Estimation et classification des temps de relaxation multi-exponentiels en IRM. Application aux tissus végétaux

Christian El Hajj

▶ **To cite this version:**

Christian El Hajj. Estimation et classification des temps de relaxation multi-exponentiels en IRM. Application aux tissus végétaux. Traitement du signal et de l'image [eess.SP]. Ecole Centale de Nantes, 2019. Français. NNT : . tel-02561480v1

HAL Id: tel-02561480

https://theses.hal.science/tel-02561480v1

Submitted on 4 May 2020 (v1), last revised 30 Sep 2020 (v2)

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.

UNIVERSITE BRETAGNE LOIRE / MATHSTIC

CENTRALE NANTES

THÈSE DE DOCTORAT DE
CHRISTIAN EL HAJJ

L'ÉCOLE CENTRALE DE NANTES
COMUE UNIVERSITÉ BRETAGNE LOIRE

ECOLE DOCTORALE N° 601
Mathématiques et Sciences et Technologies de l'Information et de la Communication
Spécialité : Traitement du signal et de l'image

Par
Christian EL HAJJ

Estimation et classification des temps de relaxation multi-exponentiels en IRM. Application aux tissus végétaux.

Thèse présentée et soutenue à Nantes, le 16/12/2019
Unité de recherche : LS2N UMR CNRS 6004

Rapporteurs :

Philippe CIUCIU — Directeur de recherche CEA
Ludovic DE ROCHEFORT — Chargé de recherche CNRS

Composition du Jury :

Président :	Prénom Nom	Fonction et établissement d'exercice *(à préciser après la soutenance)*
Examinateurs :	Su RUAN	Professeure des universités, Université de Rouen
	Olivier COMMOWICK	Chargé de recherche INRIA
	Jean-Marie BONNY	Directeur de recherche INRA
Dir. de thèse :	Saïd MOUSSAOUI	Professeur des universités, École Centrale de Nantes
Encadrante de thèse :	Guylaine COLLEWET	Ingénieur de recherche Irstea
Encadrante de thèse :	Maja MUSSE	Chargée de recherche Irstea

REMERCIEMENTS

À mes encadrants :
Saïd MOUSSAOUI, Guylaine COLLEWET et Maja MUSSE,
pour leur partages de connaissances indispensable pour la préparation de cette thèse.

À mes parents :
pour leur conseil ainsi que leur soutien inconditionnel et les motivations qui ont permis l'aboutissement de cette thèse.

À mes amis :
pour les moments de défoulement, bien nécessaire, mais aussi leur soutien moral pendant ces années d'études.

Table des matières

Introduction générale **19**

I.1 Motivations 19

I.2 Objectifs méthodologiques 21

I.2.1 Estimation des paramètres 21

I.2.2 Exploitation des cartographies 22

I.3 Plan du manuscrit 23

I.4 Production scientifique 24

Chapitre II: Modélisation du signal IRM **27**

II.1 Imagerie par résonance magnétique 29

II.1.1 La résonance magnétique nucléaire 29

II.1.2 Séquence d'écho de spin multiple 31

II.1.3 Reconstruction d'images IRM 31

II.2 IRM des tissus végétaux 33

II.2.1 Cellule végétale 34

II.2.2 Séquence d'écho de spin multiple adaptée 36

II.3 Modélisation du signal de décroissance en l'absence du bruit 36

II.3.1 Le modèle mono-exponentiel 37

II.3.2 Les modèles multi-exponentiels 37

II.3.3 Prise en compte de l'hétérogénéité du champ $\vec{B}_1$ 40

II.4 Le bruit de mesure 41

II.4.1 Statistique de la distribution ricienne 42

II.4.2 Le rapport signal sur bruit 43

II.4.3 Estimation des paramètres du bruit 45

II.5 Modèle choisi pour l'étude 45

Chapitre III: Estimation des paramètres à l'échelle du voxel **47**

III.1 Estimation sous l'hypothèse du bruit gaussien 49

III.2 Estimation des paramètres sous l'hypothèse d'un bruit ricien 50

III.2.1 Correction basée sur l'observation 51

III.2.2 Correction basée sur les statistiques riciennes du bruit 52

III.3 Algorithme d'optimisation 53

III.3.1 Algorithme de Levenberg-Marquardt 54

III.3.2 Algorithme de majoration-minimisation 55
III.4 Algorithme d'estimation proposé 56
III.4.1 Analyse de convexité du critère de maximum de vraisemblance ricien 56
III.4.2 Calcul de la majorante 57
III.4.3 Majoration-minimisation approchée 58
III.4.4 Prise en compte des contraintes 58
III.5 Analyse des performances sur des données simulées 59
III.5.1 Objets virtuels 59
III.5.2 Paramètres de l'IRM 61
III.5.3 Simulation des images 61
III.5.4 Initialisation des algorithmes 61
III.5.5 Critère d'évaluation 61
III.5.6 Résultats 62
III.6 Conclusion 64

Chapitre IV: Estimation conjointe des paramètres à l'échelle de l'image 67
IV.1 Notations 69
IV.2 Levenberg-Marquard global 70
IV.2.1 Minimisation du critère global 70
IV.2.2 Stratégie de vectorisation 70
IV.3 État de l'art de la régularisation spatiale des paramètres de relaxation . . . 71
IV.3.1 Régularisation spatiale par patch 72
IV.3.2 Régularisation spatiale par cartographies filtrées 72
IV.3.3 Récapitulatifs des méthodes d'estimation de l'état de l'art 72
IV.4 Algorithme de minimisation conjointe prenant compte la régularité spatiale 73
IV.4.1 Minimisation par majoration du critère de régularisation 74
IV.4.2 Paramètres de régularisation spatiale 74
IV.5 Résultats 78
IV.5.1 Analyse empirique de convergence 78
IV.5.2 Analyse sur le choix des poids de régularisation 79
IV.5.3 Temps de convergence 79
IV.5.4 Résultats de la minimisation des critères régularisés sur l'OV1 . . . 80
IV.5.5 Résultats sur l'OV2 82
IV.5.6 Régularisation l2-l1 sur l'OV1 82
IV.5.7 Résultats sur les données expérimentales 83
IV.6 Conclusion 84

Chapitre V: Exploitation des cartographies Multi-exponentielles 87
V.1 Algorithmes de classification 90

V.1.1 Classification supervisée . . . 90
V.1.2 Algorithme de classification non-supervisée . . . 90
V.2 Paramètres de l'algorithme . . . 93
V.2.1 Espace de descripteurs . . . 93
V.2.2 Choix de nombre de classes . . . 95
V.3 Distribution interclasse . . . 95
V.4 Validation et tests . . . 95
V.4.1 Classification sur l'OV1 . . . 96
V.4.2 Données expérimentales . . . 97
V.4.3 Résultats sur d'autres fruits . . . 101
V.5 Conclusion . . . 101

Conclusion générale **105**
VI.1 Contributions par chapitre . . . 105
VI.2 Perspectives . . . 107
VI.2.1 Initialisation des paramètres . . . 107
VI.2.2 Choix des paramètres de régularisation . . . 107
VI.2.3 Détection automatique de nombre de composantes . . . 108

Chapitre A: Calcul des Jacobiens **111**
A.1 Calcul du Jacobien dans le cas des MCCMG . . . 111
A.2 Calcul du Jacobien dans le cas des MCCER . . . 112
A.3 Calcul de la dérivée dans le cas d'un modèle multi-exponentielle . . . 113

Liste d'abréviations

ES-M	Écho de Spin Multiple
GEC	Generalised Echo Curve
IRM	Imagerie par Résonance Magnétique Nucléaire
LM	Levenberg-Marquardt
LMG	Levenberg-Marquardt Global
MCNL	Moindres Carrés Non Linéaires
MCCER	Moindres Carrés Corrigé par l'Espérance Ricienne
MCCER	Moindres Carrés Corrigé par l'Espérance Ricienne Pénalisé
MM	Majoration-Minimisation
MMG	Modèle de Mélange Gaussien
MV	Maximum de Vraisemblance
MVR	Maximum de Vraisemblance Ricien
MVRP	Maximum de vraisemblance Ricien Pénalisé
REQMN	Racine carré de l'Erreur Quadratique Moyenne Normalisée
RMN	Résonance Magnétique Nucléaire
RSB	Rapport Signal à Bruit

Liste des figures

II.1 a- Protons en l'absence du champ magnétique. b- Réorientation des protons suivant la direction du champ magnétique homogène extérieur appliqué et le phénomène de précession. 29

II.2 Phénomène de relaxation des spins une fois le champ magnétique secondaire arrêté. Ici l'angle de bascule Γ est égale à $\frac{\pi}{2}$. 31

II.3 Le phénomène de déphasage et le retour en phase des protons. 32

II.4 Séquence d'écho de spin avec les différents gradients utilisés pour la reconstruction d'une image IRM. 33

II.5 Structure de la cellule végétale 34

II.6 Le type et le grand nombre de cellules contenu dans les images IRM 35

II.7 Distribution des temps de relaxation mesurée sur un ROI de 1100 voxels sur une Tomate. 36

II.8 Exemple d'un spectre mono-exponentiel. 37

II.9 Exemple d'un spectre multi-exponentiel avec un nombre de composantes $N_c = 3$. 39

II.10 Distribution ricienne avec $\sigma = 7$ pour diverses valeurs de $(\frac{s_t}{\sigma})$. 42

II.11 Effet du bruit ricien sur les données. Le signal mesuré est donné en rouge et le signal non bruité en bleu pour un $RSB = 230$. 44

III.1 Décroissance du critère en utilisant l'approche MM. 55

III.2 Les régions d'intérêts choisis pour la simulation de l'OV1 60

III.3 Évolution du critère de MVR dans un voxel au cours des itérations MM (points en mauves) et des itérations LM interne (points en jaune) pour un nombre d'itérations internes maximal ℓ_{max}. 62

III.4 Variation de REQMN pour différents RSB sur l'OV1 simulé en minimisant les critères MC, MCCMG, MCCER, et MVR 63

III.5 Résultats de l'estimation sur l'OV2. La première colonne représente les cartographies des amplitudes A_0 de référence utilisée pour la simulation. La deuxième, la troisième et la quatrième colonne représentent respectivement les cartographies des A_0 reconstruites en utilisant les méthodes de MCCER, MVR, MCCMG. 65

III.6 Résultats de l'estimation sur l'OV2. La première colonne représente les cartographies des temps de relaxation T_2 de référence utilisée pour la simulation. La deuxième, la troisième et la quatrième colonne représentent respectivement les cartographies des T_2 reconstruites en utilisant les méthodes de MCCER, MVR, MCCMG. 65

IV.1 Évolution du critère de MVRP au cours des itérations MM (points en orange et en mauves) et des itérations LMG internes (points en bleu et en jaune). 76

IV.2 REQMN pour les 4 approches; MCCERR, MCCERP, MVR, MVRP pour différents RSB. 76

IV.3 Moyenne et écart-type des paramètres estimés dans les différentes régions du fantôme pour un RSB=150. 77

IV.4 Comparaison entre l'utilisation d'un poids de régularisation unique, β_f, pour tous les paramètres (courbe en bleu) et un poids de régularisation différent par paramètres (courbe en rouge). 79

IV.5 Cartographies d'intensités reconstruites pour les composantes à décroissance rapide (première ligne), à décroissance moyenne (deuxième ligne) et à décroissance lente (troisième ligne). Les cartographies de référence (première colonne) sont comparées aux cartographies estimées par MCCER (deuxième colonne), MCCERP (troisième colonne) et MVRP (quatrième colonne). . . 81

IV.6 Cartographies de temps de relaxation reconstruites pour les composantes à décroissance rapide (première ligne), à décroissance moyenne (deuxième ligne) et à décroissance lente (troisième ligne). Les cartographies de référence (première colonne) sont comparées aux cartographies estimées par MCCER (deuxième colonne), MCCERP (troisième colonne) et MVRP (quatrième colonne). 81

IV.7 Résultats de l'algorithme MVRP sur l'OV2 82

IV.8 Régularisation l2-l1 en utilisant la méthode MVRP avec une valeur de $\delta = 100$ 83

IV.9 Résultats de l'algorithme MVRP sur l'IE1. Comparaison entre régularisation sur 1 acquisition et 32 acquisitions 84

IV.10 Résultats de l'algorithme MVRP sur l'OV2. Comparaison entre la reconstruction sans régularisation (colonne 1), en utilisant une régularisation l2 (colonne 2), avec régularisation l2-l1 (colonne 3). 85

V.1 Exemple de partitionnement en utilisant l'algorithme K-moyennes. 91

V.2 Dendrogramme tracé à partir de 11 points dans un espace de descripteurs définis par $[A_{0_3}, T_{2_3}]$ 92

V.3 Variation spatiale de A_0 au sein d'un tissus supposé homogène (exemple du péricarpe équatorial) due aux imperfections de l'antenne RF de réception. . 94

V.4 Carte de répartition spatiale des 5 classes identifiées par K-moyennes et distributions conjointes des paramètres (T_2, A_0) des trois composantes de chaque classe sur l'OV1. Les points gris représentent les valeurs utilisées pour la simulation et les croix représentent les centroïdes des classes estimées par K-moyennes. 96
V.5 Résultat de la méthode en coude appliqué à l'IE1; le nombre de classes en abscisse et en ordonnée la valeur de la variation cumulée d'inertie ΔW en pourcentage de la valeur maximale, la valeur de seuil est fixée à 85%. . . . 97
V.6 Dendrogramme avec une connectivité 'Full-linkage' appliqué à l'IE1. Les premiers points du dendrogramme représentent un ensemble de points pour éviter l'encombrement sur la figure. 97
V.7 Carte de répartition spatiale des 3 classes identifiées par K-moyennes sur l'IE1 et distributions conjointes des paramètres (T_2, A_0) des trois composantes de chaque classe et les spectres de temps de relaxation dans chaque classe. Les croix représentent les centroïdes des classes estimées. 98
V.8 Carte de répartition spatiale des 3 classes identifiées par MMG sur l'IE1 et distributions conjointes des paramètres (T_2, A_0) des trois composantes de chaque classe et les spectres de temps de relaxation dans chaque classe. Les croix représentent les centroïdes des classes estimées. 99
V.9 Carte de répartition spatiale des 3 classes identifiées par la classification hiérarchique sur l'IE1, les distributions conjointes des paramètres (T_2, A_0) des trois composantes de chaque classe et les spectres de temps de relaxation dans chaque classe. Les croix représentent les centroïdes des classes. 99
V.10 Annotation des 5 principaux tissus végétaux de la tomate en utilisant l'outil graphique 100
V.11 Tomate partitionnée à partir de l'algorithme KNN en utilisant des voxels étiquetés. 100
V.12 Carte de répartition spatiale des 3 classes identifiées par K-moyennes sur une carotte et distributions conjointes des paramètres (T_2, A_0) des trois composantes de chaque classe et les spectres de temps de relaxation dans chaque classe. Les croix représentent les centroïdes des classes. 102
V.13 Carte de répartition spatiale des 3 classes identifiées par K-moyennes sur une pomme et distributions conjointes des paramètres (T_2, A_0) des trois composantes de chaque classe et les spectres de temps de relaxation dans chaque classe. Les croix représentent les centroïdes des classes. 102

Liste des tables

III.1 valeurs des paramètres de relaxation dans chaque région de l'OV1. 60
III.2 Nombre d'itérations et temps de calcul de la méthode MVR 63

IV.1 Tableau résumant les différents articles de la littérature qui traitent les problématiques adressées dans cette thèse 73

A.1 Tableau résumant les différents critères à minimiser 112

Combien ce fruit est gros,
et sa tige menue.
A quoi songeait dit-il,
l'auteur de tout cela ?

Jean de la fontaine

Introduction générale

L'imagerie par résonance magnétique nucléaire (IRM) est une technique d'imagerie non invasive développée pour le diagnostic médical. Elle est basée sur la résonance magnétique nucléaire (RMN). Le principe de l'IRM permet d'étudier tout objet, corps ou tissu contenant des atomes d'hydrogène[1]. C'est la modalité de référence pour la détection et l'analyse de différentes maladies. Par exemple, l'IRM est utilisée pour le diagnostic de différentes pathologies comme le cancer [1, 2, 3], la sclérose en plaques [4], les maladies cardiovasculaires [5], ou pour l'étude du fonctionnement du cerveau [6]. Cependant cette technique d'imagerie peut également être mise à profit dans d'autres domaines d'application : en génie civil, l'IRM est utilisée pour l'estimation de la teneur en eau dans des matériaux comme l'argile, le sable et le béton [7] afin de contrôler la qualité des produits. Dans le domaine agroalimentaire, l'IRM est utilisée pour la caractérisation de la qualité des aliments ou des processus de leur transformation comme, par exemple, le suivi des phases de cuisson d'aliments [8], de maturation du fruit [9] ou de caractérisation de la qualité de la viande (taux de gras dans la viande du porc ou de saumon) [10].

Le sujet de cette thèse s'insère dans ce dernier champ d'applications et en particulier sur l'étude des fruits charnus. Plus précisément, il se positionne dans un contexte multidisciplinaire impliquant le développement d'outils de traitement du signal et de l'image et l'analyse de données IRM appliquées à la caractérisation des produits végétaux. Les travaux de la thèse se sont déroulés dans le cadre d'une collaboration entre l'équipe IRM-Food de Irstea Rennes, spécialisée en étude de produits et procédés agroalimentaires par IRM, et l'équipe SIMS (Signal, Image et Son), qui développe une activité de recherche sur dans le domaine du traitement du signal et de l'image au sein du Laboratoire des Sciences du Numérique de Nantes (LS2N).

I.1 Motivations

Un des atouts de l'IRM est de proposer un grand nombre de contrastes différents qui permettent d'explorer de nombreuses propriétés tissulaires. Ces contrastes peuvent être

1. Nous nous intéressons dans ces travaux à l'imagerie du noyau d'hydrogène. C'est l'imagerie la plus largement utilisée. Cependant, il est à noter que cette technique permet d'explorer d'autres atomes qui possèdent un nombre impair de nucléons (protons + neutrons) tels que le phosphore 31, le carbone 13, le sodium 23, le potassium 39 ou le fluor 19.

utilisés pour faciliter la visualisation des différents tissus mais également pour estimer des paramètres physiques d'intérêt tel que le flux, la diffusion, le pourcentage de lipides, la porosité ou les temps de relaxation : T_1, T_2^* et T_2. Ces temps de relaxation sont caractéristiques de la technique RMN. Les temps de relaxation transversale T_2 des molécules d'eau, qui nous intéressent dans ce travail, sont mesurés à partir d'un ensemble de données temporelles correspondant à un signal de décroissance exponentielle. Les T_2 dépendent de la mobilité des molécules d'eau et de leurs interactions avec les macromolécules, et permettent donc d'accéder à des informations sur des paramètres physico-chimiques des tissus (structure, composition, transferts d'eau) [4, 1].

Un des intérêts particuliers de la relaxation RMN T_2 pour l'étude des tissus se situe dans le cas où l'eau est compartimentée car elle permet alors d'accéder aux informations sur l'état et la répartition de l'eau dans ces compartiments. En effet, dans le cas des tissus végétaux, les paramètres de relaxation des protons de l'eau se distinguent pour les différents compartiments cellulaires, du fait des différences de l'environnement physique et chimique des molécules d'eau. Ainsi, le signal de relaxation possède un caractère multi-exponentiel et les valeurs des différents T_2 et de leurs intensités associées A_0 permettent d'obtenir des informations sur chacun de ces compartiments. De plus, comme les tissus végétaux sont vivants, et donc possèdent une dynamique d'évolution, ces mesures permettent d'avoir des informations sur les transferts d'eau au niveau cellulaire et tissulaire et donc de mieux comprendre le fonctionnement hydrique.

Dans les applications au sein du domaine du végétal, plusieurs travaux de caractérisation de la structure des tissus des fruits par IRM se sont focalisés sur des cartographies de T_2 mono-exponentiels, c'est-à-dire avec un seul T_2 estimé par voxel [9, 10]. Cependant, il est nécessaire de développer une approche IRM de mesure des temps de relaxation transversale multi-exponentielle transversale afin de pouvoir accéder à des informations sur les fractions d'eau dans les principaux compartiments cellulaires. Cette mesure n'est pas simple car le rapport signal sur bruit (RSB) est relativement faible à l'échelle du voxel ce qui rend l'estimation très instable. C'est pourquoi ces études sont généralement réalisées par relaxométrie RMN qui nécessite l'échantillonnage du produit par prélèvement d'une petite quantité insérable dans l'appareil (de l'ordre du cm^3). Les T_2 multi-exponentiels ont toutefois été estimés par IRM mais seulement à l'échelle de régions d'intérêt permettant d'augmenter le RSB [11]. Ceci n'est pas satisfaisant car cela réduit l'information spatiale et, de plus, la définition de ces régions d'intérêt ne peut être qu'approximative. Un autre problème très important est la nature du bruit dans les images IRM qui introduit un biais dans les paramètres de relaxation estimés par des algorithmes classiques telles que les méthodes de moindres carrés non-linéaires. En effet, le bruit de mesure en IRM n'est pas gaussien centré, mais ricien [12]. Cela est dû au fait que les données traitées en relaxométrie sont celles du module des images complexes qui, elles, sont altérées par un bruit gaussien. C'est pourquoi, afin de pouvoir descendre à l'échelle du voxel et obtenir les paramètres de relaxation de façon plus précise et robuste, **il est nécessaire de développer des**

méthodes avancées d'estimation des paramètres du modèle multi-exponentiel à l'échelle de l'image entière et sous l'hypothèse d'un bruit ricien.

Il est à noter que cette problématique est générale à l'étude des tissus par IRM. En effet, on retrouve dans le domaine médical un intérêt pour ce modèle multi-exponentiel de relaxation transversale dans différentes applications. Par exemple, dans l'estimation du taux de la myéline dans le cerveau pour le diagnostic des maladies chroniques telles que la sclérose en plaques [4, 13, 14, 15]. Dans ce cas, l'eau intra et extra cellulaire ont des temps de relaxation plus longs que l'eau contenue dans la couche de myéline [16]. La problématique multi-exponentielle se retrouve également dans les études du cartilage ([17]) où une analyse multi-exponentielle permet d'identifier et de caractériser les différents pools d'eau dans les tissus normaux et ceux pathologiques. Enfin nous pouvons citer l'intérêt de la relaxation multi-exponentielle pour la caractérisation des muscles qui contiennent de l'eau extra-cellulaire et de l'eau en interaction avec des macro-molécules [18].

Par ailleurs, en aval de l'estimation des cartographies des valeurs des T_2 et de leurs amplitudes associées A_0 dans le modèle multi-exponentiel à l'échelle de l'image entière se pose le problème de l'exploitation de ces résultats qui représentent une quantité importante de données. En effet, cette question est moins cruciale dans le cas d'un modèle mono-exponentiel où une seule cartographie des valeurs des T_2 ou A_0 sera exploitée pour l'analyse. Cependant, dans un modèle multi-exponentiel, **il est nécessaire de proposer une démarche d'analyse permettant de faciliter l'analyse des cartographies résultantes en vue de la discrimination des différents tissus composants les objets analysés**.

I.2 Objectifs méthodologiques

L'estimation des cartographies des paramètres du modèle de relaxation transversale multi-exponentielle, T_2 et A_0, à l'échelle de l'image entière, présente plusieurs défis méthodologiques concernant la modélisation du signal, le développement d'algorithmes d'estimation paramétrique et l'exploitation des cartographies estimées pour l'analyse du contenu des images. Ces trois défis ont été à l'origine des axes développés dans cette thèse et dont les objectifs sont présentés ci-après.

I.2.1 Estimation des paramètres

L'estimation des temps de relaxation multi-exponentiels en IRM nécessite la résolution d'un problème inverse mal posé. En fait, la principale source de difficulté est le rapport signal sur bruit qui peut être faible, notamment lorsqu'on réduit la taille du voxel pour accéder à des structures fines, ou lorsqu'on veut diminuer le temps d'acquisition pour analyser un grand nombre d'échantillons. De plus, l'analyse des images de module induit des biais d'estimation dans le cas de l'utilisation d'un critère du type moindres

carrés. Ainsi, le premier objectif de ce travail de thèse est de proposer une méthode d'estimation des paramètres dans chaque voxel de l'image IRM et sous l'hypothèse du bruit ricien. De plus, la contrainte de non-négativité des paramètres devra être intégrée dans les algorithmes d'estimation développés. **La première contribution de ce travail de thèse consiste en le développement d'algorithmes d'estimation des paramètres du modèle multi-exponentiel intégrant l'hypothèse de distribution ricienne du bruit**, en adoptant une approche fondée sur l'estimation par maximum de vraisemblance et l'optimisation en couplant des techniques de majoration-minimisation avec les méthodes de moindres carrés non-linéaires.

Par ailleurs, étant donné le caractère mal-posé du problème, l'estimation des paramètres du modèle multi-exponentiel, même en prenant en considération la distribution ricienne du bruit, n'est pas efficacement résolu, notamment dans région de faible RSB dans l'image. Pour cela, une information supplémentaire doit être exploitée, comme, par exemple, la régularité spatiale des paramètres estimés. En effet, dans la majorité des cas, on trouve des régions cohérentes et la régularisation peut permettre la réduction de l'effet du bruit et aboutir à une meilleure estimation. **La deuxième contribution de la thèse réside en le développement d'un algorithme d'estimation paramétrique capable de la prise en considération la dimension image, de prendre en compte la nature ricienne du bruit et de la régularité spatiale des paramètres** afin de faire face au faible RSB.

I.2.2 Exploitation des cartographies

L'estimation des temps de relaxation multi-exponentiels fournit un nombre d'informations important par voxel (6 par exemple dans le cas d'un modèle tri-exponentiel) ce qui aura pour conséquence de rendre l'interprétation des paramètres de relaxation moins évidente que dans le cas d'un modèle mono-exponentiel. Il est nécessaire de proposer des outils d'analyse automatique afin de mettre en évidence, par exemple, des similarités de certains tissus au sein de l'image analysée et d'identifier les paramètres caractéristiques de chaque tissu. Ainsi, **la troisième contribution de cette thèse est la proposition d'une approche fondée sur les algorithmes de classification non-supervisée** afin de segmenter les différents tissus et d'identifier les distributions des temps de relaxation au sein de chaque classe. Pour y parvenir, les paramètres du modèle multi-explonentiel sont utilisés pour définir l'espace de représentation pour l'analyse multi-classes.

Les développements algorithmiques réalisés dans le cadre de cette thèse ont toujours fait l'objet de validation sur des données synthétiques réalistes mais aussi d'applications à des données IRM acquises sur des fruits (machine Magnetom, Avanto, Siemens, Erlangen, Germany de $1.5T$. Cette dernière analyse a permis d'obtenir des distributions des temps de relaxation de certains tissus assez similaires à celles obtenues dans d'autres analyse par RMN, ce qui est révélateur de la pertinence de l'approche méthodologique proposée dans

cette thèse.

I.3 Plan du manuscrit

Dans le **Chapitre 2**, la problématique de modélisation du signal sera présentée en expliquant, dans un premier temps, la technique IRM, en se focalisant sur les notions nécessaires à la compréhension de ce travail et, dans un second temps, son application à la caractérisation des tissus végétaux. Le modèle ricien du bruit sera présenté dans cette section et les différentes notions aboutissant à un tel modèle seront détaillées. Ensuite, l'état de l'art sera établi sur les différents modèles à décroissance multi-exponentielle proposés dans la littérature et cette partie se terminera par une discussion sur le choix du modèle le plus adéquat à l'étude de tissus végétaux.

Dans le **chapitre 3**, les méthodes d'estimation des paramètres du modèle mutli-exponentiel seront détaillées dans le contexte d'une approche d'estimation par voxel. Les différentes méthodes d'estimation basées sur une minimisation d'un critère sous une hypothèse ricienne du bruit seront présentées. Un algorithme de majoration-minimisation (MM) pour la minimisation d'un critère de maximum de vraisemblance sous l'hypothèse d'une distribution ricienne du bruit. À la fin de ce chapitre, des résultats seront présentés pour comparer les différentes approches et mettre en avant l'importance de prendre en compte la nature ricienne du bruit.

Ensuite, dans le **chapitre 4**, nous proposons une approche de minimisation conjointe sur l'ensemble de l'image basée sur un algorithme de Levenberg-Marquardt adapté. Différentes approches de régularisation spatiale seront présentées et nous proposerons une méthode de régularisation spatiale et, afin de résoudre le problème en grandes dimensions, une approche de séparation par MM sera proposée dans le cas de l'estimation conjointe. Ce chapitre se conclut par une présentation des résultats de l'estimation conjointe en utilisant un algorithme régularisé comparé aux méthodes non régularisées.

Dans le **chapitre 5** nous présenterons une approche d'exploitation des paramètres estimés basée sur des algorithmes de classification pour le partitionnement des cartographies en zones homogènes. D'abord, un rappel sur différentes techniques de classification supervisée (K-NN) et non-supervisée sera fourni (K-moyennes, modèle de mélange gaussien et classification hiérarchique). Ensuite, nous présenterons différentes méthodes basées sur la nature des cartographies pour la reconstruction d'un espace de descripteurs appropriés. Une étude sera proposée sur le choix de nombre de classes. Les résultats de la classification seront présentés sur des données synthétiques et réelles.

Enfin, nous présentons une **conclusion générale** sur les différentes contributions de

ce travail de thèse. Nous conclurons par une discussion sur des perspectives applicatives des méthodes proposées et sur des problématiques non traitées durant ce travail.

I.4 Production scientifique

Les travaux de cette thèse ont fait objet de plusieurs communications scientifiques au niveau international. Un article écrit sur la partie d'estimation des paramètres du modèle multi-exponentiel à l'échelle de l'image entière et sous l'hypothèse d'un bruit ricien a été soumise dans un journal international de traitement de signal [a]. Les méthodes d'estimation de paramètres de relaxation ont été présentées dans deux congrès internationaux [b,f]. Les techniques proposées pour l'exploitation des cartographies ont été présentées dans un congrès francophone [d] et deux congrès internationaux [c,e].
Durant cette thèse j'ai co-encadré le travail de deux stagiaires. Le premier travail avait pour but de développé une interface graphique sous Matlab R2016a pour la visualisation des données IRM, l'estimation et la visualisation des paramètres multi-exponentiels et la classification des cartographies estimées [g]. Le but du deuxième travail que j'ai encadré était de tester différentes méthodes de classification et de développer une méthode de choix de nombre de classes dans les algorithmes de classification [h].

[a] El Hajj C., Moussaoui S., Collewet G., & Musse M. *"Multi-exponential Transverse Relaxation Times Estimation from Magnetic Resonance Images under Rician Noise and Spatial Regularization."* Journal d'IEEE Transactions on Image Processing *soumis le 10/07/2019.*

[b] El Hajj C., Moussaoui S., Collewet G., & Musse M. *"Spatially Regularized Multi-Exponential Transverse Relaxation Times Estimation from Magnitude Magnetic Resonance Images Under Rician Noise."* dans IEEE International Conference on Image Processing (ICIP) (pp. 1143-1147). 2019.

[c] El Hajj C., Moussaoui S., Collewet G., & Musse M. *"Multi-Exponential Relaxation Times Maps Reconstruction and Unsupervised Classification in Magnitude Magnetic Resonance Imaging."* dans 27th European Signal Processing Conference (Eusipco) 2019.

[d] El Hajj C., Moussaoui S., Collewet G., & Musse M. *"Reconstruction et classification des temps de relaxation IRM multi-exponentielle."* actes du colloque francophone Gretsi sur le Traitement du Signal et des Images, Lille, 2019

[e] El Hajj C., Moussaoui S., Collewet G., & Musse M. *"Fruit Tissues Classification From Multi-Exponential T2 Maps."* dans 14th International Conference on the Applications of Magnetic Resonance in Food Science MRFood. 2018

[f] El Hajj C., Moussaoui S., Collewet G., & Musse M. *"Spatially Regularized Multi-exponential Transverse Relaxation Times Estimation from Magnitude MRI under*

Rician Noise." dans International Society of Magnetic Resonance in Medicine ISMRM, Paris, 2018

[g] YVER C. *"Classification automatique de cartographies de temps de relaxation mesurées par Imagerie par Résonance Magnétique"*, Rapport de stage EII à l'Institut Nationale Des Sciences Appliquées Rennes

[h] TOM R. *"Study of Unsupervised Clustering Methods for Vegetal Tissues Characterization Using Magnetic Resonance Images."*, Rapport de stage du Master Commande et Robotique Master in Control and Robotics à l'École Centrale de Nantes.

II

Modélisation du signal IRM

En relaxométrie IRM, la première étape consiste à mesurer des signaux émis par l'ensemble des protons dans chaque voxel de l'image. La deuxième étape concerne la représentation et la modélisation de ces différents signaux afin d'aboutir à des paramètres qui décrivent en mieux chaque signal mesuré de l'image IRM. En soi, la modélisation nécessite une compréhension de la méthode d'acquisition avec laquelle le signal a été acquis. Il est aussi nécessaire de connaître les différents phénomènes physiques et biologiques qui peuvent affecter ces mesures. Dans ce chapitre, la technique d'imagerie IRM et son importance pour les études agroalimentaires et médicales seront présentées. Ensuite, les différents modèles qui prennent en considération les aspects de mesure seront présentés. Enfin, nous présentons les statistiques du bruit de mesure affectant les données IRM.

II.1 Imagerie par résonance magnétique

L'IRM se base sur le phénomène de résonance magnétique nucléaire découvert par Edward Mills Purcell [19] et Felix Bloch[20] en fin des années 1940. Un système de codage approprié permet de localiser le signal et reconstruire ainsi des images.

II.1.1 La résonance magnétique nucléaire

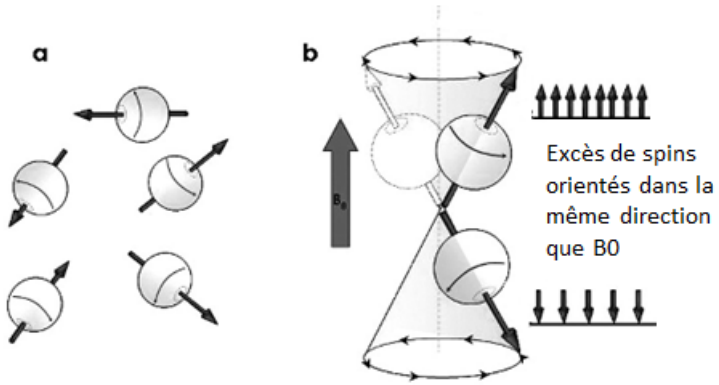

Figure II.1 – a- Protons en l'absence du champ magnétique. b- Réorientation des protons suivant la direction du champ magnétique homogène extérieur appliqué et le phénomène de précession.

La technique RMN est basée sur les propriétés magnétiques de l'hydrogène. Le noyau de cet atome est animé d'un mouvement de rotation et peut être considéré comme un aimant, caractérisé par un moment magnétique, proportionnel au moment cinétique appelé spin. Par simplification, et par abus de langage, on se réfèrera ci-après au noyau en utilisant le terme spin. Dans le corps humain, et en particulier dans les tissus mous, les atomes d'hydrogène se retrouvent en abondance dans les molécules d'eau et de lipide. C'est également le cas dans les tissus végétaux. Ainsi la RMN peut être utilisée pour caractériser ces milieux.

II.1.1.1 Création d'une aimantation macroscopique

En l'absence du champ magnétique, les spins ont une orientation aléatoire et la somme vectorielle des moments magnétiques est nulle (voir figure II.1.a). En présence d'un champ magnétique statique, B0, les spins s'orientent de façon parallèle ou antiparallèle au champ. Un excès de spin s'oriente dans la même direction de B0 ce qui a pour conséquence de créer une magnétisation macroscopique non nulle. Les spins ainsi orientés sont animés d'un mouvement de précession autour de l'axe parallèle à B0 (voir figure II.1.b). Le mouvement de précession se produit à une fréquence donnée par la formule de Larmor et connue sous le nom de fréquence de Larmor :

$$\omega_0 = \gamma \left\| \vec{B_0} \right\|, \tag{II.1}$$

avec $\gamma = 2.675 \cdot 10^8$ rad $\cdot$ s^{-1} $\cdot$ T^{-1} le rapport gyromagnétique de l'hydrogène. Néanmoins, cette magnétisation macroscopique ne possède aucune composante dans le plan perpendiculaire à $\vec{B_0}$. Cela est dû à un déphasage arbitraire dans la précession. De plus, cette magnétisation a une amplitude négligeable par rapport au champ $\vec{B_0}$ et ainsi elle n'est pas directement détectable.

II.1.1.2 Phénomène d'excitation et de relaxation

Le but de cette étape est de donner aux protons une énergie suffisante pour faire basculer leur axe d'aimantation depuis la direction de $\vec{B_0}$ vers une autre direction où on pourra observer leur aimantation.

Afin d'atteindre une résonance magnétique, cette énergie sera fournie par un champ magnétique tournant à une vitesse radio-fréquence (RF) égale à la fréquence de Larmor. Ce champ tournant $\vec{B_1}$ (appelé également champ RF) sera appliqué pendant une courte durée, afin de dévier l'axe de précession d'un angle Γ par rapport à l'axe du champ $\vec{B_0}$. Après l'arrêt du champ $\vec{B_1}$, les protons se réorientent vers l'axe initial de précession $\vec{B_0}$ suivant une trajectoire spirale (voir figure II.2). En étudiant ce phénomène de réorientation appelé relaxation, il est possible de caractériser le tissu étudié. En effet, le temps que les spins mettent à se réorienter dépend du tissu étudié.

II.1.1.3 Temps de relaxation

Une décomposition est ainsi possible du champ magnétique macroscopique résultant suivant deux composantes :

- L'axe longitudinal Z qui représente la direction du vecteur $\vec{B_0}$;
- Le plan transversal XY orthogonal à $\vec{B_0}$.

Deux temps de relaxation sont à distinguer, le temps de relaxation transversale T_2 et le temps de relaxation longitudinale T_1. Ainsi on définit la constante de temps T_1 comme étant la valeur qui représente le retour de la magnétisation longitudinale à la valeur

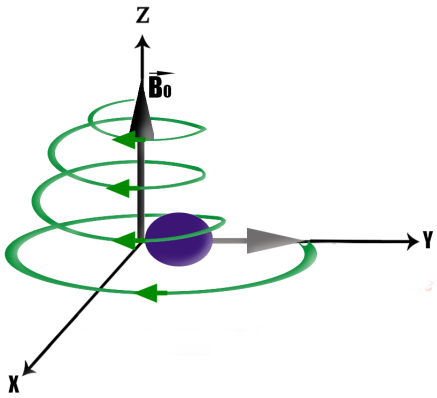

Figure II.2 – Phénomène de relaxation des spins une fois le champ magnétique secondaire arrêté. Ici l'angle de bascule Γ est égale à $\frac{\pi}{2}$.

d'équilibre (63 % de la valeur maximale) suivant l'axe Z, tandis que la constante de temps T_2 représente la disparition de la magnétisation transversale (suivant la projection sur le plan XY), cette disparition est liée au déphasage entre les différents axes d'aimantation.

La mesure des temps de relaxation est connue sous le nom de relaxométrie. En relaxométrie, on mesure, par échantillonnage temporel, la courbe de décroissance (dans le cas de T_2) ou de croissance (dans le cas de T_1).

II.1.2 Séquence d'écho de spin multiple

Théoriquement, le signal mesuré en relaxométrie T_2 traduit la perte de la cohérence de phase de l'ensemble des spins dans le plan transverse. Cela est vrai sous l'hypothèse que la fréquence de précession est constante dans l'espace de mesure. Or, le champ magnétique $\vec{B_0}$ n'est pas parfaitement homogène. Par conséquent, un déphasage résultant d'une différence de fréquence de précession se présente entre les différents spins et le signal mesuré décroit avec un temps de relaxation T_2^* plus court que T_2. Afin de mesurer le signal de décroissance en T_2, il est nécessaire d'appliquer une impulsion de refocalisation de 180°. Cette impulsion permet d'inverser l'orientation des spins dans le plan XY et de ce fait, permet aux protons de redevenir en phase. Elle est appliquée à un temps $TE/2$ et les protons se remettent en phase au temps TE et c'est à ce moment que la mesure est faite (voir figure II.3). T_E est le "temps d'écho". Afin de pouvoir estimer le temps T_2, il est nécessaire d'acquérir une succession d'échos pendant la relaxation. Afin de mesurer plusieurs échos, il est possible d'utiliser une série d'impulsion de refocalisation, cette technique de mesure est connue sous le nom d'écho de spin multiple (ES-M).

II.1.3 Reconstruction d'images IRM

Le phénomène de RMN, discuté dans la section II.1.1, nécessite l'extraction d'un échantillon du corps étudié pour effectuer les mesures. L'imagerie par résonance magnétique est l'extension de ce phénomène permettant d'aboutir à une reconstruction spatiale de

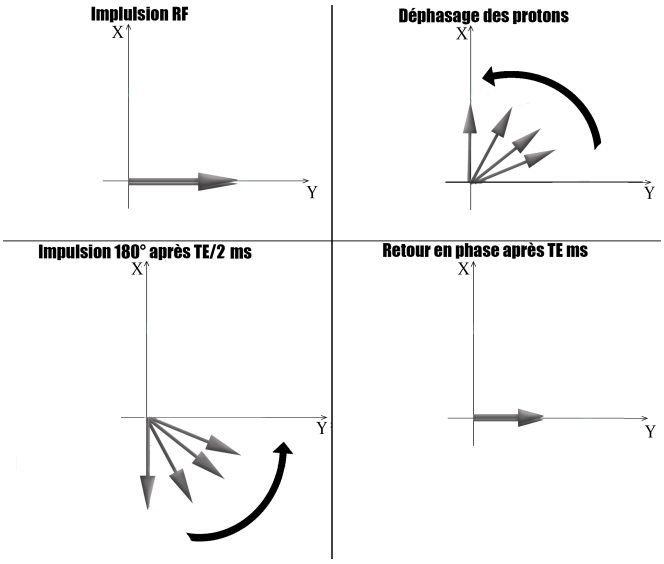

Figure II.3 – Le phénomène de déphasage et le retour en phase des protons.

plusieurs signaux de relaxation en introduisant le corps entier dans la machine. Dans cette section nous présentons le principe de reconstruction d'une image en 2D.
Dans le but de coder la position des différents signaux provenant de différentes composantes spatiales, deux étapes sont nécessaires : la sélection d'une coupe et la localisation spatiale des signaux RMN émis par les protons dans cette coupe.

II.1.3.1 Sélection de coupe

En plus du champ magnétique $\vec{B_0}$, dans l'étape de sélection de coupe, un gradient de champ magnétique est appliqué en même temps que l'impulsion RF. Ainsi, différentes parties du corps étudié reçoivent un champ magnétique de différentes intensités dépendant de leur position suivant l'axe Z (voir figure II.4). Ensuite, l'impulsion RF sera calibrée à la fréquence de Larmor correspondante à la coupe à activer:

$$\omega_{sc} = \gamma \left\| \vec{B_0} + G_{z_{sc}} \right\|, \tag{II.2}$$

$G_{z_{sc}}$ correspond à la valeur du gradient du champ dans la coupe et ω_{sc} correspond à la fréquence de Larmor qui permet de faire entrer en résonance les spins de cette coupe et eux seuls.

II.1.3.2 Localisation spatiale du signal

Une fois la coupe sélectionnée et les spins de cette coupe excités il reste à récupérer l'emplacement spatial des différentes zones de la coupe contribuant à la mesure. Pour cela, deux gradients de champs sont utilisés (voir figure II.4). Le premier est le gradient de codage en fréquence, G_ω. Ce gradient est activé durant la mesure et comme la fréquence dépend de l'intensité du champ il permet de changer la fréquence de précession des protons

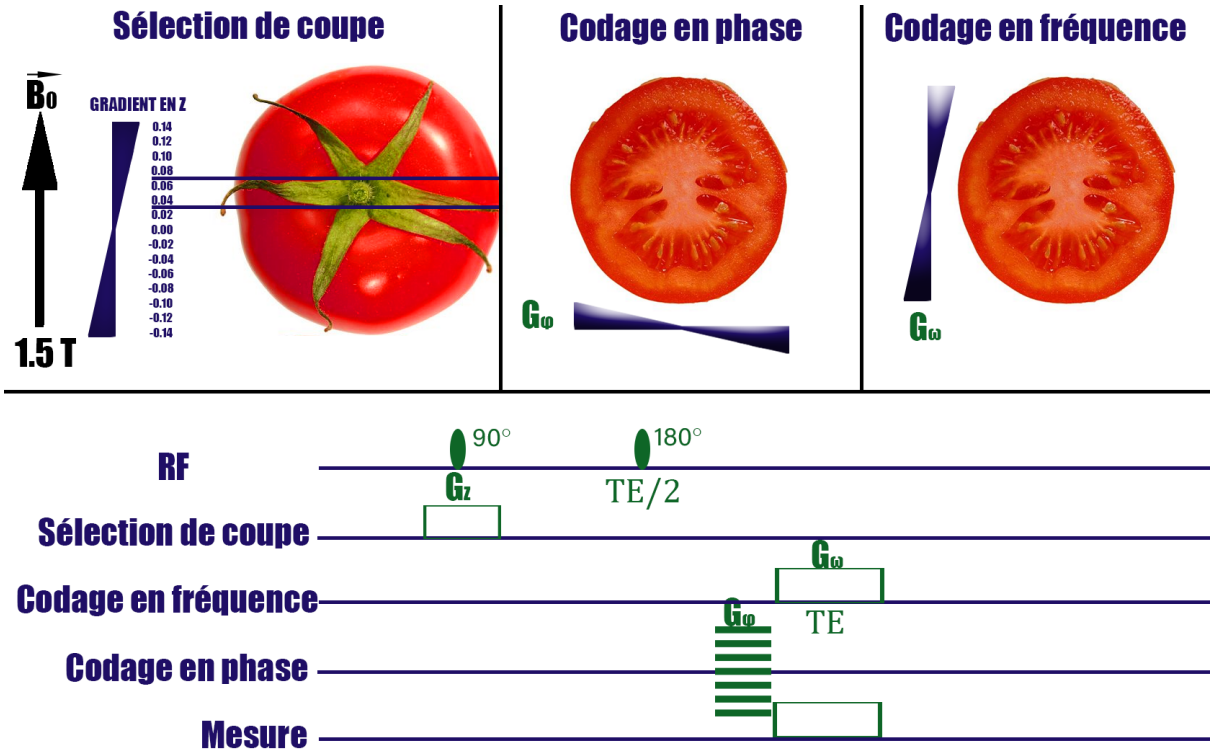

Figure II.4 – Séquence d'écho de spin avec les différents gradients utilisés pour la reconstruction d'une image IRM.

suivant l'axe sur lequel il est appliqué, soit X. Ainsi, le signal mesuré est une somme de signaux avec des fréquences de précession différentes. En appliquant une transformée de Fourier inverse, les signaux sont décomposés et chaque fréquence dans le spectre obtenu correspond à une position suivant l'axe X et son amplitude correspond à la somme suivant Y des différents signaux correspondants à cette position.
Pour le codage de la position spatiale suivant le deuxième axe de mesure, un gradient de champ G_ϕ est aussi utilisé. Par contre ce champ est activé avant la mesure et il permet ainsi de déphaser les protons suivant l'autre axe, soit Y. La mesure sera répétée plusieurs fois, mais à chaque répétition l'intensité du gradient de codage de phase sera augmentée. Entre deux mesures différentes, les protons qui se trouvent proches du centre ne subissent pas un grand changement de vitesse de précession (le gradient est toujours proche de zéro au centre) tandis que les protons qui se trouvent à l'extrémité subissent un changement d'intensité de champ beaucoup plus important. Cette information permet de coder l'emplacement spatial des protons suivant le deuxième axe. Pour cela, une deuxième transformée de Fourrier inverse est nécessaire suivant le deuxième axe de mesure.

II.2 IRM des tissus végétaux

Dans les tissus végétaux, le signal IRM mesuré provient des noyaux d'hydrogène présents dans la cellule végétale. Dans ce qui suit, nous présentons la composition de cette

cellule.

II.2.1 Cellule végétale

II.2.1.1 Composition

La cellule végétale se compose de compartiments hydriques séparés par des membranes semi-perméables (permettant des échanges des molécules d'eau). Les principaux compartiments sont :

- Le cytoplasme : il renferme divers organites, dont les chloroplastes, et les regroupe dans un milieu semi-liquide appelé le cytosol.
- La vacuole : c'est un compartiment rempli d'eau (environ 80% de l'eau cellulaire) contenant des solutés.
- la paroi cellulaire : elle donne une structure à la cellule et assure la communication de la cellule végétale avec son entourage.

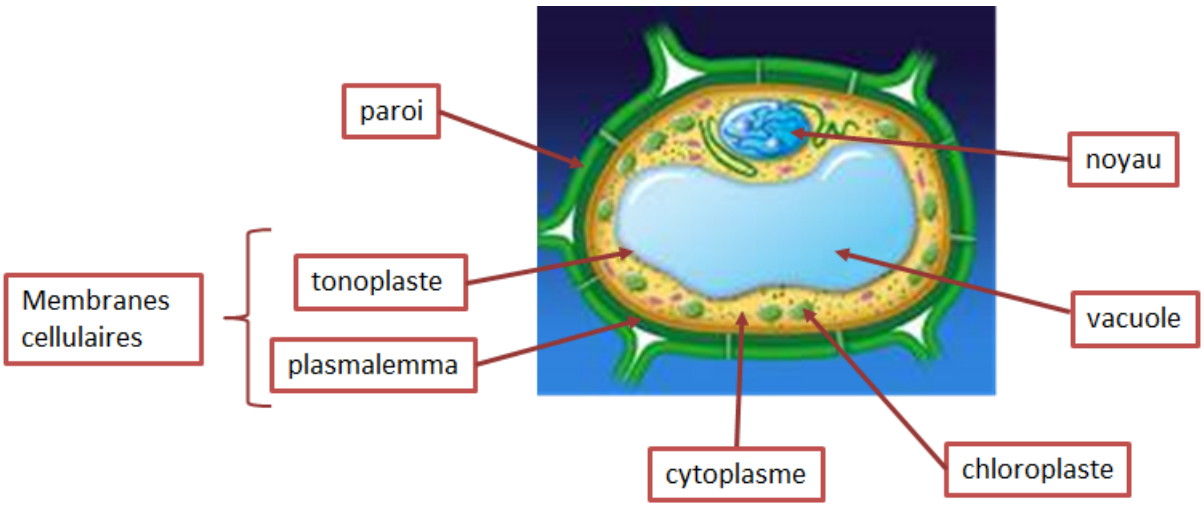

Figure II.5 – Structure de la cellule végétale

II.2.1.2 Temps de relaxation

A- à l'échelle d'un compartiment cellulaire Dans le cas d'un échantillon d'eau pure, et sous l'hypothèse d'homogénéité du champ magnétique, les différentes molécules d'eau partagent les mêmes conditions. Ainsi, les interactions chimiques entre ces molécules (par échange des protons et par interactions dipolaires) sont les mêmes. On obtient ainsi un seul temps de relaxation pour toutes ces molécules $T_{2_{eau}}$.

Par contre, une solution aqueuse présentera des noyaux d'hydrogène avec des temps de relaxation différents. Dans une soluté, on peut distinguer quatre types de populations de protons :

(a) Les protons de l'eau libre qui ne sont pas en interaction avec les molécules des solutés.

(b) Les protons de l'eau en interaction avec les molécules des solutés.

(c) Les protons échangeables des solutés.

(d) Les protons non échangeables des solutés.

Sous condition d'échange chimique rapide des protons a, b et c et si l'échange diffusionnel entre les molécules d'eau est aussi rapide, les temps de relaxation observés correspondent à un temps de relaxation « moyen » dépendant des temps de relaxation et des proportions de chaque type de protons. On peut faire l'hypothèse qu'à l'intérieur de chaque compartiment cellulaire ces échanges sont assez rapides et que chaque compartiment peut être représenté par un seul T_2.

B- à l'échelle de la cellule L'échange diffusionel entre les compartiments cellulaires à travers les membranes (tonoplaste et plasmalemme) impacte également le signal de relaxation. Cependant, on peut en première approximation considérer que ces échanges diffusionels sont suffisamment lents pour que le signal de relaxation transversale soit multi-exponentiel et permette de discriminer les différents compartiments cellulaires présentés dans la section II.2.

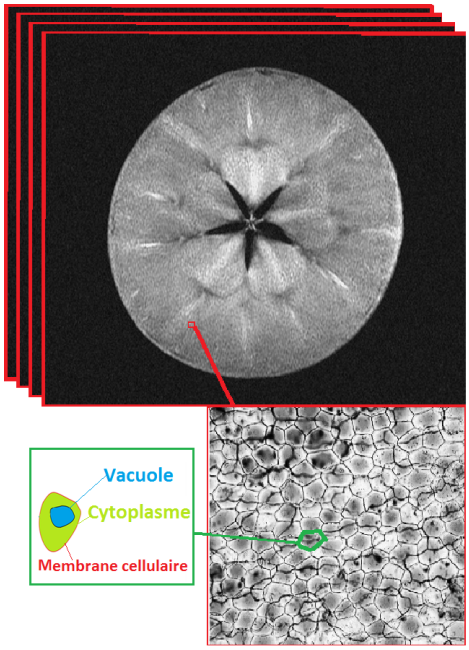

Figure II.6 – Le type et le grand nombre de cellules contenu dans les images IRM

C- à l'échelle d'un voxel La dimension de la cellule végétale est de l'ordre de 10 à 200 μm. Ainsi, à l'intérieur de chaque voxel d'un volume de l'ordre du mm^3, on trouve une multitude de cellules (voir figure II.6). On considère que ces ces compartiment subcellulaires sont similaires, et donc les temps de relaxation associés, sont similaires. Ainsi on associe à chaque voxel un seul type de cellule.

II.2.2 Séquence d'écho de spin multiple adaptée

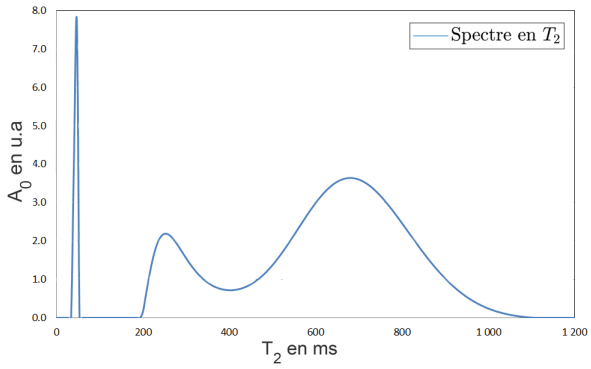

Figure II.7 – Distribution des temps de relaxation mesurée sur un ROI de 1100 voxels sur une Tomate.

Afin de prendre en compte les spécificités des temps de relaxation T_2 rencontrés dans les végétaux, il est nécessaire d'utiliser une séquence adaptée. En effet, dans la gamme mesurable par IRM compte-tenu du temps d'écho minimal disponible de l'ordre de quelques ms, les T_2 varient de quelques dizaines de ms à quelques centaines de ms. Ceci est illustré dans la figure II.7 où est représenté une distribution des T_2 dans une région d'intérêt de 1100 voxels estimée sur une tomate. Pour pouvoir estimer le T_2 dans cette large gamme il est nécessaire d'utiliser un temps d'écho le plus court possible tout en acquérant un grand nombre d'échos afin d'échantillonner la courbe de relaxation jusqu'à la fin de la relaxation des protons ayant les T_2 les plus longs. C'est pourquoi, une séquence ES-M a été développéepar l'équipe IRM-FOOD [11], permettant de mesurer, à l'aide d'une machine IRM clinique, un nombre d'échantillons élevé couvrant la totalité de la courbe de décroissance. Ce travail de thèse s'appuie sur cette séquence afin de pouvoir estimer les paramètres de la courbe de décroissance multi-exponentielle des tissus végétaux.

II.3 Modélisation du signal de décroissance en l'absence du bruit

À l'intérieur de chaque voxel j à l'instant τ_t, la contribution des pools d'eau est représentée par un spectre (distribution des temps de relaxation en T_2 sans bruit) présentant différents temps de relaxation avec différentes amplitudes :

$$s_j(\tau_t) = \int S(T_2) e^{\frac{-\tau_t}{T_2}} dT_2, \tag{II.3}$$

que nous notons s_t par la suite. $S(T_2)$ représente la fonction de distribution des temps de relaxation T_2. Ce spectre est généralement estimé par relaxométrie RMN, où le temps d'écho minimal est plus petit que dans les mesures IRM dû à l'absence des gradients de localisation spatiale.

II.3.1 Le modèle mono-exponentiel

En outre, dans les mesures IRM, des approximations représentatives de ce spectre sont utilisées. La plupart des études faites sur les courbes de décroissances acquises par IRM considèrent un signal de décroissance mono-exponentielle (voir figure II.8) :

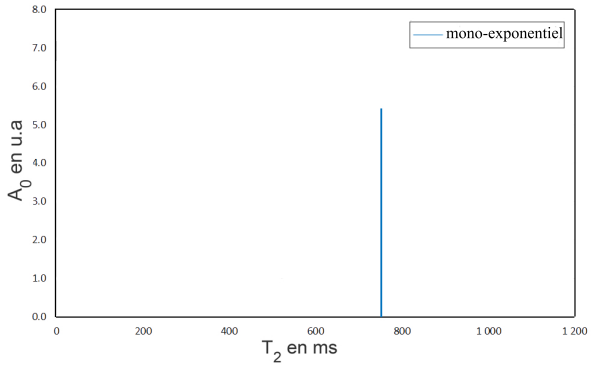

Figure II.8 – Exemple d'un spectre mono-exponentiel.

$$S(T_2) = A_{0_{mono}}\delta(T_2 - T_{2_{mono}}), \tag{II.4}$$

où $\delta(\cdot)$ représente la fonction Dirac. $A_{0_{mono}}$ et $T_{2_{mono}}$ représentent respectivement l'amplitude et le temps de relaxation du signal mono-exponentiel. Le modèle mono-exponentiel est souvent utilisé dans le domaine médical [21, 22]. Ce modèle a aussi été utilisé dans des applications sur le tissu végétal. En fait, ce modèle reste valable et intéressant pour plusieurs études dans ce domaine [23, 24].

Les temps de relaxation sont fortement sensibles à la structure des tissus et à l'intégrité des cellules. Un ajustement mono-exponentiel ne permet pas de connaitre la contribution des composantes ayant des temps de relaxation différents. Par exemple, en mesurant une variation du temps de relaxation T_2 mono-exponentiel, on peut attribuer cette variation soit à une variation des amplitudes des composantes du signal multi-exponentielles, soit à une variation des temps de relaxation multi-exponentielle soit aux deux en même temps.

II.3.2 Les modèles multi-exponentiels

Plusieurs modèles représentatifs de l'aspect multi-exponentiel du signal IRM dans chaque voxel peuvent être considérés. Dans ce qui suit nous présentons les différents types de modèles multi-exponentiels.

II.3.2.1 Modèle à distribution paramétrique

Ce modèle est basé sur l'hypothèse que la distribution des temps de relaxation à l'intérieur de chaque voxel suit une loi de probabilité précise. Ainsi, différentes valeurs de T_2 par compartiment peuvent être représentées par une distribution continue d'une fonction de probabilité. Le modèle dans ce cas dépend de la distribution choisie et il

dépend de l'application et sera choisi par expertise dans chaque domaine d'application. Dans la littérature, les fonctions de distributions qui ont été étudiées sont:

Distribution gaussienne:

$$S(T_2) = \sum_{c=1}^{N_c} A_c \cdot \mathcal{N}\left(T_2 \mid \mu_c, \sigma_c\right), \tag{II.5}$$

où $\mathcal{N}(T_2 \mid \mu_c, \sigma_c)$ est une loi normale de moyenne μ_c et d'ecart type σ_c et A_c représente la proportion de l'amplitude dans la distribution. Ici le vecteur de paramètres est $\boldsymbol{\theta} = [A_1, \mu_1, \sigma_1 \ ... \ A_{N_c}, \mu_{N_c}, \sigma_{N_c}]$. Cette distribution a été utilisée pour l'estimation de taux de myéline dans le cerveau [25]. Dans ce cas, le nombre de composantes a été fixé à $N_c = 2$. Ainsi, deux composantes représentant des distributions des temps de relaxation courts et moyens ont été représentées par une distribution gaussienne. Un Dirac représentant la composante ayant le T_2 le plus long a été utilisé et ainsi le nombre de composantes total dans ce cas était égal à 3.

Distribution gamma inverse:

$$S(T_2) = \sum_{c=1}^{N_c} A_c \cdot \frac{\beta^{(\alpha_c)}}{\Gamma\left(\alpha_c\right)} T_2^{-\alpha_c - 1} e^{-\frac{\beta_c}{T_2}}, \tag{II.6}$$

où $\Gamma\left(\alpha_c\right)$ est la fonction gamma de paramètre α_c. Ainsi, avec cette distribution le vecteur des paramètres à estimer devient: $\boldsymbol{\theta} = [A_1, \alpha_1, \beta_1, \ ... \ A_{N_c}, \alpha_{N_c}, \beta_{N_c}]$. De même que pour le cas précédent, ce modèle a été employé dans le cadre de l'estimation de taux de myéline dans le cerveau [26, 27]. Deux compartiments ont été considérés comme ayant une distribution gamma inverse des T_2, tandis que le compartiment ayant le T_2 le plus long a été représenté par un Dirac.

Distribution gaussienne inverse généralisée :

$$S(T_2) = \sum_{c=1}^{N_c} A_c \cdot \left(\frac{\lambda_c}{(2\pi R_2^3)}\right)^{\frac{1}{2}} e^{\left(\frac{-\lambda_c}{2\mu_c^2 R_2}(R_2 - \mu_c)^2\right)}, \tag{II.7}$$

où $\boldsymbol{\theta} = [A_c, \mu_c, \lambda_c]$, $R_2 = \frac{1}{T_2}$, μ_c la moyenne, λ_c un paramètre de forme et $\frac{\mu_c^3}{\lambda_c}$ la variance. La distribution gaussienne inverse généralisée a été étudiée pour un modèle à deux compartiments [28].

Dans le cas d'un modèle à distribution paramétrée, la fonction de probabilité à utiliser dépend de l'application. Ces modèles n'ont pas été testés pour des applications dans le domaine végétal.

II.3.2.2 Modèle à décroissance multi-exponentielle discrétisé

Un autre modèle représentatif du signal de décroissance est le modèle à décroissance multi-exponentielle discrétisé. Danc ce modèle, on cherche à estimer les intensités A_0 pour des valeurs prédéfinies de temps de relaxation T_2. En adoptant ce modèle, le nombre de variables est élevé mais le modèle devient linéaire par rapport aux paramètres à estimer. Ce modèle permet de mesurer des variations des paramètres à l'intérieur de chaque voxel. Il est défini par :

$$s_t(\boldsymbol{\theta}) = \sum_{T_2 \in Z_{T_2}} A_0\left(T_2\right) e^{\left(\frac{-\tau_t}{T_2}\right)}, \tag{II.8}$$

où Z_{T_2} contient N_{T2} valeurs de T_2 prédéfinies sur une échelle de valeurs définies selon l'application et $\boldsymbol{\theta} = [A_{0_1} \ \ldots \ A_{0_{N_{T2}}}]$. Le modèle discrétisé a été utilisé pour la détection du taux de myéline dans le cerveau [13, 29, 30, 31, 32]. Il faut noter que l'estimation du taux de myéline ne dépend pas de l'exactitude de la valeur de T_2, en effet elle dépend uniquement des amplitudes relatives entre les valeurs de T_2 courts et les deux autres valeurs. Dans l'application au tissu végétal, le modèle à décroissance multi-exponentielle discrétisé a été étudié en relaxométrie RMN [9, 33, 34].

II.3.2.3 Modèle à décroissance multi-exponentielle

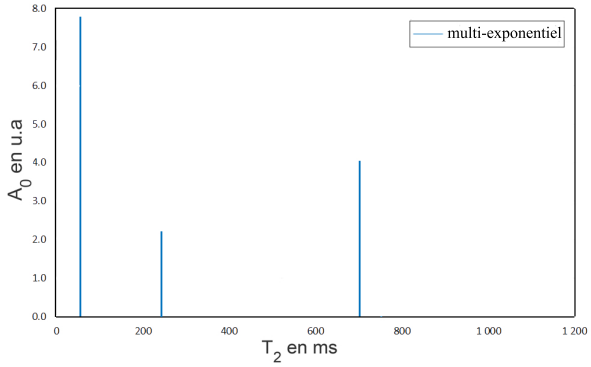

Figure II.9 – Exemple d'un spectre multi-exponentiel avec un nombre de composantes $N_c = 3$.

Le spectre multi-exponentiel peut être représenté par une discrétisation des distributions de valeurs de T_2 (voir figure II.9) :

$$S(T_2) = \sum_{c=1}^{N_c} A_{0_c} \delta\left(T_2 - T_{2_c}\right), \tag{II.9}$$

où A_{0_c} et T_{2_c} représentent respectivement les amplitudes et les temps de relaxation de la composante c du signal multi-exponetiel.

Ainsi, pour un nombre N_c de co mposantes, le modèle de décroissance est une somme

de N_c courbes de décroissance exponentielle :

$$s_t(\boldsymbol{\theta}) = \sum_{c=1}^{Nc} A_{0_c} e^{\frac{-\tau_t}{T_{2_c}}}, \tag{II.10}$$

ici $\boldsymbol{\theta} = [A_{0_1},\ T_{2_1}\ \ldots\ A_{0_c},\ T_{2_c}]$. Dans ce cas, le modèle présente une non linéarité par rapport aux paramètres du modèle.

Ce modèle multi-exponentielle présente un intérêt pour les applications où le nombre de composantes dans chaque voxel est considéré connu. Dans la littérature, le modèle multi-exponentiel est utilisé pour l'estimation de taux de myéline dans le cerveau. En effet, on peut faire l'hypothèse que dans chaque voxel on cherche à estimer les paramètres d'un modèle tri-exponentiel provenant du liquide cérébro-spinal, de l'eau intra et extra-cellulaire, et de l'eau entre les bicouches de la myéline [35, 36, 37, 38, 39]. D'autres exemples de l'application de ce modèle dans le domaine médical ont été présentés dans la littérature. Un modèle bi-exponentiel pour une application sur le cartilage nasal bovin [14] et l'œdème musculaire du rat [40]. Un modèle tri-exponentiel peut être encore représentatif pour des études du cartilage nasal bovin [17]. Pour l'étude du muscle squelettique, un modèle quadri-exponentiel a été utilisé [18]. Pour les applications sur des tissus végétaux, les composantes du modèle correspondent aux différents compartiments cellulaires.

II.3.3 Prise en compte de l'hétérogénéité du champ $\vec{B_1}$

La non-homogénéité de $\vec{B_1}$ perturbe les données mesurées et peut conduire à une estimation erronée des paramètres du modèle. À cause de cette hétérogénéité, l'angle de bascule de l'aimantation des protons de leur axe principal dépend de leur position. Par exemple, dans une séquence d'ES-M, lors de la phase de refocalisation il est nécessaire d'appliquer un angle de 180° afin d'obtenir un écho, mais en réalité l'angle appliqué n'est pas parfaitement égal à 180°.
Il existe un second problème qui est le profil de la sélection de coupe. En théorie, une coupe est un parallélépipède et en observant le profil suivant la profondeur c'est un « échelon » parfait. Pour obtenir cela, il faudrait appliquer la RF, donc $\vec{B_1}$, sur une durée infinie. Comme ce n'est pas le cas, l'angle appliqué en réalité n'est pas le même dans toute la profondeur de coupe. Dans le cas de l'acquisition d'une seule coupe, la sélection de coupe peut se faire uniquement pour la première impulsion à 90° degrés mais pas pour les 180° successifs, donc dans ce cas le problème du profil de coupe peut être négligé. En revanche, si l'on veut faire plusieurs coupes à la fois, alors il faut appliquer une sélection de coupe à chaque 180° et l'effet devient non négligeable.

Dans un article publié en 2016 [28], les deux problèmes de refocalisation et de coupe imparfaite sont abordés. Un modèle de distribution gaussienne inverse pour représenter la distribution de $R_2 = \frac{1}{T_2}$ a été utilisé. Un terme multiplicatif est ajouté au modèle pour

tenir compte de la non-homogénéité de champ $\vec{B_1}$:

$$s_t(\boldsymbol{\theta}) = A_0 \left(\frac{\lambda}{(2\pi R_2^3)} \right)^{\frac{1}{2}} \exp\left(\frac{-\lambda}{2\mu^2 R_2} (R_2 - \mu)^2 \right) GEC_t(R_2, B_1), \tag{II.11}$$

ici $\boldsymbol{\theta} = [A_0, R_2, \lambda, \mu, B_1]$. GEC ('Generalised Echo Curve') est la fonction qui contient une correction sur les échos pour réduire l'hétérogénéité de champ $\vec{B_1}$.

Dans le cadre de cette thèse, ces deux problèmes d'hétérogénéités ne seront pas traités. En effet, la séquence de mesure utilisée pour l'acquisition des images permet de faire l'hypothèse que ces effets sont négligeables grâce à l'utilisation d'impulsions de refocalisation non-sélectives et l'application de crushers optimaux [11].

II.4 Le bruit de mesure

Le bruit en IRM est issu de l'agitation thermique des protons (bruit thermique) et de la chaine de mesure (bruit électronique). Le bruit en IRM est supposé gaussien indépendant et identiquement distribué dans l'espace complexe des mesures (partie réelle et partie imaginaire) [41]. La distribution gaussienne est donnée par :

$$P_{Gauss}(y_t \mid s_t, \sigma^2) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(y_t - s_t)^2 / 2\sigma^2}, \tag{II.12}$$

où y_t représente le signal mesuré à l'instant τ_t, σ est l'écart-type du bruit indépendant du temps de mesure τ_t. Ainsi, on obtient à chaque voxel, dans l'espace complexe, les mesures suivantes :

$$\begin{cases} y_t^{re} = s_t \cos(\phi_t) + n_t^{re}, \\ y_t^{im} = s_t \sin(\phi_t) + n_t^{im}. \end{cases} \tag{II.13}$$

y_t^{re} et y_t^{im} sont les valeurs des parties réelles et imaginaires respectivement. ϕ_t représente la valeur de la phase à l'instant τ_t. Cette phase est sensible à de nombreux facteurs de mesure, notamment le type de la machine d'acquisition, la température, l'hétérogénéité du champ magnétique et la conception de la séquence d'acquisition. Elle présente ainsi une forte variation entre les instants de mesure τ_t et les voxels j [42]. Pour contourner le problème complexe d'estimation de la phase, c'est le module du signal qui est considéré :

$$y_t = \sqrt{y_t^{re^2} + y_t^{im^2}} \tag{II.14}$$

Le module des données complexes est un opérateur non linéaire qui transforme le bruit gaussien à une distribution connue sous le nom de la distribution de Rice [12] décrite par l'expression (II.15).

$$P_{Rice}(y_t \mid s_t, \sigma^2) = \frac{y_t}{\sigma^2} \exp\left(-\frac{y_t^2 + s_t^2}{2\sigma^2} \right) I_0\left(\frac{s_t y_t}{\sigma^2} \right), \tag{II.15}$$

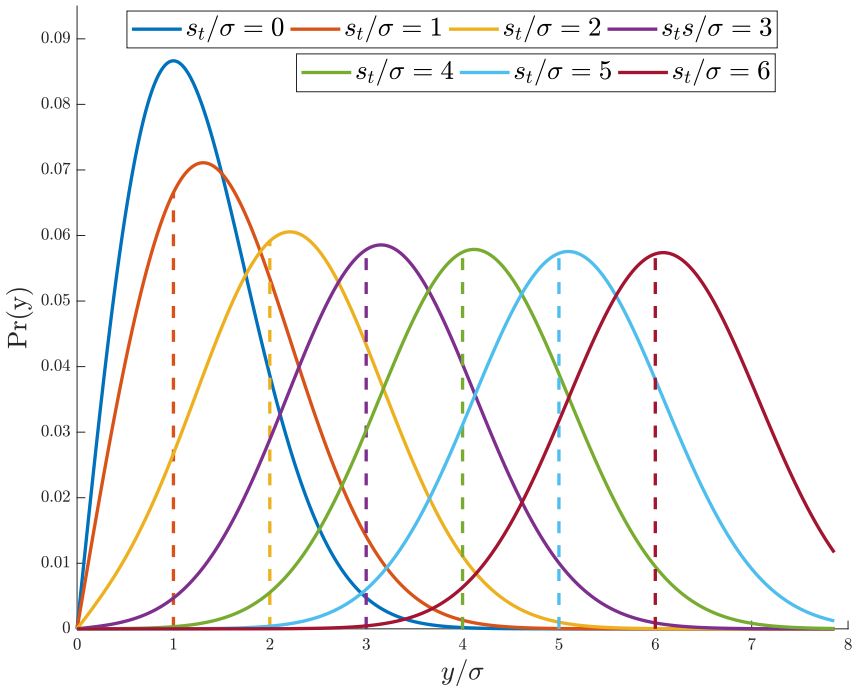

Figure II.10 – Distribution ricienne avec $\sigma = 7$ pour diverses valeurs de $(\frac{s_t}{\sigma})$.

avec $I_0(\cdot)$ la fonction de Bessel modifiée de première espèce et d'ordre zéro.

L'hypothèse d'une distribution ricienne est valide dans les cas suivants [43];

1. Les données ont été acquises avec une antenne mono-canal.

2. Les données ont été acquises avec des antennes multi-canaux non corrélées et la reconstruction a été faite avec un "filtre à correspondance spatiale" ('spatial matched filter' en anglais).

Si ces deux conditions ne sont pas respectées, la distribution du bruit suit la distribution de χ et/ou avec un modèle du bruit non stationnaire [44, 45]. Dans le cadre de cette thèse, nous nous sommes positionnés dans le cas où le bruit est stationnaire et suit la distribution ricienne $P_{Rice}(y_t \mid s_t, \sigma^2)$. En effet, ce cas correspond aux conditions d'acquisition utilisées dans le contexte de ce travail. Les extensions restent possibles au cas général.

II.4.1 Statistique de la distribution ricienne

Dans cette partie, nous présentons les statistiques riciennes du bruit en détaillant ses moments d'ordre 1 et d'ordre 2.

II.4.1.1 Moment d'ordre 1

Le moment ricien d'ordre 1 est donné par:

$$\begin{aligned} E_{Rice}[y_t] &= \int_0^{\infty} y_t P_R\left(y_t \mid s_t, \sigma^2\right) dy_t \\ &= \sigma\sqrt{\frac{\pi}{2}} e^{-\alpha_t}[(1+2\alpha_t)I_0(\alpha_t) + 2\alpha_t I_1(\alpha_t)], \end{aligned} \tag{II.16}$$

avec $\alpha_t = \frac{s_t^2}{4\sigma^2}$ et I_1 représente la fonction de Bessel modifiée de première espèce et d'ordre 1 [46].

II.4.1.2 Moment d'ordre 2

En reprenant les équations (II.13), le moment ricien de second ordre est exprimé par :

$$\begin{aligned} & E\left[y_t^2\right] = \\ & s_t^2 \cos^2\phi_t + s_t^2 \sin^2\phi_t \\ & + 2s_t \cos\left(\phi_t\right) E\left[n_t^{re}\right] \\ & + 2s_t \sin\left(\phi_t\right) E\left[n_t^{im}\right] \\ & + E\left[(n_t^{re})^2\right] + E\left[(n_t^{im})^2\right] \end{aligned} \tag{II.17}$$

Étant donné que le bruit gaussien a une moyenne nulle et une variance de σ^2, il en résulte :

$$E_{Rice}\left[y_t^2\right] = s_t^2 + 2\sigma^2. \tag{II.18}$$

II.4.2 Le rapport signal sur bruit

Le RSB en IRM dépend de plusieurs facteurs liés au matériel de mesure et au paramétrage de la séquence de mesure. Il dépend d'une façon directe de l'intensité de champ, de la séquence utilisée et des tissus mesurés mais il dépend aussi du volume de voxel, du nombre d'excitations (nombre de répétitions de la séquence de mesure), et de la bande passante de l'antenne de réception qui sont des phénomènes contrôlables. Ainsi, pour augmenter le RSB, on peut augmenter le signal en augmentant le volume du voxel ΔV (en augmentant la dimension du voxel suivant les axes X et/ou Y et/ou en choisissant une épaisseur de coupe plus large). Cela implique une dégradation de la résolution spatiale des images acquises. Ou bien on peut diminuer le bruit en :

- Augmentant le nombre d'excitations N_{EX}, cela implique une augmentation du temps de mesure.
- Diminuant la bande passante de réception BP, cela implique une augmentation de temps d'échantillonnage.

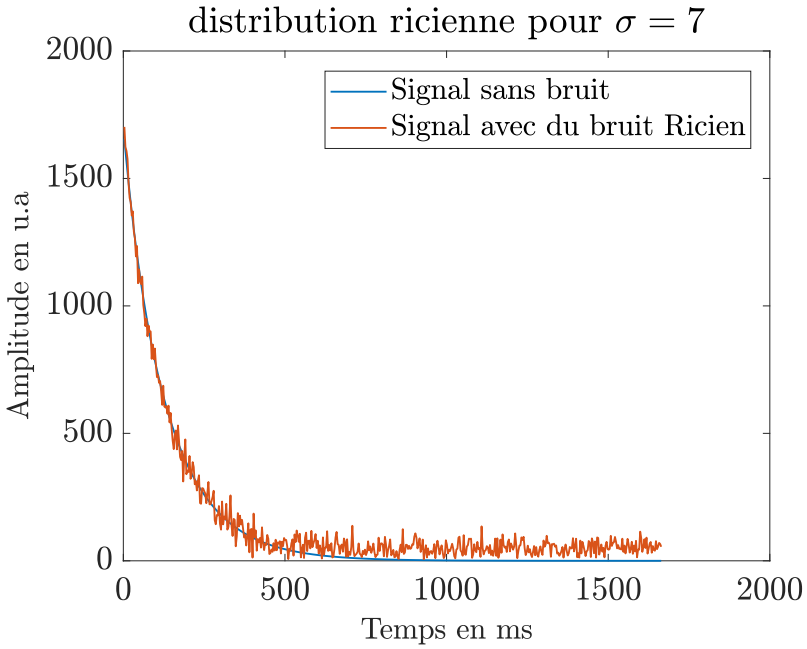

Figure II.11 – Effet du bruit ricien sur les données. Le signal mesuré est donné en rouge et le signal non bruité en bleu pour un $RSB = 230$.

- Augmentant la taille de matrice de codage spatiale T_M (dans le cas de codage en phase).

En fait le RSB dépend de ces différents facteurs suivant la formule suivante:

$$RSB \propto \Delta V \frac{\sqrt{T_M N_{EX}}}{\sqrt{BP}} \tag{II.19}$$

L'augmentation du RSB fait tendre la distribution ricienne vers une distribution gaussienne, et ainsi l'écart présent entre la moyenne de cette distribution et les vraies valeurs diminue (voir figure II.10). Le RSB d'une image IRM est calculé comme suit:

$$RSB_t = \frac{\sum_{j=1}^{N_v} y_{jt}}{\sigma N_v}, \tag{II.20}$$

où y_{jt} représente le signal du voxel j à l'instant τ_t. Au début de la relaxation, le RSB est à sa valeur maximale, il diminue au cours de la relaxation et à la fin de la relaxation le signal émis par les protons devient faible et ainsi le RSB devient extrêmement faible. Ainsi, en relaxométrie le signal mesuré décroit au cours du temps amenant à une mesure d'un signal de faible amplitude pour les derniers échantillons temporels. Dans la suite du document, le RSB correspondra à la valeur du RSB au début de la relaxation :

$$RSB = \frac{\sum_{j=1}^{N_v} y_{j1}}{\sigma N_v} \tag{II.21}$$

Par conséquent, on est toujours amené à traiter des images à RSB faibles. Cela implique

qu'une écart est toujours présente entre les valeurs mesurées à faibles amplitudes et les vraies valeurs (voir figure II.11).

II.4.3 Estimation des paramètres du bruit

Pour calculer le RSB et estimer les paramètres de relaxation d'une façon réussie, il faut estimer la variance du bruit sous l'hypothèse d'une distribution ricienne. Les images IRM contiennent des zones où il n y a que du bruit. Cela permet l'estimation des paramètres du bruit à partir de ces zones. Dans les régions du fond où aucun signal n'est présent, cette distribution peut être assimilée à une distribution de Rayleigh :

$$P_{Ray}(y_{jt}|s_{jt}=0,\sigma)=\frac{y_{jt}}{\sigma^2}\exp^{(-\frac{y_{jt}^2}{2\sigma^2})}, \tag{II.22}$$

s_{jt} représente le signal non bruité au voxel j et à l'instant τ_t. Le moment d'ordre 1 de la distribution de Rayleigh peut être utilisé pour l'estimation de la variance du bruit. Toujours sous l'hypothèse que le bruit est le même pour tous les voxels, cela amène le calcul de l'écart-type à la formulation suivante : [47]

$$\sigma=\sqrt{\frac{2}{\pi}}\times\frac{1}{N_bN_\tau}\sum_{t=1}^{N_\tau}\sum_{b=1}^{N_b}y_{bt}, \tag{II.23}$$

ici, N_b représente le nombre de voxels appartenant à la région avec un RSB extrêmement faible.

II.5 Modèle choisi pour l'étude

En relaxométrie T_2, chaque voxel de l'image IRM contient un signal décrivant le phénomène de relaxation transversale. Ces images sont corrompues par un bruit du type ricien. Nous avons présenté les conditions qui aboutissent à la distribution ricienne du bruit et les différents aspects de ce modèle. Nous avons établi qu'en relaxométrie T_2 le problème de RSB faible est toujours présent à cause des mesures effectuées à la fin de la relaxation.

En outre, nous avons considéré, en première approximation, que les signaux à l'intérieur de chaque voxel proviennent d'un ensemble de cellules végétales identiques. De plus, nous avons établi que la structure de la cellule peut être modélisée par trois compartiments hydriques présentant des différences entre eux en matière de teneur en eau, de l'interaction des molécules d'eau avec leur environnement et de l'échange diffusionnel entre les compartiments. Ces trois parties sont : l'eau dans la vacuole, le cytoplasme et l'eau d'hydratation de la paroi. Ainsi, chaque compartiment hydrique est caractérisé par un seul temps de relaxation. Un modèle de décroissance tri-exponentielle permet de modéliser le signal

acquis sur un tissu végétal ($s_t(\boldsymbol{\theta}) = \sum_{c=1}^{3} A_{0_c} e^{\frac{-\tau_t}{T_{2_c}}}$).

Dans le chapitre suivant, nous présentons une méthode d'estimation des paramètres de ce modèle à l'échelle du voxel.

III

Estimation des paramètres à l'échelle du voxel

Dans le chapitre précédent, nous avons établi que le signal de chaque voxel des images mesurées en relaxométrie IRM est modélisé par une décroissance multi-exponentielle. En effet, ce modèle représente la contribution des différents compartiments hydriques présents dans le voxel. Les paramètres de ce modèle représentent les amplitudes des composantes du signal de relaxation A_0 et les temps de relaxation transversale T_2 qui prennent des valeurs strictement positives. Néanmoins, l'estimation de ces paramètres est un problème inverse mal posé par définition. En effet, on définit un problème comme étant bien posé seulement si les conditions suivantes sont bien vérifiées :

1. Existence : une solution existe.
2. Unicité : la solution est unique.
3. Stabilité : la solution dépend d'une façon continue des mesures.

Sous les conditions initiales d'échantillonnage et de la forte présence du bruit, un problème inverse peut avoir différentes solutions pour les mêmes données mesurées. Ainsi, quel que soit le modèle adopté, le problème d'estimation des paramètres du modèle en présence du bruit est un problème mal posé. Les problèmes inverses mal posés sont généralement résolus en trouvant les paramètres qui maximisent la probabilité qu'un modèle choisi soit représentatif des données mesurées. On dit alors qu'on maximise la vraisemblance du modèle par rapport aux données mesurées. Cette vraisemblance dépend de la distribution statistique du bruit. Dans les images de module en IRM, ce bruit suit une distribution ricienne. Ainsi, toute méthode d'estimation des paramètres doit prendre en compte cette distribution.

De plus, l'algorithme d'estimation doit assurer la positivité des paramètres. En fait, l'estimation des paramètres multi-exponentiels peut aboutir à des composantes avec des valeurs de paramètre négatives. Ce résultat n'est pas cohérent puisque les paramètres de relaxation sont strictement positifs. Ainsi, une contrainte de positivité permet d'assurer la cohérence physique des solutions.

Commençons, dans le cadre de l'établissement de l'état de l'art, par présenter l'approche d'estimation des paramètres dans le cas d'une distribution gaussienne. Ensuite, nous présenterons diverses approches d'estimation des paramètres intégrant la distribution ricienne du bruit.

Pour rappel, les paramètres à estimer sont ceux du vecteur $\boldsymbol{\theta}$ d'un modèle de décroissance multi-exponentielle $s_t(\boldsymbol{\theta})$. Dans le cas du modèle multi-exponentiel le vecteur de paramètres $\boldsymbol{\theta} = [A_{0_1},\ T_{2_1}\ \ldots\ A_{0_c},\ T_{2_c}]$, cependant, d'autres modèles peuvent être considérés sans changement des méthodes d'estimation présentées dans la suite.

III.1 Estimation sous l'hypothèse du bruit gaussien

Cette méthode d'estimation basée sur une hypothèse gaussienne du bruit sera présentée dans le cadre de l'établissement de l'état de l'art. Elle a été employée dans la littérature

pour sa facilité d'implémentation. De plus, il est intéressant de présenter cette approche dans la mesure où la plupart des méthodes qui prennent en considération le modèle ricien du bruit peuvent être présentées comme étant une correction de cette méthode.
La méthode de maximisation de la vraisemblance dans le cadre du bruit gaussien est basée sur la maximisation de la fonction de densité de probabilité donnée dans l'équation (II.12) pour toutes les images mesurées. On obtient ainsi, à chaque voxel, le critère de vraisemblance à maximiser $V_G(\boldsymbol{y}, \boldsymbol{\theta})$ donné par :

$$\begin{aligned} \underset{\boldsymbol{\theta}\in\mathbb{R}_+^N}{\operatorname{argmax}}\, V_G(\boldsymbol{y},\boldsymbol{\theta}) &= \underset{\boldsymbol{\theta}\in\mathbb{R}_+^N}{\operatorname{argmax}} \prod_{t=1}^{N_\tau} P_G\left(y_t \mid s_t\left(\boldsymbol{\theta}\right), \sigma^2\right) \\ &= \underset{\boldsymbol{\theta}\in\mathbb{R}_+^N}{\operatorname{argmax}} \prod_{t=1}^{N_\tau} \frac{1}{\sqrt{2\pi\sigma^2}} e^{\frac{-(y_t - s_t(\boldsymbol{\theta}))^2}{2\sigma^2}} \\ &= \underset{\boldsymbol{\theta}\in\mathbb{R}_+^N}{\operatorname{argmax}} \prod_{t=1}^{N_\tau} e^{\frac{-(y_t - s_t(\boldsymbol{\theta}))^2}{2\sigma^2}}, \end{aligned} \tag{III.1}$$

avec $\boldsymbol{y}$ le vecteur contenant les mesures pour tous les temps de mesure τ_t. La maximisation de la fonction de vraisemblance est effectuée en maximisant le logarithme népérien de celle-ci. On obtient ainsi la fonction de log-vraisemblance suivante :

$$\begin{aligned} \underset{\boldsymbol{\theta}\in\mathbb{R}_+^N}{\operatorname{argmax}} \log\left(V_G\left(\boldsymbol{y},\boldsymbol{\theta}\right)\right) &= \underset{\boldsymbol{\theta}\in\mathbb{R}_+^N}{\operatorname{argmax}} \sum_{t=1}^{N_\tau} -\left[\frac{\left(y_t - s_t\left(\boldsymbol{\theta}\right)\right)^2}{2\sigma^2}\right] \\ &= \underset{\boldsymbol{\theta}\in\mathbb{R}_+^N}{\text{minimiser}} \frac{1}{2\sigma^2} \sum_{t=1}^{N_\tau} \left[y_t - s_t\left(\boldsymbol{\theta}\right)\right]^2 \end{aligned} \tag{III.2}$$

Le critère ainsi obtenu est connu sous le nom du critère des moindres carrés. Une autre façon de l'interpréter consiste à dire qu'on minimise la distance entre les données mesurées et les valeurs attendues d'une distribution gaussienne des données, sachant que l'espérance gaussienne $E_{Gauss}[y_t] = s_t(\boldsymbol{\theta})$. Il en résulte ainsi le critère des moindres carrés à minimiser par :

$$C_{MC}(\boldsymbol{y},\boldsymbol{\theta}) = \frac{1}{2\sigma^2} \sum_{t=1}^{N_\tau} \left[y_t - s_t\left(\boldsymbol{\theta}\right)\right]^2, \tag{III.3}$$

Ce modèle des moindres carrés a été utilisé dans la littérature dans le cas où une distribution gaussienne du bruit a été considérée pour l'estimation des paramètres de relaxation ([35, 29, 48]).

III.2 Estimation des paramètres sous l'hypothèse d'un bruit ricien

L'estimation des paramètres sous l'hypothèse du bruit ricien sera essentielle pour réduire le biais d'estimation [49]. Pour cela, plusieurs méthodes prenant en considération

ce bruit ont été proposées dans la littérature.

III.2.1 Correction basée sur l'observation

Certaines méthodes proposent des corrections basées sur l'observation (le biais sur les échantillons temporels à faible amplitude), mais ne prennent pas en considération la vraie statistique du bruit. Commençons par présenter ces différentes méthodes.

III.2.1.1 Correction de Gudjbartsson sur le signal mesuré

Gudjbartsson et Patz [12] ont proposé une correction sur le signal mesuré basée sur le biais observé en approximant la distribution ricienne par une distribution gaussienne. En fait, ils ont proposé de corriger cette erreur en admettant que la distribution ricienne puisse être approchée par une distribution gaussienne de variance σ^2, mais de moyenne $\sqrt{s_t^2 + \sigma^2}$. Cette correction est introduite pour les échantillons temporels avec un RSB faible et le signal mesuré sera remplacé par un signal corrigé comme suit :

$$y_t = \begin{cases} \sqrt{y_t^2 + \sigma^2}, & \text{si RSB faible} \\ y_t, & \text{autrement} \end{cases} \tag{III.4}$$

En se basant sur cette correction, une estimation des paramètres d'un modèle de relaxation bi-exponentiel a été proposée en minimisant un critère des moindres carrés [14].

III.2.1.2 Moindres carrés avec un paramètre représentant le biais du bruit ricien

Comme présenté dans la section II.4.2, le bruit ricien induit une divergence entre les échantillons mesurés à faible RSB et le signal non bruité. Pour réduire cette divergence, dans cette méthode un paramètre B est ajouté au critère des moindres carrés [50] :

$$C_{MCB}(\boldsymbol{y}, \boldsymbol{\theta}) = \frac{1}{2\sigma^2} \sum_{t=1}^{N_\tau} \left[y_t - s_t(\boldsymbol{\theta}) - B \right]^2, \tag{III.5}$$

où B est un paramètre à estimer. Cette technique semble intuitive du fait que la divergence des valeurs proches de zéro est prise en compte par ce terme. Le problème est que dans cette technique le biais est considéré comme constant au cours du temps (le long de la décroissance) ce qui n'est pas le cas réellement. En effet, le biais varie en fonction de l'amplitude du signal (voir figure II.11), plus l'amplitude du signal est élevée plus le biais est faible.

III.2.1.3 Moindres carrées pondérés

Une autre approche pour réduire l'effet du biais est d'utiliser un critère des moindres carrés pondérés afin d'affecter moins de poids dans le critère des moindres carrés sur les échantillons temporels à faible amplitude [51] :

$$C_{MCP}(\boldsymbol{y}, \boldsymbol{\theta}) = \frac{1}{2\sigma^2} \sum_{t=1}^{N_\tau} \omega_t^2 \left[y_t - s_t(\boldsymbol{\theta}) \right]^2, \tag{III.6}$$

où ω_t représente le poids affecté au critère du maximum de vraisemblance à l'instant τ_t. W_t est choisit tel que pour les valeurs de τ_t élevées W_t a des valeurs plus faibles.

Un cas spécial de cette méthode est la méthode de troncature où il suffit de ne pas prendre en considération les échantillons temporels ayant des valeurs inférieures à un certain seuil $W_t = 0 \; \forall \; t \, ; \tau_t > \tau_{seuil}$. Néanmoins, le problème réside dans le choix d'une stratégie de pondération qui s'adapte aux différents RSB.

Ces méthodes basées sur des corrections du signal mesuré ne seront pas retenues pour la suite de cette discussion.

III.2.2 Correction basée sur les statistiques riciennes du bruit

Comme alternative, il existe des méthodes qui prennent les statistiques du bruit ricien en considération afin de réduire son effet. Dans la section suivante, nous présenterons les différentes méthodes basées sur les statistiques riciennes du bruit que nous retenons pour la suite.

III.2.2.1 Correction de McGibney

En 1993, une observation faite indépendamment par McGibney et Smith [52] et par Miller et Joseph [53] a conduit à l'utilisation du moment d'ordre 2 ($E\left[y_t^2\right]$) du signal mesuré. En se basant sur l'équation (II.18), les auteurs proposent de minimiser la distance entre le carré du signal et son espérance ricienne. On obtient ainsi le critère des Moindres Carrés Corrigé de McGibney (MCCMG) :

$$C_{MCCMG}(\boldsymbol{y}, \boldsymbol{\theta}) = \sum_{t=1}^{N_\tau} \left[\, y_t^2 - \left(s_t(\boldsymbol{\theta})^2 + 2\sigma^2 \right) \right]^2. \tag{III.7}$$

III.2.2.2 Correction par l'espérance ricienne

Une autre approche utilisée dans la littérature [14, 54] consiste à considérer le moment d'ordre 1 de la distribution ricienne donnée dans l'équation (II.16). En fait, dans cette approche, la minimisation sera apportée d'une façon directe sur la différence entre le signal mesuré et son espérance sous l'hypothèse du bruit ricien. Cette méthode est aussi motivée

par le fait que le critère des moindres carrés peut être considéré comme une minimisation de la différence entre le signal mesuré et son espérance ricienne.
Ainsi, la méthode des moindres carrés corrigés par l'espérance ricienne est donnée par :

$$C_{MCCER}(\boldsymbol{y},\boldsymbol{\theta}) = \frac{1}{2\sigma^2}\sum_{t=1}^{N_\tau}\Big[y_t - f\left(s_t\left(\boldsymbol{\theta}\right),\sigma\right)\Big]^2, \tag{III.8}$$

où $f\left(s_t\left(\boldsymbol{\theta}\right),\sigma\right) = E_{Rice}[y_t]$ (voir l'équation II.16.

III.2.2.3 Estimation par Maximum de Vraisemblance

La méthode d'estimation la plus correcte pour prendre en considération les statistiques du bruit est la maximisation de la vraisemblance des mesures en utilisant la probabilité ricienne donnée dans l'équation (II.15).

$$V(\boldsymbol{y},\boldsymbol{\theta}) = \prod_{t=1}^{N_\tau} P_R\left(y_t \mid s_t\left(\boldsymbol{\theta}\right),\sigma^2\right). \tag{III.9}$$

La maximisation de cette fonction revient à minimiser la fonction d' anti-log-vraisemblance :

$$\begin{aligned} -\log\left(V\left(\boldsymbol{y},\boldsymbol{\theta}\right)\right) &= \sum_{t=1}^{N_\tau}\left[\frac{y_t^2}{2\sigma^2} - \log\left(\frac{y_t}{\sigma^2}\right)\right] \\ &+ \sum_{t=1}^{N_\tau}\left[\frac{s_t\left(\boldsymbol{\theta}\right)^2}{2\sigma^2} - \log\left(I_0\left(z_t\left(\boldsymbol{\theta}\right)\right)\right)\right], \end{aligned} \tag{III.10}$$

avec $z_t(\boldsymbol{\theta}) = \dfrac{y_t s_t\left(\boldsymbol{\theta}\right)}{\sigma^2}$.
En observant l'équation (III.10), le premier terme ne dépend pas des paramètres à estimer et ainsi la minimisation de cette équation revient à la minimisation du critère de maximum de vraisemblance ricien (MVR) donné par :

$$C_{MVR}\left(\boldsymbol{y},\boldsymbol{\theta}\right) = \sum_{t=1}^{N_\tau}\left[\frac{s_t\left(\boldsymbol{\theta}\right)^2}{2\sigma^2} - \log\left(I_0\left(z_t\left(\boldsymbol{\theta}\right)\right)\right)\right]. \tag{III.11}$$

Même si cet estimateur de MV présente des caractéristiques intéressantes au niveau de la statistique du bruit, la minimisation de ce critère, non-convex, reste un problème plus difficile à gérer qu'un critère basé sur les moindres carrés. Dans les approches basées sur la minimisation du critère MV un algorithme de type Newton-Raphson a été déployé [46].

III.3 Algorithme d'optimisation

Dans la partie précédente, nous avons présenté plusieurs critères qui ont été proposés dans la littérature pour l'estimation des paramètres de relaxation sous l'hypothèse ricienne. Tous ces critères, sauf celui de maximisation de vraisemblance, peuvent être ramenés à un

critère des moindres carrés non-linéaires. Néanmoins, comme établi dans l'introduction, à cause de la forte influence du bruit, le problème d'estimation des paramètres est un problème mal posé et plusieurs minima locaux peuvent se présenter dans les différentes fonctions objectives à minimiser. Ainsi, une solution analytique n'existe pas et la minimisation d'un critère des moindres carrés non-linéaires est souvent approchée en utilisant une approche itérative. Ainsi, la minimisation du critère est conduite par une série d'itérations ℓ, où, à chaque itération une direction de descente $d^{(\ell)}$ est calculée afin de garantir la décroissance d'un critère $C(\boldsymbol{y}, \boldsymbol{\theta})$:

$$C(\boldsymbol{y}, \boldsymbol{\theta}^{(\ell)} + \boldsymbol{d}^{(\ell)}) < C(\boldsymbol{y}, \boldsymbol{\theta}^{(\ell)}) \tag{III.12}$$

Ces algorithmes sont connus sous le nom d'algorithmes de minimisation des moindres carrés non linéaires (MCNL). L'algorithme MCNL le plus utilisé dans la littérature pour le problème d'estimation des paramètres de relaxation est l'algorithme de Levenberg-Marquardt (LM). Dans ce qui suit, nous présentons l'algorithme de LM.

III.3.1 Algorithme de Levenberg-Marquardt

L'algorithme LM est un des algorithmes basés sur un calcul de direction de descente en considérant le gradient $\boldsymbol{g}$ du critère à minimiser. Dans ces algorithmes, la première étape consiste à choisir des valeurs initiales des paramètres $\boldsymbol{\theta}^{(0)}$ suffisamment proches du minimum du critère. Ensuite, à chaque itération, la direction de descente est calculée en fonction du gradient d'une façon à garantir la minimisation du critère.

En 1943, Levenberg a proposé une méthode de résolution basée sur la direction de Gauss-Newton à laquelle il a introduit un paramètre λ tel que [55] :

$$\boldsymbol{d}^{(\ell)} = \left(\boldsymbol{J}^T\left(\boldsymbol{\theta}^{(\ell)}\right)\boldsymbol{J}\left(\boldsymbol{\theta}^{(\ell)}\right) + \lambda^{(\ell)}\boldsymbol{I}\right)^{-1}\boldsymbol{J}^T\left(\boldsymbol{\theta}^{(\ell)}\right)\left(\boldsymbol{r}\left(\boldsymbol{y}, \boldsymbol{\theta}^{(\ell)}\right)\right) \tag{III.13}$$

$\boldsymbol{I}$ est la matrice identité, $\boldsymbol{r}\left(\boldsymbol{y}, \boldsymbol{\theta}^{(\ell)}\right)$ est le résidu tel que : $C(\boldsymbol{y}, \boldsymbol{\theta}) = \frac{1}{2}||\boldsymbol{r}(\boldsymbol{y}, \boldsymbol{\theta})||_2^2$ et pour un nombre de paramètre N_p, le Jacobien du gradient $\boldsymbol{J}$ est donné par (une discussion sur le calcul du Jacobien pour les différents critères est présentée dans l'annexe A):

$$\boldsymbol{J}\left(\boldsymbol{\theta}^{(\ell)}\right) = \begin{bmatrix} \frac{\partial \boldsymbol{r}(y_1, \boldsymbol{\theta})}{\boldsymbol{\theta}(1)} & \cdots & \frac{\partial \boldsymbol{r}(y_1, \boldsymbol{\theta})}{\boldsymbol{\theta}(N_p)} \\ \vdots & \ddots & \vdots \\ \frac{\partial \boldsymbol{r}(y_{N_\tau}, \boldsymbol{\theta})}{\boldsymbol{\theta}(1)} & \cdots & \frac{\partial \boldsymbol{r}(y_{N_\tau}, \boldsymbol{\theta})}{\boldsymbol{\theta}(N_p)} \end{bmatrix}$$

20 ans après cette proposition, Donald W. Marquardt a contribué au développement de cette méthode en proposant de substituer la matrice identité par une matrice formée par la diagonale de $\boldsymbol{J}^T\left(\boldsymbol{\theta}^{(\ell)}\right)\boldsymbol{J}\left(\boldsymbol{\theta}^{(\ell)}\right)$ [56]. L'algorithme ainsi formé sera connu sous le nom de l'algorithme de Levenberg-Marquardt. La direction de descente de LM est donné par :

$$\boldsymbol{d}^{(\ell)} = \left(\boldsymbol{J}^T\left(\boldsymbol{\theta}^{(\ell)}\right)\boldsymbol{J}\left(\boldsymbol{\theta}^{(\ell)}\right) + \lambda^{(\ell)}\mathrm{diag}\left(\boldsymbol{J}^T\boldsymbol{J}\right)\right)^{-1}\boldsymbol{J}^T\left(\boldsymbol{\theta}^{(\ell)}\right)\left(\boldsymbol{r}\left(\boldsymbol{y}, \boldsymbol{\theta}^{(\ell)}\right)\right) \tag{III.14}$$

Le paramètre $\lambda^{(\ell)}$ assure la convergence rapide de l'algorithme en assurant un compromis entre la descente de gradient et celle de Gauss-Newton. En fait, au début de l'algorithme et puisqu'on est loin du minimum on a intérêt à choisir une grande valeur pour $\lambda^{(\ell)}$ afin de s'approcher de la direction de gradient de descente. Puis, une fois près du minimum, on diminue la valeur de $\lambda^{(\ell)}$ et la direction de descente se rapproche de celle de Gauss-Newton. Il existe différentes méthodes, itératives et directes, pour le choix de $\lambda^{(\ell)}$. Nous proposons la mise à jour de ce paramètre tel que $\lambda^{(\ell)} = min(\|\text{gradient}\|, 10^{-5} max(diag(\boldsymbol{J}^{(\ell)T}\boldsymbol{J}^{(\ell)})))$. Ensuite, à chaque itération et jusqu'à la convergence de l'algorithme, le vecteur des paramètres sera mis à jour suivant la direction calculée :

$$\boldsymbol{\theta}^{(\ell+1)} = \boldsymbol{\theta}^{(\ell)} + \alpha^{(\ell)}\boldsymbol{d}^{(\ell)} \tag{III.15}$$

$\alpha^{(\ell)}$ est un pas qui doit garantir la descente du critère et la convergence de l'algorithme. L'approche classique pour le choix du pas qui garantit une descente suffisante du critère à minimiser est d'adopter une stratégie basée sur une technique de rebroussement qui satisfait la condition d'Armijo définit par [57] :

$$C(\boldsymbol{y}, \boldsymbol{\theta}^{(\ell+1)}) \leq C(\boldsymbol{y}, \boldsymbol{\theta}^{(\ell)}) + c_1\alpha^{(\ell)}\boldsymbol{g}^{(\ell)^T}\boldsymbol{d}^{(\ell)} \text{ avec } c_1 \in]0, 1/2]. \tag{III.16}$$

$\boldsymbol{g}^{(\ell)^T}$ représente la valeur du gradient à l'itération ℓ. c_1 est une constante avec une valeur typique $c_1 = 10^{-3}$.

III.3.2 Algorithme de majoration-minimisation

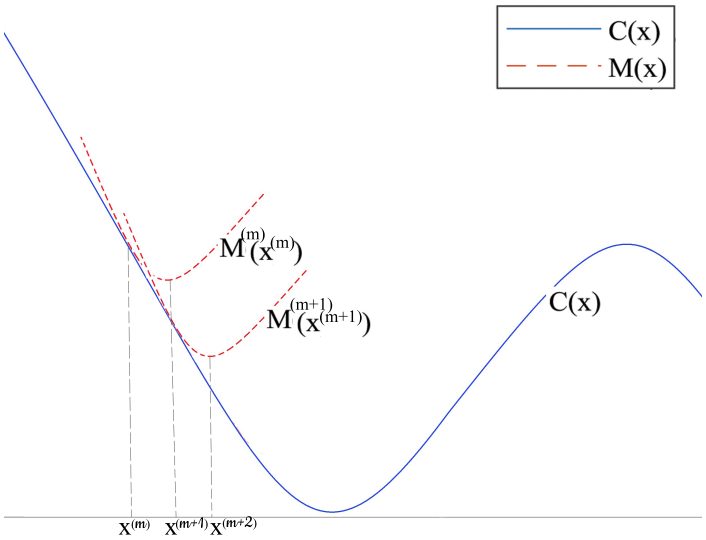

Figure III.1 – Décroissance du critère en utilisant l'approche MM.

Dans les algorithmes de majoration-minimisation (MM) la minimisation d'une fonction objective est abordée en minimisant de manière itérative une fonction majorante à celle-ci.

L'approche est surtout pratique quand la minimisation de la fonction majorante est plus facile que celle de la fonction originale. Pour le calcul de la fonction majorante, les propriétés de convexité et de concavité de la fonction objective sont exploitées. Dans cette technique, la minimisation sera assurée par une série de majorations de la fonction objective et de minimisations de la fonction majorante [58]. Prenons pour exemple un critère C défini sur $\boldsymbol{x} \in \mathrm{dom}_C$. Une définition classique de la majoration peut être formulée comme suit; à chaque itération m de l'algorithme, il existe une fonction $M^{(m)}$ telle que :

$$\begin{aligned} &M^{(m)}\left(\boldsymbol{x}^{(m)}\right) = C\left(\boldsymbol{x}^{(m)}\right), \\ &M^{(m)}\left(\boldsymbol{x}\right) \geq C(\boldsymbol{x}), \quad \forall\, \boldsymbol{x} \in \mathrm{dom}_C \end{aligned} \tag{III.17}$$

$M^{(m)}$ est la fonction majorante, elle est tangente à C au point $\boldsymbol{x}^{(m)}$. À chaque itération m de l'algorithme MM, la décroissance du critère C sera ainsi assurée par la minimisation du critère majorant $M^{(m)}$ (voir figure III.1). Ainsi la procédure MM est divisée en deux étapes; l'étape de la majoration et l'étape de la minimisation. Dans l'étape de la majoration, une fonction majorante doit être choisie d'une façon qu'elle se rapproche localement de la fonction objective C. Ensuite, dans l'étape de la minimisation, la fonction majorante choisie sera minimisée en utilisant un algorithme de minimisation adapté. La particularité de cette technique est qu'en conduisant une série d'itérations sur ces deux étapes, le critère C fini par être minimisé après la convergence de l'algorithme.

III.4 Algorithme d'estimation proposé

La première contribution de ce travail de thèse consiste en une proposition d'une approche basée sur un algorithme de Majorisation-Minimisation en adoptant une résolution approchée afin de minimiser le critère C_{MVR} d'une façon réussie et de proposer une modification sur l'étape de recherche de pas de l'algorithme LM afin d'assurer la positivité des paramètres. Le critère de MV sera majoré par une fonction majorante quadratique et la minimisation du critère global sera assurée par une série de minimisations de cette fonction quadratique majorante par minimisation MCNL [59].

III.4.1 Analyse de convexité du critère de maximum de vraisemblance ricien

Nous avons introduit dans la partie précédente l'algorithme de LM utilisé dans le cadre de la minimisation d'un critère de moindres carrés non linéaire. Or, le critère de maximum de vraisemblance ricien C_{MVR} est un critère non quadratique et non convexe. En fait, ce

critère, donné dans l'équation (III.11), peut se réécrire sous la forme :

$$C_{MVR}\left(\boldsymbol{y},\boldsymbol{\theta}\right)=\sum_{t=1}^{N_\tau}\left(C_1\left(s_t\left(\boldsymbol{\theta}\right)\right)+C_2\left(s_t\left(\boldsymbol{\theta}\right)\right)\right). \tag{III.18}$$

tel que $C_1\left(s_t\left(\boldsymbol{\theta}\right)\right)$ est une fonction quadratique strictement convexe par rapport à $s_t(\boldsymbol{\theta})$:

$$C_1\left(s_t\left(\boldsymbol{\theta}\right)\right)=\frac{s_t\left(\boldsymbol{\theta}\right)^2}{2\sigma^2}, \tag{III.19}$$

et compte tenu de fait que la fonction $I_0(.)$ est une fonction convexe par définition, $C_2\left(s_t\left(\boldsymbol{\theta}\right)\right)$ est non quadratique et strictement concave par rapport à $s_t(\boldsymbol{\theta})$, :

$$C_2\left(s_t\left(\boldsymbol{\theta}\right)\right)=-\log\left(I_0\left(z_t\left(\boldsymbol{\theta}\right)\right)\right). \tag{III.20}$$

Ainsi, la minimisation du critère $C_{MVR}\left(\boldsymbol{\theta}\right)$ n'est pas directement possible en utilisant l'algorithme LM. Nous proposons d'utiliser une approche basée sur l'algorithme de majoration-minimisation (MM) afin de résoudre le problème de minimisation de ce critère.

III.4.2 Calcul de la majorante

Nous souhaitons majorer le critère concave par une fonction strictement convexe afin d'améliorer l'estimation en obtenant un critère majorant plus simple à minimiser avec un seul minimum global (en l'absence du bruit). Étant concave, la majorante de la fonction $C_2(s_t(\boldsymbol{\theta}))$ à chaque point $\tilde{s}_t(\boldsymbol{\theta})$ est donnée par sa dérivée en ce point [59]. Ainsi on peut établir la relation :

$$-\log\left(I_0\left(z_t(\boldsymbol{\theta})\right)\right)\leq-\log\left(I_0\left(\tilde{z}_t(\boldsymbol{\theta})\right)\right)-\frac{y_t}{\sigma^2}R\left(\tilde{z}_t(\boldsymbol{\theta})\right)\left[s_t(\boldsymbol{\theta})-\tilde{s}_t(\boldsymbol{\theta})\right], \tag{III.21}$$

avec $R(\cdot)=\dfrac{I_1(\cdot)}{I_0(\cdot)}$ et $\tilde{z}_t(\boldsymbol{\theta})=\dfrac{y_t\tilde{s}_t(\boldsymbol{\theta})}{\sigma^2}$.
Ainsi, l'étape de majoration donne à chaque itération m est donnée par :

$$\begin{aligned}C_1\left(s_t\left(\boldsymbol{\theta}\right)\right)+C_2\left(s_t\left(\boldsymbol{\theta}\right)\right)\leq&\frac{s_t^2\left(\boldsymbol{\theta}\right)}{2\sigma^2}-\frac{\tilde{y}_t\left(\boldsymbol{\theta}^{(m)}\right)}{\sigma^2}s_t\left(\boldsymbol{\theta}\right)\\&-\log\left(I_0\left(\tilde{z}_t\left(\boldsymbol{\theta}^{(m)}\right)\right)\right)+\frac{\tilde{y}_t\left(\boldsymbol{\theta}^{(m)}\right)}{\sigma^2}\tilde{s}_t\left(\boldsymbol{\theta}^{(m)}\right).\end{aligned} \tag{III.22}$$

avec $\tilde{y}_t\left(\boldsymbol{\theta}^{(m)}\right)=R\left(\tilde{z}_t\left(\boldsymbol{\theta}^{(m)}\right)\right)y_t$ et $\boldsymbol{\theta}^{(m)}$ le vecteur de paramètres à l'itération m. Les deux derniers termes de l'équation (III.22) ne dépendent pas des paramètres $\boldsymbol{\theta}$. En fait à chaque itération de l'algorithme MM ces valeurs sont constantes durant la minimisation du critère. Ainsi, une formulation permet d'aboutir à un critère quadratique à chaque

itération de l'algorithme MM. le critère majorant est donné par :

$$\begin{aligned} M_{MVR}^{(m)}(\boldsymbol{y},\boldsymbol{\theta}) &= \frac{1}{2\sigma^2}\sum_{t=1}^{N_\tau}\left(s_t^2(\boldsymbol{\theta}) - 2\tilde{y}_t\left(\boldsymbol{\theta}^{(m)}\right)s_t(\boldsymbol{\theta})\right) + v_1 \\ &= \frac{1}{2\sigma^2}\sum_{t=1}^{N_\tau}\left(s_t^2(\boldsymbol{\theta}) - 2\tilde{y}_t\left(\boldsymbol{\theta}^{(m)}\right)s_t(\boldsymbol{\theta}) + \tilde{y}_t^2\left(\boldsymbol{\theta}^{(m)}\right)\right) + v_2 \\ &= \frac{1}{2\sigma^2}\sum_{t=1}^{N_\tau}\left[\tilde{y}_t\left(\boldsymbol{\theta}^{(m)}\right) - s_t(\boldsymbol{\theta})\right]^2 + v_2, \end{aligned} \tag{III.23}$$

où v_1 et v_2 contiennent les valeurs qui ne dépendent pas des paramètres $\boldsymbol{\theta}$. L'étape de majoration a abouti à une fonction majorante quadratique dans la famille des critères des MCNL. Ainsi, l'étape de minimisation de la fonction majorante, à chaque itération m, est réalisée en utilisant l'algorithme LM.

III.4.3 Majoration-minimisation approchée

De plus, la minimisation exacte de la fonction majorante obtenue dans la sous-section précédente n'est pas nécessaire, ce qui est important c'est d'assurer la descente du critère majoré de MVR [60, 61]. Nous proposons d'alterner entre la mise à jour de la fonction majorante et de prendre quelques pas de descente sur cette majorante en utilisant une version tronquée de l'algorithme de LM.

III.4.4 Prise en compte des contraintes

En plus, nous proposons d'initialiser la rétropropagation à une valeur maximale $\alpha_{\max}$ assurant la positivité des paramètres. Ainsi, on cherche pour chaque paramètre $\theta(p)^{(\ell)}$ dont la direction calculée est négative ($d^{(\ell)}(p) < 0$) que le pas maximal assure l'inégalité suivante:

$$\theta(p)^{(\ell)} + \alpha_{\max} d^{(\ell)}(p) > 0 \;\forall\; p\,; d^{(\ell)}(p) < 0. \tag{III.24}$$

Ainsi, le pas maximal qui satisfait cette condition est donné par:

$$\alpha_{\max} = \min_{p\in\mathbb{P}} -\theta^{(\ell)}(p)/d^{(\ell)}(p) \text{ avec } \mathbb{P} = \{p\,; d^{(\ell)}(p) < 0\}. \tag{III.25}$$

L'algorithme proposé de minimisation par majoration du critère de maximum de vraisemblance (MM-MV) est résumé dans l'algorithme 1.

L'algorithme est considéré avoir convergé si l'une des conditions suivantes est satisfaite :

- La variation du critère $\frac{C_{MVR}^{(m+1)} - C_{MVR}^{(m)}}{C_{MVR}^{(m)}}$ entre deux itérations de l'algorithme MM devient inférieure à une valeur de tolérance ($\epsilon_f = 10^{-3}$).
- La valeur de la norme du gradient devient inférieure à une valeur de tolérance ($\epsilon_g = 10^{-3}$).

Algorithme 1: Les étapes principales de l'algorithme MM-MV

Données : $\boldsymbol{\theta}^{(0)}, \boldsymbol{y}$;
for $m \leftarrow 0$ **to** $m_{\max}$ *ou jusqu'à convergence* **do**
    Fixer $\tilde{s}_t = s_t(\boldsymbol{\theta}^{(m)})$;
    Calculer $\tilde{y}_t = R\left(\tilde{z}_t\right) y_t$;
    Minimiser $M_{MVR}^{(m)}\left(\boldsymbol{y}, \boldsymbol{\theta}\right)$ en utilisant l'algorithme LM avec les paramètres suivants :
    - Fixer le nombre maximal d'itérations LM à une petite valeur;
    - Données : $\boldsymbol{\theta}^{(m)}$;
    - Sortie : $\boldsymbol{\theta}^{(m+1)}$;
end
Sortie $\boldsymbol{\theta}^{(m+1)}$

III.5 Analyse des performances sur des données simulées

Dans la suite de ce chapitre, une validation de l'algorithme MM-MV proposé sera présentée en le comparant aux deux méthodes de l'état de l'art qui prennent en considération le modèle ricien du bruit. Une série de tests a été faite sur deux objets virtuels construits; le premier objet virtuel (OV1) représente une tomate et le deuxième objet virtuel (OV2) représente une pomme. Pour le modèle tri-exponentiel ($N_c = 3$) six cartographies de paramètres (A_0 et T_2) sont estimées à chaque étape de la validation. Le but de cette partie est de montrer le gain en précision apporté par la prise en compte du bruit ricien et de le comparer à plusieurs méthodes de l'état de l'art sous différentes conditions.

III.5.1 Objets virtuels

Les deux jeux de données ont été construits dans le but d'avoir des images avec des régions définies et des images intégrant la complexité du fruit. L'objet virtuel 1 (OV1) est réparti spatialement en 5 régions d'une façon à avoir des valeurs des paramètres constantes dans chaque région. La répartition circulaire de ces régions est similaire à la structure des fruits. Ainsi, un critère d'erreur dans chaque région permettra de valider l'erreur pour différentes configurations des paramètres.

Le second objet virtuel (OV2) est basé sur des paramètres estimés sur une pomme. Cela amène à la construction d'une pomme virtuelle servant comme validation de l'algorithme dans des conditions proches de celles retrouvées dans les images IRM. Dans la section suivante, les deux objets virtuels seront présentés d'une façon plus détaillée.

Table III.1 – valeurs des paramètres de relaxation dans chaque région de l'OV1.

N°	Région	A_{01}	T_{21}	A_{02}	T_{22}	A_{03}	T_{23}
1	anneau extérieur	95	88	459	356	1015	716
2	région intérieure	96	76	410	433	1024	870
3	disque central	71	50	339	218	908	627
4	disques de droite et de gauche	109	78	375	303	1004	685
5	disques de haut et de bas	108	50	482	202	756	508

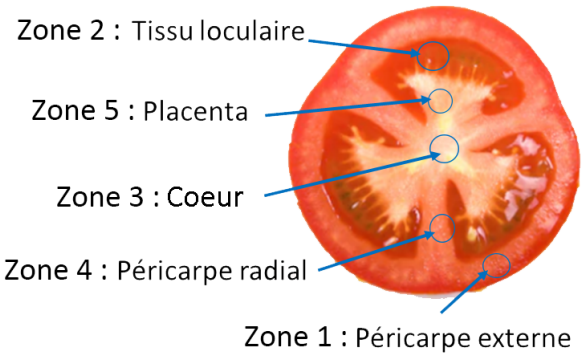

Figure III.2 – Les régions d'intérêts choisis pour la simulation de l'OV1

III.5.1.1 Objet virtuel 1

Afin de construire des images totalement synthétiques, mais qui prennent en considération les distributions des paramètres dans une tomate, des temps de relaxation multi-exponentielle ont été estimés sur les régions présentées dans la figure III.2. À partir de ces paramètres, une image a été reconstruite de dimensions 128×128. Elle est répartie en un disque central, un anneau extérieur et de 9 disques intérieurs de tailles différentes. L'objet sera divisé en 5 régions avec des paramètres estimés sur les régions d'intérêt de la figure III.2 comme détaillé dans le tableau III.1.

III.5.1.2 Objet virtuel 2

L'OV2 représente une pomme virtuelle avec une distribution de paramètres proches de celle estimées sur des images IRM acquises sur une pomme. En utilisant les paramètres de mesures de la section III.5.2. Une série de $N_{ex} = 32$ excitations a servi pour obtenir des données avec un RSB élevé($RSB \sim 630$). En utilisant ces images, l'algorithme de LM par voxel pour minimiser un critère de moindres carrés a été utilisé pour estimer les paramètres d'un modèle de décroissance tri-exponentielle. Ensuite, un filtrage spatial a été appliqué sur les cartographies estimées afin d'obtenir des cartographies de références lisses.

III.5.2 Paramètres de l'IRM

Dans l'étape d'acquisition de données, nous avons utilisé une séquence M-ES sur un scanner IRM de $1.5T$ (Magnetom, Avanto, Siemens, Erlangen, Germany) avec $\Delta TE = 6.5ms$ et un nombre d'images $N_t = 512$. Le plan médian du fruit (section transversale au milieu du fruit) a été imagé avec un nombre de voxels 128×128 et une épaisseur de coupe de 5 mm ainsi, la taille du voxel est de $1.19 \times 1.19 \times 5\ mm^3$. Ces paramètres sont utilisés pour la simulation des images et pour les données expérimentales dans le chapitre suivant.

III.5.3 Simulation des images

Afin de tester l'algorithme sous des conditions maîtrisées, les deux objets virtuels ont été utilisés en ajoutant du bruit ricien sur les images construites avec des RSB différents. À partir des objets virtuels, les images de référence sont reconstruites en reprenant les paramètres utilisés pour l'acquisition. Ensuite, du bruit gaussien a été ajouté à la partie réelle et imaginaire et le module des images complexes bruitées ainsi créées est calculé. Pour l'OV1, différents RSB ont été testés, variant entre une valeur égale à 150 jusqu'à 450. Pour l'OV2, un RSB de 150 a été utilisé, cela permet de reproduire un niveau de bruit qui correspond à celui rencontré dans les données expérimentales.

III.5.4 Initialisation des algorithmes

Une des contraintes importantes qui influence la convergence d'un algorithme d'estimation en présence du bruit à faible RSB est la stratégie d'initialisation des paramètres. En fait, le critère qu'on cherche à minimiser présente des minima locaux conduisant à une estimation erronée. La stratégie que nous avons adoptée est d'initialiser tous les voxels aux mêmes valeurs choisies par expertise pour chaque cas d'étude (tomate, pomme ...).

III.5.5 Critère d'évaluation

Le critère numérique utilisé pour l'évaluation des différentes méthodes est le critère de la racine de l'erreur quadratique moyenne normalisée (REQMN). En fait, dans le REQMN, la normalisation permet de quantifier l'erreur d'une façon équitable pour des paramètres ayant d'ordres de grandeur différents.

$$\text{REQMN} = 100\sqrt{\frac{1}{N_v N_p}\sum_{j=1}^{N_v}\sum_{p=1}^{N_p}\frac{\left(\theta_j\left(p\right)-\theta_j^*\left(p\right)\right)^2}{\theta_j^*\left(p\right)^2}} \tag{III.26}$$

Où N_p correspond au nombre de paramètres à estimer ($2N_c$ dans le cas c-exponentiel), $\boldsymbol{\theta}_j^*$ représente le vecteur des paramètres de référence au voxel j et $\boldsymbol{\theta}_j$ représente le vecteur des paramètres estimés au voxel j.

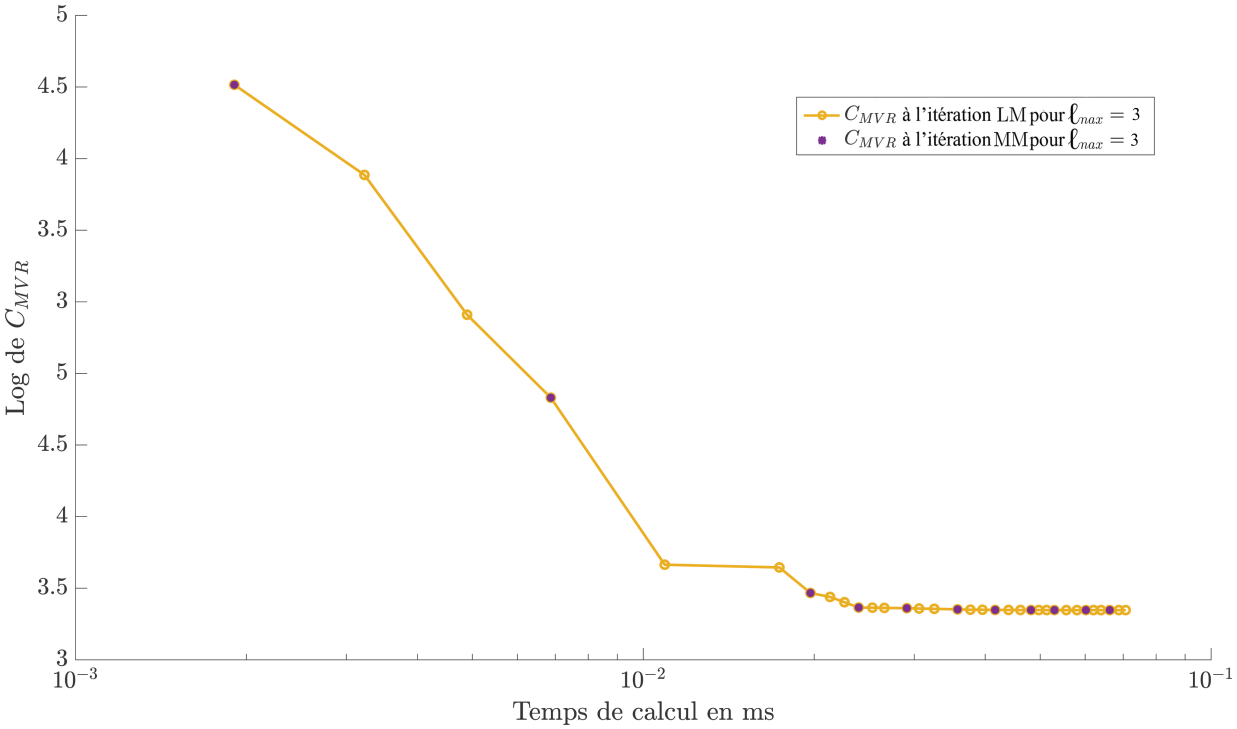

Figure III.3 – Évolution du critère de MVR dans un voxel au cours des itérations MM (points en mauves) et des itérations LM interne (points en jaune) pour un nombre d'itérations internes maximal ℓ_{max}.

III.5.6 Résultats

Dans cette section, nous présentons les résultats obtenus sur les deux objets virtuels OV1 et OV2 en appliquant l'algorithme LM pour minimiser les critères suivants :

- Critère des Moindres Carrés $C_{MC}(\boldsymbol{y}, \boldsymbol{\theta})$ (méthode MC)
- Critère des Moindres Carrés Corrigé par l'observation de McGibney $C_{MCCMG}(\boldsymbol{y}, \boldsymbol{\theta})$ (méthode MCCMG)
- Critère des Moindres Carrés corrigé par l'espérance ricienne $C_{MCCER}(\boldsymbol{y}, \boldsymbol{\theta})$ (méthode MCCER)
- Série de minimisation des critères majorant du Maximum de Vraisemblance ricien $M_{MVR}(\boldsymbol{y}, \boldsymbol{\theta})$ (méthode MVR)

Nous présenterons les résultats de l'algorithme d'estimation en comparant la méthode de minimisation par majoration du critère de maximum de vraisemblance aux critères de moindres carrés corrigés. Pour cela, des tests ont été conduits sur les deux objets virtuels simulés.

III.5.6.1 Analyse empirique de convergence

Dans un premier temps, nous avons validé, sur l'OV1 avec un RSB de 150, la convergence de l'algorithme en traçant l'évolution du critère majoré au cours des itérations (voir figure III.3). Cela nous a permis de valider que le critère C_{MVR} décroit en minimisant le critère

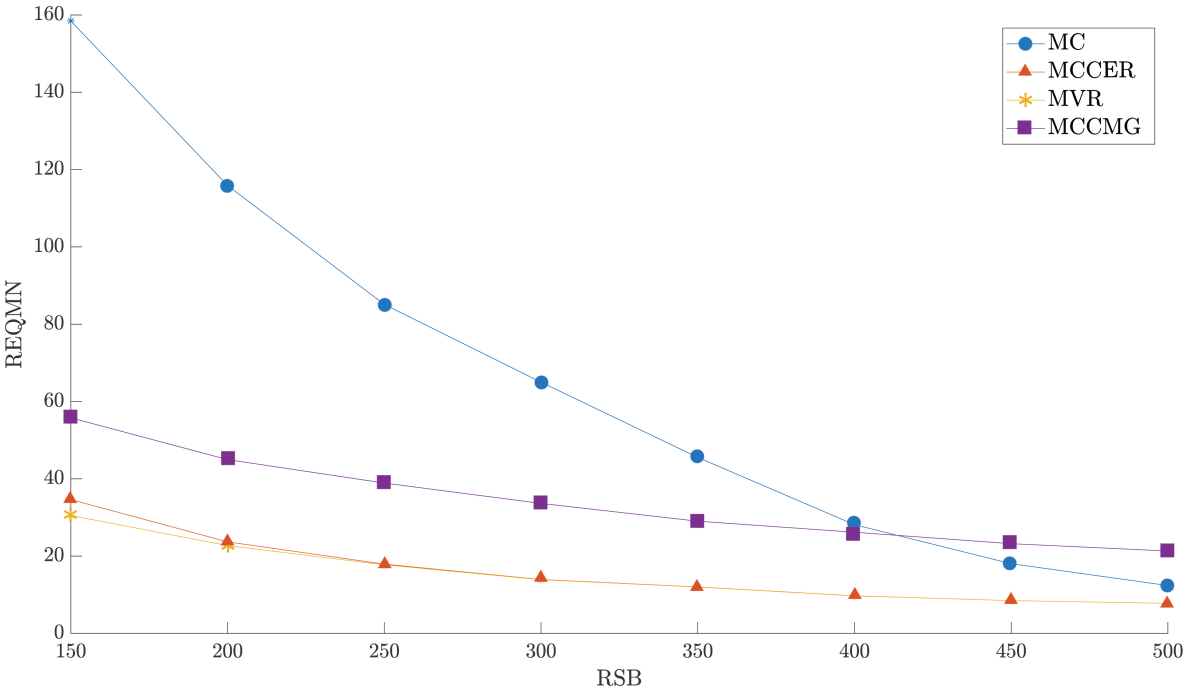

Figure III.4 – Variation de REQMN pour différents RSB sur l'OV1 simulé en minimisant les critères MC, MCCMG, MCCER, et MVR

majorant M_{MVR}. De plus, nous résumons dans le tableau III.2, le nombre d'itérations et le temps de calcul moyens par voxel ainsi que leurs valeurs maximales, minimales et leurs valeurs totales sur l'ensemble de voxels (pour 3064 voxel). On remarque que le nombre d'itérations et le temps de calcul dépendent des mesures à l'intérieur de chaque voxel.

Table III.2 – Nombre d'itérations et temps de calcul de la méthode MVR

	moyen	minimum	maximum	nombre total
Nombre d'itérations	29	6	3000	112323
Temps de calcul (en secondes)	0.05	0.01	51	207

III.5.6.2 Résultats sur l'OV1

Pour l'OV1, les paramètres ont été initialisés à une valeur identique pour tous les voxels tel que : $\boldsymbol{\theta}^{(0)} = [A_{0_1}, T_{2_1}, A_{0_2}, T_{2_2}, A_{0_3}, T_{2_3}] = [95, 75, 423, 368, 600, 605]$. La figure III.4 résume les valeurs des REQMN pour les paramètres estimés en minimisant les différents critères. Comme prévu, l'estimation par minimisation du critère des moindres carrés aboutit à une très grande erreur pour les RSB faibles et s'améliore en augmentant le RSB. Cela est bien expliqué par la nature du bruit ricien qui se rapproche d'une distribution gaussienne pour les RSB élevés. Par contre, les méthodes qui prennent la nature ricienne du bruit en considération ont abouti à une amélioration d'estimation importante. La méthode MCCMG a abouti à une bonne estimation pour les RSB faibles, mais par contre pour les RSB élevés cette méthode a abouti à une précision pire que celle basée sur l'hypothèse gaussienne.

Les deux méthodes qui ont abouti à la meilleure précision sont celles basées sur la statistique ricienne d'ordre 1 (MCCER et MVR).

III.5.6.3 Résultats sur l'OV2

Pour le cas de l'OV2, nous avons utilisé un vecteur d'initialisation tel que $\boldsymbol{\theta}^{(0)} = [A_{0_1}, T_{2_1}, A_{0_2}, T_{2_2}, A_{0_3}, T_{2_3}] = [95, 57, 322, 207, 367, 430]$. Ce vecteur d'initialisation correspond à la moyenne des vraies valeurs des cartographies de référence. Les trois méthodes qui prennent en considération la nature ricienne du bruit ont été testées. Dans les figures III.5 et III.6 représentent respectivement les cartographies reconstruites des A_0 et T_2. On peut remarquer que visuellement les deux méthodes, MCCER et MVR, aboutissent à des résultats très similaires et que ces deux méthodes présentent une meilleure reconstruction que la méthode MCCMG. De plus, numériquement nous avons obtenu un REQMN égale à 30% pour la méthode de MCCER, 30% pour la méthode de MVR et 40% pour la méthode de MCCMG. Cela valide les résultats des tests sur l'OV1, et l'on peut conclure que les méthodes de MCCER et MVR présentent les meilleurs résultats avec des reconstructions presque similaires.

III.6 Conclusion

Dans ce chapitre, nous avons établi l'état de l'art sur les méthodes d'estimation des paramètres qui prennent en compte la distribution ricienne du bruit. Ensuite, nous avons proposé une méthode de minimisation d'un critère basée sur la maximisation de la vraisemblance ricienne. Cette méthode, basée sur une minimisation partielle d'une fonction majorante quadratique, est connue sous le nom de majoration-minimisation. La minimisation partielle de la fonction majorante quadratique a été conduite en utilisant l'algorithme LM. Cela facilite l'implémentation de la minimisation du critère quadratique et en plus nous permettra d'implémenter, comme nous le verrons dans le chapitre suivant, des méthodes telles que l'estimation conjointe des paramètres de l'image et la prise en compte d'une régularité spatiale.
Nous avons démontré, en ce qui concerne l'erreur d'estimation mesurée en REQMN, l'importance de la prise en compte de la nature ricienne du bruit. De plus, la méthode proposée et celle basée sur une correction du critère des moindres carrés en prenant en compte l'espérance ricienne ont abouti à la meilleure estimation. Néanmoins, la prise en compte de la nature ricienne du bruit ne suffit pas dans les conditions des mesures aboutissant à un RSB faible. En fait, pour un RSB égal à 150 la REQMN aboutit à une erreur d'estimation de l'ordre de 30%. De plus, étant donné que le calcul est effectué par voxel, le temps de calcul est élevé. Dans le chapitre suivant, nous proposons une méthode d'estimation conjointe pour réduire le temps de calcul et une implémentation d'une méthode de régularisation spatiale afin de réduire l'erreur d'estimation.

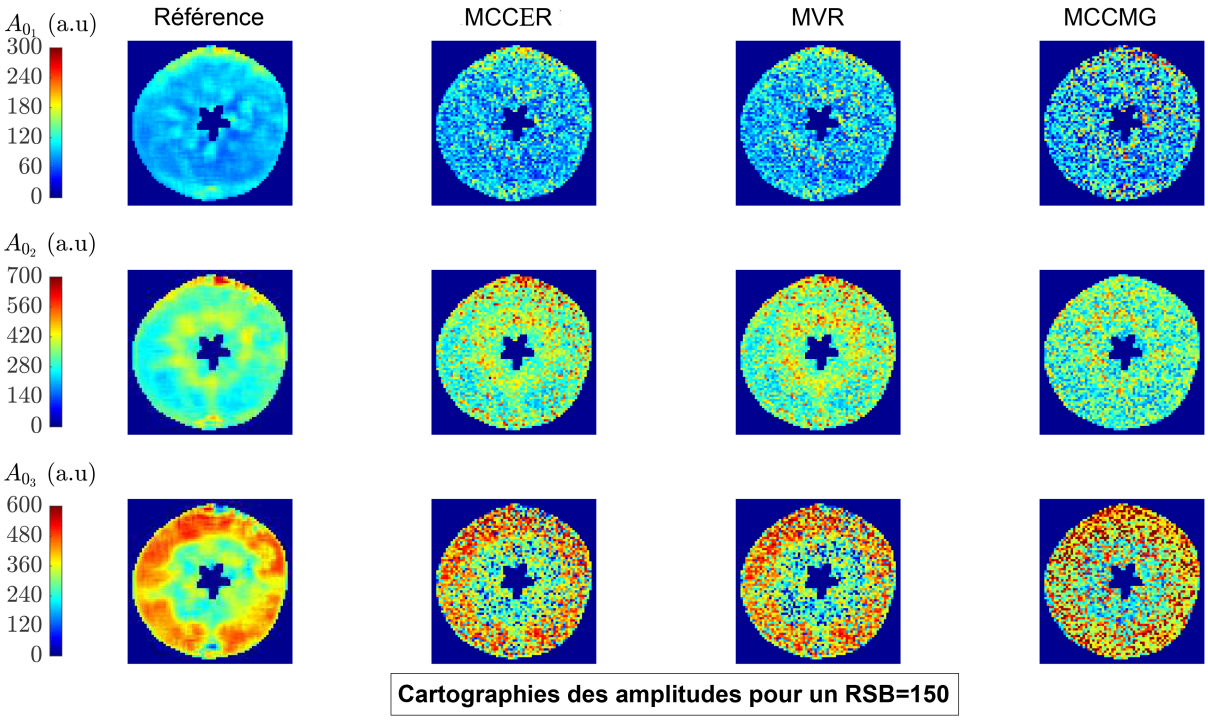

Figure III.5 – Résultats de l'estimation sur l'OV2. La première colonne représente les cartographies des amplitudes A_0 de référence utilisée pour la simulation. La deuxième, la troisième et la quatrième colonne représentent respectivement les cartographies des A_0 reconstruites en utilisant les méthodes de MCCER, MVR, MCCMG.

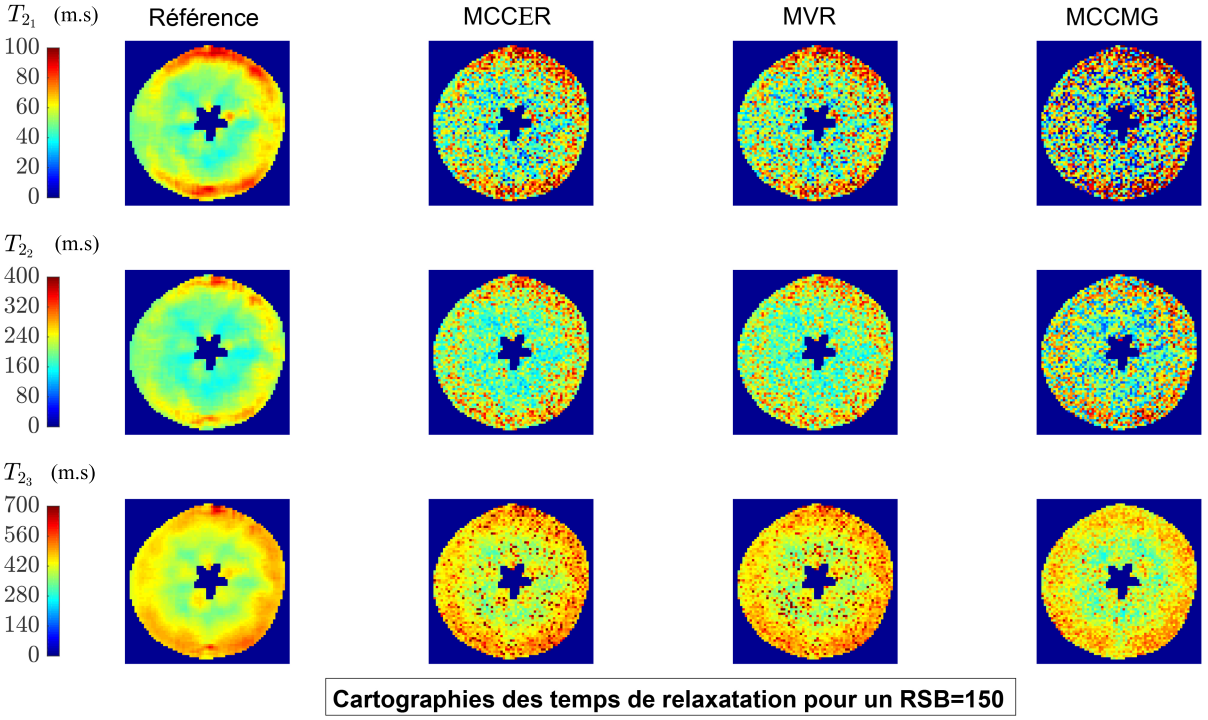

Figure III.6 – Résultats de l'estimation sur l'OV2. La première colonne représente les cartographies des temps de relaxation T_2 de référence utilisée pour la simulation. La deuxième, la troisième et la quatrième colonne représentent respectivement les cartographies des T_2 reconstruites en utilisant les méthodes de MCCER, MVR, MCCMG.

IV

Estimation conjointe des paramètres à l'échelle de l'image

Nous avons proposé dans le chapitre précédent une méthode d'estimation des paramètres qui prend en considération la nature ricienne du bruit. Nous avons démontré l'importance de l'incorporation de cette information sur la précision de l'estimation des paramètres. Néanmoins, même avec l'augmentation de précision l'erreur reste toujours importante, et cela malgré le RSB élevé. De plus, afin d'obtenir un RSB élevé il faut un nombre d'accumulations élevé cela nécessite un trop long temps d'acquisition. De plus, nous souhaitons réduire le temps de calcul. Nous proposons une alternative à l'estimation séquentielle en développant une méthode d'estimation conjointe pour, à la fois, aboutir à une diminution du temps de calcul et incorporer l'information de la régularité spatiale des paramètres.
Il n'existe pas, à notre connaissance, une approche qui prenne en compte en même temps le modèle multi-exponentiel de décroissance, la nature ricienne du bruit et la régularité spatiale des paramètres. De plus, toutes ces méthodes proposent une estimation voxel par voxel, ou bien sur des petits patchs de l'image. Cela nous a motivés pour développer une méthode qui pourrait résoudre le problème d'estimation au niveau de l'image toute entière en prenant compte des différents aspects des images en IRM.
Dans ce chapitre, nous présentons notre approche pour minimiser les critères d'une façon conjointe sur toute l'image et pour prendre la régularité spatiale des paramètres en considération. Le premier problème qui sera abordé, est celui de l'estimation conjointe des paramètres sur tous les voxels. Pour cela, nous allons apporter plusieurs adaptations à l'algorithme de Levenberg-Marquardt afin de résoudre le problème de minimisation d'un critère quadratique de grandes dimensions. Ensuite, une régularisation spatiale sera introduite et l'approche pour effectuer l'estimation d'une façon conjointe sur l'ensemble des voxels sera présentée.

IV.1 Notations

Les données IRM issues d'une étude ES-M sont stockées dans des matrices 3D de dimensions $N_X \times N_Y \times N_t$, où N_X et N_Y correspondent aux nombres de voxels suivant les axes horizontaux et verticaux respectivement et N_t représente le nombre d'échos (échantillons temporels) des signaux mesurés. Pour chaque voxel j de position donnée $[X, Y]$ à un instant τ_t le signal mesuré est défini par $y_{jt}\ \forall\ t\ \in\ 1, 2...N_t$. Comme discuté dans le chapitre précédent, nous cherchons à estimer des paramètres de décroissance multi-exponentielle d'un modèle $s_{jt}(\boldsymbol{\theta}_j)$ dans chaque voxel j de l'image. Ici $\boldsymbol{\theta}_j$ est le vecteur des paramètres dans le voxel j. Cela est assuré par minimisation d'un critère quadratique non linéaire, soit $C_j(\boldsymbol{y}_j, \boldsymbol{\theta}_j)$, à l'aide de l'algorithme LM ($\boldsymbol{y}_j$ est le vecteur contenant les mesures au voxel j pour les différents temps de mesure τ_t). La minimisation de ce critère est généralement résolue par une approche par voxel, sans prendre en considération la dimension image.

IV.2 Levenberg-Marquard global

Nous proposons d'étendre le critère afin d'obtenir un seul critère à minimiser pour tous les voxels simultanément. Cela est assuré en ajoutant une somme du critère sur tous les voxels de l'image.

$$C(\boldsymbol{\gamma}, \boldsymbol{\Theta}) = \sum_{j=1}^{N_v} C_j(\boldsymbol{y}_j, \boldsymbol{\theta}_j) \tag{IV.1}$$

$\boldsymbol{\gamma} = [\boldsymbol{y}_1, \ \ldots \ , \boldsymbol{y}_{N_v}]$ et $\boldsymbol{\Theta} = [\boldsymbol{\theta}_1 \ .. \ \boldsymbol{\theta}_{N_v}]$ représentent respectivement le vecteur des signaux mesurés et le vecteur des paramètres de modèles pour tous les voxels.

IV.2.1 Minimisation du critère global

Rappelons qu'un critère minimisé par l'algorithme de Levenberg-Marquardt doit avoir la particularité d'être reformulé de la façon suivante $C_j(\boldsymbol{y}_j, \boldsymbol{\theta}_j) = \frac{1}{2}||\boldsymbol{r}(\boldsymbol{y}_j, \boldsymbol{\theta}_j)||_2^2$. Ainsi, à chaque itération ℓ de l'algorithme de LM, les paramètres seront mis à jour suivant une direction globale $\boldsymbol{D}^{(\ell)}$:

$$\boldsymbol{\Theta}^{(\ell+1)} = \boldsymbol{\Theta}^{(\ell)} + \alpha^{(\ell)} \, \boldsymbol{D}^{(\ell)}, \tag{IV.2}$$

$\alpha^{(\ell)}$ est le pas global à prendre suivant la direction de LM.

La minimisation du critère global présente un problème de mémoire, nous proposons ainsi un algorithme de LM global (LMG) qui bénéficie de la séparabilité par voxel de la solution. En fait, dans le cas du critère global, le Jacobien a une dimension égale à $N_\tau \times (N_p N_v)$, ainsi on obtient une matrice de dimension $N_p N_v \times N_p N_v$ à inverser. Par exemple, pour une image typique de dimensions 128×128 et avec 6 paramètres à estimer, la taille de la matrice à inverser devient 98304×98304. L'inversion directe d'une telle matrice nécessite un super calculateur avec une puissance de calcul élevée. Néanmoins, si l'on observe cette matrice, la séparabilité des solutions implique que c'est une matrice bloc-diagonale. L'inversion directe d'une telle matrice est possible en conduisant l'inversion par bloc. Ainsi, le calcul de la direction sera amené en inversant N_v blocs de dimensions $N_p \times N_p$ [62]. Cette reformulation réduit considérablement l'occupation mémoire et rend le calcul de la direction globale possible avec un ordinateur de bureau.

IV.2.2 Stratégie de vectorisation

L'algorithme de minimisation globale proposé permet de diminuer le temps de calcul en adoptant des stratégies de vectorisation (valable pour les langages Matlab et Scilab). La vectorisation consiste à remplacer une tâche répétitive, normalement codée en forme des boucles 'for', par une seule opération d'addition, de multiplication ou de division entre vecteurs. Pour donner un exemple, considérons la tâche de mettre à jour, à l'itération ℓ, les valeurs des paramètres de relaxation. Dans le cas de calcul par voxel, cette opération est effectuée par l'implémentation d'une boucle 'for' qui itère sur tous les voxels et fait

avancer les paramètres de chaque voxel dans la direction correspondante. Tandis que dans une approche de minimisation globale cela pourra être fait en une seule opération de somme vectorielle (voir l'équation (IV.2)). En plus de la réduction du temps de calcul, la stratégie globale nous permet d'implémenter des stratégies à l'échelle de l'image telle que la régularisation spatiale que nous introduisons dans les sections suivantes.
Les étapes principales de l'algorithme LMG pour l'estimation des paramètres au niveau de l'image sont données par l'algorithme 2.
Dans l'étape de recherche de pas, la différence entre l'approche par voxel et l'approche

Algorithme 2: Résumé des étapes principales de l'algorithme de LMG avec la condition d'Armijo pour la recherche du pas

Données: $\boldsymbol{\Theta}^{(0)}$, $\rho \in]0\,;1[$;
for $\ell \leftarrow 0$ **to** ℓ_{max} *ou jusqu'à convergence* **do**
  Calculer $\boldsymbol{r}(\boldsymbol{\Theta}^{(\ell)})$ et $\boldsymbol{J}(\boldsymbol{\Theta}^{(\ell)})$;
  Mettre à jour $\lambda_j^{(\ell)}$;
  Calculer $\boldsymbol{D}^{(\ell)}$;
  Calculer le pas maximal $\alpha_{\max}$;
  Fixer $\alpha^{(0)} = \alpha_{\max}$;
  for $n \leftarrow 0$ **to** $n_{\max}$ *ou jusqu'à la condition d'Armijo est satisfaite* **do**
    Fixer $\alpha^{(n+1)} \leftarrow \rho\,\alpha^{(n)}$;
    Vérifier la condition d'Armijo;
  end
  Fixer $\alpha^{(\ell)} = \alpha^{(n+1)}$;
  Mettre à jour $\boldsymbol{\Theta}^{(\ell+1)} = \boldsymbol{\Theta}^{(\ell)} + \alpha^{(\ell)}\,\boldsymbol{D}^{(\ell)}$;
end
Sortie: $\boldsymbol{\Theta}^{(\ell+1)}$

globale est dans le choix d'un pas de descente global. Ainsi, dans l'algorithme LMG nous proposons de calculer un pas de descente assurant la minimisation du critère global. Par contre, dans l'approche par voxel, le pas de descente est calculé d'une façon à assurer la convergence de chaque critère associé à un voxel spécifique des différents voxels.

IV.3 État de l'art de la régularisation spatiale des paramètres de relaxation

Nous proposons d'adopter une stratégie tenant compte de la régularité spatiale dans les images IRM. Dans le domaine médical, les méthodes de régularisation spatiale ont réussi à réduire l'erreur d'estimation due au bruit de mesure [35, 29]. Nous faisons l'hypothèse qu'à l'intérieur d'un voxel les cellules végétales présentent des propriétés similaires. De même, d'un voxel à un autre, les caractéristiques des composantes subcellulaires ne varient pas d'une façon rapide. Cela se traduit par une faible variation des paramètres de relaxation d'un voxel à un autre, surtout pour les voxels appartenant au même tissu végétal. Une

méthode prenant en considération cet aspect de régularité spatiale pour diminuer l'effet du bruit consiste à ajouter une régularisation spatiale sur le critère à minimiser. Dans la littérature, cette régularisation spatiale a été ajoutée aux critères de moindres carrés basés sur la distribution gaussienne du bruit. Dans un premier temps, les différentes approches de régularisation utilisées dans la littérature seront présentées. Ensuite, nous présenterons notre proposition pour prendre en compte la régularité spatiale et aboutir à une minimisation conjointe du critère tenant compte de la distribution ricienne du bruit.

IV.3.1 Régularisation spatiale par patch

L'estimation est conduite en minimisant le critère des MC sous un modèle de décroissance multi-exponentiel discrétisé avec une régularisation de Thikonov [29]. Cette régularisation permet de régulariser le spectre discrétisé à l'intérieur de chaque voxel et ainsi réduire le bruit dans le spectre discrétisé. La régularisation spatiale est effectuée sur des patchs avec des voxels communs entre les patchs pour éviter l'effet du bord. Le choix d'effectuer l'estimation par patchs réside dans le fait que la régularisation spatiale introduit une dépendance entre les voxels et l'estimation globale n'est pas faisable d'une façon directe en raison du coût de mémoire élevé. Cette régularisation spatiale a été appliquée dans le cas d'une estimation de taux de myéline dans le cerveau. Le modèle considéré dans cet article est le modèle continu, et Z_{T_2} est construit avec 50 points logarithmiquement séparés entre $5ms$ et $600ms$ (équation (II.8)).

IV.3.2 Régularisation spatiale par cartographies filtrées

Une autre méthode de régularisation basée sur la distribution continue de T_2 a été proposée [35]. La minimisation est faite en deux étapes; la première étape est la minimisation du critère des moindres carrés, et dans la deuxième étape un filtre moyennant sur l'espace des cartographies estimées est appliqué pour avoir des cartographies spatialement régularisées. Ensuite, la différence entre cette matrice et les paramètres à estimer sera utilisée comme terme de pénalisation ajouté au critère des moindres carrés.

IV.3.3 Récapitulatifs des méthodes d'estimation de l'état de l'art

Dans le tableau IV.1, nous présentons un récapitulatif des différents articles de l'état de l'art qui abordent le problème d'estimation des paramètres d'un modèle exponentiel avec les différentes problématiques qu'ils traitent.

Ces méthodes ne proposent pas une estimation conjointe des paramètres au niveau de l'image. De plus, les méthodes proposant une régularisation spatiale se basent sur

Table IV.1 – Tableau résumant les différents articles de la littérature qui traitent les problématiques adressées dans cette thèse

Articles	Multi-exponentiel	Bruit ricien	Régularisation spatiale	Estimation conjointe
[29, 25, 63, 35, 64, 32]	✓	X	✓	X
[14, 13]	✓	✓	X	X
[46, 54, 65, 38]	X	✓	X	X
[26, 28, 36]	✓	X	X	X

l'hypothèse gaussienne du bruit. Dans la partie suivante, nous proposerons une approche de minimisation globale, tenant compte de la régularité spatiale et du modèle ricien en même temps.

IV.4 Algorithme de minimisation conjointe prenant compte la régularité spatiale

Afin de prendre en compte la régularité spatiale d'une façon directe, nous proposons d'ajouter aux critères à minimiser pour l'estimation des paramètres un critère de régularisation pénalisant la forte variation entre un voxel et ses voxels voisins. Alors, pour chaque voxel j, on définit un ensemble de voxels voisins $\mathcal{V}_j$ (par exemple une fenêtre de taille 3×3 voxels) et nous proposons d'imposer une pénalité sur la différence entre les paramètres telle que:

$$R(\boldsymbol{\Theta}) = \sum_{p=1}^{N_p} \beta_p \sum_{j=1}^{N_v} \sum_{k \in \mathcal{V}_j} \psi\left(\theta_j\left(p\right) - \theta_k\left(p\right)\right), \tag{IV.3}$$

avec ψ une fonction de pénalisation et $(\beta_1, \ldots, \beta_{N_p})$ un ensemble de paramètres de régularisation permettant d'ajuster l'influence de la régularisation sur les paramètres estimés. De plus, afin d'établir un équilibre sur le poids de la régularisation pour les différents paramètres, un paramètre β_p plus grand pourra être choisi pour les paramètres avec une faible amplitude. Le paramètre de régularisation introduit un couplage entre les voxels de l'image. Ainsi, la solution d'un voxel dépend d'une façon directe de la solution de ses voxels voisins et d'une façon indirecte des autres voxels. Dans l'algorithme de LM cela aboutit à une matrice à inverser qui n'est plus bloc diagonale et ainsi l'inversion par bloc n'est plus possible. Cette étape de régularisation nécessite de découper l'image en des patchs. Dans la partie suivante, nous présenterons une approche de séparation des variables entre les voxels et ainsi une minimisation conjointe.

IV.4.1 Minimisation par majoration du critère de régularisation

Afin de résoudre le problème de coût de mémoire élevé causé par la dépendance inter-voxel des solutions, nous proposons d'adopter une technique permettant l'estimation conjointe de tous les paramètres sans diviser l'image en patchs. Cette approche est basée sur l'algorithme MM introduit dans la partie III.3.2.
En se basant sur la propriété de convexité de la fonction de pénalisation spatiale ψ, une fonction majorante séparable pourra être introduite [66]. Pour toute variable (u, v) et à chaque itération m de l'algorithme MM la fonction majorante est donnée par :

$$\begin{aligned}\psi\left(u-v\right) &= \psi\left(\frac{1}{2}\Big(2u-\Big(u^{(m)}+v^{(m)}\Big)\Big)+\frac{1}{2}\Big(-2v+\Big(u^{(m)}+v^{(m)}\Big)\Big)\right)\\ &\leq \frac{1}{2}\psi\left(2u-\Omega_{uv}^{(m)}\right)+\frac{1}{2}\psi\left(2v-\Omega_{uv}^{(m)}\right),\end{aligned} \tag{IV.4}$$

avec $\Omega_{uv}^{(m)} = \left(u^{(m)}+v^{(m)}\right)$. La fonction de majoration est ainsi une somme de deux fonctions convexes séparables. En appliquant cette stratégie de régularisation sur le critère de régularisation globale le critère majorant suivant est obtenu :

$$M_{Reg}^{(m)}\left(\boldsymbol{\Theta}\right) = \sum_{p=1}^{N_p}\beta_p\sum_{j=1}^{N_v}\sum_{k\in\mathcal{V}_j}\psi\left(2\theta_j\left(p\right)-\delta_{jk}^{(m)}(p)\right), \tag{IV.5}$$

avec $\delta_{jk}^{(m)}(p) = \left(\theta_j^{(m)}\left(p\right)+\theta_k^{(m)}\left(p\right)\right)$ et $\theta_j^{(m)}(p)$ représente la valeur du paramètre p du voxel j à l'itération m de l'algorithme MM. Si une fonction de régularisation quadratique ψ est utilisée, la fonction majorante $M_{reg}^{(m)}$ sera aussi quadratique, dans les autres cas nous introduisons une reformulation afin d'obtenir un critère quadratique non linéaire:

$$M_{Reg}^{(m)}\left(\boldsymbol{\Theta}\right) = \sum_{p=1}^{N_p}\sum_{j=1}^{N_v}\sum_{k\in\mathcal{V}_j}\left(\beta_p^{\frac{1}{2}}\psi^{\frac{1}{2}}\left(2\theta_j\left(p\right)-\delta_{jk}^{(m)}(p)\right)\right)^2, \tag{IV.6}$$

sous condition que ψ soit strictement positive ce qui est le cas des fonctions de régularisations classiques [67]. En combinant les critères des moindres carrés (inclus le critère M_{MVR}) avec le critère de majoration M_{Reg}, l'algorithme MM pour l'approche régularisée est donné par l'algorithme 3.

IV.4.2 Paramètres de régularisation spatiale

Un choix important qui affecte la précision d'estimation est le choix des paramètres de régularisation spatiale; deux choix sont à faire:

- Le choix de la fonction de pénalisation ψ
- L'ajustement du vecteur du poids de régularisation $\boldsymbol{\beta}$

Le choix de la fonction de pénalisation affecte l'importance de la variation des paramètres entre les différents tissus (ou zones ayant des paramètres proches) tandis que le choix

Algorithme 3: L'algorithme MM pour l'approche pénalisée

Initialiser $\boldsymbol{\Theta}^{(0)}$;
for $m \leftarrow 0$ **to** $m_{\max}$ *ou jusqu'à convergence* **do**
  calculer $\delta_{jk}^{(m)}(p)$ pour tout j et k ;
  Définir $M_{Reg}^{m}(\boldsymbol{\Theta})$;
  if *MVR* **then**
    Calculer $\tilde{y}_{jt}\left(\boldsymbol{\theta}^{(m)}\right) = R\Big(z_{jt}\left(\boldsymbol{\theta}^{(m)}\right)\Big)y_{jt}$;
    Fixer $M_{Q}^{(m)}(\boldsymbol{\gamma}, \boldsymbol{\Theta}) = M_{MVR}^{(m)}(\boldsymbol{\gamma}, \boldsymbol{\Theta})$;
  end
  if *MCCER* **then**
    Fixer $M_{Q}^{(m)}(\boldsymbol{\gamma}, \boldsymbol{\Theta}) = C_{MCCER}(\boldsymbol{\gamma}, \boldsymbol{\Theta})$
  end
  Minimiser $M_{Q}^{(m)}(\boldsymbol{\gamma}, \boldsymbol{\Theta}) + M_{Reg}^{(m)}(\boldsymbol{\gamma}, \boldsymbol{\Theta})$ en utilisant l'algorithme 2 avec les paramètres suivants :
    - Fixer le nombre maximal d'itérations LM à une petite valeur;
    - Initialiser : $\boldsymbol{\Theta}^{(m)}$;
    - Sortie : $\boldsymbol{\Theta}^{(m+1)}$;
end
Output $\boldsymbol{\Theta}^{(m+1)}$

du vecteur de pénalisation affecte le degré de régularisation sur l'image. Dans ce travail, nous avons adopté une approche de régularisation quadratique en utilisant la fonction de pénalisation suivante:

$$\psi(x) = x^2 \tag{IV.7}$$

Néanmoins d'autres fonctions pourraient être considérées; une régularisation l2-l1 qui pénalise moins les variations brusques entre les tissus.

$$\psi(x) = \sqrt{x^2 + \delta^2} \tag{IV.8}$$

En outre, nous avons choisi un poids de pénalisation $\boldsymbol{\beta} = [0.4,\ 0.4,\ 0.06,\ 0.06,\ 0.01,\ 0.01]$. Ce choix a été fait par essai-erreur afin de garantir une bonne séparation entre les tissus. Si l'ordre de grandeur des couples de paramètres $[A_{0_c}, T_{2_c}]$ est proche, le même paramètre de régularisation sera choisi pour chaque couple de paramètres. Cependant, d'autres possibilités pour les valeurs de poids existent. Si l'ordre de grandeur des paramètres est différent, un poids différent pour chaque paramètre doit être considéré. Une autre stratégie est d'affecter le même poids pour tous les paramètres.

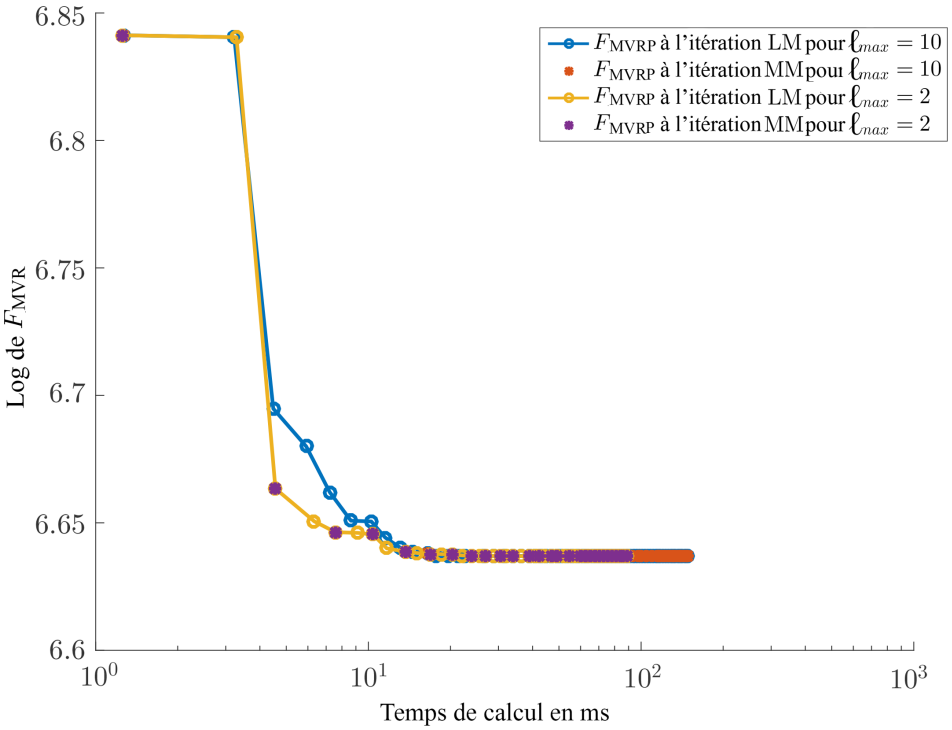

Figure IV.1 – Évolution du critère de MVRP au cours des itérations MM (points en orange et en mauves) et des itérations LMG internes (points en bleu et en jaune).

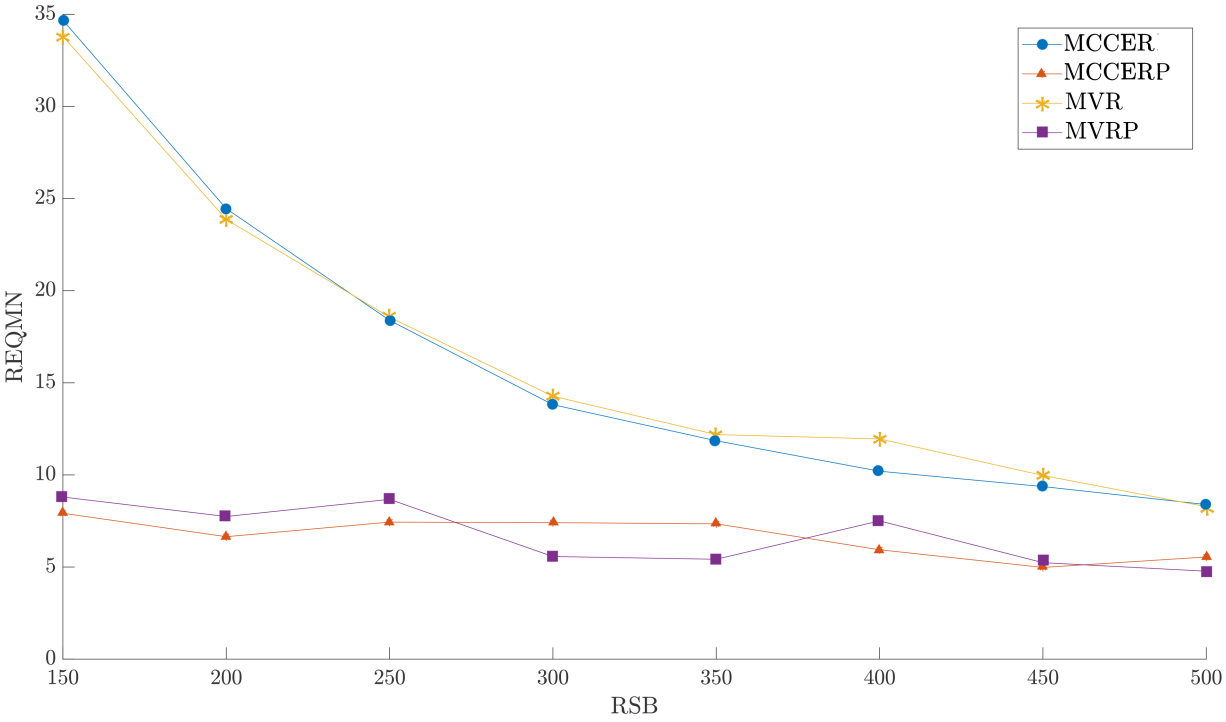

Figure IV.2 – REQMN pour les 4 approches; MCCERR, MCCERP, MVR, MVRP pour différents RSB.

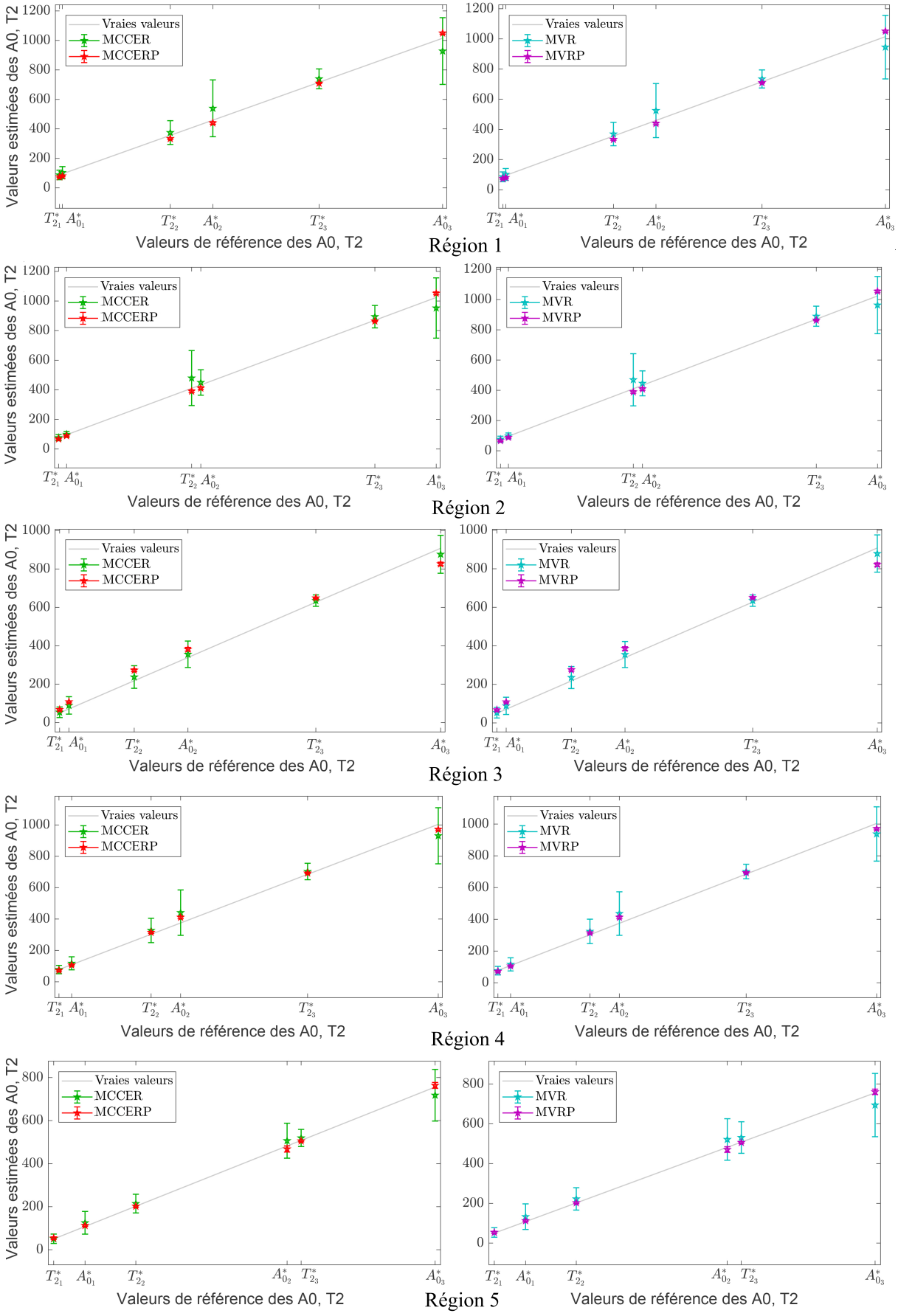

Figure IV.3 – Moyenne et écart-type des paramètres estimés dans les différentes régions du fantôme pour un RSB=150.

IV.5 Résultats

Dans cette partie nous présenterons les résultats de l'algorithme MM pour l'approche pénalisée donnée dans l'algorithme 3. Dans un premier temps, nous présenterons une analyse empirique sur la convergence de l'algorithme, ensuite nous présenterons les résultats numériques et visuels de l'estimation des paramètres sur les objets virtuels présentés dans le chapitre précédent (OV1 et OV2) et sur des images expérimentales IRM acquises sur une tomate (IE1).
Nous avons comparé, en termes d'erreur d'estimation et de biais/variance, la minimisation des quatre critères suivants:

- Critère de Maximum de vraisemblance $C_{MVR}(\boldsymbol{\gamma}, \boldsymbol{\Theta})$ (MVR).
- Critère des moindres carrés corrigé par l'espérance ricienne $C_{MCCER}(\boldsymbol{\gamma}, \boldsymbol{\Theta})$ (MCCER).
- Critère de Maximum de vraisemblance pénalisé (MVRP):

$$F_{MVRP} = C_{MVR}(\boldsymbol{\gamma}, \boldsymbol{\Theta}) + R(\boldsymbol{\Theta}) \tag{IV.9}$$

- Critère des moindres carrés corrigé par l'espérance ricienne pénalisée (MCCER) :

$$F_{MCCERP} = C_{MCCER}(\boldsymbol{\gamma}, \boldsymbol{\Theta}) + R(\boldsymbol{\Theta}) \tag{IV.10}$$

IV.5.1 Analyse empirique de convergence

Avant d'étudier la performance des algorithmes d'estimation, nous avons procédé à une étape de validation de convergence de l'algorithme régularisé MVRP. En fait, la minimisation d'un critère majorant doit amener la minimisation du critère initial dit majoré et cela dépend de l'exactitude du choix de la fonction majorante et de l'algorithme de minimisation. L'évolution du critère de maximum de vraisemblance régularisé F_{MVRP}, pour une fonction de régularisation quadratique, est présentée dans la figure IV.1. C'est bien évident qu'en minimisant le critère majorant le critère de MVRP est minimisé et par conséquent le choix de la fonction majorante est validé.

De plus, une étude sur le choix du nombre d'itérations internes de l'algorithme de LMG a été fait. Dans la même figure IV.1, nous montrons une comparaison entre le choix d'un petit nombre d'itérations internes ($\ell_{\max} = 2$) et entre le choix d'un grand nombre ($\ell_{\max} = 10$). Ce choix n'affecte pas la convergence globale de l'algorithme, et par conséquent les résultats, mais par contre cela impacte sa vitesse de convergence. Nous constatons qu'en choisissant un grand nombre d'itérations internes l'algorithme converge avec moins d'itérations MM mais avec un temps de convergence plus élevé et ainsi nous avons fixé le nombre d'itérations interne à 2.

IV.5.2 Analyse sur le choix des poids de régularisation

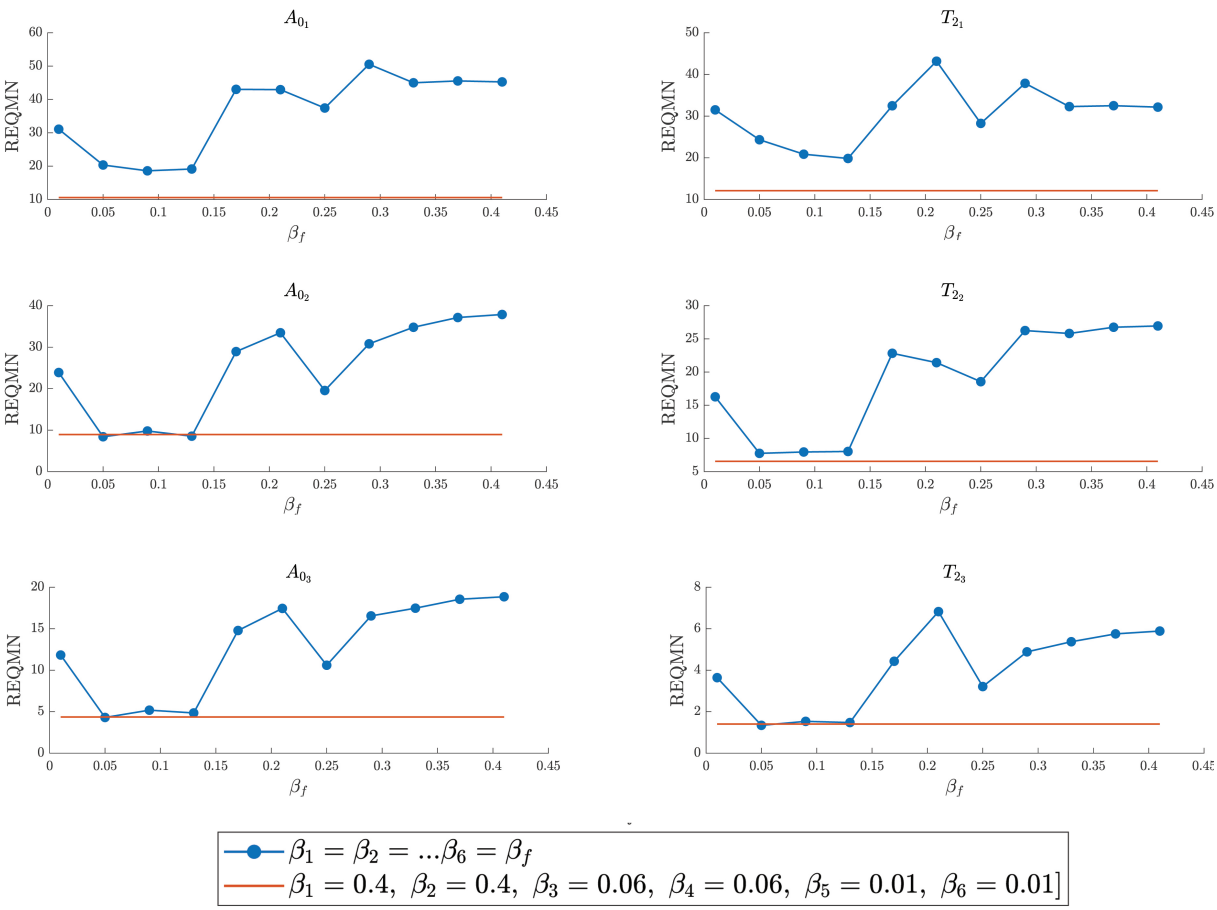

Figure IV.4 – Comparaison entre l'utilisation d'un poids de régularisation unique, β_f, pour tous les paramètres (courbe en bleu) et un poids de régularisation différent par paramètres (courbe en rouge).

Pour démontrer l'importance du choix d'un poids de régularisation différent pour les paramètres de différents ordres de grandeur, nous avons effectué des tests sur l'objet virtuel OV1 (avec un $RSB = 150$) en utilisant la méthode MVRP. Pour la stratégie où un poids de régularisation unique, β_f, pour tous les paramètres a été choisi, la valeur de l'erreur en termes de REQMN était la plus faible pour les valeur de β_f entre 0.05 et 0.14 (voir figure IV.4). La différence entre les deux stratégies est plus évidente pour les paramètres A_{0_1} et T_{2_1}, en effet, la REQMN pour la composante $c = 1$ est plus importante que dans le cas où les poids de régularisation ont été calibrés pour chaque paramètre. Cela est dû à la différence de l'ordre de grandeur entre les paramètres des différentes composantes c.

IV.5.3 Temps de convergence

Pour valider le fait que l'estimation conjointe présente un intérêt dans la minimisation du temps de calcul, nous avons testé, sur le même objet virtuel OV1 (avec un $RSB = 150$), l'algorithme MVR en comparant l'algorithme LM à l'algorithme LMG. Les deux algorithmes

ont abouti aux mêmes résultats, la méthode itérative a convergé avec un temps de calcul égal à $207s$ tandis que l'approche conjointe a convergé en $71s$.

IV.5.4 Résultats de la minimisation des critères régularisés sur l'OV1

Dans cette partie, les paramètres ont été estimés sur les images simulées à partir de l'objet virtuel OV1 en utilisant les méthodes MCCER, MCCERP, MVR et MVRP. La première étude consiste à évaluer les performances pour différents niveaux de bruit avec un RSB variant de 150 à 500. Une réduction importante d'erreur d'estimation est obtenue en estimant les paramètres par minimisation des critères régularisés pour les différents RSB (voir figure IV.2)). Si l'on prend par exemple la valeur de l'erreur pour le RSB le plus faible, une réduction de l'ordre de 24% est obtenue avec une erreur de 8% pour la version régularisée. Les deux versions régularisées ont conduit à des précisions très proches et l'estimation est très stable pour les différents RSB.

Dans la deuxième étude, nous détaillons les résultats en termes de biais d'estimation et de variance pour chaque région. Dans la figure IV.3, nous présentons, dans chaque région de l'OV1, la variance et la valeur moyenne des paramètres estimés. La valeur moyenne est comparée aux vraies valeurs et ainsi le biais est représenté par la distance de la valeur moyenne aux vraies valeurs. Étant donné que l'OV1 est constitué de régions uniformes, il est intéressant d'observer l'effet de la régularisation dans chaque région, et pour les différents paramètres. La première observation est la diminution importante de la variation en adoptant l'approche régularisée (voir figure IV.3). Étant donné que la variance des paramètres de référence à l'intérieur de chaque classe est nulle, une variance plus petite implique que l'algorithme est plus précis. En outre, la régularisation a abouti à un biais d'estimation plus faible sauf pour la région 3 (le disque central). Cela pourrait être dû à l'effet de la régularisation quadratique sur de fortes variations entre les paramètres des voxels voisins, ce qui est le cas de la variation entre les paramètres du disque central et son entourage, plus précisément pour les A_0 des composantes de relaxation rapide et moyenne.

En ce qui concerne le temps de calcul, l'algorithme MVRP a été plus rapide à converger que l'algorithme MCCERP. Sur un ordinateur de bureau à configuration classique sous Matlab, l'algorithme MVRP a convergé en $84s$ tandis que le MCCER a pris $251s$ pour converger (18144 paramètres à estimer).

Pour évaluer visuellement la reconstruction, nous présentons les cartographies reconstruites dans le cas du plus bas RSB (150) (voir figures IV.5 et IV.6). L'impact positif de la régularisation est clairement montré ; les cartographies estimées présentent beaucoup moins de bruit et la séparation entre les structures de l'OV1 est beaucoup plus claire. De plus, la différence entre les algorithmes MCCERP et MVRP est visuellement insignifiante. Ainsi, pour éviter la redondance dans les figures, nous montrerons dans la suite les résultats de l'algorithme MVRP, qui a convergé plus rapidement.

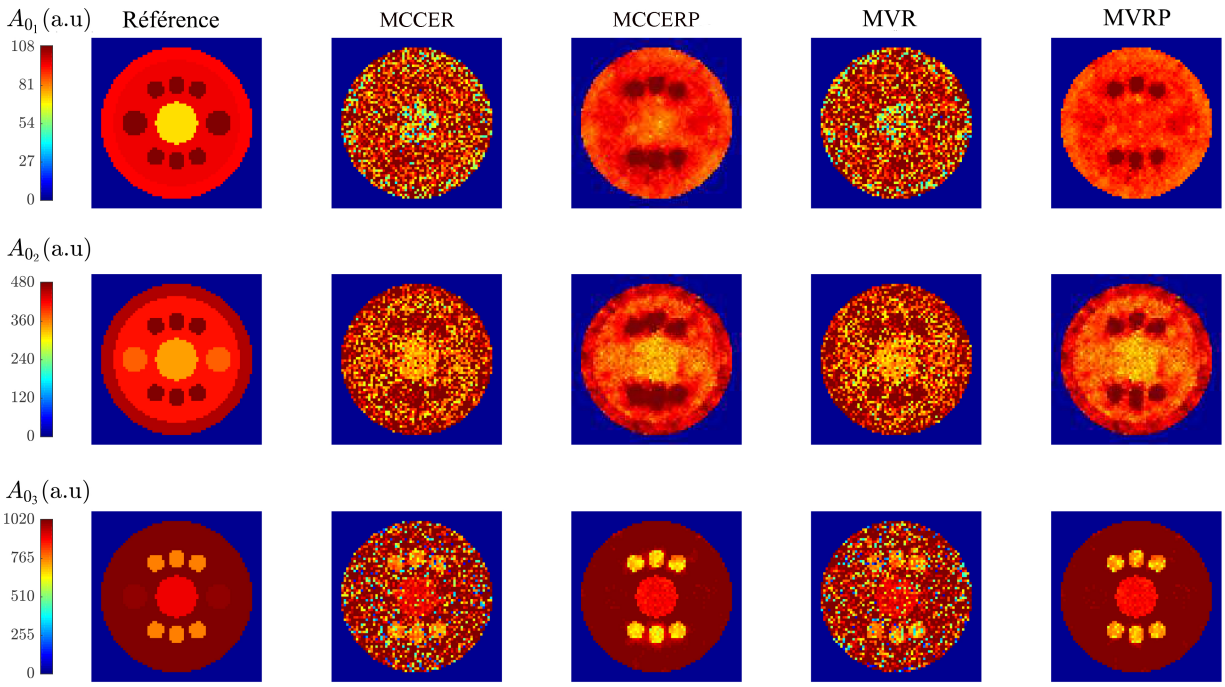

Figure IV.5 – Cartographies d'intensités reconstruites pour les composantes à décroissance rapide (première ligne), à décroissance moyenne (deuxième ligne) et à décroissance lente (troisième ligne). Les cartographies de référence (première colonne) sont comparées aux cartographies estimées par MCCER (deuxième colonne), MCCERP (troisième colonne) et MVRP (quatrième colonne).

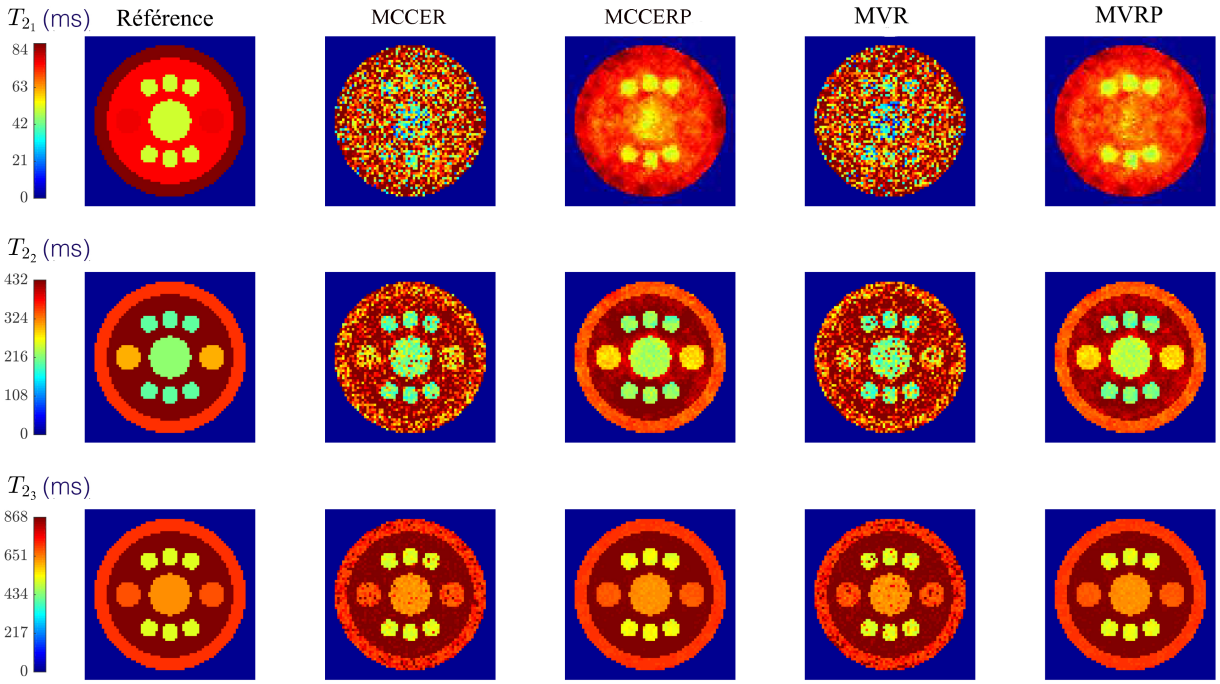

Figure IV.6 – Cartographies de temps de relaxation reconstruites pour les composantes à décroissance rapide (première ligne), à décroissance moyenne (deuxième ligne) et à décroissance lente (troisième ligne). Les cartographies de référence (première colonne) sont comparées aux cartographies estimées par MCCER (deuxième colonne), MCCERP (troisième colonne) et MVRP (quatrième colonne).

IV.5.5 Résultats sur l'OV2

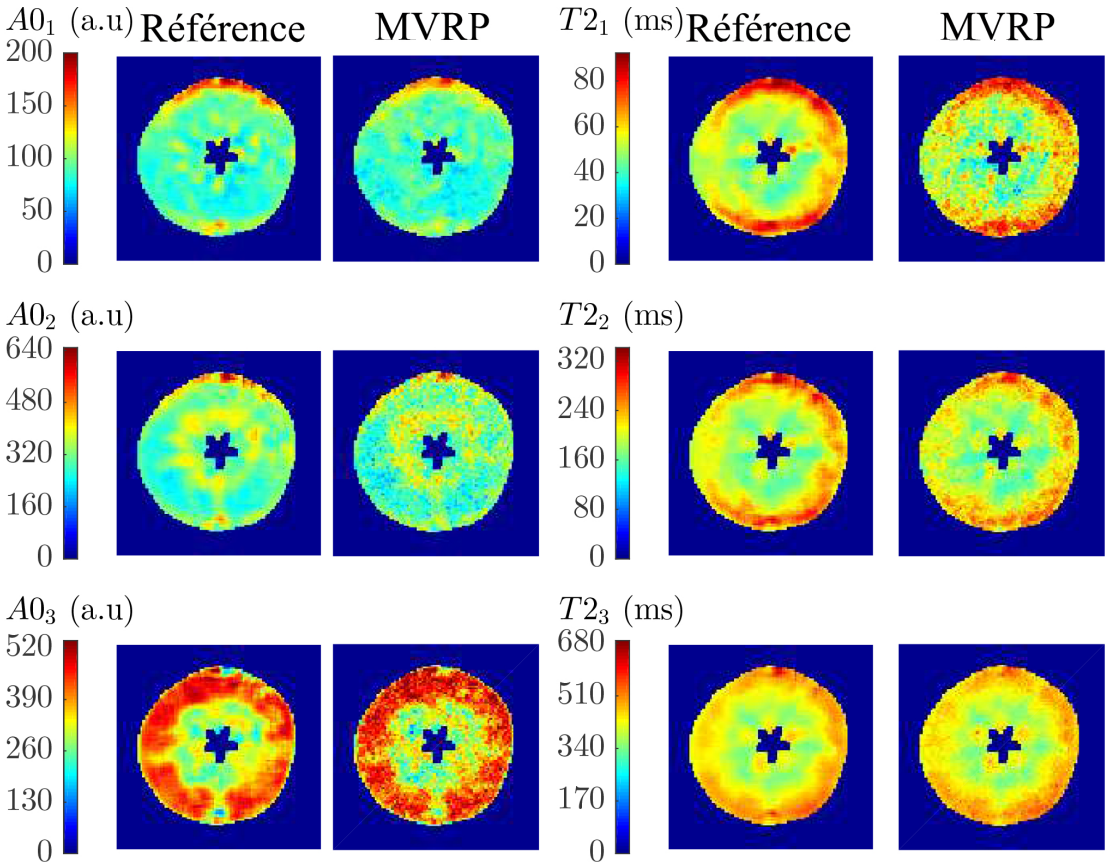

Figure IV.7 – Résultats de l'algorithme MVRP sur l'OV2

De plus, nous avons testé l'algorithme sur l'OV2 avec un RSB égal à 150. Nous avons ainsi obtenu un REQMN faible égal à 11%. Cela montre que dans des conditions de distribution de paramètres très proches de celles rencontrées dans les vraies images IRM, l'algorithme a pu aboutir à une très bonne estimation des paramètres. De plus, la facilité d'identifier les différents tissus de la pomme est observée dans la figure IV.7.

En outre, l'algorithme a donné des résultats visuellement plus satisfaisants que dans le cas du fantôme et cela peut être dû au fait que la variation entre les tissus est plus lisse dans le cas de la pomme virtuelle. Cette variation lisse est normalement rencontrée dans les études sur des fruits.

IV.5.6 Régularisation l2-l1 sur l'OV1

Afin d'aboutir à une meilleure estimation dans le cas de l'OV1 nous avons testé l'algorithme MVRP avec une fonction de régularisation l2-l1 (voir l'équation (IV.8)). La valeur du vecteur de poids de relaxation est fixée à $\boldsymbol{\beta} = [0.4,\ 0.4,\ 0.06,\ 0.06,\ 0.01,\ 0.01]$, et nous avons pris une valeur de δ de l'équation égale à 100 (même valeur pour toutes les composantes). La régularisation l2-l1 a permis d'identifier des structures dans l'OV1 tel le disque central pour le couple de paramètres $[A_{0_1}, T_{2_1}]$ (voir figure IV.8). Néanmoins,

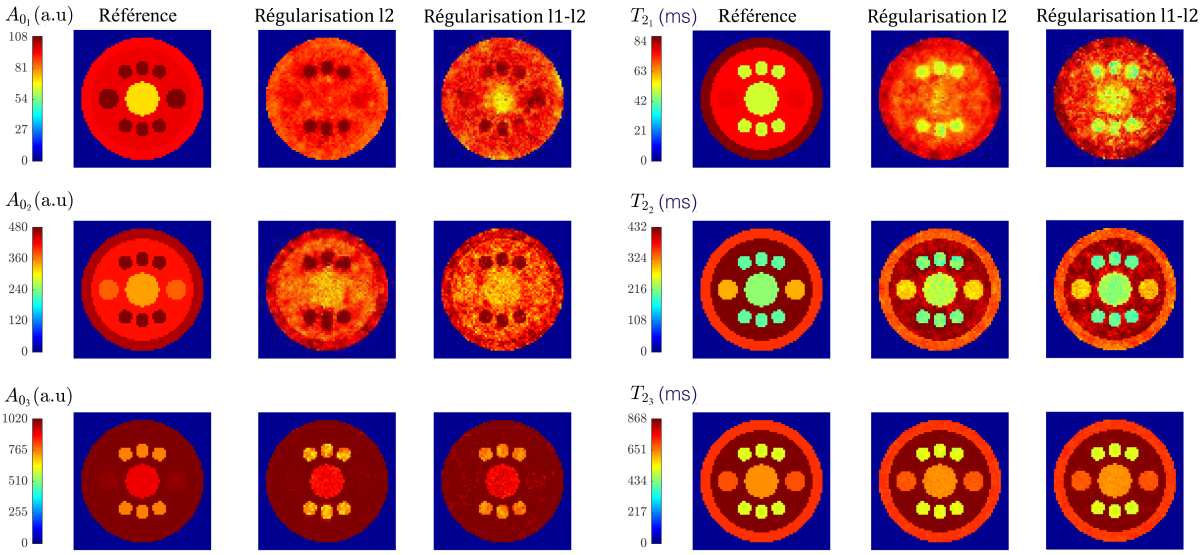

Figure IV.8 – Régularisation l2-l1 en utilisant la méthode MVRP avec une valeur de $\delta = 100$

l'erreur en termes de REQMN (égale à 8%) est égale à celle obtenue avec une régularisation quadratique. Cela est dû au fait que les images sont moins régularisées en utilisant cette stratégie (on peut observer un effet de bruitage dans les cartographies estimées).

IV.5.7 Résultats sur les données expérimentales

Pour valider l'algorithme sur des données expérimentales, nous avons utilisé des données acquises sur une tomate.

Deux ensembles d'images IRM ont été acquis avec les paramètres d'acquisition de la section III.5.2; le premier avec 32 répétitions afin d'obtenir un RSB élevé (687) qui servira de référence et le second avec une seule répétition afin de tester la robustesse de l'algorithme sur le RSB le plus faible (145). Dans la figure IV.9 nous montrons les résultats obtenus en appliquant l'algorithme MVRP sur des données acquises avec une seule répétition (RSB faible) qui seront comparés aux résultats de même algorithme appliqué aux données acquises avec 32 acquisitions. Visuellement, une très petite différence entre les deux cartographies reconstruites est présente. De plus, un REQMN de 6% a été obtenu entre les deux reconstructions. Cela valide la conclusion que la précision de l'algorithme dépend faiblement du RSB. Ainsi, nous pourrons réduire le temps d'acquisition en acquérant les données en une seule répétition. Enfin, nous avons testé la régularisation l2-l1 sur la tomate avec 1 acquisition. Dans la figure IV.10, nous montrons les résultats sans régularisation ($\beta_p = 0 \ \forall \ p$), avec une régularisation l2 et avec une régularisation l2-l1. La première déduction c'est que sans régularisation les structures de la tomate ne sont pas identifiables et les cartographies sont noyées dans le bruit. En plus, la régularisation l2-l1 permet de gagner plus de précision sur les structures fines de la tomate.

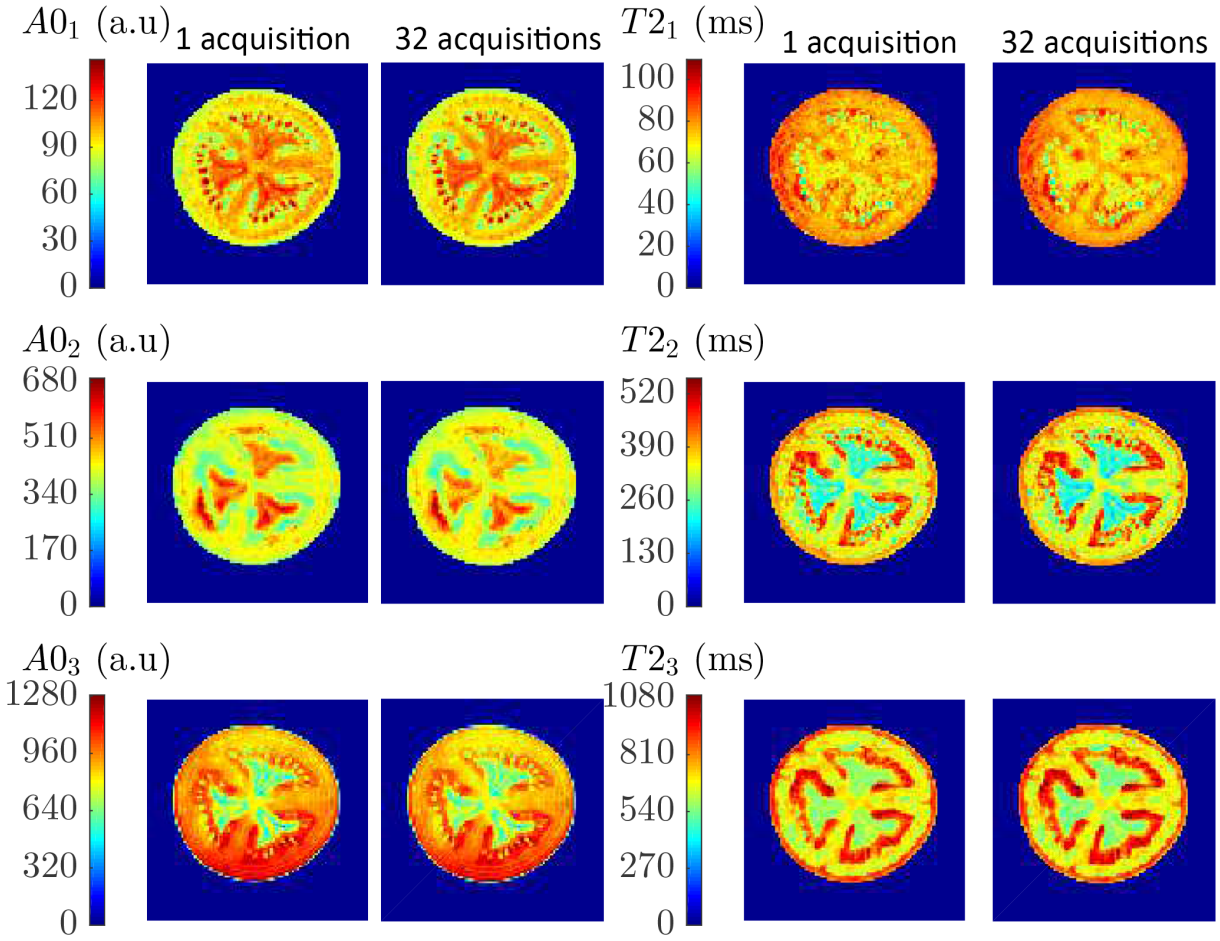

Figure IV.9 – Résultats de l'algorithme MVRP sur l'IE1. Comparaison entre régularisation sur 1 acquisition et 32 acquisitions

IV.6 Conclusion

Dans ce chapitre nous avons présenté une méthode d'estimation conjointe prenant en compte la distribution ricienne du bruit et la régularité spatiale des paramètres. L'estimation conjointe a permis de diminuer le temps de calcul d'une façon importante et a permis d'introduire la régularisation spatiale d'une façon réussie. Cette méthode d'estimation conjointe est basée sur l'algorithme de LM auquel nous avons amené des modifications afin de minimiser un critère global d'une façon réussite. Ensuite nous avons présenté notre approche d'estimation basée sur une minimisation d'un critère de maximisation de la vraisemblance ricienne couplée à un critère de pénalisation favorisant la régularité spatiale. La minimisation d'un tel critère n'est pas réalisable d'une façon directe, cela nous a poussé à proposer une méthode de minimisation afin de résoudre les problématiques suivantes de la dépendance inter-voxel de la solution du critère pénalisé et de la minimisation à grande échelle d'une façon à estimer les paramètres au niveau de l'image. Nous avons ainsi développé un algorithme basé sur une approche de minimisation par majoration, ramenant le problème à une minimisation d'un critère quadratique séparable par voxel. L'algorithme de minimisation est basé sur l'algorithme de LM avec des

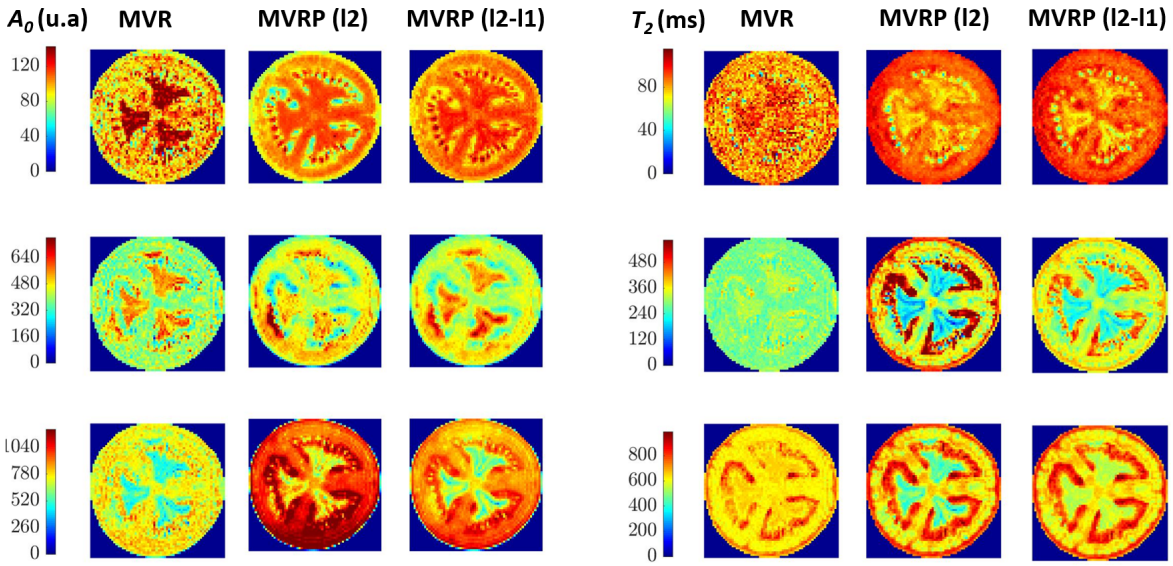

Figure IV.10 – Résultats de l'algorithme MVRP sur l'OV2. Comparaison entre la reconstruction sans régularisation (colonne 1), en utilisant une régularisation l2 (colonne 2), avec régularisation l2-l1 (colonne 3).

adaptations pour résoudre les différentes problématiques de minimisation de grande dimension.

Nous avons montré, numériquement et visuellement, que la méthode d'estimation proposée aboutie à une précision satisfaisante sur différentes données simulées et synthétiques. Cela a ouvert la porte sur plusieurs pistes dont la détection automatique des tissus qui sera présentée dans le chapitre suivant.

V

Exploitation des cartographies Multi-exponentielles

Dans le chapitre IV, nous avons introduit une méthode de reconstruction des cartographies des paramètres d'un modèle de relaxation multi-exponentiel. Cette méthode permet de reconstruire N_c cartographies de temps de relaxation T_2 et N_c cartographies d'amplitudes A_0 où chaque cartographie contient N_v voxels. Dans notre cas d'étude, caractérisation des tissus végétaux et plus précisément des fruits, le but de la reconstruction est l'exploitation de ces cartographies pour effectuer des études longitudinales. Néanmoins, la quantité d'information dans les différentes cartographies rend l'exploitation difficile. Nous souhaitons aborder le problème en divisant les paramètres estimés en groupes homogènes dans l'espace des paramètres estimés où chaque groupe correspond à un tissu végétal avec des caractéristiques uniques. À notre connaissance, cette problématique n'a jamais été traitée dans le contexte de la relaxométrie IRM. Afin d'aboutir à ce résultat, nous proposons de reformuler le problème en un problème de partitionnement d'un groupe d'observation. Ainsi, chaque voxel j des cartographies estimées est considéré comme étant une observation et les paramètres $\boldsymbol{\theta}_j$ sont les descripteurs de chaque observation. Pour le modèle multi-exponentiel nous obtenons une matrice d'observation à partitionner définie par :

$$x = \begin{bmatrix} A_{0_{1,1}} & T_{2_{1,1}} & \cdots & A_{0_{Nc,1}} & T_{2_{Nc,1}} \\ A_{0_{1,2}} & T_{2_{1,2}} & \cdots & A_{0_{Nc,2}} & T_{2_{Nc,2}} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ A_{0_{1,Nv}} & T_{2_{1,Nv}} & \cdots & A_{0_{Nc,Nv}} & T_{2_{Nc,Nv}} \end{bmatrix} \tag{V.1}$$

Pour rappel $[A_{0_{c,j}},\ T_{2_{c,j}}]$ est le couple de paramètres A_0 et T_2 pour la composante c de la courbe de décroissance multi-exponentielle au voxel j. Ainsi, chaque ligne de la matrice x représente une observation dans un espace de dimension $2N_c$. Notre but est de partitionner l'image IRM en se servant de cette matrice d'observation en N_{par} zones présentant des paramètres de relaxation proches. Un algorithme de partitionnement doit minimiser l'inertie intraclasse définie par :

$$W = \sum_{i=1}^{N_{par}} W_i \tag{V.2}$$

avec W_i défini par la somme des distances entre tous les points d'une classe i au centroïde de cette classe μ_i.

Nous proposons une approche basée sur des algorithmes de classification supervisée ou non-supervisée [68, 69, 70, 71]. Ce sont des algorithmes itératifs dans lesquels une solution de partitionnement satisfaisante, mais pas nécessairement optimale, est obtenue. Les résultats de ces algorithmes dépendent des données et de l'espace des descripteurs. Dans ce chapitre, nous présenterons des algorithmes de classification, des méthodes de projection de

l'espace des descripteurs et une approche de visualisation des distributions des paramètres dans chaque partition de l'image pour faciliter l'exploitation.

V.1 Algorithmes de classification

Dans les études visant un partitionnement des données, il existe deux approches; supervisée et non-supervisée. Dans cette partie nous présentons les différents algorithmes de partitionnement.

V.1.1 Classification supervisée

L'approche de classification supervisée nécessite des données étiquetées sur lesquelles un apprentissage sera fait pour détecter un modèle représentatif de chaque classe des données. Ces données ne sont pas disponibles dans le cadre de l'étude que nous présentons et ainsi nous avons proposé un outil graphique facilitant à un expert l'étiquetage, sur une image de contraste élevé, de quelques points représentatifs des zones à identifier.

Une fois ces données étiquetées, un algorithme de k plus proches voisins (K-NN pour 'k neighrest neighboors' en anglais) est utilisé pour attribuer les points, un par un, aux zones choisies par l'expert [68]. Dans l'algorithme K-NN, k correspond au nombre de voisins à considérer, ainsi la première étape consiste à calculer la distance entre chaque point $\boldsymbol{\theta}_j$ à classifier et tous les points étiquetés. Ensuite, les k plus proches voisins seront choisis. Enfin, le point $\boldsymbol{\theta}_j$ sera attribué à la classe qui a le plus de points parmi ces k plus proches.

V.1.2 Algorithme de classification non-supervisée

Les algorithmes de classification non-supervisée permettent de classifier d'une façon automatique des données en se basant sur des statistiques de similarités sur le plan de distance ou de probabilité d'appartenance à une classe. Ces algorithmes sont utilisés pour le partitionnement des données non étiquetées. En outre, nous avons intérêt à considérer des algorithmes aveugles à l'information spatiale (ne pas inclure une information sur la position) afin de pouvoir classifier des tissus similaires mais discontinus spatialement. Dans cette partie, nous présenterons un rappel sur les différentes méthodes de classification non supervisée. Trois algorithmes ont été étudiés, celui de K-moyennes [69], de modèle de mélange gaussien (MMG) [70], et de la classification hiérarchique [71].

V.1.2.1 K-moyennes

L'algorithme de K-moyennes est très largement utilisé pour résoudre les problèmes de partitionnement. Cet algorithme est basé sur un calcul de la distance (dans l'espace dans lequel les points sont représentés) entre les différents points à des points de référence représentant les moyennes des différentes classes. L'utilisateur a, dans le cas de cet

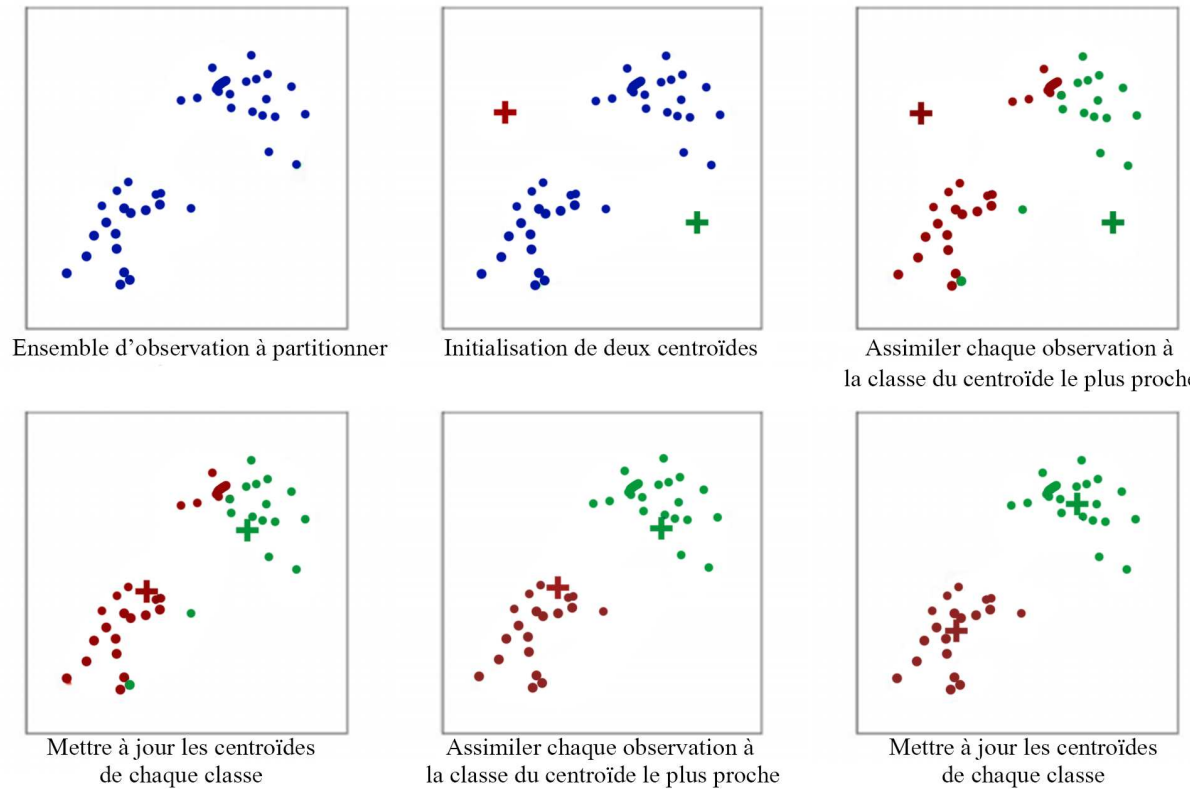

Figure V.1 – Exemple de partitionnement en utilisant l'algorithme K-moyennes.

algorithme, la responsabilité du choix de nombre de partitions N_{par} et de la fonction de mesure de la distance (euclidienne, Manhattan, Mahalanobis ... [72]). Néanmoins, il existe des méthodes d'estimation de nombre de partitions dont la méthode dite "de coude" qui est présentée dans la section V.2.2.

Les étapes principales de l'algorithme sont (voir figure V.1):

1. Initialiser N_{par} centroïdes ($\boldsymbol{\mu}_1$, $\boldsymbol{\mu}_2$, ... $\boldsymbol{\mu}_{N_{par}}$) dans l'espace des paramètres d'une façon arbitraire.
2. Mesurer la distance entre chaque observation $\boldsymbol{\theta}_j$ et les différents centroïdes.
3. Assimiler chaque observation $\boldsymbol{\theta}_j$ à la classe correspondant au centroïde $\boldsymbol{\mu}_i$ le plus proche.
4. Calculer la valeur moyenne des points dans chaque classe et mettre à jour les centroïdes ($\boldsymbol{\mu}_1$, $\boldsymbol{\mu}_2$, ... $\boldsymbol{\mu}_{N_{par}}$).
5. Répéter les étapes 2 à 4 jusqu'à ce que la variation des valeurs des centroïdes devienne faible.

Dans la première itération de cet algorithme, N_{par} points seront choisis soit d'une façon arbitraire soit avec des valeurs introduites par l'utilisateur. Dans notre cas, le choix sera fait d'une façon arbitraire, cela implique que si on lance l'algorithme plusieurs fois, plusieurs résultats sont obtenus. On parle alors de minima locaux dans l'algorithme de classification. Pour remédier à ce problème, l'algorithme est généralement lancé plusieurs fois (par exemple 60 fois) et la valeur qui aboutit à la plus petite inertie W sera retenue [73].

V.1.2.2 Modèle de mélange gaussien

L'algorithme de K-moyennes est basé sur une mesure de distances des différentes observations à un centroïde. Cela implique que les observations sont partitionnées dans des hypersphères, cependant les distributions des données dans différentes classes peuvent avoir des formes elliptiques. Ces hyperellipses peuvent être représentées par une distribution gaussienne avec une moyenne égale au centroïde de la classe et une matrice de variance qui dépend de la forme de la dispersion des observations. Pour prendre cette forme de dispersion en considération, le deuxième algorithme de partitionnement que nous présentons est celui basé sur un modèle de mélange gaussien. Il est souvent utilisé dans le cas où les observations sont considérées être issues d'un mélange de distributions gaussiennes, où chaque classe de point est représentée par une distribution gaussienne.

Dans cet algorithme, la première étape consiste à estimer à partir des observations un ensemble de distribution de probabilités gaussiennes. Pour cela, un algorithme d'entropie maximale est utilisé, il permet d'estimer les moyennes ($\boldsymbol{\mu}_1,\ \boldsymbol{\mu}_2,\ \ldots\ \boldsymbol{\mu}_{N_{par}}$), les matrices de covariance ($\boldsymbol{\Sigma}_1,\ \boldsymbol{\Sigma}_2,\ \ldots\ \boldsymbol{\Sigma}_{N_{par}}$) et la probabilité qu'un point soit attribué a une classe ($\boldsymbol{p}_1,\ \boldsymbol{p}_2,\ \ldots\ \boldsymbol{p}_{N_{par}}$) dans le cadre d'un mélange de N_{par} gaussienne. Ensuite, chaque point de vecteur de paramètres $\boldsymbol{\theta}_j$ sera affecté la classe k pour laquelle la probabilité $\mathbb{N}(\boldsymbol{\theta}_j|\mu_k, \Sigma_k)p_k$ est maximale.

De même que l'algorithme de K-moyenne, l'algorithme MMG nécessite une connaissance a priori sur le nombre de classes N_{par}.

V.1.2.3 Algorithme de classification hiérarchique

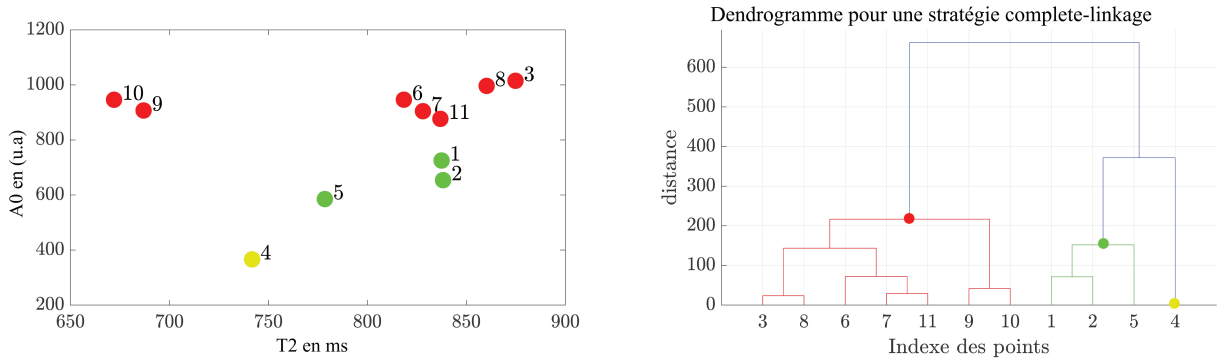

Figure V.2 – Dendrogramme tracé à partir de 11 points dans un espace de descripteurs définis par $[A_{0_3}, T_{2_3}]$

La classification hiérarchique est une méthode de partitionnement de données par regroupement itératif basée sur une distance de similarité. Le principal avantage de cette technique est de pouvoir estimer un nombre de partitions N_{par} d'une façon automatique. Dans cet algorithme, la première étape consiste à mesurer la distance entre toutes les paires des points de l'espace d'observation. Au début de l'algorithme, chaque point est considéré comme appartenant à une classe différente des autres points. Ensuite, la paire

des classes ayant la distance la plus faible fusionne pour former une nouvelle classe. Cette étape est répétée jusqu'à ce que tous les points de l'espace fusionnent dans une seule classe. Chaque étape de fusion sera représentée par une branche avec une longueur qui dépend de la distance entre les classes. Ainsi, deux classes avec une grande similarité seront fusionnées par une petite branche tandis que deux classes avec une grande distance seront fusionnées par une longue branche. Un dendrogramme sera ainsi formé (voir figure V.2).

Dans le cas de la classification hiérarchique, un choix sur la mesure entre les classes doit être fait. Il existe plusieurs stratégies dont les plus importantes sont :

- single-linkage : dans cette stratégie, la distance entre les points les plus proches de chaque paire de classes est retenue. Cela aboutit à un grand nombre de branches.
- complete-linkage : la distance entre les points les plus éloignés de chaque paire de classes est retenue. Cela favorise la fusion des classes avec le moins de points.
- average-linkage : l'average-linking permet d'avoir une mesure moyenne. Ainsi, la distance moyenne entre tous les points de chaque paire de classes sera retenue.

Étant donné qu'on ne cherche pas à trouver des points aberrants, nous avons intérêt à avoir des classes avec un nombre de points élevés et ainsi la stratégie de complete-linkage sera adoptée.

En outre, il existe, à partir du dendrogramme, plusieurs méthodes d'estimation de nombre de classes. La méthode que nous proposons d'utiliser est de découper le dendrogramme au point où la longueur de la branche devient inférieure à la moitié de la longueur de la branche la plus longue.

V.2 Paramètres de l'algorithme

Le résultat de tout algorithme de partitionnement dépend de l'espace dans lequel les paramètres sont présentés. Le choix de l'espace doit prendre en considération la nature des données et les informations qu'on connait a priori sur les descripteurs. En outre, pour les algorithmes de classification non hiérarchique, un choix doit être fait sur le nombre de classes. Dans cette section nous présenterons notre choix de l'espace de descripteurs et une approche pour le choix du nombre de classes.

V.2.1 Espace de descripteurs

Pour classifier les données d'une façon réussie, il est nécessaire de bien définir l'espace des descripteurs. Nous proposons dans ce travail d'amener la classification sur l'espace des paramètres estimés. Nous présenterons dans cette partie deux aspects importants pour le choix de l'espace de descripteurs ; la normalisation et la correction de la variation spatiale des paramètres au sein d'un tissus supposé homogène due aux imperfections de l'antenne RF de réception.

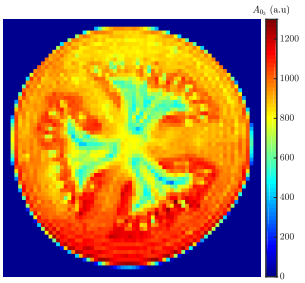

Figure V.3 – Variation spatiale de A_0 au sein d'un tissus supposé homogène (exemple du péricarpe équatorial) due aux imperfections de l'antenne RF de réception.

V.2.1.1 Artéfacts de l'antenne de mesure

Dans une mesure IRM, les antennes de mesures présentent une sensibilité qui n'est pas spatialement homogène. En fait, dans la zone où le fruit est proche de l'antenne, un signal un signal sera mesuré avec une amplitude plus élevée que dans la zone où le fruit est loin de l'antenne. Dans une première approximation, ce biais est considéré multiplicatif sur le signal mesuré $y_{jt} = B_j y^*_{jt}\ \forall t$ où y^*_{jt} représente le signal en l'absence d'inhomogénéité. Ainsi ce problème n'affecte que les valeurs de l'amplitude qui sont multipliées par le même facteur ; $B_j A_{0_{c,j}}\ \forall c$.

Pour remédier à ce problème, nous proposons de remplacer les valeurs des A_{0_c} par les pourcentages en A_0 ce qui donne :

$$\tilde{A}_{0_{c,j}} = \frac{B_j A_{0_{c,j}}}{\sum_{c=1}^{N_c} B_j A_{0_{c,j}}} = \frac{A_{0_{c,j}}}{\sum_{c=1}^{N_c} A_{0_{c,j}}} \tag{V.3}$$

Ainsi, l'effet de ce biais multiplicatif n'est pas présent sur les valeurs des amplitudes en pourcentage.

V.2.1.2 Normalisation

Les vecteurs de descripteurs sont constitués de paramètres avec différents ordres de grandeur. Par exemple, pour un T_{2_1} de l'ordre de $100ms$ et un T_{2_3} de l'ordre de $600ms$, le résultat de tout algorithme de classification est 6 fois plus affecté par les T_{2_3} que par les T_{2_1}. De plus, l'étape précédente aboutit à des valeurs faibles de $\tilde{A}_{0_c}$ et ainsi leur effet sur la classification devient négligeable. Nous proposons ainsi de normaliser la matrice x. Les données sont ainsi représentées par la matrice $\hat{x}$:

$$\hat{x} = \begin{bmatrix} \frac{\tilde{A}_{0_{1,1}}}{\max_j \tilde{A}_{0_{1,j}}} & \frac{T_{2_{1,1}}}{\max_j T_{2_{1,j}}} & \cdots & \frac{\tilde{A}_{0_{N_c,1}}}{\max_j \tilde{A}_{0_{N_c,j}}} & \frac{T_{2_{N_c,1}}}{\max_j T_{2_{N_c,j}}} \\ \vdots & \vdots & \ddots & \vdots & \\ \frac{\tilde{A}_{0_{1,N_v}}}{\max_j \tilde{A}_{0_{N_c,j}}} & \frac{T_{2_{1,Nv}}}{\max_j T_{2_{1,j}}} & \cdots & \frac{\tilde{A}_{0_{N_c,N_v}}}{\max_j \tilde{A}_{0_{N_c,j}}} & \frac{T_{2_{N_c,Nv}}}{\max_j T_{2_{N_c,j}}} \end{bmatrix} \tag{V.4}$$

V.2.2 Choix de nombre de classes

Le choix du nombre des classes dans les algorithmes de K-moyennes et de MMG peut être fait d'une façon automatique. Une des méthodes d'estimation du nombre de classes la plus utilisée est la méthode de 'coude' [74]. Le but est de trouver le point N_{par} pour lequel en passant à un nombre $N_{par}+1$ le gain en inertie interclasse devient faible. Ainsi, il faut calculer, pour un nombre de classes N_{par} variant entre 1 et un nombre maximal de classes N_{parmax}, l'inertie interclasse $W_{N_{par}}$. Ensuite, la variation d'inertie en pourcentage est calculée comme suit :

$$\Delta W_{N_{par}} = \frac{W_{N_{par}} - W_{N_{par}+1}}{\max_{N_{par}}(W_{N_{par}})} \tag{V.5}$$

En traçant la somme cumulée de la variation d'inertie en fonction du nombre de classes, une forme de 'coude' apparait. Ainsi, le nombre pour lequel le gain d'inertie devient faible correspond au point pour lequel la somme cumulée devient supérieure à une valeur seuil. Ce nombre correspond au nombre de classes choisi.

V.3 Distribution interclasse

Une fois les voxels de l'image partitionnés et les différentes classes de l'objet imagé identifiées, une étape d'exploitation des paramètres estimés doit permettre d'analyser les distributions des paramètres à l'intérieur de chaque classe. Nous proposons ainsi deux approches pour visualiser les variations des paramètres à l'intérieur des classes :

- Dans la première approche, nous proposons de tracer les valeurs des amplitudes $[A_{0_1} \cdots A_{0_{N_c}}]$ en fonction des temps de relaxation $[T_{2_1} \cdots T_{2_{N_c}}]$. Ainsi, nous obtenons des nuages de points représentant la variation des paramètres suivant les axes des A_0 et des T_2 où le centroïde de chaque classe représente les valeurs typiques des temps de relaxation dans chaque classe.
- Dans la deuxième approche, nous proposons de reconstruire un spectre de temps de relaxation. Ainsi, on définit des plages de temps de relaxation dans chaque classe reparties d'une façon logarithmique entre la valeur minimale $T_{2_{min}}$ et la valeur maximale $T_{2_{max}}$. Ensuite, la somme des amplitudes dans chaque plage est calculée. La répartition logarithmique permet d'avoir des plages plus larges pour les T_2 longs et des plages plus petites pour les T_2 courts.

V.4 Validation et tests

Afin de valider et tester les différents algorithmes de classification, nous présentons dans cette section les résultats des algorithmes de classification sur l'OV1 simulé et sur des données expérimentales des fruits.

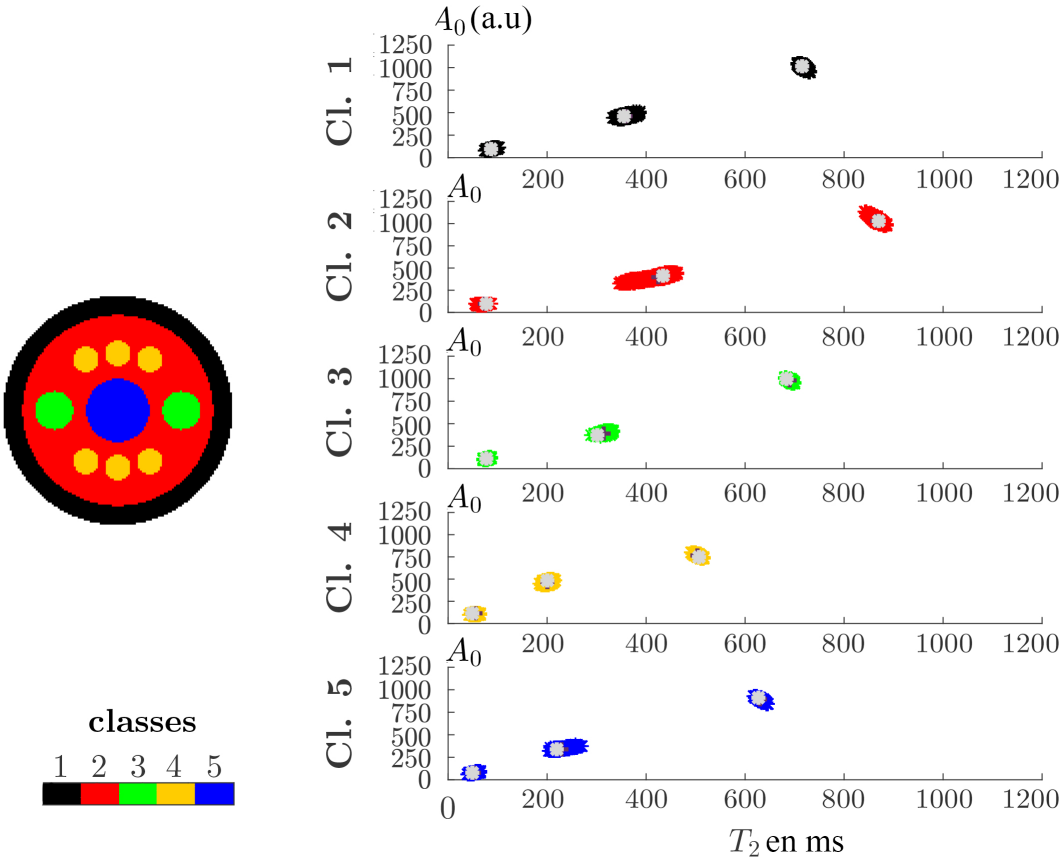

Figure V.4 – Carte de répartition spatiale des 5 classes identifiées par K-moyennes et distributions conjointes des paramètres (T_2, A_0) des trois composantes de chaque classe sur l'OV1. Les points gris représentent les valeurs utilisées pour la simulation et les croix représentent les centroïdes des classes estimées par K-moyennes.

V.4.1 Classification sur l'OV1

Nous avons utilisé les paramètres estimés à partir de l'OV1 décrit en section III.5.1 en utilisant l'approche MVRP. Le nombre de classes dans ce cas est connu, il correspond au nombre de zones dans l'OV1. Puisque ce sont des images simulées, l'inhomogénéité des A_0 causée par l'antenne de mesure n'est pas présente. Ainsi, la correction de cet artéfact présenté dans la section V.2.1.1 ne sera pas utilisée dans cette partie. Pour les différents algorithmes, la classification a été menée sur les paramètres estimés normalisés. Pour un $RSB = 150$, le partitionnement a été effectué sans aucune erreur (tous les voxels ont été attribués à leur classe initiale) pour les trois algorithmes (K-moyennes, MMG et hiérarchique). Dans la figure V.4, nous montrons les résultats de l'algorithme K-moyenne en représentant la distribution des paramètres dans chaque classe. Les croix violettes représentent la position des centroïdes et les points en gris représentent les valeurs utilisées pour la simulation. On peut remarquer que le centroïde est bien positionné et que la dispersion (correspondante à une erreur d'estimation dans ce cas) des paramètres est faible.

Ensuite, nous avons utilisé l'IE1 (images acquises sur une tomate) pour tester les différents algorithmes et présenter notre approche d'exploitation des cartographies parti-

tionnées.

V.4.2 Données expérimentales

Dans un premier temps, les algorithmes de K-moyennes, MMG, et hiérarchique ont été comparés sur les paramètres tri-exponentiels estimés à partir de l'IE1 dont les paramètres d'acquisition sont présentés dans la section III.5.2.

V.4.2.1 Classification non supervisée

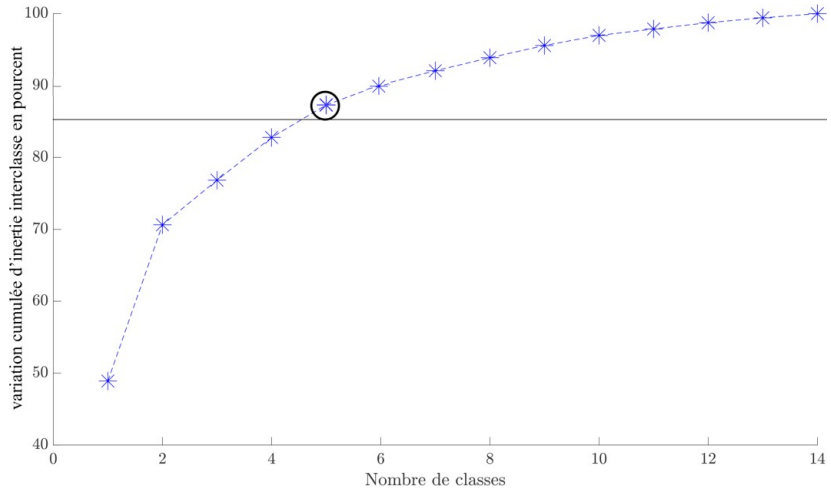

Figure V.5 – Résultat de la méthode en coude appliqué à l'IE1; le nombre de classes en abscisse et en ordonnée la valeur de la variation cumulée d'inertie ΔW en pourcentage de la valeur maximale, la valeur de seuil est fixée à 85%.

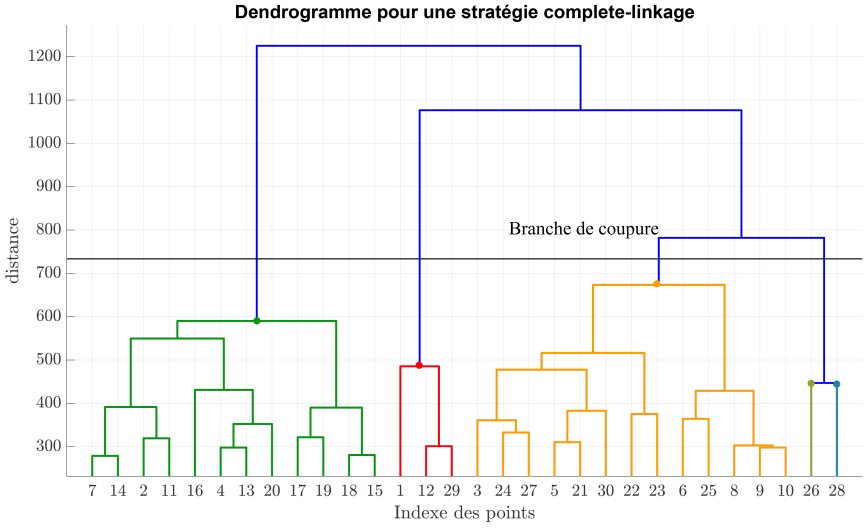

Figure V.6 – Dendrogramme avec une connectivité 'Full-linkage' appliqué à l'IE1. Les premiers points du dendrogramme représentent un ensemble de points pour éviter l'encombrement sur la figure.

En utilisant la méthode de coude, un nombre de classes égale à 5 est choisi (voir figure V.5), ce point correspond à la valeur pour laquelle l'inertie interclasse cumulée est égale à 88%. Pour l'algorithme de classification hiérarchique, le dendrogramme est coupé à la

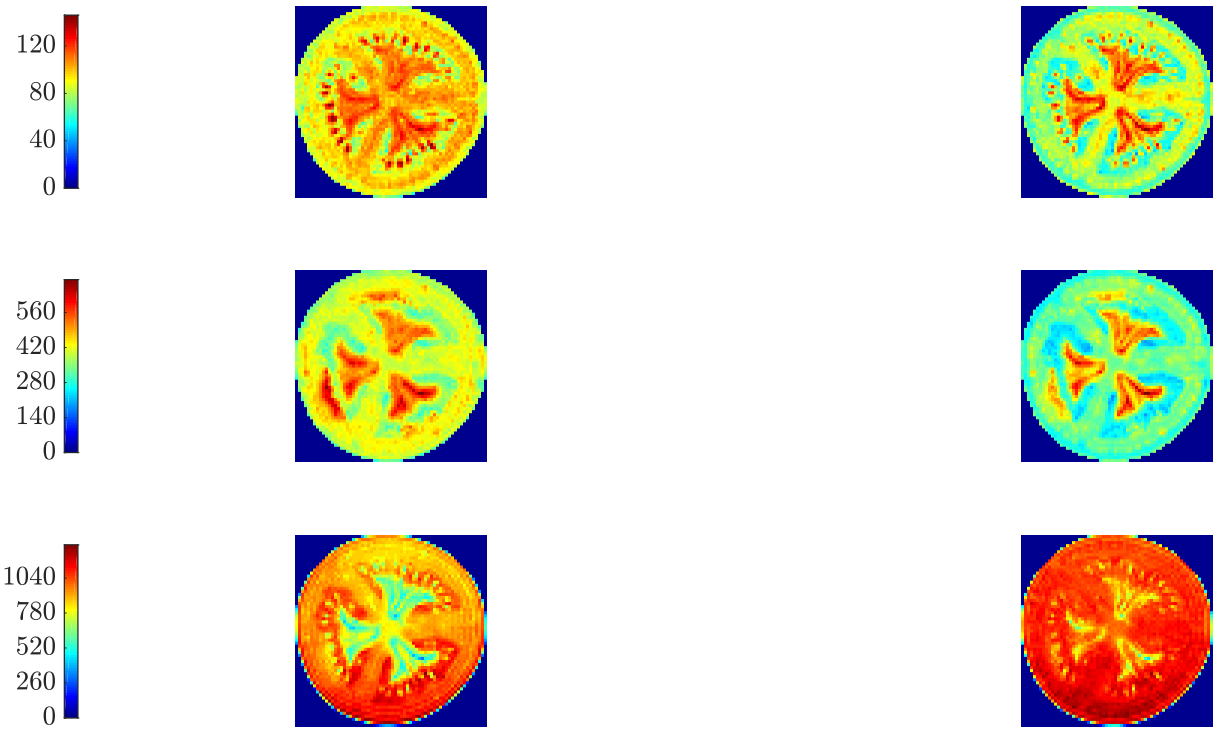

Figure V.7 – Carte de répartition spatiale des 3 classes identifiées par K-moyennes sur l'IE1 et distributions conjointes des paramètres (T_2, A_0) des trois composantes de chaque classe et les spectres de temps de relaxation dans chaque classe. Les croix représentent les centroïdes des classes estimées.

moitié de la plus grande branche aboutissant au même résultat de 5 classes (voir figure V.6).

La première observation est que, pour les trois méthodes, une classe avec un petit nombre de points a été identifiée (classe bleue dans les figures V.7, V.8 et V.9). Cette classe contient des points répartis sur l'extrémité du fruit considérés comme des volumes partiels. Ainsi, cette classe sera exclue de l'étude. Pour les autres classes, on remarque que le spectre des temps de relaxation est similaire à celui mesuré par les études de RMN. La dispersion de chaque composante de ce spectre est dû à la fois à une variation des paramètres de relaxations entre les voxels d'un tissu mais aussi au bruit de mesure. Ce résultat est très intéressant par le fait que, l'estimation par voxel en utilisant un modèle tri-exponentiel a aboutit un spectre caractérisant de différents tissus végétaux. Dans les détails, tous les algorithmes ont pu identifier au minimum trois structures de la tomate représentées par :

- Une classe rouge représentant l'union entre le cœur, le péricarpe radial et le péricarpe externe.
- Une classe verte représentant le placenta de la tomate.
- Une classe jaune représentant le tissu loculaire de la tomate.

Néanmoins, une petite variation de forme et de taille de ces composantes est observée entre les différents algorithmes. La différence majeure entre les algorithmes est dans

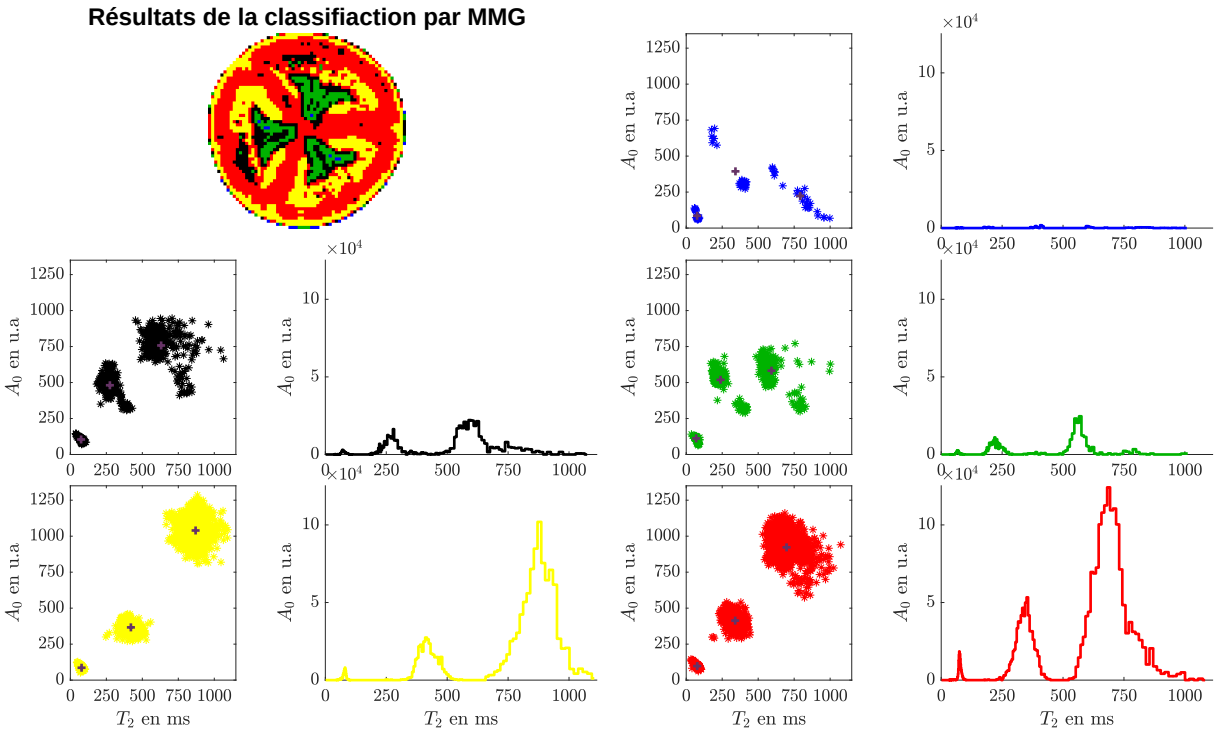

Figure V.8 – Carte de répartition spatiale des 3 classes identifiées par MMG sur l'IE1 et distributions conjointes des paramètres (T_2, A_0) des trois composantes de chaque classe et les spectres de temps de relaxation dans chaque classe. Les croix représentent les centroïdes des classes estimées.

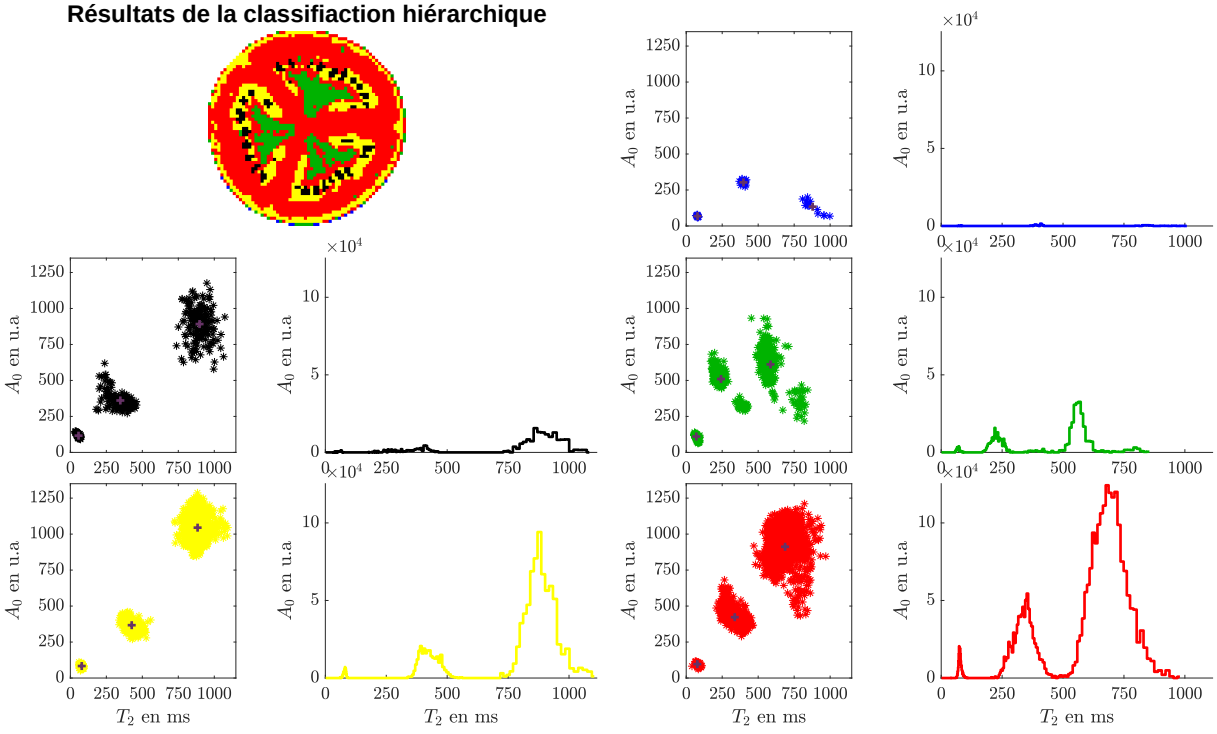

Figure V.9 – Carte de répartition spatiale des 3 classes identifiées par la classification hiérarchique sur l'IE1, les distributions conjointes des paramètres (T_2, A_0) des trois composantes de chaque classe et les spectres de temps de relaxation dans chaque classe. Les croix représentent les centroïdes des classes.

l'identification de la classe noire. En fait, l'algorithme de K-moyennes et de la classification hiérarchique ont détecté les pépins tandis que le MMG a attribué la classe noire à des changements de tissus dans la tomate et une variation de la structure du placenta. Le choix parmi ces trois méthodes n'est pas évident, et il dépend de l'étude envisagée.

V.4.2.2 Classification supervisée

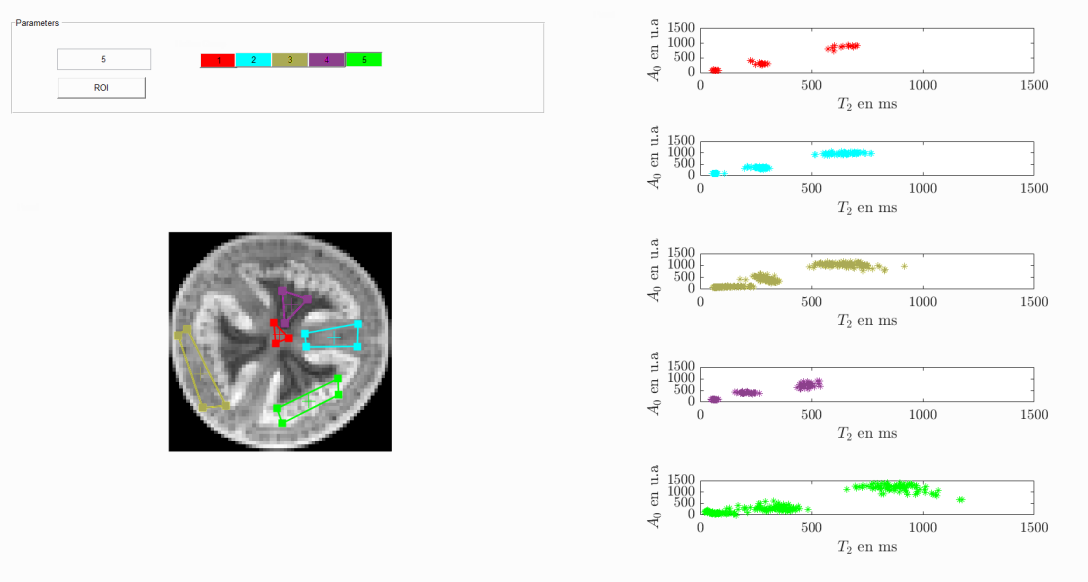

Figure V.10 – Annotation des 5 principaux tissus végétaux de la tomate en utilisant l'outil graphique

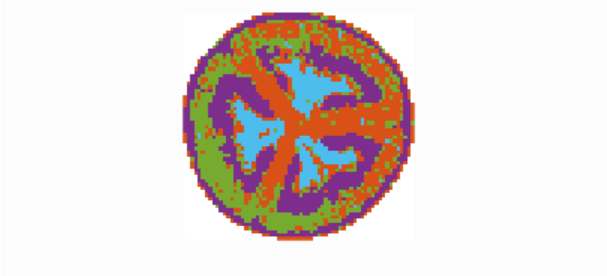

Figure V.11 – Tomate partitionnée à partir de l'algorithme KNN en utilisant des voxels étiquetés.

Dans le cas de la classification non-supervisée, une limite sur la détection apparait notamment pour le cœur, le péricarpe radial et le péricarpe externe qui ont été représentés par la même classe. ceci est probablement dû au fait que les trois tissus présentent des paramètres de relaxation T_2 très similaires. Pour vérifier cette hypothèse, nous avons appliqué la deuxième approche de partitionnement que nous proposons sur la même tomate. Cette approche est basée sur un algorithme de classification supervisée. Ainsi, la première étape consiste à annoter différentes zones d'intérêt sur une figure de fort contraste par un spécialiste en utilisant un outil graphique que nous avons développé dans ce but (voir figure V.10). Ensuite, en appliquant une classification supervisée tel l'algorithme de K-NN on

obtient une image partitionnée (voir figure V.11). En utilisant cet algorithme, seulement 4 zones ont été obtenues (une classe reste vide après convergence de l'algorithme). On remarque que même avec l'annotation, l'algorithme n'arrive pas à trouver une différence entre ces zones et ainsi ils sont regroupés dans deux zones (classe orange et classe verte). En fait, la distribution des paramètres dans les zones rouge, bleue et marron de la figure V.10 sont très proches. Ainsi, on a pu démontrer en utilisant la classification supervisée et non-supervisée que le cœur, le péricarpe radial et le péricarpe externe présentent des paramètres de relaxation transversale proches.

V.4.3 Résultats sur d'autres fruits

Afin d'éviter une redondance dans l'affichage des résultats, nous présentons dans cette partie les résultats de l'algorithme de K-moyennes. Les données ont été acquises par des mesures de relaxométrie IRM sur une carotte et une pomme avec les paramètres décrits dans la partie III.5.2. Les paramètres ont été estimés en utilisant la méthode MVRP avec un vecteur de poids de régularisation $\boldsymbol{\beta} = [0.4,\ 0.4,\ 0.06,\ 0.06,\ 0.01,\ 0.01]$. Pour les deux fruits, en utilisant une analyse de la méthode en 'coude', le nombre de classes a été fixé à 4. Sur la carotte, la classification a réussi à identifier les trois importantes structures; le cœur (classe noire) et deux zones de la chair (classes jaune et verte). La distribution montre que les amplitudes A_{0_2} sont plus grandes que celles des deux autres composantes. Sur la pomme, les trois composantes identifiées sont le cœur (classe noire), le péricarpe interne (classe verte) et le péricarpe externe (classe jaune). De plus, la distribution des paramètres montre que les composantes avec un temps de relaxation plus faible présentent des amplitudes plus faibles.

V.5 Conclusion

Nous avons présenté dans ce chapitre une approche capable de détecter pour la première fois à notre connaissance plusieurs tissus végétaux à partir de cartographies de paramètres multi-exponentiel. Cette approche facilite l'exploitation des données en résumant les cartographies par une image partitionnée en zones homogènes et différentes les unes des autres. Pour cela nous avons proposé :

- l'utilisation des algorithmes de classification non supervisée telle l'algorithme K-moyennes, MMG et la classification hiérarchique;
- de considérer les intensités en valeurs relatives pour s'affranchir de l'artéfact lié à l'antenne de mesure;
- de normaliser les paramètres pour garantir que tous les paramètres aient le même poids sur la classification;
- l'interprétation de chaque zone représentant un tissu végétal par une approche de

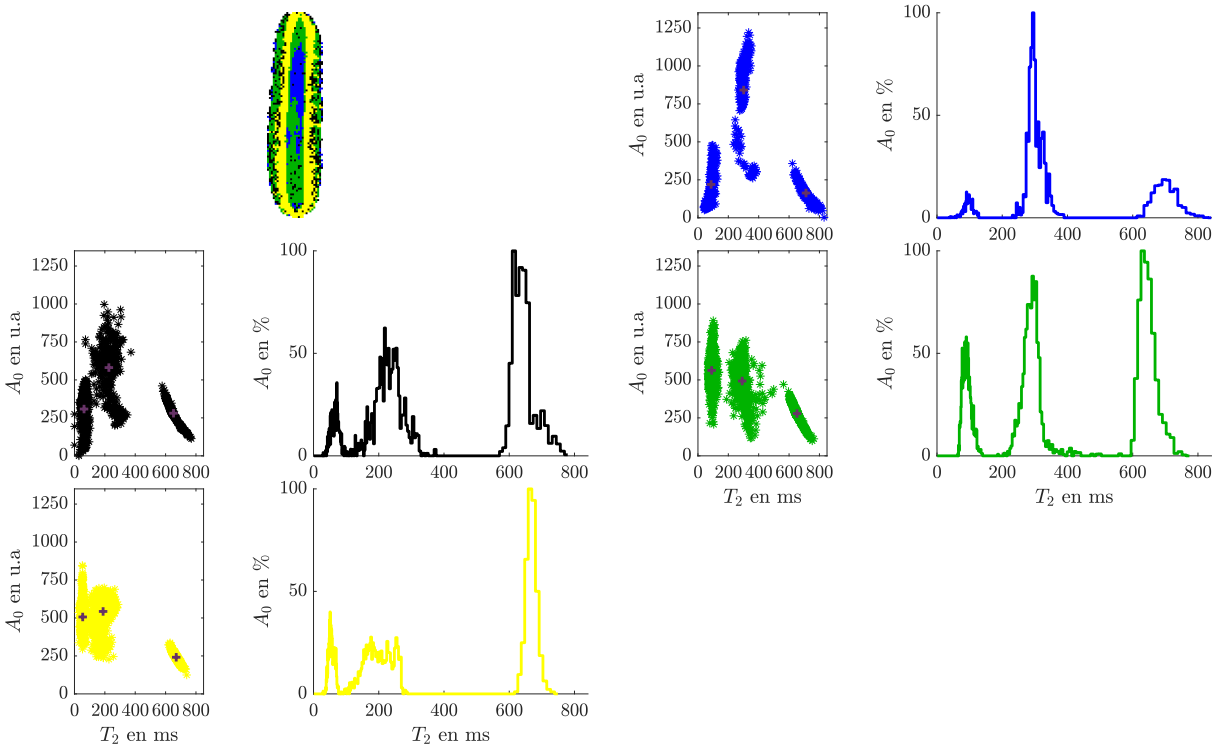

Figure V.12 – Carte de répartition spatiale des 3 classes identifiées par K-moyennes sur une carotte et distributions conjointes des paramètres (T_2, A_0) des trois composantes de chaque classe et les spectres de temps de relaxation dans chaque classe. Les croix représentent les centroïdes des classes.

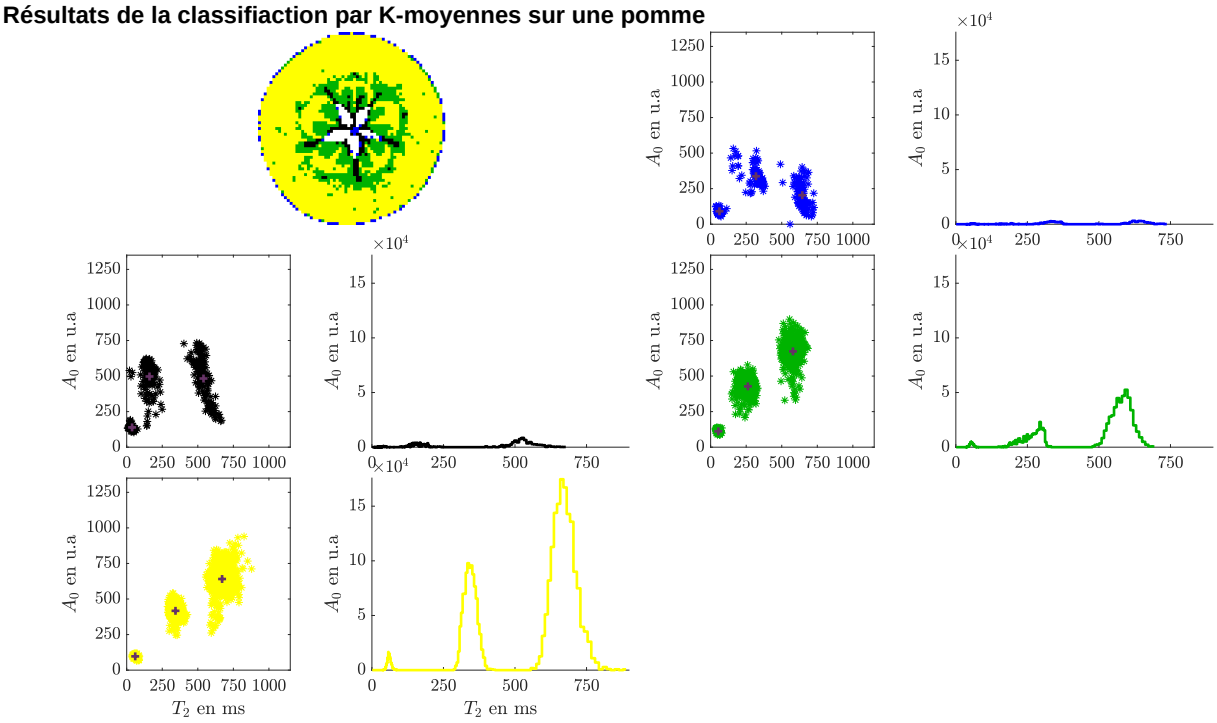

Figure V.13 – Carte de répartition spatiale des 3 classes identifiées par K-moyennes sur une pomme et distributions conjointes des paramètres (T_2, A_0) des trois composantes de chaque classe et les spectres de temps de relaxation dans chaque classe. Les croix représentent les centroïdes des classes.

visualisation de la distribution des couples de paramètres A_0 et T_2 en traçant les valeurs des amplitudes en fonctions des temps de relaxation;

- une méthode d'affichage du spectre des T_2 basée sur une discrétisation des valeurs des temps de relaxation sous une échelle logarithmique.

Nous avons validé l'algorithme sur l'OV1 et des tests sur différents fruits ont montré l'intérêt de cette approche pour la caractérisation des tissus végétaux par des spectres uniques à chaque tissu. L'exploitation de ces spectres permettra une compréhension sur les changements que subissent les tissus durant des phénomènes longitudinaux tels la maturation, la congélation et le stockage.

Conclusion générale

La relaxométrie $T2$ en IRM est une modalité de mesure permettant une acquisition des signaux de décroissances multi-exponentielles dans chaque voxel de l'image. L'estimation des paramètres de décroissance $\boldsymbol{\theta}_j$ dans chaque voxel j présente un intérêt pour la caractérisation de plusieurs phénomènes dans le domaine médical et agroalimentaire. Dans le domaine médical, on retrouve des problématiques telle l'estimation de taux de myéline dans le cerveau. Dans le domaine agroalimentaire, cadre applicatif de ce travail de thèse, ces paramètres présentent des informations permettant de caractériser différents tissus végétaux. Cependant, l'estimation se heurte à plusieurs problèmes liés au bruit de mesure et à une résolution d'un problème de grandes dimensions. De plus, l'interprétation et l'exploitation des paramètres multi-exponentiels ne sont pas évidentes au vu du nombre élevé de paramètres estimés. Dans cette thèse, nous avons présenté un algorithme permettant, en prenant en compte la nature du bruit et la régularité spatiale des paramètres, de mesurer avec une grande précision les paramètres sur l'ensemble de l'image d'une façon conjointe. De plus, nous avons proposé une méthode pour l'exploitation de ces données basée sur des algorithmes de classification supervisée et non-supervisée. Nous avons pu ainsi partitionner les images IRM en différentes zones représentant des tissus végétaux qu'on caractérise par un spectre de temps de relaxation unique.

VI.1 Contributions par chapitre

Dans le chapitre II, nous avons présenté les méthodes d'acquisition de données et les différents aspects des données mesurées. Ainsi, nous avons proposé de modéliser le signal de relaxation par un modèle à décroissance multi-exponentielle dans chaque voxel de l'image. En fait, le signal dans chaque voxel de l'image est attribué à une contribution de différents compartiments hydriques des cellules végétales présentant différents temps de relaxation. Dans le cadre applicatif de ce travail, nous avons proposé d'utiliser un modèle à décroissance tri-exponentiel. Ensuite, nous avons présenté le modèle du bruit de mesure dans les images de relaxométrie. Nous avons établi que le bruit dans les images en module suit une distribution de Rice. Cette distribution de bruit induit une divergence entre les signaux mesurés à RSB faible et le vrai signal de relaxation. Et par conséquent, cette

divergence induit un biais sur les paramètres estimés à partir des signaux de relaxation.

Dans le chapitre III, nous avons présenté différentes méthodes qui prennent en considération la distribution ricienne dans l'estimation des paramètres. Ces méthodes sont basées sur une minimisation d'un critère de moindres carrés corrigé. Levenberg-Marquardt (LM) est un algorithme de minimisation des critères de moindres carrés non linéaire souvent utilisé pour la minimisation de ces critères. Notre **première contribution** consiste dans un premier temps à proposer une approche de calcul d'un pas de descente basé sur la condition d'Armijo avec un calcul d'un pas maximal qui garantit la positivité des paramètres.

Nous avons proposé une méthode d'estimation basée sur une minimisation d'un critère de maximisation de vraisemblance ricienne. Ce critère est non quadratique et non-convexe et ainsi sa minimisation n'est pas directement possible en utilisant l'algorithme LM. Ainsi nous proposons d'adopter une approche de maximisation de ce critère par un critère quadratique. La **deuxième contribution** est de résoudre le problème de maximisation de vraisemblance ricienne par une série de minimisations approchées du critère majorant en utilisant l'algorithme LM avec un petit nombre d'itérations. Nous avons montré par des résultats sur des données simulées l'importance de la prise en compte de la distribution ricienne. Nous avons constaté que la minimisation d'un critère de moindres carrés corrigés par l'espérance ricienne et l'algorithme basé sur la maximisation de vraisemblance ricienne aboutissent aux meilleurs résultats.

Dans le chapitre IV, nous avons proposé une méthode d'estimation des paramètres d'une façon conjointe sur l'ensemble de l'image. Cette méthode est basée sur l'algorithme de LM auquel nous avons apporté des modifications pour bénéficier de la structure de l'image. Ainsi, **la troisième contribution** consiste à bénéficier du fait que les solutions sont indépendantes entre les voxels et ainsi la direction de descente globale sera amenée par une inversion de blocs de matrices de petites tailles. Cet algorithme permet de réduire le temps de calcul et en plus d'adopter des stratégies telle que la régularisation spatiale. En fait, les images IRM présentent une variation lente d'un voxel à un autre, et prendre en compte cette information aboutit à une meilleure estimation. Nous avons proposé d'intégrer cette information en ajoutant un critère pénalisant les fortes variations entre les voxels voisins. Cela implique une dépendance des solutions entre les voxels de l'image et ainsi le calcul de la direction de descente globale n'est plus possible par blocs. Pour cela, **la quatrième contribution** consiste à intégrer une approche de majoration-minimisation pour obtenir un critère majorant séparable entre les voxels de l'image. L'algorithme a été testé sur des images simulées et sur des images acquis sur une tomate. Les résultats sont satisfaisants et l'erreur d'estimation obtenue est faible.

Enfin, dans le chapitre V, nous avons présenté une méthode de partitionnement des données basée sur des algorithmes de classification supervisée et non-supervisée dans un espace défini par les paramètres estimés. Cela constitue **la cinquième contribution** de ce travail de thèse. Les méthodes de classification non-supervisée présentée dans ce chapitre

sont celui de k-moyennes, de modèle de mélange gaussien et de la classification hiérarchique. Pour la classification supervisée, nous avons développé une interface facilitant l'annotation et nous avons présenté l'algorithme de K-NN. Nous avons proposé de réduire les artéfacts causés par l'antenne de mesure en remplaçant les intensités par leurs valeurs relatives et ensuite de normaliser les données pour que tous les paramètres aient le même effet sur la classification. Enfin, nous avons montré que la classification a abouti à une détection de différents tissus dans la tomate, la carotte et la pomme. À l'intérieur de chaque classe identifiée, un spectre de temps de relaxation a été observé. Ce spectre, affecté par le bruit d'estimation, représente des distributions de temps de relaxation caractéristiques du tissu identifié.

Les résultats de cette thèse permettront d'effectuer diverses études longitudinales sur différents fruits. Par exemple, un suivi du même fruit durant la maturation, la congélation ou le stockage en mesurant la différence de taille, de temps de relaxation et des amplitudes associées du premier jour jusqu'à la fin de l'étude. De plus, les problématiques que nous avons traitées sont intéressantes pour des applications dans le domaine médical. Le type de donnée et les cas d'étude sont évidemment différents, mais la méthode développée pourra être utilisée dans ce domaine.

VI.2 Perspectives

Différentes problématiques ont été identifiées durant ce travail de thèse. Nous présentons dans cette dernière partie quelques perspectives pour des futurs travaux sur l'estimation et l'exploitation des paramètres de relaxation multi-exponentiel.

VI.2.1 Initialisation des paramètres

Nous avons opté pour une stratégie d'initialisation par expertise pour les différentes études que nous avons menées dans ce travail de thèse. Par contre, cela n'est pas toujours réalisable, en particulier pour des données traitées pour la première fois et sur lesquelles il y a un manque d'expertise. Pour cela, une stratégie d'initialisation plus complète et qui ne dépend pas de l'utilisateur complétera notre algorithme. Néanmoins, l'initialisation reste un problème ouvert, des méthodes basées sur un ajustement d'une droite au log du signal mesuré n'ont pas abouti à de meilleurs résultats qu'une initialisation experte.

VI.2.2 Choix des paramètres de régularisation

Le choix du poids de régularisation est une étape primordiale pour la précision de l'algorithme. Nous avons utilisé le même poids de régularisation pour différentes études et différents fruits et nous avons obtenu des résultats satisfaisants. Mais nous conseillons d'adapter ce paramètre visuellement pour différentes études. Par exemple, une méthode

de type EGO [75] pourra résoudre le problème en conduisant une recherche des paramètres qui aboutit à une meilleure estimation pour chaque cas d'étude.

VI.2.3 Détection automatique de nombre de composantes

Dans notre étude, nous avons choisi un modèle de décroissance tri-exponentielle, étant le modèle le plus représentatif de la cellule végétale. Cela n'est pas toujours le cas, en fait dans plusieurs tissus végétaux, un modèle bi- ou quadri-exponentiel pourra mieux représenter les variations à l'intérieur de chaque voxel. Une méthode de détection de nombre de composantes d'une façon automatique facilitera le choix du modèle de décroissance. Un modèle à distribution continue pourra résoudre ce problème surtout dans le cas où différents voxels présentent différents nombres de composantes pour le même fruit.

Annexes

A

Calcul des Jacobiens

Le calcul du Jacobien est une étape nécessaire pour l'obtention de la direction de descente de LM. Ce jacobien dépend du critère et du modèle choisi pour l'estimation des paramètres. Dans cette partie, nous considérons un modèle de décroissance c-exponentiel et le calcul du gradient sera donné pour un critère généralisé $C(\boldsymbol{y}, \boldsymbol{\theta})$. Quel que soit le critère choisi, on pourra l'écrire sous la forme suivante :

$$C(\boldsymbol{y}, \boldsymbol{\theta}) = \sum_{t=1}^{N_\tau} (\hat{y}_t - \hat{s}_t(\boldsymbol{\theta})) \tag{A.1}$$

avec les valeurs de $\hat{y}_t$ et de $\hat{s}_t(\boldsymbol{\theta})$ résumées dans le tableau A.1. On obtient ainsi le résidu $r(\boldsymbol{y}, \boldsymbol{\theta})$ avec $C(\boldsymbol{y}, \boldsymbol{\theta}) = \frac{1}{2} \|r(\boldsymbol{y}, \boldsymbol{\theta})\|_2^2$. Dans ce qui suit, une formulation sera présentée afin d'exprimer la dérivée partielle de $\hat{s}_t(\boldsymbol{\theta})$ en fonction de $s_t(\boldsymbol{\theta})$ et cela pour simplifier l'implémentation des différentes corrections. Notamment, la correction de McGibney et celle basée sur la valeur attendue ricienne nécessite cette étape.

A.1 Calcul du Jacobien dans le cas des MCCMG

Le calcul du Jacobien pour le critère de MCCMG nécessite le calcul de la dérivée reformulée de la façon suivante:

$$\begin{aligned} \frac{\partial s_t(\boldsymbol{\theta})^2}{\partial \theta(p)} &= \frac{\partial s_t(\boldsymbol{\theta})^2}{\partial s_t(\boldsymbol{\theta})} \frac{\partial s_t(\boldsymbol{\theta})}{\partial \theta(p)}, \\ &= 2 s_t(\boldsymbol{\theta}) \frac{\partial s_t(\boldsymbol{\theta})}{\partial \theta(p)}. \end{aligned} \tag{A.2}$$

Table A.1 – Tableau résumant les différents critères à minimiser

	C_{MC}	Gudbjartsson	C_{MCCMG}	C_{MCCER}	$M_{MVR}^{(m)}$
$\hat{s}_t(\boldsymbol{\theta})$	$s_t(\boldsymbol{\theta})$	$s_t(\boldsymbol{\theta})$	$s_t(\boldsymbol{\theta})^2$	$f\left(s_t\left(\boldsymbol{\theta}\right),\sigma\right)$	$s_t(\boldsymbol{\theta})$
$\hat{y}_t$	y_t	$\sqrt{y_t^2+\sigma^2}$	y_t^2	y_t	$R\left(\tilde{z}_t\right)y_t$

Cette façon d'exprimer la dérivée permet de faciliter l'implémentation, en fait il suffit maintenant de calculer la dérivée du modèle par rapport aux paramètres et de multiplier cette valeur par la valeur du modèle.

A.2 Calcul du Jacobien dans le cas des MCCER

Dans ce cas il faut considérer le moment d'ordre 1 du signal mesuré;

$$E_{Rice}[y_t] = \sigma\sqrt{\frac{\pi}{2}}e^{-\alpha_t}[(1+2\alpha_t)I_0(\alpha_t)+2\alpha_t I_1(\alpha_t)]$$

avec $\alpha_t = \frac{s_t(\boldsymbol{\theta})^2}{4\sigma^2}$. La dérivée de cette fonction par rapport aux différents paramètres peut être reformulée comme suit :

$$\frac{\partial E_{Rice}\left[y_t\right]}{\partial\boldsymbol{\theta}} = \frac{\partial E_{Rice}[y_t]}{\partial\alpha_t}\frac{\partial\alpha_t}{\partial s_t\left(\boldsymbol{\theta}\right)}\frac{\partial s_t\left(\boldsymbol{\theta}\right)}{\partial\boldsymbol{\theta}}. \tag{A.3}$$

En plus, puisque $I_0'(x) = I_1(x)$ et $I_1'(x) = I_0(x) - \frac{1}{x}I_1$, on obtient :

$$\begin{aligned}
\frac{\partial E_{Rice}\left[s_t\left(\boldsymbol{\theta}\right)\right]}{\partial\alpha_t} &= \sigma\sqrt{\frac{\pi}{2}}\Big(-e^{-\alpha_t}I_0\left(\alpha_t\right)+e^{-\alpha_t}I_1\left(\alpha_t\right)+2e^{-\alpha_t}I_0\left(\alpha_t\right)\\
&-2\alpha_t e^{-\alpha_t}I_0\left(\alpha_t\right)+2\alpha_t e^{-\alpha_t}I_1\left(\alpha_t\right)\\
&+2I_1\left(\alpha_t\right)e^{-\alpha_t}+2\alpha_t e^{-\alpha_t}\Big(I_0\left(\alpha_t\right)-\frac{1}{\alpha_t}I_1\left(\alpha_t\right)\Big)\\
&-2\alpha_t e^{-\alpha_t}I_1\left(\alpha_t\right)\Big)\\
&= \sigma\sqrt{\frac{\pi}{2}}e^{-\alpha_t}\left[I_0\left(\alpha_t\right)+I_1\left(\alpha_t\right)\right].
\end{aligned} \tag{A.4}$$

Ainsi :

$$\frac{\partial E_{Rice}[y_t]}{\partial\boldsymbol{\theta}} = \frac{s_t(\boldsymbol{\theta})}{2\sigma}\sqrt{\frac{\pi}{2}}\,e^{-\alpha_t}\times\left[I_0\left(\alpha_t\right)+I_1\left(\alpha_t\right)\right]\frac{\partial s_t\left(\boldsymbol{\theta}\right)}{\partial\theta\left(p\right)}. \tag{A.5}$$

De même que dans le cas du critère MCCMG, le calcul de la dérivée dans le cas de MCCER se ramène à la multiplication de la dérivée du modèle par rapport aux paramètres par une valeur qui dépend du modèle. Cela rend l'implémentation beaucoup plus facile en évitant un calcul compliqué de la dérivée initiale.

A.3 Calcul de la dérivée dans le cas d'un modèle multi-exponentielle

On a établi que pour les différents critères, le calcul du jacobien se ramène à un calcul de la dérivée partielle du modèle par rapport aux paramètres du modèle. Nous présenterons dans cette section cette dérivée dans le cas du modèle c-exponentiel donné par l'équation (II.10). Pour tout paramètre $\theta\,(p)$ du vecteur du paramètre

$$\boldsymbol{\theta} = [A_{0_1}\ ,\ T_{2_1}\ \ldots\ A_{0_{N_c}}\ ,\ T_{2_{N_c}}].$$

On peut noter que $\theta(p)$ correspond à une amplitude A_0 pour une valeur impaire de p et à un temps de relaxation T_2 pour une valeur paire de p.
Ainsi pour p impair :

$$\frac{\partial s_t\,(\boldsymbol{\theta})}{\partial \theta\,(p)} = e^{-\tau_t/\theta(p+1)}. \tag{A.6}$$

et pour p pair :

$$\frac{\partial s_t\,(\boldsymbol{\theta})}{\partial \theta\,(p)} = \frac{\theta\,(p-1)\,\tau_t}{\theta\,(p)^2}\,e^{-\tau_t/\theta(p)}. \tag{A.7}$$

Bibliographie

[1] J. Klostergaard, K. Parga, and R. G. Raptis, "Current and future applications of magnetic resonance imaging (MRI) to breast and ovarian cancer patient management," *Puerto Rico health sciences journal*, vol. 29, no. 3, p. 223, 2010. 19, 20

[2] H. Degani, V. Gusis, D. Weinstein, S. Fields, and S. Strano, "Mapping pathophysiological features of breast tumors by MRI at high spatial resolution," *Nature medicine*, vol. 3, no. 7, p. 780, 1997. 19

[3] S. Bakas, H. Akbari, A. Sotiras, M. Bilello, M. Rozycki, J. S. Kirby, J. B. Freymann, K. Farahani, and C. Davatzikos, "Advancing the cancer genome atlas glioma MRI collections with expert segmentation labels and radiomic features," *Scientific data*, vol. 4, p. 170117, 2017. 19

[4] P. A. Brex, O. Ciccarelli, J. I. O'Riordan, M. Sailer, A. J. Thompson, and D. H. Miller, "A longitudinal study of abnormalities on MRI and disability from multiple sclerosis," *New England Journal of Medicine*, vol. 346, no. 3, pp. 158–164, 2002. 19, 20, 21

[5] C. Rickers, N. M. Wilke, M. Jerosch-Herold, S. A. Casey, P. Panse, N. Panse, J. Weil, A. G. Zenovich, and B. J. Maron, "Utility of cardiac magnetic resonance imaging in the diagnosis of hypertrophic cardiomyopathy," *Circulation*, vol. 112, no. 6, pp. 855–861, 2005. 19

[6] J. N. Giedd, J. Blumenthal, N. O. Jeffries, F. X. Castellanos, H. Liu, A. Zijdenbos, T. Paus, A. C. Evans, and J. L. Rapoport, "Brain development during childhood and adolescence: a longitudinal MRI study," *Nature neuroscience*, vol. 2, no. 10, p. 861, 1999. 19

[7] J. Kaufmann, W. Studer, J. Link, and K. Schenker, "Study of water suction of concrete with magnetic resonance imaging methods," *Magazine of concrete research*, vol. 49, no. 180, pp. 157–166, 1997. 19

[8] M. Wagner, S. Quellec, G. Trystram, and T. Lucas, "MRI evaluation of local expansion in bread crumb during baking," *Journal of Cereal Science*, vol. 48, no. 1, pp. 213–223, 2008. 19

[9] M. Musse, S. Quellec, M. Cambert, M. Devaux, M. Lahaye, and F. Mariette, "Monitoring the postharvest ripening of tomato fruit using quantitative MRI and NMR

relaxometry," *Postharvest Biology and Technology*, vol. 53, no. 1-2, pp. 22–35, 2009. 19, 20, 39

[10] S. D. Evans, K. P. Nott, A. A. Kshirsagar, and L. D. Hall, "The effect of freezing and thawing on the magnetic resonance imaging parameters of water in beef, lamb and pork meat," *International journal of food science & technology*, vol. 33, no. 3, pp. 317–328, 1998. 19, 20

[11] H. Adriaensen, M. Musse, S. Quellec, A. Vignaud, M. Cambert, and F. Mariette, "MSE-MRI sequence optimisation for measurement of bi-and tri-exponential T_2 relaxation in a phantom and fruit," *Magnetic resonance imaging*, vol. 31, no. 10, pp. 1677–1689, 2013. 20, 36, 41

[12] H. Gudbjartsson and S. Patz, "The Rician distribution of noisy MRI data," *Magnetic Resonance in Medicine*, vol. 34, no. 6, pp. 910–914, 1995. 20, 41, 51

[13] R. Bai, C. Koay, E. Hutchinson, and P. Basser, "A framework for accurate determination of the T_2 distribution from multiple echo magnitude MRI images," *Journal of Magnetic Resonance*, vol. 244, pp. 53–63, 2014. 21, 39, 73

[14] M. Bouhrara, D. Reiter, H. Celik, J. Bonny, V. Lukas, K. Fishbein, and R. Spencer, "Incorporation of Rician noise in the analysis of biexponential transverse relaxation in cartilage using a multiple gradient echo sequence at 3 and 7 Tesla," *Magnetic Resonance in Medicine*, vol. 73, no. 1, pp. 352–366, 2015. 21, 40, 51, 52, 73

[15] S. Chatterjee, O. Commowick, O. Afacan, B. Combès, A. Kerbrat, S. Warfield, and C. Barillot, "A 3-year follow-up study of enhancing and non-enhancing multiple sclerosis (ms) lesions in ms patients demonstrating clinically isolated syndrome (cis) using a multi-compartment T_2 relaxometry (mct2) model," 2018. 21

[16] A. Lebois, *Brain microstructure mapping using quantitative and diffsusion MRI*. PhD thesis, Paris 11, 2014. 21

[17] D. A. Reiter, P.-C. Lin, K. W. Fishbein, and R. G. Spencer, "Multicomponent T_2 relaxation analysis in cartilage," *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*, vol. 61, no. 4, pp. 803–809, 2009. 21, 40

[18] G. Saab, R. T. Thompson, and G. D. Marsh, "Multicomponent T_2 relaxation of in vivo skeletal muscle," *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*, vol. 42, no. 1, pp. 150–157, 1999. 21, 40

[19] E. M. Purcell, H. C. Torrey, and R. V. Pound, "Resonance absorption by nuclear magnetic moments in a solid," *Physical review*, vol. 69, no. 1-2, p. 37, 1946. 29

[20] F. Bloch, "Nuclear induction," *Physical review*, vol. 70, no. 7-8, p. 460, 1946. 29

[21] G. S. Pell, R. S. Briellmann, A. B. Waites, D. F. Abbott, and G. D. Jackson, "Voxel-based relaxometry: a new approach for analysis of T_2 relaxometry changes in epilepsy," *Neuroimage*, vol. 21, no. 2, pp. 707–713, 2004. 37

[22] E. Alexopoulou, F. Stripeli, P. Baras, I. Seimenis, A. Kattamis, V. Ladis, E. Efstathopoulos, E. N. Brountzos, A. D. Kelekis, and N. L. Kelekis, "R2 relaxometry with MRI for the quantification of tissue iron overload in β-thalassemic patients," *Journal of Magnetic Resonance Imaging: An Official Journal of the International Society for Magnetic Resonance in Medicine*, vol. 23, no. 2, pp. 163–170, 2006. 37

[23] P. Chen, M. McCarthy, S.-M. Kim, and B. Zion, "Development of a high-speed nmr technique for sensing maturity of avocados," *Transactions of the ASAE*, vol. 39, no. 6, pp. 2205–2209, 1996. 37

[24] C. Clark, P. Hockings, D. Joyce, and R. Mazucco, "Application of magnetic resonance imaging to pre-and post-harvest studies of fruits and vegetables," *Postharvest biology and technology*, vol. 11, no. 1, pp. 1–21, 1997. 37

[25] A. Raj, S. Pandya, X. Shen, E. LoCastro, T. Nguyen, and S. Gauthier, "Multi-compartment T_2 relaxometry using a spatially constrained multi-Gaussian model," *PLoS One*, vol. 9, no. 6, p. e98391, 2014. 38, 73

[26] K. Layton, L. Johnston, P. Farrell, B. Moran, and M. Morelande, "Estimation of relaxation time distributions in magnetic resonance imaging," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 697–700, IEEE, 2012. 38, 73

[27] S. Chatterjee, O. Commowick, O. Afacan, S. K. Warfield, and C. Barillot, "Multi-compartment model of brain tissues from T_2 relaxometry MRI using gamma distribution," in *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)*, pp. 141–144, IEEE, 2018. 38

[28] A. Akhondi-Asl, O. Afacan, M. Balasubramanian, R. Mulkern, and S. Warfield, "Fast myelin water fraction estimation using 2D multislice CPMG," *Magnetic Resonance in Medicine*, vol. 76, no. 4, pp. 1301–1313, 2016. 38, 40, 73

[29] D. Kumar, T. Nguyen, S. Gauthier, and A. Raj, "Bayesian algorithm using spatial priors for multiexponential T_2 relaxometry from multiecho spin echo MRI," *Magnetic Resonance in Medicine*, vol. 68, no. 5, pp. 1536–1543, 2012. 39, 50, 71, 72, 73

[30] M. Björk, D. Zachariah, J. Kullberg, and P. Stoica, "A multicomponent T_2 relaxometry algorithm for myelin water imaging of the brain," *Magnetic Resonance in Medicine*, vol. 75, no. 1, pp. 390–402, 2016. 39

[31] T. A. Bjarnason, C. R. McCreary, J. F. Dunn, and J. R. Mitchell, "Quantitative T_2 analysis: the effects of noise, regularization, and multivoxel approaches," *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*, vol. 63, no. 1, pp. 212–217, 2010. 39

[32] Y. Yoo and R. Tam, "Non-local spatial regularization of MRI T_2 relaxation images for myelin water quantification," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 614–621, Springer, 2013. 39, 73

[33] E. Chouzenoux, S. Moussaoui, J. Idier, and F. Mariette, "Efficient maximum entropy reconstruction of nuclear magnetic resonance T_1-T_2 spectra," *IEEE Transactions on Signal Processing*, vol. 58, no. 12, pp. 6040–6051, 2010. 39

[34] L. Zhang and M. McCarthy, "Black heart characterization and detection in pomegranate using NMR relaxometry and MR imaging," *Postharvest Biology and Technology*, vol. 67, pp. 96–101, 2012. 39

[35] D. Hwang and Y. Du, "Improved myelin water quantification using spatially regularized non-negative least squares algorithm," *Journal of Magnetic Resonance Imaging*, vol. 30, no. 1, pp. 203–208, 2009. 40, 50, 71, 72, 73

[36] H. Chung, Y. Nam, D. Kim, and D. Hwang, "Three-pool model vs. nonnegative least squares algorithm for myelin water quantification," in *2011 IEEE 54th International Midwest Symposium on Circuits and Systems (MWSCAS)*, pp. 1–4, IEEE, 2011. 40, 73

[37] O. Kwon, E. Woo, Y. Du, and D. Hwang, "A tissue-relaxation-dependent neighboring method for robust mapping of the myelin water fraction," *NeuroImage*, vol. 74, pp. 12–21, 2013. 40

[38] M. Ambrosanio, F. Baselice, G. Ferraioli, F. Lenti, and V. Pascazio, "Intra voxel analysis in magnetic resonance imaging," *Magnetic Resonance Imaging*, vol. 37, pp. 70–80, 2017. 40, 73

[39] L. Soustelle, *Myelin imaging in MRI using ultra-short echo time sequences*. PhD thesis, 2018. 40

[40] Z. Ababneh, H. Beloeil, C. B. Berde, G. Gambarota, S. E. Maier, and R. V. Mulkern, "Biexponential parameterization of diffusion and T_2 relaxation decay curves in a rat muscle edema model: decay curve components and water compartments," *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*, vol. 54, no. 3, pp. 524–531, 2005. 40

[41] J. Libove and J. Singer, "Resolution and signal-to-noise relationships in nmr imaging in the human body," *Journal of physics E: Scientific instruments*, vol. 13, no. 1, p. 38, 1980. 41

[42] D. B. Rowe, "Modeling both the magnitude and phase of complex-valued fMRI data," *Neuroimage*, vol. 25, no. 4, pp. 1310–1324, 2005. 41

[43] S. Aja-Fernández and G. Vegas-Sánchez-Ferrero, "Statistical analysis of noise in MRI," *Switzerland: Springer International Publishing*, 2016. 42

[44] A. den Dekker and J. Sijbers, "Data distributions in magnetic resonance images: A review," *Physica Medica*, vol. 30, pp. 725–741, 2014. 42

[45] J. Idier and G. Collewet, "Properties of Fisher information for Rician distributions and consequences in MRI," tech. rep., LS2N/IRSTEA, 2014. 42

[46] O. T. Karlsen, R. Verhagen, and W. Bovée, "Parameter estimation from Rician-distributed data sets using a maximum likelihood estimator: application to T_1 and perfusion measurements," *Magnetic Resonance in Medicine*, vol. 41, no. 3, pp. 614–623, 1999. 43, 53, 73

[47] R. Nowak, "Wavelet-based Rician noise removal for magnetic resonance imaging," *IEEE Transactions on Image Processing*, vol. 8, no. 10, pp. 1408–1419, 1999. 45

[48] K. Whittall and A. MacKay, "Quantitative interpretation of NMR relaxation data," *Journal of Magnetic Resonance*, vol. 84, no. 1, pp. 134–152, 1989. 50

[49] J.-M. Bonny, M. Zanca, J.-Y. Boire, and A. Veyre, "T_2 maximum likelihood estimation from multiple spin-echo magnitude images," *Magnetic resonance in medicine*, vol. 36, no. 2, pp. 287–293, 1996. 50

[50] J. C. Wood, C. Enriquez, N. Ghugre, J. M. Tyzka, S. Carson, M. D. Nelson, and T. D. Coates, "MRI R2 and R2* mapping accurately estimates hepatic iron concentration in transfusion-dependent thalassemia and sickle cell disease patients," *Blood*, vol. 106, no. 4, pp. 1460–1465, 2005. 51

[51] S. S. Vasanawala, H. Yu, A. Shimakawa, M. Jeng, and J. H. Brittain, "Estimation of liver t* 2 in transfusion-related iron overload in patients with weighted least squares t* 2 ideal," *Magnetic resonance in medicine*, vol. 67, no. 1, pp. 183–190, 2012. 52

[52] G. McGibney and M. Smith, "An unbiased signal-to-noise ratio measure for magnetic resonance images," *Medical physics*, vol. 20, no. 4, pp. 1077–1078, 1993. 52

[53] A. Miller and P. Joseph, "The use of power images to perform quantitative analysis on low SNR MR images," *Magnetic Resonance Imaging*, vol. 11, no. 7, pp. 1051–1056, 1993. 52

[54] J. Raya, O. Dietrich, A. Horng, J. Weber, M. Reiser, and C. Glaser, "T_2 measurement in articular cartilage: impact of the fitting method on accuracy and precision at low SNR," *Magnetic Resonance in Medicine*, vol. 63, no. 1, pp. 181–193, 2010. 52, 73

[55] L. Kenneth, "A method for the solution of certain non-linear problems in least squares," *Quarterly of Applied Mathematics*, vol. 2, no. 2, pp. 164–168, 1944. 54

[56] D. W. Marquardt, "An algorithm for least-squares estimation of nonlinear parameters," *Journal of the society for Industrial and Applied Mathematics*, vol. 11, no. 2, pp. 431–441, 1963. 54

[57] J. Nocedal and S. J. Wright, *Numerical Optimization*. New York, NY: Springer-Verlag, 1999. 55

[58] D. Hunter and K. Lange, "A tutorial on MM algorithms," *The American Statistician*, vol. 58, no. 1, pp. 30–37, 2004. 56

[59] D. Varadarajan and J. Haldar, "A majorize-minimize framework for Rician and non-central chi MR images," *IEEE Transactions on Medical Imaging*, vol. 34, no. 10, pp. 2191–2202, 2015. 56, 57

[60] M. W. Jacobson and J. A. Fessler, "An expanded theoretical treatment of iteration-dependent majorize-minimize algorithms," *IEEE Transactions on Image Processing*, vol. 16, pp. 2411–2422, Oct. 2007. 58

[61] M. Allain, J. Idier, and Y. Goussard, "On global and local convergence of half-quadratic algorithms," *IEEE Transactions on Image Processing*, vol. 15, pp. 1130 – 1142, 06 2006. 58

[62] E. Chouzenoux, M. Legendre, S. Moussaoui, and J. Idier, "Fast constrained least squares spectral unmixing using primal-dual interior-point optimization," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 7, pp. 59–69, Jan. 2012. 70

[63] D. Kumar, S. Siemonsen, C. Heesen, J. Fiehler, and J. Sedlacik, "Noise robust spatially regularized myelin water fraction mapping with the intrinsic b1-error correction based on the linearized version of the extended phase graph model," *Journal of Magnetic Resonance Imaging*, vol. 43, no. 4, pp. 800–817, 2016. 73

[64] X. Shen, T. Nguyen, S. Gauthier, and A. Raj, "Robust myelin quantitative imaging from multi-echo T_2 MRI using edge preserving spatial priors," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 622–630, Springer, 2013. 73

[65] T. Yokoo, Q. Yuan, J. Sénégas, A. Wiethoff, and I. Pedrosa, "Quantitative R2* MRI of the liver with Rician noise models for evaluation of hepatic iron overload: Simulation, phantom, and early clinical experience," *Journal of Magnetic Resonance Imaging*, vol. 42, no. 6, pp. 1544–1559, 2015. 73

[66] H. Erdogan and J. Fessler, "Ordered subsets algorithms for transmission tomography," *Physics in Medicine & Biology*, vol. 44, no. 11, p. 2835, 1999. 74

[67] J. Idier, "Convex half-quadratic criteria and interacting auxiliary variables for image restoration," *IEEE Transactions on Image Processing*, vol. 10, pp. 1001–1009, July 2001. 74

[68] N. S. Altman, "An introduction to kernel and nearest-neighbor nonparametric regression," *The American Statistician*, vol. 46, no. 3, pp. 175–185, 1992. 89, 90

[69] J. MacQueen *et al.*, "Some methods for classification and analysis of multivariate observations," in *Proceedings of the fifth Berkeley symposium on mathematical statistics and probability*, vol. 1(14), pp. 281–297, Oakland, CA, USA, 1967. 89, 90

[70] C. E. Rasmussen, "The infinite gaussian mixture model," in *Advances in neural information processing systems*, pp. 554–560, 2000. 89, 90

[71] S. C. Johnson, "Hierarchical clustering schemes," *Psychometrika*, vol. 32, no. 3, pp. 241–254, 1967. 89, 90

[72] M. Bora, D. Jyoti, D. Gupta, and A. Kumar, "Effect of different distance measures on the performance of k-means algorithm: an experimental study in matlab," *arXiv preprint arXiv:1405.7471*, 2014. 91

[73] P. Fränti and S. Sieranoja, "How much can k-means be improved by using better initialization and repeats?," *Pattern Recognition*, vol. 93, pp. 95–112, 2019. 91

[74] T. M. Kodinariya and P. R. Makwana, "Review on determining number of cluster in k-means clustering," *International Journal*, vol. 1, no. 6, pp. 90–95, 2013. 95

[75] D. R. Jones, M. Schonlau, and W. J. Welch, "Efficient global optimization of expensive black-box functions," *Journal of Global optimization*, vol. 13, no. 4, pp. 455–492, 1998. 108

UNIVERSITE BRETAGNE LOIRE / MATHSTIC

CENTRALE NANTES

Titre : Estimation et classification des temps de relaxation multi-exponentiels en IRM. Application aux tissus végétaux.

Mots clés : Relaxométrie IRM, Bruit de Rice, Modèle Multi-exponentiel, Classification, Majoration-Minimisation, Estimation conjointe.

Résumé : L'acquisition de données de relaxation en imagerie de résonance magnétique permet une analyse très fine de la composition des tissus. L'analyse est classiquement réalisée à l'aide de modèles mono-exponentiels au sein de chaque voxel de l'image, mais des informations plus riches peuvent être obtenues à l'aide d'un modèle de décroissance multi-exponentielle. Cependant, l'obtention puis l'exploitation des cartographies des temps de relaxation multi-exponentielles à l'échelle d'une image entière, à partir des données IRM de module, nécessitent la résolution d'un problème inverse de grande taille. Ce travail de thèse propose des algorithmes de reconstruction des cartographies des temps de relaxation multi-exponentielles et des intensités relatives associées à l'échelle du voxel. Ces algorithmes de reconstruction sont fondés sur l'estimateur du maximum de vraisemblance exploitant l'hypothèse d'un bruit de Rice, inhérent aux images de module, et une régularisation spatiale favorisant la régularité des cartographies. Le problème d'optimisation en grande dimension qui en résulte est résolu en utilisant une approche de descente itérative par majoration-minimisation couplée à un algorithme de Levenberg-Marquardt avec recherche de pas. Enfin, nous proposons une méthode de caractérisation de la composition des images à partir des paramètres estimés en utilisant des algorithmes de classification. Les développements de la thèse font l'objet d'application à l'analyse de tissus végétaux.

Title : Multi-exponential relaxation times estimation and classification in MRI. Application to vegetal tissus.

Keywords : MRI relaxometry, Rician noise, Multi-exponential model, Classification, Majorization-Minimization, Joint estimaion.

Abstract : Acquired relaxation data in magnetic resonance imaging makes it possible to conduct fine analysis of tissues composition. Conventionally, the analysis is realized by adopting a mono-exponential model at each voxel of the image, yet, a multi-exponential decay model may provide richer information. However, obtaining and interpreting multi-exponential relaxation time maps at the whole image level, from magnitude MRI images, requires solving a large scale inverse problem. This thesis work proposes algorithms for multi-exponential relaxation times and their associated intensities maps reconstruction. These algorithms are based on the maximum-likelihood estimator under the hypothesis of a Rician noise distribution, case of magnitude images, and a spatial regularization favoring the regularity of the maps. The resulting large-scale optimization problem is solved using an iterative descent approach by majorization-minimization coupled with a Levenberg-Marquardt algorithm with step search. Finally, we propose a method for image composition characterization from the estimated parameters using classification algorithms. The developed algorithms in this thesis are applied to vegetal tissue analysis.