation d'intérêt liée au mouvement des globules rouges $\sigma_{flow}(\mathbf{r})$, qui correspond donc à la grandeur $\sigma[A^{\text{ideal}}](\mathbf{r})$ définie à partir de l'éq. 3.8 en conditions idéales. Le terme $\sigma[A](\mathbf{r})$ référera au calcul de la variance des images photoacoustique A_k selon k.

3.2.5 Conséquences sur les artefacts de visibilité

L'équation 3.8 montre que le calcul de l'image de fluctuation implique une convolution entre l'objet et le module de la PSF au carré, alors que le calcul de l'image conventionnelle PA implique le module de la convolution de l'objet avec la PSF. La PSF réelle ou complexe, notée $h_r(r)$ ou $h_c(r)$ est bipolaire et anisotrope : les sommations lors de la convolution sont cohérentes, conduisant à des interférences constructives et destructives ce qui entraîne la présence des artefacts. A l'inverse, lors d'une convolution avec le module du carrée de la PSF, les sommations sont toujours constructives. Il n'y a pas d'interférences destructives entre ces signaux PA, ce qui explique l'absence des artefacts de visibilité sur les images de fluctuation.

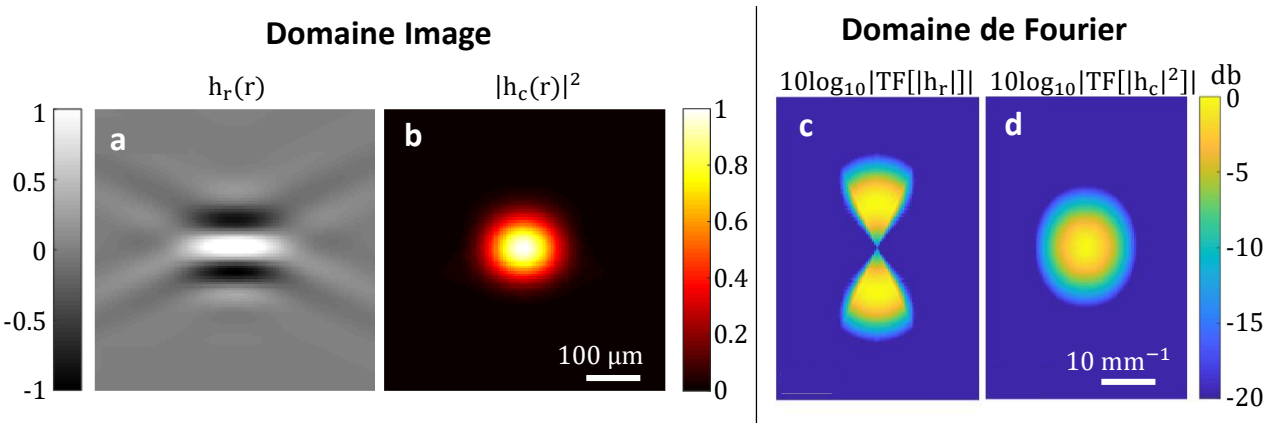

Figure 3.5 – Point spread function (PSF) de l'imagerie photoacoustique (PA) conventionnelle (a,c) et PSF de l'imagerie PA de fluctuation (b,d) à gauche dans l'espace image et à droite dans l'espace de Fourier. Dans l'espace image, la PSF de l'imagerie PA conventionnelle (a) est bipolaire, ce qui se traduit par l'absence de certaines fréquences spatiales dans l'espace de Fourier (c). A l'inverse, la PSF de l'imagerie PA de fluctuation est strictement positive (b) : les fréquences spatiales dans toutes les directions sont conservées comme le montre sa représentation dans l'espace de Fourier (d).

La figure 3.5 illustre les propos précédents. Les PSFs de l'imagerie conventionnelle et de fluctuation y sont présentées ainsi que leurs transformées de Fourier (TF). La bipolarité de la PSF $h_r(r)$ se traduit dans le domaine de Fourier par l'absence de certaines fréquences spatiales. A l'inverse $|h_c|^2(\mathbf{r})$ est strictement positive : les fréquences spatiales sont conservées dans toutes les directions, pour les basses comme les plus hautes fréquences. La convolution se traduit par une multiplication dans l'espace de Fourier. Ainsi, les basses fréquences et certaines directions sont perdues lors de la multiplication avec $TF[h_r(r)]$. Inversement, elles sont conservées lors de la multiplication avec $TF[|h_c(r)|^2]$.

3.2.6 L'image de fluctuation en conditions idéales de simulation

Des simulations ont été réalisées avec le logiciel Simsonic [15], voir annexe B pour plus de détail. On génère P = 512 acquisitions de signaux radiofréquences (RF) détectés par un transducteur linéaire (fréquence centrale 15 MHz, une bande passante de 80%, 128 éléments de largeur 108 µm, soit les caractéristiques de la sonde d'imagerie 2D présentée dans le chapitre 2) en provenance d'un objet rempli de particules. Chaque acquisition correspond à une réalisation d'un processus qui distribue aléatoirement des sources $(g_k(\mathbf{r}) \times f(\mathbf{r}))$ sur une grille de résolution 2.5 µm selon une loi de Bernouilli de valeur moyenne η, fixée à 50%. La taille de la particule est fixée à 10 µm $\times$ 10 µm (16 $\times$ 16 pixels), de l'ordre de grandeur d'un globule rouge. L'image $A_k(\mathbf{r})$ est reconstruite sur une grille de pas 20 µm par *delay and sum*.

L'image de fluctuation $\sigma[A](r)$ est obtenue en calculant l'écart type de la série d'images selon :

$$\sigma[A](r) = \sqrt{\frac{1}{P-1}\sum_{k=1}^{P}\left|A_k(r) - m[A](r)\right|^2} \tag{3.12}$$

avec

$$m[A](r) = \frac{1}{P}\sum_{k=1}^{P} A_k(r) \tag{3.13}$$

La précision de l'estimation de ces images, qui sont des grandeurs statistiques, dépend du nombre d'images P. La variance de l'estimation qui traduit cette précision, décroit généralement en $\frac{1}{\sqrt{(P)}}$. Un compromis devra donc être réalisé entre la qualité de ces estimations et donc la qualité de l'image, et le temps d'acquisition. Nous verrons dans la partie 3.5 que le nombre d'images nécessaire à l'obtention d'une représentation correcte de l'objet dépend également de la présence de bruit dans les données.

La figure 3.6 présente les images de fluctuation dans trois configurations (schématisés en figure 3.7) : une coupe longitudinale d'un tube de 100 µm en forme de boucle (première colonne, a et d), une coupe longitudinale d'un tube d'1 mm de diamètre (seconde colonne, b et e) et une coupe transversale de ce même tube (troisième colonne, c et f). Les artefacts visibles sur les images en (b) et (c) sont causés par la réponse fréquentielle limitée de la sonde. Sur ces deux images, seules les zones correspondantes à de hautes fréquences, les frontières inférieures et supérieures du tube, sont visibles. La partie centrale du tube, où se produisent les interférences destructives, n'est à l'inverse pas visible sur les reconstructions. L'artefact visible sur l'image (a) est produit par l'ouverture numérique limitée de la sonde d'imagerie : seules les composants du tube perpendiculaire à l'axe de la sonde sont visibles, les composants parallèles n'apparaissent pas. Tous ces artefacts visibles sur les images moyennes (images de la première ligne), sont parfaitement corrigés sur les images de fluctuation (image de la seconde ligne).

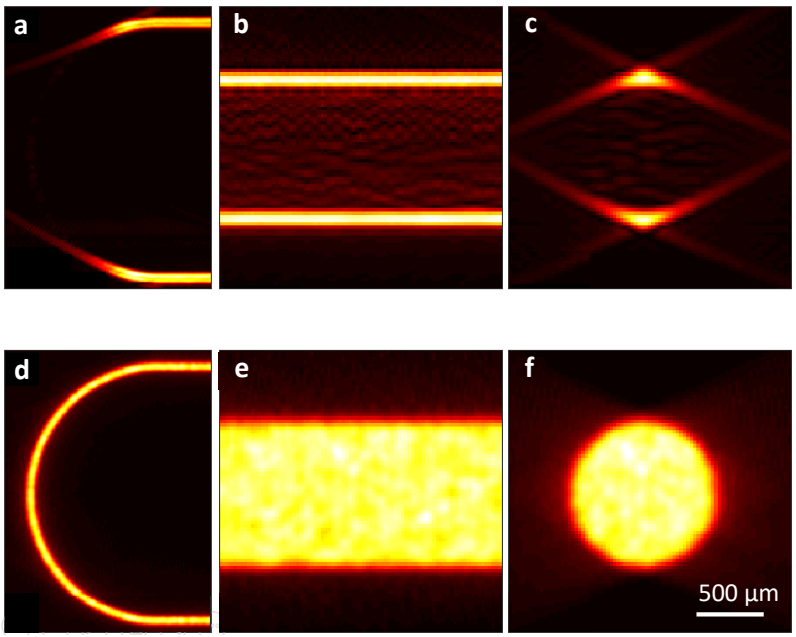

Figure 3.6 – Images photoacoustiques moyennes (première ligne) et images photoacoustiques de fluctuation (seconde ligne) reconstruites à partir de jeux de données simulant les acquisitions de trois objets 2D : une boucle (a,d), une coupe longitudinale d'un tube (b,e) et une coupe transverse d'un tube (c,f).

3.3 Imagerie photoacoustique de fluctuation en conditions expérimentales *in vitro*

3.3.1 Nécessité d'un filtrage spatio-temporel en conditions expérimentales

En réalité, si une partie de la fluctuation du signal photoacoustique entre chaque acquisition est produite par le mouvement des globules rouges, que nous appelons fluctuations d'intérêt, d'autres sources de fluctuation dites "parasites" s'ajoutent en conditions expérimentales. On peut citer la fluctuation de l'énergie émise par le laser (en anglais *pulse energy fluctuation*, abrégé "PEF") , une variation spatiale de la distribution de cette énergie, le *jitter* (défini dans le chapitre 2, qui évalue la qualité de la synchronisation entre le début de l'illumination et le début de l'acquisition), ou encore un mouvement physiologique. L'image de fluctuation n'est alors plus spécifique du signal produit par les globules rouges. Des images de fluctuation expérimentales obtenues sans tenir compte de la présence de ces fluctuations "parasites" sont observables en figure 3.7.d. Les fluctuations d'intérêt y sont masquées par les fluctuations "parasites" de plus haute énergie.

Les fluctuations "parasites" sont de natures différentes de celle d'intérêt, et présentent donc potentiellement des comportements spatio-temporels différents. Discriminer les signaux en fonction de leur dynamique spatio-temporelle pourrait donc permettre d'isoler les signaux produits par la fluctuation d'intérêt. Dans ce but, un filtrage par décomposition en valeurs singulières (SVD) est utilisé. Cette méthode a été introduite en imagerie Power Doppler ultrasonore [28] afin de discriminer le mouvement physiologique des tissus de celui du sang. La SVD permet de décomposer une pile de P images dans une base spatio-temporelle de vecteurs singuliers :

$$A_k(\mathbf{r}) = \sum_{i=1}^{P} \alpha_i U_i(\mathbf{r}) V_i^*(k) \tag{3.14}$$

avec α_i la matrice diagonale de valeurs propres, $U_i(r)$ la matrice de taille $n_\mathbf{r} \times P$ contenant les vecteurs spatiaux ($n_\mathbf{r}$ le nombre de pixels constituant l'image), et $V_i(k)$ la matrice de taille $P \times P$ contenant les vecteurs temporels. Chaque combinaison de valeurs et vecteurs propres correspond à une contribution dont le comportement spatio-temporel diffère. Ces couples de vecteurs sont ordonnés selon leur énergie, traduite par la valeur singulière.

On peut alors appliquer un filtrage sur le choix des valeurs et vecteurs singuliers à conserver pour extraire les fluctuations liées au flux sanguin, selon l'expression :

$$A_k^{SVD}(\mathbf{r}) = \sum_{i=a}^{P} \alpha_i U_i(\mathbf{r}) V_i^*(k) \tag{3.15}$$

avec a la borne inférieure du filtrage par SVD. Le choix de cette borne a est déterminant. Nous verrons dans la partie 3.6.4 qu'en simulation, lorsque l'énergie du laser fluctue entre chaque acquisition (présence de la PEF), il suffit de retirer la première valeur propre ($a = 2$) pour supprimer cette source de fluctuation "parasite" et faire apparaître les structures liées à la fluctuation d'intérêt. En conditions expérimentales, la valeur de a peut varier d'une expérience à une autre, en raison de la présence d'autres sources de fluctuation "parasites" telles que le mouvement physiologique etc ... La justesse du choix de cette borne de filtrage a est donc crucial et influencera la visibilité des structures. Si la borne a est sous-évaluée, les fluctuations d'intérêt peuvent être masquées par les fluctuations "parasites" de plus haute énergie (voir section 3.6). A l'inverse, si cette borne a est surévaluée, les fluctuations d'intérêt peuvent être supprimées.

On peut également définir une borne supérieure de filtrage b. Le filtrage par SVD est maintenant réalisé selon :

$$A_k^{SVD}(\mathbf{r}) = \sum_{i=a}^{b} \alpha_i U_i(\mathbf{r}) V_i^*(k) \tag{3.16}$$

Il peut en effet être utile de supprimer les signaux d'énergie inférieure, par exemple pour diminuer le niveau de bruit électronique. L'impact du choix des bornes a et b sera étudié dans l section 3.6.4.

L'image de fluctuation est finalement estimée selon :

$$\sigma[A^{\mathrm{SVD}}](\mathbf{r}) = \sqrt{\frac{1}{P-1}\sum_{k=1}^{P}\left|A_k^{\mathrm{SVD}}(\mathbf{r}) - m[A_k^{\mathrm{SVD}}](\mathbf{r})\right|^2} \tag{3.17}$$

3.3.2 Imagerie phototoacoustique de fluctuation *in vitro* 2D

Les résultats expérimentaux de cette sous-section ont été obtenus au cours de la thèse de Sergey Vilov [131]. Ils sont présentés à nouveau ici dans un souci de cohérence, au préalable de la présentation des résultats obtenus au cours de ma thèse. Des objets correspondant à ceux des simulations de la section 3.2.6, figure 3.6, sont imagés. Les images moyennes et les images de fluctuation, avec et sans filtrage par SVD sont présentées en figure 3.7 laquelle inclue également un schéma expérimental. Ces images sont obtenues selon les expressions de la section 3.3.1. La boucle est un capillaire en polycarbonate de diamètre intérieur 100 µm et d'épaisseur de paroi 20 µm, et les tubes de diamètre 1 mm sont en verre, et d'épaisseur de paroie 100 µm. Les objets sont immergés dans une cuve d'eau dégazée, et remplis de sang mis en mouvement à l'aide d'un pousse seringue, à une vitesse d'environ 1 cm/s. Les expériences ont été réalisées à $\lambda = 680$ nm. Environ 1000 acquisitions sont réalisées pour chaque configuration.

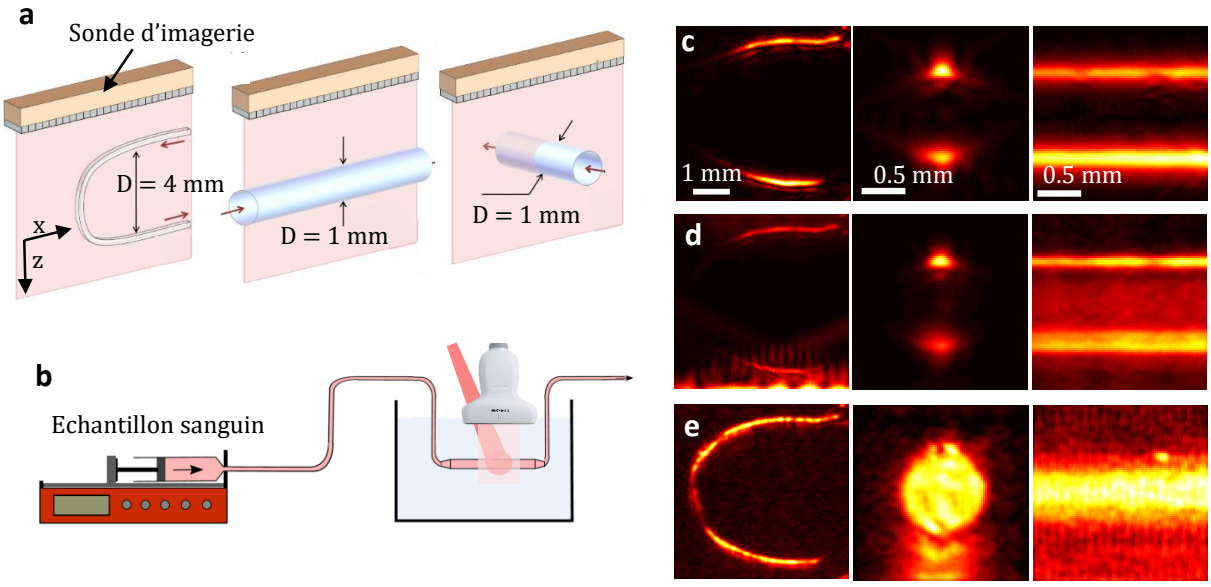

Figure 3.7 – Imagerie photoacoustique (PA) de fluctuation expérimentale en imagerie 2D.
(a) : schémas des trois configurations imagées, la boucle, le tube parallèle au plan d'imagerie et le tube transverse au plan d'imagerie. (b) : dispositif expérimental, représenté dans le cas de l'imagerie du tube parallèle au plan d'imagerie. (c) images PA moyennes. (d) : images PA de fluctuation sans filtrage par décomposition en valeurs singulières (SVD). (e) : images PA de fluctuation après filtrage par SVD.

Comme le montrent les images de la figure 3.7 (d,e), le filtrage par SVD est indispensable à la correction des artefacts de visibilité, malgré une fluctuation de l'énergie émise par le laser relativement faible, mesurée à environ 3% à l'aide d'un détecteur pyroélectrique. La présence de signal résiduel en dehors du tube dans (e), deuxième et troisième colonnes, est probablement due à des réverbérations acoustiques entre les parois du tube en verre. Ces résultats montrent que la présence de bruit n'est pas limitante en imagerie de fluctuation. Le ratio signal à bruit est ici entre 25 et 30, pour une fluctuation d'intérêt environ

2 fois inférieure à l'écart type du bruit. L'impact de ce bruit, lequel influencera notamment le nombre d'images nécessaires à la visibilité de la structure, est détaillé en section 3.5.

3.3.3 Imagerie phototoacoustique de fluctuation *in vitro* 3D

Le dispositif d'imagerie 3D présenté dans le chapitre 2 est utilisé pour l'imagerie phototoacoustique de fluctuation *in vitro* 3D. A noter que dans les expériences de ce chapitre, bien que nous soyons capables de mesurer l'énergie du laser à la surface de l'échantillon (et donc le PEF) grâce à notre photodiode, nous n'utiliserons pas cette information pour corriger nos images. Nous avons en effet remarqué que cette correction n'est pas suffisante pour supprimer l'influence des fluctuations "parasites", ce que le filtrage par SVD permet de réaliser.

Dans une première expérience, du sang est mis en circulation par un pousse seringue à une vitesse moyenne de 2.1 mm.s^{-1} dans un tube en verre (Capillary Tube Supplies Ltd, UK) de 1 mm de diamètre et de paroie 100 µm. La longueur d'onde utilisée est 700 nm, 1000 images sont acquises. Le filtrage par SVD est réalisé avec $a = 55$. Des coupes dans les plans YZ, XZ, et XY de l'image moyenne et de l'image de fluctuation sont présentées en figure 3.8. Leurs observations conduisent aux mêmes conclusions que dans le cas de l'imagerie 2D : l'intérieur du tube est rempli pour les images de fluctuation, l'artefact de visibilité lié à la réponse fréquentielle limitée de la sonde d'imagerie est donc corrigé.

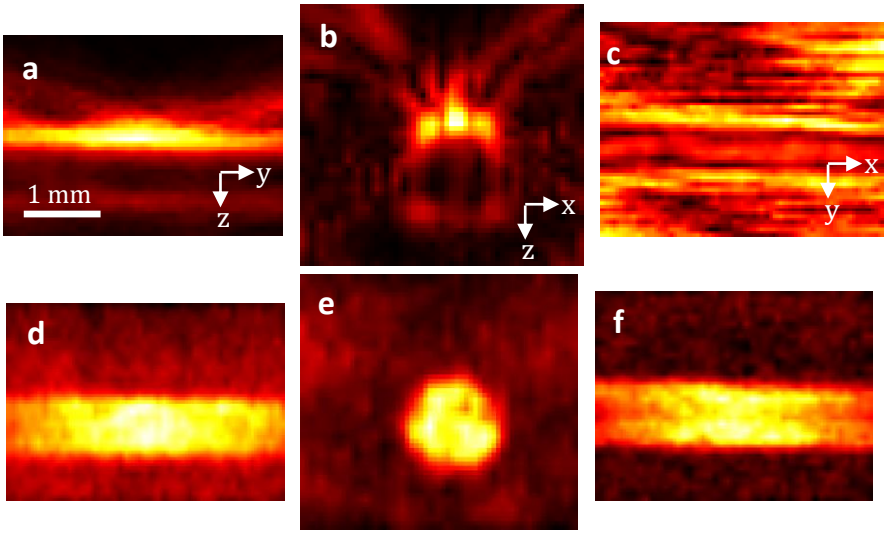

Figure 3.8 – Coupe YZ (a,d) XZ (b,e) et XY (c,f) de l'image de moyenne (première ligne) et de l'image photoacoustique de fluctuation (seconde ligne) d'un tube en verre de 1 mm de diamètre rempli de sang mis en mouvement par un pousse seringue.

Dans un second temps, du sang est placé dans un capillaire en polycarbonate (Paradigm Optics Inc., USA) de 380 µm de diamètre interne et de 460 µm de diamètre externe arrangé sous forme d'une boucle. Le sang est mis en mouvement par un pousse seringue et circule à une vitesse moyenne de 2.1 mm.s^{-1}. A nouveau, 1000 images PA sont acquises. Le filtrage par SVD est réalisé avec a = 30. Les intensités maximales projetées (*maximal intensity projection*, MIP) dans chaque direction sont présentées en figure 3.9.

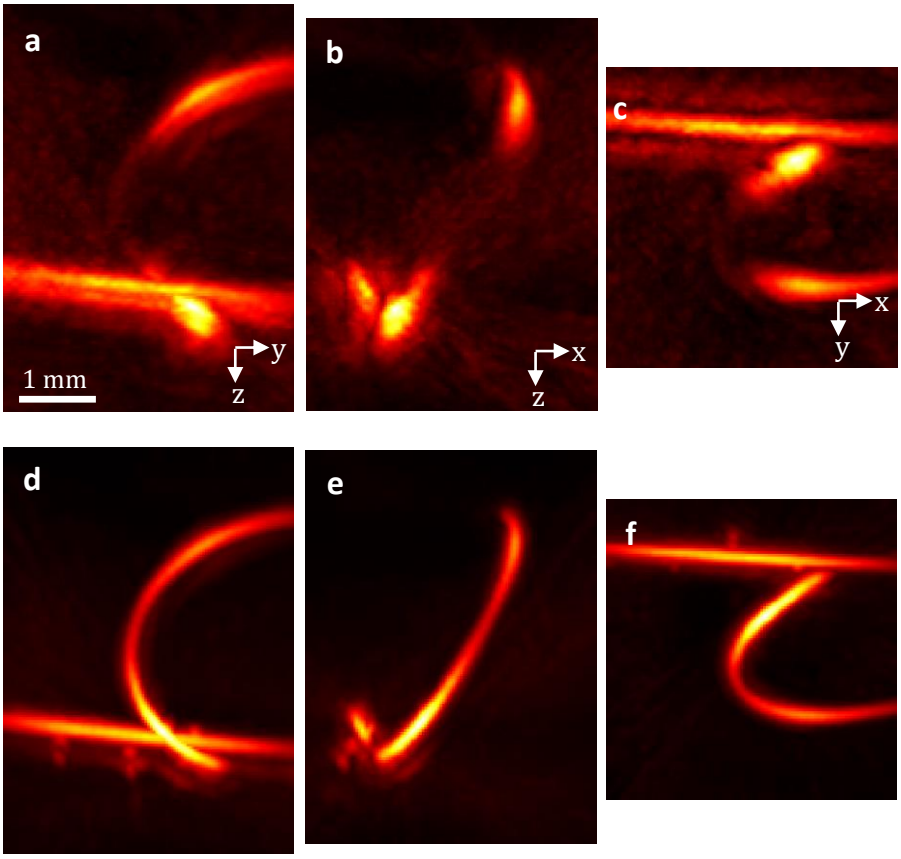

Figure 3.9 – Intensités maximales projetées dans les plans YZ (a,d), XZ (b,e) et XY (c,f) de l'image photoacoustique moyenne (première ligne) et de l'image photoacoustique de fluctuation (seconde ligne) d'un capillaire arrangé sous forme d'une boucle rempli de sang mis en mouvement par un pousse seringue.

La partie verticale, non visible sur l'image de fluctuation, apparaît : l'artefact de visibilité lié à l'ouverture numérique limitée de la sonde d'imagerie est donc corrigé. La position de la boucle n'est ici pas parfaitement verticale afin de limiter au maximum la propagation de l'illumination et de l'onde ultrasonore le long du tube qui provoquerait respectivement de l'absorption et des aberrations de ces deux ondes.

Dans cette section, nous avons donc vérifié que ce qui avait été démontré en imagerie 2D est également valable en imagerie 3D avec une sonde d'imagerie parcimonieuse composée de seulement 256 éléments. Dans la section suivante, l'approche est étendue à l'imagerie photoacoustique de fluctuation *in vivo*.

3.4 Imagerie photoacoustique de fluctuation et Power Doppler 3D *in vivo*

3.4.1 Imagerie photoacoustique de fluctuation 3D de l'embryon de poulet

L'imagerie photoacoustique de fluctuation est réalisée *in vivo*, dans un premier temps sur la membrane chorioallantoique d'un embryon de poulet âgé de 6 jours, selon le schéma expérimental présenté dans le chapitre 2. La longueur d'onde de l'illumination est de 730 nm pour une fluence à la surface d'approximativement 3 mJ/cm², soit 300 mW/cm², de l'ordre de grandeur de la limite de l'*American national standards institute* (ANSI) de 230 mW/cm². Une série de 768 images est obtenue. La figure 3.10 présente les MIP de l'image moyenne et de l'image de fluctuation obtenues après un filtrage par SVD de paramètres $a = 25$ et $b = 220$. L'image moyenne est affectée par les artefacts de visibilité, et par un fort *clutter*. On définit le *clutter* comme le signal enveloppant les structures absorbantes. Sa présence provoque une forte diminution du contraste de l'image (première ligne de la figure 3.10). L'importance du *clutter* dans nos images 3D est probablement provoquée par la parcimonie du réseau de capteurs de notre sonde d'imagerie et au sous-échantillonnage spatial associé.

On peut observer sur l'image de fluctuation la disparition des artefacts de visibilité, traduite notamment par l'apparition de vaisseaux verticaux. Le *clutter* est par ailleurs fortement réduit ce qui permet de faire apparaître de nouveaux vaisseaux. La diminution du *clutter* s'explique par sa relation avec les structures de l'image moyenne. Ces structures proviennent de signaux ne fluctuant pas, ou alors fluctuant selon les fluctuations "parasites" de haute énergie : le *clutter* est donc supprimé avec elles par le filtrage par SVD et le calcul de la variance.

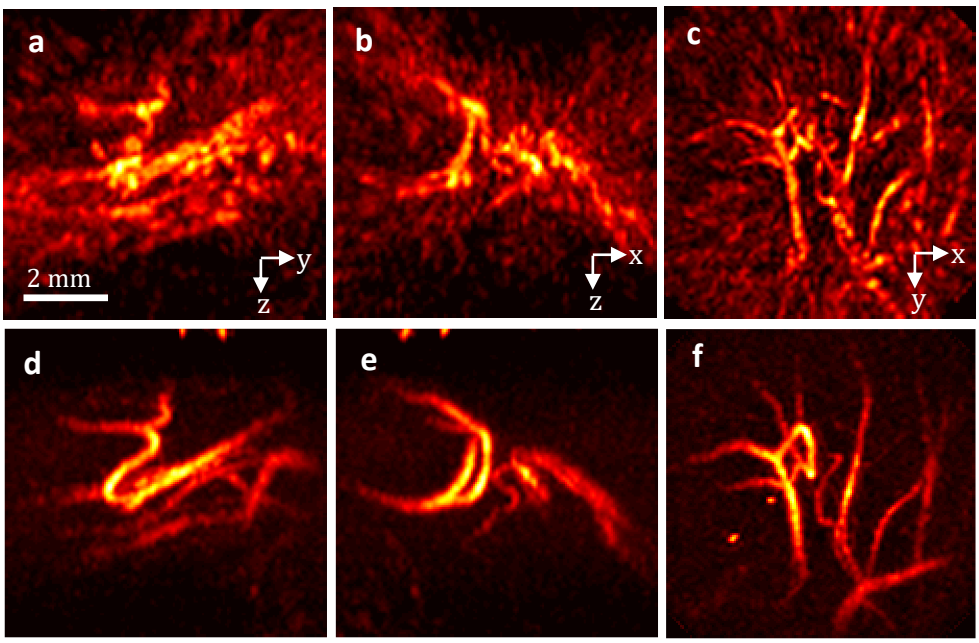

Figure 3.10 – Intensités maximales projetées dans les plans YZ (a,d) XZ (b,e) et XY (c,f) de l'image moyenne (première ligne) et de l'image de fluctuation (seconde ligne) d'un embryon de poulet.

Sur les quatre exemples de la figure 3.11, obtenus à partir du traitement de 1500 images, la correction des artefacts de visibilité et la diminution du *clutter* sont encore plus marquées. Sur l'un des vaisseaux du premier exemple pointé par une flèche, l'artefact d'ouverture numérique limité est bien visible sur l'image moyenne : seules les parties horizontales de ce vaisseau en forme de *W* sont visibles. A l'inverse les parties verticales sont visibles sur l'image de fluctuation. Sur le second exemple, l'une des flèches au centre de l'image pointe vers une structure affectée par l'artefact de réponse fréquentielle limitée. Une seconde pointe à nouveau sur un vaisseau affecté par l'ouverture numérique limitée de la sonde.

Les deux artefacts sont corrigés sur l'image de fluctuation. Les exemples 3 et 4 montrent l'importance de l'effet de réduction du *clutter*, entraînant l'apparition d'une multitude de vaisseaux et petits capillaires correspondants à la membrane chorioallantoique. Ce réseau de capillaires est repéré par le cercle sur l'exemple 4.

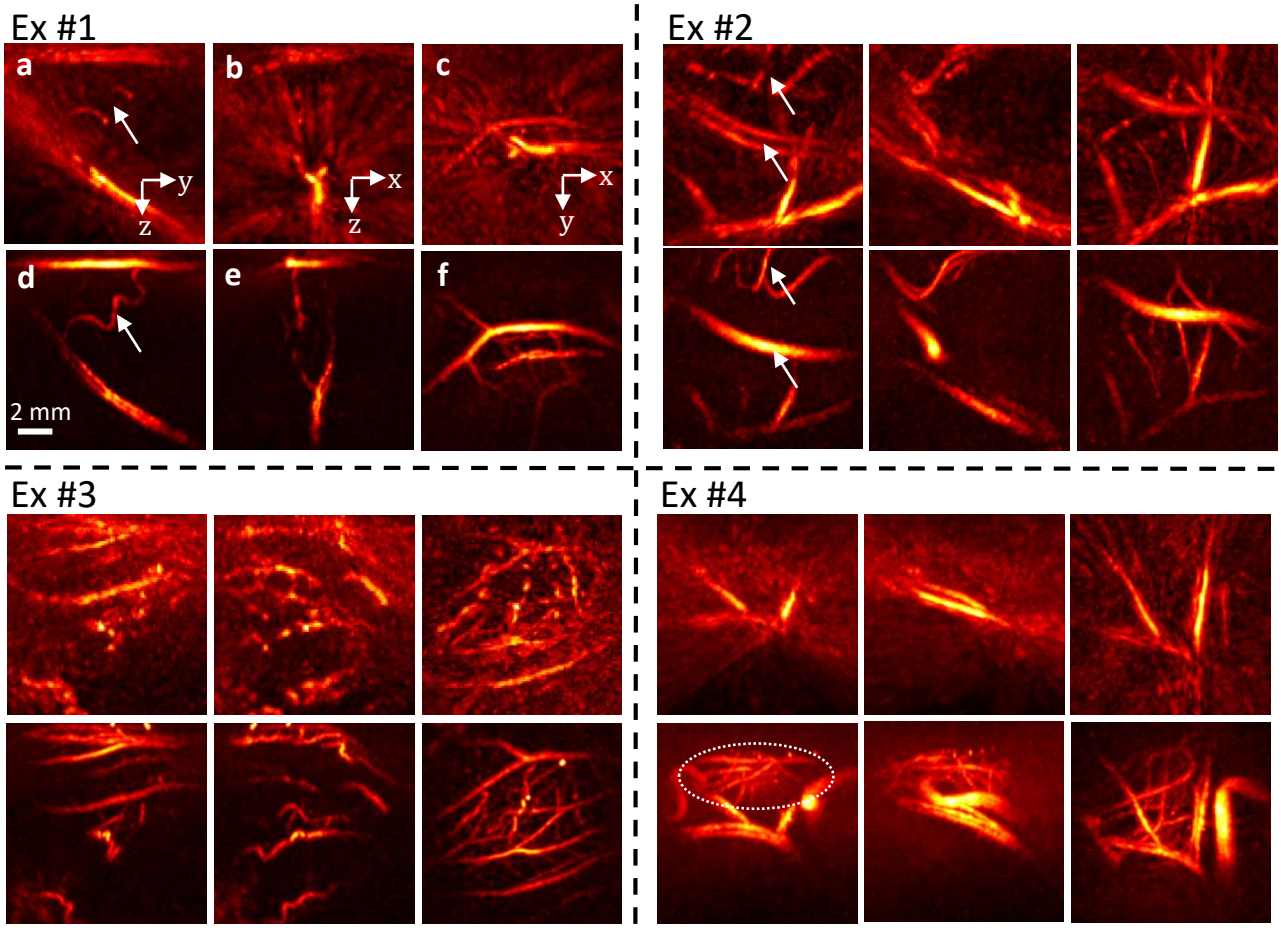

Figure 3.11 – Intensités maximales projetées dans les plans YZ (a,d) XZ (b,e) et XY (c,f) de l'image moyenne (première ligne) et de l'image de fluctuation (seconde ligne) d'un embryon de poulet. Quatre exemples sont présentés.

3.4.2 Imagerie Power Doppler ultrasonore 3D de l'embryon de poulet

Notre dispositif d'imagerie permet également de réaliser une approche de type imagerie Power Doppler ultra-rapide, comme présenté dans le chapitre 2. A partir des images ultrasonore (US) reconstruites par *delay and sum*, les images Power Doppler peuvent être obtenues. Le traitement des données est similaire à celui de l'imagerie photoacoustique de fluctuation. Les images reconstruites sont filtrées par SVD, et l'image Power Doppler est estimée selon :

$$\sigma^2[B^{SVD}](\mathbf{r}) = \sqrt{\frac{1}{P-1}\sum_{k=1}^{P}\left|B_k^{SVD}(\mathbf{r}) - m[B_k^{SVD}](\mathbf{r})\right|^2} \tag{3.18}$$

avec $B_k(\mathbf{r})$ l'image US. La zone de l'embryon de poulet dont l'image photoacoustique de fluctuation a été présentée en figure 3.10 a également été imagée en imagerie ultrasonore dans le but d'obtenir son image Power Doppler. Une série de P = 1000 images US sont acquises à une fréquence de 200 Hz. Chaque image est reconstruite à partir d'une série de 13 sous-acquisitions en changeant l'élément émetteur. Les MIP de l'image Power Doppler sont présentées en figure 3.12.

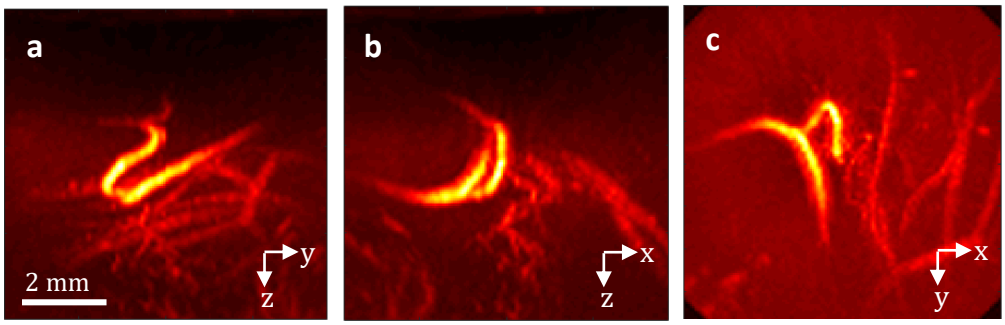

Figure 3.12 – Intensités maximales projetées dans les plans YZ (a) XZ (b) et XY (c) de l'image Power Doppler d'un embryon de poulet.

Des structures similaires à celles visibles sur l'image de fluctuation apparaissent. La présence d'un réseau de capillaire est observable dans la partie basse de la zone d'imagerie. L'applicabilité de l'approche Power Doppler ultra-rapide en imagerie 3D avec une sonde parcimonieuse composée de seulement 256 éléments est ainsi démontrée. Deux autres images Power Doppler d'embryon de poulet, reconstruites à partir de P=1000 images US, sont visibles en figure 3.13.

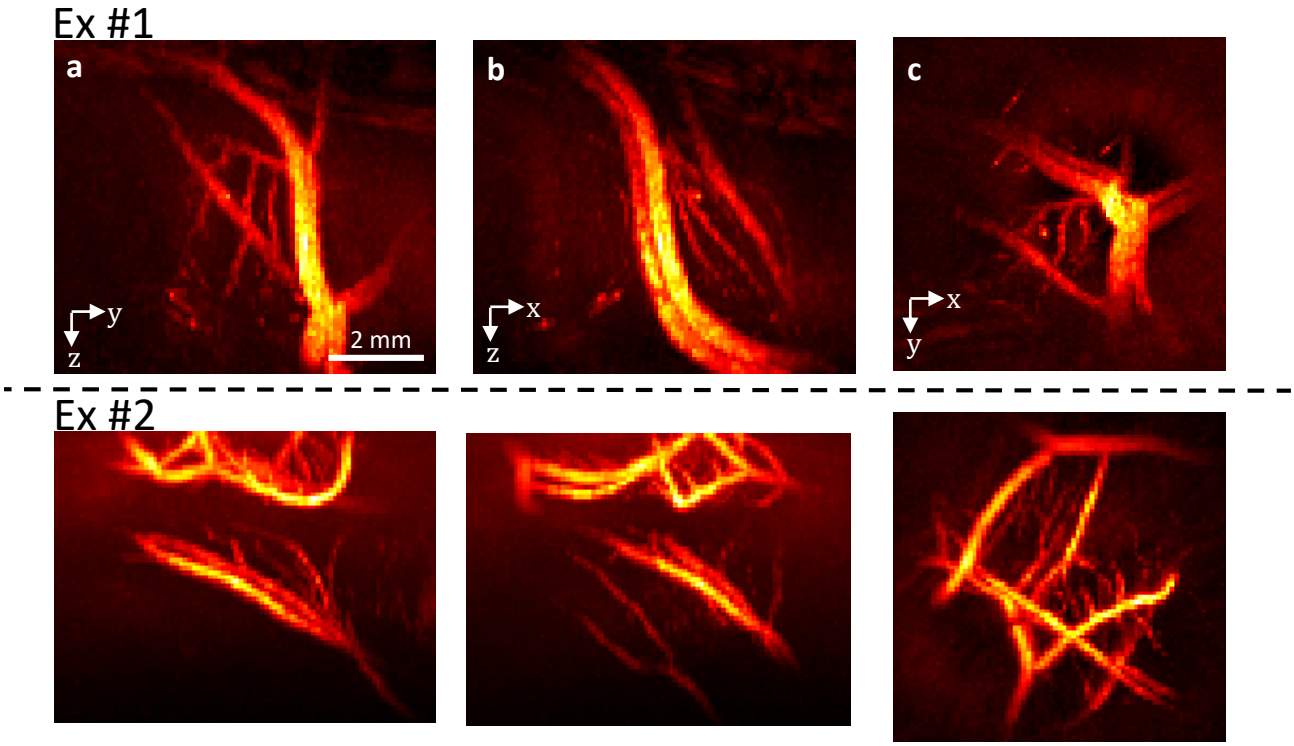

Figure 3.13 – Intensités maximales projetées dans les plans YZ (a) XZ (b) et XY (c) de l'image Power Doppler d'un embryon de poulet. Deux exemples sont présentés.

3.4.3 Imagerie photoacoustique de fluctuation et imagerie Power Doppler ultra-rapide simultanées de l'embryon de poulet

Par la suite, la séquence d'imagerie présentée dans le chapitre 2 est utilisée pour réaliser simultanément les acquisitions PA et US. Deux exemples d'images photoacoustiques de fluctuation et d'images Power Doppler résultantes de telles acquisitions sont proposées en figure 3.14. L'embryon imagé est âgé de 9 jours. Les images sont reconstruites avec P respectivement égal à 1000 et 2500 images en imagerie PA, et P = 1000 images acquises à une fréquence de 200 Hz en imagerie US pour les deux exemples. Le

nombre de sous-image US par image US est ici de 10, et la longueur d'onde en imagerie PA est 730 nm. Les représentations 3D sont obtenues à l'aide du logiciel Amira.

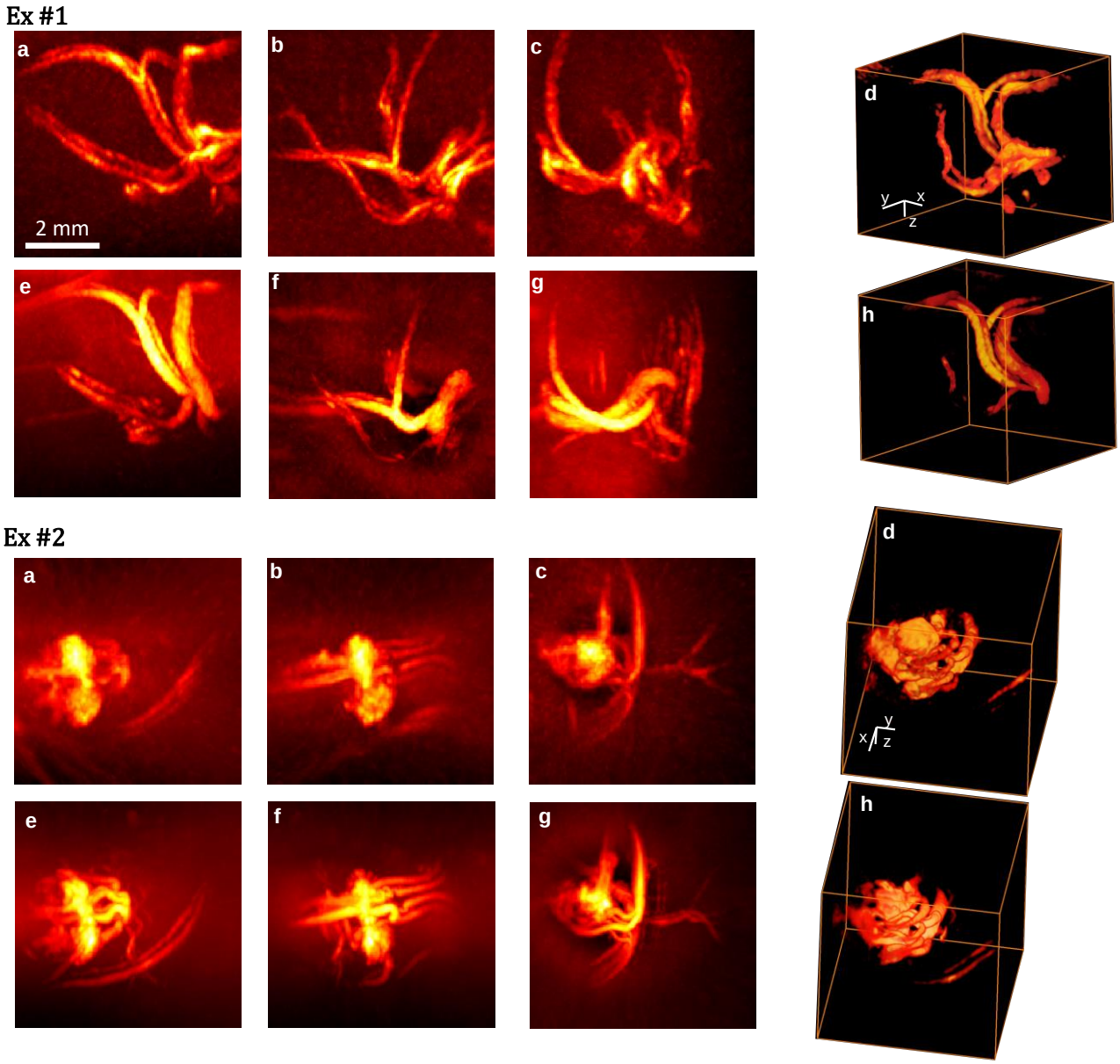

Figure 3.14 – Image photoacoustique de fluctuation (a-d, première ligne) et image Power Doppler (e,h seconde ligne). Les intensités maximales projetées dans les plans YZ (a,e) XZ (b,f) et XY (c,g) sont observables. Les représentations 3D (d,h) sont obtenues à l'aide du logiciel Amira. Deux exemples sont présentés.

L'imagerie photoacoustique de fluctuation et l'imagerie Power Doppler ultrasonore exploitent le même phénomène pour améliorer la qualité de l'image : la fluctuation du signal des globules rouges amenée par le flux sanguin. Des structures similaires sont donc observables sur les images fournies par les deux modalités. Alors que l'image US est reliée à l'amplitude des ondes rétrodiffusées par ces globules, l'image PA est directement reliée au volume sanguin et à l'absorption optique : les deux modalités produisent ainsi des informations complémentaires. La combinaison de ces deux approches *state of the art* en un unique dispositif d'imagerie est une première. Cela ouvre des perspectives intéressantes d'exploitation combinée des images issues de ces modalités.

3.4.4 Imagerie photoacoustique de fluctuation d'une souris : résultats préliminaires

Des résultats préliminaires en imagerie photoacoustique de fluctuation ont également été obtenus sur une souris, selon le protocole décrit dans le chapitre 2. La difficulté augmente ici considérablement, pour plusieurs raisons. Tout d'abord, les images de souris sont plus affectées par les artefacts de mouvement que les images d'embryons, en raison notamment du cycle respiratoire. Le milieu imagé est maintenant fortement diffusant et absorbant : la lumière est très atténuée, l'amplitude du signal est donc plus faible. Par ailleurs, dans le cas de l'imagerie du cerveau, la traversée du crane peut provoquer des aberrations de l'onde ultrasonore et sa réflexion dans le milieu.

Une souris est imagée à 730 nm, et P=1500 images sont obtenues. Les résultats présentés dans cette section sont préliminaires et doivent être pris avec précautions. L'image moyenne d'un cerveau de souris, l'image photoacoustique de fluctuation correspondante et cette image photoacoustique de fluctuation après application d'un filtre de Frangi sont présentées en figure 3.15.

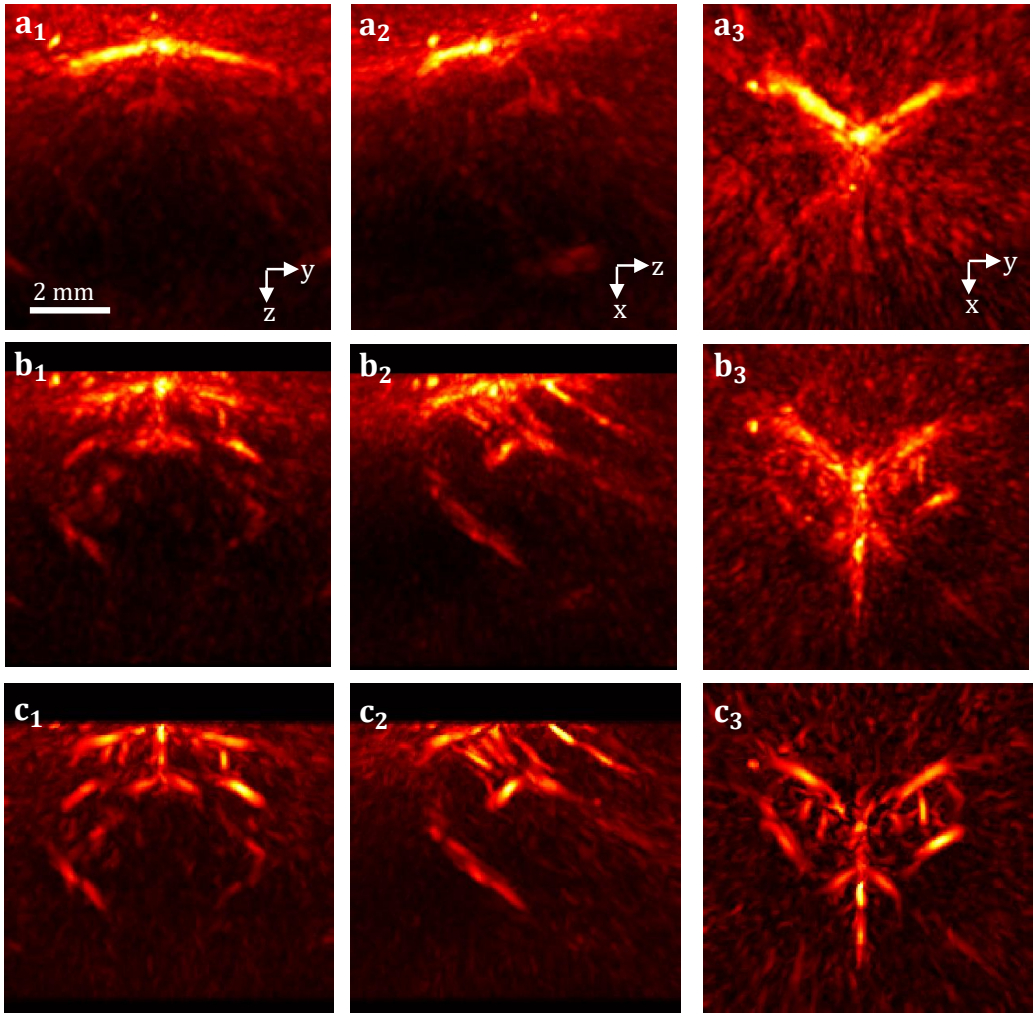

Figure 3.15 – Intensités maximales projetées dans les plans YZ, XZ et XY (indices 1, 2 et 3) pour (a) l'image moyenne, (b) l'image de fluctuation et (c) l'image de fluctuation après application d'un filtre de Frangi.

Le filtrage de Frangi est couramment utilisé pour restituer des structures allongées dans une direction telles que des vaisseaux sanguins. Les premières tranches dans la direction z sur les images de fluctua-

tions ne sont pas conservées pour la représentation des MIPs afin de supprimer le *clutter* résiduel lié au signal de la peau. Ce *clutter* diminue légèrement sur l'image de fluctuation, et des structures apparaissent à la surface du cerveau et en profondeur. Le filtrage de Frangi permet de distinguer une multitude de vaisseaux, bien qu'il convient d'être prudent avec ces méthodes de filtrage, du bruit pouvant être assimilé à des vaisseaux.

Deux autres exemples sont présentés en figure 3.16, le premier correspondant toujours à l'imagerie du cerveau de la souris et le second à l'imagerie de l'artère fémorale.

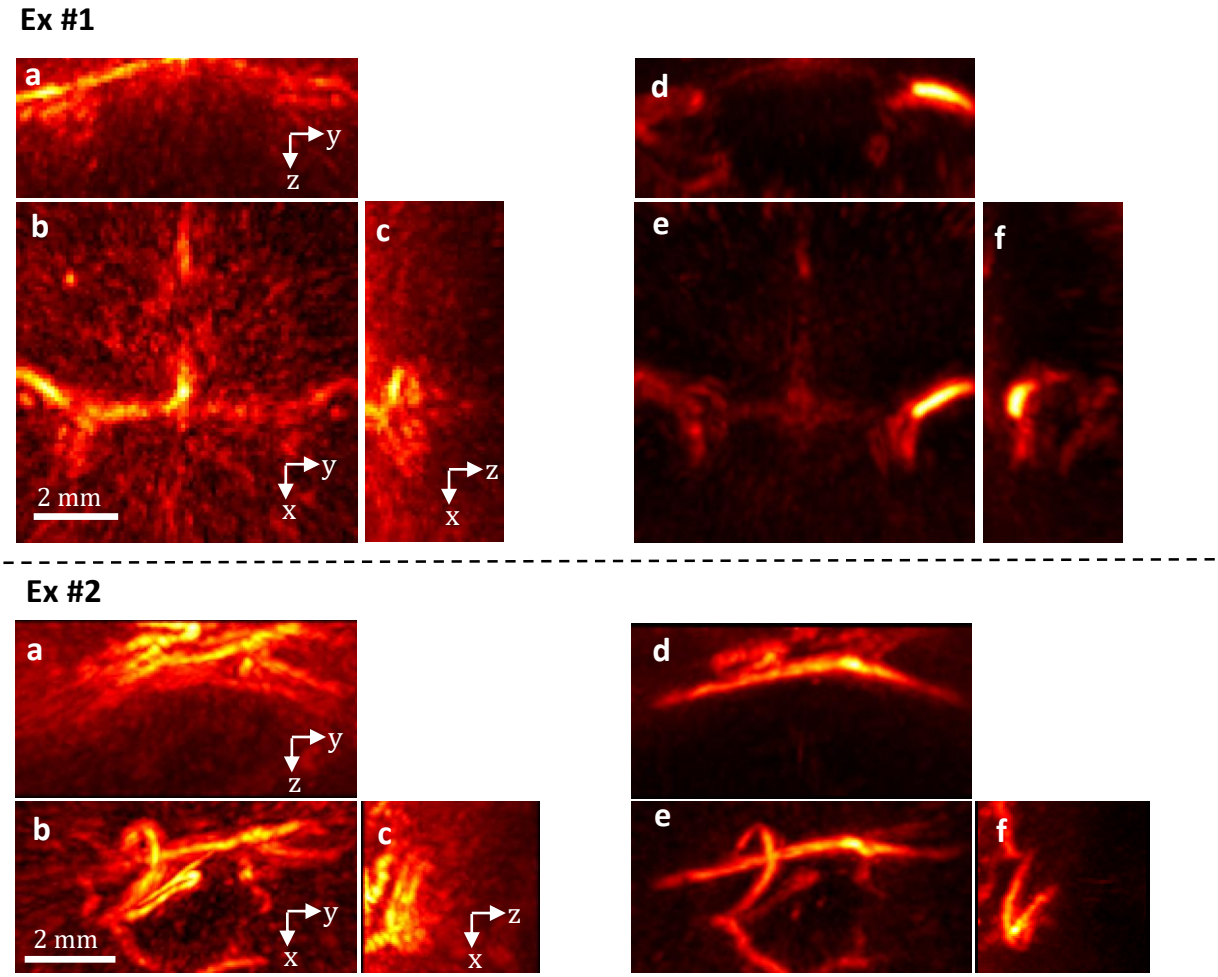

Figure 3.16 – Intensités maximales projetées dans les plans YZ (a,d) XY (b,e) et XZ (c,f) de l'image moyenne à gauche et de l'image de fluctuation à droite. L'exemple 1 correspond à une acquisition sur le cerveau de souris, l'exemple 2 à celle de l'artère fémorale.

Une forte diminution du *clutter* est observable sur les deux exemples. Les structures sont ainsi plus nettes sur l'image de fluctuation. Notamment pour l'image de l'artère fémorale, on distingue bien plus clairement l'enchevêtrement des différents vaisseaux. Néanmoins, le réseau cérébro-vasculaire est peu visible, sur l'exemple 1 de cette figure comme sur celui de la figure 3.16. Ces résultats préliminaires sont encourageant et démontrent que l'imagerie de fluctuation est applicable à l'imagerie du petit animal. Les images de fluctuations peuvent très probablement être améliorées, en modifiant les conditions de reconstruction des images PA ou en modifiant les filtres appliqués sur les signaux RF par exemple. Potentiellement, le signal PA est trop faible pour extraire les fluctuations liées au mouvement des globules rouges en profondeur. Des développements *hardware* pourront être réalisés pour améliorer la sensibilité, en utilisant par exemple un amplificateur en amont de l'enregistrement des signaux RF.

3.5 Impact du bruit électronique sur la qualité de l'image de fluctuation

La présence de bruit électronique est inhérente à l'acquisition de signaux ultrasonores. Comme le montre les résultats expérimentaux, sa présence n'est pas forcément limitante en imagerie de fluctuation. Néanmoins, il est intéressant d'observer son impact et la manière dont il contribue à l'image finale. Nous verrons notamment que son importance conditionnera en partie le nombre d'images minimales nécessaires à la formation de l'image de fluctuation et que la prise en compte du bruit sera absolument nécessaire dans la perspective d'extraire une information quantitative de l'image du fluctuation (chapitre 4).

3.5.1 Définition de l'image de fluctuation en présence de bruit

Comme indiqué en fin de première partie, on appelle la fluctuation d'intérêt liée au mouvement des globules rouges $\sigma_{flow}(\mathbf{r})$, qui correspond donc à la grandeur $\sigma[A_k^{\text{ideal}}](\mathbf{r})$ définie à partir de l'éq. 3.8 en conditions idéales (absence de bruit, unique source de fluctuation etc ...). Un bruit électronique $b_k(\mathbf{r})$ d'écart type σ_b s'ajoute maintenant aux signaux radiofréquences. L'image PA en considérant le bruit électronique $A_k(\mathbf{r})$ peut s'écrire selon :

$$A_k(\mathbf{r}) = A_k^{ideal}(\mathbf{r}) + b_k(\mathbf{r}) \tag{3.19}$$

Le calcul de la variance produit :

$$\sigma^2[A(\mathbf{r})] =< |A_k^{ideal}(\mathbf{r}) + b_k(\mathbf{r}) - < A_k(\mathbf{r}) > |^2 > \tag{3.20}$$

Ce bruit $b_k(\mathbf{r})$ est de nature additive et à moyenne nulle : pour un nombre infini d'images, $< A_k(\mathbf{r}) >=< A_k^{ideal}(\mathbf{r}) >$ (en pratique $< A_k(\mathbf{r}) >\approx< A_k^{ideal}(\mathbf{r}) >$). Les fluctuations du bruit électronique et du signal d'intérêt sont indépendantes : leurs variances s'ajoutent et on obtient l'expression :

$$\sigma[A](\mathbf{r}) = \sqrt{\sigma_{flow}^2(\mathbf{r}) + \sigma_b^2} \tag{3.21}$$

avec

$$\sigma_b^2 =< |b_k(\mathbf{r})|^2 >_k \tag{3.22}$$

3.5.2 Conséquences de la présence de bruit sur l'image reconstruite en simulation

Les estimations des grandeurs σ_b^2 et $\sigma_{flow}^2(\mathbf{r})$ dépendent du nombre d'acquisitions. La variance, qui traduit la précision de ces estimations, décroit généralement en $\frac{1}{\sqrt{(P)}}$. Si l'on disposait d'un nombre d'images P infini, les grandeurs σ_b^2 et $\sigma_{flow}^2(\mathbf{r})$ seraient parfaitement estimées. Cela impliquerait que la grandeur σ_b^2 est identique pour chaque pixel de l'image : il serait alors toujours possible d'observer les structures correspondant à $\sigma_{flow}^2(\mathbf{r})$, quelque soit l'amplitude de la fluctuation sous réserve d'une quantification électronique suffisante (de l'ordre au moins de $\sigma_{flow}^2(\mathbf{r})$). L'image $\sigma_{flow}^2(\mathbf{r})$ apparaîtrait en effet au-dessus d'un fond constant de valeur σ_b^2. En pratique, ces estimations de σ_b^2 et $\sigma_{flow}^2(\mathbf{r})$ sont évidemment imparfaites car le nombre d'images que l'on peut acquérir est limité. Cela induit notamment une variation spatiale de σ_b, visible dans la figure 3.17, panneau de droite.

On introduit ici le ratio fluctuation à bruit (*fluctuation to noise ratio*, FNR) défini tel que :

$$FNR^{RF} = \frac{\sigma_{flow,RF}}{\sigma_{b,RF}} \; et \; FNR^{BF} = \frac{\sigma_{flow}}{\sigma_b} \tag{3.23}$$

à bien différencier du ratio signal sur bruit défini tel que $PNR^{RF} = \frac{RF_{max}}{\sigma_{b,RF}}$ avec RF_{max} la valeur maximale des signaux radiofréquences. Ces grandeurs sont illustrées dans la figure 3.17, sur laquelle apparaît également une estimation de l'image de fluctuation 2D d'un disque rempli de particule, et de son profil en z. On observe clairement l'objet qui apparaît au-dessus du bruit de fond fluctuant légèrement spatialement, de valeur moyenne σ_b.

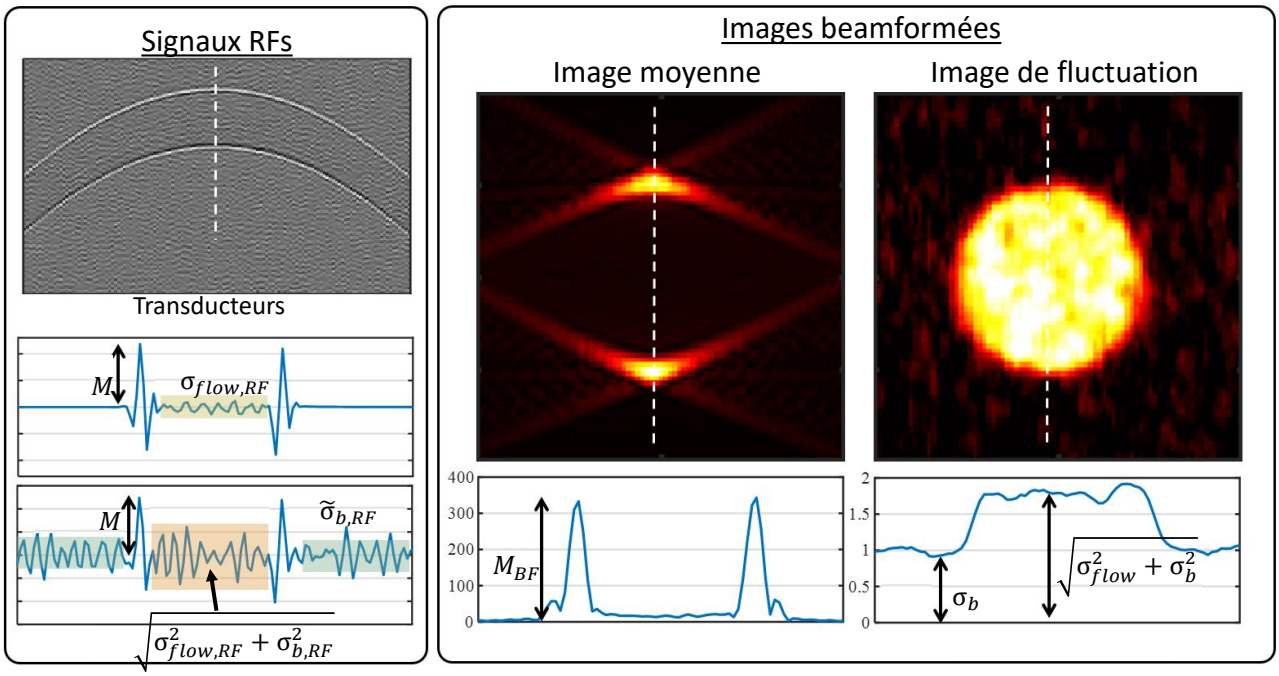

Figure 3.17 – Panneau de gauche : représentation des signaux radiofréquences, et zoom sur le signal reçu par un transducteur en présence et en l'absence de bruit électronique $\sigma_{b,RF}$. La fluctuation de ce bruit électronique s'ajoute à celle d'intérêt $\sigma_{flow,RF}$. Panneau de droite : Estimation de l'image moyenne et de l'image de fluctuation, et profils correspondants selon les lignes pointillées. L'estimation de l'image de $\sigma_{flow}(\mathbf{r})$ apparaît au-dessus d'un niveau de bruit de fond de valeur moyenne σ_b. L'estimation de σ_b varie spatialement, mais la fluctuation σ_{flow} est suffisamment grande et bien estimée pour que l'objet soit visible.

La visibilité de l'objet dépendra à la fois du FNR et du nombre d'images P, lequel influe sur la qualité des estimations et donc sur la variabilité spatiale des grandeurs $\sigma_{flow}(\mathbf{r})$ et $\sigma_b(\mathbf{r})$. Ainsi, pour un nombre d'images et un FNR trop faible, l'objet sera masqué, 'noyé' dans la variation de l'estimation de σ_b entraînant une fluctuation du bruit de fond de l'image de fluctuation. Ces considérations sont illustrées dans la figure 3.18 pour laquelle l'image de fluctuation $\sigma[A]$ est reconstruite en faisant varier le nombre d'acquisition et le FNR.

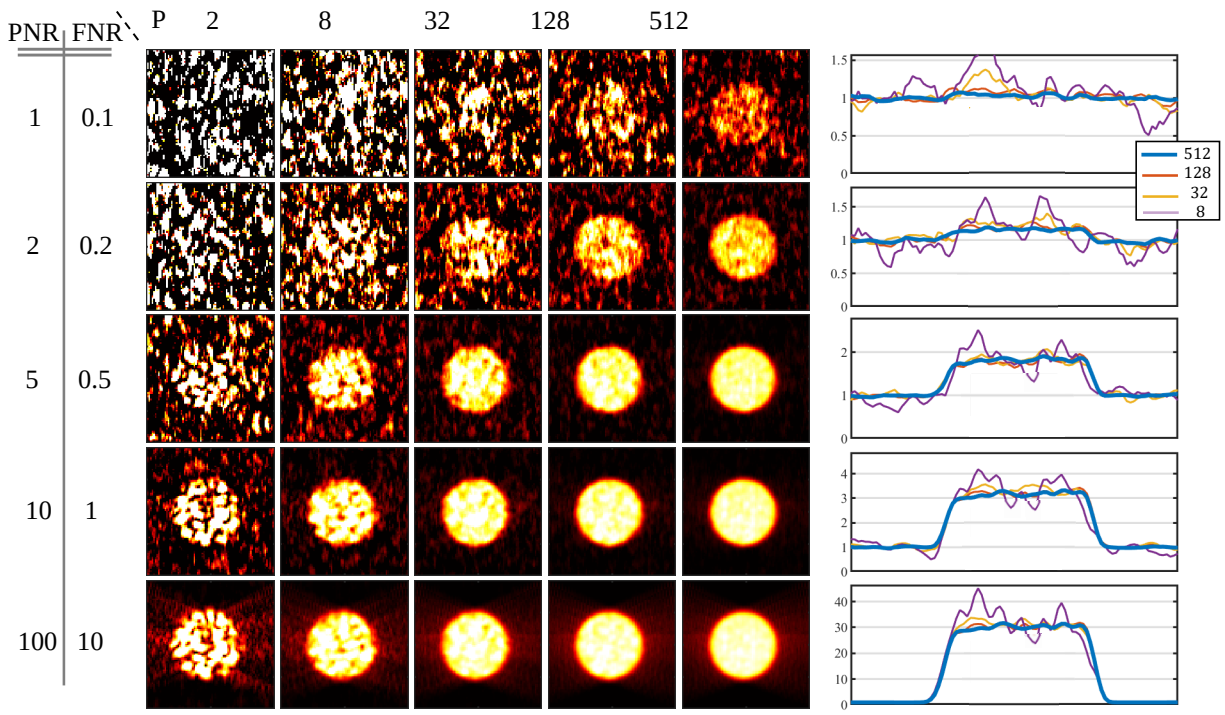

Figure 3.18 – Estimation de l'image de fluctuation pour différents couples de ratio fluctuation à bruit (fluctuation to noise ratio, FNR) et de nombre d'images utilisées pour le calcul de l'image. Les barres de couleur sont ajustées pour que le noir corresponde au minimum du bruit et le jaune au niveau moyen de l'objet. PNR correspond au rapport signal à bruit (peak to noise ratio). Les profils de ces images dans l'axe horizontal sont également présentés.

Ainsi, malgré une fluctuation très faible, dix fois inférieure au bruit ($FNR = 0.1$,$PNR = 1$), l'objet est visible si le nombre d'images est suffisamment grand. Inversement, la qualité de la reconstruction diminue lorsque le nombre d'images diminue.

3.5.3 Conséquences de la présence de bruit sur l'image reconstruite lors des expériences *in vitro* et *in vivo*

La variation de la qualité de l'image de fluctuation selon le nombre d'images acquises suit un comportement similaire en simulation et sur les données expérimentales. Dans le cas de la coupe transversale du tube (imagerie 2D), comme le montrent les profils axiaux de la figure 3.19, l'objet apparaît au-dessus du bruit.

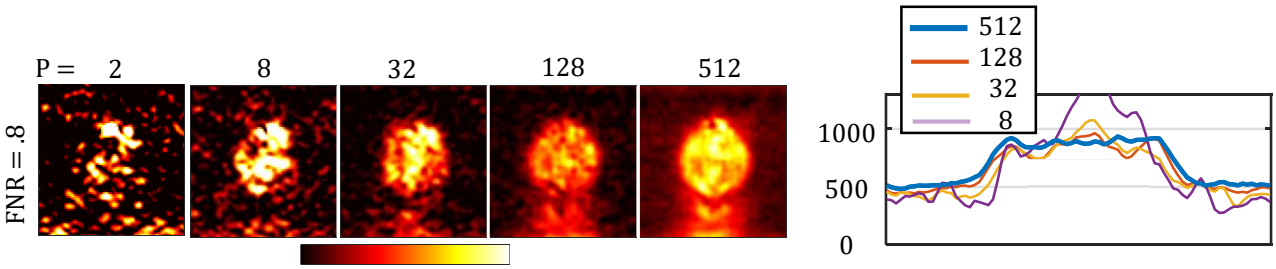

Figure 3.19 – Influence du nombre d'images P sur la visibilité de l'objet dans le cas d'acquisitions expérimentales de la coupe transversale d'un tube en imagerie 2D. Des images de fluctuation sont présentées. Les barres de couleurs sont ajustées pour que le noir corresponde au minimum du bruit et le jaune au niveau moyen de l'objet. Un profil selon l'axe x est présenté pour chaque reconstruction. Le ratio fluctuation à bruit (fluctuation to noise ratio, FNR) est de 0.8 et le ratio signal à bruit de 26.

Il est détectable lorsque l'on dispose d'un nombre d'images suffisamment grand pour discriminer les niveaux moyens de $\sigma_{flow}(\mathbf{r})$ et de σ_b.

La même étude est réalisée sur les données de l'embryon de poulet, présentée dans la figure 3.20.

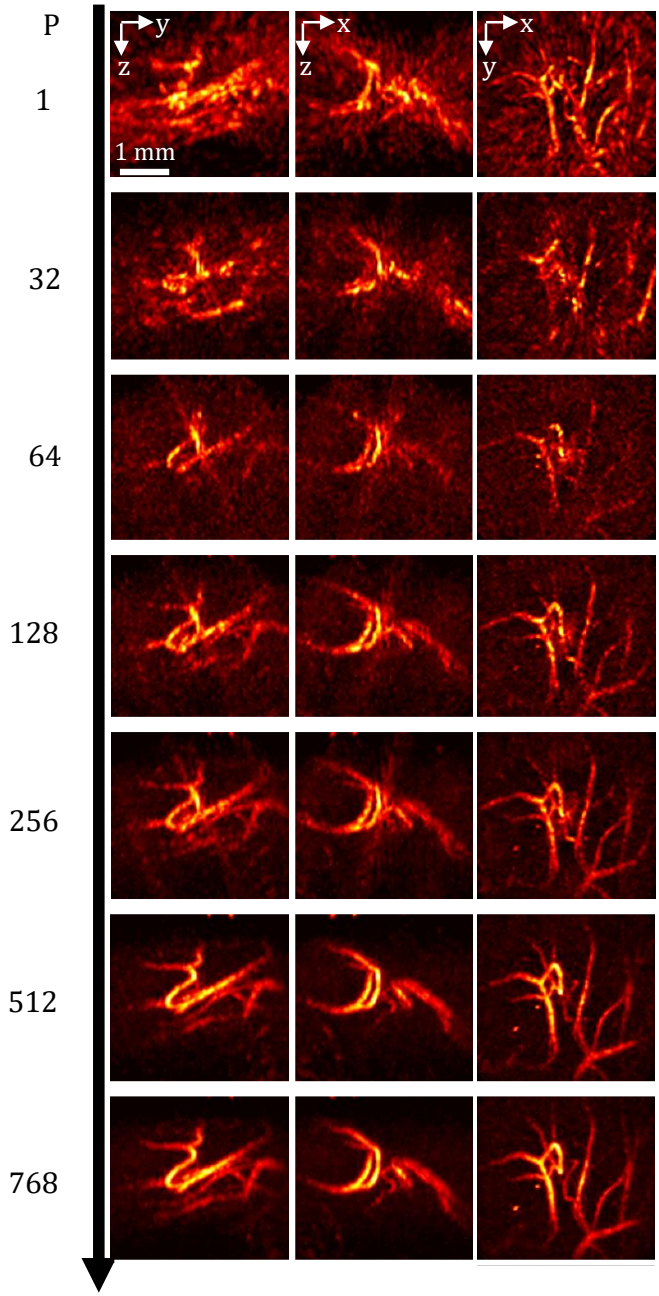

Figure 3.20 – Influence du nombre d'images P sur la qualité de l'image photoacoustique de fluctuation d'un embryon de poulet. Les intensités maximales de projection des images de fluctuation dans les plans YZ, XZ et XY sont affichées pour différentes valeurs de P

Un nombre d'images supérieur à 256 semble nécessaire pour reconstruire correctement l'image. Le compromis entre la qualité de cette image et la résolution temporelle prend ici tout son sens. L'augmentation de la résolution temporelle se fera au détriment de la qualité de l'image reconstruite. A une cadence de 100 Hz, celle dont nous disposons dans nos expériences, les 256 images sont acquises en 2,56 secondes.

3.6 Impact d'une variation tir à tir de la fluence du laser

Comme présenté dans la section 3.3, un filtrage par décomposition en valeur singulière (SVD) est en pratique indispensable car le mouvement des globules rouges n'est pas la seule source de fluctuation du signal photoacoustique. Si les origines de ces sources sont multiples et variées, nous allons illustrer l'impact de la variation tir à tir de la fluence du laser (PEF), dont nous avons constaté expérimentalement l'importance sur l'image de fluctuation.

3.6.1 Définition de l'image de fluctuation en présence d'une variation tir à tir de la fluence du laser

Notons $\Phi_k = \bar{\Phi}(1+\delta_k)$ avec $\bar{\Phi}$ la fluence moyenne du laser et δ_k la fluctuation relative à chaque tir. En reprenant l'équation 3.19, on obtient comme définition de l'image photoacoustique $A_k(\mathbf{r})$ la nouvelle expression :

$$A_k(\mathbf{r}) = (1+\delta_k)A_k^{ideal}(\mathbf{r}) + b_k(\mathbf{r}) \tag{3.24}$$

en remplaçant Φ_0 par $\bar{\Phi}$ dans la définition de $A_k^{ideal}(\mathbf{r})$. On obtient alors pour le calcul de la variance :

$$\sigma^2[A(\mathbf{r})] =< |A_k^{ideal}(\mathbf{r}) + \delta_k A_k^{ideal}(\mathbf{r}) + b_k(\mathbf{r}) - < A_k(\mathbf{r}) > |^2 > \tag{3.25}$$

La grandeur δ_k, tout comme $b_k(\mathbf{r})$ précédemment, est à moyenne nulle, $< \delta_k >_k$ peut donc également être négligée pour le calcul de la moyenne dans le cas où le nombre d'acquisition k est suffisamment grand. On peut à nouveau écrire $m[A](\mathbf{r}) =< A_k(\mathbf{r}) >_k =< A_k^{ideal}(\mathbf{r}) >_k$. L'expression devient donc :

$$\begin{aligned} \sigma^2[A(\mathbf{r})] &=< |A_k^{ideal}(\mathbf{r}) - m[A](\mathbf{r}) + \delta_k A_k^{ideal}(\mathbf{r})|^2 > + \sigma_b^2 \\ \sigma^2[A(\mathbf{r})] &=< |A_k^{ideal}(\mathbf{r}) - m[A](\mathbf{r})|^2 > + < \delta_k^2 |A_k^{ideal}(\mathbf{r})|^2 > + \sigma_b^2 \end{aligned} \tag{3.26}$$

La grandeur d'intérêt $\sigma_{flow}^2(\mathbf{r})$ est reconnaissable dans $< |A_k^{ideal}(\mathbf{r}) - m[A](\mathbf{r})|^2 >$. En nommant $\sigma_{\Phi_N}^2 =< \delta_k^2 >_k$ la variance de la fluctuation relative de la PEF, on obtient les expressions :

$$\sigma^2[A(\mathbf{r})] = \sigma_{flow}^2(\mathbf{r}) + \sigma_{\Phi_N}^2(\sigma_{flow}^2(\mathbf{r}) + |m[A](\mathbf{r})|^2) + \sigma_b^2 \tag{3.27}$$

c'est à dire :

$$\sigma^2[A(\mathbf{r})] = (1+\sigma_{\Phi_N}^2)\sigma_{flow}^2(\mathbf{r}) + \sigma_{\Phi_N}^2 |m[A](\mathbf{r})|^2 + \sigma_b^2 \tag{3.28}$$

Dans nos expérience, le terme σ_{Φ_N} est faible, de l'ordre de quelques pourcents. Si son effet dans le terme $(1+\sigma_{\Phi_N}^2)\sigma_{flow}^2(\mathbf{r})$ est insignifiant, sa présence n'est absolument pas négligeable car elle entraine la présence du second terme $\sigma_{\Phi_N}^2 |m[A](\mathbf{r})|^2$, incluant la grandeur $|m[A](\mathbf{r})|^2$ dont la valeur est largement supérieure à $\sigma_{flow}^2(\mathbf{r})$. Pour donner un ordre de grandeur, les valeurs des différents termes de l'équation 3.28 sont données dans le cas de l'une de nos expériences sur l'embryon de poulet, avec un σ_{Φ_N} typique de 3% :

$$\begin{aligned} (1+\sigma_{\Phi_N}^2)\sigma_{flow}^2(\mathbf{r}) &= 1.7 \times 10^6 a.u. \\ \sigma_{\Phi_N}^2 |m[A](\mathbf{r})|^2 &= 1.2 \times 10^7 a.u. \\ \sigma_b^2 &= 7.7 \times 10^5 a.u. \end{aligned} \tag{3.29}$$

Bien que σ_{Φ_N} soit très faible, la valeur du terme dépendant de $|m[A](\mathbf{r})|$ est conséquent, de l'ordre de 7 fois le premier terme de l'équation 3.28 incluant la fluctuations d'intérêt. La présence de la PEF provoque ainsi une fluctuation temporelle des structures initialement statiques de l'image moyenne, les faisant naturellement apparaître dans l'image de fluctuation, sous la forme du second terme $\sigma_{\Phi_N}^2 |m[A](\mathbf{r})|^2$. La présence de ce terme est problématique car les structures d'intérêts décrites par $\sigma_{flow}(\mathbf{r})$ peuvent se retrouver masquées par les structures décrites par $m[A](\mathbf{r})$. C'est notamment le cas en présence de *clutter* : cet artefact, lié au sous-échantillonnage spatial de la sonde, est principalement produit par les

structures non fluctuantes et contribue à l'image $m[A](\mathbf{r})$ sous forme de signal diffus enveloppant les vaisseaux (exemple en figure 3.10). Cette analyse met en évidence l'importance du filtrage de ce terme, en l'occurrence à l'aide d'une SVD. Ces chiffres nous montrent également l'importance du terme de bruit dans nos expériences.

3.6.2 Amplitude de la fluctuation liée au mouvement des globules rouges vis à vis du signal non fluctuant

Si le second terme de l'équation 3.28 dépend de σ_{Φ_N}, il dépend donc également du rapport entre $\sigma_{flow}(\mathbf{r})$ et $m[A](\mathbf{r})$. On peut extraire des grandeurs décrites dans la section précédente les valeurs moyennes selon $\mathbf{r}$ de $\sigma_{flow}(\mathbf{r})$ et $|m[A](\mathbf{r})|$, égale respectivement à 1.3×10^3 a.u et 1.2×10^5 a.u, soit un rapport d'environ 92, représentatif de l'une de nos expériences. Ce rapport peut également être directement calculé à partir des équations 3.5 et 3.11 selon :

$$R(\mathbf{r}) = \frac{m[A](\mathbf{r})}{\sigma_{flow}(\mathbf{r})} = \sqrt{\frac{\eta}{W(\eta)V_p}} \frac{f(\mathbf{r}) * h(\mathbf{r})}{\sqrt{f^2(\mathbf{r}) * |h|^2(\mathbf{r})}} \tag{3.30}$$

en moyennant spatialement $R(\mathbf{r})$. On peut s'intéresser à l'impact de la présence d'une structure non fluctuante (par exemple un vaisseau horizontal produisant du clutter) sur la quantification d'une fluctuation. On considère donc un vaisseau horizontal fin pour $f(\mathbf{r})$ (perpendiculaire à l'axe de la sonde, non affecté par les artefacts de visibilité) dans le calcul de $m[A](\mathbf{r})$, et un vaisseau horizontal ou vertical pour $f(\mathbf{r})$ dans le calcul de $\sigma_{flow}(\mathbf{r})$. Le rapport $R(\mathbf{r})$ dépend également de la *point spread function* (PSF) du système d'imagerie, on peut alors distinguer la cas de notre sonde d'imagerie 2D et celle d'imagerie 3D. Ces ratios sont présentés dans le tableau 3.1 en prenant pour le flux sanguin $\eta = 50\%$, $W(50\%) = 16\%$ et $V_p = 100\ \mu m^3$ et des vaisseaux de diamètre $\lambda/4$ soit 24 µm dans le cas 2D et 47 µm dans le cas 3D :

Tableau 3.1 – Rapport entre l'image moyenne et l'image de fluctuation en considérant un vaisseau horizontal pour l'image moyenne et un vaisseau vertical ou horizontal pour l'image de fluctuation. Ces rapports sont présentés pour nos sondes d'imagerie 2D et 3D.

	2D 15 MHz		3D 8 MHz	
	Horiz.	Vert.	Horiz.	Vert.
$\frac{m[A](\mathbf{r}_{horiz})}{\sigma[A](\mathbf{r}_{horiz/vert})}$	43.3	38.8	103.1	65.8

Le signal de fluctuation est 40 fois plus faible que celui de l'image moyenne à 15 MHz en imagerie 2D, et de 103 ou 66 fois en imagerie 3D à 8 MHz. Les différences observées entre les deux types de structures (verticale et horizontale) s'expliquent par l'anisotropie de la PSF dans les directions verticales et horizontales. La différence entre les deux sondes d'imagerie s'explique principalement par leurs différentes fréquences centrales. A 15 MHz, la PSF est plus fine : moins de globules rouges sont inclus dans la tâche de diffraction ce qui augmente l'amplitude de la fluctuation vis à vis du signal moyen. On observe enfin que les ratios obtenus à partir de la formule 3.30 présentés dans ce tableau pour l'imagerie 3D sont très proches de celui obtenu expérimentalement (de valeur 92).

L'expression 3.30 nous permet donc d'obtenir une information sur l'amplitude du signal de fluctuation vis à vis du signal moyen, ce qui peut être utile pour estimer la valeur du signal de fluctuation à partir de la valeur du signal moyen. En connaissance du bruit électronique, un FNR pourrait être calculé en temps réel, permettant de définir un seuil de détectabilité ainsi que le nombre d'images minimal à acquérir pour extraire les structures du bruit de fond lors d'une acquisition. Comme le montre cette expression 3.30, le rapport $R(\mathbf{r})$ dépend des caractéristiques de la sonde d'imagerie. L'équation peut donc également être

utilisée pour évaluer en amont l'intérêt pour l'imagerie de fluctuation d'une nouvelle sonde dont la PSF est connue (par simulation par exemple).

3.6.3 Mécanismes du filtrage spatio-temporel par décomposition en valeurs singulières

L'impact du filtrage par SVD sur la qualité de l'image de fluctuation est étudié au travers de simulations Simsonic, réalisées avec une fluctuation de la fluence du laser de 50% (très élevé pour "voir" son effet en simulation) que l'on suppose inconnue. Un très faible *FNR*, de 0.5 est choisi : l'écart type du bruit sur les données radiofréquences est donc deux fois supérieur à l'amplitude des fluctuations d'intérêt. Alors que les artefacts de visibilité sont présents dans l'image de fluctuation en absence de filtrage SVD (seconde ligne, d-f) car l'amplitude de $m[A]$ est de plusieurs ordres de grandeur au-dessus de σ^2_{flow}, ils disparaissent dans les images de la troisième ligne (g-i), lorsque la première valeur propre de la SVD, contenant la contribution de $m[A]$, est supprimée. Ces résultats de simulation sont similaires aux résultats expérimentaux de la figure 3.7.

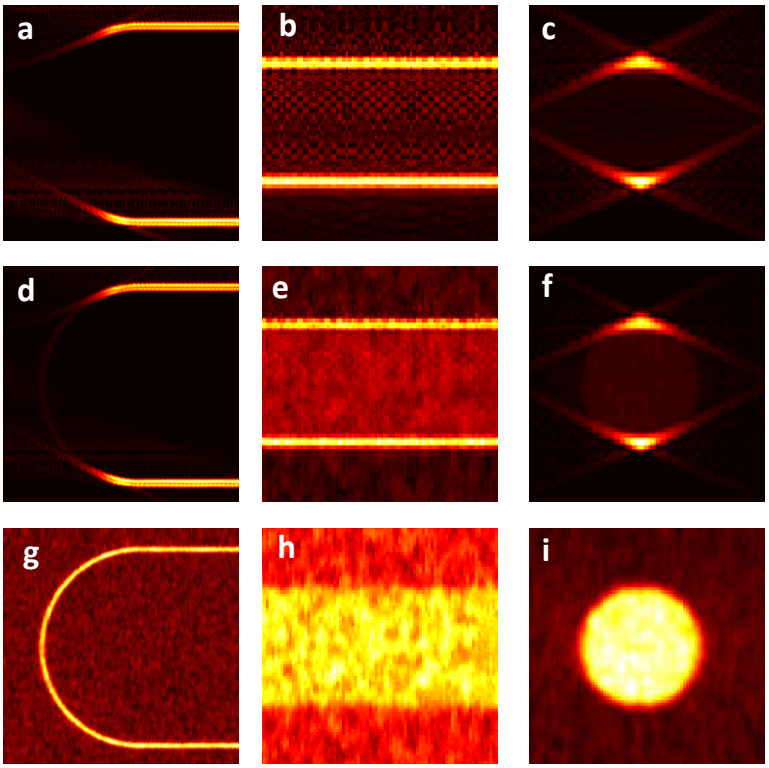

Figure 3.21 – Effet du filtrage par singular value décomposition (SVD) pour un fluctuation to noise ratio de 0.5. La première ligne (a-c) correspond à l'estimation de l'image moyenne, la seconde (d-f) et la troisième (g-i) à l'estimation de l'image de fluctuation avec et sans filtrage SVD.

La figure 3.22 permet d'appréhender l'effet de la SVD sur de telles données. La matrice de covariance des vecteurs spatiaux $U_i(\mathbf{r})$ visible en (a) permet de distinguer les familles de signaux dont les comportements spatiaux-temporels diffèrent significativement.

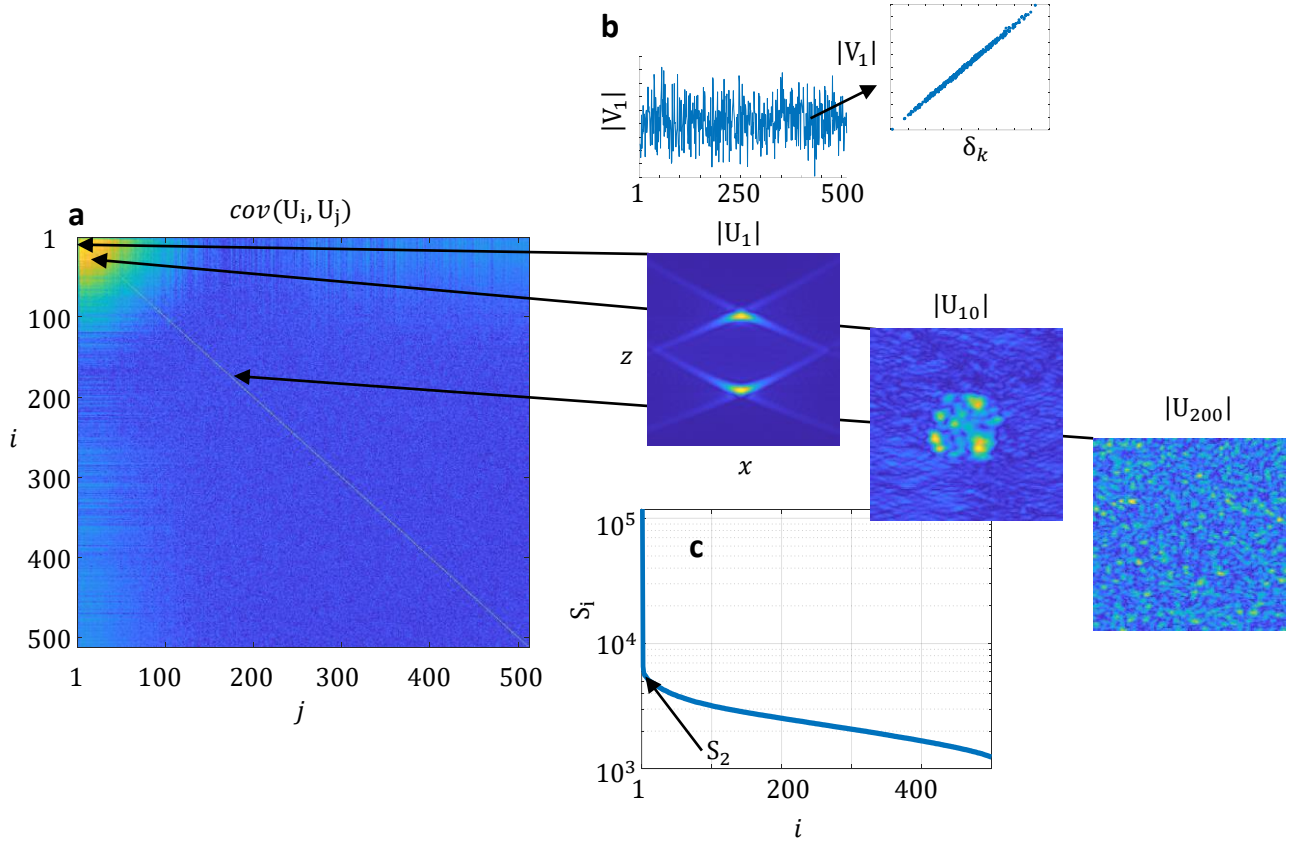

Figure 3.22 – Mécanisme et illustration de la décomposition en valeurs singulières. (a) : Matrice de covariance des vecteurs spatiaux U_i. (b) : vecteurs spatiaux $|U_{i=1}|$, $|U_{i=10}|$ et $|U_{i=200}|$ ainsi que le vecteur temporel $|V_{i=1}|$ et la représentation de sa relation linéaire avec δ_k. (c) : Valeurs propres S_i.

Dans cette simulation, trois familles se distinguent :

1. Les bords de l'objet, fluctuants principalement selon la PEF δ_k et correspondant au premier couple de vecteur. Ainsi, $|V_{i=1}|$ le vecteur décrivant la variation temporelle de cette structure, visible en (b), est proportionnelle à δ_k. C'est la structure qui correspond à l'image moyenne.
2. Le centre du disque, fluctuant principalement selon $\delta_k \times g_k$, représenté par les couples de vecteurs propres d'indice 2 à environ 100.
3. Le bruit électronique, fluctuant selon b_k, de plus faible énergie.

La courbe des valeurs propres (c) montre que l'énergie contenue dans le couple de vecteur correspondant au signal de l'image moyenne est largement supérieure à celle contenue dans les couples de vecteurs correspondant au signal de fluctuation d'intérêt. Cette figure illustre l'importance du choix des bornes de la SVD : si a est trop élevé, les vecteurs propres restants ne contiendront alors que du bruit. La section suivante présente des images de fluctuation obtenues en changeant les paramètres du filtrage par SVD, en simulation puis sur des données expérimentales.

3.6.4 Influence du choix des bornes de la décomposition en valeurs singulières sur la qualité de l'image

En simulation

Les images de fluctuation du disque pour une borne a variant de 1 à 50, et de b variant de 10 à 512 sont présentées en figure 3.23 :

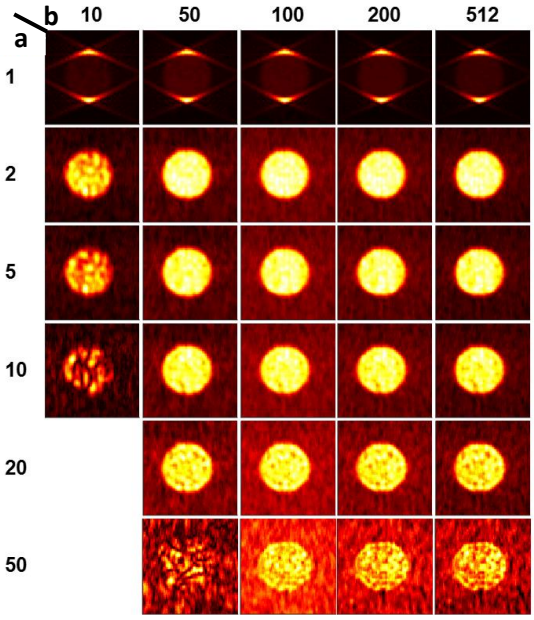

Figure 3.23 – Influence des bornes a et b du filtrage par décomposition en valeurs singulières sur la qualité de l'image reconstruite en simulation.

Il est clairement nécessaire de supprimer la première valeur propre pour faire apparaître l'objet. On note que si la borne a est trop élevée, l'objet disparaît, noyé dans le bruit. A partir de la figure 3.23, l'effet de b, la borne supérieure du filtrage par SVD, peut également être évalué. Comme présenté dans la section précédente, les derniers vecteurs propres contiennent uniquement du bruit électronique. La suppression des dernières valeurs singulières peut ainsi permettre d'améliorer la qualité de l'image en diminuant le niveau du bruit de fond électronique : l'importance relative de la fluctuation par rapport au bruit (le FNR) augmente et permet d'améliorer la visibilité des structures. Si l'effet est peu visible sur ces simulations, il est plus marqué dans d'autres situations, notamment sur certains résultats expérimentaux.

En expérience

L'impact du choix des bornes de la SVD peut également être étudié sur les données expérimentales. Contrairement aux simulations, il est nécessaire de supprimer plusieurs valeurs propres pour faire apparaître l'objet dont le nombre peut varier entre deux expériences, par exemple lorsque l'échantillon change.

Sur la coupe transversale du tube imagé avec la sonde d'imagerie 2D, l'objet semble correctement reconstruit à partir de $a = 15$ (fig. 3.24). A nouveau, si a est trop élevé, l'image est dégradée. En diminuant b, le contraste augmente très légèrement, jusqu'à ce que b soit trop faible et coupe des valeurs propres décrivant l'objet, dégradant la qualité de la reconstruction. Le CNR, calculé tel que $CNR = \frac{fluctuation\ in\ object\ -\ background\ fluctuation\ level}{spatial\ fluctuation\ of\ the\ background}$ est maximal sur une bande pour a variant de 15 à 20 et b variant de 70 à 1000, validant ces observations. Il semble possible d'utiliser le CNR pour déterminer automatiquement ces bornes : une méthode est proposée dans le chapitre 4.

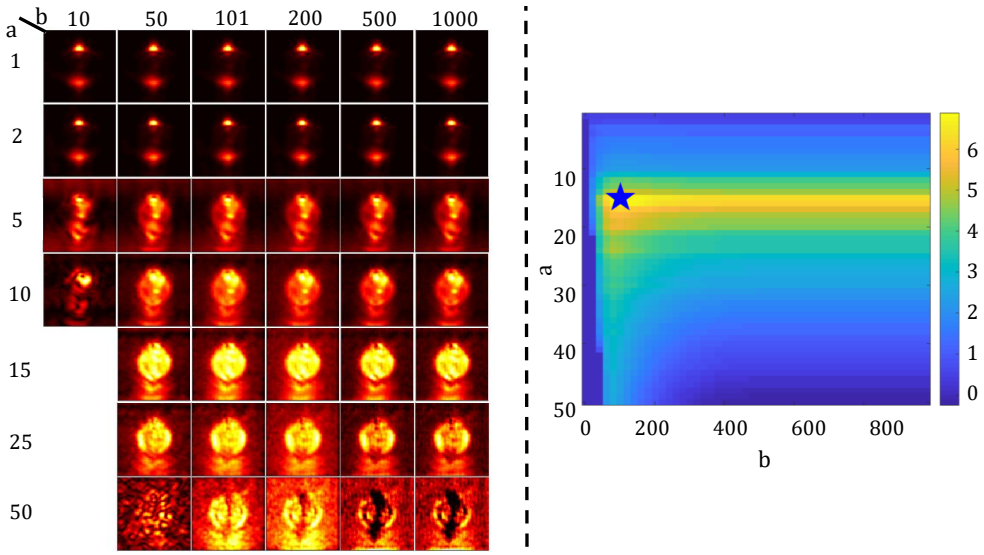

Figure 3.24 – Influence des bornes a et b du filtrage par décomposition en valeurs singulières sur la qualité de l'image de fluctuation expérimentale 2D du tube transverse au plan d'imagerie. La carte de CNR des images reconstruites pour différents couples a, b est également présentée.

Dans le cas de l'embryon imagé en 3D, les mêmes observations sont réalisées. Les structures semblent correctement reconstruites dès $a = 10$, comme le montre la figure 3.25.

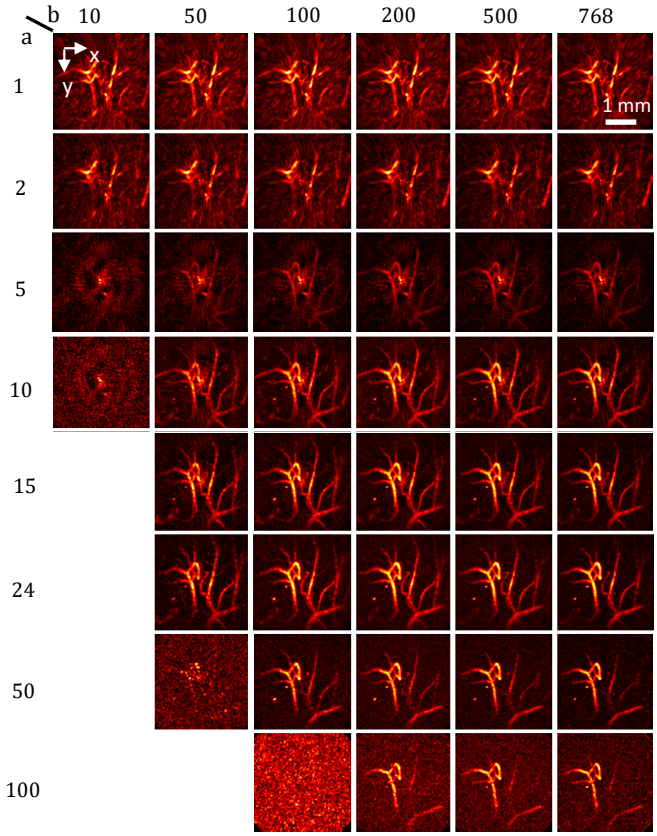

Figure 3.25 – Influence des bornes a et b du filtrage par décomposition en valeurs singulières sur la qualité de l'image de fluctuation expérimentale 3D de l'embryon de poulet. L'intensité maximale projetée sur le plan XY pour différents couples a,b est affichée.

3.7 Conclusion

Dans ce chapitre, une nouvelle méthode d'imagerie photoacoustique, appelée imagerie photoacoustique de fluctuation, a été présentée. En exploitant les fluctuations produites par le flux sanguin, celle-ci permet de corriger les artefacts de visibilité d'ouverture numérique et de bande passante limitée, et de réduire les artefacts liés au sous échantillonnage spatial en imagerie 3D. Dans un premier temps, nous avons proposé une théorie permettant de définir une expression analytique de l'image de fluctuation, et d'expliciter les mécanismes d'amélioration de la reconstruction de l'image. La faisabilité de l'approche a ensuite été démontrée *in vitro* en imagerie 3D avec notre sonde parcimonieuse présentée dans le chapitre 2. Le potentiel de l'imagerie photoacoustique de fluctuation est ensuite exploité *in vivo* en imageant en 3D le réseau vasculaire situé au niveau de la membrane chorioallantoique d'un embryon de poulet. Dans ces expériences, l'imagerie photoacoustique de fluctuation 3D est couplée à l'imagerie 3D Power Doppler, permettant d'obtenir deux images complémentaires du réseau vasculaire. Si les images sont sensiblement les mêmes en terme de morphologie, elles contiennent des informations de nature très différente, qui pourront être exploitées pour extraite des indicateurs biologiques différents comme cela est présenté dans le chapitre 4. Des résultats préliminaires d'imagerie 3D de fluctuation du cerveau d'une souris ont également été présentés, potentiellement améliorable en modifiant certains paramètres de l'algorithme de *delay-and-sum*. Enfin, nous avons réalisé des études approfondissant l'effet du bruit électronique et de la fluctuation tir à tir de la fluence du laser, et démontré l'importance du choix des bornes du filtrage par décomposition en valeurs singulières. Une étude sur simulations permettant de valider certains aspects de la théorie présentée en début de chapitre est disponible en annexe. Les nouveautés présentées dans ce chapitre et auxquelles ces travaux de thèse ont contribué sont notamment :

— Le développement d'une théorie de l'imagerie photoacoustique de fluctuation et sa validation

— La mise en oeuvre de l'imagerie photoacoustique de fluctuation en 3D avec une sonde parcimonieuse *in vitro* et *in vivo* en imagerie de l'embryon de poulet puis de manière préliminaire en imagerie de la souris

— La mise en oeuvre de l'imagerie Power Doppler en 3D avec une sonde parcimonieuse également *in vivo* en imagerie de l'embryon de poulet, réalisée simultanément à l'imagerie photoacoustique de fluctuation

— L'extension de la théorie en présence de bruit et d'une fluctuation tir à tir de la fluence, permettant les développements pour l'imagerie quantitative présentés dans le chapitre 4.

Concluons ce chapitre sur les potentielles limites de l'imagerie de fluctuation, inhérentes à son principe et fondamentalement reliées entre-elles : la résolution temporelle et la sensibilité. Si les propriétés statistiques des différents phénomènes mis en jeu sont parfaitement estimées, la fluctuation d'intérêt peut être détectée au-dessus de n'importe quel niveau du bruit (sensibilité infinie). Ce n'est bien évidemment pas le cas en pratique : la sensibilité augmente en réalité avec la puissance statistique laquelle dépend du nombre d'images à disposition, au prix donc de la résolution temporelle. La puissance de l'imagerie de fluctuation peut par ailleurs être affectée par la présence de sources de fluctuation "parasites". Si la sensibilité de l'approche peut encore être améliorée, les résultats présentés dans ce chapitre permettent de lui entrevoir un potentiel conséquent. On peut évidemment citer sa capacité à améliorer la qualité de l'image photoacoustique, mais également sa relative simplicité en terme de *post processing* et la possibilité de l'implémenter facilement sur n'importe quel dispositif d'imagerie conventionnel. Par ailleurs, nous allons voir dans le chapitre suivant que l'image de fluctuation, après certaines corrections, est proportionnelle à l'énergie absorbée et peut donc trouver des applications en imagerie photoacoustique quantitative.

4

Imagerie quantitative de fluctuation

Sommaire

4.1 Introduction **99**
- 4.1.1 Imagerie photoacoustique quantitative conventionnelle 99
- 4.1.2 Contexte et objectifs de ce chapitre 99

4.2 Théorie : l'imagerie photoacoustique quantitative de fluctuation **101**
- 4.2.1 L'imagerie de fluctuation quantitative en conditions idéales 101
- 4.2.2 Impact du bruit électronique et méthode de correction 103
- 4.2.3 Impact de la fluctuation tir à tir de la fluence du laser (PEF) et méthode de correction 104
- 4.2.4 Estimation de la saturation en oxygène en présence de bruit et de PEF 107

4.3 Imagerie photoacoustique quantitative 3D de fluctuation : imagerie d'oxygénation du sang à visibilité totale **108**
- 4.3.1 Méthode de l'imagerie photoacoustique quantitative de fluctuation en conditions expérimentales 108
- 4.3.2 Imagerie photoacoustique quantitative de fluctuation *in vitro* 110
- 4.3.3 Imagerie photoacoustique quantitative de fluctuation de l'embryon de poulet . 111

4.4 Imagerie photoacoustique quantitative de fluctuation et Doppler ultrasonore 3D simultanées **113**
- 4.4.1 Méthode de l'imagerie Doppler ultrasonore quantitative 113
- 4.4.2 Images Directional Doppler et Color Doppler 3D de l'embryon de poulet . . . 113
- 4.4.3 Analyse conjointe de la SO_2 et du flux 115

4.5 Étude sur simulation de l'impact du bruit électronique et du PEF sur l'estimation de SO_2 . . . . . **117**
- 4.5.1 Influence du bruit électronique 117
- 4.5.2 Etude : impact du filtrage par SVD sur l'estimation de SO_2, simulation 119

4.6 Estimation automatique de la borne inférieure du filtrage par SVD à partir du rapport contraste à bruit sur l'image de fluctuation **123**
- 4.6.1 Impact de la borne inférieure sur l'estimation de la SO_2 sur les données expérimentales 123
- 4.6.2 Méthode de sélection semi-automatique de la borne inférieure a 124

4.7 Conclusion **126**

4.1 Introduction

4.1.1 Imagerie photoacoustique quantitative conventionnelle

La dépendance spectrale de l'absorption optique des principaux chromophores rencontrés dans les tissus biologiques est conséquente. Il est ainsi possible, en mesurant l'absorption optique $\mu_\lambda(\mathbf{r})$ à différentes longueurs d'onde λ, de discriminer ou quantifier ces chromophores. L'imagerie photoacoustique, dont la nature du contraste provient de la distribution spatiale de $\mu_\lambda(\mathbf{r})$ dans le milieu, peut être exploitée par de telles approches multispectrales, en changeant la longueur d'onde d'illumination. Vis à vis des méthodes conventionnelles telles que la spectroscopie par transmission optique, les informations sont obtenues *in vivo* de manière localisée, et en profondeur dans les tissus biologiques. L'intérêt pour l'imagerie photoacoustique multispectrale s'est considérablement renforcé lors de l'apparition de laser accordable en longueur d'onde et modifiable à haute cadence, ouvrant la voie à des applications d'imagerie fonctionnelle [31]. Une image montrant une estimation de la concentration de chromophores peut ainsi être obtenue à l'intérieur d'un milieu diffusant. Les chromophores les plus étudiés sont des molécules endogènes telles que l'oxyhémoglobine (HbO_2) et la desoxyhémoglobine (Hbb), la mélanine ou les lipides. A partir de ces concentrations, des paramètres physiologiques d'intérêt peuvent être calculés, tels que la saturation en oxygène d'un tissu (SO_2) [23].

En imagerie photoacoustique quantitative, on recherche en général une estimation relative de $\mu_\lambda(\mathbf{r})$ à chaque λ, soit une grandeur proportionnelle à $\mu_\lambda(\mathbf{r})$ à différentes longueurs d'onde. La première étape consiste à estimer la densité d'énergie absorbée par les chromophores du milieu $\Phi_\lambda(\mathbf{r}) \times \mu_\lambda(\mathbf{r})$. L'imagerie photoacoustique conventionnelle peut en fournir une mesure relative, comme le montre l'équation 3.4 du chapitre 3. La seconde étape consiste à isoler $\mu_\lambda(\mathbf{r})$. Il est donc nécessaire de connaître la fluence déposée par le laser $\Phi_\lambda(\mathbf{r})$. En effet, le faisceau étant généralement atténué lors de sa traversée du milieu imagé, cette grandeur dépend des propriétés du milieu et de la position $\mathbf{r}$. La forte dépendance spectrale des coefficients d'absorption et de diffusion des tissus biologiques soulève une difficulté majeure : l'atténuation du faisceau est différente en fonction de la longueur d'onde d'illumination, c'est l'effet de coloration spectrale. L'estimation de $\Phi_\lambda(\mathbf{r})$ reste un problème ouvert, et un sujet d'étude important en imagerie photoacoustique. Les principales méthodes développées se basent sur la modélisation de la propagation de la lumière [63,85,136], via des approches analytiques [4,66,123] et des approches numériques de type Monte-carlo [135]. Une fois μ_λ obtenu, des analyses quantitatives peuvent être menées, telles que l'estimation de la SO_2 [23].

4.1.2 Contexte et objectifs de ce chapitre

Le principal objectif de ce chapitre est de démontrer que l'image photoacoustique de fluctuation est proportionnelle à la densité d'énergie absorbée $\Phi_\lambda(\mathbf{r}) \times \mu_\lambda(\mathbf{r})$, et donc que l'imagerie photoacoustique *quantitative* de fluctuation (*quantitative photoacoustic fluctuation imaging*, qPAFi) fournit la même information que l'imagerie photoacoustique quantitative conventionnelle. Néanmoins, cette information y est disponible sur une plus grande variété de structures, puisque l'imagerie de fluctuation permet de corriger les artefacts de visibilité et également d'augmenter le contraste. L'intérêt est majeur, au vue de l'importance de ces artefacts sur les images photoacoustiques conventionnelles. En absence de bruit électronique et de fluctuation tir à tir de la fluence du laser (*pulse energy fluctuation*, PEF), l'image de fluctuation est directement proportionnelle à l'énergie absorbée. En conditions réelles d'imagerie, il est nécessaire de tenir compte de ces deux éléments. La première partie de ce chapitre décrit en détail leur impact sur l'estimation de l'énergie absorbée et les méthodes de correction développées au cours de cette thèse.

La réalisation de la seconde étape de l'imagerie photoacoustique quantitative, l'extraction de $\mu_\lambda(\mathbf{r})$,

n'est pas l'objet de nos travaux : on suppose que toutes méthodes permettant d'estimer la fluence $\Phi_\lambda(\mathbf{r})$ en imagerie photoacoustique quantitative conventionnelle est applicable à l'image de fluctuation une fois celle-ci proportionnelle à l'énergie absorbée. Dans notre étude, nous considérons le cas simple d'un milieu quasi transparent dans lequel les effets de coloration spectrale peuvent être négligés. Il est important de mentionner ici que si cette hypothèse est valide dans les expériences *in vitro* et dans une moindre mesure sur embryon de poulet, elle ne l'est pas dans le cas général et notamment lorsqu'une souris est imagée, le faisceau traversant de multiples couches de tissu biologique absorbant et diffusant. Néanmoins, l'objectif de cette étude n'est pas de développer une méthode *quantitative* pour obtenir $\mu_\lambda(\mathbf{r})$, mais bien de la densité d'énergie absorbée $\Phi_\lambda(\mathbf{r}) \times \mu_\lambda(\mathbf{r})$.

Nous illustrons l'intérêt de l'imagerie photoacoustique quantitative de fluctuation en estimant la SO_2 à partir d'une méthode d'inversion linéaire. On appelle SO_2 le pourcentage de HbO_2 rapporté à la quantité totale d'hémoglobine Hb_{tot}. Nous nous efforcerons dans ce chapitre de démontrer que l'imagerie de fluctuation permet d'obtenir cette information sur les structures affectées ou non par les artefacts de visibilité. Il est intéressant de noter que l'imagerie quantitative de fluctuation semble particulièrement adaptée à l'estimation de la SO_2. Les méthodes de séparation spectrale nécessitent pour la plupart une connaissance *a priori* des chromophores présents dans l'échantillon. Pour réduire la complexité des systèmes de résolution, la présence des chromophores autres que l'hémoglobine, tels que la mélanine ou l'eau, est souvent négligée, de par leurs plus faibles absorptions optiques dans les gammes de longueur d'onde utilisées en imagerie des vaisseaux sanguins, et donc leurs moindres contributions aux signaux PA (voir figure 1.14 du chapitre 1). Cette hypothèse peut cependant être une source d'imprécision pour l'estimation de la SO_2. Puisque l'imagerie photoacoustique de fluctuation est uniquement sensible aux globules rouges en mouvement dans le sang, le signal en provenance des autres chromophores est naturellement filtré, rendant cette hypothèse bien plus robuste.

Dans un second temps, l'approche qPAFI est implémentée sur notre dispositif d'imagerie 3D, et l'imagerie Doppler sera réalisée simultanément. En plus de l'image Power Doppler présentée dans le chapitre précédent, des analyses de flux et de pulsatilité seront également réalisées à l'aide de la méthode Pulsed Doppler, et des images Color Doppler et Directional Doppler seront notamment présentées et étudiées. Nous proposerons enfin des études sur simulation détaillant les effets du bruit électronique, de la PEF et du filtrage par décomposition en valeurs singulières (SVD). Ces études mettent notamment en évidence l'impact du filtrage par SVD et du choix des bornes, amenant au développement d'une méthode pour déterminer ces dernières automatiquement. Cette méthode est présentée à la fin de ce chapitre.

4.2 Théorie : l'imagerie photoacoustique quantitative de fluctuation

L'imagerie photoacoustique de fluctuation a été introduite dans le chapitre 3. Nous y avons démontré sa capacité à supprimer les artefacts de visibilité et à réduire les artefacts générés par l'utilisation d'une sonde parcimonieuse tels que le *clutter*. Le chapitre 3 inclue également une description de l'impact du bruit électronique et de la présence de sources de fluctuation autres que le mouvement des globules rouges, telles que la PEF. Le formalisme qui y est employé est réutilisé tout au long de ce chapitre 4. Dans cette première partie, nous proposons une théorie de l'imagerie photoacoustique *quantitative* de fluctuation et décrivons notamment les corrections à appliquer en présence de bruit et de PEF. Des simulations accompagnent la démonstration, à partir desquelles la SO_2 est estimée.

4.2.1 L'imagerie de fluctuation quantitative en conditions idéales

Définition de l'image de fluctuation multi-longueurs d'onde

Reprenons l'équation 3.8 de l'image de fluctuation du chapitre 3 en ajoutant une dépendance en $\mathbf{r}$ de la fluence du laser et du coefficient d'absorption optique. On suppose ici que ces deux grandeurs varient lentement en fonction de $\mathbf{r}$ vis à vis de la *point spread function* (PSF) de la sonde d'imagerie et de $g_k(\mathbf{r})$. La fonction $g_k(\mathbf{r})$, décrite dans le chapitre 3 section 3.2.1, représente la répartition aléatoire des globules rouges dans les vaisseaux sanguins à chaque acquisition k. Le coefficient d'absorption dépend également de la longueur d'onde λ, et s'écrit donc $\mu_\lambda(\mathbf{r})$. Il s'agit ici du coefficient d'absorption local de l'hémoglobine pure à un certain état d'oxygénation. On définit la fluence qui varie également en fonction de λ, et en fonction de k en présence de PEF, selon $\Phi_{k,\lambda}(\mathbf{r}) = \overline{\Phi}_\lambda(\mathbf{r})(1+\delta_{k,\lambda})$ avec $\overline{\Phi}_\lambda(\mathbf{r})$ la fluence moyenne du laser qui dépend de λ et fluctue tir à tir, et $\delta_{k,\lambda}$ la PEF (pulse energy fluctuation) qui mesure la fluctuation relative tir à tir. On suppose ici une dépendance de la PEF à λ, dépendance que l'on retrouve en conditions expérimentales avec notre laser.

Considérons dans un premier temps la situation pour laquelle le bruit électronique est nul, tout comme la PEF. La fluence est donc constante entre chaque acquisition à une même longueur d'onde, c'est à dire que $\delta_{k,\lambda} = 0$ et donc $\Phi_{k,\lambda}(\mathbf{r}) = \overline{\Phi}_\lambda(\mathbf{r})$. En modifiant l'équation 3.8 du chapitre 3, on obtient :

$$\sigma_\lambda[A^{ideal}](\mathbf{r}) = \Gamma\mu_\lambda(\mathbf{r})\overline{\Phi}_\lambda(\mathbf{r})\sigma_g\sqrt{V_g}\sqrt{f^2(\mathbf{r}) * |h|^2(\mathbf{r})} \tag{4.1}$$

Comme dans le chapitre 3, A_k^{ideal} est l'image PA à chaque acquisition k en conditions idéales (absence de bruit, de PEF ...). On peut définir ici $\sigma_{\lambda,flow}(\mathbf{r})$ tel que $\sigma_\lambda[A^{ideal}](\mathbf{r}) = \overline{\Phi}_\lambda(\mathbf{r})\sigma_{\lambda,flow}(\mathbf{r})$, c'est à dire :

$$\sigma_{\lambda,flow}(\mathbf{r}) = \Gamma\mu_\lambda(\mathbf{r})\sigma_g\sqrt{V_g}\sqrt{f^2(\mathbf{r}) * |h|^2(\mathbf{r})} \tag{4.2}$$

C'est une définition légèrement différente de celle de $\sigma_{flow}(\mathbf{r})$ du chapitre précédent, de par les dépendances en $\mathbf{r}$ et λ, et de l'exclusion du terme $\overline{\Phi}_\lambda(\mathbf{r})$ de l'expression. Cette nouvelle définition est introduite pour améliorer la lisibilité des démonstrations qui vont suivre.

Correction du terme de fluence moyenne

Dans ces conditions, l'image de fluctuation, au même titre que l'image PA conventionnelle, est proportionnelle à la densité d'énergie absorbée. La correction de la fluence $\overline{\Phi}_\lambda(\mathbf{r})$ doit être réalisée pour estimer la SO_2. Dans notre étude, nous négligeons les effets de coloration spectrale, ce qui permet d'écrire $\overline{\Phi}_\lambda(\mathbf{r}) = \overline{\Phi}_\lambda(\mathbf{0})\zeta(\mathbf{r})$ avec $\overline{\Phi}_\lambda(\mathbf{0})$ la fluence moyenne à la surface de l'échantillon et $\zeta(\mathbf{r})$ une carte d'atténuation indépendante de λ. Le terme $\overline{\Phi}_\lambda(\mathbf{0})$ peut être mesuré à l'aide d'une photodiode calibrée comme celle présentée dans le chapitre 2. Une fois corrigées par l'estimation de $\overline{\Phi}_\lambda(\mathbf{0})$, les images de fluctuation sont alors proportionnelles aux $\mu_\lambda(\mathbf{r})$, et la SO_2 peut être estimée. Cette correction est réalisée implicitement pour tous les résultats présentés dans cette section. La proportionnalité des images de fluctuation aux densités d'énergie absorbée sera donc illustrée par leur proportionnalité aux coefficients d'absorption.

Condition de simulations et estimation de la saturation en oxygène

Des simulations 2D similaires à celles présentées dans le chapitre 3 sont réalisées. On génère 5 séries d'acquisitions à plusieurs longueurs d'onde de signaux radiofréquences produits par un disque dans le plan d'imagerie d'une sonde similaire à notre sonde d'imagerie 2D. Ce disque, de 1 mm de diamètre, est rempli de particules. Pour chaque acquisition k, on génère une distribution aléatoire de ces particules de taille 10 µm × 10 µm dans un milieu de fraction volumique $\eta = 90\%$. Le coefficient d'absorption $\mu_{true,\lambda}$ est modifié à chaque série d'acquisitions pour simuler un changement de λ, choisies telles que λ_1 =680 nm, λ_2 =720 nm, λ_3 =760 nm, λ_4 =800 nm et λ_5 =840 nm. Jusque 512 acquisitions sont réalisées pour chaque λ, soit 2560 acquisitions au total. On note N le nombre d'acquisitions par λ, M le nombre de λ et P le nombre total d'images. Les coefficients d'absorption choisis attribuent artificiellement une oxygénation au contenu du disque, de 95% pour le cas présenté en figure 4.1, proche de l'oxygénation du sang dans une artère. Puisque les effets de coloration spectrale ne sont pas étudiés, l'atténuation du milieu est constante vis à vis de λ (en réalité nulle dans ces simulations, soit $\zeta(\mathbf{r}) = 1$). Une variation de la fluence moyenne à chaque longueur d'onde $\overline{\Phi}_\lambda(\mathbf{0})$ est simulée en multipliant les signaux radiofréquences par des facteurs variant entre 1 et 3, supposés connus. Les images sont reconstruites sur une grille de pas 20 µm par *delay-and-sum*. L'image de fluctuation est estimée indépendamment pour chaque λ sur les séries de $N = 512$ images : on obtient ainsi $M = 5$ images de fluctuation, comme illustré dans la figure 4.1 (a).

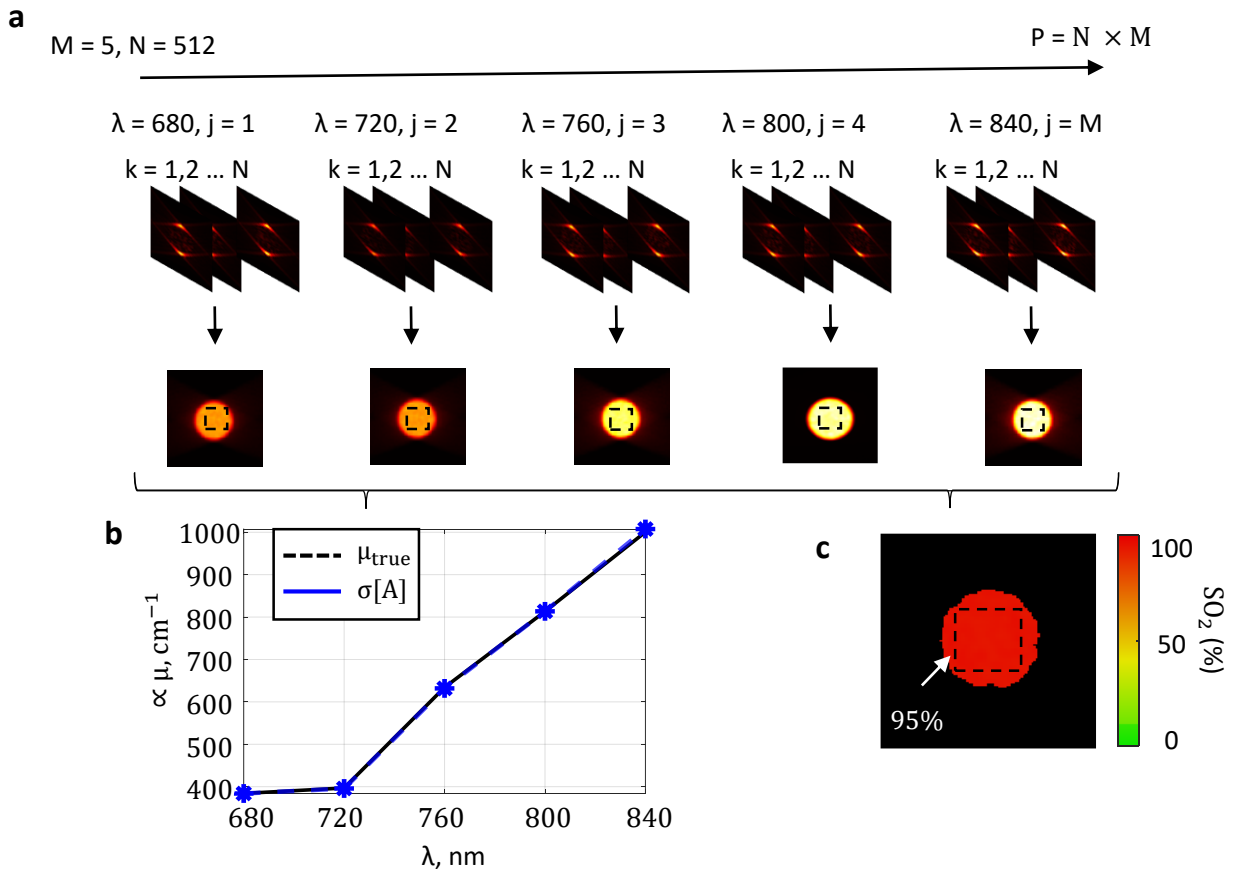

Figure 4.1 – Principe de l'imagerie photoacoustique quantitative de fluctuation .
(a) : N=512 acquisitions et M=5 longueurs d'onde différentes λ_j soit P = 2560 images représentant chacune une distribution aléatoire de particules. La variance des N images est calculée pour chaque λ_j. (b) : la courbe noire représente l'absorption de référence μ_{true}, la courbe bleue représente l'absorption estimée à partir des images de fluctuation, obtenue en moyennant les valeurs des pixels dans la ROI (rectangle pointillé) pour chaque λ_j. (c) : carte de SO_2 obtenue par l'inversion linéaire en chaque pixel. La SO_2 moyennée dans la ROI est de 95%, comme attendu.

La lettre j correspond à l'indexation selon λ. Pour valider la proportionnalité des images de fluctuation $\sigma_\lambda[A]$ aux coefficients d'absorption optique, les valeurs des pixels inclus dans une ROI (correspondant au centre du tube, rectangle pointillé sur la figure 4.1.b) sont moyennées à chaque longueur d'onde pour former la courbe bleue. Cette courbe est parfaitement superposée à la courbe noire qui représente les coefficients d'absorption de référence $\mu_{true,j}$ choisis lors de la simulation des données. A noter que la courbe bleue est divisée par la moyenne du rapport entre les deux courbes, on retrouve donc bien une quantité proportionnelle à l'absorption optique, et non pas les valeurs absolues de ces coefficients.

Pour estimer la SO_2, on considère que l'absorption optique $\mu_\lambda(\mathbf{r})$ en chaque position du milieu dépend uniquement de HbO_2 et Hbb (on néglige la présence d'autres chromophores) selon la combinaison linéaire :

$$\mu_\lambda(\mathbf{r}) = [Hb_{tot}](\mathbf{r}).(\frac{[HbO_2](\mathbf{r})}{[Hb_{tot}](\mathbf{r})}\mu_{\lambda,HbO_2} + \frac{[Hbb](\mathbf{r})}{[Hb_{tot}](\mathbf{r})}\mu_{\lambda,Hbb}) \tag{4.3}$$

avec $[HbX]$ les concentrations en g/L et $\mu_{\lambda,X}$ en cm^{-1}/(g/L). En écrivant $[Hb_{tot}](\mathbf{r}) = [HbO_2](\mathbf{r}) + [Hbb](\mathbf{r})$, on peut définir la SO_2 du tissu en chaque pixel avec le paramètre α, de valeur 0 pour un vaisseau contenant uniquement Hbb, et de valeur 1 pour pour un vaisseau contenant uniquement HbO_2 :

$$\mu_\lambda(\mathbf{r}) = [Hb_{tot}](\mathbf{r}).(\alpha(\mathbf{r})\mu_{\lambda,HbO_2} + (1-\alpha(\mathbf{r}))\mu_{\lambda,Hbb}) \tag{4.4}$$

L'image de SO_2, présentée en figure 4.1 (c) peut alors être calculée à partir du modèle d'inversion linéaire suivant :

$$\sigma_{\lambda,model}(\mathbf{r}) = K(\mathbf{r})[\alpha(\mathbf{r})\mu_{\lambda,HbO_2} + (1-\alpha(\mathbf{r}))\mu_{\lambda,Hbb}] \tag{4.5}$$

En notant $K(\mathbf{r}) = \Gamma\zeta(\mathbf{r})[Hb_{tot}](\mathbf{r})\sigma_g\sqrt{V_g}\sqrt{f^2(\mathbf{r}) * |h|^2(\mathbf{r})}$. Ce modèle est ajusté aux données $\sigma_\lambda[A](\mathbf{r})$ avec deux inconnues $\alpha(\mathbf{r})$ et $K(\mathbf{r})$, déterminées au travers de la minimisation de :

$$[\alpha_{estim}(\mathbf{r}), K_{estim}(\mathbf{r})] = \underset{\alpha(\mathbf{r}),K(\mathbf{r})}{argmin} \sum_{j=1}^{M_j} \left\|\sigma_{j,model}(\mathbf{r}) - K(\mathbf{r})(\alpha(\mathbf{r})\mu_{j,HbO_2} + (1-\alpha(\mathbf{r}))\mu_{j,Hbb})\right\|_2^2 \tag{4.6}$$

4.2.2 Impact du bruit électronique et méthode de correction

En présence de bruit électronique, $\sigma_\lambda[A]$ n'est plus directement proportionnelle à $\Phi_\lambda(\mathbf{r})\mu_\lambda(\mathbf{r})$. En reprenant l'équation 3.21 du chapitre 3, et comme b_k est indépendant de λ, on obtient directement :

$$\sigma_\lambda[A](\mathbf{r}) = \sqrt{\overline{\Phi}_\lambda^2(\mathbf{r})\sigma_{\lambda,flow}^2(\mathbf{r}) + \sigma_b^2} \tag{4.7}$$

Il est donc nécessaire de supprimer σ_b^2 pour obtenir une mesure quantitative. Une manière d'estimer σ_b^2 est de mesurer $\sigma_{b,RF}$ l'écart type de signaux radiofréquences de bruit pur, par exemple en amont d'une acquisition. Le terme σ_b^2 peut alors être estimé à partir de la formule :

$$\sigma_b^2 = 2.N_{el}.\sigma_{b,RF}^2 \tag{4.8}$$

La multiplication par N_{el}, le nombre de capteurs, est liée à l'opération de *delay-and-sum* et le facteur deux à l'utilisation de signaux complexes pour la reconstruction. La valeur σ_b^2 peut alors être retirée à l'image de variance $\sigma_\lambda^2[A](\mathbf{r})$.

Le caractère indispensable de cette correction est mis en évidence par la simulation présentée en figure 4.14, de paramètres FNR = 0.2, N = 512 images, avec FNR le *fluctuation to noise ratio*. L'une des images de fluctuation (λ = 680 nm) est présentée en (a) et les cartes de SO_2 en (b). La carte de SO_2 encadrée en bleu correspond à l'estimation en corrigeant le bruit, celle encadrée en rouge à l'estimation sans correction du bruit. La courbe noire représente l'absorption de référence, la courbe rouge l'estimation des coefficients d'absorption à partir des images de fluctuation non corrigées, et la courbe bleue cette estimation à partir des images de fluctuation corrigées.

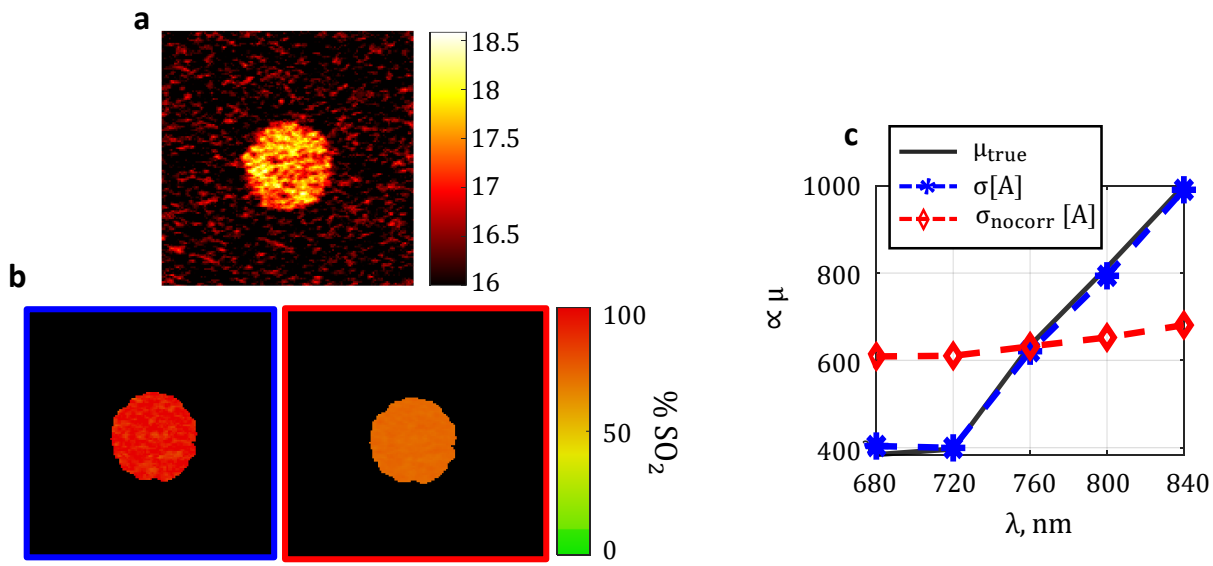

Figure 4.2 – Effet de la correction de bruit.
(a) : image de fluctuation λ = 680 nm, la barre de couleur est ajustées pour que le noir corresponde au minimum du bruit et le jaune au niveau moyen de l'objet. (b) : cartes de SO_2, en bleu estimation à partir des images de fluctuation corrigées du bruit, en rouge estimation à partir des images non corrigées. (c) : coefficients d'absorption de référence en noir, estimation des coefficients d'absorption à partir des images de fluctuation non corrigées en rouge et à partir des images de fluctuation corrigées en bleu. La SO_2 moyenne obtenue à partir des images corrigées est de $95.3 \pm 6.3\%$, contre $72.7 \pm 6.3\%$ à partir des images non corrigées.

Les images en (b) ainsi que les courbes en (c) montrent que les images de fluctuation non corrigées sont biaisées, mais que l'on obtient une estimation correcte de l'absorption optique et de la SO_2 lorsque le bruit est supprimé. La SO_2 moyenne obtenue à partir des images corrigées est ainsi de $95.3 \pm 6.3\%$, contre $72.7 \pm 6.3\%$ à partir des images non corrigées. La qualité de l'estimation de la SO_2 dépendra fortement de la qualité des estimations de σ_b^2 et $\sigma_{flow}^2(\mathbf{r})$, et donc du nombre d'images disponibles à chaque λ et du FNR. Une étude de l'impact de ces paramètres est disponible en partie 4.5. Un compromis entre nombre d'images (donc temps d'acquisition), qualité de la reconstruction et précision de l'estimation de la SO_2 devra donc être trouvé.

4.2.3 Impact de la fluctuation tir à tir de la fluence du laser (PEF) et méthode de correction

Définition de l'image de fluctuation en présence de PEF

Comme présenté dans le chapitre précédent, g_k n'est pas la seule source de fluctuation du signal photoacoustique. L'impact de la PEF, la fluctuation tir à tir de l'énergie délivrée par le laser, est à nouveau étudiée. Par rapport aux sections précédentes de ce chapitre, $\delta_{k,\lambda}$ est non nul et donc $\Phi_{k,\lambda}(\mathbf{r}) = \overline{\Phi}_\lambda(\mathbf{r})(1+\delta_{k,\lambda})$. Notons $\sigma_{\Phi,\lambda}^2 =< \delta_{k,\lambda}^2 >_k$ la variance de PEF à chaque λ. En modifiant l'équation 3.19, l'image photoacoustique à chaque longueur d'onde s'écrit :

$$A_{k,\lambda}(\mathbf{r}) = (1+\delta_{k,\lambda})A_{k,\lambda}^{ideal}(\mathbf{r}) + b_k(\mathbf{r}) \tag{4.9}$$

Les images de variance à chaque longueur d'onde sont alors obtenues en modifiant l'équation 3.28 du chapitre 3 :

$$\sigma_\lambda^2[A_{k,\lambda}(\mathbf{r})] = (1+\sigma_{\Phi_N,\lambda}^2)\overline{\Phi}_\lambda^2(\mathbf{r})\sigma_{\lambda,flow}^2(\mathbf{r}) + \sigma_{\Phi_N,\lambda}^2|m_\lambda(\mathbf{r})|^2 + \sigma_b^2 \tag{4.10}$$

où $m_\lambda(\mathbf{r}) =< [A_{k,\lambda}] >_k$ est l'image photoacoustique moyenne à chaque λ. La présence de PEF a deux impacts distincts :

1. Un terme dépendant de $m_\lambda(\mathbf{r})$ est présent dans la définition de l'image de fluctuation. En plus de masquer les structures à visibilité limitée comme décrit dans le chapitre 3, sa présence biaise la proportionnalité de l'image de fluctuation à l'énergie absorbée.
2. Le terme $(1+\sigma^2_{\Phi_N,\lambda})$ s'ajoute en facteur de $\overline{\Phi}^2_\lambda(\mathbf{r})\sigma^2_{\lambda,flow}(\mathbf{r})$, la grandeur d'intérêt dont la racine est proportionnelle à l'énergie absorbée.

Ces grandeurs ont été calculées pour l'expérience embryon de poulet dont les images de fluctuation sont visibles en figure 4.8, et ont été présentées dans le chapitre 3 (expression 3.29). Si la magnitude du terme $(1+\sigma^2_{\Phi_N,\lambda})$ est négligeable, la magnitude de $\sigma^2_{\Phi_N,\lambda}|m_\lambda(\mathbf{r})|^2$ (un ordre de grandeur supérieur au terme d'intérêt) est suffisante pour empêcher toute analyse quantitative. En plus de supprimer le bruit, il est donc nécessaire de supprimer ce second terme de l'équation 4.10.

Méthode de correction de la PEF

Décomposition en valeurs singulières

Comme présenté dans le chapitre précédent, un filtrage par décomposition en valeurs singulières (SVD) est réalisé. On considère maintenant P images photoacoustiques. Le filtrage est réalisé sur l'ensemble des P images photoacoustiques, indépendamment de λ. Il aurait été possible de réaliser un filtrage indépendant pour chaque pile de N images à chaque λ, mais cela aurait diminué le pouvoir discriminant de la SVD, lequel augmente avec le nombre d'images. L'expression de l'image photoacoustique devient alors :

$$\begin{aligned} A_{k,\lambda}(\mathbf{r}) &= \sum_{i=1}^{P} S_i U_i(\mathbf{r}) V_i^*(k) \\ A^{SVD}_{k,\lambda}(\mathbf{r}) &= \sum_{i=a}^{b} S_i U_i(\mathbf{r}) V_i^*(k) \end{aligned} \tag{4.11}$$

La pile d'images filtrées est appelée $A^{SVD}_{k,\lambda}(\mathbf{r})$. La SVD identifie dans les valeurs propres élevées les termes qui présente une forte ressemblance spatio-temporelle : la PEF conduit uniquement à une fluctuation temporelle, chaque image fluctuant tir à tir sans changement de structure spatiale, contrairement à la fluctuation d'intérêt qui elle fluctue aussi spatialement. Si la borne a est judicieusement choisie, on suppose que le terme dépendant de $m_\lambda(\mathbf{r})$ est supprimé au cours du filtrage SVD. Les images de variance sont ensuite calculées sur les piles d'images PA filtrées à chaque λ, et on obtient l'expression :

$$\sigma^2_\lambda[A^{SVD}_{k,\lambda}(\mathbf{r})] = \overline{\Phi}^2_\lambda(\mathbf{r})(1+\sigma^2_{\Phi_N,\lambda})\sigma^2_{\lambda,flow}(\mathbf{r}) + \sigma^2_{b,SVD} \tag{4.12}$$

Le filtrage par la borne haute b provoque principalement une diminution du bruit électronique. Le terme $\sigma^2_{b,SVD}$ tient compte de l'impact du filtrage par SVD sur la quantité de bruit présente dans l'image de variance. En effet, le bruit électronique se répartit sur l'ensemble des valeurs singulières, et est majoritairement décrit par les valeurs les moins énergétiques. Ainsi, si la borne b est différente du nombre d'images P (c'est à dire que des vecteurs propres de basse énergie sont supprimés au cours du filtrage), une partie du bruit présent dans les données est supprimée : l'estimation de σ_b à partir de la formule 4.8 n'est plus valide. Il est alors nécessaire de déterminer $\sigma_{b,SVD}$, le bruit restant dans l'image de fluctuation.

Correction de l'effet du filtrage par décomposition en valeurs singulières sur le bruit

Rappelons que la SVD d'une pile de P images agit tel que les vecteurs spatiaux $U_i(\mathbf{r})$ et temporels $V_i(k)$ sont unitaires. On s'attend alors à :

$$E[|U_i(\mathbf{r})|^2] = \frac{1}{N_r} \tag{4.13}$$

$E[|U_i(\mathbf{r})|^2]$ étant l'espérance selon $\mathbf{r}$ de la magnitude au carré de $U_i(\mathbf{r})$ et $N_r = N_x \times N_z$ dans le cas 2D ou $N_r = N_x \times N_y \times N_z$ dans le cas 3D. On s'attend aussi à :

$$E[|V_i(k)|^2] = \frac{1}{P} \tag{4.14}$$

$E[|V_i(k)|^2]$ l'espérance selon k de la magnitude au carré de $V_i(k)$.

En supposant que les vecteurs spatiaux de $b+1$ à P soient uniformément remplis de bruit aléatoire et que le bruit soit homogène vis à vis de la longueur d'onde, retirer les grandeurs $\alpha_i U_i(\mathbf{r}) V_i^*(k)$ avec i de $b+1$ à P revient à retirer une quantité de bruit de $\frac{1}{P \times N_r} \sum_{i=b+1}^{P} \alpha_i^2$, avec α_i la valeur singulière i. On obtient finalement :

$$\sigma_{b,SVD}^2 = \sigma_b^2 - \frac{1}{PN_r} \sum_{i=b+1}^{P} \alpha_i^2 \tag{4.15}$$

Si l'expression 4.14 est bien vérifiée dans nos simulations, la validité de l'expression 4.13 peut être questionnée. La figure 4.3 présente $|U_i(\mathbf{r})|^2$, les magnitudes au carré des vecteurs spatiaux U_i calculées sur une zone de bruit (en rouge) et une zone correspondant à l'objet (en bleue), pour la simulation de la section 4.2.2 (en présence d'un fort bruit électronique).

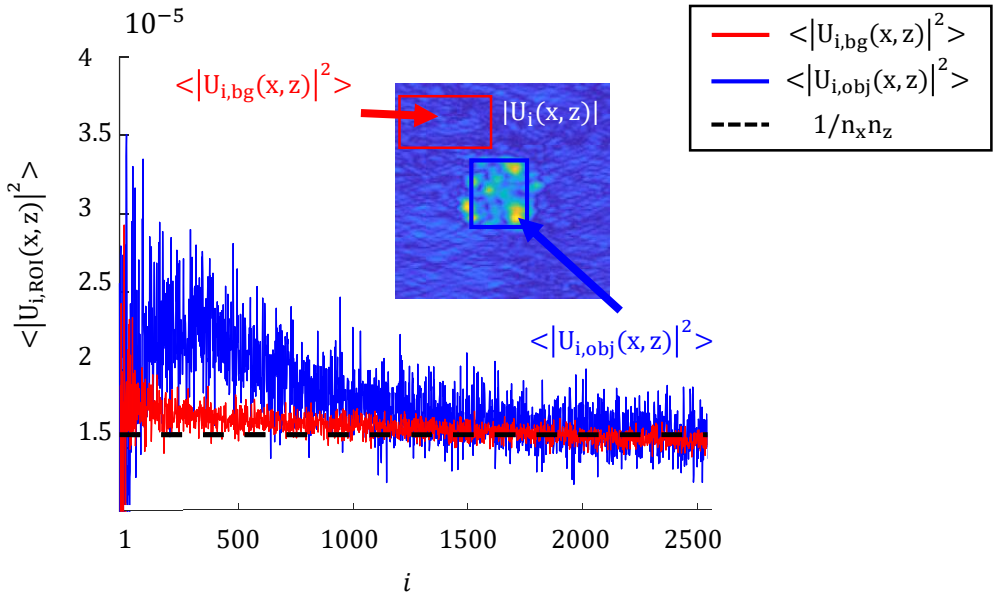

Figure 4.3 – Magnitude au carré de vecteurs spatiaux U_i moyennés sur une zone de bruit (courbe rouge) et au centre de l'objet (courbe bleue). Ces vecteurs résultent de la décomposition en valeurs singulières de P images photoacoustiques simulées en présence d'un fort bruit électronique.

Les valeurs de $|U_i(\mathbf{r})|^2$ sur la zone de de bruit et sur la zone correspondant à l'objet oscillent comme attendu autour de $1/n_x n_z$ (représentée par la droite en pointillé noir) à partir d'environ i = 1000. La fluctuation entre les différents vecteurs propres est conséquente, et une légère tendance négative est observable. Pour i inférieur à 1000, l'hypothèse 4.13 n'est plus valide. L'égalité 4.15 est donc une bonne approximation du bruit restant sur l'image de fluctuation lorsque la borne b est suffisamment élevée. Comme le montre la figure 4.4, la prise en compte de cette réduction du bruit permet de largement améliorer la proportionnalité de l'image de fluctuation à la densité d'énergie absorbée, et ainsi l'estimation de la SO_2.

Correction du terme de variance de PEF

La correction de l'effet (2) c'est à dire la prise en compte du terme $(1+\sigma^2_{\Phi_N,\lambda})$ dans l'équation 4.11, peut éventuellement être réalisée. Plusieurs stratégies sont envisageables :

1. Ce terme peut être négligé. En effet, la PEF est de l'ordre de quelques pourcents et toujours inférieure à 4 % dans nos expériences, résultant en $\sigma^2_{\Phi_N,\lambda} < 0.16\%$. Cette stratégie est appliquée pour nos expériences.
2. Le paramètre $\sigma_{\Phi_N,\lambda}$ peut être estimé pendant les acquisitions à partir d'une mesure de photodiode par exemple, ou en amont si la fluence du laser est stable entre deux expériences.
3. Les images $A^{SVD}_{\lambda}(\mathbf{r})$ soit les P images filtrées par SVD peuvent être corrigées par les valeurs mesurées à la photodiode $\Phi_{k,\lambda}(\mathbf{r})$. Plus contraignant en terme de dispositif expérimental, cette approche permet de tenir compte d'une instabilité intra longueur d'onde du laser telle qu'une dérive de la puissance moyenne par exemple.

L'image de fluctuation est maintenant proportionnelle à l'énergie absorbée.

$$\sigma^2_{corr,\lambda} = \sigma^2_{\lambda}[A^{SVD}_{k,\lambda}(\mathbf{r})] - \sigma^2_{b,SVD} \approx \overline{\Phi}^2_{\lambda}(\mathbf{r})\sigma^2_{\lambda,flow}(\mathbf{r}) \tag{4.16}$$

4.2.4 Estimation de la saturation en oxygène en présence de bruit et de PEF

L'importance de la correction apportée par l'estimation du bruit retiré par le filtrage SVD est observable sur la simulation présentée en figure 4.4, de paramètres FNR = 0.2, N = 512 images et $\sigma_{\Phi_N,\lambda}$ variant entre 20 et 40%. Ces conditions de simulation sont plus défavorables que celles des expériences. Les cartes de SO_2 avec et sans correction de l'estimation du bruit retiré par le filtrage par SVD (a=2, b=1000), et les courbes d'absorption associées y sont présentées. Sans correction de l'estimation du bruit, l'estimation de la SO_2 est complètement biaisée, de $32.7 \pm 20\%$ en moyenne : la quantité de bruit calculée par la formule 4.8 est largement supérieure à la quantité réellement présente dans l'image. En corrigeant cette quantité par l'estimation du bruit supprimé par la SVD, une SO_2 moyenne de $90.4 \pm 6.5\%$ est obtenue. Après cette correction, le biais est d'environ 5%, indiquant que notre hypothèse d'un bruit uniformément réparti sur toute l'image n'est pas parfaitement vérifiée. Elle permet néanmoins d'obtenir une amélioration plus que significative de la précision de l'estimation.

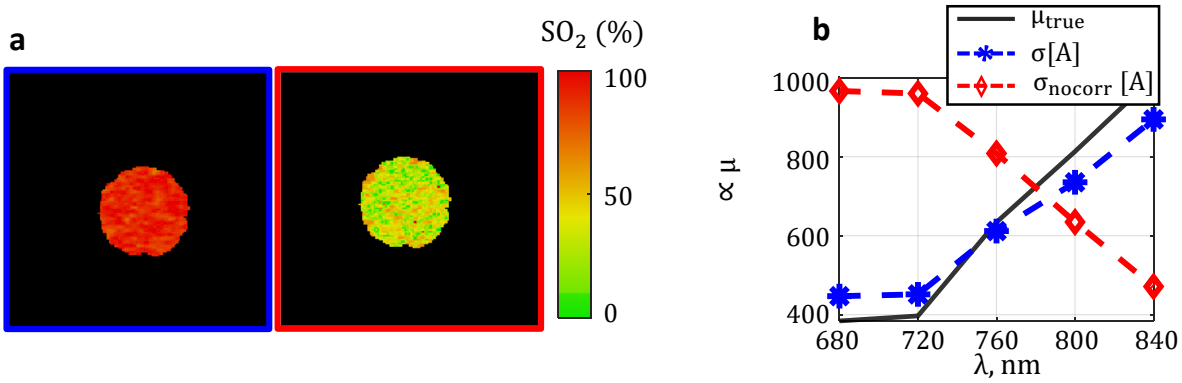

Figure 4.4 – Illustration de l'importance de la correction de l'estimation du bruit après filtrage par SVD. (a) cartes de SO_2 de fluctuation, à droite (rouge) calculée à partir des images pour lesquelles le bruit non corrigé σ_b est retiré et à gauche (bleu) à partir des images pour lesquelles le bruit corrigé de l'effet du filtrage par SVD $\sigma_{b,SVD}$ est retiré. (b) : la courbe noire représente l'absorption de référence, les courbes bleue et rouge correspondent aux estimations des coefficients d'absorption utilisés pour calculer les carte de SO_2 de (a).

Dans ces simulations, il suffit de filtrer la première valeur singulière (donc choisir $a = 2$) pour compenser l'effet de la PEF et ainsi supprimer le contribution de $\sigma^2_{\Phi_N,\lambda}|m_{\lambda}(\mathbf{r})|^2$. En conditions expérimentales, a

est généralement compris entre 10 et 40. On fait alors l'hypothèse que le bruit contenu dans les vecteurs singuliers de 1 à a est négligeable pour un a suffisamment petit devant P. Plus généralement, l'impact de le borne a n'est pas pris en compte : sa valeur devra donc être judicieusement choisie. Une étude en partie 4.5.2 décrit, en fonction de différents paramètres, l'impact du choix de ces bornes sur la qualité de l'estimation. Nous présentons également dans la partie 4.6 une méthode semi-automatique permettant de choisir une borne a suffisamment élevée pour faire apparaître les structures affectées par les artefacts de visibilité mais suffisamment basse pour que l'estimation d'oxygénation soit correcte.

Dans cette partie, une théorie de l'imagerie photoacoustique quantitative de fluctuation a été proposée, et les méthodes de correction à appliquer pour obtenir une image proportionnelle à l'énergie absorbée ont été détaillées. L'approche, appliquée sur des données de simulation en présence de bruit électronique et de PEF dans des conditions très défavorables, permet d'estimer la saturation en oxygène avec une très bonne précision. Dans la partie suivante, l'imagerie photoacoustique quantitative de fluctuation est appliquée en conditions expérimentales avec la sonde d'imagerie 3D.

4.3 Imagerie photoacoustique quantitative 3D de fluctuation : imagerie d'oxygénation du sang à visibilité totale

4.3.1 Méthode de l'imagerie photoacoustique quantitative de fluctuation en conditions expérimentales

Chaîne de traitement des images expérimentales

Deux types échantillons sont imagés : des tubes remplis de sang mis en mouvement par un pousse-seringue, et des embryons de poulet. Les protocoles et schémas expérimentaux sont similaires à ceux décrits dans le chapitre 3, en changeant la longueur d'onde de l'illumination entre chaque acquisition et en enregistrant la fluence à l'aide d'une photodiode. La séquence d'acquisition inclut l'obtention d'images US intercalées entre chaque acquisition PA, comme décrit dans le chapitre 2 section 2.3.1. La partie 4.4 traitera de l'exploitation des données US, nous nous concentrons sur l'exploitation des données PA dans cette partie. Le nombre de longueurs d'onde d'illumination M_λ est égal à 7 ou 10 selon l'exemple.

Le schéma présenté en figure 4.5 illustre la chaîne de traitement appliquée aux images PA pour obtenir l'image de SO_2, résumant la méthode décrite dans la partie précédente. Les images 3D sont reconstruites par *delay-and-sum*, puis filtrées par SVD. Le choix des bornes a et b est semi-automatique selon la méthode décrite dans la partie 4.6. Les images de variance pour chaque λ sont calculées puis le bruit est estimé et soustrait à chacune de ces images. Le terme $(1+\sigma^2_{\Phi_N,\lambda})$ est négligé : les images de fluctuation sont maintenant proportionnelles à l'énergie absorbée. La fluence est ensuite corrigée via la mesure de $\Phi_\lambda(\mathbf{0})$ par la photodiode calibrée, en négligeant donc les effets de coloration spectrale. Les images de fluctuations sont proportionnelles aux coefficients d'absorption optique. L'inversion pour obtenir la carte de SO_2 est ensuite réalisée pour chaque pixel appartenant à un masque contenant les vaisseaux, défini automatiquement (section 4.6) à partir de l'une des images de fluctuation. Pour limiter l'effet des fluctuations spatiales des grandeurs σ_{flow} et σ_b résultant du nombre fini de mesure, un masque gaussien 3D est préalablement appliqué sur les images de fluctuation.

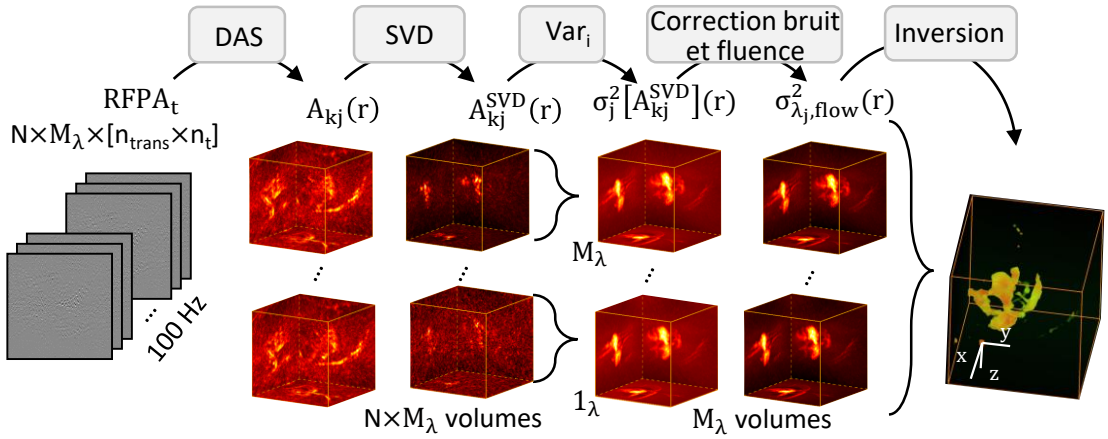

Figure 4.5 – Chaîne de traitement des images expérimentales, à partir des signaux radiofréquences photoacoustiques jusqu'à la carte de SO_2. M_λ est le nombre de longueurs d'onde λ d'indice j et N le nombre d'images par λ d'indice k. Les images photoacoustiques sont reconstruites par delay-and sum, puis une décomposition en valeurs singulières est réalisée sur l'ensemble des $N \times M_\lambda$ images. Les images de variance sont calculées pour chaque λ_j puis le bruit et la fluence sont corrigées : les images de fluctuation sont maintenant proportionnelles l'absorption optique et la carte de SO_2 peut être estimée.

Validation de la calibration de la photodiode

Nous validons l'imagerie photoacoustique quantitative de fluctuation en comparant la carte de SO_2 de fluctuation et la carte de SO_2 conventionnelle sur des structures visibles à la fois sur l'image de fluctuation et sur l'image conventionnelle. Les estimations des cartes de SO_2 conventionnelles sont obtenues à partir des images photoacoustiques moyennées à chaque λ et corrigées par la fluence mesurée par la photodiode. La validité de cette comparaison dépend de la proportionnalité des images conventionnelles à μ_λ. Cela implique que la mesure photodiode, calibrée par le détecteur pyroélectrique, est bien proportionnelle à $\Phi_\lambda(\mathbf{0})$.

Pour s'en assurer, les spectres d'absorption de deux espèces chimiques sont mesurés à la fois par un spectrophotomètre (Jasco V650 spectrometer), qui constituent nos références, et par imagerie PA multispectrale conventionnelle. Deux capillaires en verre de diamètre interne 1 mm sont remplis respectivement d'une solution de sulfate de cuivre et d'une solution de mélanine. Les deux capillaires sont imagés simultanément. On obtient une série de 2300 acquisitions pour $M_\lambda = 23$ (100 acquisitions par λ, de 680 à 900 nm par pas de 10 nm), puis on calcule les moyennes des images reconstruites par *delay-and-sum* après correction de la fluence pour chaque λ. La valeur des pixels est moyennée sur les parties inférieures et supérieures des tubes (le centre étant invisible à cause des artefacts de visibilité), permettant d'obtenir des spectres proportionnels à l'absorption optique pour les deux espèces. Ces spectres sont comparés à ceux obtenus avec le spectrophotomètre. Ils sont présentés en figure 4.6.

Les spectres photoacoustique se superposent aux spectres produits par le spectrophotomètre. L'écart quadratique moyen est ainsi de 3% pour la solution de mélanine, et de 1.6% pour la solution de sulfate de cuivre. Notre image photoacoustique conventionnelle est donc bien proportionnelle à l'énergie absorbée. La calibration de la fluence du laser à la sortie de la fibre est correcte ainsi que l'enregistrement par la photodiode, ce qui nous permet de retrouver une grandeur proportionnelle à l'énergie déposée et d'extraire les coefficients d'absorption.

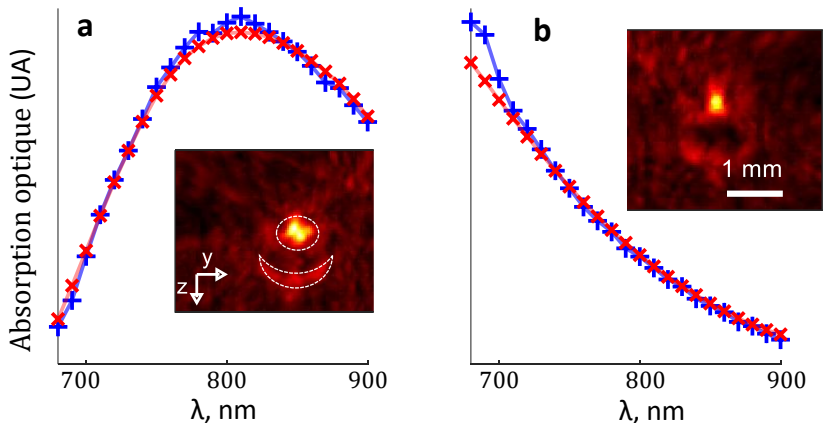

Figure 4.6 – Validation de la calibration de la photodiode.
Spectres d'absorption obtenus par imagerie photoacoustique (bleu) et à l'aide d'un spectrophotomètre (rouge) pour des solutions de mélanine (a) et de sulfate de cuivre (b). Les deux espèces chimiques sont imagées simultanément, une coupe YZ de chacun de ces tubes est visible. Les zones en pointillés blancs correspondent grossièrement aux extrémités des tubes, dans lesquels la valeur des pixels est moyennée pour former les marqueurs bleus.

4.3.2 Imagerie photoacoustique quantitative de fluctuation *in vitro*

Deux échantillons de sang humain sont imagés dans deux capillaires en verre de 1 mm de diamètre et de paroi 0.1 mm. Le sang est mis en mouvement à l'aide d'un pousse seringue induisant une vitesse moyenne de flux dans ces capillaires de 2 mm/s. On utilise $M_\lambda = 10$ longueurs d'onde pour l'illumination, de 700 à 880 nm par pas de 20 nm, et $N = 400$ acquisitions sont réalisées par λ. Chaque échantillon a préalablement été mis au contact d'un sac rempli respectivement d'O_2 pure ou de CO_2 pendant 3h à 36°C pour obtenir des niveaux d'oxygénation différents dans chacun des deux tubes. Des coupes selon l'axe YZ des images 3D de fluctuation et des images 3D conventionnelles (à 700 nm) ainsi que les cartes de SO_2 correspondantes sont présentées en figure 4.7.

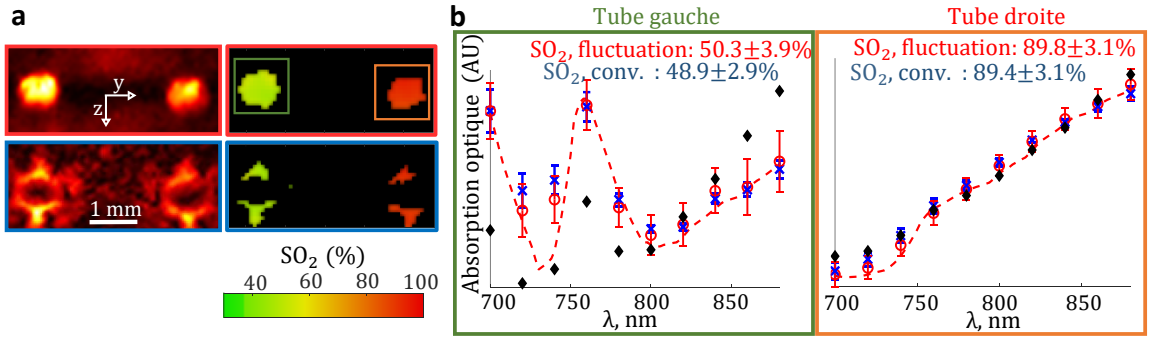

Figure 4.7 – Imagerie 3D quantitative de deux tubes, l'un contenant du sang oxygéné et l'autre du sang désoxygéné.
(a) : à gauche, coupe YZ de l'image de fluctuation en haut et de l'image conventionnelle en bas à 700 nm. A droite, cartes de SO_2 associées. (b) : Les marqueurs représentent la moyenne des pixels inclus dans des masques du tube pour les images de fluctuation avec le bruit corrigé (cercles rouges), sans corriger le bruit (losanges noirs) et pour les images conventionnelles (croix bleues). La SO_2 moyenne dans chacun des tubes est en légende. Les courbes en pointillés correspondent aux coefficients d'absorption théoriques des SO_2 moyennes (modèle) des images de fluctuation corrigées.

Les marqueurs représentent la valeur moyenne des pixels et l'écart type associés à l'intérieur de masques pour chacun des deux tubes à chaque λ. Les masques correspondent aux centres des tubes pour les images de fluctuation et aux extrémités des tubes pour les images conventionnelles. Les marqueurs représentant les images de fluctuation en corrigeant le bruit sont en rouge, ceux représentant les images de fluctuation sans corriger le bruit en noir, et ceux représentant les images conventionnelles en bleu. Les valeurs des pixels sont préalablement divisées par leurs valeurs de $K(r)$ respectives obtenues lors de l'estimation de la SO_2. Cela permet de normaliser les images de sorte que la valeur de chaque pixel ne dépende plus que de μ_λ, et que la moyenne spatiale puisse être calculée. Enfin, les courbes en pointillé rouge correspondent aux coefficients d'absorption théoriques (le modèle) pour les SO_2 moyennes estimées.

Alors que seules les extrémités du tube sont visibles sur l'image conventionnelle, l'intérieur est bien reconstruit sur l'image de fluctuation, indiquant que les artefacts de visibilité sont corrigés (figure 4.7.a, gauche). Les cartes de SO_2 de fluctuation sont uniformes et les valeurs sont similaires pour les images conventionnelles et les images de fluctuation. Ces observations sont confirmées par les valeurs moyennes de SO_2 pour chaque tube, très proches (tube de gauche, $50.3 \pm 3.9\%$ pour l'image de fluctuation et $48.9 \pm 2.9\%$ pour l'image conventionnelle, tube de droite, $89.8 \pm 3.1\%$ pour l'image de fluctuation et $89.4 \pm 3.1\%$ pour l'image conventionnelle). Les courbes d'absorption sont par ailleurs très similaires. Il est intéressant d'observer que la correction du bruit est indispensable pour obtenir une bonne estimation des coefficients d'absorption. C'est notamment le cas pour le tube contenant le sang moins oxygéné (cadre vert) comme le montre l'écart entre les marqueurs noirs et les marqueurs bleus ou rouges, ce qui confirme l'importance de la correction du bruit. Sans la correction, la SO_2 est ainsi estimée à 66.3% pour ce tube au lieu d'environ 50%. Enfin, on note que les coefficients d'absorption des images de fluctuation sont en adéquation avec ceux obtenus par le modèle.

Ces résultats sont très encourageants : ils démontrent que l'imagerie photoacoustique quantitative de fluctuation permet bien de retrouver une grandeur proportionnelle à la densité d'énergie absorbée et d'en extraire une estimation de l'oxygénation, sur des zones affectées par les artefacts de visibilité et donc invisibles sur les images conventionnelles.

4.3.3 Imagerie photoacoustique quantitative de fluctuation de l'embryon de poulet

Un embryon de poulet est imagé selon le schéma expérimental disponible dans le chapitre 2, avec cette fois $N = 150$ et $M_\lambda = 7$ (700 nm à 820 nm par pas de 20 nm). L'échantillon imagé est celui présenté en figure 3.14 ex#1 du chapitre 3. Les cartes de SO_2 associées sont présentées en figure 4.8. Les estimations sont réalisées dans des masques obtenus en seuillant les projections d'intensité maximale (MIP) d'une image de fluctuation à et d'une image conventionnelle. Les représentations 3D sont obtenues avec le logiciel Amira. Le bénéfice apporté par l'imagerie de fluctuation est évident : la SO_2 peut être estimée sur un nombre bien plus important de structures vis à vis de l'imagerie conventionnelle. Par ailleurs, une plus grande variété de niveau d'oxygénation est visible sur l'image de fluctuation.

Comme attendu, les estimations de SO_2 sont similaires lorsque les structures sont visibles sur les deux types d'images. Une zone de l'image pour laquelle la SO_2 est stable dans une structure visible à la fois sur l'image conventionnelle et sur l'image de fluctuation est plus précisément étudiée. Les mêmes traitements que ceux décrits dans la section 4.3.2 sont appliqués pour extraire les courbes d'absorption ainsi que les estimations moyennes de SO_2. Il en résulte à nouveau des profils très similaires et donc des estimations d'oxygénation très proches lorsque l'image de fluctuation est corrigée.

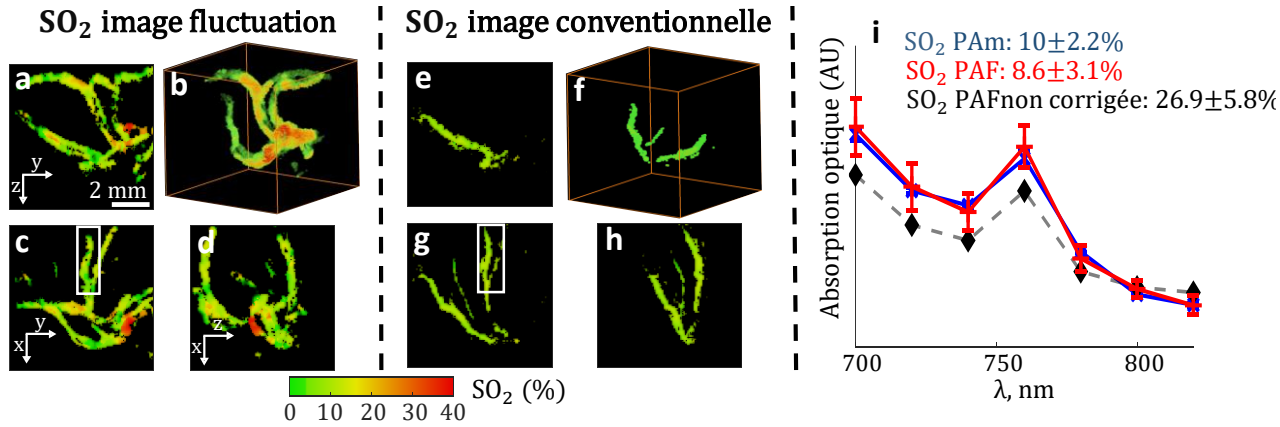

Figure 4.8 – Imagerie 3D quantitative de l'embryon de poulet.
Représentation 3D et projections de la carte de SO_2 de fluctuation (a-d) et conventionnelle (e-h). (i) : courbe d'absorption de la structure inclue dans la ROI (carré blanc) calculée à partir des images de fluctuation (rouge), des images moyennes (bleu), ou des images de fluctuation sans corriger le bruit (noir). Les valeurs de SO_2 moyennes sont en légende.

Des comparaisons similaires ont été réalisées sur d'autres zones d'embryons, confirmant ces observations. Les cartes de SO_2 de deux autres échantillons sont proposées en figure 4.9, ainsi que les analyses sur des ROIs incluant des vaisseaux visibles sur l'image conventionnelle et l'image de fluctuation.

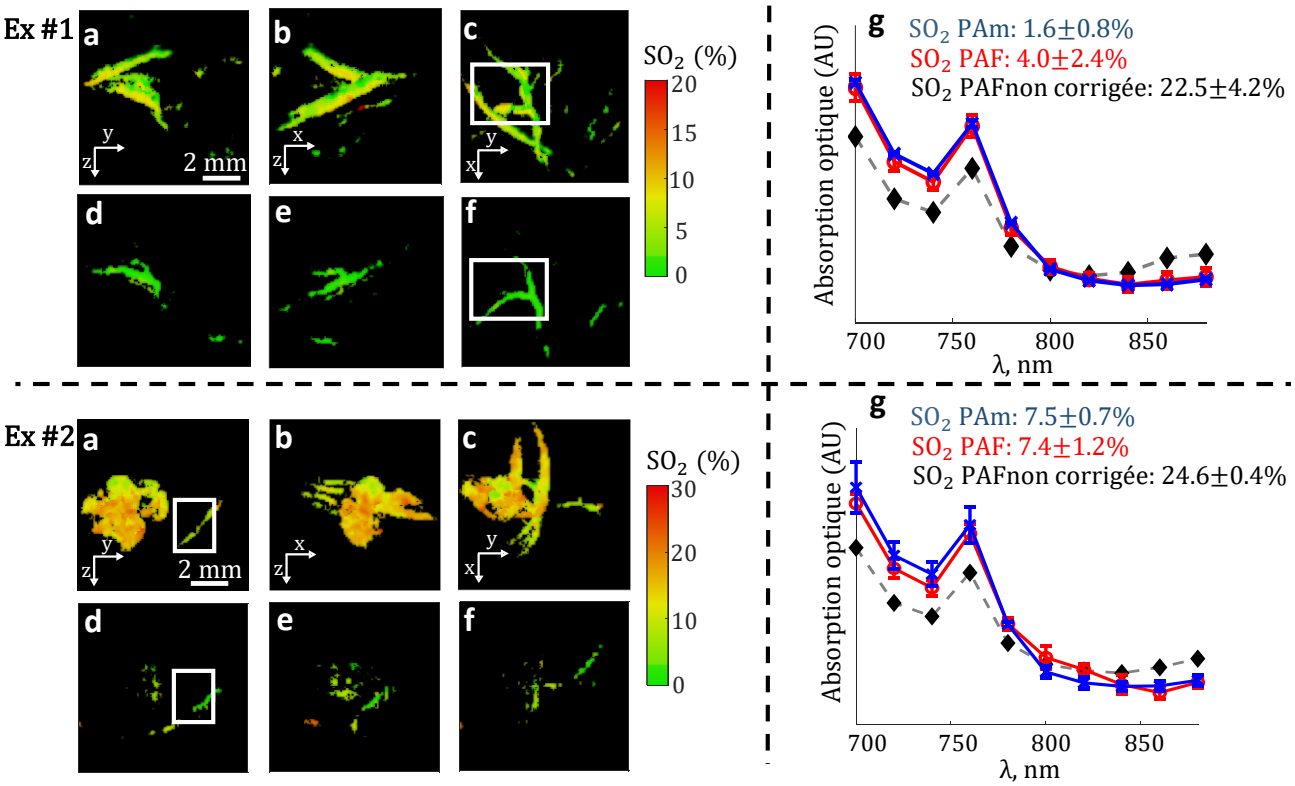

Figure 4.9 – Imagerie 3D quantitative de l'embryon de poulet, deux exemples.
Carte de SO_2 de fluctuation (a-c) et conventionnelle (d-f). (g) : courbe d'absorption de la structure inclue dans la ROI (carré blanc) calculée à partir des images de fluctuation (rouge), des images moyennes (bleu), et des images de fluctuation sans corriger le bruit (noir). Les valeurs de SO_2 moyennes sont en légende.

Sur le premier exemple, des observations similaires peuvent être réalisées : des structures plus oxygénées apparaissent. Sur le second exemple, le coeur de l'embryon est imagé. La carte de SO_2 conventionnelle est très pauvre car les images de fluctuation sont polluées par un très fort *clutter* (l'image de fluctuation de cet exemple est visible dans le chapitre 3, figure 3.14 ex#2). A l'inverse, sur la carte de SO_2 de fluctuation, une structure présentant une oxygénation complexe, globalement plus élevée apparaît. Dans les deux cas, les courbes d'absorption calculées pour des structures visibles sur l'image conventionnelle et l'image de fluctuation sont similaires.

Dans nos expériences *in vivo*, une SO_2 globalement très basse est observée, en particulier sur les images conventionnelles : il aurait été intéressant de comparer les mesures de SO_2 sur des structures plus oxygénées. Il est possible que le contact du cône de couplage de la sonde d'imagerie avec la surface de l'embryon limite l'apport en oxygène, ou simplement que l'embryon ne s'oxygène plus correctement lorsque la coquille de l'oeuf est brisée, comme c'est le cas au cours de ces expériences. Les résultats *in vitro* présentés précédemment démontrent néanmoins la possibilité d'estimer correctement la SO_2 pour de haute oxygénation. Les résultats présentés dans cette section nous permettent donc de conclure que l'imagerie photoacoustique quantitative de fluctuation est applicable *in vivo* et permet de générer des cartes de SO_2 bien plus riches que l'imagerie photoacoustique quantitative conventionnelle.

4.4 Imagerie photoacoustique quantitative de fluctuation et Doppler ultrasonore 3D simultanées

4.4.1 Méthode de l'imagerie Doppler ultrasonore quantitative

La séquence d'imagerie présentée dans le chapitre 2 section 2.3.1 est utilisée pour réaliser les acquisitions PA et US. Les 20 sous-acquisitions nécessaires à la formation de deux images US sont réalisées entre chaque acquisition PA. Les sous-images correspondantes à ces sous-acquisitions sont reconstruites par *delay-and-sum* et sommées entre elles pour former l'image US et un filtrage par SVD est ensuite appliquée sur la pile d'images résultante, avec des bornes a et b choisies empiriquement. L'image Directional Doppler (DPD), fournissant la direction du flux et le *blood volume*, et l'image Color Doppler (CD) fournissant la vitesse du flux sont ensuite calculées après une transformée de Fourier de la pile d'images filtrées. L'image DPD résulte de la superposition de deux images : celle du sang circulant en direction de la sonde, définie comme l'intégrale en chaque pixel des fréquences de la partie négative du spectre, et celle du sang circulant dans la direction opposée, définie comme l'intégrale des fréquences appartenant à la partie positive du spectre. Un paramètre de transparence est choisi sur un critère visuel. L'image CD est proportionnelle à la fréquence Doppler moyenne calculée à partir du spectre temporel en chaque position.

4.4.2 Images Directional Doppler et Color Doppler 3D de l'embryon de poulet

La figure 4.10 présente les images CD et DPD 3D de l'échantillon dont l'image Power Doppler est observable dans le chapitre 3, figure 3.14 ex#1, et dont l'estimation de la SO_2 a été réalisée dans la figure 4.8. Ces images CD et DPD sont obtenues à partir du traitement de P=1000 images US.

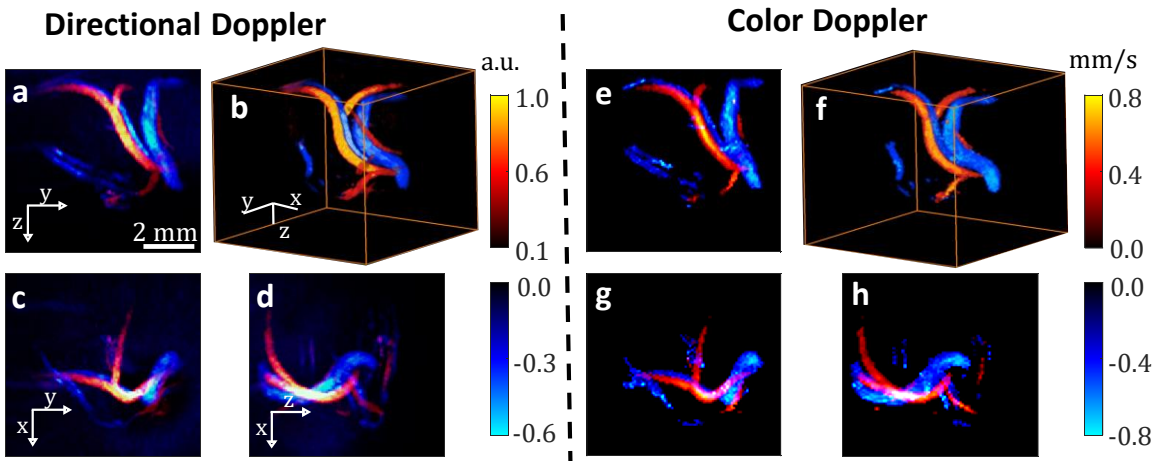

Figure 4.10 – Analyses de flux.
Partie gauche (a-d) : représentation 3D et projection de l'image Directional Doppler. La couleur encode la direction du flux et le blood volume, en rouge le flux en direction de la sonde et en bleu le flux en direction opposée. Partie droite (e-h) : représentation 3D et projection de l'image Color Doppler. La couleur encode la direction du flux et sa vitesse.

La couleur encode la direction du flux sanguin vis à vis de l'axe Z de la sonde. Deux vaisseaux voisins sont clairement identifiables sur la projection YZ, montrant des directions de flux opposées et pouvant ainsi être apparentés à une veine et une artère. Sur l'image CD, le profil de la vitesse, plus élevé au centre que sur les extrémités du vaisseau, est similaire à un écoulement de Poiseuille.

Sur un autre échantillon dont les images sont acquises selon le même protocole, une analyse de pulsatilité est présentée dans la figure 4.11.

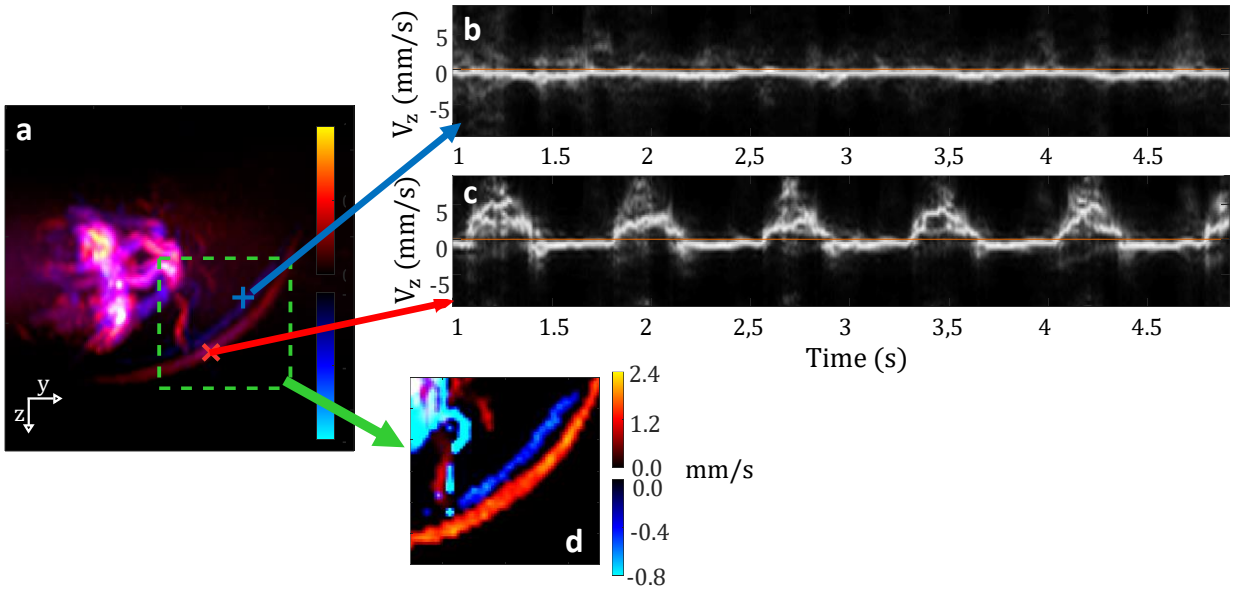

Figure 4.11 – Analyse de flux, pulsatilité.
(a) : projection YZ de l'image Directionnal Doppler. (b) et (c) : spectres des deux vaisseaux contenus dans la ROI verte. (d) image Color Doppler des deux vaisseaux dans la ROI verte.

L'image DPD est représentée en (a). La couleur violette indique la présence de flux complexes et de mouvement de tissus générant de l'aliasing, résultant d'une fréquence d'acquisition trop faible. Néanmoins, cette fréquence d'acquisition est suffisamment élevée pour étudier les deux vaisseaux à l'intérieur du rectangle vert. Leurs images CD (d) ainsi que les spectres (b) et (c) sont affichés. La vitesse du flux dans le vaisseau bleu est faible, et son spectre est plat ce qui semble indiquer la présence d'un flux revenant de la membrane chorioallantoique, permettant d'identifier ce vaisseau comme une veine. A l'inverse, le vaisseau rouge peut être considéré comme une artère : la vitesse du flux y est plus élevée, et son spectre décrit une forte activité pulsatile en raison probablement de la proximité avec le coeur.

L'obtention de telles images Doppler ultrasonores en 3D avec une sonde d'imagerie parcimonieuse composée de seulement 256 éléments est inédite, et ouvre la voie à une utilisation conjointe de l'imagerie Doppler ultrasonore quantitative et de l'imagerie photoacoustique quantitative de fluctuation.

4.4.3 Analyse conjointe de la SO_2 et du flux

Le système est complètement exploité lorsque les estimations de SO_2 et les analyses de flux sont réalisées conjointement. La figure 4.12 présente des résultats obtenus sur un nouvel échantillon provenant d'un embryon âgé de 9 jours. Il est imagé avec 10 longueurs d'onde en imagerie PA (700 nm à 880 nm par pas de 20 nm), $N = 250$ images par λ. En imagerie US, 10 sous-acquisitions par image sont réalisées, P=1000 images US et la fréquence d'imagerie est de 200 Hz. Une morphologie similaire est observable sur la carte de SO_2 et les images Doppler, bien que certaines structures n'apparaissent que dans l'une ou l'autre des deux modalités. Une analyse fonctionnelle peut être réalisée, comme le montre la partie basse de la figure. L'image DPD et la carte de SO_2 d'une ROI incluant deux vaisseaux y sont affichées. Ces deux vaisseaux ont des directions de flux opposées et présentent des oxygénations mesurées respectivement à $5.7 \pm 2.7\%$ et $13.1 \pm 3.1\%$, une différence significative.

Les résultats présentés dans cette partie démontrent que l'imagerie photoacoustique quantitative de fluctuation 3D est applicable *in vivo* et permet d'obtenir une estimation de la SO_2 sur un grand nombre de structures, quant l'imagerie photoacoustique quantitative conventionnelle est limitée par les artefacts de visibilité et le *clutter*. Ces résultats démontrent également l'applicabilité des méthodes d'imagerie Doppler ultrasonores quantitatives avec notre sonde d'imagerie 3D parcimonieuse. L'utilisation conjointe des deux modalités, permise par le système d'imagerie mis en oeuvre au cours de cette thèse, permet d'obtenir simultanément de nombreuses informations d'intérêt (morphologie des vaisseaux sanguins, saturation en oxygène, vitesse et direction du flux sanguin, pulsatilité) pour une multitude d'applications pré-cliniques et cliniques. Une dernière étape reste à franchir, non des moindres : la démonstration de notre système en milieu turbide. Les challenges sont nombreux, principalement pour l'imagerie photoacoustique quantitative de fluctuation, de par la forte diffusivité du milieu et les effets de coloration spectrale dont il faudra tenir compte pour fournir une estimation de la SO_2.

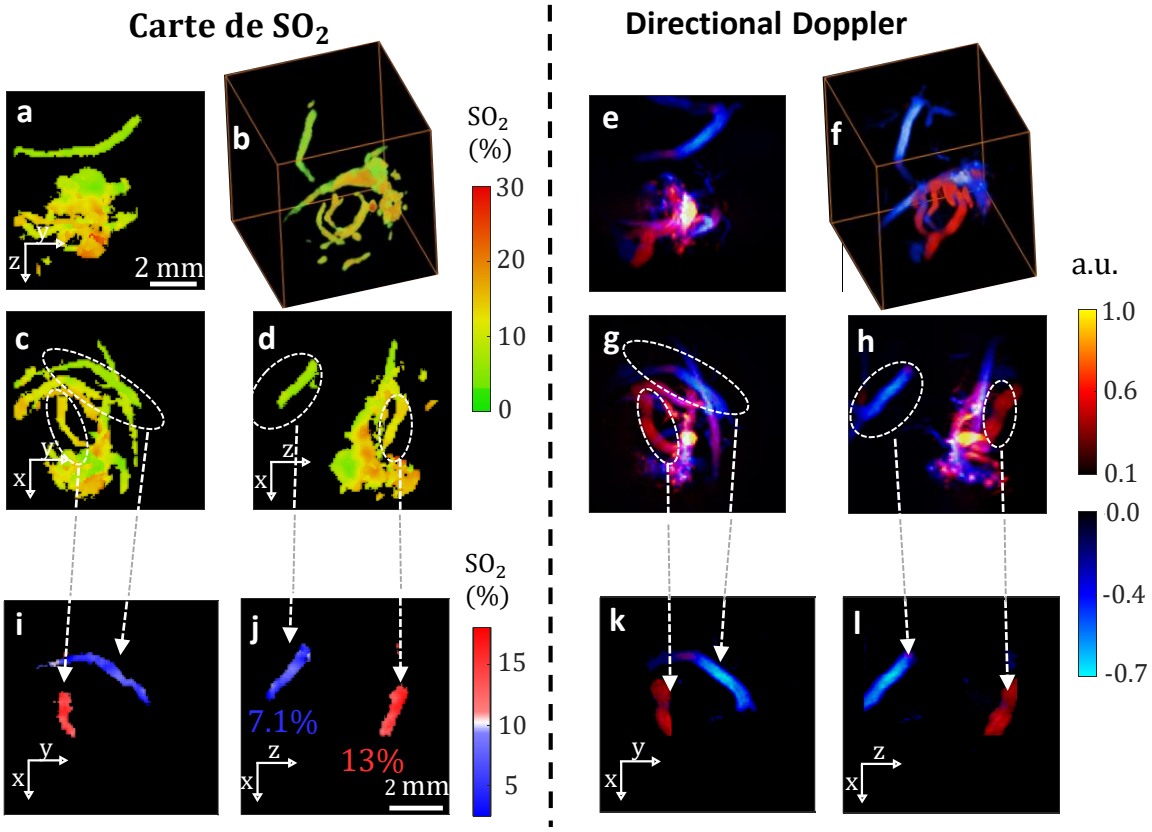

Figure 4.12 – SO_2 et analyse de flux conjointe.
La partie gauche (a-d) représente la carte de SO_2, la partie droite (e-h) l'image Directional Doppler. Les vaisseaux repérés par les cercles pointillés sont extraits et représentés en (i-j) pour la SO_2 avec une barre de couleur saturée et en (k-l) pour le Directional Doppler.

4.5 Étude sur simulation de l'impact du bruit électronique et du PEF sur l'estimation de SO_2

La qualité de l'estimation de la SO_2 dépendra fortement de la qualité de l'estimation de σ_b^2 et $\sigma_{flow}^2(\mathbf{r})$, et donc du nombre d'images disponibles à chaque λ et du ratio fluctuation sur bruit (FNR). L'effet du filtrage par SVD et le choix des bornes a et b a également un impact conséquent. Des analyses sur simulation sont réalisées pour évaluer l'impact de ces paramètres.

4.5.1 Influence du bruit électronique

Pour observer l'impact du nombre d'images disponibles à chaque λ et du FNR sur la qualité de la reconstruction, du bruit aléatoire de variance σ_b^2 est ajouté au jeu de données de simulation, obtenu de manière similaire à celui de la première partie. Les coefficients d'absorption sont choisis de sorte à traduire une SO_2 de 74.4%, proche du niveau d'une veine. La figure 4.13 présente, pour différents couples de FNR / nombre d'images, la carte de SO_2 de l'objet avec correction du bruit (en retirant σ_b^2 selon l'équation 4.8) ainsi que les courbes correspondantes aux coefficients d'absorption au sein de l'objet à chaque longueur d'onde, estimés à partir des images corrigées (bleu) et des images non corrigées (rouge). Les courbes noires représentent les coefficients d'absorption théoriques utilisés pour la simulation. Les courbes de la dernière ligne représentent, pour un nombre d'images donné, l'évolution de l'estimation de l'oxygénation moyenne dans l'objet en fonction du FNR pour les images corrigées et non corrigées du bruit. La zone bleutée correspond à l'écart type de l'estimation de cette oxygénation pour les images corrigées.

Dans des conditions idéales, la SO_2 est correctement estimée (FNR 10, 512 images). La diminution du nombre d'images et du FNR impactent la qualité des estimations et il est clair que notre méthode de correction n'est pas suffisante dans les cas les moins favorables, car les estimations de $\sigma_{flow,\lambda}^2$ et/ou de σ_b^2 sont alors trop imprécises (faible nombre d'images) et que la variation spatiale du bruit est grande devant le signal de fluctuation (faible FNR). Lorsque le nombre d'images est suffisant, les estimations des μ_λ et de la SO_2 sont néanmoins très proches des valeurs recherchées y compris pour les faibles FNR (dernière colonne, première ligne). Dans ce cas, l'application de notre méthode de correction permet de très bonnes estimations des μ_λ et donc de la SO_2.

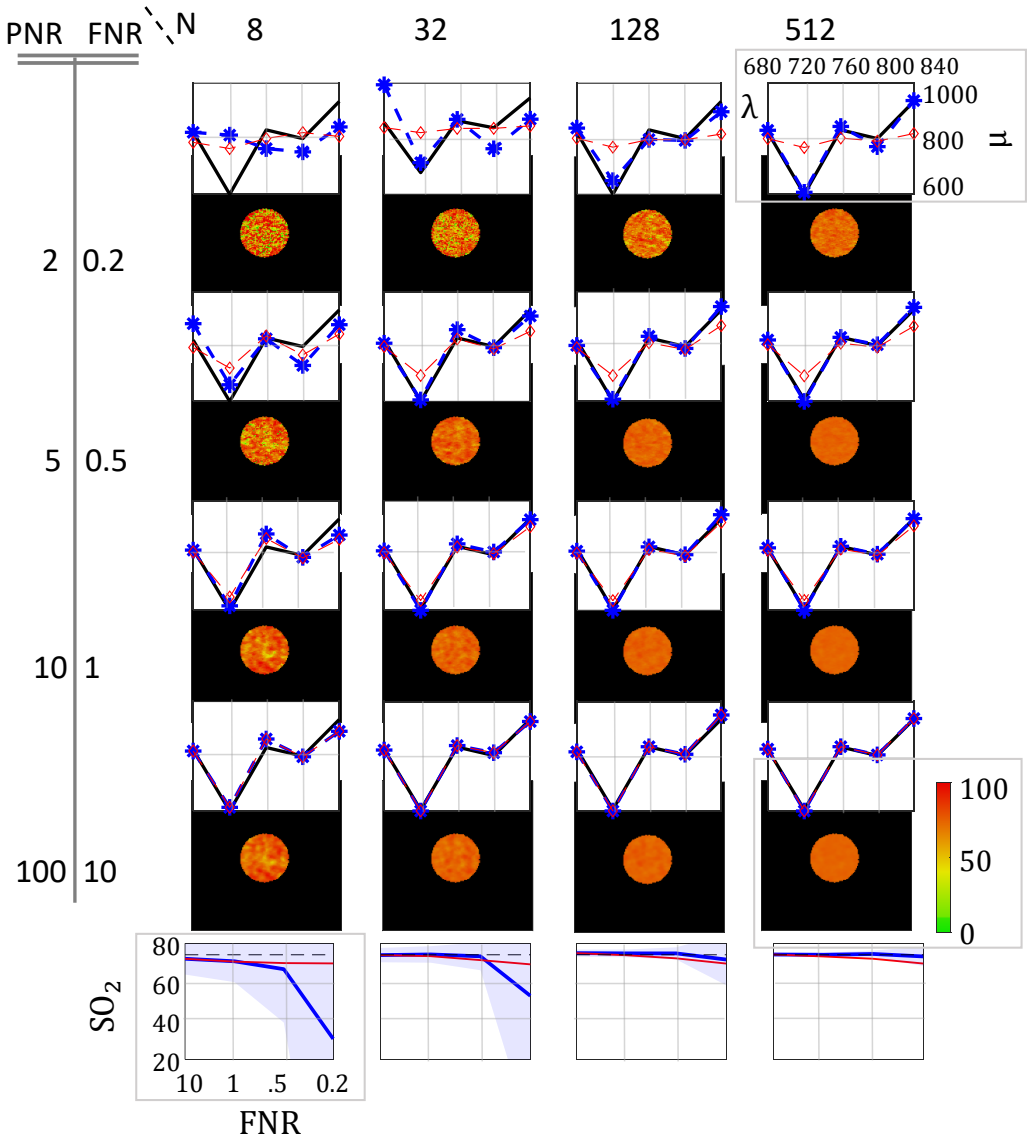

Figure 4.13 – Influence du fluctuation to noise ratio (FNR) et du nombre d'images sur l'estimation de la SO_2. La carte de SO_2 de l'objet avec correction du bruit ainsi que les courbes correspondant aux coefficients d'absorption estimés à partir des images corrigées (bleues) et des images non corrigées (rouges) sont présentées. Les courbes noires représentent les coefficients μ_λ utilisés pour la simulation, c'est à dire les valeurs à retrouver. Les courbes de la dernière ligne représentent, pour un nombre d'images donné, l'évolution de l'estimation de l'oxygénation en fonction du FNR pour les images corrigées et non corrigées.

4.5.2 Etude : impact du filtrage par SVD sur l'estimation de SO_2, simulation

Filtrage par SVD : intuition

La SVD décompose les données dans une base de vecteurs singuliers spatiaux U_i et temporels V_i, chaque couple de vecteurs singuliers caractérisant donc un comportement spatio-temporel. Lors du filtrage, seules les couples représentant l'objet d'intérêt, celui fluctuant selon g_k, sont conservés. Cette SVD est réalisée sur la totalité des acquisitions c'est à dire sur la pile de P images, ce qui inclue la variation temporelle selon λ. Une autre approche aurait été de réaliser une SVD indépendamment pour chaque λ sur N images, mais cette option n'a pas été retenue car elle diminuerait le pouvoir discriminant de la SVD, qui augmente avec la dimension de l'ensemble. On obtient donc :

$$A_k^{SVD}(\mathbf{r}) = \sum_{i=a}^{b} \alpha_i U_i(\mathbf{r}) V_i^*(k) \tag{4.17}$$

Avec a et b les bornes de filtrage, V_i de taille P et U_i de taille N_r. Ces vecteurs sont ordonnés selon leur énergie, traduite par α_i.

Il est ici intéressant d'observer les contenus de ces vecteurs, dans le cas simple décrit par les simulations d'un disque d'absorption uniforme, avec du bruit (FNR = 0.5), en l'absence d'autres sources de fluctuation que g_k (pas de PEF). La fluence du laser est constante vis à vis de λ. La figure 4.14 présente quatre de ces couples de vecteurs d'index 1, 8, 60 et 100, correspondant à des sources de fluctuations distinctes.

Chaque colonne de cette figure est décrite successivement.

1. La première valeur singulière $i = 1$ décrit la structure la plus énergétique : c'est l'image moyenne. La structure est observable dans $|U_1|$ (a, première colonne). Ici, le signal fluctue temporellement principalement selon μ_λ. On visualise cette fluctuation de μ_λ dans $|V_1|$ (b, première colonne). Dans la ligne c, les vecteurs $|V_k|$ sont moyennés selon k à chaque λ ce qui génère les courbes bleues. Pour $i = 1$, on obtient une courbe bleue parfaitement superposée au spectre d'absorption de référence en noir (c, première colonne). Si du bruit b_k est également décrit par ce premier couple de vecteurs singuliers, son amplitude est faible vis à vis de celle des structures décrites par $|U_1|$. La contribution du bruit, visible par la variation entre chaque k (b, première colonne) est négligeable.

2. Le couple i = 8 décrit la structure variant selon g_k (a, deuxième colonne). Cette structure variant également selon μ_λ, les deux fluctuations se combinent. Le spectre d'absorption de référence se retrouve bien dans $|V_8|$ (b, deuxième colonne) comme le montre sa moyenne sur k à chaque λ (c, deuxième colonne). Cependant, l'amplitude du bruit b_k, visible sur le vecteur spatial $|U_8|$ (a, deuxième colonne) est bien plus grande vis à vis de l'amplitude des structures fluctuant selon g_k, en comparaison du couple de vecteur à $i = 1$. Il en résulte une dépendance moins marquée des vecteurs temporels (b,c deuxième colonne) à μ_λ.

3. Le couple i = 60 décrit toujours la structure variant selon g_k. Néanmoins, la contribution de b_k s'intensifie, diminuant d'autant plus la qualité de l'estimation de μ_λ, ici fortement dégradée comme le montre la différence entre les courbes bleue et noire en c.

4. Enfin, le couple i = 1000 ne décrit plus que le bruit électronique : la fluctuation résultante est alors complètement indépendante de k.

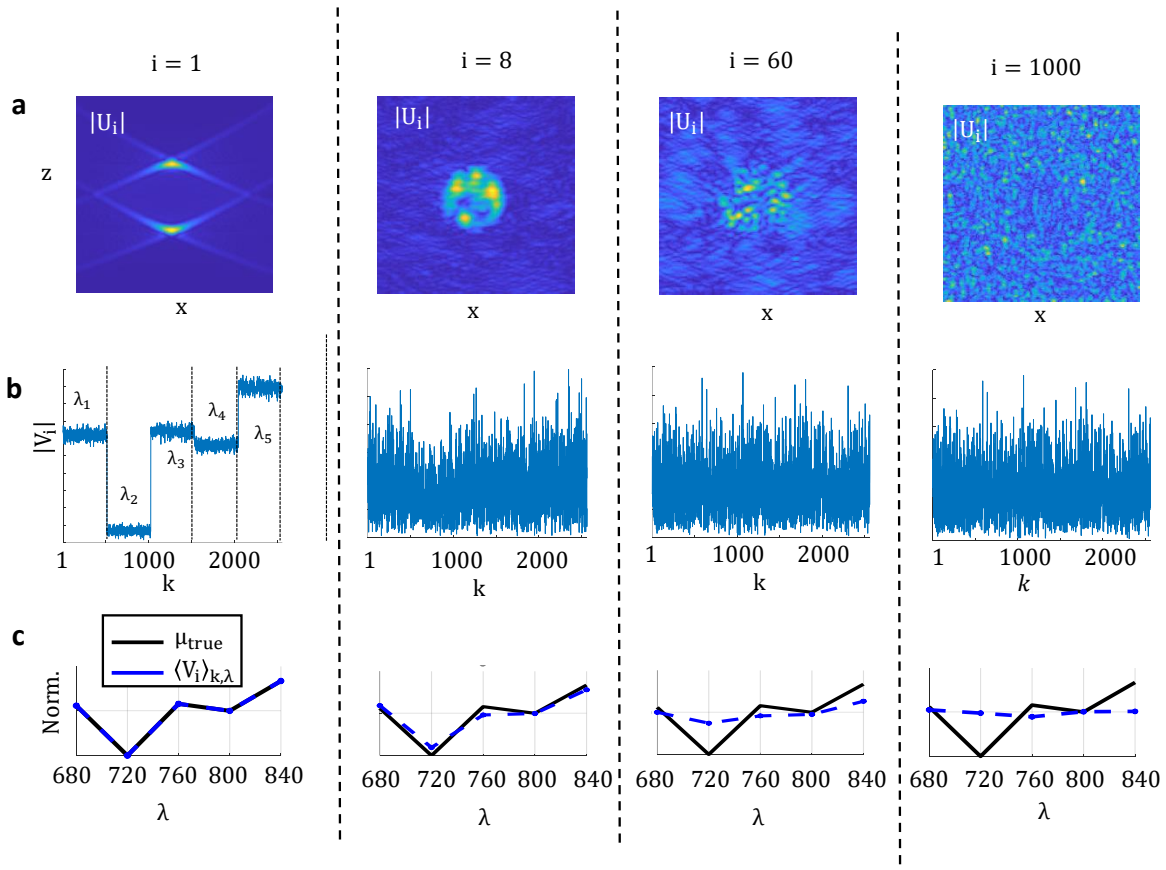

Figure 4.14 – Description des vecteurs temporels V_i et spatiaux U_i, pour les indices i =1, 8, 60 et 1000.
(a) : Représentation des vecteurs spatiaux $|U_i|$. En fonction de i, la structure décrite change : structure de l'image moyenne pour i=1, structure de l'image de fluctuation pour i=8 et i=60, bruit électronique pour i = 1000.
(b) : Les vecteurs $|V_i|$ contiennent l'information temporelle. Pour i=1, la fluctuation selon λ est bien visible. Elle est moins nette pour i = 10 en raison de l'importance du bruit b_k, et semble absente dès i = 60.
(c) : La grandeur par la courbe bleue représentée est la moyenne de $|V_i|$ selon k pour chaque λ. On peut ainsi visualiser la proximité entre la variation temporelle moyenne à chaque longueur d'onde et $\mu_{true,\lambda}$ (courbe noire), les coefficients d'absorption de référence à chaque λ. Cette proximité diminue lorsque i augmente.

Cette analyse montre que la borne a doit être suffisamment basse pour que l'estimation finale de μ_λ, moyennée sur les vecteurs propres de a à b, soit correcte. En effet, lorsque a augmente, l'objet est décrit par de moins en moins de vecteurs et ces vecteurs sont de plus en plus affectés par le bruit électronique.

Impact de la SVD sur l'estimation de la SO_2 en simulation

La SVD est nécessaire pour filtrer les fluctuations ne dépendant pas du flux sanguin. Elle a ainsi été systématiquement appliquée lors des expériences. Dans cette partie, nous essayons d'évaluer l'effet du filtrage par SVD sur la proportionnalité des images de fluctuation à $\mu_\lambda(\mathbf{r})$. Pour cela, on simule 512 acquisitions par longueur d'onde en présence d'un bruit électronique (FNR = 0.5), d'une PEF d'en moyenne 30 % et d'une fluence moyenne $\overline{\Phi}_\lambda(\mathbf{0})$ variant d'un facteur allant de 1 à 3. La fluence résultante du laser à chaque acquisition est visible en 4.15 (c).

Cette figure présente l'une des images de fluctuation (à 720 nm) et la carte de SO_2 associée pour différents couples de bornes a et b. Comme tout au long de ce chapitre, les images sont corrigées par $\overline{\Phi}_\lambda(\mathbf{0})$, supposée connue. Une étude plus approfondie est réalisée dans une ROI au centre du disque. La valeur des pixels au sein de cette ROI est moyennée à chaque λ pour former une courbe représentant le spectre d'absorption. Ces courbes sont générées pour plusieurs types d'images : la bleue est obtenue en corrigeant les images de fluctuation avec la méthode complète décrite dans ce chapitre, la rouge correspond à une absence totale de correction du bruit, et la cyan à une absence de correction de l'estimation du bruit (et donc de l'effet du filtrage SVD). La courbe noire correspond toujours aux coefficients d'absorption théoriques connus. Des informations supplémentaires sont disponibles dans les barres de couleur des images de fluctuations. Les nombres en orange et magenta correspondent respectivement à la valeur des pixels moyennés spatialement dans une zone de bruit et au centre de l'objet (la valeur magenta correspond ainsi à la racine de la somme du bruit et du signal de fluctuation). Enfin, en (b), on peut observer l'évolution en fonction de la borne a de l'estimation de la SO_2 moyennée dans la ROI, pour les différentes bornes b. La couleur de la courbe encode à nouveau le type d'images.

On observe à nouveau que, sans correction du bruit, les images de fluctuation ne sont plus proportionnelles à l'absorption (courbes rouges), l'estimation de l'oxygénation est biaisée. De même, la ressemblance entre les courbes et l'estimation de la SO_2 sont de moins en moins correctes lorsque b diminue en l'absence de prise en compte de l'effet du filtrage par SVD sur la quantité de bruit présente dans les images de fluctuation (courbes cyans, de gauche à droite). On peut maintenant s'intéresser aux courbes bleues, qui correspondent aux images pour lesquelles notre méthode de correction complète est appliquée. L'effet de la diminution de b (de gauche à droite) semble marginal, comme attendu. L'effet de l'augmentation de a (de haut en bas) est cependant bien visible. La ressemblance entre les courbes bleues et noires diminue, bien que le disque soit encore visible (par exemple à la ligne 3 correspondant à $a = 30$). Cette ressemblance est quasi inexistante à $a = 50$ (ligne 4). Dans cette simulation, la gamme de valeurs de borne a utilisable est vraisemblablement de 2 à 20.

Il est également intéressant d'observer l'évolution de la valeur moyenne des pixels dans la zone du bruit (nombres en orange) et du niveau de l'objet (nombres en magenta) en fonction des bornes du filtrage par SVD. Lorsque b diminue (de gauche à droite, notamment entre les colonnes 2 et 3), la valeur du bruit diminue fortement à l'inverse de celle de l'objet, conduisant à un meilleur FNR, ce qui se traduit visuellement par un objet ainsi qu'un fond plus lisse. Malgré cette augmentation artificielle du FNR, la qualité de l'estimation de la SO_2 diminue tout de même légèrement lorsque b diminue, en raison des approximations réalisées sur l'estimation du bruit supprimé par le filtrage par SVD. Lorsque a augmente, c'est principalement la valeur de l'objet qui diminue (exemple sur la première colonne, de 3.89 à 2.03 pour l'objet, de 2.39 à 2.03 pour le bruit), jusqu'à atteindre une valeur proche du bruit. En effet, le nombre de vecteurs propres décrivant l'objet diminue et ceux restant contiennent de plus en plus de bruit. Cela biaise la proportionnalité à μ_{true}, et l'estimation de la SO_2.

Une analyse similaire est disponible en annexe D sur une géométrie de type boucle, pour s'assurer que nos observations se généralisent au cas des artefacts liés à l'ouverture numérique limitée de la sonde d'imagerie.

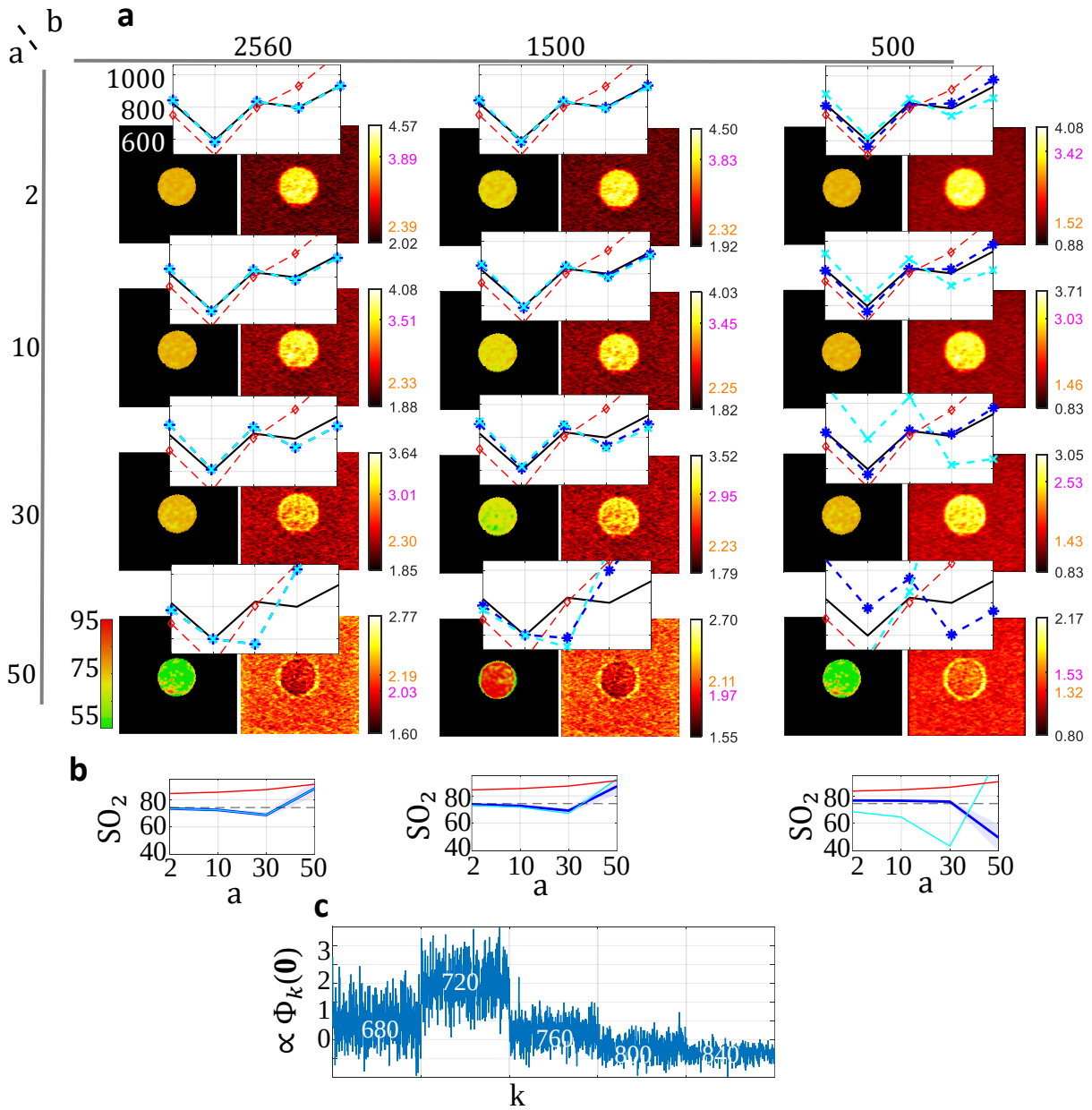

Figure 4.15 – Influence du choix des bornes a et b : simulation en présence d'une pulse energy fluctuation et d'une variation de la fluence moyenne. Ratio fluctuation à bruit à 0.5.
(a) : pour différents couples de bornes de filtrage, les images de fluctuation et les cartes de SO_2 associées sont présentées. Les nombres en magenta et orange sur la barre de couleur des images de fluctuation correspondent respectivement à la valeur des pixels moyennés spatialement au centre de l'objet, et dans une zone de bruit. Les courbes représentent, en noir les coefficients d'absorption de référence, en rouge, cyan et bleu les estimations des coefficients d'absorption obtenus à partir des images de fluctuation calculées respectivement sans correction du bruit, sans correction de l'effet de la SVD sur l'estimation du bruit et en appliquant la méthode de correction complète. (b) : les valeurs d'oxygénation en fonction de a pour les différents b sont présentées. (c) : représentation de la fluctuation de la fluence du laser.

Pour conclure cette partie, l'impact du bruit et du filtrage par SVD sont conséquents et doivent être pris en compte. Un FNR et un nombre d'images trop faible ne permettent pas d'estimer correctement la SO_2 malgré nos méthodes de correction. L'effet du filtrage et notamment de la borne *a* a été mis en évidence, dans des conditions de bruit très défavorables. Une borne *a* surestimée conduit à une baisse de la qualité de l'estimation de l'énergie absorbée et donc de celle des coefficients d'absorption et de la SO_2.

4.6 Estimation automatique de la borne inférieure du filtrage par SVD à partir du rapport contraste à bruit sur l'image de fluctuation

4.6.1 Impact de la borne inférieure sur l'estimation de la SO_2 sur les données expérimentales

Idéalement, le choix des bornes de la SVD ne doit pas impacter l'estimation de la SO_2, tant que la structure est suffisamment visible. Les simulations ont montré que dans certaines conditions de bruit et de nombre d'images disponibles, bien que l'objet soit visible, l'estimation de la SO_2 est affectée par le choix de ces bornes et notamment par la borne a, son influence n'étant pas compensée dans notre méthode contrairement à celle de la borne b. Cette borne a doit donc être judicieusement choisie : suffisamment haute pour faire apparaître les structures et filtrer l'image moyenne, et suffisamment basse pour préserver l'intégrité des structures et obtenir une estimation correcte de la SO_2. Si en simulation, les structures affectées par les artefacts de visibilité apparaissent dès que la première valeur propre est supprimée, ce n'est pas le cas dans nos expériences, car d'autres sources de fluctuation s'ajoutent à la PEF du laser.

L'influence du choix de ces bornes sur l'estimation de la SO_2 dans le cas de l'imagerie de l'embryon de poulet est présentée dans la figure 4.16. La SO_2 est moyennée dans les structures incluses dans chacune des trois ROIs (vert, jaune et rouge) en utilisant différents couples de bornes a et b lors du filtrage par SVD. La couleur des cellules représente ce niveau de SO_2. Pour cet exemple, les images semblent visuellement correctement filtrées à partir de $a = 20$ jusqu'à $a = 100$, borne pour laquelle l'image commence à être significativement dégradée.

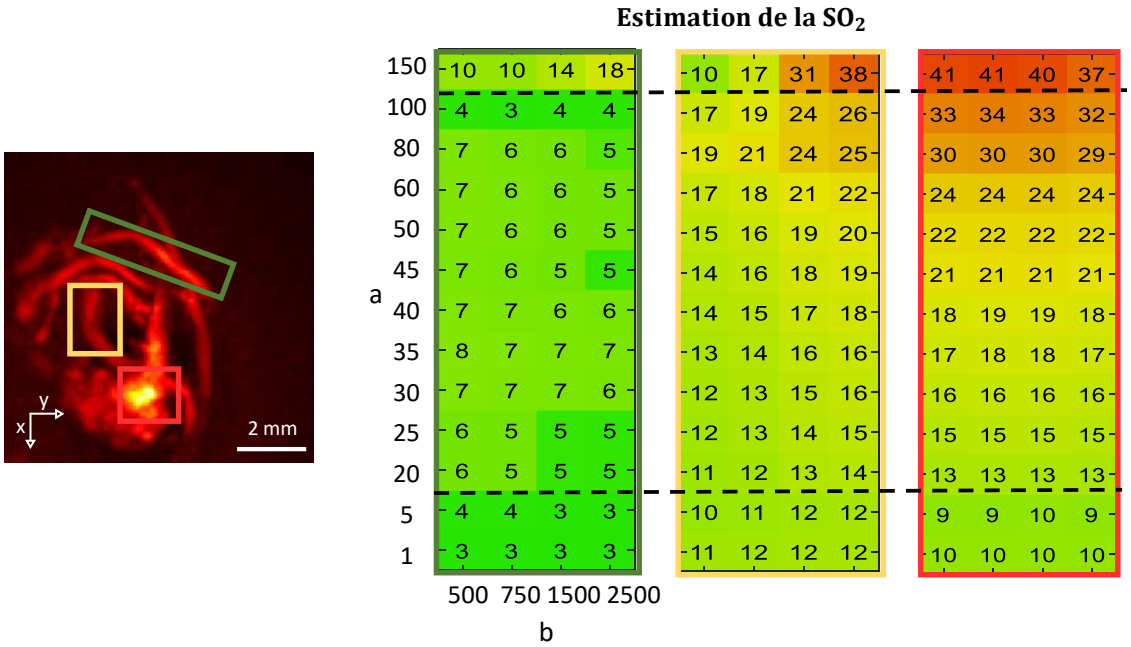

Figure 4.16 – Estimation de la SO_2 moyenne dans trois ROI pour différents couples a et b, les bornes du filtrage par décomposition en valeurs singulières.

Comme attendu, l'influence de b est négligeable. En revanche, l'impact de a est plus significatif, notamment pour la troisième ROI : la SO_2 estimée augmente avec a, à mesure que la structure est de plus en plus altérée par l'effet du filtrage. Pour cet exemple, La valeur optimale de a semble être comprise entre 20 et 40, intervalle pour lequel l'estimation de la SO_2 est relativement stable. Néanmoins, sur un critère

uniquement visuel, l'utilisateur aurait pu choisir une valeur de a beaucoup plus élevée. Une méthode automatique pour sélectionner a est donc nécessaire pour s'assurer que l'estimation de la SO_2 n'est pas biaisée par le filtrage par SVD.

4.6.2 Méthode de sélection semi-automatique de la borne inférieure a

Dans toutes les expériences, la borne b est fixée à $P/5$, un bon compromis pour conserver les vecteurs contenant de l'objet, tout en supprimant suffisamment de bruit.

La sélection de la borne a_{final}, qui sera utilisée pour le filtrage par SVD et l'estimation de l'oxygénation fournie à l'utilisateur, est basée sur le calcul d'un rapport contraste à bruit (*contrast to noise ratio*, CNR) pour différentes valeurs possibles de a. Dans nos expériences, on choisit de tester un éventail relativement grand de valeur de a possible, de 2 à 100. La borne a_{final} correspond à la valeur de a maximisant le CNR. Le CNR dépend du signal des structures fluctuantes de l'image de fluctuation, et du niveau moyen du bruit ainsi que sa variance, tel que :

$$CNR = \sum_{j}^{M} \frac{< structures\ fluctuantes(\mathbf{r}) >_{\vec{r}} - < bruit\ de\ fond(\mathbf{r}) >_{\vec{r}}}{std_{\vec{r}}(bruit\ de\ fond(\mathbf{r}))} \tag{4.18}$$

Les grandeurs définies par $structures\ fluctuantes$ et $bruit\ de\ fond$ sont des images correspondant à une longueur d'onde indexées par j. Le CNR est moyenné sur l'ensemble des M longueurs d'onde. L'obtention des grandeurs $structures\ fluctuantes$ et $bruit\ de\ fond$ est décrite dans le paragraphe suivant.

Dans un premier temps, les images de fluctuations sont calculées (à toutes les longueurs d'onde) à partir des jeux d'images PA filtrées selon les différentes bornes a envisagées. L'image de fluctuation à la longueur d'onde λ_j obtenue après filtrage par la borne inférieure a est appelée $\sigma_{a,j}$. Pour obtenir la grandeur $structures\ fluctuantes$, les structures affectées par les artefacts de visibilité et qui apparaissent dans l'image de fluctuation doivent être segmentées. Pour cela, on utilise l'image de variance $\sigma^2_{a_0,j}$, calculée à partir des images PA filtrées avec une estimation "grossière" de la borne a, a_0 définie telle que $a_0 = g_a \times P$, g_a étant un paramètre à déterminer par l'utilisateur. La valeur de a_0 est généralement plus élevée que celle de a_{final}, l'objectif étant ici d'extraire une majorité de structures fluctuantes. L'image de variance à la longueur d'onde fournissant le meilleur SNR est extraite ($\sigma^2_{a_0,j_{max}}$). Un masque, appelé $Mask_\sigma$, est calculé à partir de cette image au travers d'une binarisation dont le seuil th_σ dépend du bruit théorique σ_b^2, de l'écart type spatial de l'image et d'un second paramètre appelé g_b :

$$thr_\sigma = \sigma_b^2 + g_b \cdot std_{\vec{r}}(\sigma^2_{a_0,j_{max}}) \tag{4.19}$$

On calcule ensuite un masque plus restrictif se concentrant sur les structures non visibles sur l'image conventionnelle m_λ. Cela permet de ne pas sous estimer a lorsque la majorité des structures fluctuantes sont visibles sur l'image conventionnelle. Pour cela, le masque $Mask_m$ est calculé à partir de l'image $m_{j_{max}}$, l'image conventionnelle à la longueur d'onde fournissant le meilleur SNR. Une binarisation est réalisée avec un seuil th_m déterminé selon :

$$thr_m =< m_{j_{max}} >_{\vec{r}} + g_b \cdot std_{\vec{r}}(m_{j_{max}}) \tag{4.20}$$

La grandeur $structures\ fluctuantes(\mathbf{r})$ correspond ainsi à l'image de fluctuation $\sigma^2_{a,j}$ masquée par $Mask_\sigma - Mask_\sigma \cap Mask_m$. Cette image est principalement constituée des structures fluctuantes non visibles sur l'image conventionnelle. La grandeur $bruitdefond(\mathbf{r})$ est calculée à partir de $\sigma^2_{a,j}$ masquée par le complémentaire de $Mask_\sigma$, $\overline{Mask_\sigma}$. Elle correspond, comme son nom l'indique, au bruit de fond des images de fluctuation.

Pour une borne a donnée, la formule 4.21 résume cette définition du CNR :

$$CNR = \sum_{j}^{M} \frac{< (Mask_\sigma - Mask_\sigma \cap Mask_m) \cdot \sigma^2_{a,j} >_{\vec{r}} - < \overline{Mask_\sigma} \cdot \sigma^2_{a,j} >_{\vec{r}}}{std_{\vec{r}}(\overline{Mask_\sigma} \cdot \sigma^2_{a,j})} \tag{4.21}$$

Le CNR est ensuite calculé pour chaque borne a de la gamme prédéfinie. Les images de fluctuation calculées à partir de la borne a_{final} fournissant le meilleur CNR sont conservées et utilisées pour estimer la SO_2. La méthode est schématisée en figure 4.17.

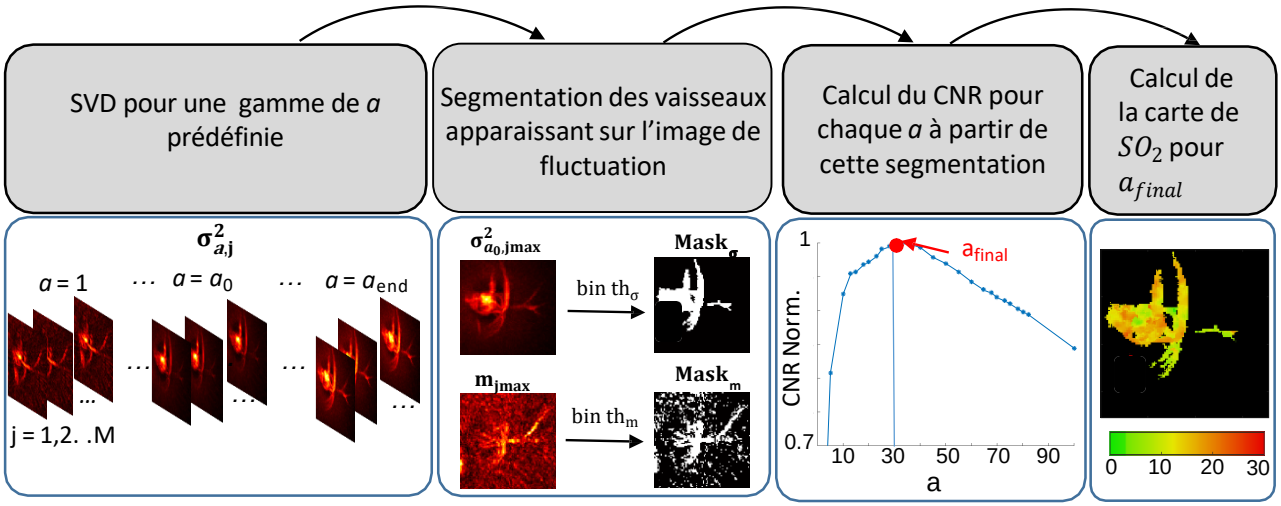

Figure 4.17 – Méthode de sélection semi-automatique de la borne inférieure a du filtrage par décomposition en valeurs singulières. Les images de fluctuation $\sigma_{a,j}$ sont calculées pour différents a à toutes les longueurs d'onde indexées par j. On extrait l'image de fluctuation filtrée avec a_0, dont la valeur est estimée préalablement par l'utilisateur. A partir de cette image et de l'image moyenne, on calcule deux masques, $Mask_\sigma$ représentant les structures fluctuantes et $Mask_m$ les structures de l'image moyenne. A partir de ces masques, une segmentation des vaisseaux fluctuants non visibles sur l'image moyenne est obtenue. Le contrast to noise ratio (CNR) est calculé pour les différents a, et le a_{final} fournissant le CNR maximal est conservé pour extraire par la suite la carte de SO_2.

Deux paramètres dépendent donc de l'utilisateur : g_a, qui détermine l'image de fluctuation "grossière" sur laquelle le masque $Mask_\sigma$ est calculé, et g_b qui détermine le seuil de binarisation lors du calcul des masques. Ces paramètres dépendent des conditions expérimentales et de la classe d'objets imagés et devront donc être déterminés manuellement par l'utilisateur. Une fois leur valeur fixée, la carte de SO_2 peut être obtenue de manière complètement non supervisée. Sur notre série d'expériences (5 zones provenant de deux embryons différents, et l'expérience *in vitro* sur les tubes), g_a est fixé à 0.03 (soit $a = 30$ pour $P = 1500$ ou $a = 75$ pour $P = 2500$). Le paramètre g_b est fixé à 3.5. Des études supplémentaires sont néanmoins nécessaires pour valider la reproductibilité de la méthode, et sa sensibilité aux paramètres g_a et g_b.

4.7 Conclusion

Dans ce chapitre, nous avons proposé une théorie et des méthodes de correction dans le but d'extraire une image proportionnelle à la densité d'énergie absorbée des images photoacoustiques de fluctuation. Vis à vis de l'imagerie photoacoustique quantitative conventionnelle, cette information est obtenue sur une image morphologiquement bien plus riche, qui fait apparaître les vaisseaux masqués par le *clutter* et les artefacts de visibilité. A partir des images de fluctuation à plusieurs longueurs d'onde, et en corrigeant par une méthode conventionnelle le terme de fluence, une grandeur proportionnelle à l'absorption optique est obtenue. Nous avons illustré l'intérêt de notre approche, appelée imagerie photoacoustique quantitative de fluctuation, en estimant la saturation en oxygène à partir des images de fluctuation, dans des situations pour lesquelles les effets de coloration spectrale peuvent être négligés. L'impact des bornes du filtrage par décomposition en valeurs singulières a par ailleurs été étudié en détail. En particulier, lorsque la borne inférieure a est surestimée, la qualité de l'estimation de l'énergie absorbée diminue. Une méthode permettant d'estimer automatiquement cette borne a a ainsi été élaborée pour éviter une telle situation.

L'imagerie photoacoustique quantitative de fluctuation a d'abord été démontrée *in silico* dans des conditions de bruit électronique et d'une fluctuation de fluence du laser très défavorables. Elle est ensuite appliquée avec succès *in vitro* et *in vivo* sur l'embryon de poulet avec notre sonde d'imagerie 3D. Des approches Doppler quantitatives sont réalisées simultanément, permettant d'estimer la direction et la vitesse du flux sanguin et d'étudier la pulsatilité. La combinaison de ces deux modalités, l'une inédite et l'autre appliquée pour la première fois en imagerie 3D avec une sonde parcimonieuse composée de seulement 256 éléments, est l'un des résultats majeurs de ces travaux de thèse.

Notre système d'imagerie et les méthodes implémentées fournissent donc plusieurs informations d'intérêt considérable dans le cadre d'applications cliniques et précliniques. Avant de telles applications, il sera nécessaire de démontrer l'applicabilité de l'approche d'imagerie photoacoustique quantitative de fluctuation dans un milieu turbide avec notre sonde d'imagerie 3D. La forte diffusivité du milieu pourra potentiellement être limitante en raison de la diminution du signal photoacoustique en profondeur et donc du ratio fluctuation à bruit. Les effets de coloration spectrale devront également être pris en compte.

5

Imagerie photoacoustique assistée par apprentissage profond

Sommaire

5.1 Contexte **130**
5.2 *Deep learning* : généralités **133**
5.2.1 Réseaux de neurones 133
5.2.2 Réseaux convolutifs 134
5.2.3 Sur et sous-apprentissage 136
5.2.4 Hyper-paramètres et leur utilisation dans les réseaux de neurones 137
5.3 Description du réseau de neurones utilisé et de la construction des jeux de données **137**
5.3.1 Constitution du jeu de données expérimentales 137
5.3.2 Constitution du jeu de données de simulation 139
5.3.3 Métriques utilisées pour évaluer les performances du réseau de neurones 140
5.3.4 *Deep learning framework* 140
5.3.5 Méthode de reconstructions alternatives 141
5.4 Correction des artefacts de visibilité par une approche *deep learning* : résultats . . **142**
5.4.1 Entraînement du réseau et évaluation sur données de simulation 142
5.4.2 Entraînement du réseau et évaluation sur données d'expérimentales 142
5.4.3 Pré-entraînement du réseau sur des données de simulation 143
5.4.4 Discussion, perspectives et conclusions 145
5.5 Quantification de l'incertitude en *deep learning* **147**
5.5.1 Sources de l'incertitude 147
5.5.2 Méthodes d'estimation existantes 148
5.5.3 Prédiction de l'incertitude aléatoire 149
5.5.4 Application des méthodes d'estimation de l'incertitude à nos données : résultats et discussion 149
5.5.5 Conclusion 151
5.6 Études complémentaires **152**
5.6.1 Influence du type de données d'entrée sur la qualité de la prédiction 152
5.6.2 Propriétés de généralisation de notre réseau de neurones 152
5.6.3 Influence de la taille du réseau 154
5.6.4 Influence de l'architecture du réseau de neurones 154

5.6.5 Data augmentation : utilisation de GAN 156
5.7 Conclusion 157

5.1 Contexte

Apprentissage profond en imagerie médicale

Dans ce chapitre, nous montrons qu'une approche basée sur l'apprentissage profond (*deep learning*), peut être utilisée pour corriger les artefacts de visibilité en imagerie photoacoustique.

Le *machine learning*, soit la découverte et l'apprentissage de motifs dans des données afin d'effectuer de manière autonome une tâche, ou réaliser des prédictions, s'est fortement développé ces dernières années dans le domaine biomédical [118]. Jusque récemment, les caractéristiques (les données) utiles à la réalisation de la tâche était principalement extraites et conçues 'à la main' par des experts grâce à leur connaissance du domaine. L'introduction de l'apprentissage profond [10] a permis d'incorporer ces étapes d'extraction et de conception dans la phase d'apprentissage. Le *deep learning* s'appuie sur un algorithme appelé réseau de neurones artificiels s'inspirant du cerveau humain. Ce réseau est composé de dizaines de couches de neurones, chacune recevant et interprétant les informations de la couche précédente. Son récent succès a plusieurs origines : améliorations technologiques des CPUs et GPUs permettant d'accélérer les calculs, augmentation de la quantité de données disponibles, et développement de nouveaux algorithmes d'apprentissage [76]. Les approches *deep learning* sont ainsi devenues incontournables dans de nombreux domaines de l'intelligence artificielle, notamment pour résoudre des problématiques de vision par ordinateur telles que la segmentation [73], le recalage [142], la fusion d'images [126], l'annotation d'images [119], ou encore la détection de lésion [98].

Apprentissage profond en imagerie photoacoustique

L'utilisation d'algorithme de deep learning (DLA) s'est également développée très récemment en imagerie PA. La figure 5.1 résume cette intérêt grandissant.

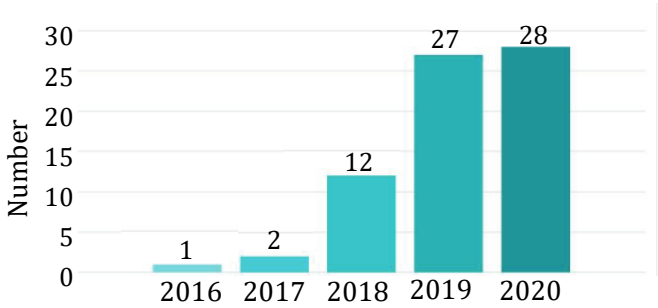

Figure 5.1 – Nombre de productions scientifiques incluant la combinaison d'au moins un des termes suivants : learning/network/Net/learned/convolution/GAN/supervise/unsupervise avec l'un des termes suivant : photoacoustic/optoacoustic [149]. En 2020, les productions scientifiques ont été comptabilisées jusqu'en septembre. Au commencement de ma thèse (2018) 6 articles avaient été publiés.

Des résultats intéressants ont été obtenus pour la réalisation de diverses tâches, telles que la reconstruction directe de la pression initiale [133], la correction d'artefacts liés à un sous échantillonnage des données [3, 24, 56], la suppression d'artefacts de réflexion [45], la localisation de point source [1, 104] ou encore en imagerie quantitative [17, 72]. Certaines études, menées simultanément à notre étude présentée dans ce chapitre, ont également montré qu'il était possible de réduire les artefacts de visibilité de type *limited view* [3, 24, 47]. Cependant, ces études présentes des limites : l'approche est utilisée principalement en simulation, et appliquée sur des données expérimentales dans de rares cas mais sans *ground truth* [47]. On définit ici la *ground truth* comme l'image sans artefacts. Les études [3] et [24] disposent d'une *ground truth* expérimentale mais visent principalement à résoudre un problème de sous échantillonnage spatial, dans des situations où les artefacts de visibilité sont peu discernables (image XY

dans [3], objet très complexe dans [24]). Dans ces études, des systèmes de détection non conventionnels sont par ailleurs utilisés (de type *ring array* dans [24] et de type fabry-perot dans [3]).

Différentes stratégies d'inclusion du réseau de neurones en imagerie photoacoustique

A partir des signaux radiofréquences de l'objet, acquis ou simulés, plusieurs stratégies ont été proposées pour obtenir une image améliorée en incluant un algorithme de *deep learning* dans le processus de reconstruction. Ces stratégies sont résumées dans la figure 5.2.

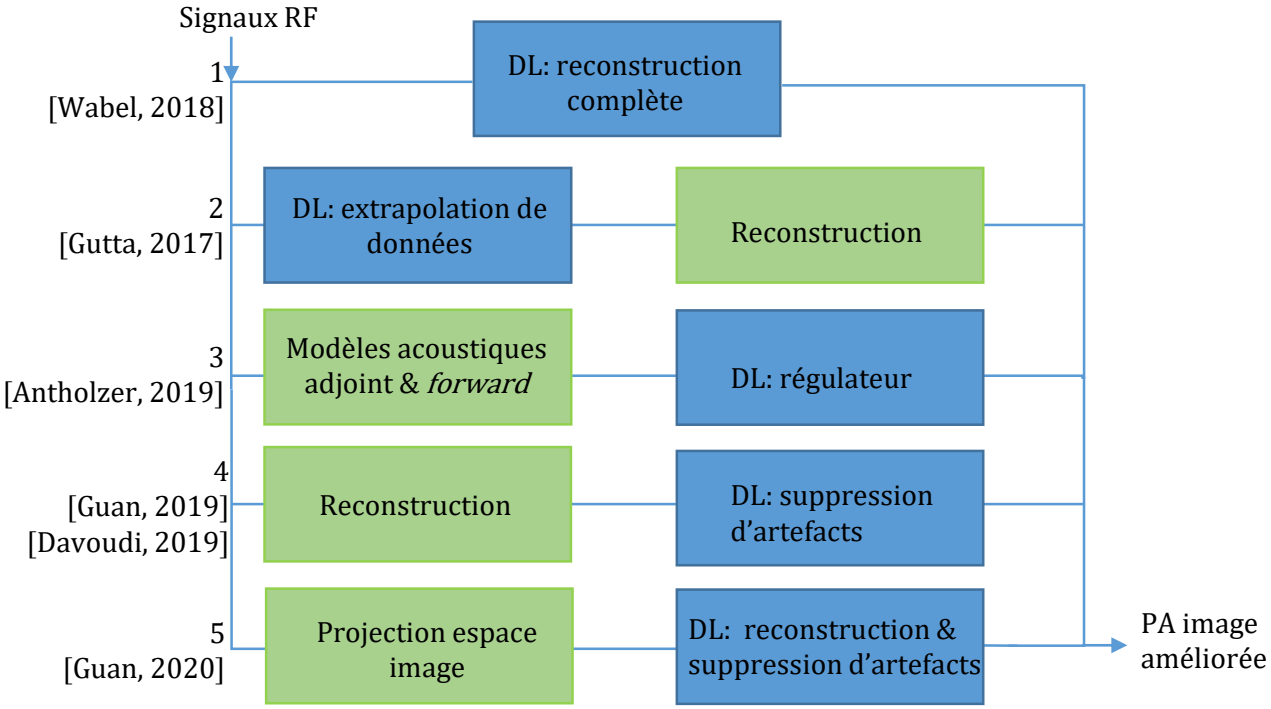

Figure 5.2 – Différentes approches pour inclure l'algorithme de deep learning (DL) dans le processus de reconstruction de l'image photoacoustique.

1. L'approche la plus directe consiste à confier au DLA la réalisation de la reconstruction à partir des données brutes [133] : aucune information, telle que la position des capteurs ou la vitesse du son, n'est utilisée. L'opérateur de propagation est donc estimé par le DLA. Si des résultats intéressants en simulation ont été produits par cette approche, elle semble être l'une des moins performante.
2. Dans [53], les auteurs proposent d'augmenter la réponse fréquentielle de leur sonde d'imagerie à l'aide d'un réseau de neurone, avant de reconstruire l'image de manière conventionnelle.
3. Le réseau de neurones peut également être utilisé pour produire le terme de régularisation d'une méthode de reconstruction itérative [3]. Cette approche nécessite néanmoins d'importantes connaissances à priori, notamment sur le milieu de propagation et la *point spread function* du système.
4. La méthode la plus couramment utilisée consiste à réaliser une première reconstruction avec un algorithme conventionnel tel que le *delay-and-sum*, et d'utiliser le réseau avec cette image en entrée pour en améliorer les caractéristiques. C'est cette approche qui est privilégiée dans notre étude.
5. Alternativement, l'information reçue par chaque capteur est projetée dans l'espace image, et l'ensemble de ces projections est fourni au réseau [49].

Objectifs de ce chapitre

Nous allons démontrer que les problèmes de visibilité de type *limited view* et de type *limited bandwidth* peuvent être corrigés à l'aide d'un algorithme de *deep learning* appliqué à des données expérimentales avec une sonde d'imagerie conventionnelle, ce qui n'avait jamais été établi avant ces travaux. Par ailleurs, une approche inédite en imagerie photoacoustique permettant d'estimer l'incertitude de la prédiction est proposée, dans le but de localiser les erreurs produites par l'algorithme lorsque les données sont mal interprétées, et ainsi améliorer la fiabilité des prédictions. Nos approches sont validées sur des signaux photoacoustiques produits par une famille d'objets complexes mimant des réseaux vasculaires, d'abord à partir de données de simulation puis sur des données expérimentales. L'intérêt d'un pré-entraînement sur les données de simulation pour reconstruire des images expérimentales est également démontré, permettant de réduire le taille du jeu de données d'entraînement expérimental. L'influence du type de données d'entrée sera également étudiée, ainsi que celle de la taille et du type du réseau. Une extension du jeu de données d'entraînement de simulation à l'aide d'un *Generative adversarial network* (GAN) est également proposée.

Dans une première partie, les principes généraux du *deep learning* sont introduits. Dans une seconde partie, le réseau utilisé dans notre étude et la construction des jeux de données sont décrits. Les résultats mettant en évidence l'amélioration de la visibilité sont présentés et discutés dans une troisième partie. Nous étudions également dans cette troisième partie l'intérêt d'une session de pré-entrainement sur des données de simulation pour prédire des données expérimentales. La quatrième partie se concentre sur l'estimation de l'incertitude de prédiction, en présentant les méthodes retenues et les résultats obtenus. Une dernière partie résume les résultats des diverses études supplémentaires réalisées sur notre jeu de données, notamment : l'influence du type de données d'entrée, l'évaluation de différents types de réseau, la capacité de généralisation de notre réseau, et la possibilité d'utiliser un *generative adversarial network* pour augmenter la taille du jeu de données de simulation.

5.2 *Deep learning* : généralités

Dans cette section, les principes du *deep learning* sont présentés. En particulier, les réseaux convolutifs, utilisés dans notre étude, sont décrits. Plusieurs notions, utiles à la compréhension de ce chapitre seront également introduites. Une partie des résultats présentés dans ce chapitre a été publiée dans [43].

5.2.1 Réseaux de neurones

Le *machine learning* est un domaine de l'intelligence artificielle qui consiste en la programmation d'une machine pour que celle-ci apprenne à réaliser des tâches de manière autonome en analysant et en s'entraînant sur des données. En *deep learning*, le modèle est un réseau de neurones artificiel. L'unité de base est le neurone lequel représente une fonction de transfert, dont le fonctionnement est inspiré des neurones biologiques. Il est schématisé en figure 5.3. Dans une première étape, les sources d'information, les données d'entrée du neurone, sont multipliées par des coefficients, lesquels correspondent aux paramètres du modèle, avant d'être sommées entre elles. Le résultat de cette somme est ensuite transformé par une fonction dite d'activation. L'objectif est alors de trouver les poids W tels que la sortie "y" réponde à l'objectif voulu.

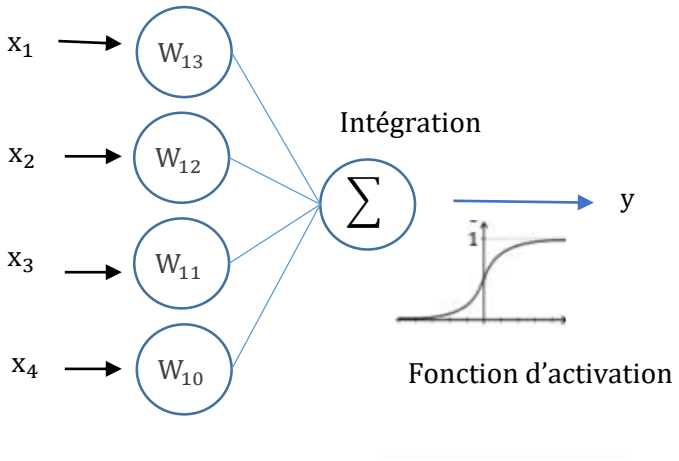

Figure 5.3 – Représentation d'un neurone artificiel, modèle du perceptron.

Le modèle du neurone isolé appelé perceptron [108] fonctionne comme une régression linéaire et ne permet donc pas de résoudre des problèmes non linéaires complexes. Cette limitation fut supprimée par l'apparition du perceptron multicouches à rétro-propagation [111] associée à l'utilisation de fonctions d'activation non linéaires. Dans ce cas, le théorème d'approximation universel garantit l'existence des poids W [59], quelque soit l'objectif assigné.

L'entraînement du réseau se déroule en plusieurs étapes, présentées en figure 5.4. Tout d'abord, la *forward propagation* : chaque neurone du réseau reçoit des données d'entrée, les traitent, et leurs données de sortie sont transmises à la couche suivante, jusqu'à produire "y_{final}", la sortie du réseau. En apprentissage supervisé, le "y_{true}" est connu : c'est la donnée de référence que l'on appellera *ground truth*. La distance, ou l'erreur entre la sortie du réseau "y_{final}" et cette sortie désirée y_{true}" peut alors être calculée au travers d'une fonction de coût. L'étape suivante est la rétro-propagation, au cours de laquelle les gradients de cette fonction de coût sont calculés pour chaque neurone, de la dernière couche à la première. Les paramètres W sont alors mis à jour, dans la direction inverse du gradient, plus ou moins significativement en fonction de l'erreur engendrée. Ce processus se répète pendant un certain nombre d'itérations jusqu'à ce que le réseau converge. Le paramètre α, le *learning rate*, permet de contrôler la vitesse d'adaptation du modèle.

Le réseau est entraîné à partir du jeu de données d'entraînement, le *training set*. Pour détecter un éventuel sur-apprentissage (cf section 5.2.3), un jeu de données de validation, le *validation set*, est souvent défini. Les données le composant ne sont pas utilisées pour l'entraînement, mais lors de la phase de conception pour évaluer de manière non biaisée les performances du modèle en fonction de diverses combinaisons d'hyper-paramètres, les variables déterminant la structure du réseau et la manière dont il est entraîné. L'efficacité du modèle final sera évaluée sur le jeu de données de test, le *test set*, constitué d'exemples non utilisés pour l'entraînement du réseau et le choix de ses hyper-paramètres. Cette distinction permet de s'assurer de la capacité de généralisation du réseau, c'est à dire de réaliser la tâche confiée sur des données d'entrée totalement inconnues : le *test set* est ainsi représentatif des performances atteintes par le réseau en production. Ces trois jeux de données, bien qu'indépendants, sont supposés être constitués d'exemples générés à partir d'une même distribution aléatoire.

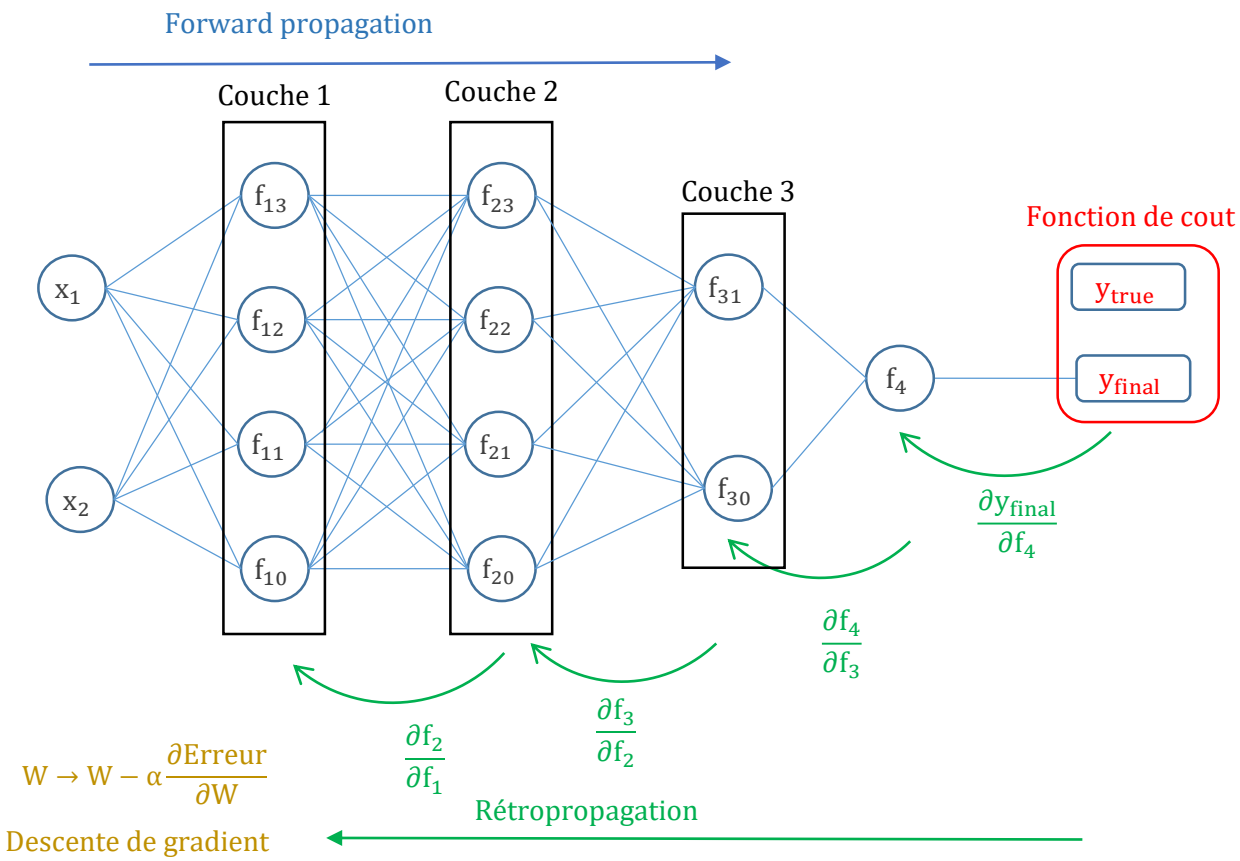

Figure 5.4 – Représentation d'un réseau de neurones artificiel composé de trois couches cachées, et des quatre étapes résumant son entraînement : la forward propagation, le calcul du coût, la rétro-propagation et la descente de gradient. x représente les données d'entrée, f les opérations réalisées par les neurones, y_{true} la ground truth et y_{final} la sortie du réseau.

5.2.2 Réseaux convolutifs

L'apparition des réseaux convolutifs (en anglais *convolutional neural network*, CNN) dans les années 90 a révolutionné l'utilisation du *deep learning* en traitement d'image. Plutôt que de vectoriser les données d'entrée de chaque neurone avant d'appliquer les poids W, entraînant inévitablement la destruction de l'information structurelle partagée par des pixels voisins, des opérations de convolution sont réalisées dans le but de capturer ces dépendances spatiales. Les valeurs des noyaux de convolution correspondent

aux paramètres W à optimiser. Le résultat de la convolution est appelé "carte de caractéristique", laquelle est transformée par une fonction d'activation. La seconde opération distinctive des CNNs est le *pooling* consistant à sous échantillonner ces cartes de caractéristiques, dans le but d'extraire les informations les plus importantes, et de diminuer la puissance de calcul nécessaire à l'entraînement du réseau. La manière la plus classique de réaliser ce sous échantillonnage est de subdiviser ces cartes en sous zones et de conserver la valeur maximale : c'est le *max pooling*. La sortie d'un CNN peut prendre de nombreuses formes : un unique chiffre indiquant par exemple la présence ou non d'un objet sur une image (classification), un vecteur de taille $N_{class} \times 1$ indiquant la présence ou non de N_{class} objets sur une image (classification multiclasse), ou encore une image. La figure 5.5 résume le fonctionnement d'un CNN. Seule la partie encodage est présentée, c'est à dire la partie où sont principalement extraites les caractéristiques représentant les données d'entrée. Les trois points sur le schéma peuvent remplacés par n'importe quelle(s) opération(s) permettant d'obtenir la sortie souhaitée.

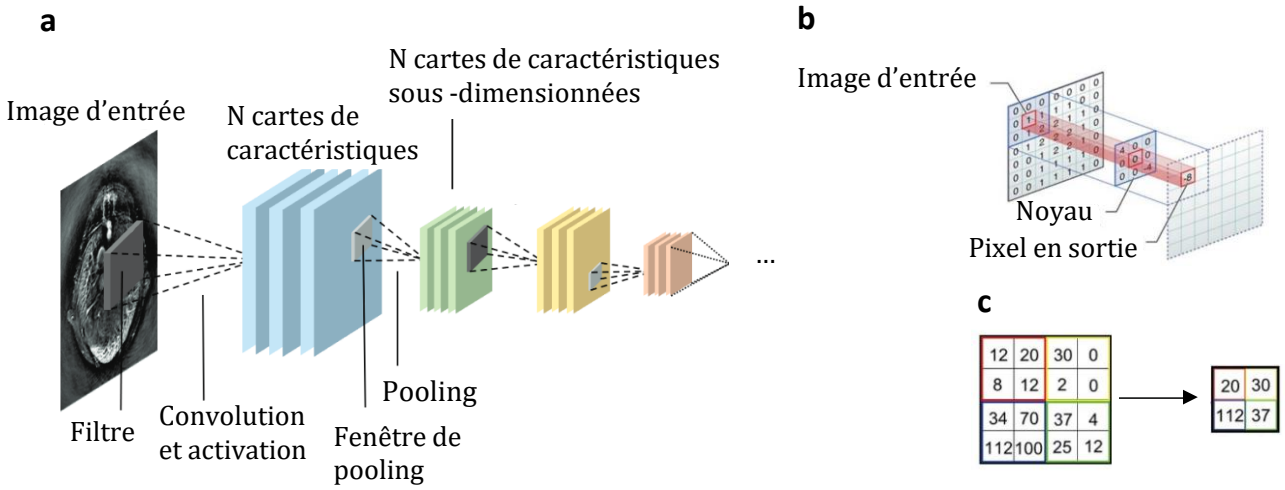

Figure 5.5 – (a) : Représentation de la partie encodage d'un réseau convolutif. Les trois points peuvent être remplacés par n'importe quelle(s) opération(s) permettant d'obtenir la sortie souhaitée. (b) : opération de convolution. (c) : opération de max-pooling.

Plus le réseau est profond, plus les cartes de caractéristiques sont complexes et abstraites, comme le montre la figure 5.6 représentant les cartes de caractéristiques d'un réseau avec une image de chat en entrée.

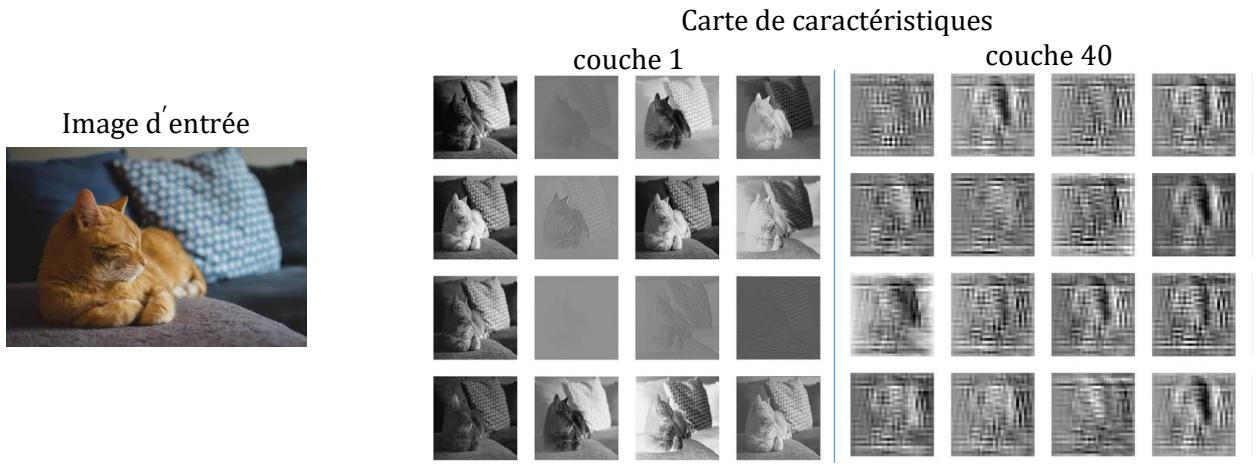

Figure 5.6 – Cartes de caractéristiques d'un réseau convolutif avec une image de chat en entrée. Source : https ://debuggercafe.com/visualizing-filters-and-feature-maps-in-convolutional-neural-networks-using-pytorch/.

Les cartes correspondant à la première couche se concentrent sur différents aspects de l'image, tels que des variations de teinte, de coloration, de forme ou encore sur les zones de haute fréquence. Dans la couche plus profonde, ces cartes semblent illisibles, mais contiennent en réalité des motifs que le réseau de neurones sera capable de reconnaître et d'utiliser pour par exemple prédire l'espèce de l'animal présent sur l'image d'entrée.

5.2.3 Sur et sous-apprentissage

La mauvaise performance d'un algorithme de DL, qui est traduite par l'écart entre la prédiction et la *ground truth*, a souvent pour cause un sous ou sur apprentissage (*underfitting* et *overfitting*). La figure 5.7 illustre ces notions dans le cadre d'un problème de séparation entre deux classes d'objets, le modèle est représenté par la courbe rouge. Dans le premier cas, les performances de l'algorithme sont insuffisantes sur le *training set* : l'écart entre le modèle et les données ne diminue pas suffisamment pendant la phase d'entraînement. Cela peut être due à l'utilisation d'un modèle trop simple vis à vis de la complexité des données et de la tâche à réaliser, où lorsque la régularisation est trop importante (figure 5.7, a). Dans le second cas, le réseau est bien moins performant sur les données de test/validation que sur celles d'entraînement. Le modèle est trop représentatif des données d'entraînement (figure 5.7, b) et ne peut pas réaliser de prédiction fiable à partir de nouvelles données qui n'ont pas été vues pendant la phase d'entraînement.

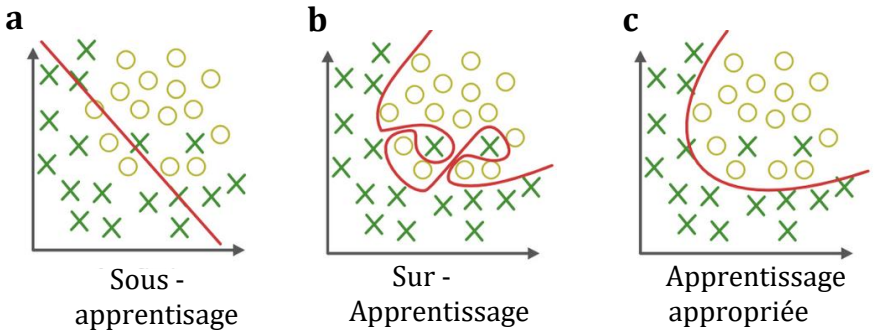

Figure 5.7 – Illustration de l'underfitting et de l'overfitting sur un problème de classification visant à séparer deux classes d'objets représentés par des ronds et des croix. Le modèle est représenté par le courbe rouge. (a) : le modèle underfit, la séparation est peu précise. (b) : le modèle overfit, le modèle est trop dépendant des données, et sujet aux outliers. (c) : le modèle semble bien entraîné : la qualité de la séparation est correcte et les outliers ne sont pas considérés.

L'*overfitting* est l'une des principales problématiques en *deep learning*. Il est possible de le détecter au cours de l'entraînement lorsque les performances commencent à stagner ou diminuer sur le *validation set*, mais continuent d'augmenter sur le *training set*. Une manière de l'éviter est de stopper l'entraînement avant que cette divergence apparaisse (*early stopping*). La limitation de l'*overfitting* peut également être obtenue par la diminution de la complexité du modèle, en diminuant le nombre de couches ou de neurones par exemple. L'augmentation de la taille du *training set* peut aussi réduire l'*overfitting*. Une multitude de méthodes de régularisation ont également été développées dans ce but. L'utilisation de couche *dropout* [124] est l'une des plus populaires. Une partie des neurones, choisis aléatoirement à chaque itération, sont désactivés. Cela limite une trop forte coadaptation des neurones, forçant l'apprentissage de caractéristiques plus robustes. La méthode de *batch normalization* [62] consiste à normaliser les données d'entrée de chaque neurone en soustrayant la moyenne du *batch* et en le divisant par son écart type. Un *batch* correspond à un certain nombre d'exemples fournis simultanément au réseau. En plus d'accélérer la convergence de l'entraînement, ce procédé injecte comme l'approche *dropout* du bruit dans le modèle.

5.2.4 Hyper-paramètres et leur utilisation dans les réseaux de neurones

Le taux d'apprentissage (*learning rate*) qui définit la vitesse à laquelle est réalisée la descente de gradient, la profondeur du réseau (le nombre de couche), le nombre de neurone, la fonction de coût, les fonctions d'activation ou encore la méthode d'initialisation des paramètres et les méthodes de régularisation utilisées, caractérisent un réseau, et constituent les hyper-paramètres du modèle. Ils dépendent de nombreux éléments (tâche à effectuer, taille du jeu de données, puissance de calcul ..) et seront optimisés, généralement manuellement, lors de la conception du réseau.

La taille de l'ensemble de données fournies à chaque itération, le *batch size*, est l'un de ces hyper-paramètres. Si la mémoire disponible est suffisante, la totalité du jeu de données d'entraînement est généralement transmise et propagée dans le réseau à chaque itération pour obtenir la meilleure estimation de la distance entre le modèle et nos données, fournie par la fonction de coût. En pratique, à chaque itération, un sous-ensemble d'exemples appartenant au *training set* est généralement fourni au réseau : l'estimation de l'erreur et donc la mise à jour des paramètres en ressort bruitée. La présence de ce bruit n'est pas nécessairement négatif, offrant au contraire une manière de régulariser le modèle et d'améliorer la généralisation du réseau.

5.3 Description du réseau de neurones utilisé et de la construction des jeux de données

L'objectif de notre étude est de produire, à partir d'une image affectée par les artefacts de visibilité, une image corrigée. Pour réaliser des prédictions sur un *test set* de simulation, le réseau est entraîné sur un *training set* de simulation. Pour réaliser des prédictions sur un *test set* expérimental, le réseau peut être entraîné sur un *training set* expérimental, ou sur un *training set* de simulation, ou encore sur les deux successivement par une approche de *transfer learning*. Dans cette partie, les méthodes de préparation des jeux de données de simulation et d'entraînement seront présentées, ainsi que les spécificités du réseau de neurones employé dans ce travail de thèse.

5.3.1 Constitution du jeu de données expérimentales

L'objet imagé est un squelette de feuille, dont les capillaires ont été remplis d'encre pour obtenir un fort signal photoacoustique. Le limbe est dissous par traitement chimique. Les plus petits capillaires sont coupés manuellement pour supprimer les détails non résolvables, en dessous de la limite de résolution du système d'imagerie. Comme le montre la figure 5.8, un objet complexe, multi-échelle et mimant un réseau vasculaire est obtenu.

Dans notre étude, les jeux de données (*training set*, *validation set* et *test set*), expérimentaux ou de simulation, sont composés d'une série de paire d'images de taille 5 cm $\times$ 5 cm. Chaque paire est constituée d'une image reconstruite par *delay-and-sum* (que l'on nomme "image *beamformée*"), la donnée d'entrée du réseau, et d'une photographie de l'objet prise à l'aide d'une camera CMOS (X-E2, Fujifilm, Japan), la *ground truth*. Pour construire les jeux de données expérimentaux, les signaux radiofréquences permettant de reconstruire chaque image d'entrée sont obtenus à partir du protocole suivant. Un squelette de feuille est maintenu horizontalement dans un gel d'agarose à l'intérieur d'une cuve remplie d'eau dégazée et déionisée. Notre sonde d'imagerie 2D est couplée à la cuve à l'aide d'une fine membrane de Mylar, de propriétés acoustiques proches de celles de l'eau. La feuille est donc dans le plan d'imagerie de la sonde. La longueur d'onde d'illumination est 532 nm. Ce protocole est schématisé en figure 5.8 (b).

En général, l'image photoacoustique en entrée du réseau de neurones est l'image photoacoustique conventionnelle, c'est à dire l'image reconstruite en appliquant l'algorithme de *delay-and-sum* aux signaux radiofréquences sous leur forme en quadrature, opération qui produit une image complexe dont le module est calculé. Comme présentée dans le chapitre 1 section 1.3.4, cette démarche permet de supprimer les oscillations induites par la réponse fréquentielle de la sonde, et ainsi produire une reconstruction plus conforme à l'objet. Nous appelons cette image "image dmBF" pour "image *beamformée* démodulée". Cependant, dans notre étude, l'image d'entrée est reconstruite en appliquant l'algorithme de *delay-and-sum* aux signaux temporels bruts. Cela produit donc une image qui contient les oscillations axiales produites par la réponse impulsionnelle de la sonde. Nous choisissons de nommer cette image "image mBF" pour "image *beamformée* modulée". Une comparaison des performances du DLA en fonction du type de données d'entrée (image mBF ou image dmBF) est présentée en section 5.6.1. Les deux types d'images sont visible en figure 5.8 (a).

Les photographies sont converties en échelles de gris (8 bits) et les pixels en dessous d'un seuil sont fixés à 0 pour supprimer le bruit de fond. Un recalage de la photographie avec l'image photoacoustique est nécessaire pour disposer d'une correspondance entre les pixels de l'image en entrée et ceux de la *ground truth*. L'amplitude des transformations (rotation, translation, mise à l'échelle) nécessaires pour réaliser ce recalage est obtenue automatiquement. Pour cela, on calcule les coefficients de corrélation entre l'image photoacoustique dmBF et les photographies transformées pour chaque combinaison possible d'angles, de pas de décalage et de facteurs de grandissement dont les plages de possibilités sont définies préalablement. On utilise pour cette opération l'image photoacoustique dmBF car celle-ci plus proche de l'objet que l'image mBF. La photographie transformée pour laquelle ce coefficient de corrélation est maximal constitue la *ground truth* de la paire d'images.

Une première série de 593 paires d'images est obtenue en scannant deux squelettes de feuille. Ce scan peut être représenté par le décalage du carré rouge situé sur la photographie de la figure 5.8 dans l'une des directions x et z, décalage qui correspond en réalité au décalage de l'objet par rapport à la sonde. Entre chaque acquisition, ce décalage est d'une demi-dimension de l'image dans l'une des deux directions, soit 2.5 cm. Pour augmenter le nombre d'acquisitions possibles, l'objet est tourné de 90 degrés. Chaque zone de l'objet peut finalement être "vue" jusque 8 fois sous une forme différente. Le jeu de données est séparé en un *training set* de 500 images et un *validation set* de 93 images. Pour constituer le *test set*, une deuxième série de 15 paires est obtenue à partir de deux autres squelettes de feuille, donc différents de ceux utilisés pour l'entraînement.

La constitution d'un tel jeu de données a demandé un effort conséquent en terme de développements expérimentaux, mais aussi algorithmiques et donc de temps. Nous montrons plus tard dans ce chapitre (section 5.4.3) qu'il est possible de réduire la taille du *training set* en utilisant des données de simulation.

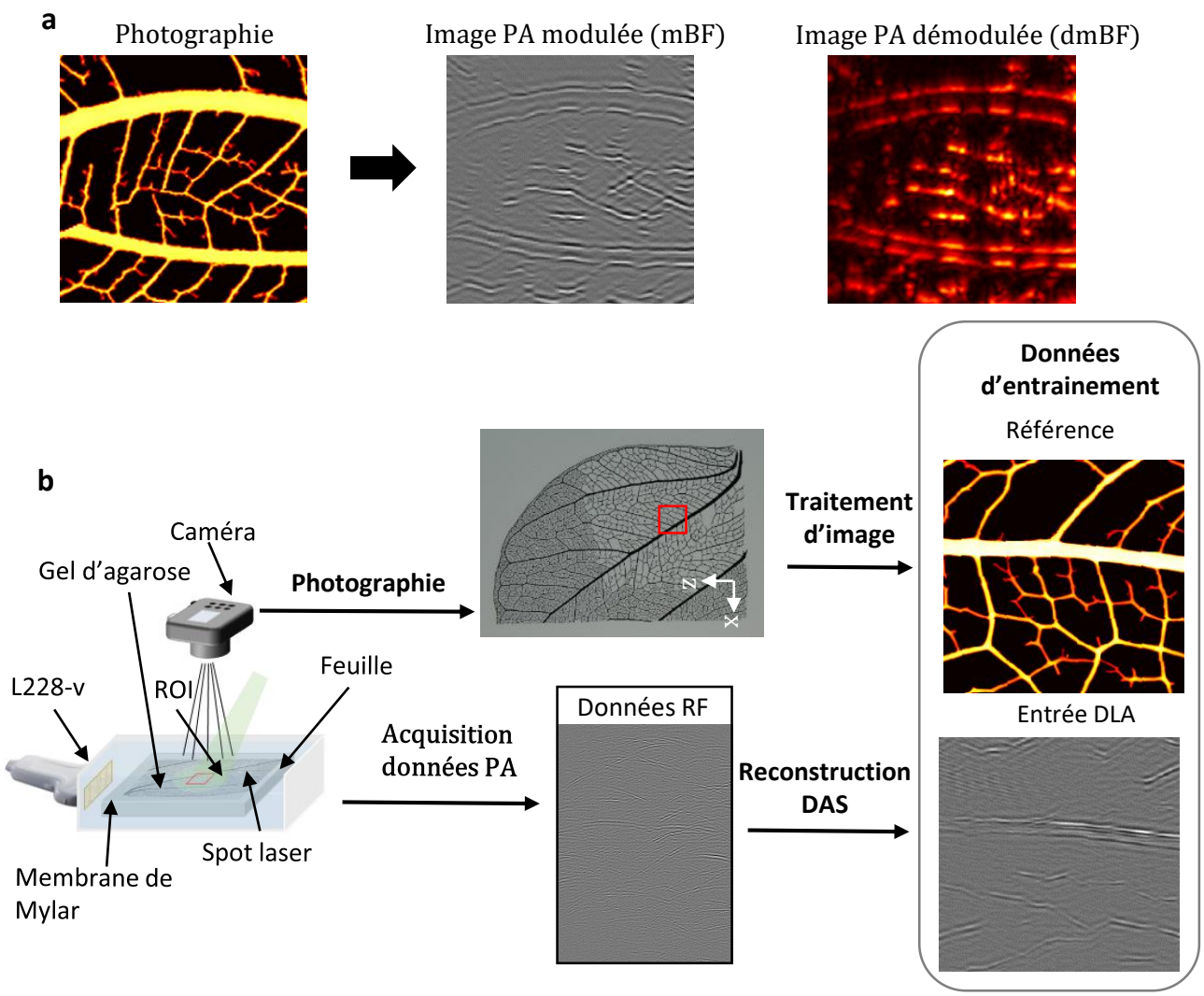

Figure 5.8 – (a) : présentation de l'image photoacoustique dite "modulée" (mBF) et de l'image photoacoustique dite "démodulée" (dmBF). (b) : constitution des jeux de donnée. Une série d'acquisitions photoacoustiques est réalisée pour plusieurs zones de l'objet, une zone est représentée par le carré rouge. L'image est reconstruite par delay-and-sum puis recalée avec la zone de la photographie correspondante pour former une paire du jeu de données.

5.3.2 Constitution du jeu de données de simulation

Les photographies des squelettes de feuille sont utilisées pour simuler des images photoacoustiques à partir de la méthode de décalage des signaux radiofréquences décrite dans l'annexe B. Une augmentation artificielle du dataset est réalisée en appliquant de manière aléatoire des rotations, des transformations miroirs, des déformations de type cisaillement et des déformations de type contraction/expansion. Chaque zone du squelette de feuille peut être vue jusque 48 fois sous une forme différente. Une première série de 1400 paires d'images est simulée à partir des photographies des deux squelettes de feuille du *training set* expérimental pour constituer le *training set* de simulation. Une seconde série de 15 paires d'images est simulée à partir des photographies des deux squelettes de feuille du *test set* expérimental pour constituer le *test set* de simulation.

5.3.3 Métriques utilisées pour évaluer les performances du réseau de neurones

Comme mentionné dans la section précédente, les mêmes squelettes de feuille sont utilisés pour créer le *test set* expérimental et celui de simulation. Les mêmes zones de ces objets vont être utilisées pour décrire et présenter les résultats obtenus, ce qui permet de comparer visuellement les résultats expérimentaux et de simulation sur les mêmes structures. Toutes les images sont normalisées par leur maximum et représentées avec une barre de couleur de 0 à 1. La performance est calculée à l'aide de deux indices : une cross-corrélation 2D normalisée (NCC) [151] et un indice de similarité structurel (*scaled and shifted structured similarity index*, sSSIM), calculés entre la sortie du réseau (la prédiction) et la *ground truth* (la photographie). Ces scores sont également calculés pour l'image photoacoustique dmBF, c'est à dire l'image conventionnelle. Le SSIM [139] est une métrique très utilisée pour évaluer la qualité perçue d'une image. Il quantifie localement les similitudes en terme de contraste, de structure et d'intensité. Le sSSIM [116], calculé selon l'équation 5.1 est un dérivé du SSIM, qui permet d'obtenir un score non biaisé ne défavorisant pas l'image photoacoustique dmBF vis à vis de la prédiction du réseau.

$$sSSIM(Y, Y_{true}) = SSIM(\alpha(Y) - \beta, Y_{true}) \tag{5.1}$$

avec Y la prédiction du réseau ou l'image photoacoustique dmBF, et Y_{true} la *ground truth*. Les valeurs de α et β sont obtenues en maximisant $sSSIM(Y_{pred}, Y_{true})$.

5.3.4 *Deep learning framework*

Propriétés et caractéristiques du Réseau U-Net conventionnel

Le type de réseau utilisé dans notre étude est un U-Net, proposé pour la première fois en 2015 [107] pour la segmentation d'images de cellules. Il est particulièrement adapté aux jeux de données de taille limitée, une situation fréquente en imagerie médicale qui s'applique à notre étude. Le U-Net est un CNN qui présente une architecture symétrique décomposable en deux sous-parties, l'encodeur et le décodeur, reliées entre elles par des connections. L'encodeur fonctionne comme les CNN présentés dans la section 5.2.2, et a donc pour but de capturer le contexte de l'image, c'est à dire d'extraire des caractéristiques de plus en plus complexes et abstraites à mesure que l'on progresse dans les couches. Le décodeur permet, à partir de ces cartes de caractéristiques, de reconstruire l'image par une succession de convolutions. Le bloc de base du décodeur inclut une déconvolution et une opération de sur-échantillonnage pour retrouver la résolution de l'image d'entrée, qui remplace l'opération de *pooling* de l'encodeur. Des connections entre l'encodeur et le décodeur permettent de transférer efficacement les informations entre ces deux sous-réseaux. Ces connections consistent en la combinaison des cartes de caractéristiques de l'encodeur avec celles du décodeur par une convolution, à un même niveau de résolution (voir figure 5.9, flèches noires entre les deux parties du réseau). Comme pour la plupart des CNN, les fonctions d'activation après chaque convolution sont des ReLU (*Rectified linear unit*), qui présentent l'avantage vis à vis d'autres fonctions d'activation, de réduire le risque de fuite et d'explosion des gradients et d'accélérer la convergence.

Description du réseau utilisé dans ces travaux de thèse

Dans le cadre de ce travail de thèse, l'architecture originale a été légèrement modifiée. Des opérations de *batch-normalisation* ont été introduites ainsi que des couches *dropout*, dans le but d'améliorer les capacités de généralisation du réseau et d'accélérer la convergence. Les couches *dropout* sont par ailleurs nécessaires à l'estimation de l'incertitude, dont la méthode sera décrite dans la section 5.5. Le réseau U-Net produit originellement une carte de segmentation binaire. Puisque nous souhaitons reconstruire une image non binaire, une ReLU est utilisée comme dernière fonction d'activation plutôt qu'une sigmoïde (la sigmoïde favorisant une sortie binaire). L'opérateur de sur-échantillonnage est une interpolation au plus proche voisin (en anglais *nearest interpolation*) suivie d'une convolution. La fonction de coût est la moyenne des écarts au carré (MSE) entre l'image prédite par le réseau et la ground truth,

plus adaptée au problème de reconstruction que la fonction de coût originellement utilisée. L'algorithme d'optimisation est ADAM [71], une extension de la méthode de descente de gradient stochastique, avec un taux d'apprentissage de 5×10^{-4}. Une représentation du réseau est proposée en figure 5.9.

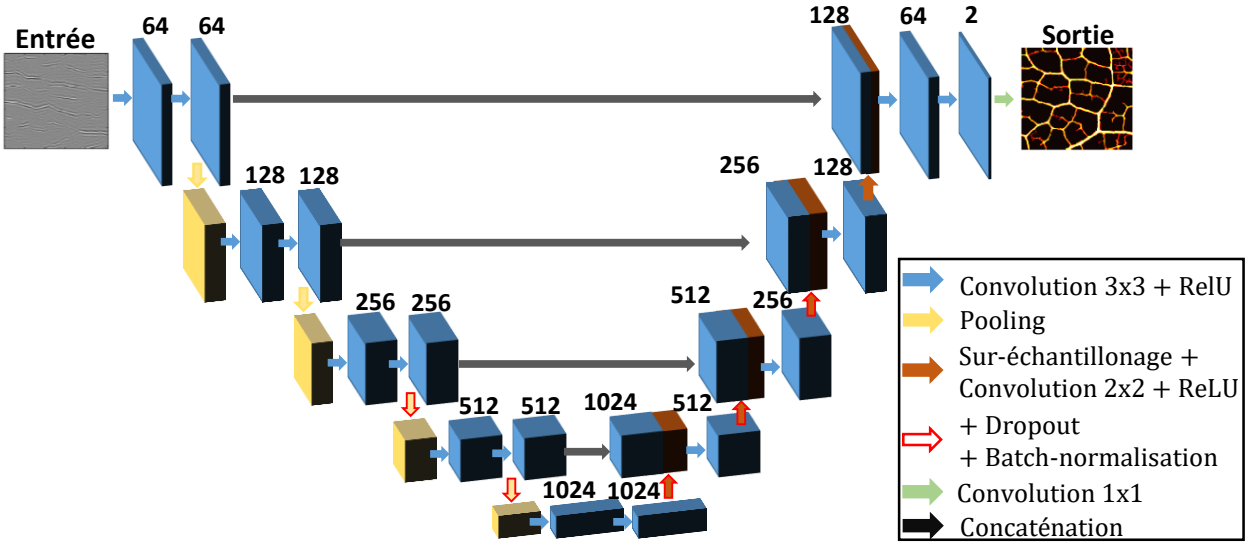

Figure 5.9 – Réseau de type U-Net utilisé dans cette étude. On distingue l'encodeur (à droite, flèches "descendantes"), le décodeur (à gauche flèches "montantes").

Cette architecture U-Net, adaptée à notre problématique, reste néanmoins très conventionnelle. Le réseau est implémenté en Python à l'aide de la librairie Tensorflow. Une grande partie des hyper-paramètres, notamment la profondeur du réseau et le nombre de filtres par couche, ont été déterminés à l'aide d'une optimisation bayésienne, implémentée à l'aide de la librairie GPyOpt. C'est une méthode automatisée efficace pour optimiser rapidement ces hyper-paramètres car particulièrement adaptée à ce type de problème non convexe et bruité pour lequel chaque évaluation est très coûteuse en temps.

5.3.5 Méthode de reconstructions alternatives

A titre de comparaison, l'image photoacoustique dmBF ainsi qu'une image déconvoluée de type L2 sont fournies. La déconvolution est réalisée par une approche de minimisation moindre carré, régularisée par norme L2, à l'aide de l'algorithme FISTA (Fast iterative shrinkage thresholding algorithm) [8]. L'inversion est définie selon :

$$\hat{X} = \underset{X}{argmin} \frac{1}{2} \|Y - AX\|_2^2 + \alpha^2 \|X\|_2 \tag{5.2}$$

avec $\hat{X}$ l'objet à reconstruire, X l'objet à chaque itération, Y les données (les signaux RF), et A la matrice de propagation, contenant la réponse impulsionnelle du système à chaque position de la grille de reconstruction. α est le paramètre de régularisation, choisi manuellement en comparant visuellement l'image reconstruite avec la *ground truth*. La matrice A est construite selon une approche similaire à celle utilisée pour les simulations décrites dans l'annexe B, premier paragraphe. On utilise par la suite le terme de "reconstructions alternatives" pour décrire les reconstructions par *delay-and-sum* et par déconvolution.

5.4 Correction des artefacts de visibilité par une approche *deep learning* : résultats

5.4.1 Entraînement du réseau et évaluation sur données de simulation

Les résultats obtenus sur le jeu de données de simulation sont présentés en figure 5.10. Les images reconstruites à l'aide des méthodes alternatives (b,c,f,g) sont clairement affectées par les artefacts de visibilité (flèches) : les vaisseaux verticaux sont absents et le centre des larges vaisseaux apparaît vide. L'épaisseur des vaisseaux est par ailleurs sur-estimée sur la reconstruction par *delay-and-sum*, et sous-estimée sur la reconstruction par déconvolution. A l'inverse, la reconstruction par DLA produit des images quasi-parfaites (d,h) pour lesquelles la totalité des artefacts de visibilité sont corrigés. Seuls quelques petits appendices, et vaisseaux sur le bord de l'image, ne semblent pas correctement reconstruits (étoiles). Les performances de la reconstruction par *delay-and-sum* et DLA, évaluées sur les 15 exemples du *test set* et présentées dans le tableau 5.1 confirment ces impressions qualitatives. Les NCC et SSIM calculés pour la reconstruction par DLA sont environ trois fois plus élevés que ceux calculés pour la reconstruction par *delay-and-sum*. Les scores pour la méthode de déconvolution sont du même ordre de grandeur que ces derniers. Ces résultats, très encourageants, montrent que l'approche DLA est très efficace en simulation.

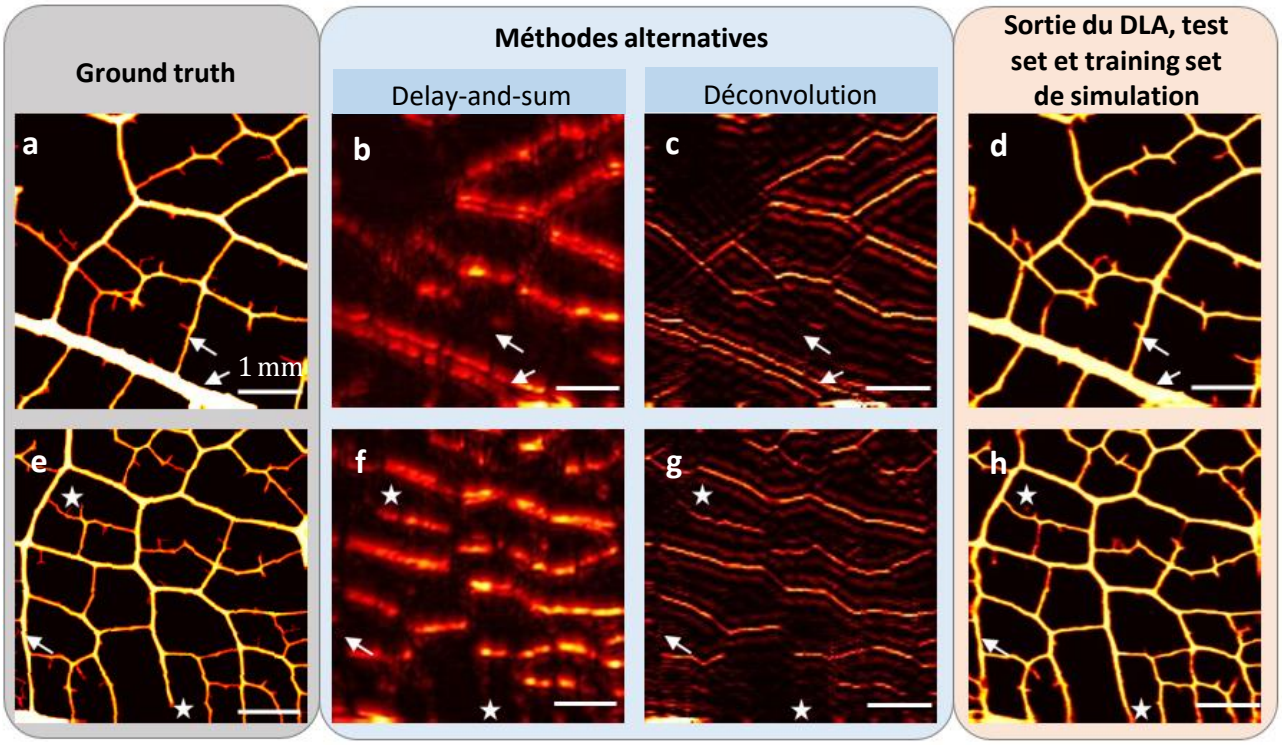

Figure 5.10 – Reconstruction de deux exemples du test set de simulation par delay-and-sum, par déconvolution et par l'approche deep learning (DLA). Le réseau est entrainé sur un training set de simulation.

5.4.2 Entraînement du réseau et évaluation sur données d'expérimentales

Un réseau est ensuite entraîné sur le *training set* expérimental et ses performances sont évaluées sur le *test set* expérimental. Les méthodes de reconstruction alternatives ne permettent toujours pas de corriger les artefacts de visibilité sur les données expérimentales (b,c,f,g). A l'inverse, la plupart des structures verticales sont visibles sur les images reconstruites par l'approche DLA (d,h), et l'épaisseur des vaisseaux est correctement estimée. Les scores obtenus, présentés dans le tableau 5.1, reflètent à nouveau la supériorité de l'approche DLA, pour laquelle le sSSIM et le NCC sont environ 2 fois plus élevés que ceux des reconstructions par *delay-and-sum*.

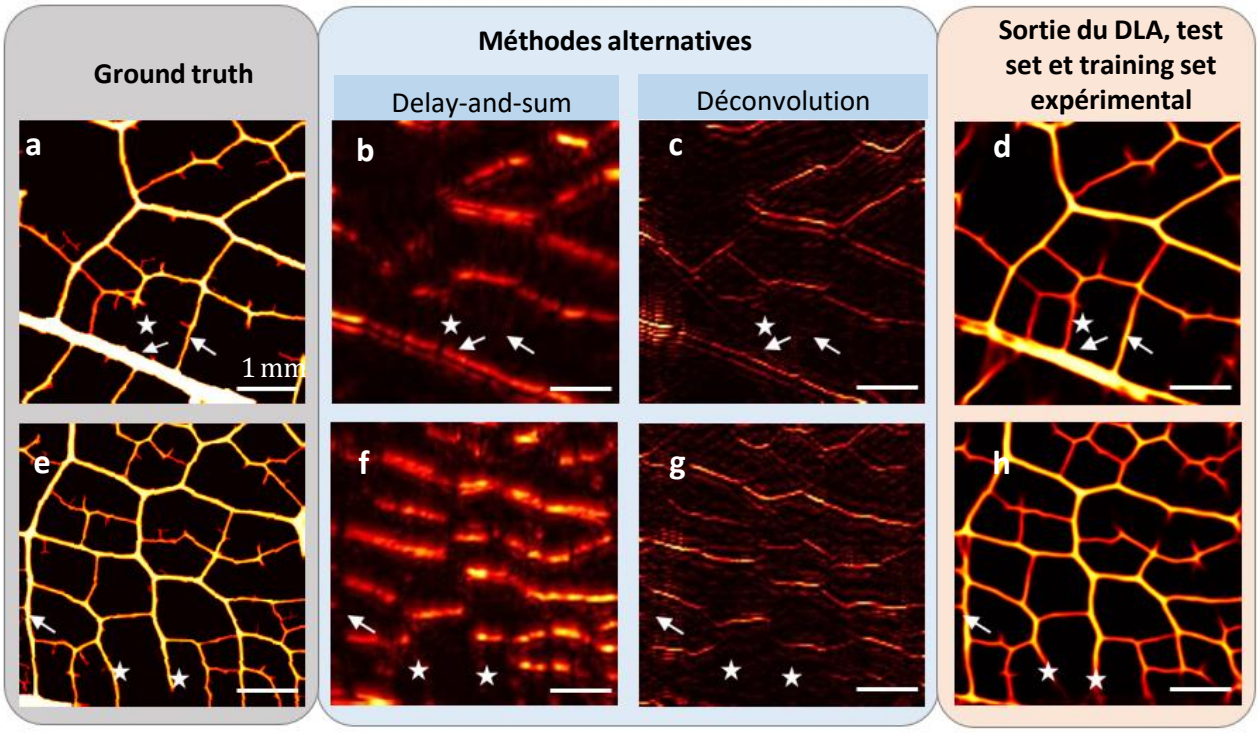

Figure 5.11 – Reconstruction de deux exemples du test set expérimental par delay-and-sum, par déconvolution et par l'approche deep learning (DLA). Le réseau est entrainé sur le training set expérimental.

Tableau 5.1 – Performances des méthodes de reconstruction par delay-and-sum et deep learning pour des données de simulation et expérimentales.

	Simulation		Expérience	
	delay-and-sum	*deep learning*	*delay-and-sum*	*deep learning*
NCC	$0.31 \pm .02$	$0.89 \pm .01$	$0.44 \pm .06$	$0.80 \pm .03$
sSSIM	$0.29 \pm .02$	$0.87 \pm .01$	$0.38 \pm .05$	$0.76 \pm .03$

Les erreurs restantes concernent majoritairement quelques vaisseaux verticaux qui restent non visibles. Dans quelques cas, une structure est mal reconstruite (avec un angle, une mauvaise direction ...) et peut même être inventée par le réseau. Ces situations sont repérées par des étoiles sur la figure 5.11. Si les premières erreurs (non-reconstruction de certains vaisseaux verticaux), communes à de nombreuses méthodes de reconstruction conventionnelles, sont acceptables, les secondes sont plus problématiques, car elles limitent notre confiance en la prédiction du réseau. C'est pourquoi une méthode permettant de mesurer l'incertitude de prédiction est présentée dans la partie suivante. Les images (d,h) sont néanmoins impressionnantes et d'une qualité largement supérieure à celle des images conventionnelles. Elles démontrent qu'une approche DLA peut être utilisée pour corriger les artefacts de visibilité sur des données expérimentales acquises avec une sonde d'imagerie conventionnelle.

5.4.3 Pré-entraînement du réseau sur des données de simulation

Dans un contexte clinique, constituer un jeu de données expérimental contenant suffisamment d'exemples est un processus qui peut s'avérer complexe, long et coûteux. Il serait alors tentant d'utiliser des données

de simulation, plus faciles et systématiques à obtenir, pour entraîner le réseau avant d'utiliser celui-ci sur des données réelles. Dans notre étude, les résultats obtenus en utilisant un réseau entraîné uniquement sur des données de simulation pour reconstruire une image à partir de données réelles ne sont pas satisfaisants. Comme le montre la figure 5.12, les images expérimentales reconstruites sont fortement aberrées, bien que certaines structures verticales apparaissent tout de même.

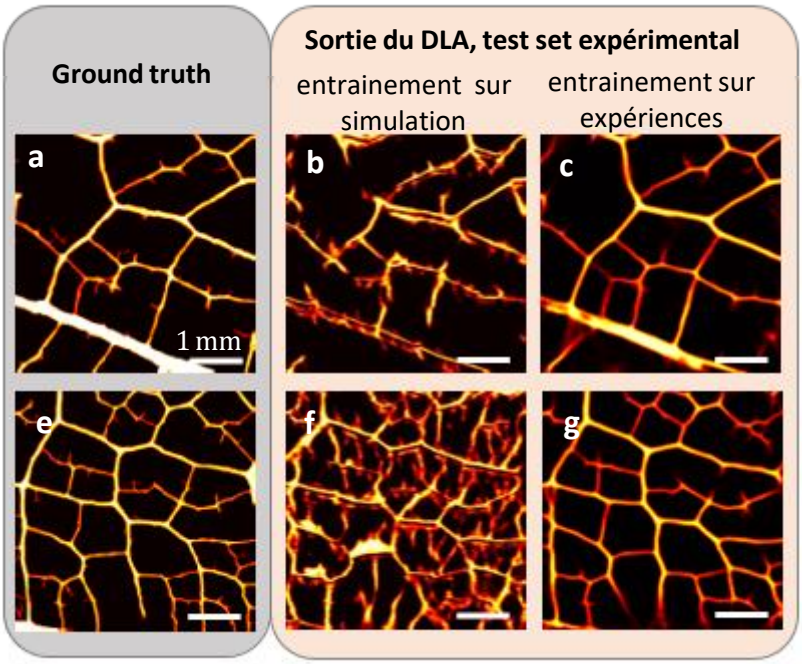

Figure 5.12 – Images expérimentales du test set reconstruites avec un réseau de neurones (DLA) entraîné avec des données de simulation uniquement ou avec des données expérimentales uniquement.

Un entraînement avec uniquement des données de simulation semble donc insuffisant, ce qui avait par ailleurs déjà été observé dans [24]. Dans le but de réduire la taille du *training set* expérimental, une approche intermédiaire consiste à pré-entraîner un réseau sur des données de simulation, puis de poursuivre l'entraînement sur des données expérimentales. Grâce à ce pré-entraînement, l'objectif est alors de réduire le taille du *training set* expérimental. Une telle étude est réalisée sur nos données. Des réseaux de neurones pré-entraînés sur notre jeu de simulation, ou non pré-entraînés, sont (ré)entraînés sur des jeux de données expérimentales constitués de 10, 20, 50, 100, 200, 300, 400 et 550 exemples. Pour chacune de ces tailles, 30 entraînements sont réalisés avec des jeux de données d'entraînement et de validation différents, et les résultats calculés sur le *test set* sont moyennés. Ce protocole permet de limiter un éventuel impact d'un *training set* peu représentatif des données, notamment lorsque celui-ci est de très petite taille. La figure 5.13 présente ces résultats.

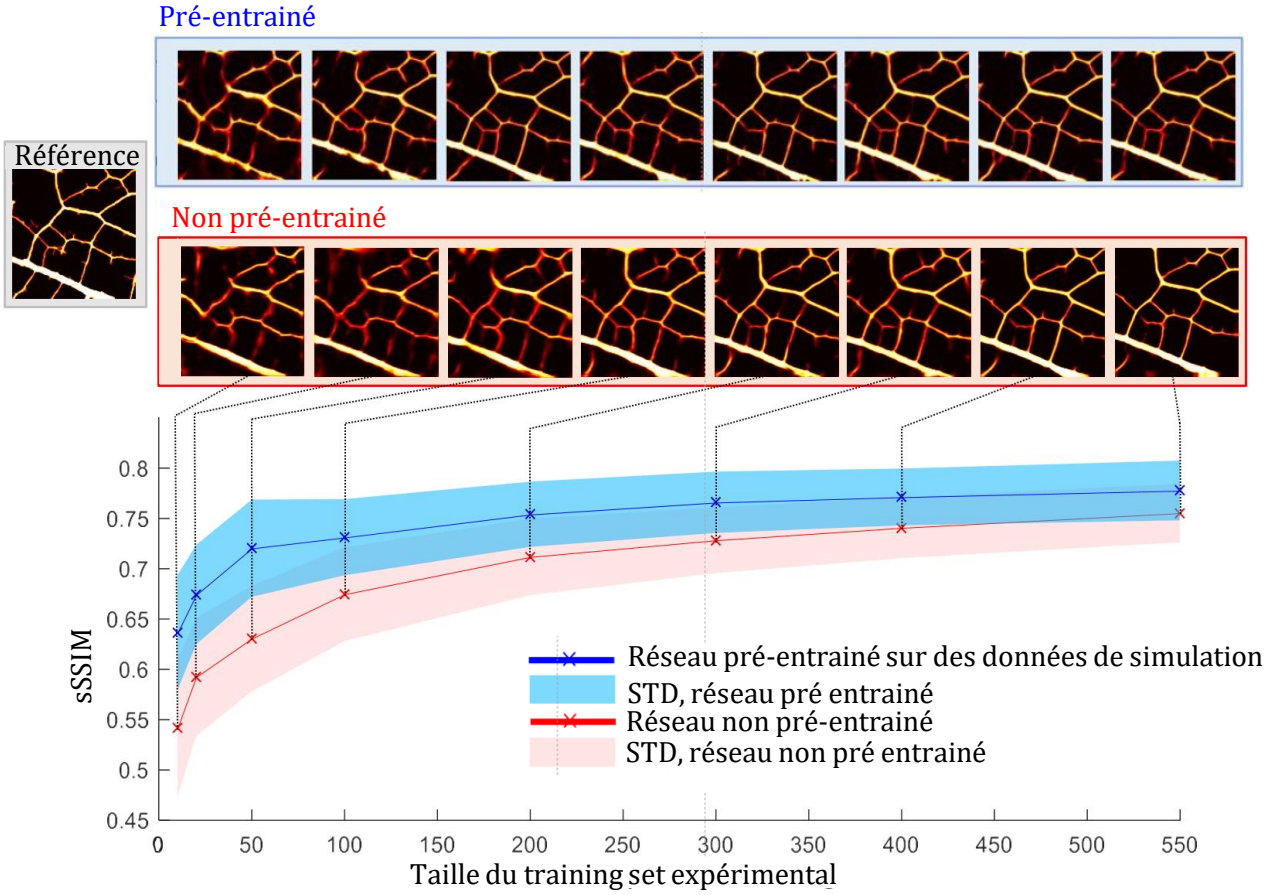

Figure 5.13 – Performance des réseaux de neurones sur le test set expérimental en fonction de la taille du training set expérimental. En bleu, ces performances pour des réseaux de neurones pré-entraînés et en rouge ces performances pour des réseaux de neurones non pré-entraînés. Pour chaque taille (chaque croix sur les courbes), 30 réseaux de neurones sont entraînés indépendamment. Ce sont les performances moyennées sur ces 30 réalisations et l'écart type associé qui sont représentées sur les courbes.

Comme attendu, la performance augmente avec la taille du *training set*. Lorsque le réseau n'a pas subi de pré-entraînement, la convergence, définie selon un critère visuel, est atteinte vers 200 paires. L'image est en effet qualitativement moins bien reconstruite pour un *training set* composé de moins de 200 paires, et est visuellement stable au-delà, bien que le sSSIM continue d'augmenter. Avec une session de pré-entraînement, la performance globale du réseau est meilleure, avec une valeur de sSSIM légèrement supérieure lorsque la totalité du *training set* est utilisée. Surtout, le réseau converge plus vite : dès 50 paires environ, l'image est de qualité comparable à celle obtenue avec la totalité du *training set* expérimental. Ainsi, sur notre jeu de données, un pré-entraînement permet de diminuer la taille du *training set* d'environ un facteur 4.

5.4.4 Discussion, perspectives et conclusions

Discussion

Une différence fondamentale entre la simulation et l'expérience concerne la nature de la *ground truth*. En simulation, la *ground truth* représente exactement la distribution de l'énergie absorbée. En expérience, la *ground truth* est une projection photographique 2D d'une distribution d'absorption optique tridimensionnelle. C'est donc l'intégration de l'absorption optique le long de l'épaisseur de la feuille qui est représentée. Cela explique vraisemblablement en partie la différence entre les performances en simulation et en conditions réelles, ainsi que la mauvaise généralisation d'un réseau entraîné sur simulations à

des données expérimentales. Les conditions de simulation pourraient également être améliorées pour se rapprocher des conditions expérimentales, en modélisant plus fidèlement la réponse impulsionnelle de la sonde ou encore en tenant compte du bruit électronique. Malgré ces approximations, l'incorporation de données de simulation par l'approche de *transfer learning* (figure 5.13) permet de réduire considérablement la taille du *training set* expérimental. Un telle procédure présente un grand potentiel notamment dans le contexte clinique, pour réduire par exemple le nombre de patients nécessaires à la formation du jeu de données tout en produisant un réseau de neurones performant.

Les erreurs présentes sur les résultats expérimentaux sont souvent localisées sur les bords des images reconstruites. En effet, dans ces zones, le réseau dispose d'une moins grande quantité d'informations sur les structures environnantes, limitant la qualité de la reconstruction. Une manière d'améliorer ces résultats serait de reconstruire à partir d'une zone plus large que la région d'intérêt, et de couper les bordures. La majorité des erreurs restantes se situent au niveau des coupes manuelles, le réseau tentant de joindre les vaisseaux lorsque des appendices sont présents. Il est probable que le nombre d'erreurs serait moindre sur un objet plus régulier. Globalement, un *training set* mieux construit avec une *ground truth* représentant réellement la distribution d'énergie permettrait probablement d'améliorer les performances.

Perspectives

L'objectif de cette étude était de démontrer que, expérimentalement, une approche DLA était envisageable pour corriger les artefacts de visibilité. Plusieurs problématiques devront néanmoins être prises en compte en cas d'application pratique.

Avant tout, notre photographie doit être remplacée par une *ground truth* mesurable dans un environnement réaliste, notamment en profondeur dans les tissus. Si une telle *ground truth* peut être obtenue avec une autre modalité d'imagerie telle que le scanner à rayon X ou l'imagerie par résonnance magnétique, il serait préférable d'obtenir une *ground truth* représentant l'absorption optique, dans le but de reconstruire une image fidèle à la morphologie de l'objet dont des informations quantitatives peuvent être extraites. Une telle *ground truth* peut être obtenue en utilisant un système d'acquisition photoacoustique plus sophistiqué [24], ou encore à l'aide d'une méthode de reconstruction quantitative permettant de corriger les artefacts de visibilité. Comme nous le verrons dans le chapitre suivant, une très bonne candidate est l'imagerie de fluctuation, décrite dans les chapitres 3 et 4.

L'influence du bruit sur les données devra également être étudiée. Un ratio signal à bruit de 60 a été mesuré pour nos données expérimentales. Si cette valeur est élevée, elle représente le ratio signal à bruit des structures horizontales, alors que nos travaux se focalisent sur la reconstruction des structures affectées par les problèmes de visibilité pour lesquelles le signal est presque inexistant. De plus, les images reconstruites par *delay-and-sum* sont affectées par le *clutter* qui se superpose au signal émis par les structures non visibles. Le ratio signal à bruit n'est donc pas forcément représentatif de la difficulté de la tâche. Dans le chapitre 6, nous avons étudié l'effet de la diminution du ratio signal à bruit en simulation et montré qu'il influençait peu la qualité de la prédiction.

La qualité de la prédiction est fortement influencée par la classe de l'objet à reconstruire, et l'homogénéité du jeu de donnée. Cette limitation est inhérente aux approches *deep learning*. En fonction de l'étendue et de la complexité de l'espace décrivant les données, l'architecture du réseau et ses hyperparamètres devront être adaptés.

Une dernière limitation, commune à de nombreuses approches DLA dans le domaine biomédical, est le manque de fiabilité des prédictions, lequel se traduit par la difficulté à appréhender les agissements du réseau, et donc d'interpréter les résultats et les éventuelles erreurs. Cette problématique se retrouve dans nos résultats : dans quelques situations, le réseau interprète incorrectement les données, et modi-

fie ou même invente des structures. Dans la partie suivante (5.5), des méthodes visant à estimer cette incertitude de la prédiction du réseau sont présentées et appliquées à nos données.

Conclusion

Notre algorithme entraîné et testé sur des données de simulation est capable de corriger parfaitement les artefacts de visibilité (figure 5.10). Entrainé sur des données expérimentales, le réseau est capable de produire des images expérimentales de très bonne qualité, en corrigeant la majorité des artefacts de visibilité (figure 5.11). Si des données expérimentales sont indispensables pour entraîner le réseau, la taille de ce jeu de données est fortement réduite en incorporant des données de simulation. Nous avons donc démontré qu'une approche DLA peut être utilisée pour corriger les artefacts de visibilité sur des données expérimentales acquises avec une sonde d'imagerie conventionnelle en utilisant un nombre restreint de données expérimentales d'entraînement. Cette étude est inédite de part son caractère expérimental, l'utilisation d'une sonde d'imagerie conventionnelle, et la démonstration de la correction de l'ensemble des artefacts de visibilité (*limited view* et *limited bandwidth*). Les résultats sont prometteurs et permettent d'envisager sereinement le passage à des données *in vivo*.

5.5 Quantification de l'incertitude en *deep learning*

Dans cette partie, nous décrivons les sources d'incertitude de prédiction en *deep learning* et présentons les méthodes développées pour en tenir compte. L'une d'elle est ensuite appliquée à notre jeu de données, et les résultats sont commentés.

5.5.1 Sources de l'incertitude

On distingue deux sources d'incertitude : l'incertitude épistémique appelée aussi incertitude du modèle, et l'incertitude intrinsèque au processus de génération des données aussi appelée incertitude statistique [5].

Incertitude épistémique

L'incertitude épistémique est causée par un manque de données d'entraînement dans certaines zones de l'espace décrivant nos données : si les réseaux de neurones sont excellents pour interpoler, ils ne le sont pas pour extrapoler. Ainsi, il est important que le *training set* soit suffisamment complet pour que les données utilisées en production soit bien décrites par le modèle : une réponse totalement arbitraire sera obtenue pour une donnée n'appartenant pas à cet espace, et une réponse peu précise sera obtenue pour des zones de l'espace peu décrite par le *training set*.

Incertitude aléatoire

L'incertitude aléatoire est produite par la stochastisticté naturelle des observations (les données d'entrée des réseaux) : des approximations sont réalisées lors de la génération des données, les protocoles et procédés utilisés pour leurs acquisitions ne sont pas parfaitement maîtrisés, et du bruit électronique peut être présent. Une même *ground truth* peut donc correspondre à plusieurs données d'entrée différentes, et inversement. Cette incertitude n'est pas explicable et ne peut pas être réduite, même si l'on dispose de plus de données. La figure 5.14 permet d'appréhender ces deux notions.

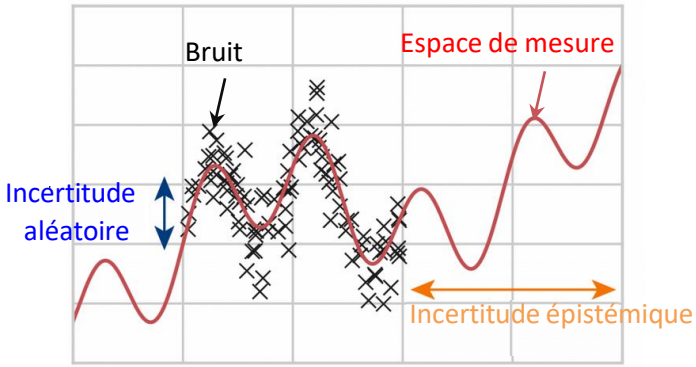

Figure 5.14 – Représentation des deux sources d'incertitude de prédiction. L'incertitude aléatoire correspond à une zone de l'espace de mesure pour laquelle des observations sont réalisées, mais ces observations sont affectées par du bruit. L'incertitude épistémique correspond à des zones de l'espace de mesure où aucune observation n'est réalisée.

5.5.2 Méthodes d'estimation existantes

Bien qu'il ne soit pas possible de garantir que les prédictions d'un réseau de neurones soient systématiquement exactes, il existe des méthodes permettant de leur attribuer un degré de fiabilité, inspirées des systèmes bayésiens.

Prédiction de l'incertitude épistémique

L'idée principale est de considérer une distribution pour chaque paramètre W du réseau de neurones, plutôt qu'une valeur fixe. Dans une première approche, appelée "ensemble de modèles", plusieurs modèles d'architecture identique sont entraînés indépendamment [75]. La phase d'entraînement d'un réseau de neurones n'est généralement pas déterministe de part par exemple l'initialisation aléatoire des paramètres W. Les paramètres W finaux de chacun des modèles entraînés indépendamment sont donc légèrement différents. En phase de test/production, l'inférence est réalisée sur tous les modèles en parallèle. Plusieurs prédictions sont donc générées pour une même donnée d'entrée. Des différences existent entre ces prédictions car chaque modèle est composé de paramètres de valeurs légèrement différentes. Chaque prédiction peut alors être vue comme un échantillon d'une loi de probabilité dont des paramètres, tels que la moyenne ou l'écart type, peuvent être estimés. Cet écart type peut alors être utilisé comme une mesure de l'incertitude.

L'utilisation de l'approche précédente nécessite des ressources computationnelles conséquentes car plusieurs modèles sont entraînés indépendamment. Une alternative appelée *Monte-Carlo Dropout* et basée sur l'utilisation du *Dropout*, a ainsi été proposée [37] pour obtenir une approximation bayésienne de "l'ensemble de modèles". Le *Dropout*, comme présenté dans la section 5.2.3 de ce chapitre, est une méthode de régularisation très populaire en *deep learning*, consistant à désactiver une partie des neurones de manière aléatoire à chaque itération pendant la phase d'entraînement pour améliorer les capacités de généralisation du réseau. Dans l'approche *Monte Carlo Dropout*, ce procédé est également appliqué pendant la phase de test/production. Cela permet d'y injecter du bruit : pour une même donnée d'entrée, plusieurs prédictions différentes peuvent être réalisées, le processus n'étant plus déterministe. Il n'est donc pas nécessaire d'entraîner un ensemble de modèles pour obtenir des prédictions différentes, comme pour l'approche précédente. Une série de prédiction est ainsi réalisée, et des estimateurs des paramètres de la loi de probabilité sont obtenus.

5.5.3 Prédiction de l'incertitude aléatoire

Les deux approches précédentes permettent de quantifier l'incertitude épistémique. L'incertitude aléatoire se subdivise en deux catégories : l'incertitude hétéroscedastique, et l'incertitude homéoscédastique. La seconde est constante pour toutes les données et dépend globalement de la tâche confiée au réseau, alors que la première dépend de chaque donnée et peut donc être prédite par le modèle. Ces données sont supposées appartenir à une distribution normale d'une certaine moyenne $\hat{y}$ et d'une certaine variance σ^2. On cherche à construire un modèle pouvant prédire ces paramètres. Pour cela, une nouvelle fonction de coût est utilisée pour incorporer l'estimateur de variance dans l'équation. Une approche de maximum de vraisemblance pour une distribution Gausienne [75] peut par exemple être utilisée [70] :

$$\frac{1}{D}\sum_{i=1}^{D} 0.5 \times \frac{|y_i - \hat{y}_i(x)|^2}{\hat{\sigma}_i(x)^2} + 0.5 \times \log \hat{\sigma}_i(x)^2 \tag{5.3}$$

avec x la donnée d'entrée, $\hat{y}_i$ l'estimation de la moyenne en chaque pixel i, σ_i^2 l'estimation de la variance en chaque pixel i, et D le nombre de pixel. Intuitivement, le terme $|y_i - \hat{y}_i(x)|^2$ encourage la prédiction $\hat{y}$ a être proche de la *ground truth*, alors que le dénominateur $\hat{\sigma}_i(x)^2$ s'assure que la variance augmente lorsque ces deux grandeurs s'éloignent. Le terme $\log \hat{\sigma}_i(x)^2$ permet d'éviter une croissance infinie de cette variance.

5.5.4 Application des méthodes d'estimation de l'incertitude à nos données : résultats et discussion

Approche Monte-Carlo Dropout

L'approche *Monte-carlo Dropout* a été appliquée à nos données. La figure 5.15 en présente le principe, décrit dans la section précédente.

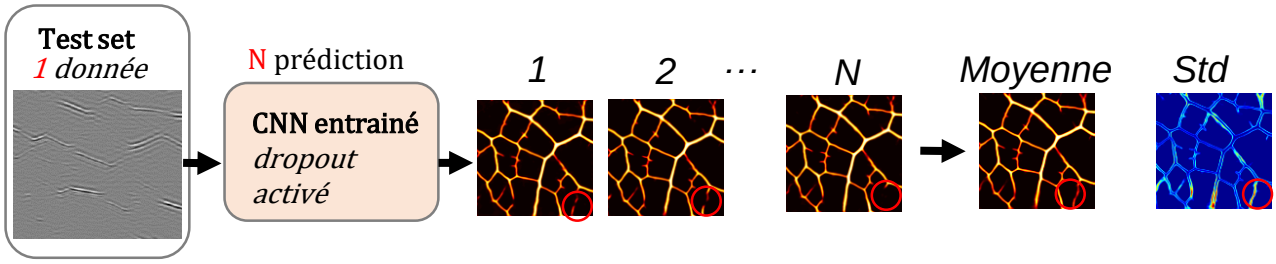

Figure 5.15 – Approche Monte-Carlo Dropout. Sur le test set, une série de prédictions est réalisée à partir d'une même donnée d'entrée en utilisant le Dropout, désactivant de manière aléatoire une partie des neurones. Chaque prédiction est donc légèrement différente. Une image moyenne et son écart type peuvent alors être estimés.

Le taux de *Dropout* est de 50% (la moitié des neurones sont donc désactivés à chaque prédiction), et 20 inférences sont réalisées pour chaque exemple du *test set*. La moyenne et l'écart type de ces 20 inférences sont calculés en chaque pixel. Un faible écart type indiquera que la prédiction est robuste pour la zone considérée : celle-ci reste stable pour l'ensemble des réalisations.

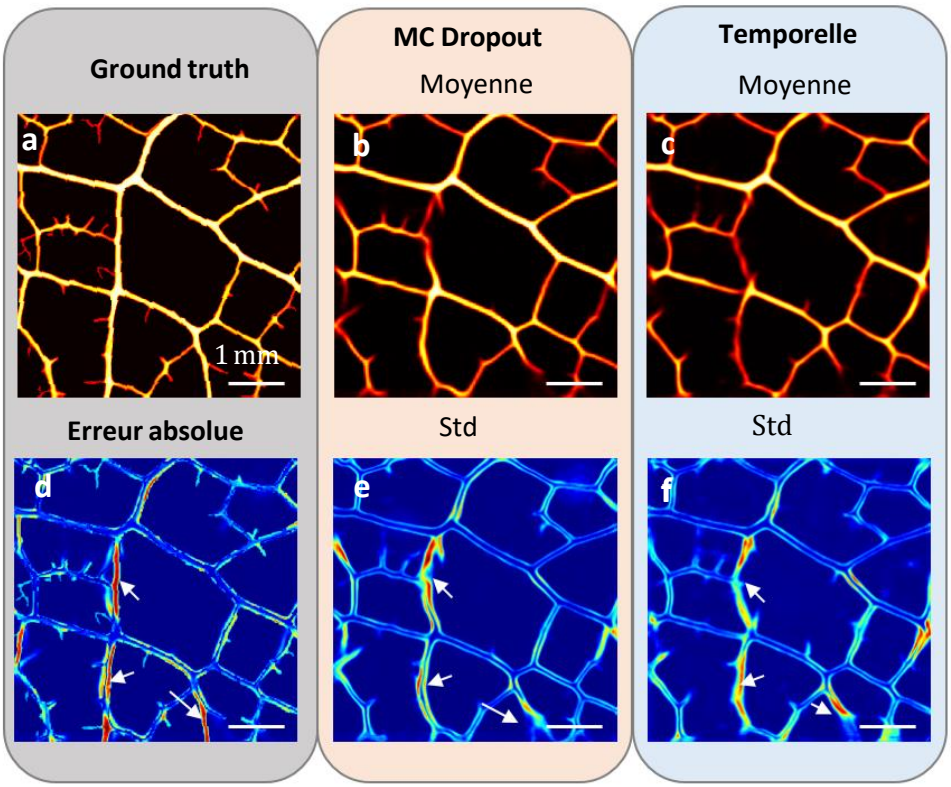

Figure 5.16 – (a) ground truth. (b,e) : moyenne d'une série de prédication réalisée à partir d'une unique acquisition (Monte-Carlo Dropout), et l'écart type correspondant. (c,f) : moyenne d'une série de prédiction réalisée à partir de différentes acquisitions de l'objet, et l'écart type correspondant. (d) : Erreur absolue.

Ayant à notre disposition plusieurs acquisitions photoacoustique de la zone imagée, une série de prédictions est réalisée avec le réseau de neurones, en désactivant le *Dropout* et en changeant la donnée d'entrée. La stochasticité provient dans ce cas directement des données et non plus du processus de prédiction, ici déterministe. La moyenne de 20 prédictions (fig. 5.16,c) et son écart type (fig. 5.16,f) sont alors calculées. La carte d'écart type dite "temporelle" est très proche de celle obtenue avec la méthode *Monte-Carlo Dropout*. Il semble donc que l'approche *Monte-Carlo Dropout* permet, à partir d'une unique acquisition, d'obtenir une information similaire à celle obtenue par une légère variation de la donnée d'entrée. Cette caractéristique est très intéressante dans le cadre de la navigation en temps réel, ou de mouvement de tissu.

Estimation de l'incertitude aléatoire

L'estimation de l'incertitude aléatoire consiste à générer en sortie du réseau de neurones une image moyenne et une image d'incertitude pour chaque donnée d'entrée. On change donc la fonction de coût pour modéliser l'incertitude aléatoire et obtenir deux données en sortie du réseau. L'équation 5.3 est légèrement modifiée pour que le modèle soit entraîné à prédire le logarithme de la variance $s_i(x) = log(\sigma_i(x)^2)$, rendant le processus plus stable et limitant le risque de division par 0.

$$\frac{1}{D}\sum_{i=1} 0.5 \times exp(-s_i(x))|y_i - \hat{y}_i(x)|^2 + 0.5 \times s_i(x) \tag{5.4}$$

Les images moyennes et d'écart type obtenues sur un exemple du *test set* expérimental sont présentées en figure 5.17. Un masque de l'image moyenne est préalablement calculé et appliqué pour supprimer la variabilité des pixels n'appartenant pas à l'objet.

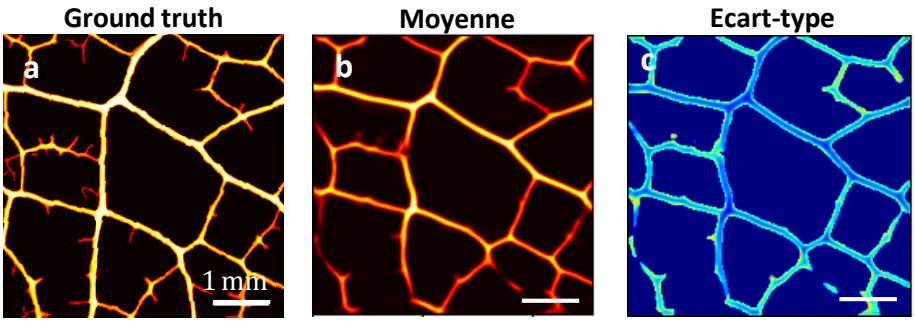

Figure 5.17 – Prédiction de l'incertitude aléatoire. (a) : ground truth. (b) : image moyenne prédite. (c) : image d'écart type prédite.

L'incertitude aléatoire est plus importante pour les vaisseaux dont l'amplitude est mal estimée, ce qui peut s'expliquer par la différence de nature entre la *ground truth* et l'image photoacoustique et l'absence d'une correspondance quantitative entre ces deux images, complexifiant la tâche du modèle. Pour notre étude, l'intérêt de cette information est limité, puisque l'objectif ici est de détecter des vaisseaux mal reconstruits, lesquels ne semblent pas spécifiquement capturés dans la carte d'écart-type.

Il est possible de combiner l'approche décrite dans cette section avec l'approche *Monte-Carlo Dropout*, en activant les *Dropout* en phase de test/production et en utilisant la fonction de coût dont l'expression est donnée par l'équation 5.4. Nous avons obtenus les mêmes cartes d'incertitude que celles présentées en figure 5.17 et 5.16.

5.5.5 Conclusion

Pour la première fois en imagerie photoacoustique, l'estimation de l'incertitude de prédiction d'un réseau de neurones a été réalisée avec succès. L'utilisation notamment de l'approche *Monte-Carlo Dropout* permet de clairement localiser les erreurs de prédiction correspondant à des structures inventées ou mal reconstruites par le réseau. De manière notable, l'incertitude estimée reste faible dans les zones bien reconstruites. Cette carte d'incertitude, permet donc d'estimer la fiabilité de la prédiction du réseau. L'intérêt est considérable dans le cadre de la reconstruction d'images médicales, dont l'observation est à l'origine de décisions aux impacts lourds de conséquences pour les patients. Cette carte d'incertitude est par ailleurs obtenue à partir d'une unique image en entrée, un atout dans le cadre de la navigation en temps réel.

5.6 Études complémentaires

5.6.1 Influence du type de données d'entrée sur la qualité de la prédiction

Pour les résultats présentés à ce stade, l'image en entrée du réseau est l'image mBF, c'est à dire l'image résultante de l'application de l'algorithme de *delay-and-sum* aux signaux temporels bruts. En effet, nous avons constaté de meilleurs performances pour ce type de données, comparativement à l'image dmBF (l'image conventionnelle). Des réseaux de neurones ont été entraînés avec les deux types de données, et les prédictions correspondantes sont présentées en figure 5.18. Le sSSIM moyen sur le *test set* est de 0.76 pour une entrée de type mBF, contre 0.72 pour une entrée de type dmBF. Plus d'artefacts sont en effet visibles lorsque l'image dmBF est utilisée. Ces artefacts sont indiqués par des flèches sur la figure 5.18. L'épaisseur des vaisseaux est également plus fréquemment mal estimée. L'image dmBF, bien que plus éloignée de la vraie structure de l'objet, semble donc contenir plus d'information (amplitude et phase), que le réseau sera capable d'utiliser.

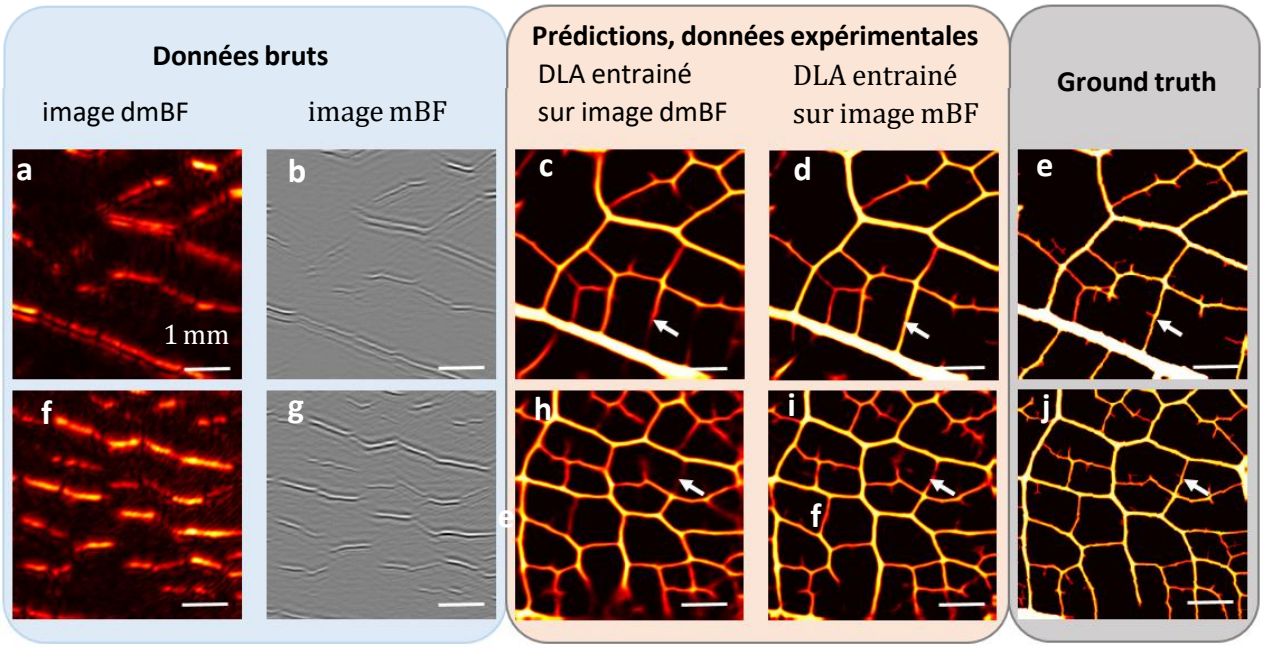

Figure 5.18 – Prédictions du réseau de neurones en fonction du type de données d'entrée. Deux exemples. (a,f) : images dmBF, c'est à dire des images reconstruites par delay-and-sum à partir des signaux radiofréquences sous leur forme en quadrature. (b,g) : images mBF, c'est à dire des images reconstruites par delay-and-sum à partir des signaux radiofréquences bruts. (c,h) : prédictions d'un réseau entraîné avec des images dmBF. (d,k) : prédictions d'un réseau entraîné avec des images mBF. (e,j) : ground truth.

5.6.2 Propriétés de généralisation de notre réseau de neurones

La capacité du réseau à généraliser à de nouvelles données est fondamentale. Pour s'assurer que le réseau ne sur-apprend pas le jeu de données d'entraînement, le résultat de la fonction de coût et du calcul de la métrique (le NCC) peuvent être calculés pour le jeu de données de validation et le jeu de données d'entraînement à chaque itération. Leurs évolutions en fonction de l'itération sont visibles en figure 5.19. Une forte divergence entre les deux courbes d'erreur est visible à partir de 38 itérations : l'erreur sur le jeu de données de validation se stabilise tandis que l'erreur sur le jeu de données d'entraînement continue de diminuer. Cette divergence apparaît dès 20 itérations sur les courbes de NCC, avec un maximum à environ 38 itérations pour le jeu de données de validation.

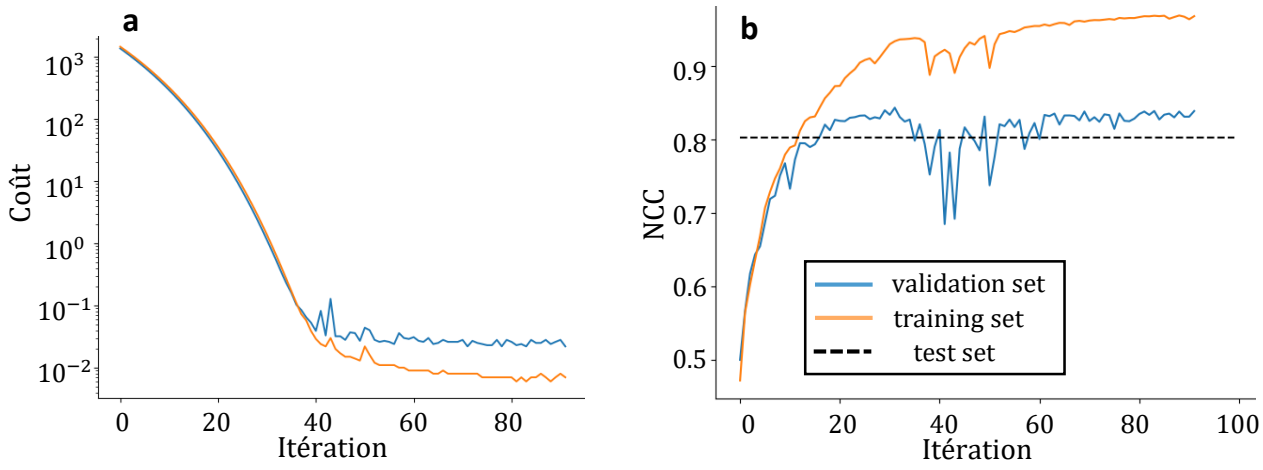

Figure 5.19 – Évolution de la fonction de coût et du résultat de la cross-corrélation normalisée (NCC) en fonction de l'itération pour le jeu de données de validation et le jeu de données d'entraînement.

Le modèle utilisé pour réaliser les prédictions est donc celui obtenu à l'itération 38, pour laquelle un NCC de 0.92 est atteint pour le jeu de données d'entraînement, et d'environ 0.82 pour le jeu de données de validation. Il semble donc que le réseau soit sujet à un léger sur-apprentissage. Malgré ce sur-apprentissage, notre réseau est capable de généraliser à nos données de test.

On peut étendre la définition du terme "généralisation" à la prédiction de données s'éloignant de l'espace décrit par notre jeu de données d'entraînement (objet différent, conditions expérimentales différentes etc...). Bien que nos travaux aient pour principal but de démontrer la capacité d'un DLA à corriger les artefacts de visibilité sur une classe d'objet spécifique, on peut s'intéresser à cette capacité de généralisation de notre réseau. Pour illustrer la problématique, un réseau entraîné sur notre jeu de données de simulation est utilisé pour reconstruire l'image photoacoustique d'une structure mimant des vaisseaux (provenant du package k-wave , http ://www.k-wave.org/). Bien que l'objet soit clairement différent de la feuille, la reconstruction est de très bonne qualité comme le montre la figure 5.20, suggérant une certaine capacité à généraliser à de nouvelles données. Il est évident qu'une étude plus approfondie est nécessaire pour définir les limites du réseau, en particulier pour des données expérimentales.

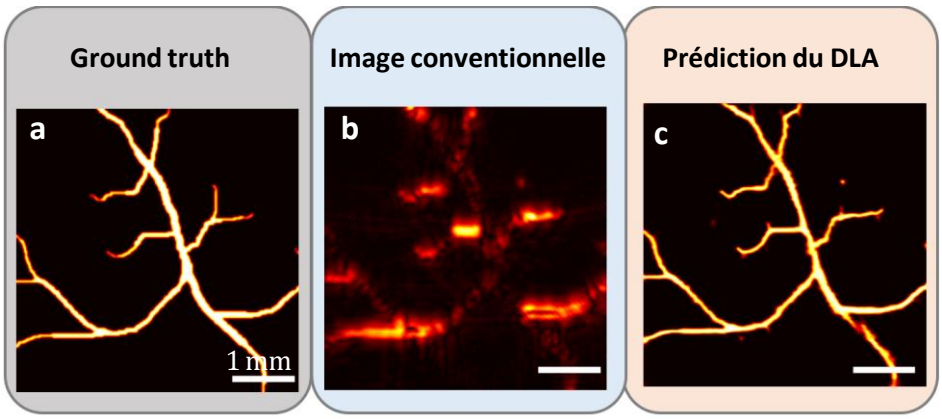

Figure 5.20 – Prédiction d'une structure mimant un vaisseau sanguin par un réseau entraîné sur le jeu de données de simulation constitué des squelettes de feuille.

5.6.3 Influence de la taille du réseau

Le réseau utilisé dans notre étude est composé de 30 millions de paramètres. Ce nombre dépend principalement du nombre de neurones et de couches, lesquels, dans notre étude, furent déterminés automatiquement pour maximiser les métriques sur le *validation set*. Néanmoins, diminuer ce nombre permet de diminuer le temps et les ressources nécessaires à l'entraînement et à la prédiction, et pourrait potentiellement augmenter les capacités de généralisation du réseau. Des réseaux de neurones contenant respectivement 30 millions, 7 millions, 1 900,000 et 460 000 paramètres sont entraînés sur notre jeu de données. Des images prédites par les réseaux ainsi que les NCC calculés sur l'ensemble du *test set* sont présentés dans la figure 5.21.

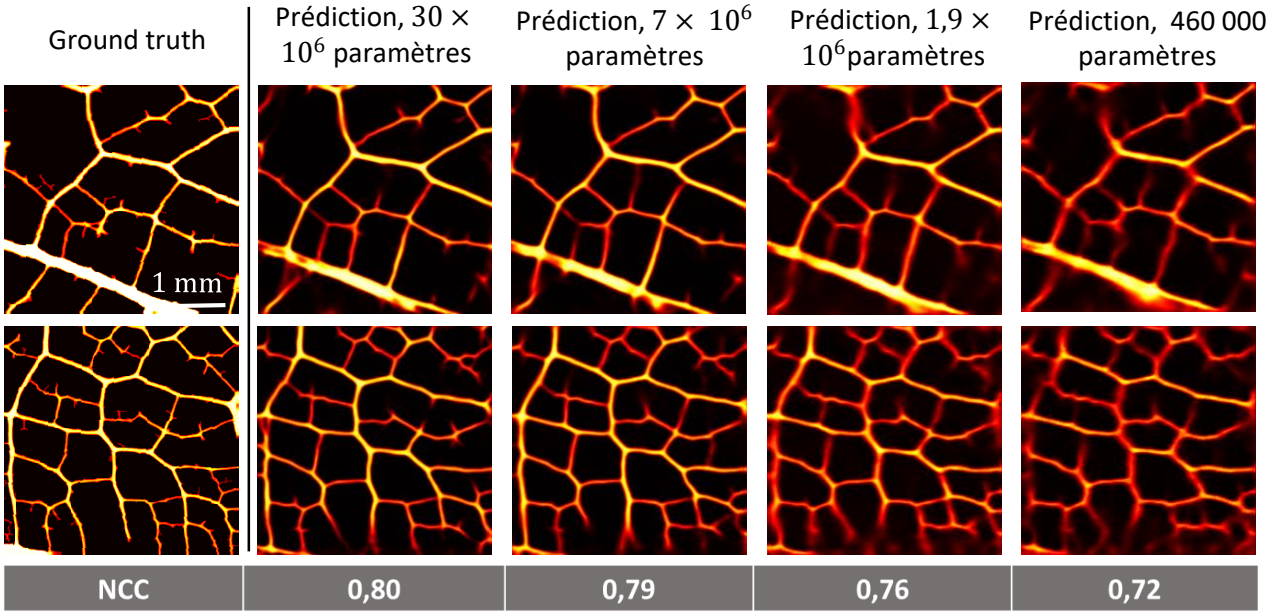

Figure 5.21 – Prédictions de réseaux de tailles différentes d'une image du test set expérimental et résultat de la cross-correlation normalisée (NCC) sur l'ensemble du test set.

L'image obtenue avec un réseau de 7 millions de neurones est similaire à celle obtenue avec 30 millions, indiquant qu'il aurait été possible de diminuer d'au moins un facteur 5 la taille du réseau. A partir de 1 900 000, l'image est visuellement dégradée et le NCC diminue.

5.6.4 Influence de l'architecture du réseau de neurones

De nombreuses architectures ont été développées à partir du modèle U-Net. Dans un réseau de type U-Net, chaque bloc est constitué de deux convolutions, avec éventuellement des étapes de *batch normalisation* (BN) avant l'opération de max-pooling/sur échantillonnage et le passage au bloc suivant. Cet enchaînement est représenté en figure 5.22.a.

Le ResUNet [154] est un modèle très utilisé, dérivé du U-Net. Des blocs résiduels remplacent ceux décrits précédemment : l'entrée et la sortie du bloc sont connectées, comme le montre la figure 5.22.b. Cette connexion facilite l'entraînement et limite les risques d'explosion et de disparition de gradient.

Le Dense U-Net [48] se caractérise par des connections denses entre les neurones. Dans chaque bloc, une série de convolutions et de concaténations sont réalisées, dont l'enchaînement est défini par le taux d'agrandissement. Dans la figure 5.22.c représentant un bloc de type Dense U-Net, un taux d'agrandissement de 4 est utilisé, pour un bloc comptant 32 cartes de caractéristiques en entrée. Ces connections permettent l'amélioration du transfert d'information entre les neurones, et la diminution du risque d'apprentissage de caractéristiques redondantes.

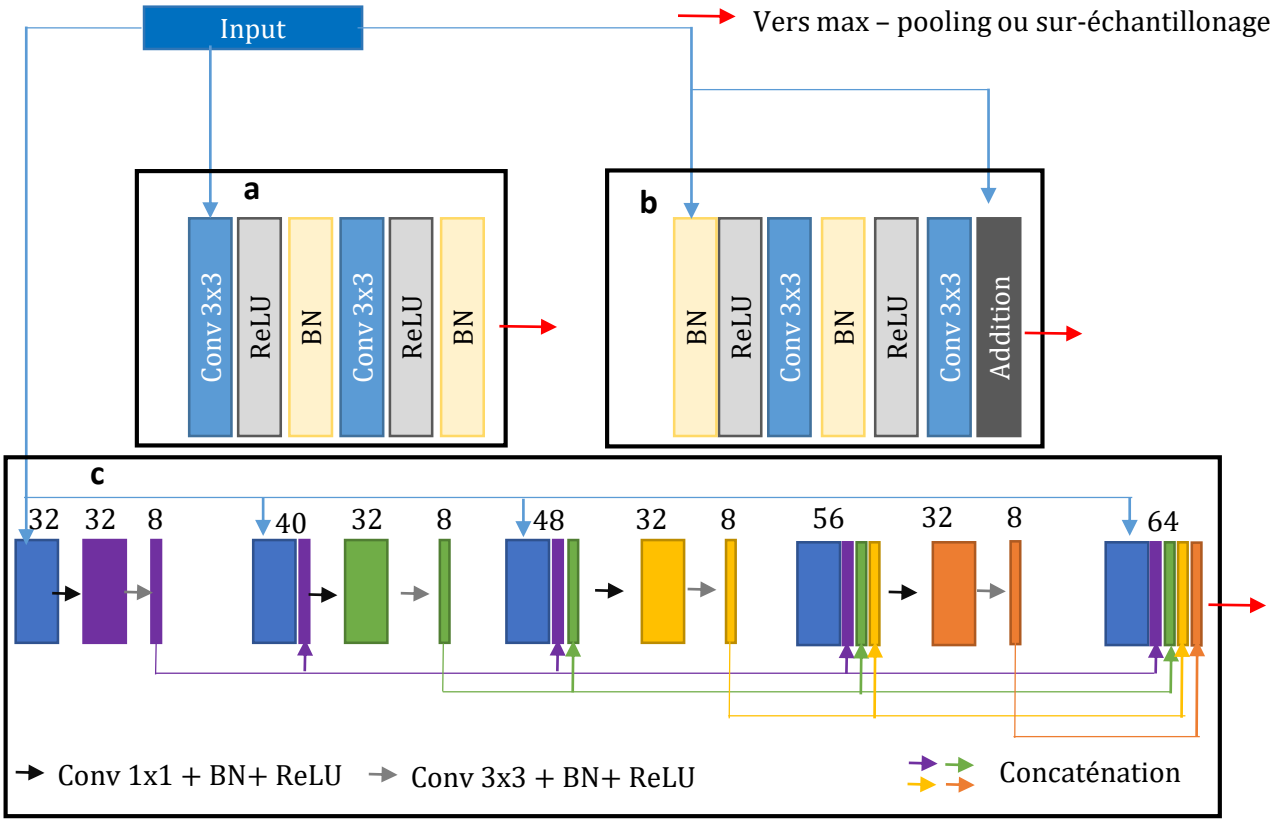

Figure 5.22 – Les différents types de blocs utilisés pour construire les réseaux de neurone. BN pour batch normalization, conv pour convolution. (a) : bloc U-Net. (b) : bloc Resunet, la donnée d'entrée du bloc est ajoutée à la sortie du bloc. (c) : bloc FD U-Net pour un taux d'agrandissement de 4 et une donnée d'entrée à 32 canaux. Les chiffres indiqués représentent le nombre de canaux après chaque opération.

Ces architectures ont été implémentées et testées sur nos données expérimentales. Une légère amélioration des métriques calculées sur le *test set* expérimental fut observée lorsque le Dense U-Net fut utilisé (+1.5%), et de (+0.5%) dans le cas du Resunet. L'impact visuel étant négligeable, nous avons conservé l'architecture U-Net.

5.6.5 Data augmentation : utilisation de GAN

Des transformations géométriques ont été appliquées au jeu de données de simulation pour augmenter sa taille. L'impact de ce type de transformation reste néanmoins limité, car l'image résultante reste relativement similaire à l'image initiale. Des méthodes plus efficaces existent, notamment basées sur l'utilisation de réseaux de neurones adverses (GAN) [44]. Cette approche repose sur la mise en compétition de deux réseaux. Un réseau générateur est entraîné à produire un nouvel exemple à partir d'une réalisation de bruit, et le réseau discriminateur détermine si cet exemple est réel ou non. Il se base pour cela sur un jeu de données d'entraînement composé d'images appartenant à l'espace décrivant la classe d'objet à générer. L'objectif du réseau générateur est de tromper le réseau discriminateur et que celui-ci classifie le nouvel exemple généré comme appartenant à l'espace décrit par les données d'entraînement. A l'inverse, l'objectif du réseau discriminateur est de détecter les faux exemples.

L'approche utilisée dans nos travaux repose sur un GAN à croissance progressive (*Progressive Growing* GAN) [69]. Un réseau est entraîné à produire une nouvelle image en plusieurs étapes, chacune correspondant à un niveau de résolution. L'entraînement d'un GAN est souvent difficile, et il arrive parfois que le réseau ne génère que des exemples identiques à ceux du *training set*. Nous allons nous assurer que les images prédites par le GAN sont suffisamment différentes du jeu de données d'entraînement. Pour une image générée par le GAN, on calcule sa ressemblance au travers du NCC avec l'ensemble des images constituant le *training set*, et on conserve l'image pour laquelle le NCC est maximale (c'est à dire l'image du *training set* la plus ressemblante à l'image générée selon le NCC). Pour l'ensemble des données générées par le GAN (4000 images), on obtient un NCC maximal de 0.59 et une moyenne de 0.22 : il semble donc que notre GAN soit capable de générer de nouveaux exemples différents de ceux composant le *training set*. La figure 5.23 présente deux exemples générés par le GAN (a), et l'image du *training set* original (b) qui lui ressemble le plus selon le NCC. Si on remarque visuellement des similitudes, la différence est très nette entre les images de la colonne de droite et celle de la colonne de gauche.

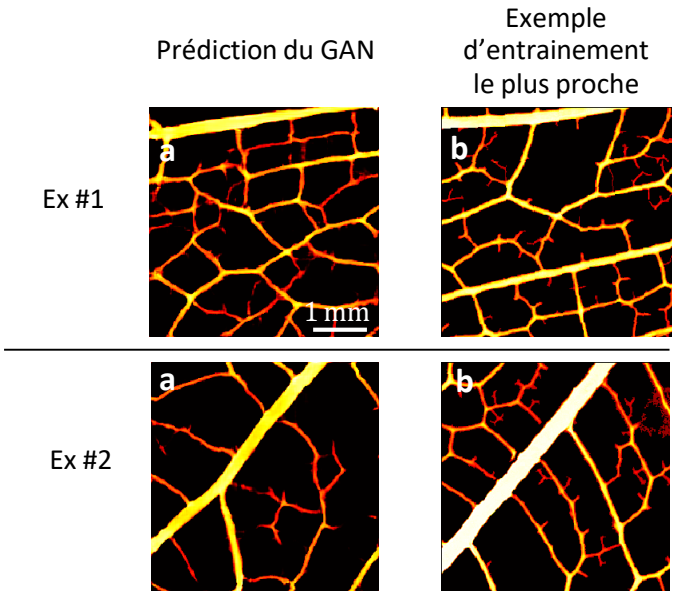

Figure 5.23 – Prédictions du GAN à gauche (a), et exemple appartenant au training set (b) le plus ressemblant selon le NCC à la prédiction en (a).

L'utilisation d'un *training set* augmenté par les images produites par le GAN, plutôt que celui augmenté par les transformations géométriques classiques, a permis d'obtenir un gain de 3% du NCC sur le *test set* de simulation.

5.7 Conclusion

La possibilité de corriger les artefacts de visibilités de type *limited view* et *limited bandwidth* a été démontrée sur des données de simulation et expérimentales. Les objets imagés sont des squelettes de feuilles, mimant un réseau vasculaire complexe et multi-échelles. La qualité des images prédites est largement supérieure à celle des images reconstruites par des méthodes conventionnelles : la majorité des vaisseaux verticaux et des larges vaisseaux y sont visibles. Ces résultats sont inédits de part l'utilisation d'une sonde d'imagerie conventionnelle, la démonstration expérimentale, et la correction des deux artefacts de visibilité. Nous avons également montré qu'un pré-entrainement du réseau de neurones sur des données de simulation permet de diminuer la taille du *training set* expérimental, une caractéristique intéressante lorsque l'on considère la difficulté à constituer de larges jeux de données en imagerie médicale. Certaines erreurs sont néanmoins présentes sur nos prédictions. Notamment, le réseau reconstruit incorrectement certaines structures. Ces erreurs sont capturées par une approche *Monte-Carlo Dropout* qui fournit une carte d'incertitude de la prédiction du réseau. Ces erreurs y sont clairement identifiables, et l'incertitude estimée est basse lorsque la prédiction est correcte. Un tel indicateur renseignant sur le degré de fiabilité de la reconstruction en chaque position de l'image est d'un intérêt considérable pour une utilisation clinique.

Pour être transposée à l'imagerie *in vivo*, notre approche doit être adaptée. Notamment il est nécessaire de pouvoir obtenir une *ground truth* dans un milieu turbide. Il serait préférable que cette *ground truth* soit proportionnelle à l'absorption optique : l'image prédite le sera alors également et pourra être utilisée à des fins d'analyses quantitatives. Nous allons montrer dans le chapitre 6 que l'image de fluctuation décrite dans les chapitres 4 et 5 est une très bonne candidate pour constituer cette *ground truth*.

6

Imagerie photoacoustique quantitative 3D assistée par apprentissage profond

Sommaire

6.1 Contexte **160**
6.2 Etudes sur des données de simulations **162**
6.2.1 Construction des jeux de données de simulation et présentation de l'architecture du réseau de neurones utilisée 162
6.2.2 Performances du réseau de neurones 166
6.2.3 Présence de bruit 168
6.3 Expériences **170**
6.3.1 Méthode 170
6.3.2 Résultats 171
6.4 Conclusion **173**

6.1 Contexte

Dans ce chapitre, nous montrons que l'image de fluctuation est une candidate idéale pour constituer la référence (*ground truth*) d'un jeu de données d'entraînement d'un réseau de neurones visant à corriger les artefacts de visibilité et de sous-échantillonnage spatial. Une méthode originale, validée sur simulation, permet de prédire simultanément une série d'images à plusieurs longueurs d'onde proportionnelles à l'énergie absorbée, sans artefacts. Ces images peuvent ensuite être utilisées pour estimer la SO_2 après une éventuelle correction de la fluence.

Intérêt de l'image de fluctuation comme *ground truth*

Alors qu'il est nécessaire que le milieu soit transparent pour utiliser la photographie du chapitre précédent comme *ground truth*, l'image de fluctuation peut être obtenue dans un milieu turbide si le FNR (le niveau de la fluctuation vis à vis du bruit) est suffisant (chapitre 3 section 3.5). Tout comme la photographie, l'image ne présente ni l'artefact d'ouverture numérique limitée, ni celui de bande passante fréquentielle limitée dans le cas de l'imagerie des vaisseaux sanguins en présence d'un flux de globules rouges, et peut être obtenue avec une sonde d'imagerie conventionnelle . Les artefacts liés au sous-échantillonnage spatial (clutter) sont également réduits. Contrairement à la photographie qui est nécessairement une version projetée de l'objet absorbant, l'image de fluctuation en fournit une représentation tridimensionnelle, qui plus est quantitative (voir chapitre 4).

A ce jour, aucune solution ne permettant d'obtenir une telle *ground truth* n'a été proposée dans la littérature. Comme présenté dans le chapitre précédent, les réseaux de neurones sont principalement entraînés à partir de jeux de données de simulation, et parfois testés sur des données expérimentales mais sans *ground truth*, ne permettant pas de conclure sur les performances du réseau. Quelques études se distinguent néanmoins par la présence d'une *ground truth*. On peut citer les études [122] et [84] dont l'objectif est cependant de produire, à partir d'une image acquise avec une faible fluence, une image à plus haute fluence pour améliorer le SNR/contraste. La *ground truth* est alors obtenue par moyennage ou en modulant la fluence du laser, et reste donc affectée par les artefacts de visibilité et le *clutter*. Dans les études [24] et [56], une *ground truth* expérimentale est également disponible. Ces deux études se focalisent sur l'amélioration d'une image reconstruite à partir de signaux enregistrés par une sonde spatialement sous-échantillonnée. La correction des artefacts de visibilité est mentionnée, mais peu évidente sur les résultats présentés. La *ground truth* correspond dans ces études à l'image reconstruite à partir des signaux acquis avec le même système d'imagerie en utilisant l'ensemble des capteurs disponibles. Ce sont des systèmes non conventionnels, de type Fabry-Pérot [56] et *ring array* [24]. Seul [56] propose de reconstruire une image 3D expérimentale. Cette étude se démarque néanmoins de la nôtre sur plusieurs aspects, notamment par l'utilisation du capteur Fabry-Perot, lequel produit des images non affectées par l'artefact de sensibilité fréquentielle limitée. Les résultats présentés ne permettent pas de conclure sur la correction des artefacts d'ouverture numérique limitée, une unique projection dans l'axe z est en effet disponible alors que cet artefact s'observe principalement sur les projections x et y. L'utilisation du réseau de neurones est par ailleurs très différente : il y est utilisé comme une manière de régulariser une fonction de coût au sein d'une méthode conventionnelle d'optimisation. A l'inverse de notre approche, il est nécessaire de disposer de la matrice de propagation dont la connaissance est parfois impossible et souvent approximative. Par ailleurs, le temps de calcul nécessaire à la génération de l'image est généralement important lorsqu'une méthode d'optimisation est utilisée pour reconstruire l'image.

Enfin, comme présenté dans le chapitre 4, l'imagerie de fluctuation est particulièrement adaptée à l'estimation de la SO_2 dans le cadre de l'imagerie à plusieurs longueurs d'onde. Après l'application des corrections décrites dans le chapitre 4, l'image de fluctuation est proportionnelle à l'énergie absorbée, contrairement à la photographie du chapitre 5. De plus, sa sensibilité aux absorbeurs en mouvement et donc à l'hémoglobine présente dans les globules rouges lui confère la capacité de discriminer le signal

provenant des autres absorbeurs statiques.

Prédiction simultanée d'images à plusieurs longueurs d'onde proportionnelles à l'énergie absorbée

Dans la littérature, lorsqu'une *ground truth* expérimentale est présente, celle-ci est généralement obtenue à partir d'un dispositif d'imagerie PA [24, 56, 84, 122] plus performant, et correspond à une image proportionnelle à l'énergie absorbée au même titre que l'image de fluctuation. Cependant, dans toutes ces études, les performances des réseaux de neurones sont uniquement évaluées selon leur capacité à reconstruire correctement la morphologie des structures imagées : le caractère quantitatif et la proportionnalité de l'image prédite à l'énergie absorbée ne sont pas étudiés.

Notre objectif final est d'obtenir en sortie du réseau de neurones des images sans artefacts de visibilité proportionnelles à l'énergie absorbée à plusieurs longueurs d'onde, qui pourront ensuite être utilisées pour estimer par exemple une carte de SO_2 après une éventuelle correction de la fluence. Plutôt que de prédire les images à chaque longueur d'onde indépendamment, notre stratégie consiste à prédire en une unique inférence une série d'images de fluctuation à plusieurs longueurs d'onde à l'aide du réseau de neurones. Comme décrit dans la partie 6.2, le réseau est entraîné à prédire simultanément une série d'images de fluctuation à différentes longueurs d'onde à partir d'une série d'images conventionnelles à ces mêmes longueurs d'onde.

Si des travaux spécifiques à l'utilisation de réseaux de neurones en imagerie PA quantitative existent, ils se concentrent sur l'estimation directe d'une carte de SO_2 tout en corrigeant des effets de coloration spectrale. Le réseau de neurones apprend donc dans ce cas à réaliser à la fois la correction de fluence et l'extraction de la SO_2, et éventuellement à améliorer la visibilité des structures [9, 57, 82, 148]. Dans ces études, les réseaux sont entraînés avec l'image de SO_2 comme *ground truth*. Constituer un large jeu de données expérimentales avec une telle approche est difficilement envisageable, l'obtention dans des milieux biologiques d'une carte de SO_2 étant en effet complexe. Dans ces études, les jeux de données d'entraînement sont ainsi exclusivement générés à partir de simulations. Les réseaux sont également testés sur des données de simulation. Seule [46] propose un test sur une image expérimentale avec un réseau entraîné sur des données de simulation, mais sans *ground truth* disponible pour valider la performance du réseau.

Travaux réalisés dans le cadre de la thèse

Dans une première partie, nous montrerons en simulation que le réseau de neurones est capable de corriger les artefacts de visibilité en utilisant une série d'images de fluctuation en *ground truth* pendant la phase d'entraînement, et donc de retrouver la morphologie des structures imagées. Nous montrerons ensuite que les images inférées par ce réseau de neurones sont proportionnelles à l'énergie absorbée $\Phi_\lambda(\mathbf{r})\mu_\lambda(\mathbf{r})$. Tout comme dans le chapitre 4, nous ne nous intéresserons pas ici aux effets de coloration spectrale. Dans nos simulations, la fluence est supposée constante spatialement et spectralement, soit $\Phi_\lambda(\mathbf{r}) = cst$. Les images prédites par le réseau à différentes longueurs d'onde sont donc supposées directement proportionnelles à μ_λ. Dans le cas plus général de milieux diffusants et comme pour l'imagerie de fluctuation quantitative, des méthodes de correction de fluence pourront être appliquées aux images prédites par le réseau de neurones, pour passer de la distribution de l'énergie absorbée à la distribution de coefficient d'absorption, comme en imagerie conventionnelle. Nous validerons le caractère quantitatif des images prédites en estimant la SO_2 par la méthode d'inversion linéaire décrite dans le chapitre 4.

Dans un second temps, des résultats expérimentaux préliminaires sont présentés. Seul l'aspect concernant l'amélioration de la visibilité sera ici étudié : le jeu de données expérimentales est acquis à une seule longueur d'onde et ne permet donc pas d'appliquer l'approche présentée dans la section 6.1. Le réseau

de neurones est donc entraîné à prédire une unique image de fluctuation, sans conclure sur l'aspect quantitatif bien que cette image puisse théoriquement être proportionnelle à l'énergie absorbée.

Deux nouveautés sont ainsi proposées dans ce chapitre :

1. L'utilisation de l'image de fluctuation à une longueur d'onde comme *ground truth*, étudiée sur des données de simulation et de manière préliminaire sur des données expérimentales.
2. La prédiction simultanée d'images de fluctuation à plusieurs longueurs d'onde proportionnelles à l'énergie absorbée, étudiée uniquement sur des données de simulation.

6.2 Etudes sur des données de simulations

6.2.1 Construction des jeux de données de simulation et présentation de l'architecture du réseau de neurones utilisée

Framework deep learning

Un réseau de convolution de type Resunet 3D [82] est utilisé dans cette étude. Comme présenté dans le chapitre 5 (section 5.6.4), l'entrée et la sortie de chaque bloc de convolution sont connectées, limitant les risques d'explosion et de fuite de gradient et améliorant le transfert d'information vers les couches les plus profondes. Puisque le réseau est utilisé avec des images 3D, des convolutions 3D sont réalisées. Les fonctions d'activation sont des ReLU. L'algorithme d'optimisation est Adam avec un *learning rate* de 1×10^{-4} et la fonction de coût est la MSE. Un *batch size* de 8 est utilisé.

La donnée d'entrée est une matrice de taille $N_x \times N_y \times N_z \times N_\lambda$ contenant les images beamformées conventionnelles à chaque longueur d'onde λ avec N_λ la dimension correspondant au nombre canaux/longueurs d'onde. Pour ces simulations, l'image PA conventionnelle démodulée (image dmBF, voir chapitre 5 section 5.3.1) est utilisée, plutôt que l'image non démodulée (image mBF) comme c'était le cas dans le chapitre 5. Des résultats très proches sont néanmoins obtenus avec ces deux types de données d'entrée. Le réseau est entraîné à générer simultanément des images de fluctuation à chaque longueur d'onde, donc également une matrice de taille $N_x \times N_y \times N_z \times N_\lambda$.

La prédiction d'images de fluctuation à plusieurs longueurs d'onde en une seule inférence a plusieurs avantages vis à vis d'une approche dans laquelle un réseau serait entraîné à prédire une unique image, et les images à chaque longueur d'onde inférées successivement. Tout d'abord, une quantité d'information plus importante est fournie au réseau puisque les images à chaque longueur d'onde sont acquises à des instants différents et correspondent donc à des répartitions différentes d'absorbeurs. De plus, le réseau est incité à conserver les relations entre les images à chaque longueur d'onde. En contrepartie, le nombre de longueurs d'onde est fixe, ce qui implique de disposer, lors de l'inférence en phase de test/production, des images aux longueurs d'onde utilisées au cours de l'entraînement.

Génération d'objets 3D mimant des réseaux vasculaires

Le dataset est constitué d'images photoacoustique simulées à partir d'assemblages de vaisseaux 3D d'oxygénations différentes. Ces vaisseaux proviennent à l'origine d'images de segmentation 2D de vaisseaux de rétines (dataset STAR et DRIVE). Une extraction manuelle grossière de vaisseaux individuels est réalisée en sélectionnant des zones des images. Les étapes suivantes sont automatisées. Chaque vaisseau est d'abord transformé en une structure 3D plane à partir de l'image 2D initiale laquelle constituera la coupe XY intermédiaire : on cherche ici à "épaissir" le vaisseau. Dans chaque direction, les coupes XY

successives sont formées en soustrayant à la coupe précédente le contour du vaisseau qui y est contenu, déterminé à partir des gradients de cette image. Ces gradients sont calculés dans chaque direction par des filtres de Roberts. Le vaisseau s'amincit donc à mesure que l'on s'éloigne de la coupe centrale, jusqu'à sa disparition complète. Une fois la structure 3D plane obtenue, une rotation 3D aléatoire est appliquée. Une oxygénation aléatoire est ensuite attribuée au vaisseau pour la gamme de longueur d'onde choisi, en l'occurence 680 nm, 720 nm, 760 nm, 800 nm et 840 nm, ($M_\lambda = 5$). L'ensemble de cette procédure est répétée pour quelques vaisseaux qui sont ensuite assemblés pour former la structure 3D finale dont l'image photoacoustique 3D sera simulée. La méthode est schématisée en figure 6.1.

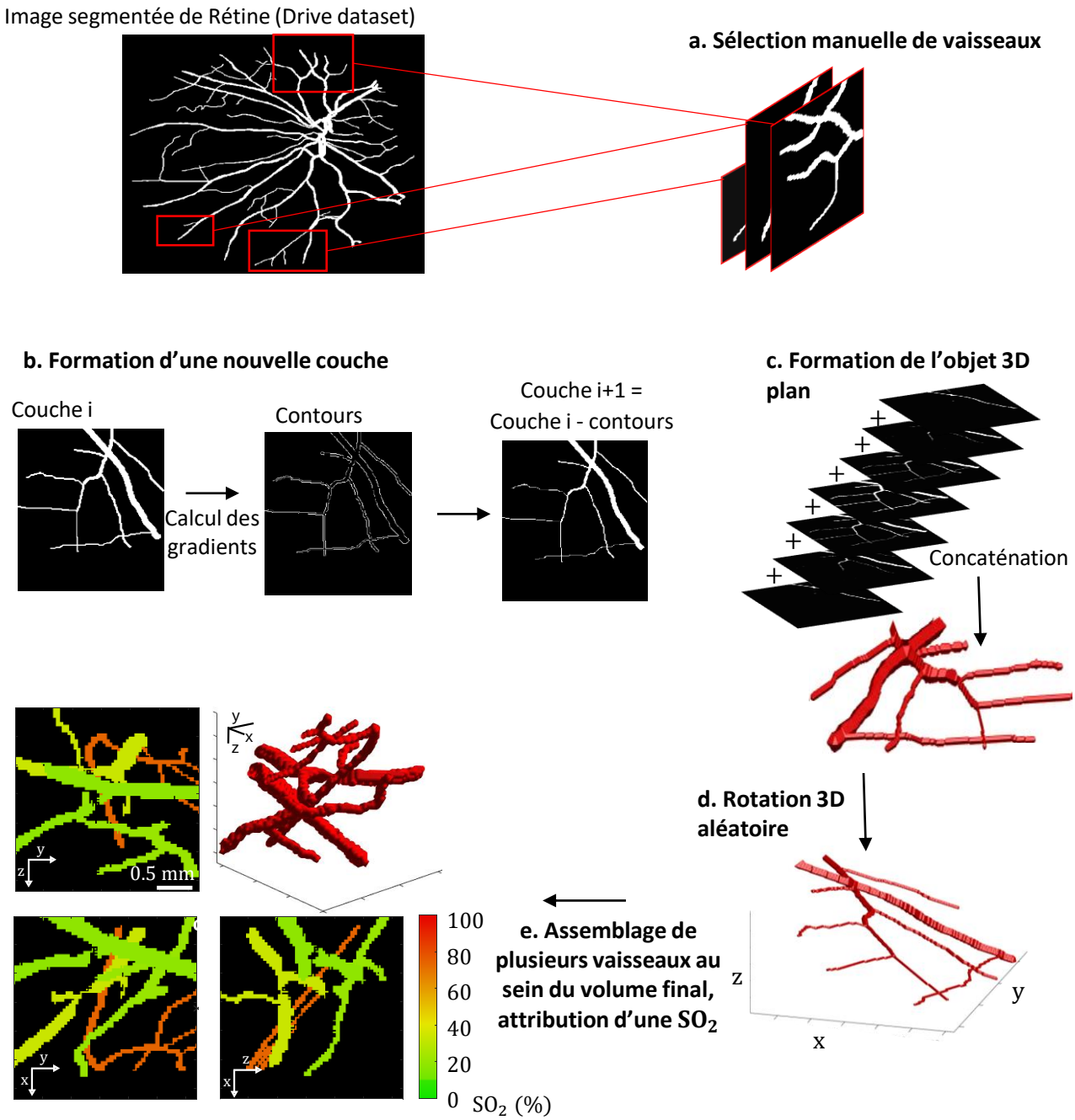

Figure 6.1 – Méthode d'obtention d'un objet 3D mimant un réseau vasculaire constitué de plusieurs vaisseaux d'oxygénations différentes.

Obtention du jeu de données

Une fois la structure 3D obtenue, des images photoacoustiques 3D correspondant chacune à des répartitions différentes d'absorbeurs sont simulés afin de calculer l'image de fluctuation. La structure 3D est remplie de particules sphériques de rayon r=5 µm représentant des globules rouges, dont les positions à l'intérieur des vaisseaux sont tirées successivement. Pour chaque nouvelle position, il est vérifié que la particule ne se superposerait pas à une particule déjà placée dans la structure, dont l'espace occupé est défini par son volume. Si c'est le cas, une nouvelle position est tirée. Sinon, celle-ci est conservée et on passe à une nouvelle particule. le processus se répète jusqu'à ce qu'une fraction volumique $\eta = 30\%$ soit atteinte. Une série de $N = 200$ répartitions de particules est générée, et les 200 acquisitions de données radiofréquences sont simulées selon la méthode décrite en annexe B. Pour limiter le temps de simulation, qui augmente avec le nombre de particules présentes dans la structure, le volume de reconstruction est limité à $2.5 \times 2.5 \times 2.5$ mm^3 (ce nombre dépend également de η et r). Le nombre de réalisation $N = 200$ est choisi empiriquement en observant la qualité de l'image de fluctuation résultante.

Le schéma en figure 6.2 décrit les étapes permettant de constituer le jeu de données.

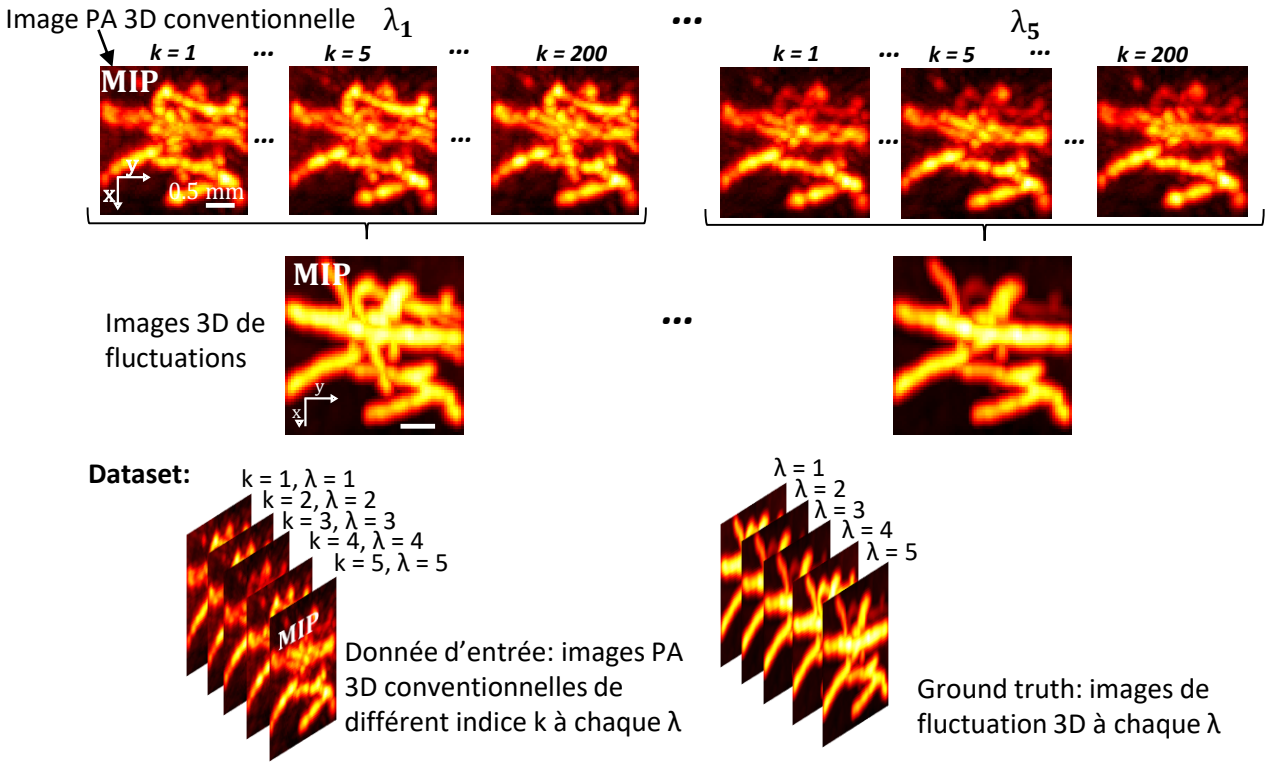

Figure 6.2 – Construction du jeu de données. Les intensités maximales projetées (MIP) selon la direction z des images 3D sont présentées. Les 5×200 images photoacoustiques 3D sont reconstruites à partir des 200 jeux de signaux radiofréquences. On calcule l'écart type de ces séries d'images 3D pour obtenir les images 3D de fluctuation à chaque longueur d'onde λ, qui formeront la ground truth. Les données d'entrée sont constituées d'une image photoacoustique 3D à chaque longueur d'onde d'indice k différent.

En conditions expérimentales, chaque acquisition de signaux radiofréquences correspond à une répartition aléatoire unique de globules rouges : elles sont donc indépendantes inter comme intra longueur d'onde. Pour s'y conformer, il serait donc nécessaire de générer $N \times M_\lambda$ distributions aléatoires et réaliser d'autant de simulations. Cependant, dans notre étude, les $N = 200$ simulations sont identiques pour chaque longueur d'onde afin de limiter la durée de la simulation. Entre chaque longueur d'onde, seuls les coefficients d'absorption attribués aux vaisseaux changent. Cette approximation est relativement correcte, puisque l'image de fluctuation est supposée indépendante de g_k dès lors qu'elle est calculée sur un

nombre suffisamment grand d'images. Pour rappel g_k représente le processus aléatoire décrivant la répartition des globules rouges à chaque acquisition k. A noter que les images en entrée du réseau à chaque λ correspondent cependant bien à des réalisations de g_k distinctes.

Du bruit électronique peut éventuellement être ajouté aux signaux radiofréquences. Les images 3D sont reconstruites par *delay-and-sum* avec un pas de reconstruction de $\lambda_{ac}/4 \approx 0.04$ mm (λ_{ac} la longueur d'onde acoustique de la sonde d'imagerie 3D), conduisant à des images de taille $N_x = 64 \times N_y = 64 \times N_z = 64$. On obtient les images de fluctuation 3D en calculant l'écart type des séries d'images 3D à chaque longueur d'onde.

Validation de la ressemblance de l'image de fluctuation à l'objet d'origine

On s'assure ici que nos images de fluctuation 3D simulées, qui constituent la *ground truth* de nos jeux de données, sont fidèles aux structures 3D d'origine. On choisit de quantifier leur ressemblance par l'écart entre les cartes de SO_2 générées à partir des images de fluctuation et les SO_2 connues des vaisseaux de la structure 3D. Les cartes de SO_2 sont obtenues par la méthode d'inversion linéaire décrite dans le chapitre 4. On moyenne pour chaque vaisseau les valeurs de SO_2 à l'intérieur de masques générés à partir de la structure 3D connue. On peut alors calculer notre écart, lequel correspond à la différence absolue entre la SO_2 estimée à partir des images de fluctuation et la SO_2 connue de chaque vaisseau. Moyennée sur l'ensemble des vaisseaux constituant une trentaine d'images de fluctuation simulée, l'erreur est de $0.6 \pm 0.8\%$, avec un maximum à 4.7%. La relation entre la SO_2 initiale et celle estimée à partir des images de fluctuations est présentée en figure 6.3, chaque vaisseau est représenté par une pastille. L'erreur irréductible de notre jeu de donnée, évaluée par rapport à l'estimation de SO_2, est donc en moyenne de 0.6%.

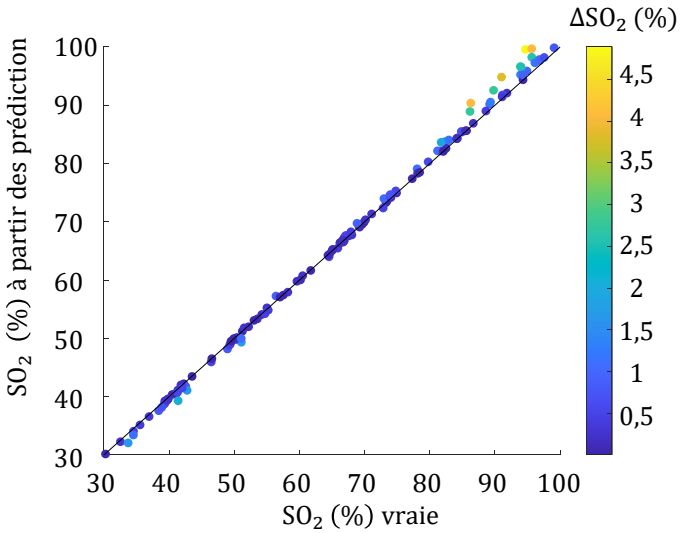

Figure 6.3 – Relation entre la saturation en oxygène (SO_2) connue de chaque vaisseau et celle estimée à partir des images de fluctuation. La couleur des points encode l'écart ΔSO_2 entre ces estimations.

6.2.2 Performances du réseau de neurones

Correction des artefacts de visibilité et diminution du *clutter*

Dans un premier temps, le réseau est entraîné sur des données non bruitées pendant environ 80 itérations avec un *training set* constitué de 600 exemples. Le *test set* est composé de 30 paires.

Les prédictions obtenues à partir de deux exemples du *test set* sont proposées en figure 6.4. Bien que les prédictions à toutes les longueurs d'onde soient obtenues par l'inférence, seule la prédiction $\lambda = 680$ nm est visible sur la figure. La ligne (a) correspond aux projections d'intensité maximale (MIP) de la donnée d'entrée, la ligne (b) à celles de la prédiction et la ligne (d) à l'image de fluctuation c'est à dire la *ground truth*. L'objet 3D d'origine est représenté à la ligne (c).

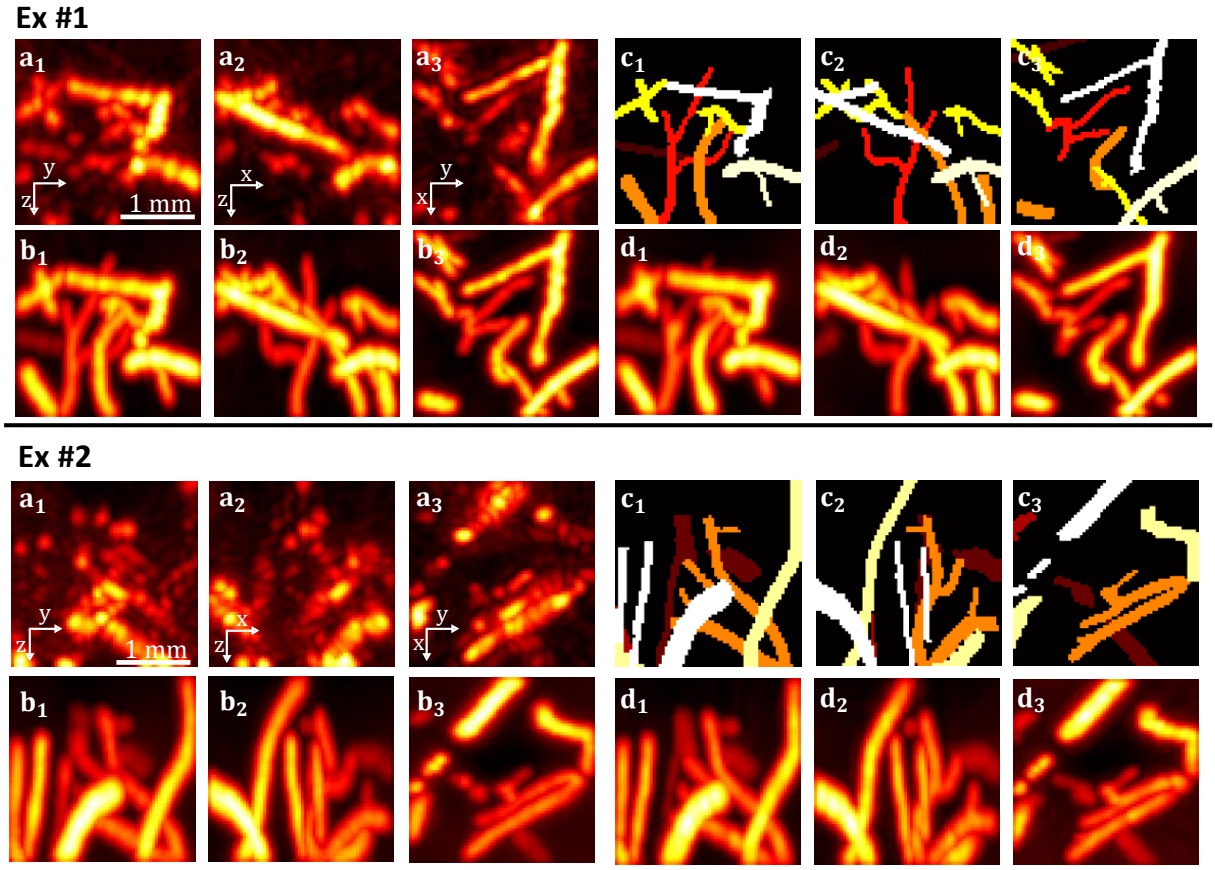

Figure 6.4 – Projections d'intensité maximale dans les plans YZ (1), XZ (2) et XY(3). L'image à 680 nm est présentée (a) : image conventionnelle composant la donnée d'entrée du réseau. (b) : image prédite par le réseau. (c) : la structure 3D d'origine, la couleur correspond au coefficient d'absorption. (d) : image de fluctuation constituant la ground truth.

Les artefacts de visibilité sont parfaitement corrigés par le réseau de neurones, et la prédiction est quasi identique à l'image de fluctuation, comme dans le cas des simulations du chapitre 5. On observe notamment l'apparition des structures verticales et la disparition du *clutter* enveloppant les vaisseaux. Les métriques utilisées pour valider quantitativement les performances du réseau sont les mêmes que celles du chapitre précèdent, la cross-corrélation normalisée (NCC) et l'index de similarité normalisé (sSSIM). Ces scores sont calculés entre les prédictions et les images de fluctuation pour chaque exemple du *test set* à chaque longueur d'ondes puis moyennées. Des performances impressionnantes sont atteintes : le sSIM est de $97.5\% \pm 0.6$ et le NCC de $98.2\% \pm 0.4$.

Estimation de la saturation en oxygène à partir des prédictions du réseau

On s'assure ensuite que les images prédites permettent de mesurer correctement la SO_2. La figure 6.5 présente pour deux exemples du *test set* les cartes de SO_2 obtenues à partir des prédictions (b), à partir des images de fluctuation (d), à partir des images conventionnelles (a) et enfin l'objet initial avec les oxygénations connues (c).

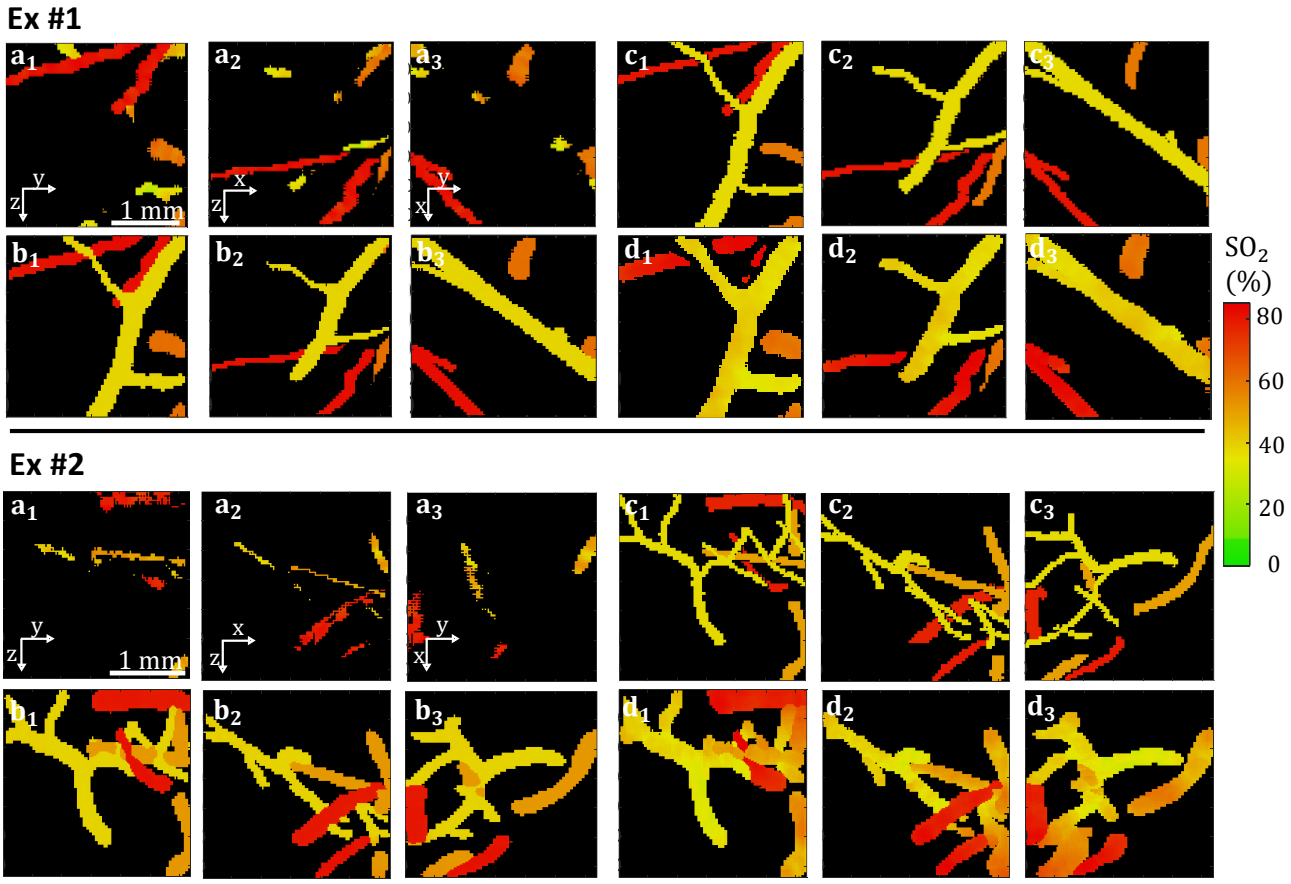

Figure 6.5 – Estimation de la saturation en oxygène (SO_2) des pixels appartenant aux projections d'intensité maximale de :
(a) : l'image conventionnelle, donnée d'entrée du réseau. (b) : l'image prédite par le réseau. (c) : la structure utilisée pour les simulations. (d) : l'image de fluctuation (ground truth). Deux exemples du test set.

Qualitativement, la carte de SO_2 estimée à partir des images produites par le réseau est similaire à celle calculée à partir des images de fluctuations. L'oxygénation est par ailleurs relativement uniforme au sein de chaque vaisseau. Sur ces données, il semble donc que le DLA soit capable de produire des images proportionnelles à l'énergie absorbée, sur des vaisseaux affectés ou non par les artefacts de visibilité, et que ces images peuvent être utilisées pour prédire la SO_2.

Pour analyser quantitativement la qualité de cette estimation sur les données de test, on compare la SO_2 moyenne de chacun des vaisseaux contenus dans les images de fluctuation avec celle des vaisseaux contenus dans les images prédites. Les masques à l'intérieur desquels est moyennée la SO_2 sont calculés à partir de la structure 3D connue (représenté en c dans les figure 6.4 et 6.5). Une erreur absolue (ΔSO_2) moyenne de $3.0 \pm 2.2\%$ avec un maximum de 9.3% est obtenue, ce qui est très satisfaisant. La relation pour chaque vaisseau entre la valeur moyenne de la SO_2 calculée à partir des images prédites et les images de fluctuation est présentée en figure 6.6.

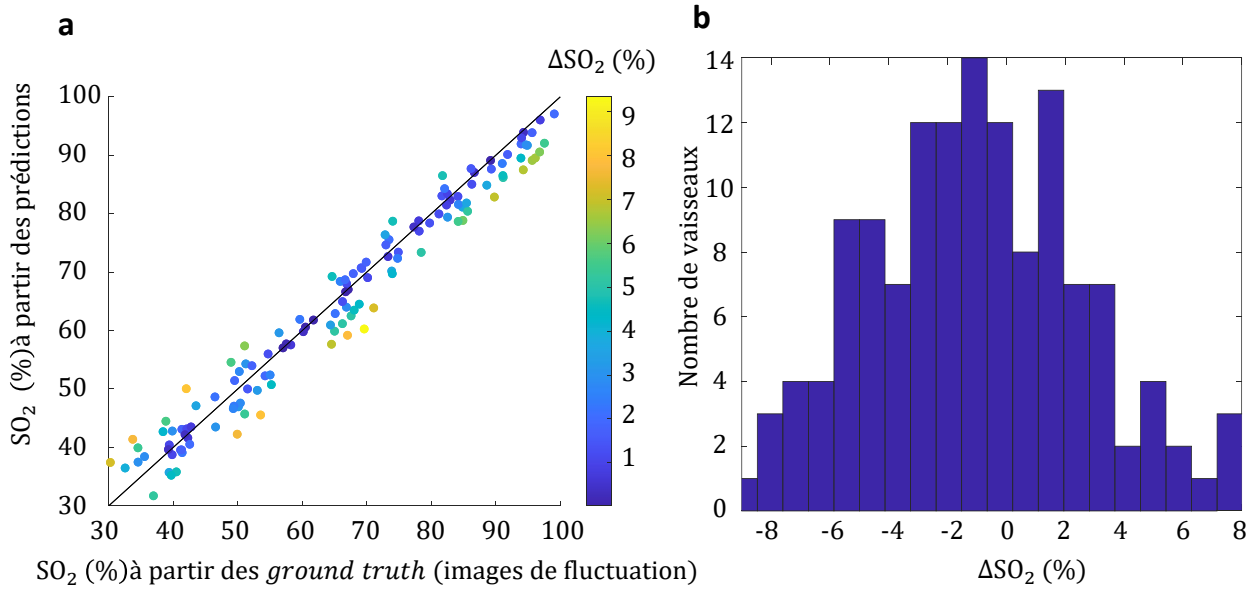

Figure 6.6 – (a) : Relation entre la valeur moyenne de l'estimation de la SO_2 de chaque vaisseau calculée à partir des images prédites et à partir des images de fluctuations. La couleur des points encode l'écart ΔSO_2 entre ces estimations. (b) : histogramme représentant le nombre de vaisseaux en fonction du pourcentage d'erreur, un bin est égal à $\Delta 0.5\%$.

6.2.3 Présence de bruit

Du bruit électronique est ensuite ajouté aux signaux radiofréquences (RF) pour obtenir des ratio signal à bruit (SNR) moyens de 20, 10 et 5. Ce dernier cas correspond à un FNR *fluctuation to noise ratio*) de 2, ce qui correspond aux conditions expérimentales. Le SNR est défini comme le rapport de la moyenne des maxima à chaque λ et l'écart type de bruit $\sigma_{b,RF}$, soit $SNR = \frac{1/M_\lambda \sum_\lambda max(RF_\lambda)}{\sigma_{b,RF}}$. La figure 6.7 présente les signaux RFs pour ces différents SNR.

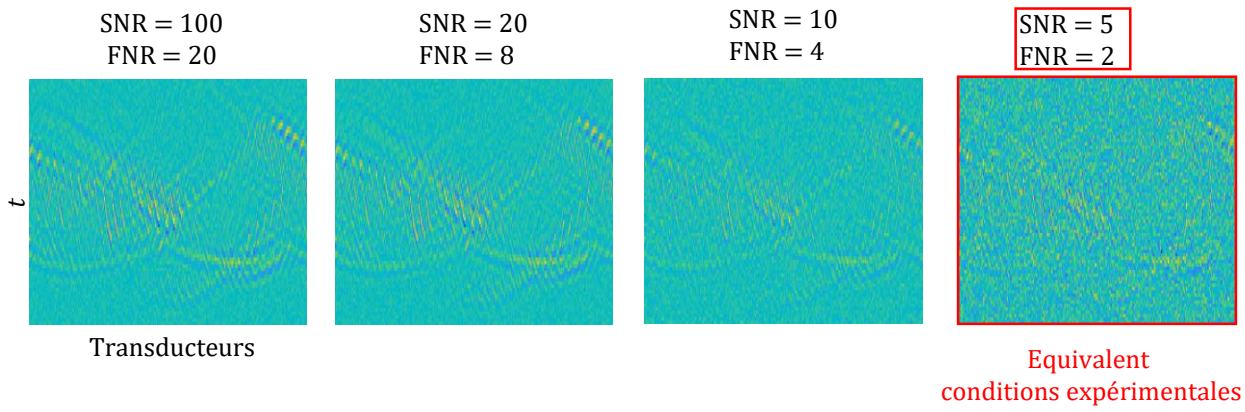

Figure 6.7 – Jeux de signaux radiofréquences de simulation pour les différents ratio de signal à bruit (signal to noise ratio SNR) et de ratio fluctuation à bruit (fluctuation to noise ratio, FNR).

Lors de l'entraînement, *la ground truth*, c'est à dire les images de fluctuation, sont corrigées du bruit (voir chapitre 4). Le réseau est donc entraîné à produire des images directement proportionnelles à l'énergie absorbée : les images sont prédites corrigées du bruit, la connaissance de $\sigma_{b,RF}$ n'est plus nécessaire. Comme dans la section précédente, on calcule l'estimation de la SO_2 moyenne de chacun des vaisseaux

contenus dans les images de fluctuation et les images prédites. La relation pour chaque vaisseau entre ces valeurs est présentée en figure 6.8 pour les différents SNR, ainsi que l'histogramme de ces erreurs.

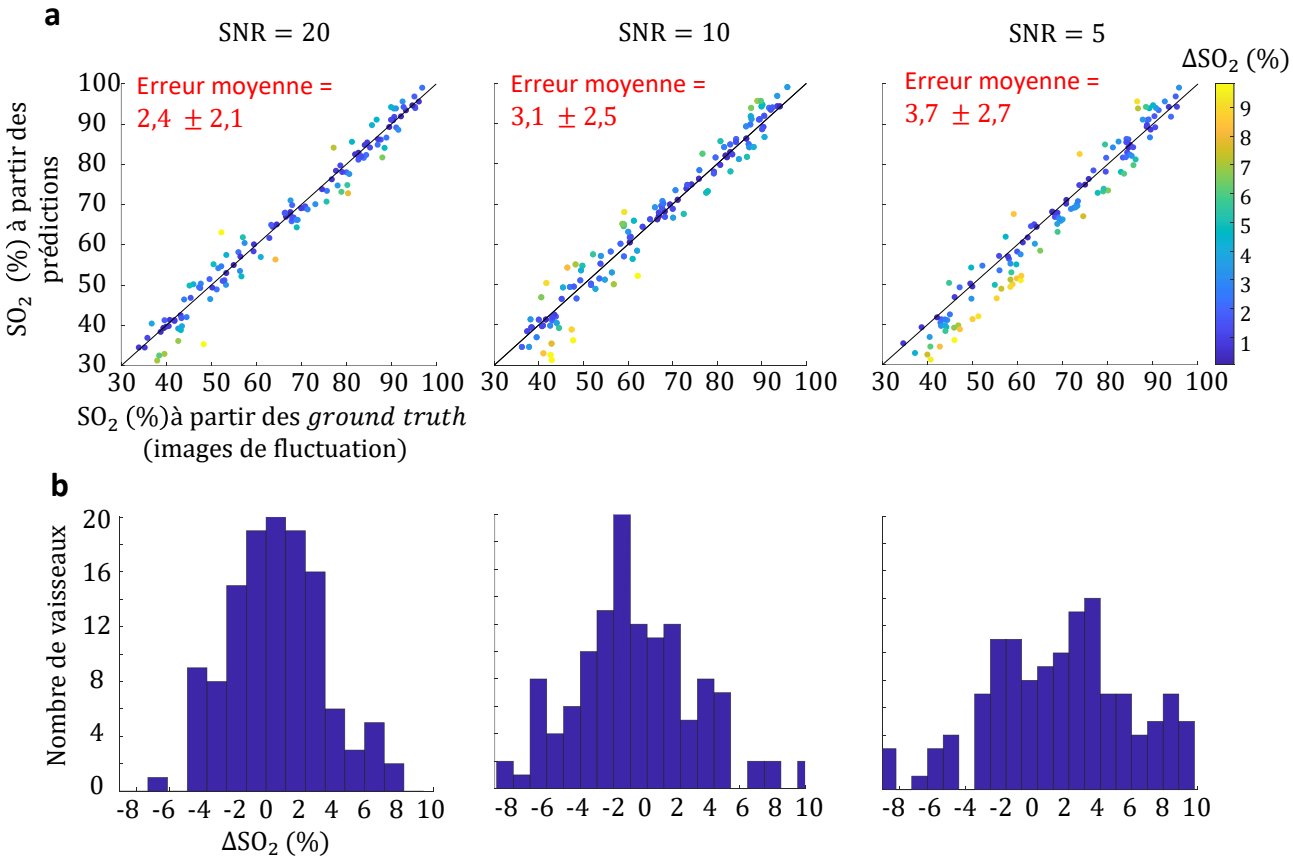

Figure 6.8 – (Pour des ratios signal à bruit (signal to noise ratio, SNR) de 20, 10 et 5 :
(a) : Relation entre la valeur moyenne de l'estimation de la SO_2 de chaque vaisseau calculée à partir des images prédites et à partir des images de fluctuations. La couleur des points encode l'écart ΔSO_2 entre ces estimations. (b) : histogramme représentant le nombre de vaisseaux en fonction du pourcentage d'erreur, un bin est égal à $\Delta 0.5\%$.

Une légère dégradation de l'estimation de la SO_2 est observable lorsque le SNR diminue, de part l'élargissement de la répartition des points autour de l'identité en (a) et de l'étalement des histogrammes en (b). Cette estimation reste néanmoins satisfaisante avec une erreur moyenne sur l'ensemble des vaisseaux du jeu de données de test de 3.7% pour le SNR de 5, et une erreur maximale de 9.8%. Les sSSIM et NCC présentés dans le tableau 6.1, restent très élevés malgré la diminution du SNR.

Tableau 6.1 – Performance du réseau de neurones en fonction du SNR évaluée à l'aide des métriques NCC et sSSIM.

	SNR 100	SNR 20	SNR 10	SNR 5
NCC	$0.982 \pm .004$	$0.982 \pm .008$	$0.975 \pm .008$	$0.964 \pm .017$
sSSIM	$0.975 \pm .006$	$0.976 \pm .006$	$0.968 \pm .006$	$0.958 \pm .025$

Ces résultats sont très encourageants. Ils démontrent que l'image de fluctuation est utilisable comme *ground truth* d'un jeu de données d'entraînement dans une approche DLA, et que la prédiction à plusieurs longueurs d'onde en une unique inférence fonctionne et permet d'estimer correctement la SO_2 à partir des images prédites. La robustesse au bruit permet d'envisager sereinement le passage à des données expérimentales.

6.3 Expériences

Dans cette section, des résultats expérimentaux préliminaires sur embryons de poulet sont présentés. Nous nous intéressons uniquement à l'amélioration de la visibilité des structures car le jeu de données expérimentales est acquis à une longueur d'onde.

6.3.1 Méthode

Création du *training set* expérimental

Le réseau de neurones ainsi que les hyperparamètres utilisés sont identiques à ceux présentés dans la section 6.2.1 pour les simulations. Puisque les données d'entrée et de sortie ainsi que la *ground truth* sont composées d'une unique image à une longueur d'onde, les dimensions sont maintenant $N_x \times N_y \times N_z \times 1$. Chaque paire d'exemples est constituée d'une image photoacoustique conventionnelle et de l'image de fluctuation calculée à partir de 1500 acquisitions. Les protocoles d'acquisition et de reconstruction de l'image de fluctuation (*delay-and-sum*, filtrage par SVD et écart type de la série d'images filtrées) sont similaires à ceux présentés dans le chapitre 3. La fluence n'est pas corrigée. A l'inverse des simulations, l'image mBF (modulée) est utilisée en entrée du réseau, les résultats étant inférieurs lorsque l'image dmBF est utilisée, ce qui correspond aux conclusions du chapitre 5. Au total, 115 paires d'images sont obtenues à partir de 35 embryons différents. Les cents premières paires sont utilisées pour constituer les jeux de données d'entraînement et les 15 restantes celui de test. On s'assure que ces 15 paires correspondent à des zones des embryons différentes de celles présentes dans le *training set*. Les images sont de taille $9 \times 9 \times 9$ 'mm^3. Le champ de vue est donc bien plus grand que celui des simulations. La création de ce jeu de données a été réalisée par Ivana Falco au cours de son stage ingénieur.

Création du *training set* de simulation pour le pré-entraînement

Les simulations présentées dans la section 6.2 produisent des images de taille $2.56 \times 2.56 \times 2.56$ mm^3. Pour pré-entrainer le réseau sur des données de simulation dans le cas des images expérimentales, de taille $9 \times 9 \times 9$ mm^3, il est nécessaire de procéder à de nouvelles simulations sur ce large champ de vue. Le volume étant 43 fois supérieur, il n'était plus envisageable de simuler suffisamment d'images photoacoustique correspondantes à différentes réalisations de g_k pour calculer l'image de fluctuation. La procédure suivante est ainsi employée. Une unique simulation est réalisée pour reconstruire l'image photoacoustique, constituant la donnée d'entrée. La *ground truth* correspond à l'image de fluctuation calculée à partir de son expression analytique (chapitre 3 section 3.2.4), consistant à la convolution de l'objet avec le module de la *point spread function* de la sonde d'imagerie 3D au carré. La racine de l'image résultante est calculée pour passer de l'image de variance à l'image de fluctuation. Cette méthode présente l'avantage de produire une image de fluctuation très rapidement et exacte, alors que l'approche statistique n'en fournit qu'une estimation. La convolution est réalisée à partir de cette *point spread function* et de l'image 3D sur-échantillonnée à $\lambda_{ac}/8$. Une représentation est proposée en figure 6.9.

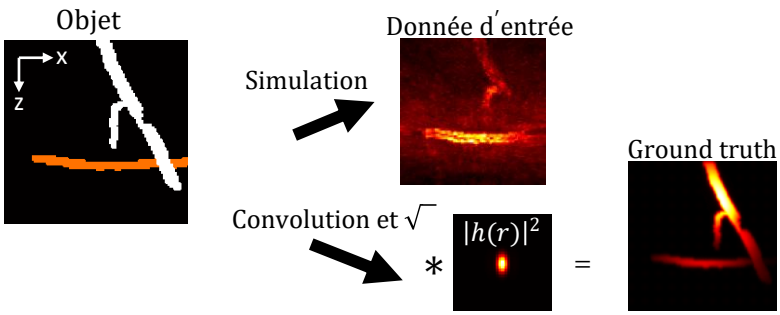

Figure 6.9 – Construction du training set de simulation utilisé pour pré-entrainer le réseau de neurones. Les intensités maximales projetées dans le plan XZ sont présentées pour représenter les images 3D. La donnée d'entrée correspond à une image photoacoustique 3D simulée par décalage des signaux radiofréquences (annexe B). La ground truth est obtenue en convoluant l'objet 3D avec le module de la point spread function de la sonde d'imagerie 3D au carré, puis en calculant la racine de l'image 3D résultante.

6.3.2 Résultats

Le réseau est pré-entraîné sur un jeu de données simulées par la méthode présentée dans la section précédente (6.3.1) et composé de 250 exemples. Un NCC de 0.91 est atteint sur le jeu de données de validation (30 exemples). Ce même réseau est ensuite ré-entraîné sur les 100 paires expérimentales, 140 itérations sont nécessaires pour que le réseau converge. Deux exemples de prédiction à partir du *test set* sont présentés en figure 6.10.

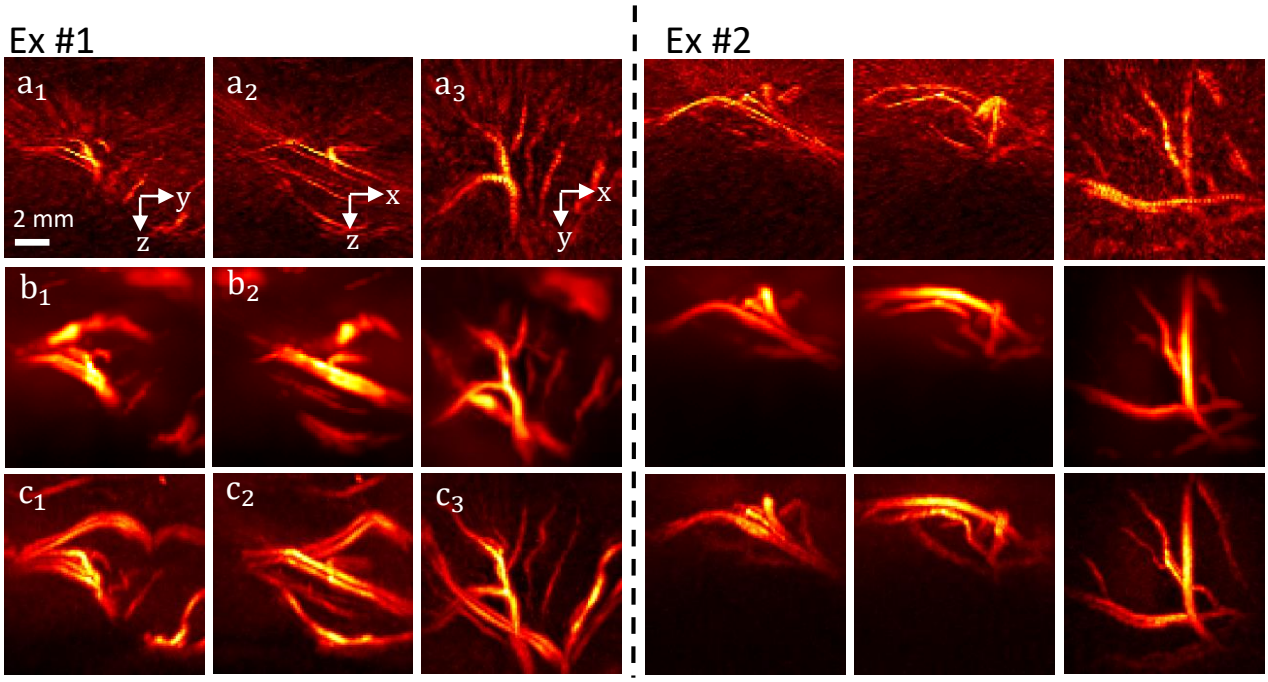

Figure 6.10 – Résultats expérimentaux de l'approche réseau de neurones avec l'image de fluctuation 3D comme ground truth. Deux exemples du test set sont présentés. La première ligne correspond à l'image conventionnelle, la seconde à l'image prédite par le réseau, et la troisième à l'image de fluctuation. les indices 1,2 et 3 correspondent respectivement à l'intensité maximale projetée dans le plan YZ, XZ et XY.

La qualité des images prédites est nettement supérieure à celles des images conventionnelles. Le clutter est réduit et des structures peu visibles sur l'image conventionnelle apparaissent. Des NCC et sSSIM moyennés sur les 10 exemples du jeu de données de test de 0.73 et 0.74 sont atteints, contre 0.24 et 0.26 pour les images conventionnelles. Bien que ces résultats soient encourageants, ils sont potentiellement perfectibles. De nombreux artefacts sont en effet présents sur l'image prédite. Notamment pour l'exemple 1, l'épaisseur et la forme des vaisseaux ne sont pas parfaitement estimées. Lorsqu'un vaisseau est complètement invisible sur l'image conventionnelle (vaisseau orienté dans l'axe de la sonde, ou noyé dans le *clutter*), il est difficilement reconstruit par le réseau.

Plusieurs raisons peuvent expliquer la baisse des performances sur les données expérimentales d'embryon de poulet 3D vis à vis des données de simulation, ou encore vis à vis des squelettes de feuille 2D du chapitre précédent, ouvrant la voie à de nombreuses perspectives d'amélioration.

1. Une première raison est la quantité de données disponibles pour l'entraînement, de 5 à 6 fois inférieure dans le cas des embryons de poulet.
2. Dans le cas des squelettes de feuille 2D et des simulations, la *ground truth* représente assez fidèlement l'ensemble des structures imagées. Ce n'est pas le cas pour les données expérimentales d'embryons de poulet. De nombreuses structures, parfois visibles sur l'image conventionnelle, n'apparaissent pas sur l'image de fluctuation, pour diverses raisons : FNR trop faible, absence de flux ... Des artefacts liés au mouvement des tissus affectent la *ground truth*, et un biais est apporté par le choix des bornes du filtrage par décomposition en valeurs singulières, puisqu'elles déterminent les signaux qui seront rejetés au cours du filtrage. Il peut en résulter des *ground truth* inconsistantes.
3. L'impact de la taille du champ de vue est également probablement conséquent. En 2D, les images conventionnelles ne sont pas affectées par les artefacts de sous-échantillonnage spatial. Dans les simulations, cet artefact est également peu présent, car le champ de vue est petit et situé autour du point focal de la sonde d'imagerie. De manière générale, les réseaux vasculaires sont très parcimonieux, résultant en un nombre conséquent de pixels uniquement composés de bruit électronique ou d'artefacts, sans structures d'intérêt ce qui complique l'apprentissage du réseau.
4. Enfin, 5 images représentant des distributions de particules différentes sont utilisées dans les simulations de la partie 6.2. A l'inverse, une unique image est fournie au réseau dans le cas expérimental, très certainement insuffisant pour séparer les structures dont le signal photoacoustique fluctue grâce à la présence de flux, de celles dont la fluctuation du signal photoacoustique est liée à d'autres causes.

On note par ailleurs que les performances du réseau entraîné sur les simulations avec le large champ de vue ($9 \times 9 \times 9$ mm^3) ne contenant qu'une image en entrée sont inférieures à celles du réseau entraîné à partir de la série de 5 simulations sur le petit champ de vue (.91 contre .98), ce qui pourrait confirmer l'impact des points 3 et 4.

Pour améliorer les performances, plusieurs options sont donc envisageables :

1. Augmenter la taille du jeu de données expérimentales.
2. Améliorer la consistance de la *ground truth*, en corrigeant les artefacts liés au mouvement, en améliorant la détectabilité ou encore en automatisant le choix des bornes de la SVD, par exemple via la méthode présentée en fin de chapitre 4.
3. Fournir au réseau une série d'images acquises successivement, en variant ou non la longueur d'onde d'illumination. Dans ce cas, de nouvelles architectures de réseaux de neurones peuvent être envisagées, plus adaptées au traitement de séries temporelles d'images telles que les réseaux de type convolution - *Long Short Term Memory* (LSTM) [146] en incluant des mécanismes d'attention, ou encore utiliser des *transformers* [32] pour améliorer l'extraction des caractéristiques de l'image.

6.4 Conclusion

La possibilité d'utiliser l'image de fluctuation comme *ground truth* dans un jeu de données pour entraîner un réseau de neurones a été démontrée en simulation, dans des conditions de bruit électronique proche de la réalité. La prédiction simultanée d'une série d'images à plusieurs longueurs d'onde permet d'estimer la SO_2 avec une erreur moyenne de l'ordre de 3.5 %.
Des résultats expérimentaux préliminaires encourageants ont été obtenus, en entraînant un réseau à reconstruire une image de fluctuation d'embryon de poulet à partir d'une unique image PA conventionnelle en entrée.

Conclusion générale

Au cours de cette thèse, nous avons exploré deux approches ayant pour principal objectif de corriger les artefacts de visibilité et d'améliorer le contraste de l'image photoacoustique : l'imagerie de fluctuation, exploitant les fluctuations du signal photoacoustique produit par les globules rouges en mouvement dans les vaisseaux sanguins, et l'utilisation du *deep learning*. Nous avons également mis en oeuvre un dispositif d'imagerie 3D permettant de réaliser simultanément l'imagerie photoacoustique de fluctuation et l'imagerie Doppler ultra-rapide avec une sonde d'imagerie parcimonieuse composée de 256 éléments. Cette sonde ainsi que le schéma expérimental, la séquence d'imagerie et l'interface de contrôle sont présentés dans le chapitre 2.

Dans le chapitre 3, nous avons proposé une théorie de l'imagerie photoacoustique de fluctuation, permettant notamment d'éclaircir les mécanismes de correction des artefacts de visibilité. L'approche a été appliquée avec succès en 3D *in vitro* et *in vivo*, sur l'embryon de poulet et la souris : une multitude de structures invisibles sur l'image conventionnelle apparaissent sur l'image de fluctuation, en raison de la correction des artefacts de visibilité et de la diminution du *clutter*. Dans ce chapitre, l'imagerie photoacoustique de fluctuation 3D est couplée à l'imagerie Power Doppler 3D, permettant d'obtenir simultanément deux images complémentaires du réseau vasculaire.

Nous avons ensuite proposée dans le chapitre 4 une théorie de l'imagerie photoacoustique quantitative de fluctuation, laquelle permet d'obtenir une image de fluctuation proportionnelle à l'énergie absorbée en présence de bruit électronique et d'une fluctuation de la fluence délivrée par le laser. L'approche est démontrée en simulation puis appliquée en 3D *in vitro* puis *in vivo* sur l'embryon de poulet. Dans ce chapitre, les effets de coloration spectrale sont négligés : l'image de fluctuation est alors proportionnelle au coefficient d'absorption optique. Cela nous a permis de calculer des cartes de saturation en oxygène à partir de ces images de fluctuation. Vis à vis de l'imagerie photoacoustique conventionnelle, la saturation en oxygène est obtenue sur une plus grande quantité de structures, permettant la réalisation d'analyses bien plus fines de l'échantillon imagé. Nous avons couplé l'imagerie photoacoustique de fluctuation quantitative avec des approches Doppler quantitatives en utilisant la sonde parcimonieuse de 256 éléments. Des cartes 3D de saturation en oxygène, de vitesse et de direction de flux sont ainsi obtenues simultanément sur l'embryon de poulet.

Dans le chapitre 5, nous avons démontré en simulation et expérimentalement, sur des fantômes mimant des réseaux vasculaires, la possibilité de corriger les artefacts de visibilité en imagerie photoacoustique avec une sonde conventionnelle d'imagerie 2D. Nous avons également mis en évidence l'intérêt d'un pré-entrainement sur des données de simulation pour diminuer la taille du jeu de données expérimentales, une caractéristique intéressante en raison de la difficulté à constituer de larges jeux de données expérimentales en imagerie médicale. La qualité des images produites par le réseau de neurones est largement supérieure à celle des images conventionnelles, en raison notamment de l'apparition des structures verticales. Quelques erreurs sont néanmoins présentes sur les images inférées, certains vaisseaux sont notamment incorrectement reconstruits (angles, formes ...). Ces phénomènes, liés à une mauvaise interprétation des données d'entrée, sont l'une des principales limitations à l'utilisation de réseau de neurones pour des applications sensibles telles que l'imagerie médicale. Dans ce chapitre, nous appliquons une approche *Monte Carlo Dropout* permettant d'obtenir une carte d'incertitude de la prédiction

du réseau sur laquelle ces erreurs sont clairement identifiables. Cette carte, qui fournit donc un degré de fiabilité de la prédiction en chaque position de l'image, est d'un intérêt considérable pour une utilisation clinique.

L'approche proposée dans le chapitre 5 est limitée par l'utilisation d'une photographie de l'objet comme *ground truth*, laquelle n'est ni proportionnelle à l'énergie absorbée, ni obtenable dans un milieu turbide. Dans le chapitre 6, nous proposons de résoudre cette problématique par l'utilisation de l'image de fluctuation comme *ground truth*, laquelle est obtenable en milieu turbide et proportionnelle à l'énergie absorbée. Cette approche est démontrée en simulation dans des conditions de bruit défavorables, et des résultats expérimentaux préliminaires dans le cadre de l'imagerie d'embryon de poulet sont présentées.

Les perspectives à ces travaux sont nombreuses. Il sera notamment nécessaire de démontrer l'applicabilité de l'imagerie de fluctuation quantitative en milieu turbide, notamment dans le cadre de l'imagerie du petit animal tel que la souris. La forte diffusivité du milieu pourra être limitante en raison de la diminution du signal photoacoustique en profondeur. Les effets de coloration spectrale devront également être pris en compte pour obtenir une estimation correcte de la carte de saturation en oxygène. De nombreuses pistes d'amélioration peuvent également être explorées pour améliorer les prédictions du réseau de neurones entraîné sur des images expérimentales d'embryon de poulet : augmentation de la taille du jeu de données, amélioration de la consistance de la *ground truth* ou encore modification de l'architecture du réseau dans le but d'exploiter une série d'images. L'obtention d'une image expérimentale proportionnelle à l'énergie absorbée doit également être démontrée.

Bibliographie

[1] Derek Allman, Austin Reiter, and Muyinatu A Lediju Bell. Photoacoustic source detection and reflection artifact removal enabled by deep learning. *IEEE transactions on medical imaging*, 37(6) :1464–1477, 2018.

[2] Stephan Antholzer, Markus Haltmeier, and Johannes Schwab. Deep learning for photoacoustic tomography from sparse data. *Inverse problems in science and engineering*, 27(7) :987–1005, 2019.

[3] Stephan Antholzer, Markus Haltmeier, and Johannes Schwab. Deep learning for photoacoustic tomography from sparse data. *Inverse problems in science and engineering*, 27(7) :987–1005, 2019.

[4] SR Arridge, M Schweiger, M Hiraoka, and DT Delpy. A finite element approach for modeling photon transport in tissue. *Medical physics*, 20(2) :299–309, 1993.

[5] Simon Bachstein. *Uncertainty Quantification in Deep Learning*. PhD thesis, MA thesis, 2019.

[6] Peter AJ Bascom and Richard SC Cobbold. On a fractal packing approach for understanding ultrasonic backscattering from blood. *The Journal of the Acoustical Society of America*, 98(6) :3040–3049, 1995.

[7] Paul Beard. Biomedical photoacoustic imaging. *Interface focus*, 1(4) :602–631, 2011.

[8] Amir Beck and Marc Teboulle. A fast iterative shrinkage-thresholding algorithm for linear inverse problems. *SIAM journal on imaging sciences*, 2(1) :183–202, 2009.

[9] Ciaran Bench, Andreas Hauptmann, and Ben T Cox. Toward accurate quantitative photoacoustic imaging : learning vascular blood oxygen saturation in three dimensions. *Journal of Biomedical Optics*, 25(8) :085003, 2020.

[10] Yoshua Bengio, Ian Goodfellow, and Aaron Courville. *Deep learning*, volume 1. Citeseer, 2017.

[11] Jeremy Bercoff, Gabriel Montaldo, Thanasis Loupas, David Savery, Fabien Mézière, Mathias Fink, and Mickael Tanter. Ultrafast compound doppler imaging : Providing full blood flow characterization. *IEEE transactions on ultrasonics, ferroelectrics, and frequency control*, 58(1) :134–147, 2011.

[12] NE Berger, RJ Lucas, and V Twersky. Polydisperse scattering theory and comparisons with data for red blood cells. *The Journal of the Acoustical Society of America*, 89(3) :1394–1401, 1991.

[13] Steinar Bjaerum, Hans Torp, and Kjell Kristoffersen. Clutter filter design for ultrasound color flow imaging. *IEEE transactions on ultrasonics, ferroelectrics, and frequency control*, 49(2) :204–216, 2002.

[14] Sarah E Bohndiek, Laura S Sasportas, Steven Machtaler, Jesse V Jokerst, Sharon Hori, and Sanjiv S Gambhir. Photoacoustic tomography detects early vessel regression and normalization during ovarian tumor response to the antiangiogenic therapy trebananib. *Journal of Nuclear Medicine*, 56(12) :1942–1947, 2015.

[15] Emmanuel Bossy, Maryline Talmant, and Pascal Laugier. Effect of bone cortical thickness on velocity measurements using ultrasonic axial transmission : A 2d simulation study. *The Journal of the Acoustical Society of America*, 112(1) :297–307, 2002.

[16] Hans-Peter F Brecht, Richard Su, Matthew P Fronheiser, Sergey A Ermilov, Andre Conjusteau, and Alexander A Oraevsky. Whole-body three-dimensional optoacoustic tomography system for small animals. *Journal of biomedical optics*, 14(6) :064007, 2009.

[17] Chuangjian Cai, Kexin Deng, Cheng Ma, and Jianwen Luo. End-to-end deep neural network for optical inversion in quantitative photoacoustic imaging. *Optics letters*, 43(12) :2752–2755, 2018.

[18] Thomas Chaigne, Bastien Arnal, Sergey Vilov, Emmanuel Bossy, and Ori Katz. Super-resolution photoacoustic imaging via flow-induced absorption fluctuations. *Optica*, 4(11) :1397–1404, 2017.

[19] Jung Woo Choe, Ömer Oralkan, and Pierre T Khuri-Yakub. Design optimization for a 2-d sparse transducer array for 3-d ultrasound imaging. In *2010 IEEE International Ultrasonics Symposium*, pages 1928–1931. IEEE, 2010.

[20] Kirsten Christensen-Jeffries, Olivier Couture, Paul A Dayton, Yonina C Eldar, Kullervo Hynynen, Fabian Kiessling, Meaghan O'Reilly, Gianmarco F Pinton, Georg Schmitz, Meng-Xing Tang, et al. Super-resolution ultrasound imaging. *Ultrasound in medicine & biology*, 46(4) :865–891, 2020.

[21] Benjamin T Cox, Simon R Arridge, and Paul C Beard. Photoacoustic tomography with a limited-aperture planar sensor and a reverberant cavity. *Inverse Problems*, 23(6) :S95, 2007.

[22] Benjamin T Cox, Simon R Arridge, and Paul C Beard. Estimating chromophore distributions from multiwavelength photoacoustic images. *JOSA A*, 26(2) :443–455, 2009.

[23] Benjamin T Cox, Jan G Laufer, Paul C Beard, and Simon R Arridge. Quantitative spectroscopic photoacoustic imaging : a review. *Journal of biomedical optics*, 17(6) :061202, 2012.

[24] Neda Davoudi, Xosé Luís Deán-Ben, and Daniel Razansky. Deep learning optoacoustic tomography with sparse data. *Nature Machine Intelligence*, 1(10) :453–460, 2019.

[25] X Luís Deán-Ben, Lu Ding, and Daniel Razansky. Dynamic particle enhancement in limited-view optoacoustic tomography. *Optics letters*, 42(4) :827–830, 2017.

[26] X Luís Dean-Ben and Daniel Razansky. Localization optoacoustic tomography. *Light : Science & Applications*, 7(4) :18004–18004, 2018.

[27] XL Deán-Ben, S Gottschalk, B Mc Larney, S Shoham, and D Razansky. Advanced optoacoustic methods for multiscale imaging of in vivo dynamics. *Chemical Society Reviews*, 46(8) :2158–2198, 2017.

[28] Charlie Demené, Thomas Deffieux, Mathieu Pernot, Bruno-Félix Osmanski, Valérie Biran, Jean-Luc Gennisson, Lim-Anna Sieu, Antoine Bergel, Stephanie Franqui, Jean-Michel Correas, et al. Spatiotemporal clutter filtering of ultrafast ultrasound data highly increases doppler and fultrasound sensitivity. *IEEE transactions on medical imaging*, 34(11) :2271–2285, 2015.

[29] Charlie Demené, Elodie Tiran, Lim-Anna Sieu, Antoine Bergel, Jean Luc Gennisson, Mathieu Pernot, Thomas Deffieux, Ivan Cohen, and Mickael Tanter. 4d microvascular imaging based on ultrafast doppler tomography. *Neuroimage*, 127 :472–483, 2016.

[30] Thomas Dertinger, Ryan Colyer, Gopal Iyer, Shimon Weiss, and Jörg Enderlein. Fast, background-free, 3d super-resolution optical fluctuation imaging (sofi). *Proceedings of the National Academy of Sciences*, 106(52) :22287–22292, 2009.

[31] Gael Diot, Stephan Metz, Aurelia Noske, Evangelos Liapis, Barbara Schroeder, Saak V Ovsepian, Reinhard Meier, Ernst Rummeny, and Vasilis Ntziachristos. Multispectral optoacoustic tomography (msot) of human breast cancer. *Clinical Cancer Research*, 23(22) :6912–6922, 2017.

[32] Alexey Dosovitskiy, Lucas Beyer, Alexander Kolesnikov, Dirk Weissenborn, Xiaohua Zhai, Thomas Unterthiner, Mostafa Dehghani, Matthias Minderer, Georg Heigold, Sylvain Gelly, et al. An image is worth 16x16 words : Transformers for image recognition at scale. *arXiv preprint arXiv :2010.11929*, 2020.

[33] Matthaios Doulgerakis, Adam T Eggebrecht, and Hamid Dehghani. High-density functional diffuse optical tomography based on frequency-domain measurements improves image quality and spatial resolution. *Neurophotonics*, 6(3) :035007, 2019.

[34] Robert Ellwood, Edward Zhang, Paul Beard, and Ben Cox. Photoacoustic imaging using acoustic reflectors to enhance planar arrays. *Journal of Biomedical optics*, 19(12) :126012, 2014.

[35] Pinhas Ephrat, Lynn Keenlislide, Adam Seabrook, Frank S Prato, and Jeffrey JL Carson. Three-dimensional photoacoustic imaging by sparse-array detection and iterative image reconstruction. *Journal of Biomedical Optics*, 13(5) :054052, 2008.

[36] Claudia Errico, Juliette Pierre, Sophie Pezet, Yann Desailly, Zsolt Lenkei, Olivier Couture, and Mickael Tanter. Ultrafast ultrasound localization microscopy for deep super-resolution vascular imaging. *Nature*, 527(7579) :499–502, 2015.

[37] Yarin Gal and Zoubin Ghahramani. Dropout as a bayesian approximation : Representing model uncertainty in deep learning. In *international conference on machine learning*, pages 1050–1059, 2016.

[38] John Gamelin, Anastasios Maurudis, Andres Aguirre, Fei Huang, Puyun Guo, Lihong V Wang, and Quing Zhu. A real-time photoacoustic tomography system for small animals. *Optics express*, 17(13) :10489–10498, 2009.

[39] Alejandro Garcia-Uribe, Todd N Erpelding, Arie Krumholz, Haixin Ke, Konstantin Maslov, Catherine Appleton, Julie A Margenthaler, and Lihong V Wang. Dual-modality photoacoustic and ultrasound imaging system for noninvasive sentinel lymph node detection in patients with breast cancer. *Scientific reports*, 5(1) :1–8, 2015.

[40] Jérôme Gateau, Thomas Chaigne, Ori Katz, Sylvain Gigan, and Emmanuel Bossy. Improving visibility in photoacoustic imaging using dynamic speckle illumination. *Optics letters*, 38(23) :5188–5191, 2013.

[41] Hartmut Gemmeke, Torsten Hopp, Michael Zapf, Clemens Kaiser, and Nicole V Ruiter. 3d ultrasound computer tomography : Hardware setup, reconstruction methods and first clinical results. *Nuclear Instruments and Methods in Physics Research Section A : Accelerators, Spectrometers, Detectors and Associated Equipment*, 873 :59–65, 2017.

[42] Gary H Glover. Overview of functional magnetic resonance imaging. *Neurosurgery Clinics*, 22(2) :133–139, 2011.

[43] Guillaume Godefroy, Bastien Arnal, and Emmanuel Bossy. Compensating for visibility artefacts in photoacoustic imaging with a deep learning approach providing prediction uncertainties. *Photoacoustics*, 21 :100218, 2021.

[44] Ian J Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial networks. *arXiv preprint arXiv :1406.2661*, 2014.

[45] Sushanth Govinahallisathyanarayana, Bo Ning, Rui Cao, Song Hu, and John A Hossack. Dictionary learning-based reverberation removal enables depth-resolved photoacoustic microscopy of cortical microvasculature in the mouse brain. *Scientific reports*, 8(1) :1–12, 2018.

[46] Janek Gröhl, Thomas Kirchner, Tim Adler, and Lena Maier-Hein. Estimation of blood oxygenation with learned spectral decoloring for quantitative photoacoustic imaging (lsd-qpai). *arXiv preprint arXiv :1902.05839*, 2019.

[47] Steven Guan, Amir Khan, Siddhartha Sikdar, and Parag Chitnis. Fully dense unet for 2d sparse photoacoustic tomography artifact removal. *IEEE journal of biomedical and health informatics*, 2019.

[48] Steven Guan, Amir A Khan, Siddhartha Sikdar, and Parag V Chitnis. Fully dense unet for 2-d sparse photoacoustic tomography artifact removal. *IEEE journal of biomedical and health informatics*, 24(2) :568–576, 2019.

[49] Steven Guan, Amir A Khan, Siddhartha Sikdar, and Parag V Chitnis. Limited-view and sparse photoacoustic tomography for neuroimaging with deep learning. *Scientific reports*, 10(1) :1–12, 2020.

[50] James A Guggenheim, Jing Li, Edward Z Zhang, and Paul C Beard. Frequency response and directivity of highly sensitive optical microresonator detectors for photoacoustic imaging. In *Photons Plus Ultrasound : Imaging and Sensing 2015*, volume 9323, page 93231C. International Society for Optics and Photonics, 2015.

[51] Zijian Guo, Changhui Li, Liang Song, and Lihong V Wang. Compressed sensing in photoacoustic tomography in vivo. *Journal of biomedical optics*, 15(2) :021311, 2010.

[52] Zijian Guo, Li Li, and Lihong V Wang. On the speckle-free nature of photoacoustic tomography. *Medical physics*, 36(9Part1) :4084–4088, 2009.

[53] Sreedevi Gutta, Venkata Suryanarayana Kadimesetty, Sandeep Kumar Kalva, Manojit Pramanik, Sriram Ganapathy, and Phaneendra K Yalavarthy. Deep neural network-based bandwidth enhancement of photoacoustic data. *Journal of biomedical optics*, 22(11) :116001, 2017.

[54] Jay S Han, Daniel M Mandell, Julien Poublanc, Alexandra Mardimae, Marat Slessarev, Cheryl Jaigobin, Joseph A Fisher, and David J Mikulis. Bold-mri cerebrovascular reactivity findings in cocaine-induced cerebral vasculitis. *Nature clinical practice Neurology*, 4(11) :628–632, 2008.

[55] Kenji Hara, Junichi Sakano, Mutsuhiro Mori, Satoshi Tamano, Ryuichi Sinomura, and Kouichi Yamazaki. A new 80v 32x32ch low loss multiplexer lsi for a 3d ultrasound imaging system. In *Proceedings. ISPSD'05. The 17th International Symposium on Power Semiconductor Devices and ICs, 2005.*, pages 359–362. IEEE, 2005.

[56] Andreas Hauptmann, Felix Lucka, Marta Betcke, Nam Huynh, Jonas Adler, Ben Cox, Paul Beard, Sebastien Ourselin, and Simon Arridge. Model-based learning for accelerated, limited-view 3-d photoacoustic tomography. *IEEE transactions on medical imaging*, 37(6) :1382–1393, 2018.

[57] Kevin Hoffer-Hawlik and Geoffrey P Luke. Abso2luteu-net : tissue oxygenation calculation using photoacoustic imaging and convolutional neural networks. 2019.

[58] Eliel Hojman, Thomas Chaigne, Oren Solomon, Sylvain Gigan, Emmanuel Bossy, Yonina C Eldar, and Ori Katz. Photoacoustic imaging beyond the acoustic diffraction-limit with dynamic speckle illumination and sparse joint support recovery. *Optics express*, 25(5) :4875–4886, 2017.

[59] Kurt Hornik, Maxwell Stinchcombe, and Halbert White. Multilayer feedforward networks are universal approximators. *Neural networks*, 2(5) :359–366, 1989.

[60] Gabriel P Howles, Ketan B Ghaghada, Yi Qi, Srinivasan Mukundan Jr, and G Allan Johnson. High-resolution magnetic resonance angiography in the mouse using a nanoparticle blood-pool contrast agent. *Magnetic Resonance in Medicine : An Official Journal of the International Society for Magnetic Resonance in Medicine*, 62(6) :1447–1456, 2009.

[61] Nam Huynh, Felix Lucka, Edward Zhang, Marta Betcke, Simon Arridge, Paul Beard, and Ben Cox. Sub-sampled fabry-perot photoacoustic scanner for fast 3d imaging. In *Photons Plus Ultrasound : Imaging and Sensing 2017*, volume 10064, page 100641Y. International Society for Optics and Photonics, 2017.

[62] Sergey Ioffe and Christian Szegedy. Batch normalization : Accelerating deep network training by reducing internal covariate shift. *arXiv preprint arXiv :1502.03167*, 2015.

[63] Akira Ishimaru. *Wave propagation and scattering in random media*, volume 2. Academic press New York, 1978.

[64] Jørgen Arendt Jensen. Field : A program for simulating ultrasound systems. In *10TH NORDICBALTIC CONFERENCE ON BIOMEDICAL IMAGING, VOL. 4, SUPPLEMENT 1, PART 1 : 351–353*. Citeseer, 1996.

[65] Jørgen Arendt Jensen and Niels Bruun Svendsen. Calculation of pressure fields from arbitrarily shaped, apodized, and excited ultrasound transducers. *IEEE transactions on ultrasonics, ferroelectrics, and frequency control*, 39(2) :262–267, 1992.

[66] J H_ Joseph, WJ Wiscombe, and JA Weinman. The delta-eddington approximation for radiative flux transfer. *Journal of Atmospheric Sciences*, 33(12) :2452–2459, 1976.

[67] Kerem Karadayi, Ravi Managuli, and Yongmin Kim. Three-dimensional ultrasound : From acquisition to visualization and from algorithms to systems. *IEEE Reviews in Biomedical Engineering*, 2 :23–39, 2009.

[68] Mustafa Karaman, Pai-Chi Li, and Matthew O'Donnell. Synthetic aperture imaging for small scale systems. *IEEE transactions on ultrasonics, ferroelectrics, and frequency control*, 42(3) :429–442, 1995.

[69] Tero Karras, Timo Aila, Samuli Laine, and Jaakko Lehtinen. Progressive growing of gans for improved quality, stability, and variation. *arXiv preprint arXiv :1710.10196*, 2017.

[70] Alex Kendall and Yarin Gal. What uncertainties do we need in bayesian deep learning for computer vision ? *arXiv preprint arXiv :1703.04977*, 2017.

[71] Diederik P Kingma and Jimmy Ba. Adam : A method for stochastic optimization. *arXiv preprint arXiv :1412.6980*, 2014.

[72] Thomas Kirchner, Janek Gröhl, and Lena Maier-Hein. Context encoding enables machine learning-based quantitative photoacoustics. *Journal of biomedical optics*, 23(5) :056008, 2018.

[73] Jens Kleesiek, Gregor Urban, Alexander Hubert, Daniel Schwarz, Klaus Maier-Hein, Martin Bendszus, and Armin Biller. Deep mri brain extraction : A 3d convolutional neural network for skull stripping. *NeuroImage*, 129 :460–469, 2016.

[74] Robert A Kruger, William L Kiser Jr, Daniel R Reinecke, and Gabe A Kruger. Thermoacoustic computed tomography using a conventional linear transducer array. *Medical physics*, 30(5) :856–860, 2003.

[75] Balaji Lakshminarayanan, Alexander Pritzel, and Charles Blundell. Simple and scalable predictive uncertainty estimation using deep ensembles. *arXiv preprint arXiv :1612.01474*, 2016.

[76] Yann LeCun, Léon Bottou, Yoshua Bengio, Patrick Haffner, et al. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11) :2278–2324, 1998.

[77] Changyeop Lee, Wonseok Choi, Jeesu Kim, and Chulhong Kim. Three-dimensional clinical handheld photoacoustic/ultrasound scanner. *Photoacoustics*, 18 :100173, 2020.

[78] Meng-Lin Li, Jung-Taek Oh, Xueyi Xie, Geng Ku, Wei Wang, Chun Li, Gina Lungu, George Stoica, and Lihong V Wang. Simultaneous molecular and hypoxia imaging of brain tumors in vivo using spectroscopic photoacoustic tomography. *Proceedings of the IEEE*, 96(3) :481–489, 2008.

[79] GR Lockwood and F Stuart Foster. Design of sparse array imaging systems. In *1995 IEEE Ultrasonics Symposium. Proceedings. An International Symposium*, volume 2, pages 1237–1243. IEEE, 1995.

[80] U-Wai Lok, Joshua D Trzasko, Ping Gong, Chengwu Huang, Shanshan Tang, Panagiotis Korfiatis, Daniel J Blezek, and Shigao Chen. Deep variational network for high quality 3d ultrasound imaging using sparse array. In *2020 IEEE International Ultrasonics Symposium (IUS)*, pages 1–4. IEEE, 2020.

[81] Jian-yu Lu. 2d and 3d high frame rate imaging with limited diffraction beams. *IEEE transactions on ultrasonics, ferroelectrics, and frequency control*, 44(4) :839–856, 1997.

[82] Geoffrey P Luke, Kevin Hoffer-Hawlik, Austin C Van Namen, and Ruibo Shang. O-net : a convolutional neural network for quantitative photoacoustic image segmentation and oximetry. *arXiv preprint arXiv :1911.01935*, 2019.

[83] Emilie Mace, Gabriel Montaldo, Bruno-Felix Osmanski, Ivan Cohen, Mathias Fink, and Mickael Tanter. Functional ultrasound imaging of the brain : theory and basic principles. *IEEE transactions on ultrasonics, ferroelectrics, and frequency control*, 60(3) :492–506, 2013.

[84] Rayyan Manwar, Xin Li, Sadreddin Mahmoodkalayeh, Eishi Asano, Dongxiao Zhu, and Kamran Avanaki. Deep learning protocol for improved photoacoustic brain imaging. *Journal of biophotonics*, 13(10) :e202000212, 2020.

[85] Fabrizio Martelli, Samuele Del Bianco, and Andrea Ismaelli. Light propagation through biological tissue and other diffusive media : theory, solutions, and software. Society of Photo-Optical Instrumentation Engineers, 2009.

[86] Y Matsumoto, Y Asao, A Yoshikawa, H Sekiguchi, M Takada, M Furu, S Saito, M Kataoka, H Abe, T Yagi, et al. Label-free photoacoustic imaging of human palmar vessels : a structural morphological analysis. *Scientific reports*, 8(1) :1–8, 2018.

[87] Elena Merčep, Xosé Luís Deán-Ben, and Daniel Razansky. Combined pulse-echo ultrasound and multispectral optoacoustic tomography with a multi-segment detector array. *IEEE transactions on medical imaging*, 36(10) :2129–2137, 2017.

[88] Gabriel Montaldo, Mickaël Tanter, Jérémy Bercoff, Nicolas Benech, and Mathias Fink. Coherent plane-wave compounding for very high frame rate ultrasonography and transient elastography. *IEEE transactions on ultrasonics, ferroelectrics, and frequency control*, 56(3) :489–506, 2009.

[89] Todd W Murray, Markus Haltmeier, Thomas Berer, Elisabeth Leiss-Holzinger, and Peter Burgholzer. Super-resolution photoacoustic microscopy using blind structured illumination. *Optica*, 4(1) :17–22, 2017.

[90] Shuai Na, Jonathan J Russin, Li Lin, Xiaoyun Yuan, Peng Hu, Kay B Jann, Lirong Yan, Konstantin Maslov, Junhui Shi, Danny J Wang, et al. Massively parallel functional photoacoustic computed tomography of the human brain. *Nature Biomedical Engineering*, pages 1–9, 2021.

[91] Kenichi Nagae, Yasufumi Asao, Yoshiaki Sudo, Naoyuki Murayama, Yuusuke Tanaka, Katsumi Ohira, Yoshihiro Ishida, Atsushi Otsuka, Yoshiaki Matsumoto, Susumu Saito, et al. Real-time 3d photoacoustic visualization system with a wide field of view for imaging human limbs. *F1000Research*, 7, 2018.

[92] R. Nagaoka, T. Tabata, R. Takagi, S. Yoshizawa, S. I. Umemura, and Y. Saijo. Development of real-time 3-d photoacoustic imaging system employing spherically curved array transducer. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 64(8) :1223–1233, 2017.

[93] Joel J Niederhauser, Michael Jaeger, Robert Lemor, Peter Weber, and Martin Frenz. Combined ultrasound and optoacoustic system for real-time high-contrast vascular imaging in vivo. *IEEE transactions on medical imaging*, 24(4) :436–440, 2005.

[94] Svetoslav Ivanov Nikolov and Jørgen Arendt Jensen. Application of different spatial sampling patterns for sparse array transducer design. *Ultrasonics*, 37(10) :667–671, 2000.

[95] Vasilis Ntziachristos. Going deeper than microscopy : the optical imaging frontier in biology. *Nature methods*, 7(8) :603–614, 2010.

[96] Nikhila Nyayapathi, Rachel Lim, Huijuan Zhang, Wenhan Zheng, Yuehang Wang, Melinda Tiao, Kwang W Oh, X Cynthia Fan, Ermelinda Bonaccio, Kazuaki Takabe, et al. Dual scan mammoscope (dsm)—a new portable photoacoustic breast imaging system with scanning in craniocaudal plane. *IEEE Transactions on Biomedical Engineering*, 67(5) :1321–1327, 2019.

[97] Guenther Paltauf, JA Viator, SA Prahl, and SL Jacques. Iterative reconstruction algorithm for optoacoustic imaging. *The Journal of the Acoustical Society of America*, 112(4) :1536–1544, 2002.

[98] Sérgio Pereira, Adriano Pinto, Victor Alves, and Carlos A Silva. Brain tumor segmentation using convolutional neural networks in mri images. *IEEE transactions on medical imaging*, 35(5) :1240–1251, 2016.

[99] Lorena Petrusca, François Varray, Rémi Souchon, Adeline Bernard, Jean-Yves Chapelon, Hervé Liebgott, W Apoutou N'Djin, and Magalie Viallon. A new high channels density ultrasound platform for advanced 4d cardiac imaging. In *2017 IEEE International Ultrasonics Symposium (IUS)*, pages 1–4. IEEE, 2017.

[100] Romain Pierrat, Jean-Jacques Greffet, Rémi Carminati, and Rachid Elaloufi. Spatial coherence in strongly scattering media. *JOSA A*, 22(11) :2329–2337, 2005.

[101] Neeraj Prabhakar and Jessica M Rosenholm. Nanodiamonds for advanced optical bioimaging and beyond. *Current Opinion in Colloid & Interface Science*, 39 :220–231, 2019.

[102] Jean Provost, Clement Papadacci, Juan Esteban Arango, Marion Imbault, Mathias Fink, Jean-Luc Gennisson, Mickael Tanter, and Mathieu Pernot. 3d ultrafast ultrasound imaging in vivo. *Physics in Medicine & Biology*, 59(19) :L1, 2014.

[103] Jean Provost, Clement Papadacci, Charlie Demene, Jean-Luc Gennisson, Mickael Tanter, and Mathieu Pernot. 3-d ultrafast doppler imaging applied to the noninvasive mapping of blood vessels in vivo. *IEEE transactions on ultrasonics, ferroelectrics, and frequency control*, 62(8) :1467–1472, 2015.

[104] Austin Reiter and Muyinatu A Lediju Bell. A machine learning approach to identifying point source locations in photoacoustic data. In *Photons Plus Ultrasound : Imaging and Sensing 2017*, volume 10064, page 100643J. International Society for Optics and Photonics, 2017.

[105] Christoph Risser, Hans Joachim Welsch, Heinrich Fonfara, Holger Hewener, and Steffen Tretbar. High channel count ultrasound beamformer system with external multiplexer support for ultrafast 3d/4d ultrasound. In *2016 IEEE International Ultrasonics Symposium (IUS)*, pages 1–4. IEEE, 2016.

[106] Justine Robin, Ali Ozbek, Michael Reiss, Xose-Luis Dean-Ben, and Daniel Razansky. Dual-mode volumetric optoacoustic and contrast enhanced ultrasound imaging with spherical matrix arrays. *arXiv preprint arXiv :2101.00622*, 2021.

[107] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net : Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention*, pages 234–241. Springer, 2015.

[108] Frank Rosenblatt. The perceptron : a probabilistic model for information storage and organization in the brain. *Psychological review*, 65(6) :386, 1958.

[109] Emmanuel Roux, François Varray, Lorena Petrusca, Christian Cachard, Piero Tortoli, and Hervé Liebgott. Experimental 3-d ultrasound imaging with 2-d sparse arrays using focused and diverging waves. *Scientific reports*, 8(1) :1–12, 2018.

[110] Jonathan M Rubin. Power doppler. 1999.

[111] David E Rumelhart, Geoffrey E Hinton, and Ronald J Williams. Learning representations by back-propagating errors. *nature*, 323(6088) :533–536, 1986.

[112] Julia L Sandell and Timothy C Zhu. A review of in-vivo optical properties of human tissues and its impact on pdt. *Journal of biophotonics*, 4(11-12) :773–787, 2011.

[113] Laurent Sandrin, Stefan Catheline, Mickaël Tanter, X Hennequin, and Mathias Fink. Time-resolved pulsed elastography with ultrafast ultrasonic imaging. *Ultrasonic imaging*, 21(4) :259–272, 1999.

[114] Jack Sauvage, Martin Flesch, Guillaume Férin, An Nguyen-Dinh, Jonathan Porée, Mickael Tanter, Mathieu Pernot, and Thomas Deffieux. A large aperture row column addressed probe for in vivo 4d ultrafast doppler ultrasound imaging. *Physics in Medicine & Biology*, 63(21) :215012, 2018.

[115] David Savery and Guy Cloutier. A point process approach to assess the frequency dependence of ultrasound backscattering by aggregating red blood cells. *The Journal of the Acoustical Society of America*, 110(6) :3252–3262, 2001.

[116] Johannes Schwab, Stephan Antholzer, Robert Nuster, and Markus Haltmeier. Real-time photoacoustic projection imaging using deep learning. *arXiv preprint arXiv :1801.06693*, 2018.

[117] Chandra M Sehgal and James F Greenleaf. Scattering of ultrasound by tissues. *Ultrasonic imaging*, 6(1) :60–80, 1984.

[118] Dinggang Shen, Guorong Wu, and Heung-Il Suk. Deep learning in medical image analysis. *Annual review of biomedical engineering*, 19 :221–248, 2017.

[119] Hoo-Chang Shin, Kirk Roberts, Le Lu, Dina Demner-Fushman, Jianhua Yao, and Ronald M Summers. Learning to read chest x-rays : Recurrent neural cascade model for automated image annotation. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 2497–2506, 2016.

[120] Tae-Hyun Shin, Pan Ki Kim, Sunghwi Kang, Jiyong Cheong, Soojin Kim, Yongjun Lim, Wookjin Shin, Joon-Yong Jung, Jungsu D Lah, Byoung Wook Choi, et al. High-resolution t 1 mri via renally clearable dextran nanoparticles with an iron oxide shell. *Nature biomedical engineering*, 5(3) :252–263, 2021.

[121] Koping Kirk Shung. On the ultrasound scattering from blood as a function of hematocrit. *IEEE Transactions on Sonics and Ultrasonics*, 29(6) :327–330, 1982.

[122] Mithun Kuniyil Ajith Singh, Kathyayini Sivasubramanian, Naoto Sato, Fumiyuki Ichihashi, Yoshiyuki Sankai, and Lei Xing. Deep learning-enhanced led-based photoacoustic imaging. In *Photons Plus Ultrasound : Imaging and Sensing 2020*, volume 11240, page 1124038. International Society for Optics and Photonics, 2020.

[123] Thorsten Spott and Lars O Svaasand. Collimated light sources in the diffusion approximation. *Applied optics*, 39(34) :6453–6465, 2000.

[124] Nitish Srivastava, Geoffrey Hinton, Alex Krizhevsky, Ilya Sutskever, and Ruslan Salakhutdinov. Dropout : a simple way to prevent neural networks from overfitting. *The journal of machine learning research*, 15(1) :1929–1958, 2014.

[125] Zbigniew Starosolski, Carlos A Villamizar, David Rendon, Michael J Paldino, Dianna M Milewicz, Ketan B Ghaghada, and Ananth V Annapragada. Ultra high-resolution in vivo computed tomography imaging of mouse cerebrovasculature using a long circulating blood pool contrast agent. *Scientific reports*, 5(1) :1–11, 2015.

[126] Heung-Il Suk, Seong-Whan Lee, Dinggang Shen, Alzheimer's Disease Neuroimaging Initiative, et al. Hierarchical feature representation and multimodal fusion with deep learning for ad/mci diagnosis. *NeuroImage*, 101 :569–582, 2014.

[127] S Torquato. *Random Heterogeneous Materials : Microstructure and Macroscopic Properties*. Springer-Verlag, New York, 2002.

[128] Victor Twersky. Acoustic bulk parameters in distributions of pair-correlated scatterers. *The Journal of the Acoustical Society of America*, 64(6) :1710–1719, 1978.

[129] Victor Twersky. Low-frequency scattering by correlated distributions of randomly oriented particles. *The Journal of the Acoustical Society of America*, 81(5) :1609–1618, 1987.

[130] Benjamin J Vakoc, Ryan M Lanning, James A Tyrrell, Timothy P Padera, Lisa A Bartlett, Triantafyllos Stylianopoulos, Lance L Munn, Guillermo J Tearney, Dai Fukumura, Rakesh K Jain, et al. Three-dimensional microscopy of the tumor microenvironment in vivo using optical frequency domain imaging. *Nature medicine*, 15(10) :1219–1223, 2009.

[131] Sergey Vilov. *Imagerie photoacoustique au-delà de la limite de diffraction acoustique*. PhD thesis, Grenoble Alpes, 2019.

[132] Sergey Vilov, Guillaume Godefroy, Bastien Arnal, and Emmanuel Bossy. Photoacoustic fluctuation imaging : theory and application to blood flow imaging. *Optica*, 7(11) :1495–1505, 2020.

[133] Dominik Waibel, Janek Gröhl, Fabian Isensee, Thomas Kirchner, Klaus Maier-Hein, and Lena Maier-Hein. Reconstruction of initial pressure from limited view photoacoustic images using deep learning. In *Photons Plus Ultrasound : Imaging and Sensing 2018*, volume 10494, page 104942S. International Society for Optics and Photonics, 2018.

[134] Lidai Wang, Guo Li, Jun Xia, and Lihong V Wang. Ultrasonic-heating-encoded photoacoustic tomography with virtually augmented detection view. *Optica*, 2(4) :307–312, 2015.

[135] Lihong Wang and Steven L Jacques. Monte carlo modeling of light transport in multi-layered tissues in standard c. *The University of Texas, MD Anderson Cancer Center, Houston*, pages 4–11, 1992.

[136] Lihong V Wang and Hsin-i Wu. *Biomedical optics : principles and imaging*. John Wiley & Sons, 2012.

[137] Xueding Wang, Xueyi Xie, Geng Ku, Lihong V Wang, and George Stoica. Noninvasive imaging of hemoglobin concentration and oxygenation in the rat brain using high-resolution photoacoustic tomography. *Journal of biomedical optics*, 11(2) :024015, 2006.

[138] Yu Wang, Zijian Guo, Lihong V Wang, Todd N Erpelding, Ladislav Jankovic, Jean-Luc Robert, and Guillaume David. In vivo three-dimensional photoacoustic imaging based on a clinical matrix array ultrasound probe. *Journal of Biomedical Optics*, 17(6) :061208, 2012.

[139] Zhou Wang, Alan C Bovik, Hamid R Sheikh, and Eero P Simoncelli. Image quality assessment : from error visibility to structural similarity. *IEEE transactions on image processing*, 13(4) :600–612, 2004.

[140] Winfried A Willinek, Dariusch R Hadizadeh, Marcus von Falkenhausen, Horst Urbach, R Hoogeveen, Hans H Schild, and Jürgen Gieseke. 4d time-resolved mr angiography with keyhole (4d-trak) : more than 60 times accelerated mra using a combination of centra, keyhole, and sense at 3.0 t. *Journal of Magnetic Resonance Imaging*, 27(6) :1455–1460, 2008.

[141] Maciej Wojtkowski. High-speed optical coherence tomography : basics and applications. *Applied optics*, 49(16) :D30–D61, 2010.

[142] Guorong Wu, Minjeong Kim, Qian Wang, Brent C Munsell, and Dinggang Shen. Scalable high-performance image registration framework by unsupervised deep feature representations learning. *IEEE Transactions on Biomedical Engineering*, 63(7) :1505–1516, 2015.

[143] Jun Xia, Muhammad R Chatni, Konstantin I Maslov, Zijian Guo, Kun Wang, Mark A Anastasio, and Lihong V Wang. Whole-body ring-shaped confocal photoacoustic computed tomography of small animals in vivo. *Journal of biomedical optics*, 17(5) :050506, 2012.

[144] Jun Xia, Junjie Yao, and Lihong V Wang. Photoacoustic tomography : principles and advances. *Electromagnetic waves (Cambridge, Mass.)*, 147 :1, 2014.

[145] Liangzhong Xiang, Bo Wang, Lijun Ji, and Huabei Jiang. 4-d photoacoustic tomography. *Scientific reports*, 3(1) :1–8, 2013.

[146] SHI Xingjian, Zhourong Chen, Hao Wang, Dit-Yan Yeung, Wai-Kin Wong, and Wang-chun Woo. Convolutional lstm network : A machine learning approach for precipitation nowcasting. In *Advances in neural information processing systems*, pages 802–810, 2015.

[147] Yuan Xu, Lihong V Wang, Gaik Ambartsoumian, and Peter Kuchment. Reconstructions in limited-view thermoacoustic tomography. *Medical physics*, 31(4) :724–733, 2004.

[148] Changchun Yang and Fei Gao. Eda-net : dense aggregation of deep and shallow information achieves quantitative photoacoustic blood oxygenation imaging deep in human breast. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 246–254. Springer, 2019.

[149] Changchun Yang, Hengrong Lan, Feng Gao, and Fei Gao. Review of deep learning for photoacoustic imaging. *Photoacoustics*, 21 :100215, 2021.

[150] Diwu Yang, Da Xing, Sihua Yang, and Liangzhong Xiang. Fast full-view photoacoustic imaging by combined scanning with a linear transducer array. *Optics express*, 15(23) :15566–15575, 2007.

[151] Jae-Chern Yoo and Tae Hee Han. Fast normalized cross-correlation. *Circuits, systems and signal processing*, 28(6) :819, 2009.

[152] Roger J Zemp, Rachel Bitton, K Kirk Shung, Meng-Lin Li, George Stoica, and Lihong V Wang. Photoacoustic imaging of the microvasculature with a high-frequency ultrasound array transducer. *Journal of biomedical optics*, 12(1) :010501, 2007.

[153] EZ Zhang, JG Laufer, RB Pedley, and PC Beard. In vivo high-resolution 3d photoacoustic imaging of superficial vascular anatomy. *Physics in Medicine & Biology*, 54(4) :1035, 2009.

[154] Zhengxin Zhang, Qingjie Liu, and Yunhong Wang. Road extraction by deep residual u-net. *IEEE Geoscience and Remote Sensing Letters*, 15(5) :749–753, 2018.

Production scientifique

Publications dans des revues internationnales à comité de lecture

Godefroy, G., Arnal, B. Bossy, E. Simultaneous 3D imaging of SO2 and blood flow dynamics by quantitative photoacoustic fluctuation imaging (qPAFI) and ultrasound Doppler imaging. *in writting*.

Godefroy, G., Arnal, B. Bossy, E. Compensating for visibility artefacts in photoacoustic imaging with a deep learning approach providing prediction uncertainties. **Photoacoustics** 21, 100218 (2021).

Vilov, S., Godefroy, G., Arnal, B. Bossy, E. Photoacoustic fluctuation imaging : theory and application to blood flow imaging. **Optica** 7, 1495–1505 (2020).

Dépot de demande de brevet

Analyse quantitative de fluctuations dans des tissus biologiques par imagerie photoacoustique multispectrale (2021).

Participation à des conférences en tant qu'orateur

conférences internationales :

Présentation invitée : Godefroy, G., Arnal, B., Bossy, E. Correcting Visibility Artefacts in Photoacoustic Imaging With a Deep Learning Approach (conference presentation). Algorithm and Reconstruction Methods (Optical Society of America, 2021). **OSA European Conferences on Biomedical Optics 2021**,

Godefroy, G., Arnal, B., Vilov, S. Bossy, E. 3D Photoacoustic Fluctuation Imaging Provides Visibility Artefacts Removal and Enhanced Contrast. Simultaneous Implementation With Ultrasound Doppler Imaging (conference presentation). Algorithm and Reconstruction Methods (Optical Society of America, 2021). **OSA European Conferences on Biomedical Optics 2021**

Godefroy, G., Arnal, B., Vilov, S. Bossy, E. 3D Photoacoustic and Ultrasound imaging of chicken embryo vasculature using a sparse array (conference presentation). In Photons Plus Ultrasound : Imaging and Sensing 2020, vol. 11240,112402I (International Society for Optics and Photonics, 2020). **SPIE Photonic West 2020**

conférences nationales :

Godefroy, G., Arnal, B ., Bossy, E. Compensating for visibility artefacts in photoacoustic imaging with a deep learning approach providing prediction uncertainties. **GT6 GDR Ondes 2021**

A

Apodisation lors de l'opération de *delay-and-sum*

Dans cette annexe, nous présentons dans la figure A.1 (b) et (c) des reconstructions par *delay-and-sum* d'une image photoacoustique 3D d'embryon de poulet avec deux matrices de coefficients d'apodisation $W_j(\mathbf{r})$ différentes ($\mathbf{r}$ la position du pixel et j le transducteur). Nous les comparons à l'image photoacoustique reconstruite par *delay-and-sum* sans apodisation ($W_j(\mathbf{r}) = 1$) en (a) et à l'image photoacoustique de fluctuation en (d).

La première matrice $W_j(\mathbf{r})$ correspond à une fenêtre de Tukey (fraction cosinus 0.35) visible en figure A.1 (e). Les éléments les plus éloignés du centre de la sonde se voient attribuer un coefficient plus faible (12% du max aux bords). Ce coefficient dépend uniquement de la position du transducteur, il ne dépend pas de $\mathbf{r}$ la position du pixel. En réduisant la contribution des signaux provenant de points éloignés de l'axe de la sonde, les lobes secondaires sont théoriquement diminués.

La seconde matrice d'apodisation correspond aux champs de vue de chaque élément, seuillés par une fonction sigmoïde dont la coupure est à environ 15% du maximum de chaque champ normalisé à 1. La matrice $W_j(\mathbf{r})$ dépend maintenant à la fois de $\mathbf{r}$ et de j. Cette matrice pour $j = 1$ est représenté en figure A.2 par ses MIPs (b). La fonction sigmoïde est également représentée en (a) dans cette figure.

L'effet des deux types d'apodisation semble négligeable sur ces images. La qualité des trois images photoacoustiques en (a) (b) et (c) de la figure A.1 sont similaire, et largement inférieure à celle de l'image photoacoustique de fluctuation.

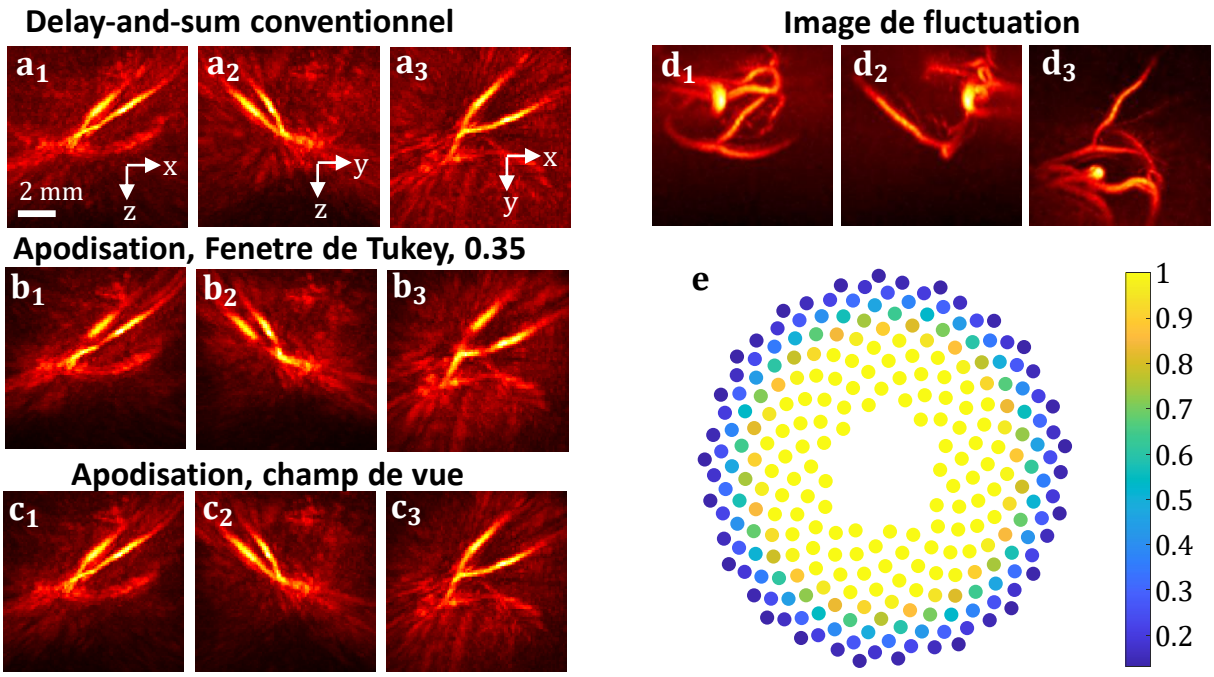

Figure A.1 – Intensités maximales projetées dans les plans XZ, YZ et XY (indice 1,2 et 3) d' images d'embryon de poulet, reconstruites par delay-and-sum sans apodisation (a), en diminuant le signal reçu par les transducteurs en bordure de la sonde selon une fenêtre de Tukey 2D de fraction cosinus 0.35 (b) et en apodisant par les coefficients correspondant à la sensibilité spatiale de chaque transducteur, seuillé à 0.15 par une fonction sigmoïde. Les coefficients de la fenêtre de Tuckey appliqués aux signaux reçus chaque transducteur sont visibles en (e). L'image de fluctuation de cet embryon est visible en (d).

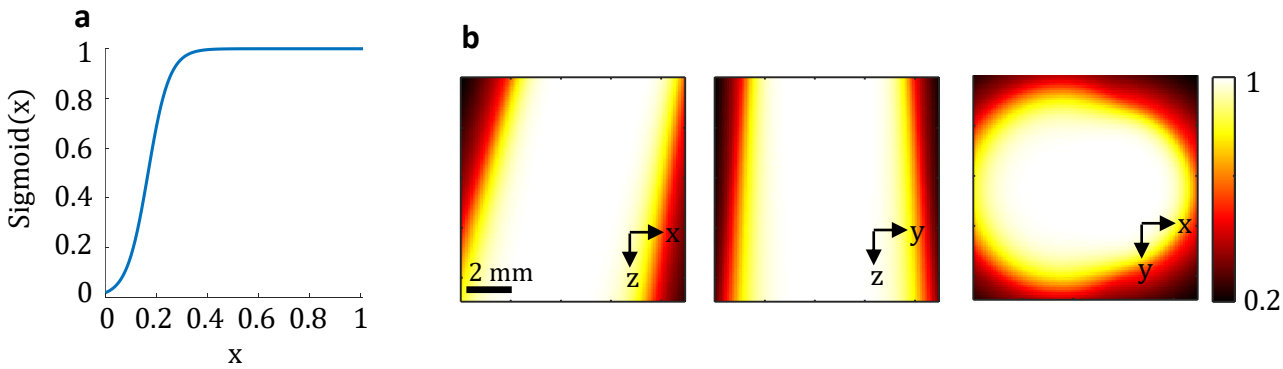

Figure A.2 – La fonction sigmoïde utilisée pour seuiller les champs de vue est visible en (a), x correspond à la valeur initial du champ de vue et Sigmoid(x) à cette valeur seuillée. Les projections des intensités maximales dans les trois directions de la matrice $W_j(\mathbf{r})$ pour j = 1 (soit le champ de vue de cet élément seuillé) est visible en (b).

B

Simulation de signaux radiofréquences

De nombreuses simulations de signaux radiofréquences (RF) photoacoustiques et ultrasonores ont été réalisées au cours de cette thèse. En fonction de l'objectif, trois méthodes différentes ont été utilisées : une méthode analytique que l'on appelle simulation par convolution avec la *point spread function* (PSF), une méthode semi-analytique utilisant le logiciel FIELD II, et une approche par différences finies avec le logiciel Simsonic [15].

Simulation par convolution avec la *point spread function*

Cette approche se prête à la simulation de jeux de signaux radiofréquences PA et US. Elle est néanmoins décrite uniquement dans le cadre de l'imagerie PA, ce qui correspond à son champ d'utilisation dans cette thèse. Elle est appliqué à la constitution des jeux de données utilisés pour l'entraînement des réseaux de neurones.

La méthode de simulation par convolution avec la PSF se base sur l'hypothèse que l'image PA est le résultat d'une convolution entre la distribution d'énergie absorbée et la PSF du système. Elle suppose également que le jeu de signaux RFs générés par le milieu imagé correspond à la somme des jeux de signaux RF générés par chacun des absorbeurs le composant. Le principe général est le suivant : les jeux de signaux RF générés par chaque absorbeur, appelés $S_k(t,j)$ avec k l'indice de l'absorbeur, sont simulés indépendamment à partir d'un jeu de signaux RF de référence correspondant aux réponses impulsionnelles des éléments d'une sonde échographique en son point focal. On appelle ce jeu de signaux de référence, obtenu par simulation ou expérimentalement, $S_p(t,j)$ avec t l'indice de temps, j l'indice du transducteur et p pour "point focal". Les signaux RF finaux représentant une acquisition de la totalité du milieu sont obtenus en sommant selon k les jeux de signaux simulés pour chaque absorbeur $S_k(t,j)$.

L'objet peut être discrétisé sur une grille de pas inférieur à la longueur d'onde acoustique λ_{ac}, et on considère alors que chaque pixel représentant l'objet correspond à un absorbeur. Les signaux RF générés en l'un de ces pixels (de position x_k, y_k, z_k) sont calculés en décalant les réponses impulsionnelles de référence $S_p(t,j)$ (du point focal x_p, y_p, z_p) de la sonde, par translation temporelle de ces signaux pour chaque transducteur (de position x_j, y_j, z_j). Pour un transducteur, cette durée est définie selon :

$$\Delta_{k,j} = \frac{1}{c}\left[\sqrt{(x_i-x_p)^2+(y_i-y_p)^2+(z_i-z_p)^2} - \sqrt{(x_k-x_j)^2+(y_k-y_j)^2+(z_k-z_j)^2}\right] \tag{B.1}$$

Ce décalage est réalisé, pour tous les transducteurs, dans le domaine de Fourier pour toutes les fréquences. Le signal provenant de deux points sources distincts diffèrent donc jusqu'ici seulement par leur temps de vol, on néglige les effets de directivité. Ce jeu de signaux décalés est ensuite multiplié par la densité d'énergie absorbée en ce pixel et l'on obtient $S_k(t, j)$. Ces jeux de signaux sont ensuite sommés entre eux pour former les données RF simulant l'acquisition PA de l'ensemble de l'objet. Le principe est représenté en figure B.1.

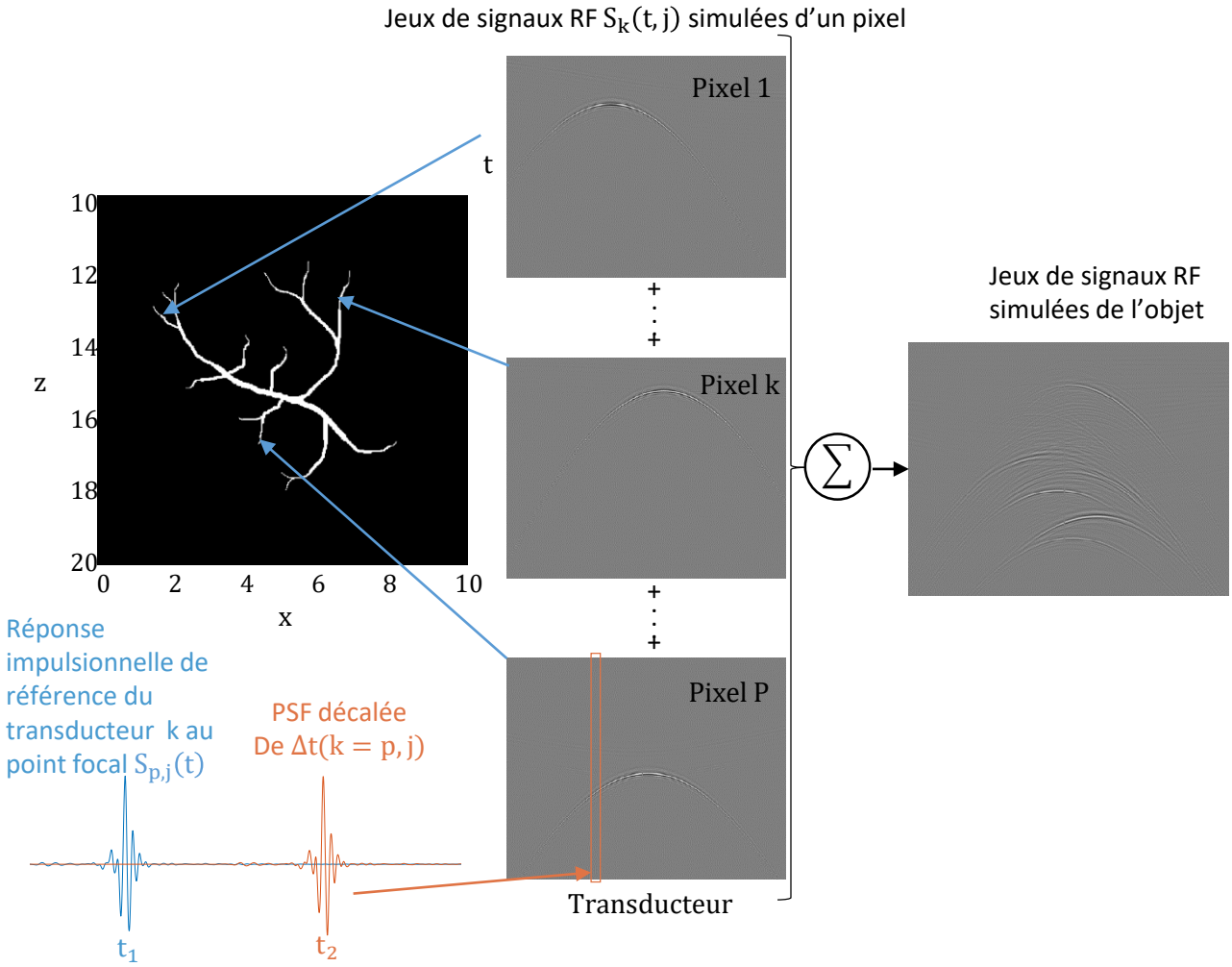

Figure B.1 – Simulation des données radiofréquences (RF) en imagerie photoacoustique. Pour chaque pixel de l'image, les signaux RF produits par un point source à la position du pixel sont obtenus en décalant temporellement un jeu de signaux de référence constitués des réponses impulsionnelles mesurées au point focal de la sonde. Les signaux RF simulés pour chacun de ces pixels sont ensuite sommés entre eux pour former le jeux de signaux RF produit par l'ensemble de l'objet.

La durée d'une simulation avec cette approche est relativement courte, une caractéristique indispensable pour constituer de grands jeux de données dans le but d'entraîner les réseaux de neurones. L'approche permet également de simuler des signaux RF à partir d'un unique jeu de signaux de référence. Cela permet donc de générer des simulations à partir de données expérimentales sans que la mesure des réponses impulsionnelles des transducteurs en tout point de la zone d'imagerie ne soit nécessaire. Cela implique de négliger les effets de directivité, ce qui est valable uniquement pour de petits champs de vue. Pour les prendre en compte, il est possible d'acquérir une série de jeu de signaux de référence en différents points de la zone de reconstruction. Les jeux de signaux RFs de chaque pixel sont alors dérivés à partir du jeu

de signaux de référence mesuré en la position la plus proche. Cela permet de mieux tenir compte de ces effets.

L'objet peut également être décrit comme une somme de particules dont la position est connue exactement, plutôt que comme une somme de pixel. Les données RFs de l'objet sont simulées de la même manière en prenant (x_k, y_k, z_k) la position de la particule.

L'approche décrite dans cette section est également appliquée pour construire la matrice A utilisée dans le chapitre 5. Cette matrice permet de réaliser la reconstruction par déconvolution de type L2. La matrice A est de taille $N_x \times N_z \times N_t \times N_{el}$ avec $N_x \times N_z$ les dimensions de l'image, N_{el} le nombre de transducteurs et N_t le nombre de pas temporel des signaux radiofréquences. Pour chaque pixel de l'image, on considère un point source dont des signaux radiofréquences photoacoustiques ($N_t \times N_e l$) sont simulés à partir de la méthode décrite dans cette section.

Simulation semi-analytique à l'aide du logiciel FIELD II

Pour reconstruire les PSF et champs de vue des sondes, les signaux radiofréquences sont obtenus de manière semi-analytique à l'aide du logiciel Field II, lequel permet de simuler un signal émis/reçu de n'importe quelle géométrie ou apodisation de transducteur, en divisant la surface de ce dernier en sous éléments cubiques. Cette approche est plus fidèle à la réalité, notamment dans le cas de large champ de vue pour lequel la PSF ne peut plus être considérée comme invariante spatialement. Cependant, la durée de la simulation est allongée, puisque les signaux temporels sont simulés indépendamment en chaque point de l'espace pour chaque transducteur.

Simulation par différences finies à l'aide du logiciel Simsonic

Enfin, certaines simulations ont été réalisées à l'aide du logiciel Simsonic [15], basé sur une approche par différences finies dans le domaine temporel qui permet de simuler la propagation d'une onde acoustique dans un milieu complexe, au prix d'une durée plus longue, notamment dans le cas d'une propagation dans un volume 3D.

C

Aspects quantitatifs des milieux aléatoires

La proportionnalité de l'image de fluctuation à μ a été exploitée dans le chapitre 4. Il est également intéressant de se pencher sur sa dépendance au phénomène g_k, et de vérifier certaines hypothèses de la première partie du chapitre 3. Ces travaux ont été réalisés par Martin Guillemaud au cours d'un stage.

Génération de milieux aléatoires

On se place dans le cadre de l'imagerie 2D. Des milieux composés de particules sous forme de disque représentant sommairement des globules rouges sont simulés par un processus qui distribue aléatoirement des sources $(g_k(\mathbf{r}) \times f(\mathbf{r}))$ sur une grille 2D de résolution suffisamment fine, pour différentes fractions volumiques η. Lors de ce processus, les positions des particules sont tirées successivement, et pour chaque nouvelle particule, il est vérifié que celle-ci ne se superpose pas à une particule déjà placée, dont l'espace occupé est défini par sa surface dans le cas 2D. Si ce n'est pas le cas, cette particule est ajoutée au milieu et ce procédé se poursuit jusqu'à ce que la valeur de η visée soit atteinte.

On peut alors vérifier que cette simulation de $g_k(\mathbf{r})$ suit bien une loi de valeur moyenne η et de σ_g. Le système est supposé ergodique, on retrouve alors $< g_k(\mathbf{r}) >_k = < g_k(\mathbf{r}) >_{\mathbf{r}} = \eta$ et $< (g_k(\mathbf{r}) - \eta)^2 >_k = < (g_k(\mathbf{r}) - \eta)^2 >_{\mathbf{r}} = \sigma_g^2 = \eta(1-\eta)$.

Puisque le cas 2D est considéré, on fixe $V_p = 2\pi r_0$ avec r_0 le rayon de la particule. La fonction d'autocovariance du processus g_k est définie selon $C(\mathbf{r}_1, \mathbf{r}_2) = < (g_k(\mathbf{r}_1) - \eta)(g_k(\mathbf{r}_2) - \eta) >_k$. Dans ce type de milieu, $C(\mathbf{0}) = \sigma_g^2$. On définit la fonction d'autocorrelation $C_n(\mathbf{r}_1, \mathbf{r}_2)$ comme la fonction d'autocovariance normalisée par $C(\mathbf{0})$, soit $C_n(\mathbf{r}_1, \mathbf{r}_2) = \frac{C(||\mathbf{r}||)}{\sigma_g^2}$.

Validation des hypothèses de l'imagerie de fluctuation

Objectifs

Dans la première partie du chapitre 3, il a été émis l'hypothèse que la fonction d'autocorrelation est une fonction avec un volume fini V_g, le volume d'autocorrelation de la particule, et que ce volume est de l'ordre de grandeur de V_p, le volume de la particule. L'intégrale de cette fonction, I_{cor} peut ainsi s'écrire selon :

$$I_{cor}(\eta) = \int_{\mathbb{R}^n} C_n(||\mathbf{r}||)d\mathbf{r} = V_g \approx V_p \tag{C.1}$$

Cette relation sera validée sur nos simulations.

Nous reviendrons également sur l'hypothèse d'une fonction $C(\mathbf{r}_1, \mathbf{r}_2)$ proportionnelle à un dirac $\delta(\mathbf{r}_1 - \mathbf{r}_2)$ lorsque la taille caractéristique de la *point spread function* (PSF) de la sonde d'imagerie $h(\mathbf{r})$ est suffisamment grande devant la taille caractéristique de la particule (dans ces simulations, $2r_0$).

1. Représentation de la fonction d'autocovariance $c(\delta_r)$

Dans nos simulations, la fonction d'autocovariance $C(\delta_r)$ est calculée dans l'espace de Fourier selon $C(\delta_r) = TF^{-1}(TF|(g)|^2)$. Le volume d'intégration $V_{integration}$ est défini selon $V_{integration} = TF^{-1}(TF|(f)|^2)$ (voir section C). On définit alors $C_p(\delta_r) = \frac{C(\delta_r)}{V_{integration}}$. Cette fonction, qui dépend de η est représentée en figure C.1.a pour $\eta = 50\%$.

Les profils selon $x+$ de cette fonction pour différents η sont représentés en figure C.1.b. Ils sont divisés par σ_g^2 produisant des fonctions proches de 1 pour $x = 0$ soit $C_p(\mathbf{0}) \approx \sigma_g^2$ comme attendu. La courbe en pointillé noire correspond au cas d'un milieu à une particule. On peut observer que les courbes colorées, représentant $C_p(\delta_r)$ à différents η, se rapprochent de cette courbe lorsque η tend vers 0 : la fonction tend donc vers l'autocorrelation d'une particule seule. Un rebond est observable, de plus en plus marqué lorsque η augmente, traduisant le rapprochement des particules et une longueur typique de corrélation.

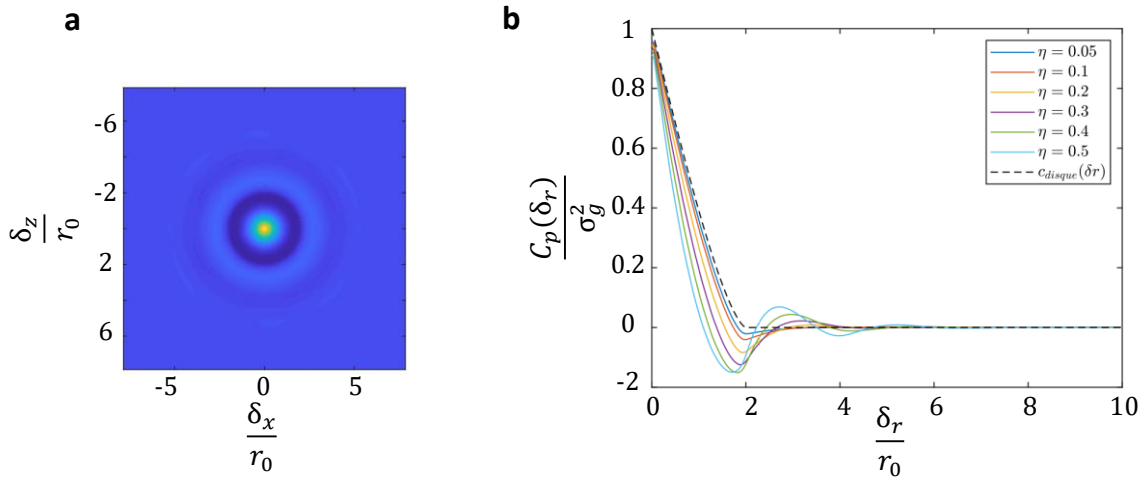

Figure C.1 – (a) : $C_p(\delta r)$ pour $\eta = 0.5$. Les axes x et z sont normalisés par le rayon d'une particule r_0. (b) : Profils de $C_p(\delta r)$ selon x_+ normalisé par r_0 pour différent η. Ces profils sont divisés par σ_g^2

2. Proportionnalité de $C(\mathbf{r}_1, \mathbf{r}_2)$ à $\delta(\mathbf{r}_1 - \mathbf{r}_2)$

I_{cor}, l'intégrale de l'autocorrélation $C_n(\delta r)$ peut maintenant être calculée pour différentes bornes d'intégration R_{max} et différents η. Les courbes résultantes sont représentées en figure C.2, dont les valeurs sont normalisées par V_p. Au maximum (pour $\eta = 50\%$), la valeur de l'intégrale se stabilise à $R_{max} \approx 7r_0$. Ainsi, lorsque le volume d'intégration est suffisamment grand (supérieur à $7r_0$ ici), la valeur de l'intégrale ne dépend plus des variations locales du milieu mais uniquement de η. Ce résultat permet de justifier l'approximation du chapitre 3 3.2.4 : lorsque la PSF est suffisamment grande devant r_0, la convolution entre celle-ci et le milieu ne dépend plus que de son volume d'intégration V_g et de ses propriétés η et σ_g. Pour $\eta = 50\%$ et un milieu composé de GRs, la taille caractéristique de la PSF doit être supérieure à 70 µm, ce qui est bien notre cas.

3. $\mathbf{I_{cor}}(\eta) = \mathbf{V_g} \approx \mathbf{V_p}$

La valeur d'intérêt est donc celle du plateau : on vérifie tout d'abord que la courbe dans le cas d'une unique particule (courbe noire) atteint une valeur proche de 1, soit $I_{cor} = V_g = V_p$, ce qui est bien le cas comme attendu (démonstration de l'égalité en C). De même, lorsque la fraction volumique diminue, I_{cor} tend bien vers V_p et reste de l'ordre de grandeur de V_p. On s'attend en réalité à ce que les valeurs de I_{cor} au niveau du plateau soient égales à $V_g(\eta) = \frac{W(\eta)}{1-\eta} V_p$ (démonstration de l'égalité en C), bien que ce résultat n'ait pas été vérifié.

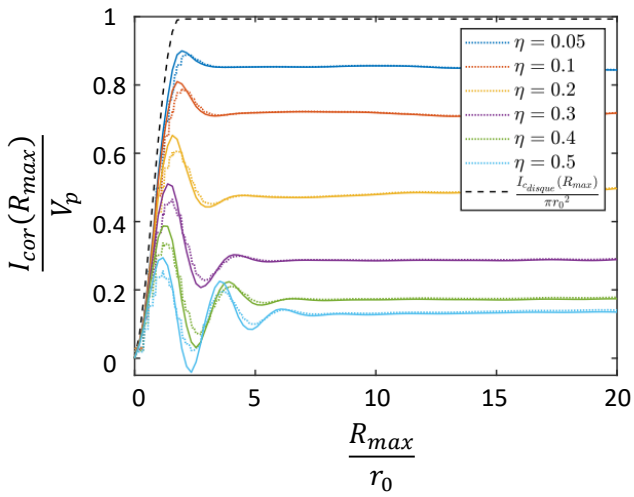

Figure C.2 – I_{corr} l'intégrale de $C_n(r)$, divisée par V_p pour différent η en fonction de la borne haute d'intégration R_{max} normalisée par r_0

Relation entre l'image de fluctuation et le phénomène aléatoire g_k

D'après l'équation du chapitre 3 3.8, l'amplitude de la fluctuation dépend fondamentalement de ses propriétés statistiques, sa variance σ_g, sa moyenne η et sa taille caractéristique V_g. Les simulations de milieux aléatoires présentées précédemment vont permettre de s'assurer que l'image de fluctuation est

bien proportionnelle à ces grandeurs et que les approximations réalisées sont valables.

Les simulations des images PA sont réalisées en utilisant la PSF 2D de la sonde linéaire L228-v (256 éléments, fréquence centrale 15.75 MHz). On se place à l'intérieur d'un objet $f(r)$ bien plus grand que la PSF, de forme similaire à celui de la figure 3.2 du chapitre 3. Dans ce cas, $\sqrt{f^2(\mathbf{r}) * |h|^2(\mathbf{r})} \simeq \sqrt{\int_{\mathbf{r}'} |h|^2(\mathbf{r}')d\mathbf{r}'} = ||h||_2$. Ainsi $\sigma[A]$ est donné par :

$$\sigma[A] = \Gamma\mu_0 F_0 \sigma_g \sqrt{V_g}||h||_2 \tag{C.2}$$

Les valeurs de l'image moyenne sont dans ce cas très proche de 0 (equation 3.6 du chapitre 3).

On peut directement étudier la relation entre $\sigma[A]$ et $I_{cov}(\eta)$. La valeur de $I_{cov}(\eta)$ est accessible via les mesures de caractérisation faites sur le milieu, à partir de $I_{cov} = \sigma_g^2 V_g = \eta W(\eta) V_p$: cette grandeur dépend donc uniquement de η en présence de particules identiques.

Environ 300 acquisitions d'images PA d'un milieu de type $f(r) \times g_k(r)$ sont ensuite simulées pour différentes fractions volumiques. Ce processus est répété trois fois pour définir une incertitude sur l'estimation de $\sigma^2[A](\mathbf{r})$. Cette estimation est obtenue en moyennant l'image sur une zone au centre de l'objet, pour s'affranchir des effets de bord. On a donc :

$$\sigma^2[A] = \left\langle < \sigma^2[A]_i(\mathbf{r}) >_r \right\rangle_3 \tag{C.3}$$

L'incertitude sur l'image de fluctuation correspond à l'écart-type moyen calculé sur ces images. Sur la fig.C.3.a, les deux courbes, normalisées par leurs maxima respectifs, se superposent. La théorie semble confirmée : le signal de fluctuation est bien proportionnel à I_{cov}. La fluctuation est maximale pour $\eta = 0.2$. La fig.C.3.b confirme la relation linéaire entre les deux grandeurs.

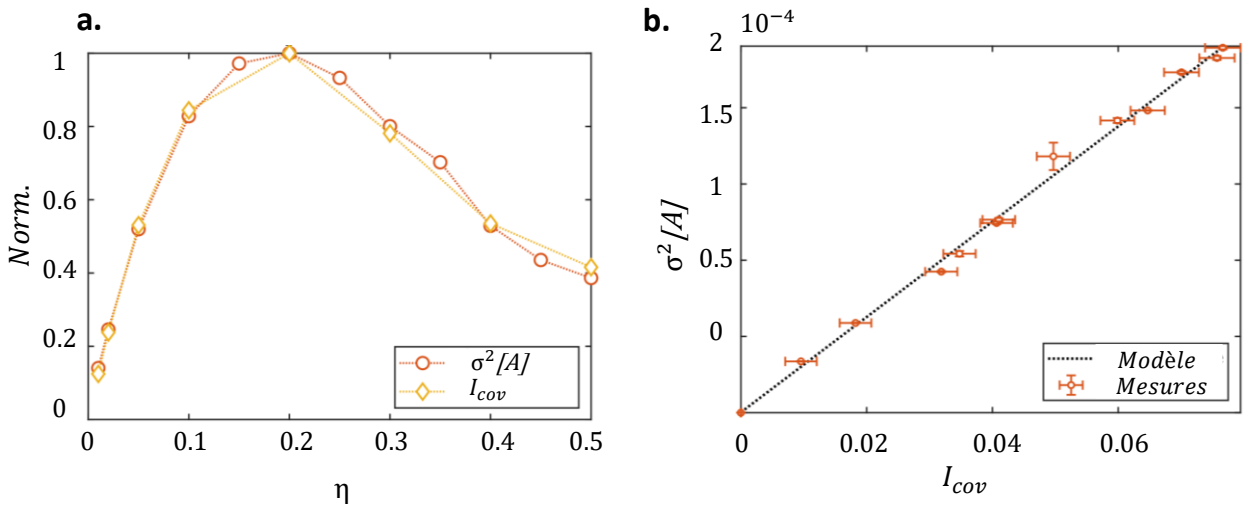

Figure C.3 – a. Moyenne de l'image de variance sur une zone au centre de l'objet et sur plusieurs réalisations, et valeur de I_{cov} pour différents η. Les courbes sont normalisées par leur maximum. b. Représentation de la relation linéaire entre ces deux grandeurs.

Ces résultats valident les approximations réalisées dans la théorie. Si on s'intéresse aux dépendances de l'imagerie de fluctuation par rapport à des grandeurs statistiques du milieu, cette étude montre que le cal-

cul de I_{cov} suffit et qu'il n'est pas forcément nécessaire de simuler tout le processus d'imagerie. Il serait intéressant d'étudier jusqu'à quelle fréquence le résultat C.3 est valide. Il est à noter que contrairement à une distribution de Bernouilli (tirage avec remise) où la fluctuation serait maximale à $\eta = 0.5$, la taille finie des particules induit un décalage du maximum de fluctuation autour de $\eta = 0.2$. Ceci est du à l'apparition de corrélations spatiales dans les distributions denses de particules qui limite le renouvellement de configurations très différentes de particules et réduit alors la fluctuation du signal photoacoustique. Ces faits, déjà observés en imagerie ultrasonore montrent que le sang à une hématocrite normale n'est pas dans une configuration idéale pour l'imagerie de fluctuations.

Relation entre V_g et V_p, milieu à une particule

Montrons tout d'abord que pour un milieu à une particule, l'intégrale de la fonction d'autocovariance normalisée, I, est égale V_p, le volume de la particule. Soit $S_p(\mathbf{r})$ la fonction valant 1 dans la particule et 0 ailleurs. On remarque que $S_p(\mathbf{r})^n = S_p(\mathbf{r})$, on peut alors écrire :

$$V_p = \int S_p(\mathbf{r}) d\mathbf{r} \tag{C.4}$$

Soit $u_p(r)$ l'autocorrelation de la fonction S_p, ici celle de la particule $\delta_\mathbf{r}$, égal à :

$$u_p(\delta_\mathbf{r}) = \int S_p(\mathbf{r}) S_p(\mathbf{r} + \delta_\mathbf{r}) d\mathbf{r} \tag{C.5}$$

En normalisant par V_p, on a $u_p^n(\delta_\mathbf{r}) = \frac{u_p(\delta_\mathbf{r})}{V_p}$. L'intégrale I de cette fonction est alors égale à

$$I = \int_{\mathbb{R}^n} u_p^n d\mathbf{r}_1 = \frac{1}{Vp} \int_{\mathbf{r}_1} \int_{\mathbf{r}_2} S_p(\mathbf{r}_2) S_p(\mathbf{r}_2 + \mathbf{r}_1) d\mathbf{r}_1 d\mathbf{r}_2 \tag{C.6}$$

Le théorème de Fubini nous permet d'écrire :

$$I = \frac{1}{V_p} \int_{\mathbf{r}_2} S_p(\mathbf{r}_2) \left[\int_{\mathbf{r}_1} S_p(\mathbf{r}_2 + \mathbf{r}_1) d\mathbf{r}_1 \right] d\mathbf{r}_2 \tag{C.7}$$

et finalement

$$I = \frac{1}{V_p} \int_{\mathbf{r}_2} S_p(\mathbf{r}_2) \left[V_p \right] d\mathbf{r}_2 = V_p \tag{C.8}$$

Relation entre V_g et V_p, Milieu de particule aléatoirement réparties

Le milieu est maintenant composé de particules identiques réparties aléatoirement. Les définitions de s_p, u_p et V_p sont conservées. Pour décrire un tel milieu, la densité aléatoire microscopique $N_k(\mathbf{r}) = \sum_i \delta(\mathbf{r}-\mathbf{r}_i)$ est introduit, décrivant la position de toutes les particules à la réalisation k. La fonction g_k peut être écrite comme $g_k = N_k * S_p$. La fonction d'autocovariance, définie précédemment telle que :

$$C(\mathbf{r}_1,\mathbf{r}_2) =< g_k(\mathbf{r}_1) \times g_k(\mathbf{r}_2) >_k -\eta^2 \tag{C.9}$$

peut s'écrire comme la fonction d'auto covariance de la densité microscopique aléatoire :

$$C_g(\mathbf{r}_1,\mathbf{r}_2) =< [N_k * S_p](\mathbf{r}_1) \times [N_k * S_p](\mathbf{r}_2) >_k -\eta^2 \tag{C.10}$$

En introduisant $A_N(\mathbf{r}_1,\mathbf{r}_2) =< N_k(\mathbf{r}_1) \times N_k(\mathbf{r}_2)) >_k$, on peut montrer, sous l'hypothèse de stationnarité :

$$C_g = A_N * U_p - \eta^2 \tag{C.11}$$

En intégrant l'équation ci-dessus et en appelant $\int A_n * u_p d\mathbf{r} = \int A_n(\mathbf{r})d\mathbf{r} \times \int u_p(\mathbf{r})d\mathbf{r} = \int A_n(\mathbf{r})V_p^2 d\mathbf{r}$, V_g peut être exprimé selon :

$$V_g = \frac{1}{\sigma_g^2}\int [A_N \times V_p^2 - \eta^2]d\mathbf{r} \tag{C.12}$$

La fonction de corrélation paire $B(\mathbf{r})$, très utilisée en physique de la matière condensée [127], et précédemment introduite en physique ultrasonore [6, 12, 115, 121, 128, 129], est reliée à la densité microscopique aléatoire [115] :

$$A_N(\mathbf{r}) = m^2 B(\mathbf{r}) + m\delta(\mathbf{r}) = \left(\frac{\eta}{V_p}\right)^2 B(\mathbf{r}) + \frac{\eta}{V_p}\delta(\mathbf{r}) \tag{C.13}$$

avec $m = \left(\frac{\eta}{V_p}\right)$ la concentration moyenne de particule. En combinant Eqs. C.12, C.13, et $\sigma_g^2 = \eta \times (1-\eta)$, on obtient l'expression de V_g suivante :

$$V_g = V_p \times \frac{1}{1-\eta}[1 + \frac{\eta}{V_p}\int (B(\mathbf{r})-1)d\mathbf{r})] \tag{C.14}$$

avec $W(\eta) = [1 + \frac{\eta}{V_p}\int (B(\mathbf{r})-1)d\mathbf{r})]$ correspondant à la valeur pour $\mathbf{q} = \mathbf{0}$ du facteur de structure $S(q)$:

$$S(\mathbf{q}) = 1 + \frac{\eta}{V_p}\int [B(\mathbf{r})-1]e^{-i\mathbf{q}\cdot\mathbf{r}}d\mathbf{r}$$

connu sous le nom de facteur d'encombrement (*packing factor*) [6, 12, 121, 128, 129] par la communauté ultrasonore. Finalement, la relation reliant entre V_g, V_p, η et $W(\eta)$ peut être obtenue :

$$V_g(\eta) = \frac{W(\eta)}{1-\eta}V_p \tag{C.15}$$

Méthode spectrale du calcul d'autocorrélation

Les calculs de corrélation sont réalisés dans l'espace des fréquences. L'autocorrelation spatiale de g est définie selon :

$$\begin{aligned}\mathbf{TF}(g \otimes g) &= \mathbf{TF}(g) \times \overline{\mathbf{TF}(g)} \\ &= |\mathbf{TF}(g)|^2\end{aligned} \tag{C.16}$$

On peut finalement écrire :

$$C(\delta r) = \mathbf{TF}^{-1}(|\mathbf{TF}(g)|^2) \tag{C.17}$$

Pour obtenir $C_p(\delta r)$, il est nécessaire de déterminer l'espace de calcul de l'intégrale $V_{integration}$. La fonction g étant une fonction à support fini, l'espace de recouvrement dépend de δr. En reprenant la définition de la fonction f, valant 1 dans la zone imagée et 0 ailleurs, on a pour δ_r :

$$V_{integration} = (f \otimes f)(\delta_r) = \int_{espace} f(r) f(r + \delta r) dr = \mathbf{TF}^{-1}(|\mathbf{TF}(f)|^2)(\delta_r) \tag{C.18}$$

D

Impact de la décomposition en valeurs singulières sur l'estimation de la saturation en oxygène en simulation sur une seconde géométrie : la boucle

Cette annexe fait suite à la section 4.5.2 du chapitre 4. Une seconde géométrie, de type boucle, est étudiée pour s'assurer que les observations faites sur le disque, sujet aux artefacts causés par la réponse fréquentielle limitée de la sonde, se généralisent aux artefacts liés à l'orientation verticale des structures vis à vis à la sonde.

Le panneau de gauche de la figure D.1 présente une simulation avec un FNR de 0.5, et une fluctuation selon $\Phi_{k,\lambda}$ de mêmes paramètres que dans le cas du disque, pour des bornes de filtrage a = 2 et b = 2560. Deux zones de la boucle y sont étudiées : l'une visible sur l'image moyenne (partie horizontale, cadre cyan) et l'autre invisible (partie verticale, cadre vert). Les coefficients d'absorption sont correctement estimés aussi bien sur les structures affectées par les problèmes de visibilité que sur celles ne l'étant pas, et les estimations de SO_2 sont similaires.

Le deuxième panneau propose une étude similaire à celle présentée dans la figure 4.15 du chapitre 3. Les courbes d'absorption sont calculées sur la ROI correspondant à la partie verticale de la boucle. Un comportement similaire à celui décrit dans le cas du disque est observé, néanmoins une variation spatiale de la SO_2 est visible pour a = 50 (4ème ligne, première et seconde colonne). Cette variation s'explique par le fait que le signal provenant de la partie basse de la boucle est plus faible, car celle-ci est plus éloignée de la sonde. Ainsi, cette structure est presque complètement noyée dans le bruit comme le montre l'image de fluctuation, tandis que la partie haute reste bien visible. Il en résulte un effet de a largement supérieur sur la partie basse que sur la partie haute, et donc une estimation de la SO_2 moins précise. Toujours sur cette ligne, on observe que lorsque b diminue (notamment pour b = 500), cette variation spatiale de l'estimation de la SO_2 disparaît. En effet, le bruit diminue, l'objet est donc mieux estimé (et donc plus visible, comme le montre l'image de fluctuation $a = 50$, $b = 500$ vis à vis de celle pour $a = 50$, $b = 2560$) et l'impact de la borne a en est diminué.

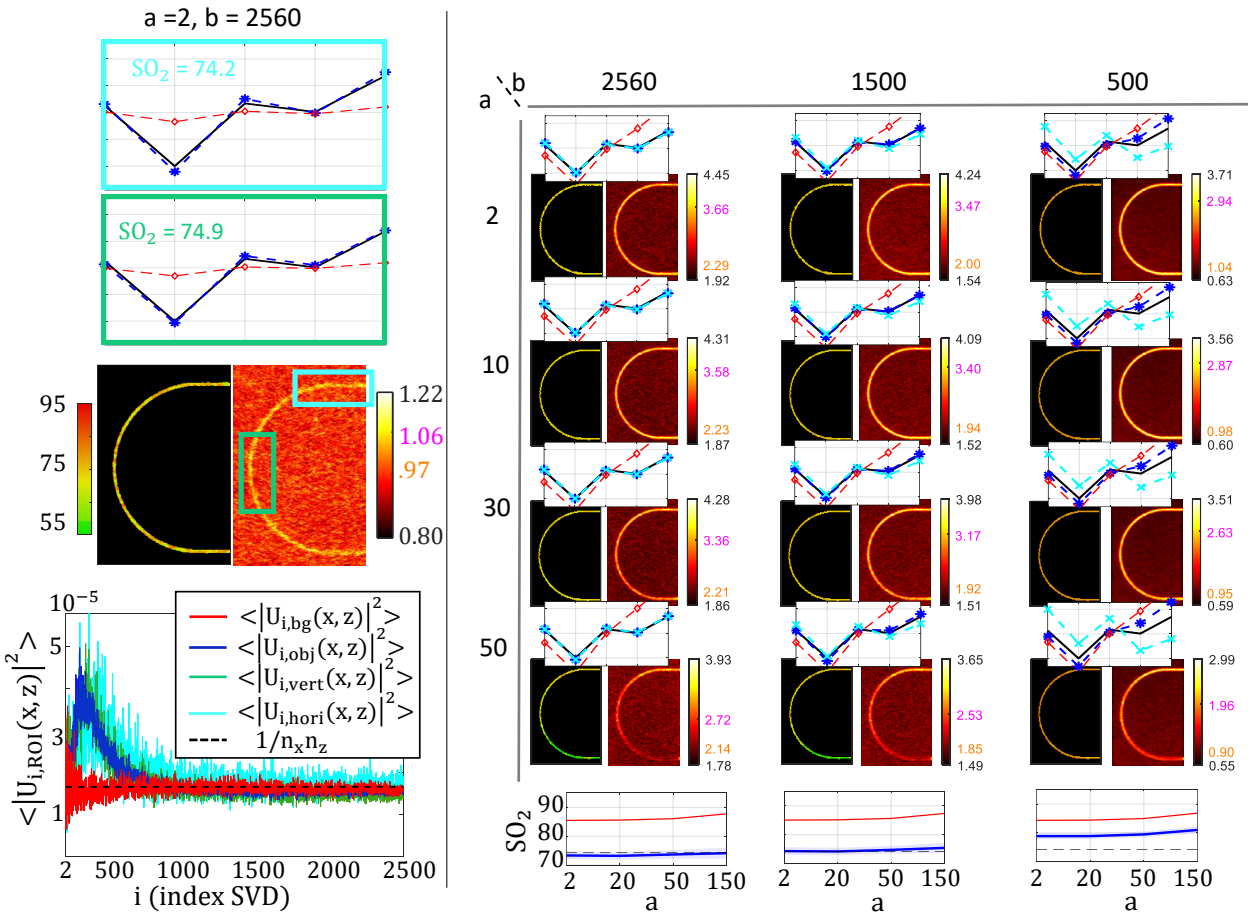

Figure D.1 – Étude de l'impact des bornes a et b sur un objet de type boucle. Cas avec bruit et variation de la fluence selon $\Phi_{k,\lambda}$. Les courbes représentent les coefficients d'absorption théorique en noir, et ceux estimés à partir des images de fluctuation sans correction du bruit (rouge), sans correction de l'effet de la décomposition en valeurs singulières sur l'estimation du bruit (cyan) et en corrigeant cet effet (bleu). Les nombres en magenta et orange sur la barre de couleur des images de fluctuation correspondent respectivement à la valeur moyenne des pixels au centre de l'objet et dans une zone de bruit.
Panneau de gauche : pour a= 2 et b = 2560, analyse de deux zones de la boucle : la partie horizontale supérieure, non affectée par les artefacts de visibilité, et la partie verticale, affectée par l'artefact d'ouverture numérique limitée. Les magnitudes au carré des vecteurs spatiaux moyennées en différentes zones de l'image (partie horizontale en cyan, partie verticale en vert, totalité de l'objet en bleu et bruit de fond en rouge) sont représentées en fonction de l'index i de la SVD.
Panneau de droite : pour différents couples de bornes de filtrage, l'image de fluctuation et l'image de SO_2 sont présentées, ainsi que les différentes courbes représentant les coefficients d'absorption. Les valeurs d'oxygénation en fonction de a pour les différents b sont affichées dans la dernière ligne de la figure.

Nouvelles approches pour l'imagerie photoacoustique des vaisseaux sanguins : imagerie quantitative 3D de fluctuations et reconstruction assistée par apprentissage profond

Résumé

L'imagerie de la vascularisation présente un intérêt considérable pour de multiples applications de santé, incluant le diagnostic de pathologies, le guidage d'intervention chirurgicale, le suivi de traitement ou encore la recherche biomédicale. Si plusieurs modalités permettent aujourd'hui d'imager la vascularisation, l'utilisation de dispositifs bimodaux combinant l'imagerie photoacoustique (PA) et l'imagerie ultrasonore (US) constitue une approche prometteuse. La sensibilité de l'imagerie PA à l'hémoglobine et l'apparition de méthodes sensibles au flux en imagerie US permettent de cartographier avec précision le réseau vasculaire au sein des tissus biologiques à des profondeurs de l'ordre de plusieurs cm et des résolutions allant jusqu'à la dizaine de µm. L'excellente résolution temporelle de l'imagerie US permet d'obtenir une multitude d'informations quantitatives sur la dynamique du flux sanguin. Au travers d'approches spectroscopiques, l'imagerie PA fournit quant à elle de précieuses données sur les molécules imagées telles que leurs concentrations, et des indicateurs biologiques d'intérêt tels que la saturation en oxygène peuvent être cartographiés. Les travaux menés au cours de cette thèse consistent en la mise en place d'un dispositif permettant de réaliser simultanément des images 3D avec ces deux modalités, et le développement de nouvelles méthodes de reconstruction et de traitement en imagerie PA. Ces méthodes ont pour but de corriger sur les images PA les artefacts dits de visibilité et ceux liés à l'utilisation de réseaux de capteurs parcimonieux. Dans une première approche, l'analyse statistique de la fluctuation du signal PA liée au flux sanguin permet de corriger ces artefacts et d'obtenir une image proportionnelle à l'énergie absorbée. L'approche est appliquée in vivo sur des embryons de poulet et une mesure quantitative de la saturation en oxygène est obtenue. Dans une seconde approche s'appuyant sur l'apprentissage profond (deep learning), un réseau de neurone a été implémenté, permettant de résoudre ces artefacts à partir d'une unique acquisition une fois le réseau entrainé. L'architecture utilisée permet d'inclure une estimation de l'incertitude de la prédiction, fournissant la localisation des erreurs dites d'hallucination. D'abord démontrée en 2D sur des échantillons tests, l'approche est ensuite appliquée en 3D et de premiers résultats expérimentaux sur l'embryon de poulet sont présentés.

Mots-clés : Imagerie médicale, Photoacoustique, Apprentissage Profond, Echographie, Fluctuation, Oxygénation

Abstract

Vascular imaging is of considerable interest for many health applications, including pathology diagnosis, surgical guidance, medical monitoring and biomedical research. While several modalities are currently used to image blood vessels, bimodal devices combining photoacoustic imaging (PA) and ultrasound imaging (US) is a promising approach. The sensitivity of PA imaging to hemoglobin and the emergence of flow-sensitive methods in US imaging allow accurate mapping of the vascular network within biological tissues at depths of several cm and resolutions down to the tens of µm. The excellent temporal resolution of US imaging provides several quantitative information on blood flow dynamics. Through spectroscopic approaches, PA imaging provides valuable data on the imaged molecules such as their concentrations, and biological indicators of interest such as oxygen saturation can be mapped. The work carried out during this thesis involves the implementation of a device allowing simultaneous acquisition of 3D images with these two modalities, and the development of new reconstruction and processing methods for PA imaging. These methods aim to correct the so-called visibility artifacts and those linked to the use of sparse sensor networks on PA images. In a first approach, the statistical analysis of the fluctuation of the PA signal related to the blood flow allows to correct these artefacts and to obtain an image proportional to the absorbed energy. The approach is applied in vivo on chicken embryos and a quantitative measurement of the oxygen saturation is obtained. In a second approach relying on deep learning, a neural network is used to resolve these artifacts from a single acquisition once trained. The used architecture allows for the estimation of the prediction uncertainty, providing the localization of so-called hallucination errors. Firstly demonstrated in 2D on test samples, the approach is then applied in 3D and the first experimental results on chicken embryos are presented.

Keywords : Medical Imaging, Photoacoustic, Deep learning, Echography, Fluctuation, Oxygenation