

UNIVERSITÉ DE SHERBROOKE

Faculté de génie

Département de génie chimique et de génie biotechnologique

DÉVELOPPEMENT DE NOUVEAUX OUTILS
STATISTIQUES POUR L'ANALYSE D'IMAGES
SPECTRALES À FAIBLES COMPTES

Mémoire de maîtrise

Spécialité : génie chimique

Francis B. LAVOIE

Jury : Ryan GOSSELIN (directeur)

Nadi BRAIDY (co-directeur)

Bernard MARCOS

Martin COUILLARD

RÉSUMÉ

Le projet de maîtrise était axé sur l'utilisation de traitements multivariés afin d'analyser adéquatement des ensembles de données provenant d'imagerie spectrale. Ce mémoire met premièrement en relief la particularité des détecteurs utilisés dans les microscopes, ainsi que les caractéristiques du bruit dans les données acquises.

L'analyse en composantes principales est une méthode qui se trouve dans plusieurs méthodologies de traitements de données d'imagerie spectrale. Celle-ci comprend certains désavantages, dont l'impossibilité d'extraire de réels spectres et des distributions de concentrations des composés en présence dans l'échantillon.

La résolution multivariée de courbes – moindres carrés alternatifs est une méthodologie qui a été popularisée dans les années 2000. Celle-ci permet de contrer les désavantages de l'analyse en composantes principales en extrayant des spectres physiquement cohérents et en créant des cartes de concentrations associées à ces spectres.

Cependant, ce mémoire démontre que cette méthode est inefficace lorsque les données ont un très faible ratio signal sur bruit et que plusieurs composants sont à extraire. Des améliorations à la résolution multivariée de courbes – moindres carrés alternatifs sont donc apportées. Notamment, la caractéristique du bruit des données, connue et documentée à la suite de nombreuses études, est utilisée afin d'améliorer la convergence de l'algorithme vers la bonne solution.

Ce mémoire démontre que ces améliorations sont appliquées avec succès sur des ensembles de données d'imagerie spectrale provenant de spectrométrie photoélectronique X, d'analyse dispersive en énergie et de spectroscopie des pertes d'énergie.

Mots clefs : Microscopie, Spectroscopie, Multivarié, Logvraisemblance, Poisson, XPS, EDX, EELS

TABLE DES MATIÈRES

| | |
|--|----|
| 1. INTRODUCTION | 1 |
| 1.1 Mise en contexte, définitions et problématique | 1 |
| 1.1.1 Type d'ensembles de données étudiés | 2 |
| 1.1.2 Spectroscopie et microscopie..... | 3 |
| 1.1.3 Problématique | 4 |
| 1.2 Définition du projet de recherche..... | 5 |
| 1.2.1 But d'une analyse par imagerie spectrale en microscopie | 5 |
| 1.2.2 Méthodes de traitements des données couramment utilisées..... | 7 |
| 1.2.3 Désavantages des méthodes couramment utilisées..... | 8 |
| 1.3 Objectifs du projet de recherche..... | 9 |
| 1.4 Structure du mémoire | 10 |
| 2. CARACTÉRISTIQUES DE LA SPECTROMÉTRIE PHOTOÉLECTRONIQUE X (XPS) ET CONCEPTS PHYSIQUES ASSOCIÉS | 11 |
| 2.1 Caractéristiques et fonctionnement du détecteur | 11 |
| 2.1.1 Plaque de microcanaux | 11 |
| 2.1.2 Détecteur à délai de ligne..... | 12 |
| 2.1.3 Particularités du détecteur..... | 13 |
| 2.2 Comportement des photoélectrons..... | 14 |
| 2.2.1 De photon à photoélectron | 14 |
| 2.2.2 Énergie analysée | 14 |
| 2.2.3 Efficacité du détecteur | 15 |
| 2.3 Bruit de fond | 18 |

| | | |
|-------|--|----|
| 2.3.1 | Signification physique du bruit de fond..... | 18 |
| 4.3.2 | Bruit de fond de Shirley..... | 19 |
| 2.3.3 | Bruit de fond de Tougaard..... | 21 |
| 2.3.4 | Comparaison entre les deux principales méthodes du calcul du bruit de fond | 23 |
| 2.3.5 | Approximation de bruit de fond d'un mélange de différents composants..... | 23 |
| 2.3 | Résumé du Chapitre 2..... | 25 |
| 3. | ALGORITHMES ET MÉTHODES..... | 27 |
| 3.1 | Analyse multivariable..... | 27 |
| 3.1.1 | Prétraitement des données..... | 29 |
| 3.1.2 | Analyse en composantes principales..... | 33 |
| 3.1.3 | Résolution multivariée de courbes – Moindres carrés alternatifs (<i>MCR-ALS</i>) | 35 |
| 3.1.4 | <i>SIMPLISMA</i> | 38 |
| 3.1.5 | Algorithme des k-moyennes..... | 41 |
| 3.1.6 | Comparaison et sélection des méthodes d'initialisation..... | 43 |
| 3.2 | Optimisation..... | 47 |
| 3.2.1 | Méthode de Newton..... | 47 |
| 3.2.2 | Optimisation sous contrainte..... | 49 |
| 3.2.3 | Algorithme <i>Active Set</i> | 50 |
| 3.3 | Traitement de signal..... | 52 |
| 3.3.1 | Fonction de vraisemblance..... | 52 |
| 3.3.2 | Algorithme espérance-maximisation..... | 53 |
| 3.4 | Résumé du Chapitre 3..... | 57 |
| 4. | NOUVELLE MÉTHODE DE TRAITEMENTS DE DONNÉES PROPOSÉE..... | 60 |
| 4.1 | Données..... | 60 |
| 4.1.1 | Caractéristiques des données..... | 60 |

| | | |
|-------|--|----|
| 4.1.2 | Prétraitement des données sélectionnée..... | 61 |
| 4.2 | Méthode développée (<i>MCR-LLM</i>)..... | 62 |
| 4.2.1 | Initialisation..... | 62 |
| 4.2.2 | Cœur de la méthode..... | 63 |
| 4.3 | Discussions..... | 67 |
| 4.3.1 | Maximisation de la fonction de vraisemblance versus minimisation de l'erreur sous contraintes..... | 67 |
| 4.3.2 | Les mathématiques sous φ | 68 |
| 4.3.3 | Valeur de l'exposant φ optimale..... | 72 |
| 4.3.4 | <i>MCR-LLM</i> par rapport à d'autres algorithmes..... | 74 |
| 4.4 | Résumé du Chapitre 4..... | 76 |
| 5. | APPLICATION DE LA MÉTHODE DÉVELOPPÉE SUR DES DONNÉES XPS | 79 |
| 5.1 | Liens entre la méthode développée et les données XPS..... | 79 |
| 5.1.1 | Comportement du signal..... | 79 |
| 5.1.2 | Spectres..... | 79 |
| 5.2 | Ensemble de données des demi-cercles..... | 80 |
| 5.2.1 | Description de l'échantillon analysé..... | 80 |
| 5.2.2 | Paramètres d'acquisition..... | 82 |
| 5.2.3 | Analyses <i>PCA</i> | 82 |
| 5.2.4 | Analyses par <i>MCR-ALS</i> et <i>MCR-LLM</i> | 84 |
| 5.3 | Ensemble de données de trois métaux..... | 87 |
| 5.3.1 | Description de l'échantillon..... | 87 |
| 5.3.2 | Paramètres d'acquisition..... | 88 |
| 5.3.3 | Analyses <i>PCA</i> | 89 |
| 5.3.4 | Aperçus des résultats à travers les itérations de <i>MCR-ALS</i> et <i>MCR-LLM</i> | 94 |

| | | |
|----------|--|-----|
| 5.3.5 | Indices de performance | 102 |
| 5.3.6 | Performance entre les différents algorithmes | 104 |
| 5.4 | Retour sur le Chapitre 5 | 106 |
| 6. | APPLICATION DE LA MÉTHODE SUR D'AUTRES TYPES DE DONNÉES | 109 |
| 6.1 | Analyse dispersive à rayons X | 109 |
| 6.1.1 | Explications de l'analyse | 109 |
| 6.1.2 | Comportement de <i>MCR-ALS</i> versus <i>MCR-LLM</i> | 112 |
| 6.1.3 | Résultats avec $\varphi > 1$ | 117 |
| 6.1.4 | Résultats avec φ variable automatiquement | 120 |
| 6.2 | Spectroscopie des pertes d'énergie | 121 |
| 6.2.1 | Explications de l'analyse | 121 |
| 6.2.2 | Résultats | 123 |
| 6.3 | Retour sur le Chapitre 6 | 128 |
| 7. | CONCLUSION..... | 130 |
| 7.1 | Retour sur les objectifs..... | 130 |
| 7.2 | Contributions originales..... | 132 |
| 7.3 | Travaux futurs..... | 133 |
| 7.3.1 | Linéarité des spectres..... | 133 |
| 7.3.2 | Rapidité de l'algorithme | 134 |
| ANNEXE A | PRÉSENTATION DU LOGICIEL DÉVELOPPÉ | 135 |
| AA.1 | Importation de données | 135 |
| AA.2 | Analyse..... | 136 |
| AA.3 | Visualisation des résultats | 138 |
| ANNEXE B | PSEUDOCODES | 140 |
| AB.1 | Analyse en composantes principales (<i>PCA</i>)..... | 140 |

| | | |
|----------|--|-----|
| AB.2 | Résolution multivariée de courbes – Moindres carrés alternatifs (<i>MCR-ALS</i>) | 141 |
| ANNEXE C | DÉMONSTRATION DE LA LINÉARITÉ DU BRUIT DE FOND AU XPS | 143 |
| AC.1 | Développement et démonstration | 143 |
| AC.2 | Exemple numérique | 149 |
| ANNEXE D | FABRICATION DES ÉCHANTILLONS | 152 |
| AD.1 | Conception des masques | 152 |
| AD.2 | Procédé de fabrication | 153 |
| AD.2.1 | Découpage | 153 |
| AD.2.2 | Nettoyage du substrat | 154 |
| AD.2.3 | Déposition du vanadium | 154 |
| AD.2.4 | Déposition de résine | 154 |
| AD.2.5 | Alignement et exposition | 155 |
| AD.2.6 | Développement | 155 |
| AD.2.7 | Déposition du titane et du chrome | 155 |
| AD.2.8 | Soulèvement | 155 |
| ANNEXE E | CARACTÉRISTIQUES SUPPLÉMENTAIRES DU XPS UTILISÉ | 156 |
| AE.1 | Processus de création des photoélectrons | 156 |
| AE.1.1 | Sections de l'instrument | 156 |
| AE.1.2 | Génération des rayons X | 157 |
| AE.1.3 | Système de lentilles | 158 |
| AE.2 | Mode spectral versus mode imagerie | 159 |
| AE.2.1 | Analyseur hémisphérique concentrique | 159 |
| AE.2.2 | Analyseur en miroir sphérique | 162 |
| ANNEXE F | | 165 |

LISTE DES RÉFÉRENCES 168

LISTE DES FIGURES

| | |
|---|----|
| Figure 1.1 Représentation d'un ensemble de données généré par caméra photographique | 2 |
| Figure 1.2 Représentation d'un ensemble de données multispectrales | 3 |
| Figure 1.3 Résolutions spatiales et spectrales des microscopes [4] | 6 |
| Figure 2.1 Plaque de microcanaux [11] | 12 |
| Figure 2.2 Détecteur à délai de ligne [10] | 13 |
| Figure 2.3 Schématisation de l'énergie du photoélectron du matériau à l'analyseur [17] | 15 |
| Figure 2.4 Facteurs géométriques expliquant [18] | 16 |
| Figure 2.5 Sensibilité par rapport à la moyenne de sensibilité entre les pixels pour le <i>Kratos XPS Ultra DLD</i> de l'Université de Sherbrooke | 17 |
| Figure 2.6 Exemple de spectre XPS [19] | 18 |
| Figure 2.7 Exemple de courbe de probabilité de perte d'énergie due à la collision inélastique | 19 |
| Figure 2.8 Définition de variables dans l'équation du fond de Shirley | 20 |
| Figure 2.9 Définition des variables dans les équations pour calculer le fond de Tougaard | 21 |
| Figure 2.10 Comparaison entre des fonds de Shirley et de Tougaard [28] | 23 |
| Figure 3.1 Représentation générale d'un ensemble de données | 27 |
| Figure 3.2 Représentation d'un ensemble de données d'imagerie multispectrale | 28 |
| Figure 3.3 Représentation de la transformation de données d'imagerie spectrale en matrice à deux dimensions | 28 |
| Figure 3.4 Exemple de spectre sans prétraitement d'une analyse par spectrométrie photoélectronique X | 32 |
| Figure 3.5 Exemples de prétraitement d'une analyse par spectrométrie photoélectronique X par (à gauche) donnée centrées et réduites, <i>Optimal Scaling</i> et (à droite) par normalisation spectrale | 32 |
| Figure 3.6 Décomposition d'une matrice de données par PCA [32] | 35 |
| Figure 3.7 Schématisation de l'algorithme de MCR-ALS [39] | 37 |
| Figure 3.8 Création de la matrice $M_{test, n}$ | 39 |

| | |
|---|----|
| Figure 3.9 Algorithme <i>SIMPLISMA</i> | 40 |
| Figure 3.10 Exemple d'extraction de trois spectres d'un ensemble de spectres par <i>SIMPLISMA</i> [35] | 41 |
| Figure 3.11 Algorithme des k-moyennes afin de segmenter les données \mathbf{X} dans K groupes régis par les centroïdes \mathbf{O} | 43 |
| Figure 3.12 Exemple d'initialisation problématique avec <i>SIMPLISMA</i> | 44 |
| Figure 3.13 Exemple d'initialisation par l'algorithme des k-moyennes | 45 |
| Figure 3.14 Comparaison des spectres obtenus entre deux méthodes d'initialisation | 46 |
| Figure 3.15 Comparaison des spectres obtenus entre deux méthodes d'initialisation | 46 |
| Figure 3.16 Cas typiques dans l'algorithme <i>Active Set</i> | 51 |
| Figure 3.17 Algorithme Espérance-Maximisation | 55 |
| Figure 4.1 Cœur de la méthode proposée | 63 |
| Figure 4.2 Exemple d'optimisation de la fonction de vraisemblance | 65 |
| Figure 4.3 Schématisation de la matrice calculée dans la première partie de la régression multilinéaire | 69 |
| Figure 4.4 Matrice \mathbf{B} simulée | 71 |
| Figure 4.5 Spectres résultant après normalisation avec (1) une matrice \mathbf{A} dont les nombres hors diagonale sont à 0 et avec (2) une matrice \mathbf{A} dont les nombres hors diagonale ont une intensité notable | 71 |
| Figure 4.6 Exemple de la quantification de la modification des spectres selon différentes valeurs de φ | 73 |
| Figure 4.7 Comparaison entre une itération de MCR-ALS standard et une itération de <i>MCR-LLM</i> | 77 |
| Figure 5.1 Représentation de l'échantillon des demi-cercles | 81 |
| Figure 5.2 Spectres de référence [49] | 82 |
| Figure 5.3 Matrice des <i>scores</i> de l'ensemble des données des demi-cercles, avec prétraitement <i>Optimal Scaling</i> | 83 |
| Figure 5.4 <i>Loadings</i> de l'ensemble de données des demi-cercles, avec prétraitement <i>Optimal Scaling</i> | 83 |
| Figure 5.5 Spectres calculés par <i>MCR-ALS</i> avec l'ensemble de données des demi-cercles (Ti/TiO ₂), acquisition de 300 secondes par niveau d'énergie | 84 |

| | |
|---|----|
| Figure 5.6 Contributions calculées par <i>MCR-ALS</i> avec l'ensemble de données des demi-cercles (Ti/TiO ₂), acquisition de 300 secondes par niveau d'énergie..... | 85 |
| Figure 5.7 Spectres calculés par <i>MCR-LLM</i> , $\varphi = 1$, avec l'ensemble de données des demi-cercles (Ti/TiO ₂), acquisition de 300 secondes par niveau d'énergie..... | 86 |
| Figure 5.8 Contributions calculées par <i>MCR-LLM</i> , $\varphi = 1$, avec l'ensemble de données des demi-cercles (Ti/TiO ₂), acquisition de 300 secondes par niveau d'énergie | 86 |
| Figure 5.9 Représentation en trois dimensions de l'échantillon V/Cr/Ti | 87 |
| Figure 5.10 Nombre de photoélectrons comptés moyen pour un pixel pour un niveau d'énergie | 89 |
| Figure 5.11 Spectres de référence pour (a) le chrome et (b) le vanadium [49], [50]..... | 90 |
| Figure 5.12 Matrice des <i>scores</i> obtenues par <i>PCA</i> à partir de l'ensemble des données avec trois métaux (Ti/Cr/V), prêtatement <i>Optimal Scaling</i> , acquisition de 1 seconde par niveau d'énergie | 90 |
| Figure 5.13 <i>Loadings</i> obtenus par <i>PCA</i> à partir de l'ensemble des données avec trois métaux (Ti/Cr/V), prêtatement <i>Optimal Scaling</i> , acquisition de 1 seconde par niveau d'énergie | 91 |
| Figure 5.14 Matrice des <i>scores</i> obtenues par <i>PCA</i> de l'ensemble des données avec trois métaux (Ti/Cr/V), données centrées-réduites, acquisition de 1 seconde par niveau d'énergie | 92 |
| Figure 5.15 <i>Loadings</i> obtenus par <i>PCA</i> de l'ensemble des données avec trois métaux (Ti/Cr/V), données centrées-réduites, acquisition de 1 seconde par niveau d'énergie..... | 92 |
| Figure 5.16 Matrice des <i>scores</i> obtenues par <i>PCA</i> de l'ensemble des données avec trois métaux (Ti/Cr/V), données normalisées au niveau spectral et ensuite centrées-réduites, acquisition de 1 seconde par niveau d'énergie | 93 |
| Figure 5.17 <i>Loadings</i> obtenus par <i>PCA</i> de l'ensemble des données avec trois métaux (Ti/Cr/V), données normalisées au niveau spectral et ensuite centrées-réduites, acquisition de 1 seconde par niveau d'énergie | 93 |
| Figure 5.18 Résultats obtenus par <i>MCR-LLM</i> à partir de l'ensemble des données avec trois métaux (Ti/Cr/V) au fil des itérations, $\varphi = 2$, acquisition de 5 secondes par niveau d'énergie | 95 |

| | |
|--|-----|
| Figure 5.19 Résultats obtenus par <i>MCR-ALS</i> standard à partir de l'ensemble des données avec trois métaux (Ti/Cr/V) au fil des itérations, acquisition de 5 secondes par niveau d'énergie | 96 |
| Figure 5.20 Résultats à convergence obtenus par <i>MCR-ALS+φ</i> à partir de l'ensemble de données avec trois métaux (Ti/Cr/V), acquisition de 5 secondes par niveau d'énergie ... | 98 |
| Figure 5.21 Problématiques au niveau des spectres obtenus par <i>MCR-ALS+φ</i> à partir de l'ensemble de données avec trois métaux (Ti/Cr/V), $\varphi = 1,1$, acquisition de 5 secondes par niveau d'énergie | 99 |
| Figure 5.22 Résultats à convergence de différents algorithmes avec une acquisition d'une seconde par énergie de liaison | 100 |
| Figure 5.23 Résultats à convergence de <i>MCR-LLM</i> avec φ optimisé à chaque itération avec une acquisition d'une seconde par énergie de liaison..... | 101 |
| Figure 5.24 Évolution du φ dans l'algorithme de <i>MCR-LLM</i> avec φ optimisé à chaque itération avec une acquisition d'une seconde par énergie de liaison | 102 |
| Figure 5.25 Spectres de référence obtenus avec un temps d'éclairement (<i>dwell time</i>) de 20 s par niveau d'énergie | 103 |
| Figure 5.26 Modèle de contributions | 103 |
| Figure 5.27 Indices de performance calculés à convergence de <i>MCR-ALS</i> et <i>MCR-LLM</i> , avec $\varphi = 1$ et φ optimisé à chaque itération, avec l'ensemble de données des trois métaux obtenus à différents temps d'acquisition..... | 105 |
| Figure 6.1 Principes physiques de l' <i>EDX</i> | 109 |
| Figure 6.2 Représentation spatiale des sections de l'analyse <i>EDX</i> | 110 |
| Figure 6.3 Contributions calculées de InAs et de GaAs par intégration par plage d'énergie et par optimisation..... | 111 |
| Figure 6.4 Représentation de l'ensemble de données <i>EDX</i> et exemple de spectre..... | 111 |
| Figure 6.5 Évolution des spectres et des contributions au fil des itérations, avec <i>MCR-ALS</i> | 113 |
| Figure 6.6 Évolution des spectres et des contributions au fil des itérations, avec <i>MCR-LLM</i> , $\varphi = 1$ | 114 |
| Figure 6.7 Contributions à convergence, avec <i>MCR-LLM</i> , $\varphi = 1$, en retirant le composant associé au vide | 115 |

| | |
|--|-----|
| Figure 6.8 Comparaison de sections de spectres entre l'initialisation et la 50 ^e itération de <i>MCR-LLM</i> dans (a) les environs de 1400 eV et dans (b) les environs de 3700 eV | 115 |
| Figure 6.9 Comparaison entre les valeurs de contributions calculées entre les deux méthodes pour un même pixel | 116 |
| Figure 6.10 Comparaison entre les contributions calculées <i>MCR-LLM</i> $\varphi = 1$, et <i>MCR-LLM</i> , $\varphi = 1,5$ | 118 |
| Figure 6.11 Comparaison entre les spectres calculés avec <i>MCR-LLM</i> , $\varphi = 1$, et <i>MCR-LLM</i> , $\varphi = 1,5$ | 118 |
| Figure 6.12 Méthodologie de l'utilisation de <i>MCR-LLM</i> suivi de <i>MCR-ALS</i> standard.. | 119 |
| Figure 6.13 Différences ciblées entre les spectres calculés pour le 1 ^{er} composant | 119 |
| Figure 6.14 φ calculé à travers les itérations | 120 |
| Figure 6.15 Schématisation..... | 122 |
| Figure 6.16 Intégration du signal sur des plages d'énergie caractéristiques..... | 123 |
| Figure 6.17 Contributions calculée avec les données <i>EELS</i> , avec utilisation de toutes les modifications apportées par <i>MCR-LLM</i> , avec trois composants et $\varphi = 1,5$ | 124 |
| Figure 6.18 Contributions calculée avec les données <i>EELS</i> , sans utilisation de la maximisation du logvraisemblance, avec trois composants et $\varphi = 1,5$ | 124 |
| Figure 6.19 Spectres <i>EELS</i> calculés pour <i>MCR-LLM</i> , mais sans maximisation du logvraisemblance, avec trois composants et $\varphi = 1,5$ | 125 |
| Figure 6.20 Contributions calculée avec les données <i>EELS</i> , avec utilisation de toutes les modifications apportées par <i>MCR-LLM</i> , avec quatre composants et $\varphi = 1,5$ | 126 |
| Figure 6.21 Contributions calculée avec les données <i>EELS</i> , sans utilisation de la maximisation du logvraisemblance, avec quatre composants et $\varphi = 1,5$ | 126 |
| Figure 6.22 Spectres calculés pour <i>MCR-LLM</i> et sans maximisation du logvraisemblance, avec trois composants et $\varphi = 1,5$ | 127 |
| Figure AA.1 Fenêtre d'importation de données dans le logiciel | 135 |
| Figure AA.2 Fenêtre d'analyse | 137 |
| Figure AA.3 Information apparaissant pendant les calculs | 138 |
| Figure AA.4 Fenêtre de visualisation des résultats..... | 138 |
| Figure AC.1 Exemples de spectres et bruit de fond calculé selon ceux-ci..... | 150 |

| | |
|---|-----|
| Figure AC.2 Spectres calculés de façon numérique (méthode 1) et par l'équation élaborée (méthode 2) | 151 |
| Figure AD.1 Motifs des masques..... | 152 |
| Figure AD.2 Procédé de fabrication des échantillons..... | 153 |
| Figure AE.1 Principe global du XPS [55] | 156 |
| Figure AE.2 Production de rayons X [56] | 157 |
| Figure AE.3 Représentation d'un cristal monochromateur [57]..... | 157 |
| Figure AE.4 Schématisation de la lentille d'immersion magnétique [58]..... | 158 |
| Figure AE.5 Analyseur hémisphérique concentrique [17] | 159 |
| Figure AE.6 Schématisation du FWHM et du ΔE dans les spectres | 161 |
| Figure AE.7 Acquisition en mode imagerie de <i>Kratos Analytical</i> [60]..... | 163 |
| Figure AE.8 Déflecteur dans l'analyseur en miroir sphérique [60]..... | 164 |

LISTE DES TABLEAUX

| | |
|---|-----|
| Tableau 2.1 Acronyme utilisé dans le Chapitre 2 | 26 |
| Tableau 2.2 Variables utilisées dans le Chapitre 2 | 26 |
| Tableau 3.1 Comparaison entre les prétraitements étudiés | 33 |
| Tableau 3.2 Avantages et désavantages des différentes méthodes d'initialisation des spectres..... | 47 |
| Tableau 3.3 Acronymes utilisés dans le Chapitre 3 | 58 |
| Tableau 3.4 Variables utilisées dans le Chapitre 3 | 58 |
| Tableau 4.1 Exemples d'erreurs absolues calculées..... | 67 |
| Tableau 4.2 Acronyme utilisés dans le Chapitre 4 | 77 |
| Tableau 4.3 Variables utilisées dans le Chapitre 4 | 78 |
| Tableau 5.1 Acronyme utilisés dans le Chapitre 5 | 107 |
| Tableau 5.2 Variables utilisées dans le Chapitre 5 | 108 |
| Tableau 6.1 Acronyme utilisés dans le Chapitre 6 | 129 |
| Tableau 6.2 Variables utilisées dans le Chapitre 6 | 129 |
| Tableau 7.1 Caractéristiques des donnés | 132 |

1. INTRODUCTION

1.1 Mise en contexte, définitions et problématique

Que ce soit en recherche expérimentale, en développement ou en procédé, recueillir des données à l'aide d'un quelconque instrument est devenu pratique courante, grâce à l'avènement des détecteurs ultras rapides et de plus en plus performants. Au cours des dernières décennies, plusieurs microscopes, tels que les microscopes électroniques à balayage se sont démarqués par rapport à leur capacité et surtout par rapport à la quantité de données qu'ils génèrent.

Par exemple, une analyse dispersive en énergie génère un ensemble de données avec 2048 canaux d'énergie (*sous-section 6.1*), et ce pour chaque pixel. Une expression courante québécoise dit : « Trop, c'est comme pas assez ». Extirper de l'information pertinente dans une très grande quantité de données s'avère souvent une opération fastidieuse.

Plusieurs domaines utilisent les données spectrales et ont développé des techniques qui leur sont propres. Entre autres, dans l'industrie de l'imagerie aérienne, alors qu'on acquiert principalement des images dans les couleurs visibles et l'infrarouge, les techniques développées sont axées sur la détection d'objets et la classification des pixels.

En microscopie, les objectifs à atteindre sont presque les mêmes qu'en imagerie aérienne. On cherche à extraire des spectres caractéristiques du matériau analysé, ainsi que les pixels liés à chacun de ces spectres. En imagerie aérienne, aucun phénomène n'est provoqué en tant que tel. On capture passivement le rayonnement émis et réfléchi par les objets analysés. Au contraire, en microscopie, un faisceau est envoyé sur l'échantillon analysé afin de créer une émission électromagnétique caractéristique de cet échantillon.

La sensibilité des échantillons aux faisceaux vient compliquer l'imagerie spectrale en microscopie. Alors que les détecteurs se sont amplement améliorés au cours des années,

on cherche à analyser et caractériser des matériaux de plus en plus petits, ce qui cause un nouveau problème. La qualité de l'analyse est désormais limitée par la sensibilité de l'échantillon au faisceau. Ainsi, la quantité d'information extirpée par pixel est très limitée et, par conséquent, les données sont grandement bruitées.

De plus, au niveau spectral, les spectres générés par les différents composés en présence sur l'échantillon peuvent avoir des différences subtiles, mais à la fois très caractéristiques. Ainsi, en plus d'être inondé par une quantité astronomique de données, ces données sont extrêmement bruitées et les différents composés sont difficiles à distinguer au niveau spectral. Bien qu'un ensemble de données demeure un défi à analyser, peu importe le domaine, les données provenant d'un microscope s'avèrent un défi encore plus relevé.

1.1.1 Type d'ensembles de données étudiés

On parle d'ensemble de données multispectrales tout ensemble de données qui contient des images provenant de différents niveaux d'énergie, que ce soit par l'analyse de différentes énergies de photons ou différentes énergies d'électrons. Il y a plusieurs exemples d'instruments générant un ensemble de données multispectrales. Le plus connu de tous est la simple caméra couleur photographique.

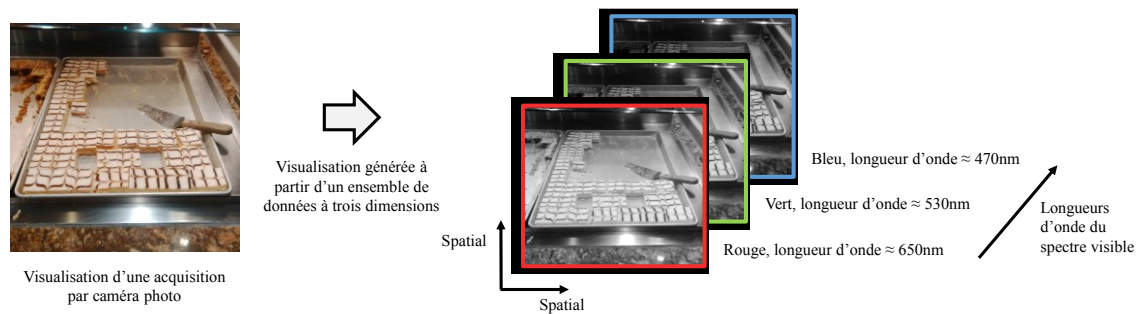


Figure 1.1 Représentation d'un ensemble de données généré par caméra photographique

En imagerie photo, on mesure l'intensité de la lumière perçue (photons) en appliquant trois filtres laissant passer différentes couleurs (rouge, vert et bleu), ce qui se traduit à laisser passer trois différentes plages de longueurs d'onde. Pour chacune des trois différentes plages de longueurs d'onde, on acquiert une matrice en deux dimensions présentant

l'intensité perçue pour la couleur correspondante. On regroupe ces trois matrices pour former une seule matrice, mais en trois dimensions, dans laquelle les deux premières dimensions sont les dimensions spatiales et la troisième est reliée aux plages de longueurs d'onde. On a donc une acquisition de données multispectrales.

Ainsi, tout ensemble de données d'imagerie multispectrale possède trois dimensions. Les deux premières dimensions sont spatiales, alors que la troisième dimension représente un spectre (e.g. longueurs d'onde ou niveaux d'énergies analysées). Ainsi, dans ce type d'ensemble de données, chaque pixel de l'image contient un spectre. (Figure 1.2)

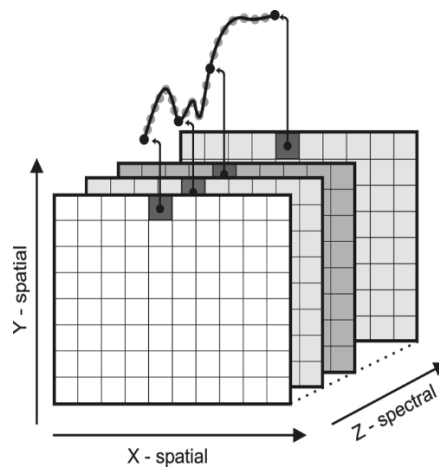


Figure 1.2 Représentation d'un ensemble de données multispectrales

1.1.2 Spectroscopie et microscopie

La spectroscopie est une technique d'analyse par laquelle on acquiert et on étudie un spectre afin de caractériser un matériau ou afin d'observer un phénomène. Un spectre est, quant à lui, une dispersion de photons selon leur énergie. Par exemple, dans le cas d'une caméra photo, des photons sont réfléchis par un objet. Ces photons ont différentes énergies (ou longueurs d'onde). En appliquant un filtre, on ne sélectionne que les photons dans un certain intervalle d'énergie et on compte le nombre de photons traversant ce filtre. Le nombre de photons comptés pour chaque pixel résultera donc en une image.

Dans le cas des spectromètres dont il sera mention dans ce mémoire, le but est d'acquérir un spectre afin de caractériser des matériaux. Pour être en mesure de générer un spectre,

un signal est créé par une sonde dirigée sur le matériau. Des signaux caractéristiques provenant du matériau sont alors détectés et quantifiés, selon leur énergie, par l'instrument.

À la base, un spectromètre fait l'acquisition d'un seul spectre. Il y a donc génération de données qui peuvent être représentées sous forme de vecteur ou d'un spectre pour une position (x,y) donnée. Lorsqu'on parle d'imagerie spectrale, le spectromètre fait l'acquisition de plusieurs spectres (en parallèle ou en série) et génère un ensemble de données d'imagerie multispectrale.

La microscopie se définit comme un ensemble de techniques qui permet d'analyser des objets de très petites dimensions. En associant ainsi la microscopie à la spectroscopie, il sera donc possible de produire une carte chimique du matériau à l'échelle du micron et même du nanomètre.

1.1.3 Problématique

Le principal problème à associer microscopie et imagerie spectrale est le fait que les signaux émis par le matériau sont généralement de très faible intensité. Ce phénomène est dû au fait que le volume sondé est généralement très petit. Afin d'obtenir des spectres pour chaque pixel (imagerie multispectrale), avec un ratio signal sur bruit suffisant, un très long temps d'acquisition, soit au-delà d'une dizaine d'heures pour certaines techniques, peut être requis, et ce, pour une seule analyse.

De plus, on peut être totalement noyé par la quantité de données. Un ensemble de données peut avoir des dimensions spatiales dépassant les 1000 pixels et des dimensions spectrales dépassant les 2000 canaux. Aussi, les signatures spectrales peuvent être très similaires entre les différents composés. Par exemple, dans les spectres résultant d'une spectroscopie des pertes d'énergie, la majorité de l'intensité perçue fait partie du bruit de fond. Les pics caractéristiques du matériau sont de très faible intensité comparativement à l'omniprésence de bruit de fond. Analyser et extraire de l'information pertinente de ces données s'avère souvent une tâche très coriace.

1.2 Définition du projet de recherche

1.2.1 But d'une analyse par imagerie spectrale en microscopie

Plusieurs raisons poussent les scientifiques à effectuer une analyse avec instrument permettant l'imagerie spectrale. Par exemple, dans l'industrie alimentaire, on recense beaucoup d'utilisations de l'imagerie infrarouge jumelée à des méthodes de chimométrie. Entre autres, on utilise ce jumelage de techniques pour la détection spatiale de contaminants ou de zones abîmées sur les pommes [1]. Les zones abîmées qui étaient auparavant très difficiles à observer sont maintenant très bien décelées grâce au développement et à l'amélioration de ces techniques. On utilise aussi ces mêmes techniques afin de détecter la présence de tumeurs sur les carcasses de poulet [2].

En microélectronique, on utilise la spectrométrie photoélectronique X, ainsi que la spectroscopie par électrons Auger, le tout jumelé à des techniques de chimométrie, afin de comprendre la délamination interfaciale dans la production de composantes électriques. [3]

Il existe des centaines d'instruments, que ce soit le microscope électronique en transmission ou le spectroscope de masse d'ions secondaires ou le spectroscope de fluorescence X, permettant l'imagerie multispectrale. Chacun des instruments a une résolution spatiale qui lui est propre. De plus, alors que certains instruments n'analysent que la surface de l'échantillon au niveau des molécules chimiques en présence, tel le spectroscope photoélectronique X, d'autres analysent la structure de l'échantillon, tel que la diffraction par rayons X. (*Figure 1.3*)

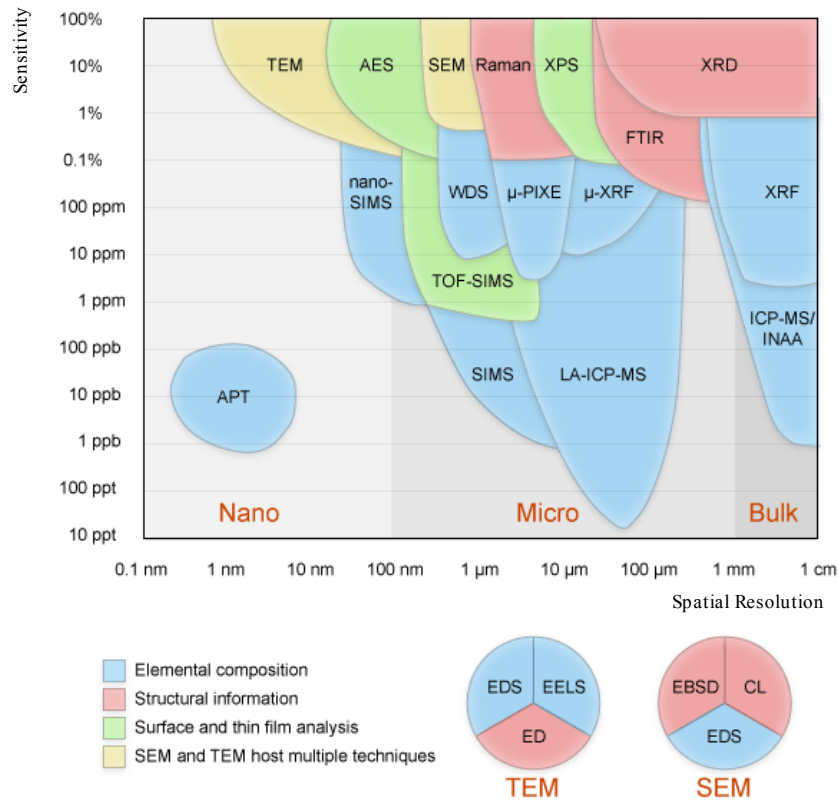


Figure 1.3 Résolutions spatiales et spectrales des microscopes [4]

Tous ces instruments, du principe le plus simple au plus complexe, ont chacun des centaines d'exemples d'utilisation, et ce dans différents domaines, tels qu'en pharmaceutique, en minéralurgie ou en plasturgie, pour ne nommer que ceux-là.

Or, dans toutes les analyses par imagerie multispectrale, le but est de trouver des zones spatiales cohérentes selon une signature spectrale tout aussi cohérente. Ainsi, d'une grande quantité de données qui sont générées par l'instrument, on veut en quelque sorte regrouper et synthétiser l'information. Cette transformation d'une grande quantité de données en information concise et cohérente permet ensuite à toute personne d'être en mesure d'interpréter les résultats, tâche impossible à accomplir sans des outils numériques adaptés à la grande quantité de données brutes générées par l'instrument.

1.2.2 Méthodes de traitements des données couramment utilisées

Différentes techniques sont couramment utilisées afin de traiter les données obtenues d'une analyse par imagerie multispectrale. Certains instruments génèrent généralement des données avec un bon ratio signal sur bruit avec un court temps d'acquisition. Dans ces cas-ci, bien qu'extirper l'information de la grande quantité de données demeure difficile, il y a beaucoup moins d'embûches liées à une grande présence de bruit.

En microscopie, plusieurs techniques fournissent des données avec un très faible ratio signal sur bruit. Le traitement des données ne peut donc s'effectuer par des techniques classiques. Souvent, le signal est tellement faible qu'on sacrifie la résolution spatiale en additionnant les pixels entre eux afin d'obtenir un signal relativement plus élevé. [5]

Parmi les techniques classiques, l'analyse en composantes principales (*sous-section 3.1.2*) est largement utilisée comme technique permettant d'extraire de l'information cohérente de données relativement bruitées. L'information est considérée comme étant cohérente quand celle-ci contient des spectres typiques du microscope utilisé et que chacun des spectres est relié à une cartographie de concentrations.

Plus traditionnellement, les données d'imagerie multispectrale sont analysées par élément chimique. Dans cette analyse, l'utilisateur sélectionne une plage d'énergie qui correspond à la présence de pics caractéristiques associés à des éléments chimiques d'intérêts. Ainsi, pour chaque pixel, tous les comptes compris dans cette plage d'énergie sont sommés, ce qui résulte en une carte en deux dimensions de l'intensité perçue pour un élément chimique donné.

Depuis les deux dernières décennies, certaines techniques d'analyse multivariées se sont globalement répandues à travers plusieurs méthodologies d'analyse de données, telles que dans les exemples d'utilisation d'imagerie spectrale et de chimiométrie présentés précédemment. De plus, ces algorithmes sont implémentés dans plusieurs logiciels, tels que *CasaXPS* pour l'analyse de données d'imagerie multispectrale provenant de spectroscopie photoélectronique X [6], ou les logiciels libres *HyperSpy* et *Cornell Spectrum Imager*, qui permettent, entre autres, le traitement de données provenant de spectroscopie des pertes d'énergie [7], [8]. Ces deux logiciels permettent d'effectuer, entre autres, une analyse en composantes principales (*sous-section 3.1.2*).

1.2.3 Désavantages des méthodes couramment utilisées

La sous-section précédente a fait part de recours couramment utilisés pour permettre d'extraire de l'information parmi une centaine de données d'un faible ratio signal sur bruit. Notamment, il va sans dire que l'addition de pixels entre eux occasionne une perte de résolution spatiale. Ainsi, des éléments qui seraient spatialement à la limite détectable par l'instrument ne pourront être détectés à cause de la perte de résolution.

De plus, les résultats fournis par une méthode chimiométrique ne permettent pas tout le temps une interprétation aisée. Par l'analyse par composantes principales (*Principal Component Analysis - PCA*), qui est une méthode couramment appliquée à la chimiométrie et dont il sera mentionné dans ce mémoire, on ne peut obtenir directement des signatures spectrales et des cartographies de concentrations reliées aux signatures spectrales. Ce type d'analyse demande à l'utilisateur des manipulations de données supplémentaires, ce qui diminue grandement la robustesse de cette technique. [9]

Un composant consiste à une zone spatiale reliée à un spectre. Dans le cas des analyses par élément chimique, on ne peut extraire de composants purs. On ne fait que cartographier spatialement l'intensité d'un élément chimique pour ensuite, par association, déduire les composants. De plus, cette technique nécessite une intervention humaine, dans laquelle on doit définir la plage d'énergie dans laquelle le pic relié à l'élément chimique devrait apparaître, ce qui ajoute de la subjectivité à l'analyse.

Enfin, les techniques de traitement des données couramment utilisées n'utilisent pas les données à leur plein potentiel. Notamment, le comportement du bruit, qui est généralement connu au niveau de sa distribution statistique, n'est nullement pris en compte dans les techniques précédemment mentionnées. De plus, le bruit de fond des spectres est considéré comme une nuisance, alors que ce bruit de fond peut représenter jusqu'à 90 % de l'intensité perçue.

1.3 Objectifs du projet de recherche

La section 1.2 a fait part de la difficulté d'analyser des données d'imagerie spectrale en microscopie. Les principales difficultés viennent du fait que les ensembles contiennent une très grande quantité de données et que celles-ci ont un très faible ratio signal sur bruit.

Ainsi, ces difficultés causent des problématiques à extraire des spectres physiquement cohérents et dont chacun d'entre eux est relié à une cartographie spatiale. Considérant ces problématiques et considérant qu'elles ne peuvent être résolues par les méthodes couramment utilisées, les objectifs du projet de recherche sont les suivants :

- Caractériser les ensembles de données d'imagerie multispectrale générés par divers instruments de spectrométrie ;
- Développer une méthodologie basée le traitement de signal et sur des méthodes chimiométrique existantes afin de traiter les données recueillies en tenant compte de leurs caractéristiques ;
- Comparer la méthode développée aux méthodologies déjà existantes et généralement utilisées selon leur capacité à prédire les spectres et à prédire les concentrations liées à ces spectres.

La principale technique étudiée lors de ce projet de recherche est la spectrométrie photoélectronique X (*XPS*). Ce microscope permet des acquisitions de surface d'échantillon de relativement faibles résolutions spatiales, mais fournit une très grande résolution spectrale (*Figure 1.3*). Il permet par exemple de différencier différents oxydes d'un même métal. Or, le plus grand défi de ce type d'appareil est d'être en mesure de différencier les spectres des différents composés en présence.

Deux autres techniques seront aussi étudiées. La première est l'analyse dispersive en énergie (*EDX* ou *EDS*). Ce type de microscope a une grande résolution spatiale (*Figure 1.3*). Le défi dans l'analyse de données provenant de cet appareil est principalement d'être en mesure d'extraire les distributions spatiales reliées aux différents composés.

La deuxième est la spectroscopie des pertes d'énergie (*EELS*). Cette technique permet à la fois d'obtenir des ensembles de données d'une grande résolution spectrale et d'une grande

résolution spatiale (*Figure 1.3*). Or, les spectres, ainsi que les distributions spatiales sont très difficiles à extraire de ce type d'ensemble de données, ce qui en fait un bon défi pour les algorithmes qui seront présentés dans ce mémoire.

1.4 Structure du mémoire

Le Chapitre 2 résumera le fonctionnement du *XPS* et surtout les principes physiques associés et les liens entre ces principes et la méthode de traitement de données proposée.

Le Chapitre 3 fait l'état de l'art et de la théorie qui ont été à la base du développement de la nouvelle méthode de traitement de données proposée et présentée au Chapitre 4.

Une fois les fondements de la nouvelle méthode de traitement expliqués, l'application de celle-ci est démontrée. Ainsi, le Chapitre 5 démontre aussi les approximations qui peuvent être posées dans l'analyse de données *XPS* et l'effet que peuvent causer ces approximations. Le Chapitre 5 est le cœur de ce mémoire, dans lequel les résultats de l'application de l'algorithme sur des ensembles de données *XPS* sont montrés et sont comparés aux méthodes d'analyse plus traditionnelles.

Enfin, le Chapitre 6 démontre l'applicabilité de la méthode sur d'autres instruments que le *XPS*.

Chaque fin de chapitre contient une table des variables et une table des acronymes utilisés tout au long du chapitre en question.

Les annexes contiennent des compléments d'information par rapport à ce qui est présenté à travers les chapitres. Entre autres, codes et pseudocodes sont présentés afin de compléter l'explication des algorithmes expliqués dans le texte. De plus, le logiciel développé permettant une utilisation facile de l'algorithme expliqué dans ce mémoire est présenté en annexe. Enfin, une méthodologie de fabrication d'échantillons standards par photolithographie est aussi présentée en annexe.

2. CARACTÉRISTIQUES DE LA SPECTROMÉTRIE PHOTOÉLECTRONIQUE X (XPS) ET CONCEPTS PHYSIQUES ASSOCIÉS

Ce chapitre portera principalement sur les sections du *XPS* qui permettent l'acquisition et la lecture des photoélectrons. De plus, il sera question de la caractérisation et du comportement du bruit de fond dans les spectres *XPS*. La compréhension de ces deux derniers éléments est essentielle pour comprendre pourquoi l'algorithme développé (*Chapitre 3*) peut s'appliquer aux données *XPS*.

Les principes du *XPS* liés à la génération des photoélectrons ne seront pas abordés dans le chapitre, mais peuvent être consultés dans l'*Annexe E*.

2.1 Caractéristiques et fonctionnement du détecteur

2.1.1 Plaque de microcanaux

La première partie de ce détecteur consiste en une plaque de microcanaux. Chaque canal de cette plaque correspond à une coordonnée spatiale de l'image résultante. Les photoélectrons émis par le matériau, et ayant traversé l'analyseur hémisphérique concentrique ou l'analyseur en miroir sphérique dans le cas du *XPS* (*Annexe E*), atteignent l'entrée de ces différents microcanaux.

Que ce soit pour la détection d'ions, d'électrons ou de photons, les microcanaux fonctionnent tous sur le même principe. La particule analysée entre en contact avec la paroi du microcanal et, par cette collision, fait dégager un électron de la paroi, qu'on appelle électron libre. (*Figure 2.1*)

Cet électron libre entrera de nouveau en collision avec la paroi du microcanal. Ce phénomène fera dégager davantage d'électrons de la paroi, qu'on nommera électrons secondaires. Le microcanal qui est considéré comme une dynode continue. Ainsi, les électrons, qui sont accélérés par un champ magnétique, entreront eux aussi en collision avec la paroi, ce qui générera encore plus d'électrons. Cette multiplication d'électrons se produit jusqu'à la fin du microcanal. On parle d'une amplification du signal entrant par un facteur d'au moins 10^7 . [10]

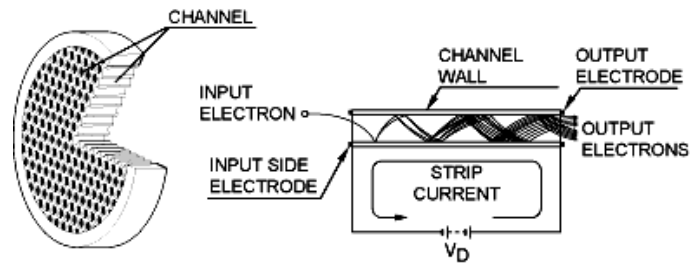


Figure 2.1 Plaque de microcanaux [11]

Les microcanaux ont donc pour fonction d'amplifier le signal, pour que celui-ci soit analysable.

2.1.2 Détecteur à délai de ligne

C'est dans le début des années 2000 que le détecteur à délai de ligne, plus connu sous son nom anglais *Delay Line Detector* ou son acronyme *DLD*, est apparu dans le marché et a commencé à être implémenté pour la spectroscopie des photoélectrons. [12]

Ce système a comme rôle de recueillir les tensions provenant des différents microcanaux et de transformer ces tensions en comptes de photoélectrons. [13]

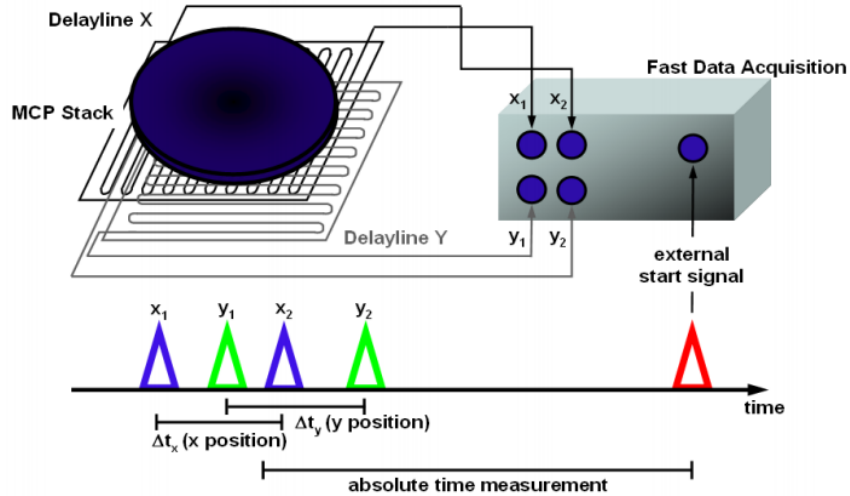


Figure 2.2 Détecteur à délai de ligne [10]

Une impulsion de départ est générée par les photoélectrons arrivant sur la plaque de microcanaux, ce qui génère un processus de synchronisation. Chaque tension accumulée dans un microcanal est dégagée dans deux serpentins. Ces deux serpentins font en sorte que les tensions provenant des différents canaux arrivent à des temps différents à la boîte d'acquisition. Cette dernière enregistre donc toutes les tensions reçues selon les différents temps à partir du début de l'acquisition. Par la suite, elle associe chaque tension perçue selon le temps pour calculer la tension totale perçue par chaque microcanal, et donc de chaque pixel. Enfin, la tension mesurée est directement proportionnelle à l'intensité du signal incident. Il est donc possible de retrouver le nombre de comptes incidents. [10], [12]

2.1.3 Particularités du détecteur

Bien que l'amplification comporte une certaine composante aléatoire [14], la tension générée est assez amplifiée pour être en mesure de bien calculer le nombre d'électrons réellement incidents. Le nombre d'électrons mesuré est donc directement proportionnel au nombre d'électrons incidents.

Au niveau du courant sombre (« dark current », en anglais, qui consiste à un léger signal électrique qui passe à travers les tubes multiplicateurs alors que l'obturateur est déployé), un fabricant de détecteurs à délai de ligne, *Surface Concept GmbH*, stipule que celui-ci est

« extrêmement » bas. Selon le manuel de leurs détecteurs, le nombre de comptes par secondes généré par le courant sombre serait d'au maximum 50 pour toute la surface pour le temps total d'acquisition. Il est à noter que plus le temps d'acquisition est long et plus ce courant est considérable dans la mesure. [15]

Il en va de même pour le bruit de lecture. Ce bruit, contrairement au courant sombre, n'est pas dépendant du temps et est surtout dépendant de la conception du détecteur. Encore là, ce bruit n'est aucunement considérable. [16]

2.2 Comportement des photoélectrons

2.2.1 De photon à photoélectron

Tel que mentionné précédemment, au *XPS*, un photon atteint donc le matériau analysé et interagit avec un électron de cœur en lui transmettant son énergie. Par ce transfert d'énergie, l'électron se détache de l'atome auquel il est attaché. Ce phénomène est appelé effet photoélectrique qui consiste, par définition, à toute émission d'électrons causée par une radiation. Un électron résultant de ce phénomène est appelé photoélectron. Le photoélectron a une énergie moindre que le photon incident au matériau, car une partie de son énergie est utilisée afin de promouvoir l'électron dans le vide, loin du pouvoir attractif des protons. [17]

2.2.2 Énergie analysée

Le photoélectron émis à une énergie cinétique E_{kin} qui est approximativement calculable par la différence de l'énergie de liaison entre l'atome et l'électron et l'énergie contenue dans le photon. Une certaine quantité d'énergie est perdue en considérant le travail de sortie de l'échantillon Φ_S , c'est-à-dire l'énergie nécessaire pour extraire l'électron du niveau de Fermi du matériau analysé. Une autre partie de l'énergie est perdue en considérant le travail de sortie de l'analyseur Φ_A . (*Figure 2.3*) [17]

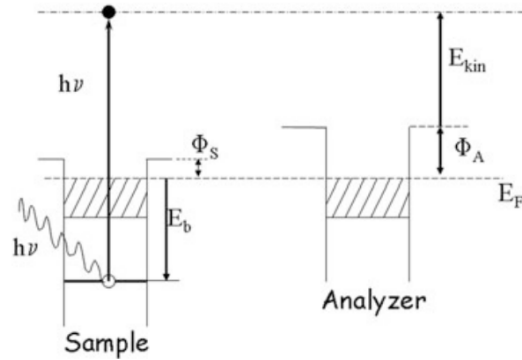


Figure 2.3 Schématisation de l'énergie du photoélectron du matériau à l'analyseur [17]

La Figure 2.3 illustre les énergies en jeu du matériau à l'analyseur. On constate qu'il n'est pas nécessaire de connaître le travail de sortie de l'échantillon, Φ_S , pour être en mesure de calculer l'énergie de liaison E_b . Seulement le travail de l'analyseur, Φ_A , est essentiel à connaître :

$$E_b = h\nu - E_{kin} - \Phi_A \quad (2.1)$$

Le travail de l'analyseur est constant et peut donc être calibré. $h\nu$, qui représente la quantité d'énergie contenue dans le rayon X, est aussi constant. [17]

2.2.3 Efficacité du détecteur

Tel que mentionné dans la sous-section précédente, une radiation est constituée de photons d'une énergie définie qui sont propagés. La propagation de ces photons ne se fait pas de façon continue. Les photons sont plutôt propagés de façon aléatoire. On parle alors de bruit de grenaille, de bruit quantique ou, en anglais, de *shot noise*. Ce bruit vient directement de la nature quantique de la lumière et se compare à une distribution de Poisson. [18] La principale caractéristique de cette distribution est que la moyenne d'un signal μ_s perçu égale sa variance σ_s^2 :

$$\mu_s = \sigma_s^2 \quad (2.2)$$

Les photons arrivent donc de façon aléatoire sur le matériau. Les photoélectrons seront donc émis eux aussi de façon aléatoire et ce, selon la même distribution de Poisson. Cependant, ce n'est pas la totalité des particules incidentes au microcanal qui généreront un électron libre. Une probabilité est associée à la création d'un électron libre par une particule et se nomme l'efficacité quantique. [18] Il ne faut pas confondre cette probabilité à la probabilité qu'un photoélectron ait une énergie donnée à la suite de l'excitation du matériau analysé. Cette dernière sera expliquée dans la sous-section 2.3.1 *Signification physique du bruit de fond*.

Aussi, l'aire associée à un pixel sur la plaque de microcanaux n'est pas totalement sensible aux rayons. Ainsi, une partie des rayons incidents au pixel ne généreront aucun électron libre (*Figure 2.4*) [18]. Pour contrer le premier point, plusieurs détecteurs comportent des microlentilles afin de faire converger les rayons vers la partie sensible, soit le microcanal. Cependant, la sensibilité variera selon l'angle du rayon incident. De plus, la transmittance d'une microlentille à une autre variera beaucoup, phénomène qui est causé par la non-homogénéité lors de leur fabrication. (*Figure 2.4*) [18]

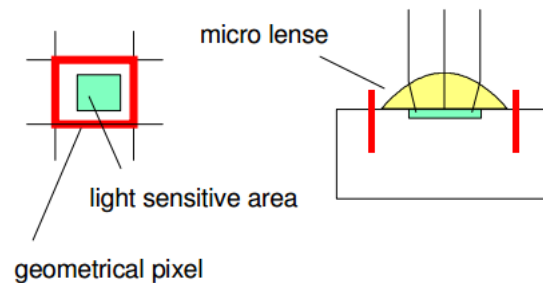


Figure 2.4 Facteurs géométriques expliquant [18]

Ce qui est important de retenir est que la quantité de rayons incidents varie spatialement et qu'il est même possible de constater un certain patron dans la sensibilité spatiale.

À titre d'exemple, la Figure 2.5 présente le patron de variation de sensibilité entre les pixels pour le Kratos XPS Ultra DLD de l'Université de Sherbrooke et dont les analyses de certains ensembles de données acquis par cet appareil seront présentées dans le Chapitre 5. Sur cet appareil, certains pixels sont deux fois moins sensibles que la sensibilité moyenne

des pixels, alors que d'autres pixels sont 50 % plus sensibles par rapport à la même moyenne.

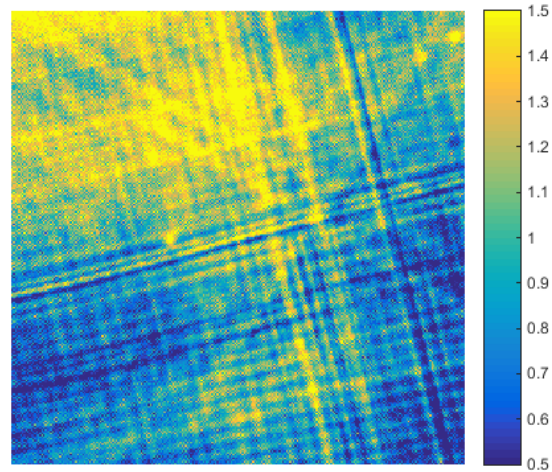


Figure 2.5 Sensibilité par rapport à la moyenne de sensibilité entre les pixels pour le *Kratos XPS Ultra DLD* de l'Université de Sherbrooke

Dans les méthodes traditionnelles d'analyse, il incombe de corriger les données brutes par un enregistrement de ce patron. Dans le jargon, on nomme habituellement ce patron *gain reference* ou *flat field correction*.

Si on combine l'efficacité quantique aux facteurs géométriques mentionnés dans les précédents paragraphes, on parle alors d'efficacité quantique totale, nommée par la variable η . [18]

Malgré ce facteur d'efficacité η , la moyenne du signal à la sortie du microcanal (μ_{sortie}) conservera une variance (σ_{sortie}^2) égale à sa moyenne, tel que le démontre le développement présenté à l'équation (2.3). [18]

$$\sigma_{sortie}^2 = \mu_{sortie} = \eta\mu_{signal} = \eta\sigma_{signal}^2 \quad (2.3)$$

En utilisant l'équation (2.3), il faut poser l'hypothèse que le bruit de lecture, ainsi que le courant sombre, sont nuls. Tel que mentionné dans la sous-section 2.1.3, ces deux bruits sont négligeables par rapport au bruit de grenaille. L'utilisation de cette hypothèse dans les algorithmes développés (*Chapitre 4*) ne crée donc qu'un très léger biais.

2.3 Bruit de fond

2.3.1 Signification physique du bruit de fond

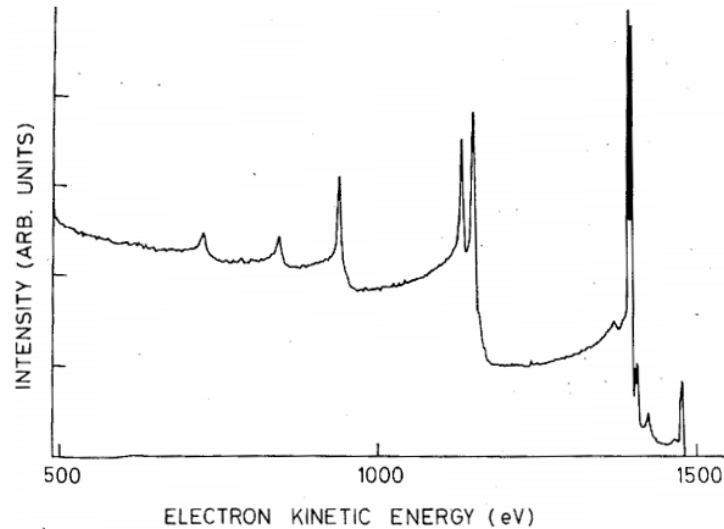


Figure 2.6 Exemple de spectre XPS [19]

Sur la Figure 2.6, on peut s'apercevoir, entre deux seuils, que l'intensité, à une énergie cinétique plus élevée, est moins importante qu'à une énergie cinétique plus faible. Cette observation résulte du fait que les pics sont toujours superposés à un fond. Ce fond est dû à la diffusion inélastique que les photoélectrons subissent entre leur point d'excitation et la surface du solide. [20], [21]

Les rayons X émis proviennent d'une couche ayant une profondeur d'environ 10 nm de la surface de l'échantillon analysé. Or, les électrons émis à une profondeur plus importante entreront en collision avec d'autres électrons jusqu'à la surface du matériau. On parle alors de diffusion inélastique qui se décrit comme une diffusion où l'énergie propagée par le système est moindre que celle incidente. [20], [21]

La probabilité qu'un électron à une énergie donnée E_0 puisse perdre une certaine quantité donnée ΔE peut être exprimée par une courbe dont l'allure est illustrée par la Figure 2.7.

Bien que cette courbe s'apparente à une distribution de Fischer. Une distribution Beta serait davantage appropriée. [20]

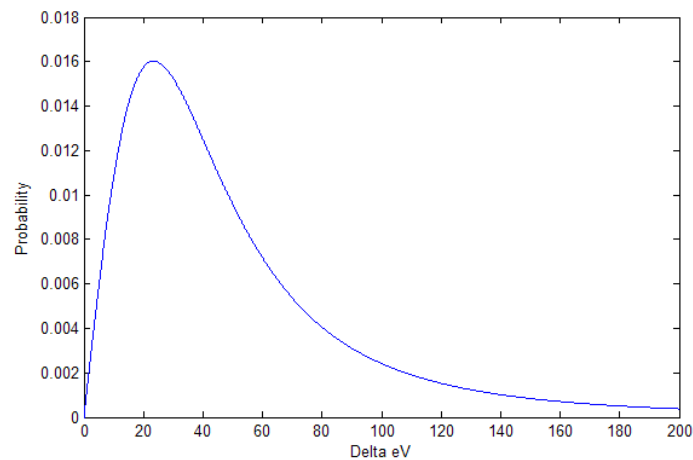


Figure 2.7 Exemple de courbe de probabilité de perte d'énergie due à la collision inélastique

Le phénomène de collision inélastique vient donc en quelque sorte brouiller les cartes. Les photoélectrons, qui perdent une partie de leur énergie de façon aléatoire, créent donc des spectres qui ne correspondent pas exactement à ce qui devrait être acquis. En analysant la *Figure 2.6*, on peut constater que le fond continu de part et d'autre des pics est asymétrique : l'intensité du fond continu est plus élevée à des énergies cinétiques plus faibles qu'à des énergies cinétiques plus élevées. [22]

Le fond des spectres, dû à la diffusion inélastique, doit donc être retiré en le calculant à partir de spectres acquis. Deux principales méthodes sont utilisées dans la littérature. L'une s'appelle le fond de Shirley, alors que l'autre se nomme le fond de Tougaard. [23]

4.3.2 Bruit de fond de Shirley

Ce calcul de fond a été utilisé pour la première fois en 1972, soit une quinzaine d'années avant les publications de Tougaard. [24] Dans les prochaines lignes sera expliquée la méthode de calcul du bruit de fond de Shirley. La figure ci-dessous définit les différentes variables utilisées :

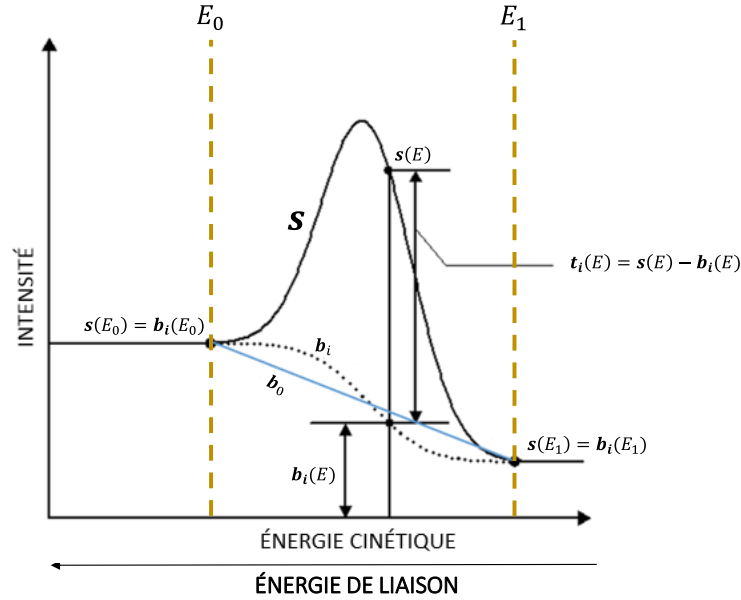


Figure 2.8 Définition de variables dans l'équation du fond de Shirley

Le fond de Shirley n'est qu'une méthode empirique et n'a aucune signification physique. À une énergie cinétique E donnée, le fond à cette énergie se calcule selon les valeurs du fond et du spectre aux énergies cinétiques inférieures :

$$\mathbf{b}_i(E) = s(E_1) + l_{i-1} \int_E^{E_1} [s(E') - \mathbf{b}_{i-1}(E')] dE' \quad (2.4)$$

Dans l'équation précédente, l est une constante. Le calcul du fond de Shirley est basé sur un algorithme itératif. Au départ, ni l , ni \mathbf{b} ne sont définis. La constante l est donc initialisée selon l'intensité moyenne du spectre \mathbf{s} et \mathbf{b} est initialisé comme une droite entre les points E_0 et E_1 . (Figure 2.8) Ainsi, à chaque itération i , un nouveau fond \mathbf{b}_i est calculé et l est mis à jour selon la valeur de \mathbf{b}_i au point E_0 . On déclare l'algorithme comme ayant convergé lorsque $\mathbf{b}(E_0) \approx \mathbf{s}(E_0)$. L'intégrale dans l'équation (2.4) est la plupart du temps approximé par la méthode des trapèzes. [22], [24]–[26]

2.3.3 Bruit de fond de Tougaard

Le fond de Tougaard nécessite un calcul quelque peu différent à celui du fond de Shirley. Il ne se calcule pas itérativement et met en jeu certaines constantes caractéristiques du matériau analysé.

Dans l'équation (2.4), la constante l est modifiée de façon itérative et est considérée comme constante, peu importe la valeur de E' dans l'intégrale. Dans le calcul du bruit de fond de Tougaard, on ne considère pas l comme étant une constante, mais plutôt une variable qui prend en compte la probabilité d'une perte d'énergie (Figure 2.8) et du type de matériau. La figure ci-dessous présente les différentes variables utilisées :

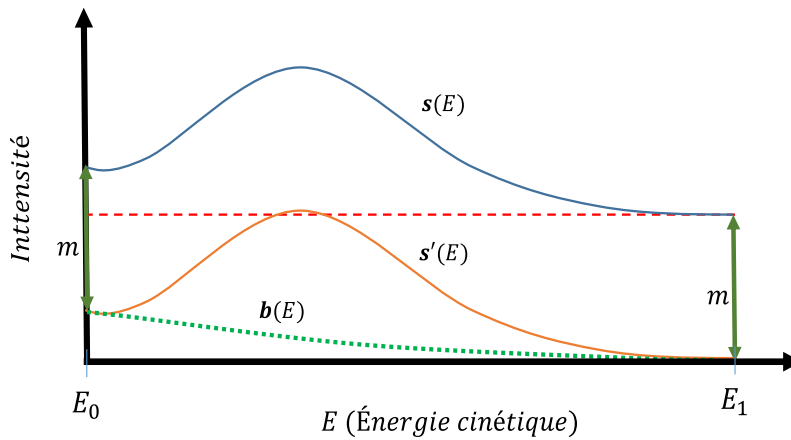


Figure 2.9 Définition des variables dans les équations pour calculer le fond de Tougaard

Dans le calcul du spectre de Tougaard, on soustrait toutes les valeurs de s par la valeur la plus basse comprise dans s . Cette opération consiste à retirer toute l'intensité mesurée qui serait causée par des collisions inélastiques de pics étant à plus haute énergie cinétique. Après cette opération de soustraction effectuée, on définira le nouveau spectre comme étant s' . [19], [20], [27]

L'équation principale permettant de calculer le bruit de fond de Tougaard est la suivante :

$$\mathbf{b}(E) = \int_E^{E_1} [W(E' - E)\mathbf{s}'(E')] dE' \quad (2.5)$$

$W(E' - E)$ dans l'équation (2.5) est la variable qui représente la probabilité de recevoir un électron d'une énergie cinétique initiale supérieure à E . (*Figure 2.7*) On utilise généralement l'équation de section efficace universelle à deux paramètres, soit l'équation (2.6), afin de calculer W pour la plupart des métaux et de leurs oxydes. [21]

$$W(T) = \frac{BT}{(C + T^2)^2} \quad (2.6)$$

Dans l'équation (2.6), T correspond à une différence en énergie cinétique, alors que B et C sont des constantes caractéristiques d'un type de matériau. $W(T)$ permet donc de calculer une simili probabilité (*Figure 2.7*) de recevoir un électron à une énergie E , mais qui aurait eu une énergie cinétique initiale de E' . [20], [25]

L'équation de section efficace universelle à trois paramètres est prescrite pour les polymères, le silicium, le germanium et l'aluminium. Pour chacun de ces composés ou familles de composés, en plus de devoir utiliser l'équation (2.7), les trois constantes recommandées sont différentes dans chacun des cas. À noter que tous les autres métaux et oxydes requièrent les mêmes constantes.

$$W(T) = \frac{BT}{(C - T^2)^2 + DT^2} \quad (2.7)$$

C'est en considérant toutes les énergies supérieures au point E et en considérant l'intensité perçue (s'), qu'il est possible de définir la valeur du bruit de fond à ce point E , tel que l'équation (2.5) le décrit.

Enfin, les constantes B et C peuvent être obtenues par expérimentation ou peuvent facilement être trouvées dans la littérature. On considère que le bruit de fond pour des spectres représentant des métaux ou leurs oxydes peut être calculé par les mêmes valeurs de B et C , et ce avec l'équation (2.6). [21]

2.3.4 Comparaison entre les deux principales méthodes du calcul du bruit de fond

Le fond de Shirley varie beaucoup en fonction de la sélection du début et de la fin du fond en termes d'énergie de liaison, alors que le fond de Tougaard est beaucoup plus stable, peu importe la sélection de l'étendue du fond. (Figure 2.10) [28]

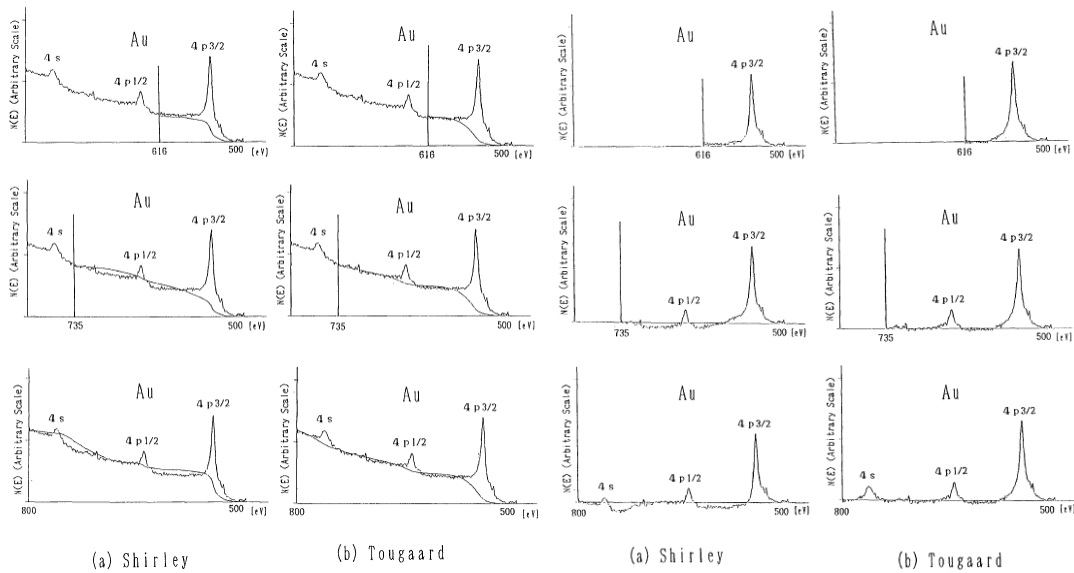


Figure 2.10 Comparaison entre des fonds de Shirley et de Tougaard [28]

En plus d'une meilleure stabilité, et tel que mentionné auparavant, le bruit de fond de Tougaard se base sur de réels fondements physiques. Le bruit de fond de Shirley n'est qu'une méthode empirique, mais qui ne nécessite aucune information au préalable. Plusieurs erreurs d'interprétation et d'analyse quantitative causées par l'utilisation du fond de Shirley ont été repérées. [20], [28]

2.3.5 Approximation de bruit de fond d'un mélange de différents composants

La section 3.1 *Analyse multivariable*, de même que le *Chapitre 4*, présenteront des méthodes d'analyse de données dans lesquelles on extrait un spectre d'un ensemble de données très bruitées. On associe chaque spectre à un composant. Ces spectres incluent le bruit de fond.

Dans les méthodes qui seront présentées, on considère que chaque spectre d'un composant est « pur », c'est-à-dire qu'on considère qu'il correspond physiquement à un seul composé

sur le matériau. En spectrométrie photoélectronique X, on définit un composé comme étant une molécule ou un mélange de molécules qui ne peuvent être dissociées.

Toujours dans ces mêmes méthodes, on calcule un spectre résultant d'un mélange de différents composants de façon linéaire. Il est à rappeler que les spectres contiennent un bruit de fond. On peut donc se demander si le bruit de fond peut être représenté comme une contribution linéaire du bruit de fond des composants individuels. L'*Annexe C* fait la démonstration mathématique que le calcul linéaire d'un spectre résultant d'un mélange de différents composants est approprié en posant certaines hypothèses :

- Un bruit de fond calculé par la méthode de Tougaard ou de Shirley est considéré comme correctement approximé;
- Le spectre, avec bruit de fond, calculé de chacun des composants est considéré comme pur;
- Le spectre de chacun des composants dont on enlève le bruit de fond (***b***), est considéré comme ne comportant aucune erreur;
- Le spectre sans bruit de fond résultant d'un mélange de composés se calcule linéairement selon les concentrations des différents composants en présence dans le mélange et selon les spectres purs sans bruit de fond (***b***) de ces composants.

Si on utilise la méthode de Tougaard pour le calcul du bruit de fond, qui est d'ailleurs plus robuste et plus physiquement significatif, on doit considérer cette contrainte afin d'approximer le calcul d'un spectre de différents composants de façon linéaire :

- Tous les composés en présence doivent avoir une section efficace semblable (mêmes constantes).

Tel qu'affirmé dans la sous-section 2.5.3, on considère que l'ensemble des collisions inélastiques sont semblables dans des composés de même type. Ainsi, si on analyse, par exemple, un matériau qui ne consiste que de métaux ou d'oxydes de métaux, $W(E' - E)$ dans l'équation (2.5) se calcule de la même façon avec les mêmes constantes *B* et *C*. On peut donc considérer le calcul linéaire d'un spectre d'un mélange de composants comme étant une approximation valable. Si on analyse un matériau qui, par exemple, est composé

de métaux et de polymères, l'approximation est beaucoup moins valable, car $W(E' - E)$ se calcule différemment entre un métal et un polymère.

Toujours avec la méthode de Tougaard, on considère que la valeur $s_{mél}(E)$ résultant d'un mélange de composants se calcule de façon linéaire par rapport à la contribution c_k de chacune des K composants et aux valeurs spectrales de $s_k(E)$ de chacun des K composants :

$$s_{mél}(E) = \sum_{k=1}^K c_k \times s_k(E) \quad (2.8)$$

À noter que la somme des valeurs de c pour toutes les K composants doit être de 1 et que toutes les valeurs doivent être positives.

2.3 Résumé du Chapitre 2

Le Chapitre 2 s'est premièrement penché sur le fonctionnement des détecteurs à délai de ligne, dont ses principales caractéristiques est sa rapidité à acquérir de l'information tout en gardant le dimensionnement spatial des différentes données, en plus d'avoir un très faible bruit de lecture.

Quant aux photoélectrons, on retient qu'ils sont créés par effet photoélectrique par les rayons X incidents au matériau analysé. Tout comme les photons, leur incidence n'est pas continue, mais se caractérise plutôt par une loi de Poisson.

Enfin, le bruit de fond est présent dans un spectre *XPS*. Deux principales méthodes existent dans la littérature pour retirer ce bruit de fond, soit la méthode de Shirley et celle de Tougaard. Cependant, dans le cadre de ce travail, on démontrera qu'un spectre résultant d'un mélange de différents composants peut être adéquatement expliqué par une combinaison linéaire de la contribution de chacun des composants.

Tableau 2.1 Acronyme utilisé dans le Chapitre 2

| Acronyme | Correspondance en anglais | Correspondance en français |
|------------|---|--------------------------------------|
| <i>XPS</i> | <i>X-Ray Photoelectric Spectroscopy</i> | Spectrométrie photoélectronique X |

Tableau 2.2 Variables utilisées dans le Chapitre 2

| Variable | Définition |
|---------------------|---|
| Φ_S, Φ_A | Travail de sortie de l'échantillon (<i>S</i>) ; de l'analyseur (<i>A</i>) |
| $h\nu$ | Énergie des rayons X incidents |
| E_b | Énergie de liaison |
| E_{kin} | Énergie cinétique |
| μ_{signal} | Moyenne de l'intensité d'un signal |
| σ_{signal}^2 | Variance de l'intensité d'un signal |
| η | Efficacité quantique totale |
| s | Spectre mesuré |
| s' | Spectre mesuré dont toutes les valeurs ont été soustraites par la plus petite valeur du spectre s |
| t | Spectre sans bruit de fond |
| b | Bruit de fond spectral |
| E | Énergie cinétique |
| E_0 | Plus basse énergie cinétique à laquelle le bruit de fond est considéré |
| E_1 | Plus haute énergie cinétique à laquelle le bruit de fond est considéré |
| l | Constante utilisée dans le calcul du bruit de fond de Shirley |
| K | Nombre de composants à extraire |
| W | Fonction de section efficace |
| B, C | Constantes dans la fonction efficace W |
| c_k | Contribution du composant k |

3. ALGORITHMES ET MÉTHODES

3.1 Analyse multivariable

Tel qu'il a été mentionné dans l'introduction, certaines techniques sont couramment utilisées afin d'être en mesure d'extraire de l'information pertinente dans une marée de données. Les prochaines sous-sections feront part de techniques qui sont implémentées dans plusieurs logiciels et méthodologies.

Les prochaines sous-sections expliqueront différents concepts et méthodes afin d'analyser un ensemble de données \mathbf{X} . Concrètement, \mathbf{X} est une matrice de deux dimensions, dont la première dimension a une taille M et correspond au nombre d'observations. La deuxième dimension est de taille N et correspond au nombre de variables. (Figure 3.1)

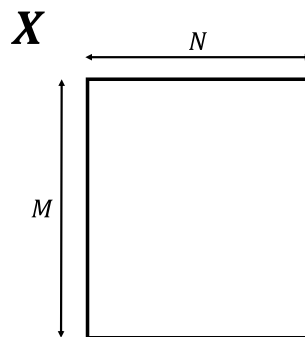


Figure 3.1 Représentation générale d'un ensemble de données

Dans le cas d'un ensemble de données d'imagerie multispectrale, c'est plutôt une matrice en trois dimensions, $\underline{\mathbf{X}}$, qui représente l'ensemble des données acquises et dont les différentes dimensions sont définies par la Figure 3.2.

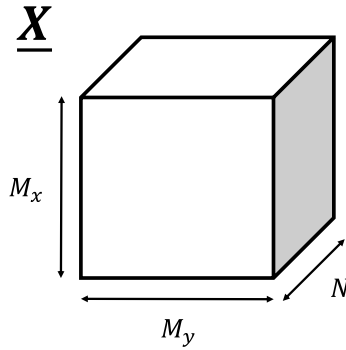


Figure 3.2 Représentation d'un ensemble de données d'imagerie multispectrale

Dans la Figure 3.2, M_x et M_y correspondent aux dimensions spatiales, alors que N correspond aux différents niveaux d'énergie.

Les analyses multivariées dont il sera question dans les prochaines sous-sections permettent d'analyser des données qui sont réparties dans une matrice de deux dimensions (Figure 3.1). L'ensemble de données d'imagerie spectrale, $\underline{\mathbf{X}}$, doit donc être converti en matrice ne possédant que deux dimensions \mathbf{X} . Pour ce, il faut fusionner les deux dimensions spatiales, M_x et M_y , en une seule dimension. Ainsi, à chaque niveau d'énergie est associée une image et celle-ci est « déroulée » pour créer un vecteur vertical. La Figure 3.3 illustre un exemple de transformation de matrice en trois dimensions en matrice à deux dimensions.

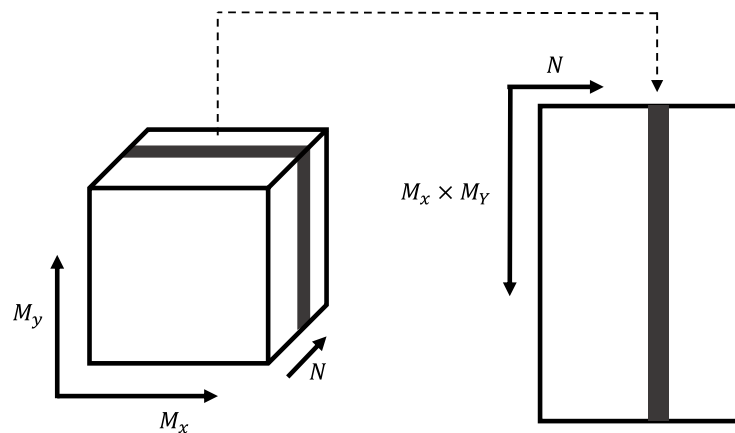


Figure 3.3 Représentation de la transformation de données d'imagerie spectrale en matrice à deux dimensions

En effectuant cette transformation de matrice de trois dimensions en matrice de deux dimensions, l'information spatiale est mise de côté. Or, connaissant la répartition spatiale des pixels à travers chaque vecteur, l'information spatiale peut par la suite être réintroduite à la fin des calculs.

3.1.1 Prétraitement des données

En analyse multivariable, le prétraitement des données est primordial afin d'être en mesure de bien extraire l'information et de réduire les impacts que pourraient engendrer de différentes variabilités entre les variables ou entre les observations. Si un prétraitement est inadéquat, il se peut que certaines variables ou observations soient favorisées dans l'analyse multivariable au détriment des autres variables ou observations.

L'équation (3.1) démontre le prétraitement le plus couramment utilisé qui est de centrer-réduire les données par rapport à chacune des variables. Pour utiliser ce prétraitement, on doit poser l'hypothèse que les données sont normalement distribuées. [29]

$$\tilde{X}_{ij} = \frac{X_{ij} - X_{.j}}{\text{std}(X_j)} \quad (3.1)$$

Dans l'équation (3.1), l'indice i correspond à la position sur la première dimension de \mathbf{X} , soit celle ayant comme dimension $M_X \times M_Y$ (Figure 3.3), alors que l'indice j correspond à la position sur la deuxième dimension de \mathbf{X} , qui a comme dimension N (Figure 3.3). Les « . » signifient une moyenne. Ainsi, $X_{.j}$ correspond à la moyenne des données pour la variable j . Enfin, $\text{std}(X_j)$ correspond à l'écart-type entre toutes les données de la variable j .

Ce prétraitement met toutes les variables sur le même pied d'égalité. La moyenne devient nulle entre toutes les données associées à une variable donnée, alors que l'écart-type entre ces mêmes données sera de 1. Ainsi, une différence significative entre les variables (dans l'ordre de grandeur des variables ou dans leurs unités) n'aura donc pas d'impact sur le calcul. [29]

Par ailleurs, l'équation (3.2) est un des prétraitements qui a été développé pour la spectrométrie de masse à ionisation secondaire et diffère de prétraitement de l'équation (3.1). [30], [31]

$$\tilde{\mathbf{X}}_{ij} = \frac{\mathbf{X}_{ij}}{\sqrt{\mathbf{X}_{i.}}\sqrt{\mathbf{X}_{.j}}} \quad (3.2)$$

Tel que le décrit l'équation (3.2), chaque donnée \mathbf{X}_{ij} est divisée par la moyenne du spectre i correspondant à la donnée et par la moyenne de la variable j correspond à la donnée.

Ce prétraitement est connu sous le nom d'*Optimal Scaling*. Dans la même optique que le prétraitement centrant et réduisant les données, ce prétraitement permet de mettre sous un même pied d'égalité toutes les données, mais en considérant un bruit poissonnien, si les données analysées sont bruitées selon ce type de bruit. D'ailleurs, ce prétraitement est basé sur l'approximation de la variance d'une donnée basée sur l'équation (3.3). [30], [31]

$$var(\mathbf{X}_{ij}) = \sqrt{\mathbf{X}_{i.}}\sqrt{\mathbf{X}_{.j}} \quad (3.3)$$

Tel qu'expliqué précédemment, la variance d'une mesure corrompue par un bruit poissonnien égale la moyenne de cette mesure. L'équation (3.3) approxime à la fois la moyenne et la variance d'une donnée en considérant la moyenne de la variable correspondant ($\mathbf{X}_{.j}$) et la moyenne de l'observation/pixel correspondant ($\mathbf{X}_{i.}$). En mettant ces deux moyennes au carré et en les multipliant ensemble, on obtient une moyenne et variance approximée. Par l'équation (3.2), la donnée \mathbf{X}_{ij} est donc divisée par cette variance approximée $var(\mathbf{X}_{ij})$. C'est par ces approximations que les variances entre toutes les données sont égalisées. [30], [31]

Il ne serait pas approprié d'utiliser ce prétraitement dans des données bruitées par une autre loi que la loi de Poisson. Avec une autre loi, la variance entre les données n'égale pas la moyenne entre ces données. Généralement, la valeur de la variance est inférieure à celle de la moyenne. Si, par exemple, une variable j contient des données en moyenne plus élevées

que les autres, alors, en se fiant à l'équation à l'équation (3.3), la variance serait surestimée. En se fiant maintenant à l'équation (3.2), toutes les données associées à la variable j en question seraient biaisement divisées par un plus gros nombre et, par conséquent, auraient moins d'impact dans les calculs subséquents.

Un autre prétraitement qui sera utilisé à travers les différents algorithmes et méthodes étudiées est décrit par cette équation :

$$\tilde{X}_{ij} = \frac{X_{ij}}{NX_i} \quad (3.4)$$

Ce prétraitement, qu'on nommera normalisation spectrale, a pour objectif de normaliser chacun des spectres. Cela permet donc de diminuer les erreurs causées par les variations de sensibilité entre les pixels et ainsi d'appliquer le même poids entre tous les pixels. Or, dans ce cas-ci, il n'y a pas de normalisation de variable. Par ce fait, certaines variables auront donc plus d'importance que d'autres dans l'analyse des données. Cependant, toutes les observations auront le même poids.

Lorsqu'une variable est centrée et réduite, environ la moitié des données deviennent négatives, si ces données sont bel et bien normalement distribuées. Un spectre, quel que soit l'instrument par lequel il a été acquis, ne contient généralement pas de données négatives. Ainsi, en transformant les données selon l'équation (3.1), on perd son sens physique et l'interprétation devient plus complexe. En utilisant les équations (3.2) ou (3.4), au contraire de l'équation (3.1), toutes les données restent positives, malgré leur transformation.

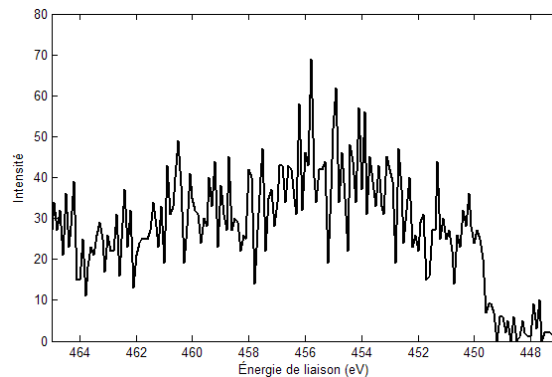


Figure 3.4 Exemple de spectre sans prétraitement d'une analyse par spectrométrie photoélectronique X

La Figure 3.4 illustre un exemple de spectre provenant d'un ensemble de données d'imagerie spectrale *XPS*. La Figure 3.5 démontre les différents prétraitements qui ont été présentés dans les dernières lignes et qui sont appliqués au spectre de la Figure 3.4.

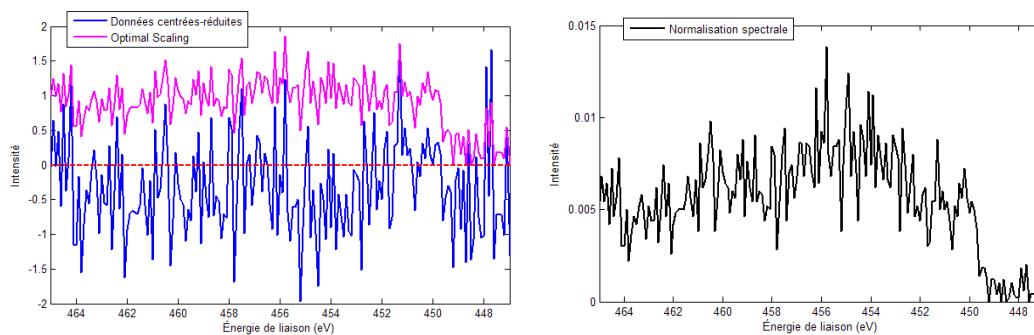


Figure 3.5 Exemples de prétraitement d'une analyse par spectrométrie photoélectronique X par (à gauche) donnée centrées et réduites, *Optimal Scaling* et (à droite) par normalisation spectrale

Ainsi, en analysant les différentes courbes de la Figure 3.4, on se rend compte que les spectres obtenus en centrant et réduisant les données ou en utilisant l'*Optimal Scaling* sont totalement méconnaissables par rapport aux spectres originaux. De plus, on constate que la grande majorité des données du spectre provenant des données centrées et réduites sont négatives. Si les données avaient réellement été bruitées selon une loi normale, environ la moitié des données auraient été négatives. Ce fait démontre bien qu'il est important de connaître la distribution statistique qui régit les données.

En analysant le spectre provenant de la normalisation spectrale, on constate que celui-ci garde tout à fait la même allure que le spectre original. La différence se constate en analysant l'axe de l'intensité. Comme la somme de toutes les valeurs d'un spectre prétraité par l'équation (3.4) est de un, les valeurs spectrales sont donc très petites.

Le tableau suivant résume les différents prétraitements qui seront utilisés tout au long de ce document :

Tableau 3.1 Comparaison entre les prétraitements étudiés

| Nom du prétraitement | Équation | Avantages | Désavantages |
|---------------------------|----------|--|---|
| Données centrées-réduites | (3.1) | <ul style="list-style-type: none"> - Toutes les variables sont prises en compte de la même façon - S'applique à tout type de données | <ul style="list-style-type: none"> - Ne prend pas en compte la variation de sensibilité d'un pixel à l'autre - Environ 50% des données deviennent négatives |
| <i>Optimal Scaling</i> | (3.2) | <ul style="list-style-type: none"> - Prises en compte du comportement poissonnien des données - Les données restent tous positives | <ul style="list-style-type: none"> - Ne s'applique que pour des données corrompues par un bruit de Poisson |
| Normalisation spectrale | (3.4) | <ul style="list-style-type: none"> - Atténue la différence de sensibilité entre les données - Les données demeurent positives | <ul style="list-style-type: none"> - Certaines variables pourraient avoir un poids plus élevé dans le modèle |

3.1.2 Analyse en composantes principales

L'analyse en composantes principales, *Principal Component Analysis* en anglais, et plus connue sous l'acronyme *PCA*, permet de décomposer un ensemble de données à plusieurs dimensions dans un espace qui comporte un nombre moindre de dimensions. Cette décomposition est obtenue en effectuant une, ou plusieurs, combinaison linéaire des variables d'origine afin de les représenter sous forme d'une seule, ou plusieurs, nouvelles variables. Cette transformation est particulièrement utile en présence de variables fortement corrélées et permet de limiter l'information redondante. Dans la plupart des cas, l'analyse est effectuée par des données qui ont été centrées et réduites a priori, soit avec l'équation (3.1). [29]

Cette décomposition s'effectue grâce à divers algorithmes, tels que l'algorithme *NIPALS* (*Non-linear iterative partial least squares*) ou la décomposition en valeurs singulières (*SVD*). [29]

La base de l'analyse en composantes principales (*voir Annexe B.1 pour pseudo-codes*) est la décomposition de la matrice de données \mathbf{X} (*Figure 3.1*) par l'équation présentée ci-dessous :

$$\mathbf{X} = \mathbf{T} \times \mathbf{\Sigma} \times \mathbf{P}^T + \mathbf{E} \quad (3.5)$$

Dans l'équation ci-dessus, \mathbf{X} représente une matrice de dimension $M \times N$ (*Figure 3.1*), \mathbf{T} une matrice de dimensions $M \times K$, où K est le nombre de composantes extraites, $\mathbf{\Sigma}$ une matrice diagonale de dimensions $K \times K$, \mathbf{P} une matrice de dimensions $K \times N$ et \mathbf{E} une matrice de même dimensions que \mathbf{X} . À noter que \mathbf{X} peut autant être une matrice avec ou sans prétraitements.

Une autre représentation de l'équation (3.5) consiste à intégrer « $\mathbf{T} \times \mathbf{\Sigma}$ » en une seule matrice. On nommera cette matrice résultante \mathbf{T}^* , qui possède les mêmes dimensions que \mathbf{T} , soit $M \times K$.

$$\mathbf{X} = \mathbf{T}^* \times \mathbf{P}^T + \mathbf{E} \quad (3.6)$$

En *PCA*, on nomme la matrice \mathbf{P} la matrice des *loadings*. Chaque ligne de cette matrice exprime la combinaison linéaire des variables de \mathbf{X} . Les *loadings* permettent en même temps de représenter la relation entre les variables. La matrice \mathbf{T} se nomme la matrice des *scores*. Chaque colonne de cette matrice représente les coordonnées des observations suivant le nouveau système de variables défini par le *loading* correspondant. Les valeurs contenues dans la diagonale de $\mathbf{\Sigma}$ représentent les valeurs singulières et se retrouvent en ordre décroissant. Ces valeurs expriment la variance expliquée de chaque composant. La matrice \mathbf{E} est la matrice d'erreur représentant l'erreur de prédiction par rapport à \mathbf{X} . [29], [32], [33]

Une façon d’imaginer la décomposition *PCA* est illustrée par la Figure 3.6, où les opérations matricielles ont été remplacées par la somme du produit du vecteur de *scores*, p , et du vecteur de *loadings*, t^* . Cette représentation exprime bien le fonctionnement de l’algorithme *NIPALS* (*Annexe B.1 pour pseudo-code*), dans lequel, à chaque itération, un seul vecteur de *loadings* et un seul vecteur de *scores* est calculé. [29], [32], [33]

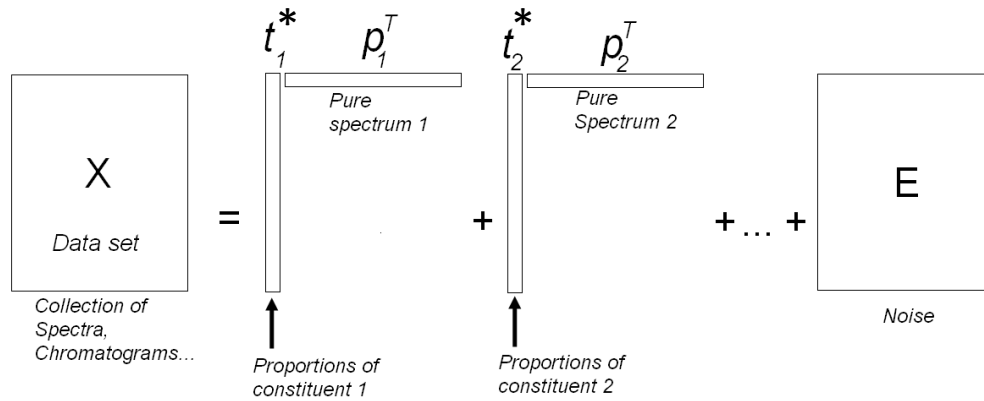


Figure 3.6 Décomposition d’une matrice de données par PCA [32]

3.1.3 Résolution multivariée de courbes – Moindres carrés alternatifs (*MCR-ALS*)

MCR-ALS est un acronyme signifiant *Multivariate Curve Resolution – Alternating Least Squares*. Pour utiliser cette méthode, il faut de nouveau une matrice 2D, \mathbf{X} (*Figure 3.1*), dont la dimension verticale de la matrice représente toutes les observations/pixels (M) et la dimension horizontale de la matrice représente les variables (N) (différents niveaux d’énergie). La décomposition *MCR-ALS* se traduit par cette équation :

$$\mathbf{X} = \mathbf{C} \times \mathbf{S}^T + \mathbf{E} \quad (3.7)$$

Dans l’équation (3.7), \mathbf{C} correspond à la matrice des contributions et a comme dimensions $M \times K$, où K correspond au nombre de composants à extraire. \mathbf{S} correspond à la matrice des spectres et a comme dimensions $N \times K$. Enfin, \mathbf{E} est la matrice d’erreur et a les mêmes dimensions que \mathbf{X} . À noter que \mathbf{X} peut autant être une matrice sans prétraitements qu’avec prétraitements.

La forme de l'équation est semblable entre (3.6) et (3.7). Les matrices ont les mêmes dimensions, mais sont définies par différentes variables. Il est important de bien cibler les différentes caractéristiques entre les matrices. Premièrement, \mathbf{S} contient de réels spectres, alors que \mathbf{P} n'exprime que des combinaisons des variables de \mathbf{X} . Deuxièmement, \mathbf{C} contient des contributions. Il n'y a donc pas présence de nombres négatifs, alors que \mathbf{T}^* représente les coordonnées résultantes des observations pour chaque composant selon les combinaisons linéaires des variables contenues dans \mathbf{P} . Il y a donc présence de nombres négatifs.

En plus des caractéristiques des matrices finales qui sont différentes entre *MCR* et *PCA*, les algorithmes se distinguent. Peu importe le nombre de composants K à extraire, les résultats obtenus avec *PCA* sont toujours les mêmes. En d'autres mots, qu'on extrait 3 ou 5 composants, le premier composant aura toujours les mêmes vecteurs de *score* et *loading*. Au contraire, avec *MCR*, le nombre de composants K à extraire influence beaucoup les résultats obtenus. L'algorithme *MCR-ALS* nécessitant une matrice spectrale initiale, le calcul de celle-ci est très important et peut influencer le résultat final. Il existe une infinité de solutions permettant de décrire \mathbf{X} par \mathbf{C} et \mathbf{P} . Le but de *MCR-ALS* est donc de converger vers la meilleure solution tout en respectant certaines contraintes.

Afin de bien initialiser, on tente d'extraire des courbes montrant le plus de distinction entre elles. Pour cela, on cherche à identifier les variables qui démontrent cette distinction. Prenons un exemple d'un spectre *XPS*. Si, à un certain niveau d'énergie de liaison, un composé démontre un pic, alors que l'autre n'en démontre pas, il y aura beaucoup de distinction en termes d'intensité entre ces deux composés. Il existe plusieurs algorithmes permettant cette approche, dont l'algorithme *SIMPLISMA*. (*sous-section 3.1.4*) [34]–[36]

Or, la principale différence entre *PCA* et *MCR-ALS* est le fait que ce dernier algorithme modifie continuellement les matrices \mathbf{S} et \mathbf{C} afin que ces matrices respectent des contraintes imposées.

Ainsi, suite à l'initialisation de la matrice spectrale, on atteint la boucle itérative. La matrice des spectres \mathbf{S} et la matrice des contributions \mathbf{C} sont continuellement calculées par rapport aux données originales contenues dans \mathbf{X} . Plus précisément, la matrice \mathbf{S} est calculée par

régression à partir de la matrice \mathbf{C} et de la matrice \mathbf{X} , alors que la matrice \mathbf{C} est calculée par régression à partir de la matrice \mathbf{S} et de la matrice \mathbf{X} . Ce processus itératif qui alterne entre deux différents calculs de régression se nomme la méthode d'alternance des moindres carrés. La méthode est plus connue sous son nom anglais, soit *Alternating Least Squares* (*ALS*). L'arrêt du processus itératif est basé sur la stabilisation des résidus de \mathbf{X} , qui diminuent habituellement au fil des itérations. [36]–[38]

En faisant le parallèle à l'analyse en composantes principales, expliquée à la section précédente, on peut relier la matrice \mathbf{S} à la matrice des *loadings* et la matrice \mathbf{C} à la matrice des *scores*.

Or, après chaque calcul, que ce soit \mathbf{C} ou \mathbf{S} , selon les contraintes mises en place, les matrices sont modifiées. Dans les contraintes les plus couramment utilisées, on peut imposer que toutes les valeurs des spectres soient positives, en plus de pouvoir imposer cette non-négativité aux valeurs de contributions. Une autre contrainte permet d'imposer une contribution totale égale pour chaque pixel, selon le nombre de composés à extraire. Enfin, une autre, appelée corrélation, permet d'imposer certaines valeurs suite à une calibration. [36]–[38] (*Figure 3.7*) Les contraintes à imposer varient selon le type d'analyse. Les contraintes appropriées dans le cadre de cette recherche seront énoncées et défendues dans le Chapitre 4.

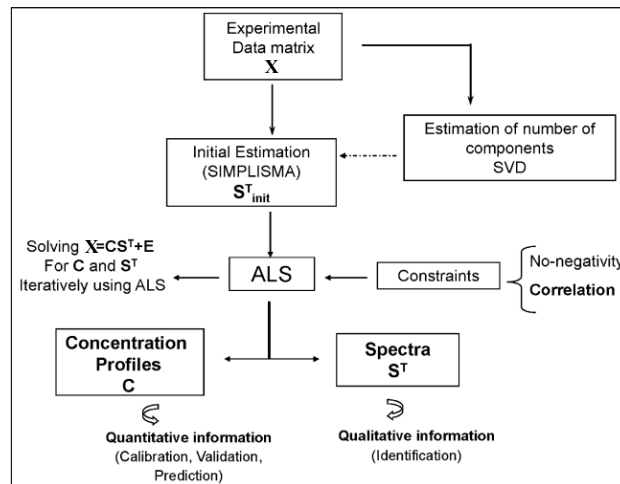


Figure 3.7 Schématisation de l'algorithme de MCR-ALS [39]

Tel que dit précédemment, les résultats obtenus par *MCR-ALS* (*Annexe B.2 pour pseudo-code*) différeront selon le nombre de composants à extraire, contrairement à *PCA* où on sélectionne le nombre de composantes à conserver après les calculs. Pour automatiser ce choix, des algorithmes de validation croisée permettent d'extraire le nombre approprié de composants selon les données. La validation croisée est une technique permettant de mesurer la robustesse et la précision d'un modèle.

L'algorithme *Element-wise k-fold*, appelé sous l'acronyme *EKF*, est fiable et peut être jumelé au *MCR-ALS*. Celui est couramment utilisé à des fins de validation croisée en analyse en composantes principales. [40]

D'autres méthodes que des algorithmes de validation croisée peuvent être utilisées. Une de ces méthodes est la sélection du nombre de composants par l'analyse des valeurs singulières en décomposition *SVD*. [39], [41]

3.1.4 SIMPLISMA

Tel que mentionné, *SIMPLISMA*, qui signifie *Simple-to-use Interactive Self-modeling Mixture Analysis*, est un algorithme qui permet de détecter les variables qui expliquent le plus de variance entre les données, en plus de calculer des spectres à partir des données de ces variables. [34], [35] Cet algorithme est souvent utilisée afin d'initialiser la matrice spectrale dans *MCR-ALS* (*sous-section 3.1.3*).

L'algorithme fonctionne de la façon suivante. On veut extraire K composants, donc K spectres. On initialise une matrice, \mathbf{M}_{purs} , qui est vide au départ et qui regroupera les données correspondant aux variables expliquant le plus de variance. [34], [35]

Pour chaque composant k , on ajoutera un vecteur-colonne, correspondant aux données associées à une variable, à la matrice \mathbf{M}_{purs} . On recherche la variable dont les données associées ont une grande variance, tout en se différenciant des autres données incluses dans \mathbf{M}_{purs} . [34], [35]

Tel que dit précédemment, \mathbf{M}_{purs} est vide au départ. Ainsi, pour trouver le 1^{er} vecteur-colonne à ajouter à cette matrice, on sélectionne tout simplement la variable dont les données associées ont la plus grande variance entre elles. [34], [35]

Pour les autres composants, c'est quelque peu plus complexe. Pour chaque variable n , on ajoute le vecteur-colonne correspondant aux données de la variable n à la matrice \mathbf{M}_{purs} . On définit cette matrice comme $\mathbf{M}_{test,n}$. (Figure 3.8)

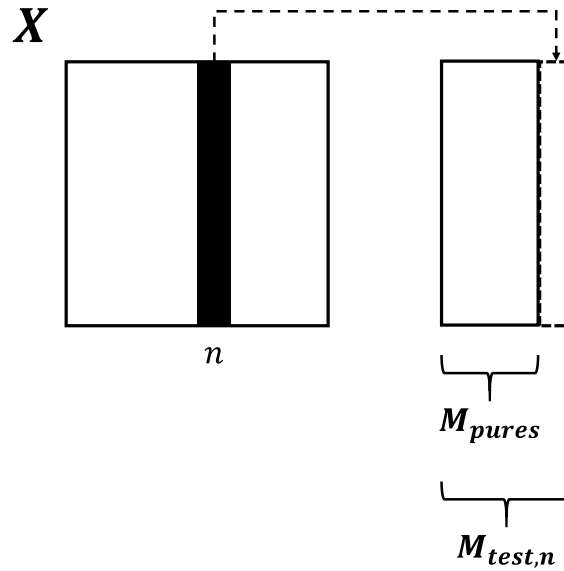


Figure 3.8 Création de la matrice $\mathbf{M}_{test,n}$

À partir de cette matrice $\mathbf{M}_{test,n}$, on calcule la somme des carrés des données et la somme des produits croisés ($\mathbf{M}_{test,n}^T \times \mathbf{M}_{test,n}$). Enfin, à partir du dernier calcul, on calcule le déterminant. On définira ce résultat par la variable w_n , où n correspond au numéro de variable dont les données associées à cette variable ont été ajoutées à \mathbf{M}_{purs} . Ces calculs sont illustrés par l'équation (3.8). [34], [35]

$$w_n = |\mathbf{M}_{test,n}^T \times \mathbf{M}_{test,n}| \quad (3.8)$$

Plus le déterminant a une valeur élevée et plus cela signifie que les données ont une grande variance et se différencient d'une variable à l'autre. Une fois que \mathbf{M}_{purs} contient K

colonnes, les spectres résultant se calculent par cette équation matricielle, dans laquelle \mathbf{X} correspond à la matrice des données :

$$\mathbf{S}_{purs} = (\mathbf{M}_{purs}^T \times \mathbf{M}_{purs})^{-1} \times \mathbf{M}_{purs}^T \times \mathbf{X} \quad (3.9)$$

La Figure 3.9 schématise le fonctionnement de l'algorithme *SIMPLISMA*, selon les explications qui ont été données dans les derniers paragraphes.

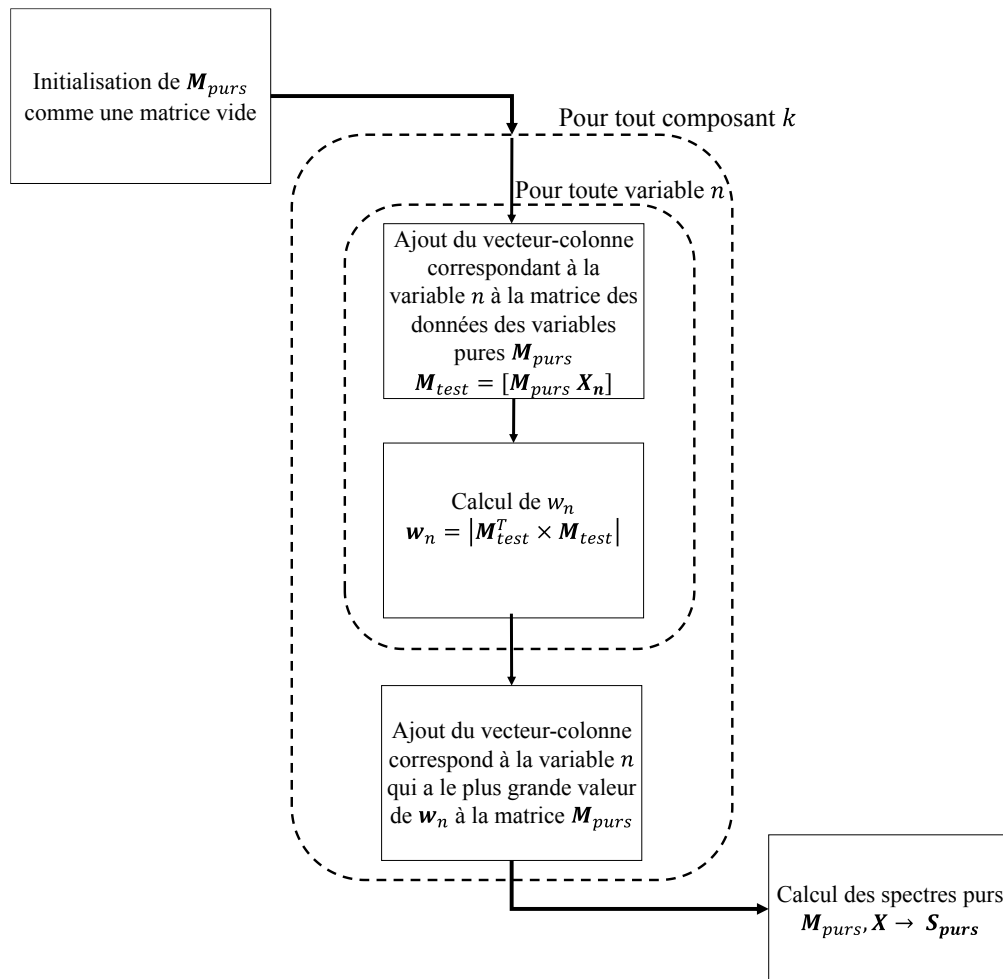


Figure 3.9 Algorithme *SIMPLISMA*

La Figure 3.10 montre le traitement de spectres par *SIMPLISMA*. On recherche trois composants ($K = 3$) et l'algorithme a été en mesure de déceler les trois variables (lignes verticales) qui permettent d'expliquer le plus de variance entre les résultats.

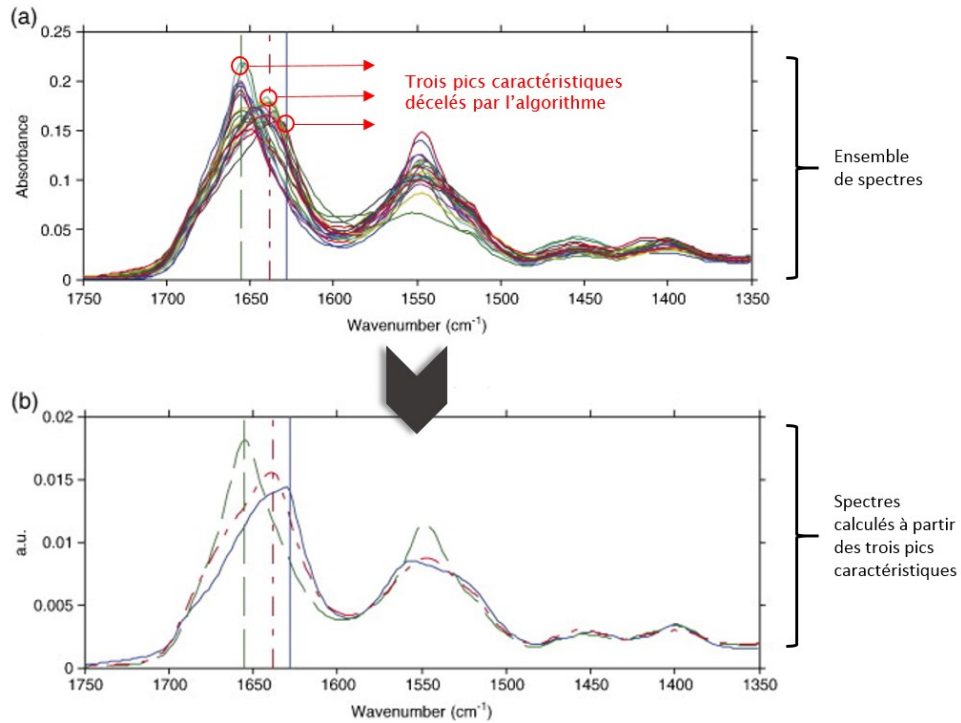


Figure 3.10 Exemple d'extraction de trois spectres d'un ensemble de spectres par *SIMPLISMA* [35]

3.1.5 Algorithme des k-moyennes

L'algorithme des k-moyennes, plus connu sous le nom anglais de l'algorithme des *k-means* est un algorithme itératif très largement utilisé et qui permet, selon un nombre donné de composants à extraire K , de grouper M observations dans K différents groupes. Les observations peuvent avoir une ou plusieurs dimensions, mais ne peuvent appartenir qu'à un seul groupe. [42] Cet algorithme est utilisé dans plusieurs domaines. Dans ce cas-ci, l'algorithme est utilisé pour l'initialisation de la matrice spectrale dans *MCR-ALS*.

Chaque groupe k est défini par un centroïde. Un centroïde est en fait les valeurs correspondant au centre du groupe (moyenne), et ce pour chaque dimension. On regroupera l'ensemble des valeurs représentant les centroïdes des K groupes dans la matrice \mathbf{O} . Cette matrice est de dimensions $K \times N$, N correspondant au nombre de variables dans l'ensemble de données. [42]

Pour associer une observation à un groupe ou à un autre, on calcule la distance de l'observation par rapport au centroïde de chacun des groupes. On utilise généralement la distance euclidienne. Ce type de distance se décrit par l'équation suivante :

$$D_{m,k} = \sqrt{\sum_{n=1}^N (X_{m,n} - O_{k,n})^2} \quad (3.10)$$

Dans l'équation précédente, $D_{m,k}$ correspond à la distance euclidienne entre l'observation m et le composant/groupe k . $X_{m,n}$ correspond à la donnée acquise pour l'observation m et la variable n , alors que $O_{k,n}$ correspond à la valeur du centroïde pour le groupe k et la variable n .

L'algorithme fonctionne sur le principe suivant. Pour initialiser le centroïde, K observations sont sélectionnées au hasard parmi les M observations. Ensuite, chacune des M observations est associée à une autre l'autre des K groupes dépendamment de la distance entre l'observation et le centroïde de ces différents groupes. [42]

Cette première classification effectuée, de nouvelles valeurs de centroïdes sont calculées. Pour ce, pour un groupe k donné, on calcule la moyenne de toutes les observations comprises dans ce groupe k pour chacune des variables.

$$O_{k,n} = \frac{\sum_{q=1}^{Q_k} X_{q,n}}{Q_k} \quad (3.11)$$

Dans l'équation (3.11), Q_k correspond au nombre de données incluses dans le composant/groupe k .

Ainsi, à partir des nouvelles valeurs des centroïdes, les observations sont reclassifiées, pour ensuite de nouveau calculer des valeurs de centroïdes. Il y a donc ici un processus itératif qui s'arrêtera lorsqu'il n'y aura plus de déplacement d'au moins une observation d'un groupe à un autre. (Figure 3.11)

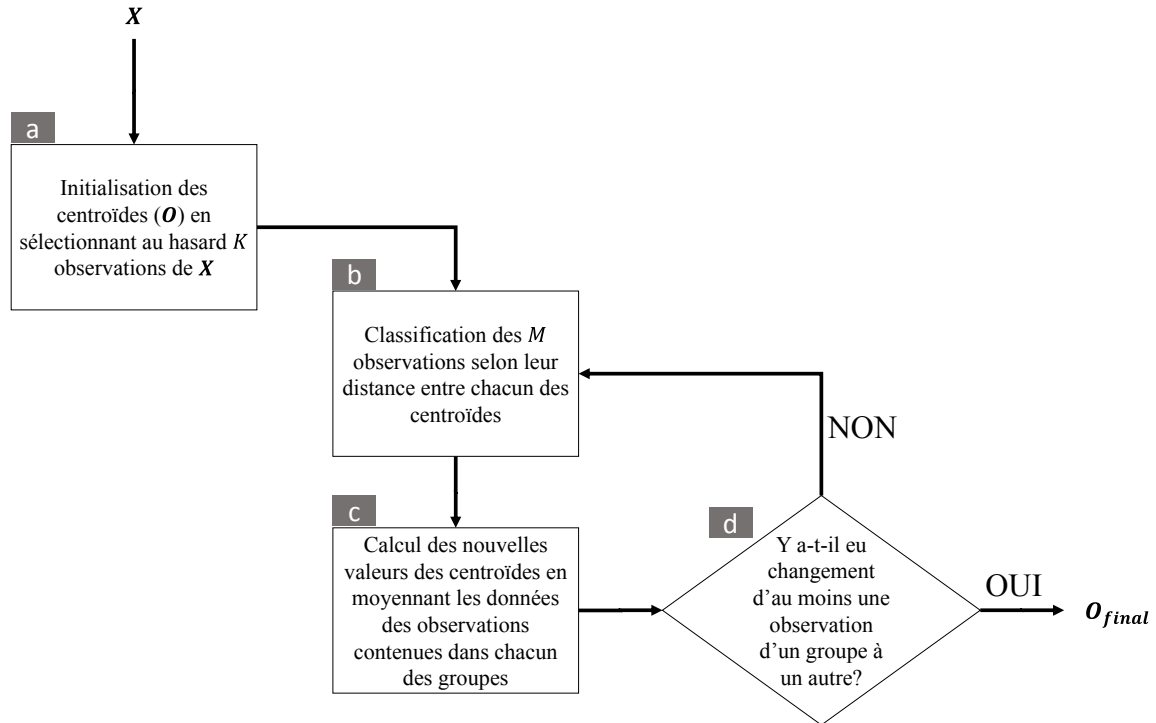


Figure 3.11 Algorithme des k-moyennes afin de segmenter les données \mathbf{X} dans K groupes régis par les centroïdes \mathbf{O}

La matrice des centroïdes finale, \mathbf{O}_{final} , peut être utilisée comme matrice spectrale (\mathbf{S}) initiale.

3.1.6 Comparaison et sélection des méthodes d'initialisation

Les sous-sections 3.1.4 et 3.1.5 ont présenté deux différents algorithmes pouvant être utilisés pour l'initialisation. *SIMPLISMA* a comme principal avantage de n'utiliser aucune fonction aléatoire, contrairement à l'algorithme des *k-moyennes* qui utilise une fonction aléatoire pour initialiser les centroïdes. Nonobstant ce point, il est intéressant de constater l'allure générale des spectres résultant de ces deux méthodes, surtout quand ceux-ci sont créés avec des données extrêmement bruitées.

Dans le cas de *SIMPLISMA*, l'équation (3.9) montre comment calculer les spectres finaux. Si \mathbf{M}_{pures} contient des données très bruitées, par cette opération, il y a de fortes probabilités

qu'une valeur spectrale initiale calculée soit beaucoup trop éloignée des valeurs recherchées.

La Figure 3.12 montre une section des spectres obtenus par une initialisation pour deux composants par *SIMPLISMA* à partir d'un ensemble de données acquises par analyse dispersive en énergie. Ces données ont un ratio signal sur bruit extrêmement faible.

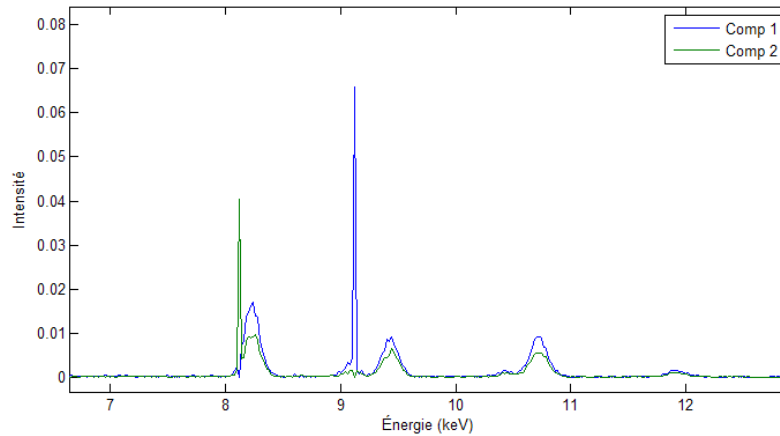


Figure 3.12 Exemple d'initialisation problématique avec *SIMPLISMA*

À vue d'œil, sur la Figure 3.12, on remarque quelques pics, vers 8,1 keV pour le composant 2 et 9,1 keV pour le composant 1, qui semblent être totalement discontinus par rapport aux autres valeurs spectrales, ce qui doit être évité. Ceci est courant, car *SIMPLISMA* n'utilise que très peu d'information afin de calculer les spectres. Malgré que ce ne soient que des spectres d'initialisation, ces pics, qui sont totalement non caractéristiques du matériau analysé, peuvent tout de même saboter l'algorithme de *MCR-ALS* ou significativement miner ses probabilités de convergence.

L'utilisation de l'algorithme des k-moyennes, se comportant différemment comparativement à *SIMPLISMA*, générera des spectres initiaux tout à fait différents (*Figure 3.12 et Figure 3.13*). Or, dans le cas d'un ensemble de données très bruitées, avec beaucoup d'observations et beaucoup de variables, il est fort probable que l'algorithme des k-moyennes ne converge jamais. Un certain nombre d'observations transigeront continuellement d'un groupe à l'autre. Malgré ce fait, les résultats obtenus, après un

nombre d'itérations donné, demeurent très intéressants, tel qu'il le sera démontré et expliqué dans les prochaines lignes.

La Figure 3.13 démontre l'initialisation obtenue avec l'algorithme des k-moyennes pour deux composants à partir des mêmes données que celles utilisées pour générer la Figure 3.12.

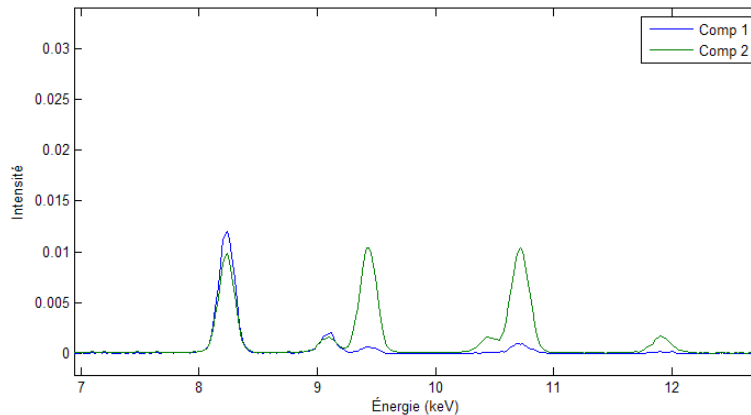


Figure 3.13 Exemple d'initialisation par l'algorithme des k-moyennes

La Figure 3.13 démontre des spectres qui n'ont aucun pic anormal, contrairement à la Figure 3.12. Il faut rappeler que les valeurs de centroïdes, donc des spectres, dans l'algorithme des k-moyennes sont calculées par moyenne entre les données. Les anomalies de la Figure 3.12 ne peuvent donc pas être calculées. Aussi, par le fait que les valeurs des centroïdes soient calculées par moyenne, aucune valeur négative ne peut être calculée. On respecte donc automatiquement la contrainte de non négativité.

La Figure 3.14 compare les spectres obtenus par les deux différents algorithmes à partir de données possédant un relativement haut ratio signal sur bruit (majorité des données corrompues par du bruit de Poisson étant au-dessus de 50).

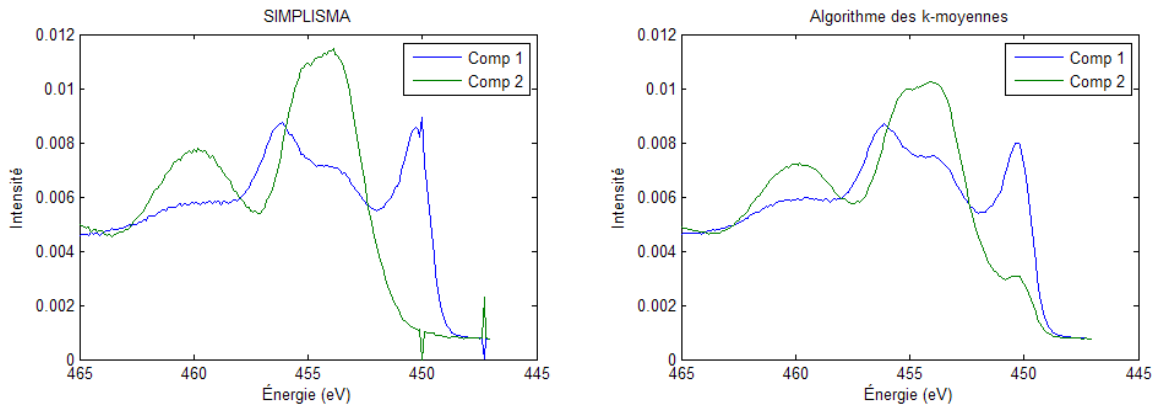


Figure 3.14 Comparaison des spectres obtenus entre deux méthodes d'initialisation

La Figure 3.14 montre encore là deux résultats plutôt différents. Encore là, l'utilisation de *SIMPLISMA* amène la création de quelques pics anormaux. Or, en analysant les résultats de l'algorithme des k-moyennes, et en les comparant à ceux de *SIMPLISMA*, on constate que les spectres correspondent plutôt à un mélange de spectres. En effet, l'algorithme des k-moyennes permet une moins grande séparation des spectres (*Figure 3.15*).

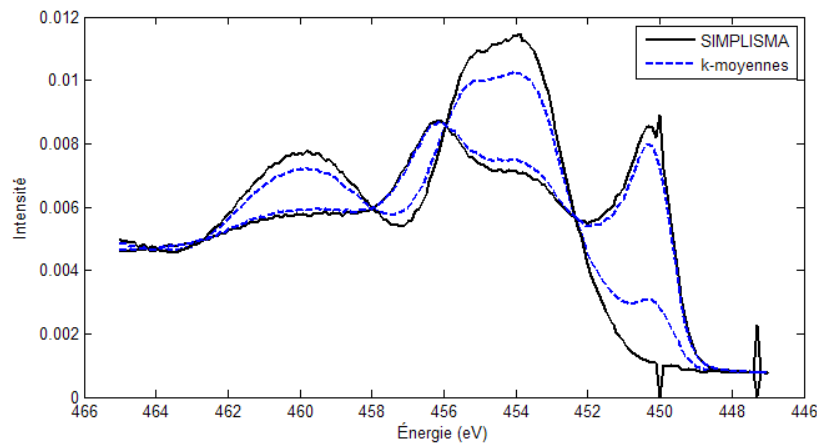


Figure 3.15 Comparaison des spectres obtenus entre deux méthodes d'initialisation

La Figure 3.15 démontre bien l'avantage de *SIMPLISMA* par rapport à l'algorithme des k-moyennes. Les spectres obtenus par une initialisation par *SIMPLISMA* s'approcheront de beaucoup des spectres des composants purs, alors que les spectres obtenus par l'algorithme des k-moyennes représenteront davantage des mélanges de spectres des composants purs. Une initialisation par *SIMPLISMA* permet donc généralement de s'approcher des résultats

finaux et diminue donc le nombre d'itérations nécessaires des algorithmes subséquents lorsque les discontinuités spectrales ne sont pas trop importantes.

Ainsi, dans cette sous-section, il a été question des avantages et désavantages de chacune des méthodes. Le Tableau 3.2 résume donc les avantages et désavantages des deux méthodes d'initialisation étudiées :

Tableau 3.2 Avantages et désavantages des différentes méthodes d'initialisation des spectres

| Nom de la méthode | Avantages | Désavantages |
|---------------------------|---|--|
| <i>SIMPLISMA</i> | <ul style="list-style-type: none"> - Les spectres sont très proches de la solution finale | <ul style="list-style-type: none"> - Création d'anomalies non caractéristiques - Valeurs spectrales pouvant être négatives |
| Algorithme des k-moyennes | <ul style="list-style-type: none"> - Aucun pic non caractéristique - Aucune valeur négative | <ul style="list-style-type: none"> - Les spectres représentent davantage des mélanges de composants purs |

Dans le reste de ce mémoire, la technique des k-moyennes sera utilisée, *SIMPLISMA* implique des désavantages qui ont rendu à plusieurs reprises inefficaces les algorithmes.

3.2 Optimisation

L'objectif de l'optimisation est de trouver les valeurs de variables x qui permettent de minimiser ou de maximiser une fonction f dont les variables peuvent être soumises à certaines contraintes ψ . L'optimisation comprend donc un ensemble de techniques et d'algorithmes qui permettent d'atteindre l'objectif.

3.2.1 Méthode de Newton

La méthode de Newton est une technique d'optimisation basée ici sur le développement d'une série de Taylor de second ordre. On a donc une fonction $f(\mathbf{x})$, où \mathbf{x} est un vecteur

contenant les variables dont les valeurs doivent être trouvées par optimisation. \mathbf{x}_z sont les valeurs actuelles de \mathbf{x} à une itération donnée et \mathbf{x}_{z+1} représente les valeurs estimées à l'itération suivante à partir des valeurs de \mathbf{x}_z . « $\mathbf{x}_{z+1} - \mathbf{x}_z$ » se définit comme le vecteur $\Delta\mathbf{x}$. Le développement de Taylor est le suivant :

$$f(\mathbf{x}_{z+1}) \simeq f(\mathbf{x}_z) + \mathbf{j}(\mathbf{x}_z)^T \times (\mathbf{x}_{z+1} - \mathbf{x}_z) + \frac{1}{2}(\mathbf{x}_{z+1} - \mathbf{x}_z)^T \times \mathbf{H}(\mathbf{x}_z) \times (\mathbf{x}_{z+1} - \mathbf{x}_z) \quad (3.12)$$

Dans l'équation (3.12), $\mathbf{j}(\mathbf{x}_z)$ correspond au vecteur gradient de f à \mathbf{x}_z et $\mathbf{H}(\mathbf{x}_z)$ à la matrice hessienne (dérivées secondes) de f à \mathbf{x}_z . L'approximation de $f(\mathbf{x}_{z+1})$ est ensuite retransformée afin d'isoler les termes (\mathbf{x}_{z+1}) d'ordres 0, 1 et 2.

$$f(\mathbf{x}_{z+1}) \simeq f(\mathbf{x}_z) + (\mathbf{j}(\mathbf{x}_z) - \mathbf{H}(\mathbf{x}_z)\mathbf{x}_z)^T \mathbf{x}_{z+1} + \frac{1}{2}\mathbf{x}_{z+1}^T \mathbf{H}(\mathbf{x}_z)\mathbf{x}_{z+1} \quad (3.13)$$

Par définition, le but d'une optimisation est de minimiser f . Afin de trouver les valeurs de \mathbf{x}_{z+1} permettant d'atteindre ce but, on calcule la dérivée première de l'équation (3.13) par rapport à \mathbf{x}_{z+1} :

$$f(\mathbf{x}_{z+1})' \simeq (\mathbf{j}(\mathbf{x}_z) - \mathbf{H}(\mathbf{x}_z)\mathbf{x}_z)^T + \mathbf{H}(\mathbf{x}_z)\mathbf{x}_{z+1} \quad (3.14)$$

Afin d'obtenir les valeurs de \mathbf{x}_{z+1} qui minimisent la fonction approximée, on résout l'équation (3.14) en faisant égalier $f(\mathbf{x}_{z+1})'$ à 0.

$$\mathbf{x}_{z+1} = -\mathbf{H}(\mathbf{x}_z)^{-1} + (\mathbf{j}(\mathbf{x}_z) - \mathbf{H}(\mathbf{x}_z)\mathbf{x}_z) = \mathbf{x}_z - \mathbf{H}(\mathbf{x}_z)^{-1}\mathbf{j}(\mathbf{x}_z) \quad (3.15)$$

La fonction (3.15) permet de calculer \mathbf{x}_{z+1} minimisant l'approximation $f(\mathbf{x}_{z+1})$. Elle peut être aussi être réécrite de ces façons :

$$-H(\mathbf{x}_z) \times \Delta \mathbf{x} = j(\mathbf{x}_z) \quad (3.16)$$

$$\Delta \mathbf{x} = -H(\mathbf{x}_z)^{-1} j(\mathbf{x}_z) \quad (3.17)$$

Cependant, on parle bien d'approximation. Cela ne garantit pas que \mathbf{x}_{z+1} optimise bel et bien la fonction f non approximée. Ainsi, par un simple calcul itératif, on refait l'équation (3.15) jusqu'à stabilisation de \mathbf{x}_{z+1} . De plus, c'est en calculant la seconde dérivée de l'approximation de f qu'on s'assure si on a bel et bien affaire à un minimum, et non à un maximum ou un point de selle. [43]–[45]

3.2.2 Optimisation sous contrainte

Les contraintes d'égalité se définissent comme suit :

$$\mathbf{G} \times \mathbf{x} = \mathbf{c} \quad (3.18)$$

Par exemple, dans un système à trois variables, on impose que la somme de x_1 et de x_2 soit égale à 1. Dans ce cas-ci, \mathbf{c} serait égale à 1 et \mathbf{G} se définirait comme suit :

$$\mathbf{G} = [1 \quad 1 \quad 0] \quad (3.19)$$

Si on ajoute une deuxième contrainte stipulant que la somme x_2 et de x_3 doit équaler 0, \mathbf{c} et \mathbf{G} se définissent alors comme suit :

$$\mathbf{c} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad (3.20)$$

$$\mathbf{G} = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} \quad (3.21)$$

Chaque ligne de \mathbf{G} indique quelles variables doivent être pris en compte dans la contrainte et \mathbf{c} indique l'égalité désirée. Ainsi, \mathbf{G} est une matrice de taille $R \times K$, où R est le nombre

de contraintes imposées et K le nombre de variables \mathbf{x} , alors que \mathbf{c} est un vecteur de longueur R .

Pour effectuer une optimisation sous contrainte, il faut introduire les multiplicateurs de Lagrange ($\boldsymbol{\lambda}$) à l'intérieur de l'équation (3.16). Ces multiplicateurs de Lagrange sont de nouvelles variables, en complémentarité de \mathbf{x} , qui permettent de prendre en compte des contraintes et sont intégrés de cette façon :

$$\begin{bmatrix} \mathbf{H}(\mathbf{x}_{z+1}) & \mathbf{G}^T \\ \mathbf{G} & 0 \end{bmatrix} \begin{bmatrix} \Delta \mathbf{x} \\ \boldsymbol{\lambda} \end{bmatrix} = \begin{bmatrix} -\mathbf{j}(\mathbf{x}_{z+1}) \\ \mathbf{c} \end{bmatrix} \quad (3.22)$$

À chaque itération, le vecteur $\begin{bmatrix} \Delta \mathbf{x} \\ \boldsymbol{\lambda} \end{bmatrix}$ est résolu jusqu'à sa stabilisation. [43]–[45]

3.2.3 Algorithme *Active Set*

Les sous-sections précédentes ont démontré comment prendre en compte des contraintes d'égalité. Or, il peut arriver qu'on veuille imposer des contraintes d'inégalité :

$$\mathbf{G}\mathbf{x} \leq \mathbf{c} \quad (3.23)$$

Certaines méthodes permettent de tenir en compte des inégalités telles que définies par l'équation (3.23). L'une d'elles se nomme l'algorithme *Active Set*. Se basant sur l'équation (3.22), l'algorithme tire son nom du fait que les contraintes imposées sont successivement activées et désactivées. [43]–[45]

Dans l'algorithme, à la base, aucune contrainte n'est considérée. Cependant, si les valeurs de \mathbf{x} touchent la limite d'une contrainte, cette contrainte d'inégalité est activée, est transformée en contrainte d'égalité et sera considérée dans l'équation (3.22).

Si, à une itération donnée, la variation des valeurs de \mathbf{x} , définie par $\Delta \mathbf{x}$, fait dépasser une contrainte d'inégalité, le déplacement $\Delta \mathbf{x}$ est alors atténué afin que les valeurs de \mathbf{x} restent à la limite de la contrainte d'inégalité. Encore là, cette contrainte est par la suite considérée dans l'équation (3.22). [43]–[45]

Le signe des multiplicateurs lagrangiens, λ , calculés au fil des itérations, a une signification. Si le multiplicateur lagrangien est positif, cela signifie que la contrainte d'égalité a été respectée. Au contraire, s'il est négatif, cela signifie qu'il n'y a pas respect de la contrainte d'égalité. À ce moment-là, dans l'algorithme *Active Set*, cela signifie que l'imposition de la contrainte n'est plus nécessaire et qu'elle doit être désactivée. [43]–[45]

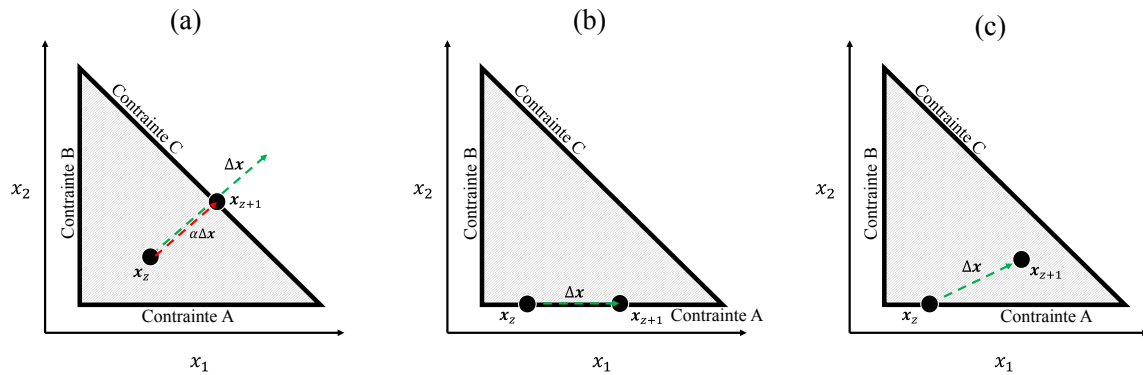


Figure 3.16 Cas typiques dans l'algorithme *Active Set*

La Figure 3.16 illustre un exemple de contraintes d'inégalité dans un système à deux variables. L'intérieur du triangle correspond au respect des contraintes d'inégalité.

Dans le cas (a), le déplacement calculé Δx amène x à l'extérieur des limites d'inégalité. Le déplacement doit donc être atténué par un facteur α et la contrainte C doit être activée afin de tronquer le déplacement dépassant la limite imposée.

Dans le cas (b), la contrainte A est activée et le déplacement calculé reste accroché à cette contrainte. Cela signifie donc que le multiplicateur lagrangien associé à cette contrainte est positif. La contrainte A reste alors activée.

Dans le cas (c), la contrainte A est activée et le déplacement calculé se décroche de la contrainte. Cela signifie donc que le multiplicateur lagrangien associé à la contrainte est négatif. La contrainte A est alors désactivée.

L'algorithme *Active Set* offre plusieurs avantages. Premièrement, le fait que des contraintes soient désactivées permet de simplifier les calculs matriciels et ainsi de réduire le temps de calcul. Deuxièmement, tel qu'illustré dans le cas (a) de la Figure 3.16, il est impossible de calculer, à une itération donnée, des valeurs de x qui soient à l'extérieur des contraintes

d'inégalité. Certaines fonctions dont on veut minimiser ou maximiser peuvent être problématiques à calculer lorsque les contraintes d'inégalité ne sont pas respectées. L'utilisation de l'algorithme *Active Set* permet donc d'éviter d'obtenir, à une itération donnée, des valeurs de \mathbf{x} problématiques. [43]–[45]

3.3 Traitement de signal

3.3.1 Fonction de vraisemblance

La fonction de vraisemblance, l , ou *likelihood* en anglais, est une fonction qui permet d'exprimer d'associer une série de valeurs, indépendantes l'une de l'autre, à une distribution statistique, dont les paramètres régissant cette distribution sont connus. [46] De façon générale, cette fonction s'exprime de cette façon :

$$l(x_1, \dots, x_N | \theta_1, \dots, \theta_N) = \prod_{n=1}^N p(x_n; \theta_n) \quad (3.24)$$

Dans l'équation (3.24), x_n correspond à la valeur de la variable n de la série de N variables, alors que θ_n contient le ou les paramètres régissant la distribution statistique à la variable n . Enfin, $p(x_n; \theta_n)$ correspond à la fonction de densité calculée pour la valeur x_n selon les paramètres θ_n .

La fonction de densité se calcule différemment selon le type de distribution statistique en jeu. Dans ce mémoire, la distribution de Poisson sera largement utilisée. La fonction de densité de cette distribution est la suivante :

$$p_{\text{Poisson}}(x; \mu) = \frac{\mu^x e^{-\mu}}{x!} \quad (3.25)$$

Dans l'équation (3.25), μ représente la valeur moyenne d'une variable et est ainsi le seul paramètre régissant la distribution de Poisson. Une deuxième distribution largement vue, mais qui ne sera pas utilisée dans ce mémoire est la distribution gaussienne. La fonction de densité de cette distribution est la suivante :

$$p_{Gaussienne}(x; \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x - \mu)^2}{2\sigma^2}\right) \quad (3.26)$$

Dans l'équation (3.26), μ correspond à la valeur moyenne d'une variable et σ correspond à l'écart-type de la variable.

L'équation (3.24) démontre que les densités calculées, pour chacune des variables, sont multipliées entre-elles. Cette multiplication est effectuée en considérant que les variables sont indépendantes l'une de l'autre. En spectroscopie, les acquisitions pour un niveau d'énergie donné n'ont aucune influence sur l'acquisition à un autre niveau d'énergie. On peut donc considérer les variables (niveaux d'énergie) comme étant indépendantes.

Or, plus il y a présence de variables dans un système, plus la valeur de l calculée sera petite. On peut donc atteindre les limites de calcul de l'ordinateur ou manquer de précision sur le résultat final. C'est pourquoi on utilise souvent, dans plusieurs algorithmes de traitement de signaux, le logvraisemblance, $\log(L)$, plutôt que la vraisemblance. [46] Du fait même, cette transformation permet de transformer la multiplication de l'équation (3.24) en addition :

$$\log(l(x_1, \dots, x_N | \theta_1, \dots, \theta_N)) = \sum_{n=1}^N \log(p(x_n; \theta_n)) \quad (3.27)$$

3.3.2 Algorithme espérance-maximisation

L'algorithme d'espérance-maximisation permet de prendre en compte une distribution aléatoire des données. Par exemple, si on pose l'hypothèse que les valeurs d'un même groupe, pour chaque dimension, sont réparties selon une distribution gaussienne,

l'algorithme d'espérance-maximisation permettra d'ajuster les valeurs des centroïdes dans le but que les valeurs faisant partie d'un même groupe suivent une distribution gaussienne dans chaque dimension.

Tel que son nom l'indique, l'algorithme repose sur deux étapes différentes, soit l'étape d'espérance et celle de maximisation. À l'étape d'espérance, on assigne chaque observation à un groupe ou à plusieurs groupes par fractionnement. À l'étape de maximisation, pour chaque groupe et pour chaque dimension d'un groupe, on calcule les paramètres régissant la distribution selon les observations faisant partie du groupe. Par exemple, avec l'hypothèse d'une distribution gaussienne, chaque dimension de chaque groupe aura une valeur d'espérance mathématique associée et une valeur d'écart-type, ces deux paramètres étant nécessaires pour régir une fonction gaussienne. [46]–[48]

À chaque itération de l'algorithme, on calcule le logvraisemblance. Cette valeur permet d'exprimer à quel point les paramètres de chaque groupe permettent de bien représenter les observations qui sont associées à l'un ou l'autre des groupes. Le calcul prend en compte la probabilité, selon la distribution posée en hypothèse, pour chaque observation de se retrouver dans le groupe auquel elle est associée. Plus il y a d'itérations effectuées, plus la valeur du logvraisemblance augmente jusqu'à se stabiliser après plusieurs itérations. C'est cette valeur qui est utilisée comme critère de convergence dans le processus itératif. [46]–[48]

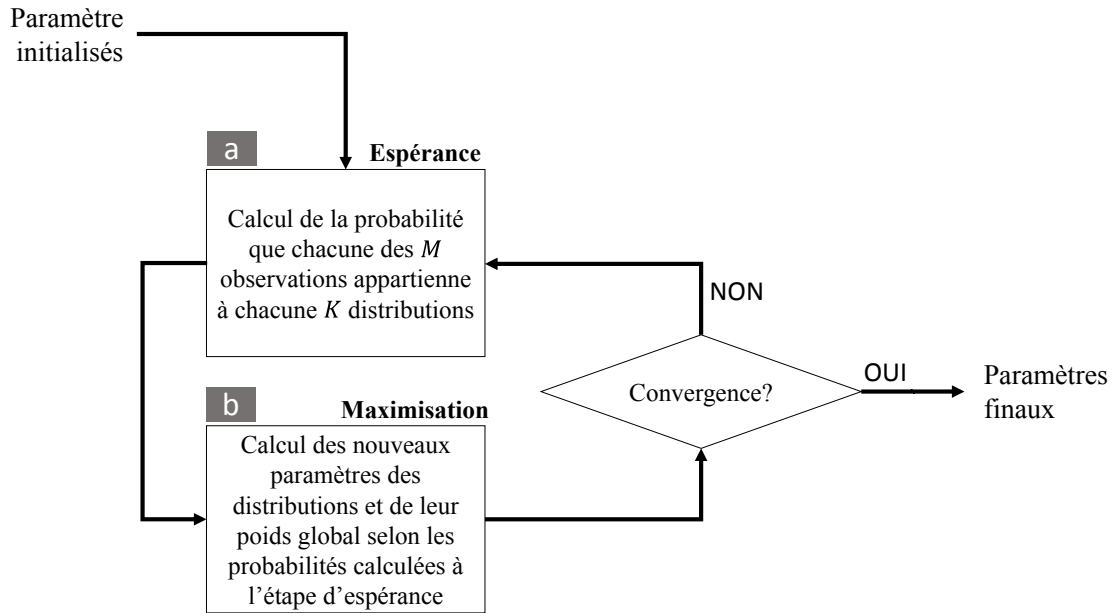


Figure 3.17 Algorithme Espérance-Maximisation

L'algorithme d'espérance-maximisation vise à maximiser cette équation :

$$\log(L(\mathbf{X}, \mathbf{W}, \boldsymbol{\theta})) = \sum_{m=1}^M \sum_{k=1}^K \mathbf{W}_{mk} \log(\boldsymbol{\pi}_k \times l(\mathbf{X}_m; \boldsymbol{\theta}_k)) \quad (3.28)$$

Dans l'équation (3.28), M correspond au nombre total d'observations et K au nombre de distributions à trouver. \mathbf{X}_m consiste aux données de l'observation m pour toutes les variables. \mathbf{W}_{mk} correspond à la probabilité que l'observation m appartienne à la distribution k . $\boldsymbol{\pi}_k$ correspond au poids global que la distribution k a sur l'ensemble des observations. La fonction l correspond à la fonction de vraisemblance (*sous-section 3.3.1*) et diffère selon le type de distribution qu'on cherche à trouver et paramétrer. $\boldsymbol{\theta}_k$ contient donc le ou les paramètres, dépendant du type de distribution, pour chaque distribution à chaque variable.

Dans l'étape d'espérance, on calcula la probabilité que chaque observation appartienne à l'une ou l'autre des distributions.

$$\mathbf{W}_{mk} = \frac{l(\mathbf{X}_m; \boldsymbol{\theta}_k) \boldsymbol{\pi}_k}{\sum_{l=1}^K l(\mathbf{X}_m; \boldsymbol{\theta}_l) \boldsymbol{\pi}_k} \quad (3.29)$$

Ainsi, la somme des valeurs de \mathbf{W}_m , en prenant en compte les K distributions, est de 1. Plus la distribution est importante globalement pour toutes les observations ($\boldsymbol{\pi}_k$), plus celle-ci aura d'impact sur la probabilité que chacune des observations soit associée à cette distribution.

Une fois ces probabilités calculées pour chacune des observations, on passe à l'étape de maximisation. Dans cette étape, on calcule premièrement le poids global que chaque observation a parmi les observations.

$$\boldsymbol{\pi}_k = \frac{\sum_{m=1}^M \mathbf{W}_{mk}}{M} \quad (3.30)$$

La somme de toutes les valeurs de $\boldsymbol{\pi}_k$ est aussi de 1. Ensuite, on calcule les paramètres des distributions selon les données de chacune des observations et du poids qu'une distribution k a sur cette observation. Dans des distributions de Poisson, le seul paramètre est la moyenne/variance et se calcule ainsi pour une variable n :

$$\boldsymbol{\theta}_{kn} = \frac{\sum_{m=1}^M \mathbf{W}_{mk} \mathbf{X}_{mk}}{\sum_{m=1}^M \mathbf{W}_{mk}} \quad (3.31)$$

En fin de convergence, soit à la stabilisation de $\boldsymbol{\theta}$, dans la majorité des cas, la plupart des valeurs contenues dans la matrice \mathbf{W} seront près de 0 ou de 1. Si une observation m n'appartient pas à une distribution k , $l(\mathbf{X}_m; \boldsymbol{\theta}_k)$ sera très petit, surtout avec un grand nombre de variables N , alors que le résultat l sera très largement supérieur si cette observation fait réellement partie de cette distribution. Ainsi, par ces faits, dans l'équation (3.29), une observation faisant partie de la distribution aura une valeur résultant très près de 1, alors que les autres valeurs seront de 0. L'algorithme d'espérance-maximisation n'est pas approprié pour attribuer une observation à un mélange de distributions.

3.4 Résumé du Chapitre 3

Le Chapitre 3 a passé en revue les algorithmes et méthodes qui seront utilisés au cours des prochains chapitres. Au niveau de l'analyse multivariable, différents prétraitements des données peuvent être utilisés. *Optimal Scaling* est recommandé pour des données bruitées par un bruit de Poisson, tandis que la normalisation spectrale permet de contourner les problèmes dus à la variation de sensibilité entre les pixels.

Toujours au niveau des analyses multivariées, *MCR-ALS*, contrairement à *PCA*, permet d'extraire directement des spectres et des contributions reliées à ces spectres. Or, dans *MCR-ALS*, l'initialisation est d'une grande importance. *SIMPLISMA* permet de se rapprocher davantage des spectres finaux calculés par le cœur itératif de *MCR-ALS*. Cependant, cette technique peut causer d'importantes discontinuités dans les spectres, alors que l'algorithme des k-moyennes ne cause pas ce type d'erreur, mais calcule des spectres qui sont davantage éloignés de ceux finaux calculés par *MCR-ALS*.

Au niveau de l'optimisation, l'algorithme *Active Set* est une technique dans la famille de la programmation quadratique. Cet algorithme, assez robuste, permet de maximiser ou de minimiser une fonction tout en imposant des contraintes d'égalité et d'inégalité au niveau des variables. Au fil des itérations, les valeurs des différentes variables respectent toujours les contraintes d'inégalité.

Enfin, au niveau du traitement de signal, la fonction de vraisemblance permet de quantifier à quel point une série de donnée est reliée statistiquement, selon, par exemple, une distribution de Poisson, à une autre série de donnée. L'algorithme d'espérance-maximisation se base sur ce concept de vraisemblance afin de calculer, à partir d'un ensemble de données à plusieurs dimensions, les paramètres de plusieurs distributions statistiques et d'associer chacune des observations à l'une ou l'autre des distributions calculées.

Tableau 3.3 Acronymes utilisés dans le Chapitre 3

| Variable | Correspondance en anglais | Correspondance en français |
|------------------|--|--|
| <i>NIPALS</i> | <i>Non-linear iterative partial least squares</i> | Moindres carrés partiels itératifs non linéaires |
| <i>SVD</i> | <i>Singular values decomposition</i> | Décomposition en valeurs singulières |
| <i>MCR-ALS</i> | <i>Multivariate Curve Resolution – Alternating Least Squares</i> | Résolution multivariée de courbes – Moindres carrés alternatifs |
| <i>PCA</i> | <i>Principal Component Analysis</i> | Analyse en composants principales |
| <i>EKF</i> | <i>Element-Wise k-fold</i> | |
| <i>SIMPLISMA</i> | <i>Simple-to-Use Interactive Self-modeling Mixture Analysis</i> | Analyse de mélange interactive facile à utiliser par auto-modélisation |

Tableau 3.4 Variables utilisées dans le Chapitre 3

| Variable | Définition |
|----------------------|--|
| \mathbf{X} | Matrice des données de dimension $M \times N$ |
| $\tilde{\mathbf{X}}$ | Matrice des données ayant subi un prétraitement |
| \mathbf{X}_{ij} | Valeur acquise à l'observation i et à la variable j |
| M | Nombre d'observations/pixels dans l'ensemble de données |
| M_x, M_y | Les deux dimensions spatiales dans un ensemble de données en trois dimensions |
| N | Nombre de variables dans l'ensemble de données |
| K | Nombre de composants à extraire |
| \mathbf{T} | Matrice des <i>scores</i> de dimension $M \times K$ |
| \mathbf{P} | Matrice des <i>loadings</i> de dimension $K \times N$ |
| $\mathbf{\Sigma}$ | Matrice diagonale des valeurs propres de dimension $K \times K$ |
| \mathbf{E} | Matrice d'erreur de prédiction de dimension $M \times N$ |
| \mathbf{C} | Matrice des contributions de dimension $M \times K$ |
| \mathbf{S} | Matrice spectrale de dimension $N \times K$ |
| \mathbf{S}_{pures} | Matrice spectral obtenue à partir de SIMPLIMA de dimension $N \times K$ |
| \mathbf{M}_{pures} | Matrice contenant les données associées aux variables pures |
| \mathbf{M}_{test} | Matrice permettant de tester l'ajout de données à la matrice \mathbf{M}_{pures} |
| w_n | Déterminant de la matrice résultant de l'ajout des données associées à la variable n à la matrice \mathbf{M} |
| \mathbf{O} | Matrice contenant les valeurs des centroïdes de dimension $K \times N$ |
| \mathbf{D} | Matrice contenant les distances euclidiennes entre les observations et les centroïdes |
| Q_k | Nombre de données incluses dans le groupe k |
| \mathbf{x} | Vecteur contenant les valeurs des variables d'optimisation |

| | |
|---|---|
| \mathbf{x}_z | Vecteur contenant les valeurs « actuelles » |
| \mathbf{x}_{z+1} | Vecteur contenant les valeurs de l'itération suivante |
| $\Delta \mathbf{x}$ | Vecteur contenant la variation de chacune des variables |
| $f(\mathbf{x})$ | Fonction dont on veut optimiser |
| $f'(\mathbf{x})$ | Dérivée première de la fonction dont on veut optimiser |
| $f''(\mathbf{x})$ | Dérivée seconde de la fonction dont on veut optimiser |
| $V(\mathbf{x}, \boldsymbol{\lambda})$ | Fonction à optimiser comprenant aussi les multiplicateurs lagrangiens |
| $F(\mathbf{x})$ | Fonction à minimiser en programmation quadratique |
| $g(\mathbf{x})$ | Fonction qui permet de prendre en compte les contraintes d'égalité |
| \mathbf{G} | Matrice qui indique quelles variables prendre en compte dans les contraintes d'égalité |
| $\boldsymbol{\lambda}$ | Vecteur contenant les multiplicateurs de Lagrange |
| \mathbf{c} | Vecteur contenant les scalaires définissant à quelles valeurs devraient évaluer les fonctions comprises dans $g(\mathbf{x})$ |
| \mathbf{j} | Vecteur contenant les dérivées premières (jacobien) |
| \mathbf{H} | Matrice contenant les dérivées secondes (hessienne) |
| α | Scalaire atténuant $\Delta \mathbf{x}$ |
| l | Fonction de vraisemblance |
| L | Fonction globale de vraisemblance |
| $p(\mathbf{x}; \boldsymbol{\theta})$ | Densité de probabilité de \mathbf{x} considérant les paramètres de distribution inclus dans $\boldsymbol{\theta}$ |
| $p_{Poisson}(\mathbf{x}; \mu)$ | Densité de probabilité de \mathbf{x} considérant le paramètre de distribution μ , considérant une distribution de Poisson |
| $p_{Gaussienne}(\mathbf{x}; \mu, \sigma)$ | Densité de probabilité de \mathbf{x} considérant les paramètres de distribution μ et σ , considérant une distribution Gaussienne |
| $\boldsymbol{\theta}_k$ | Paramètres régissant la distribution k |
| W_{mk} | Proportion attribuée à la distribution k de l'observation m |
| π_k | Proportion globale que prend la distribution k |

4. NOUVELLE MÉTHODE DE TRAITEMENTS DE DONNÉES PROPOSÉE

Ce chapitre contient le cœur de cette maîtrise, soit la méthode développée afin de traiter des données d'imagerie multispectrale. Le développement de cette méthode s'est basé sur les principes physiques liés au détecteur (*Chapitre 2*), ainsi que toutes les notions incluses dans le Chapitre 3, que ce soit en analyse multivariable, en optimisation ou en traitement de signal.

Tel qu'il le sera démontré, la méthode développée au cours de cette maîtrise permet d'utiliser à fond toutes les caractéristiques des données comparativement aux méthodes d'analyse de données d'imagerie multispectrale plus traditionnelles. De plus, cette méthode développée offre à la communauté scientifique un outil qui permet d'extraire facilement et directement des spectres et des contributions d'un ensemble de données d'imagerie multispectrale, ce que les méthodes plus traditionnelles ne sont pas en mesure de faire.

4.1 Données

4.1.1 Caractéristiques des données

Afin d'utiliser les méthodes qui seront présentées, les données qui sont en entrée dans l'algorithme doivent être brutes. Elles ne doivent donc pas avoir subi aucun prétraitement ou compression. Un prétraitement ou une compression crée habituellement des nombres décimaux. Vérifier ce fait est un moyen simple et efficace de diagnostiquer l'état des données.

Certains logiciels de spectroscopie enregistrent les données en prenant en compte une certaine calibration. Il est donc impératif de s'assurer qu'aucune modification des données ne soit effectuée. D'autres instruments modifient les comptes bruts à même le détecteur. Dans ce cas-ci, enregistrer les données brutes s'avère impossible. Le processus de modification des comptes doit donc être connu pour retrouver ces dites données brutes. S'il n'est pas connu, un modèle du comportement du bruit prenant en compte les données modifiées devra être développé. Si les données sont brutes, tel que discuté au Chapitre 2, en posant certaines hypothèses, le bruit peut être approximé comme suivant une loi de Poisson.

Ainsi, en spectroscopie, on s'attend à des données correspondant à des nombres entiers. Ces nombres entiers représentent le nombre brut de photons ou d'électrons comptés.

4.1.2 Prétraitement des données sélectionnée

Une des calibrations les plus souvent appliquées est la correction du gain à chaque pixel, plus connue sous le terme anglais *Flat Field Correction* (*sous-section 2.2.3*). Or, cette correction reste importante à considérer dans les méthodes qui seront présentées.

Ainsi, tout en gardant les valeurs des données brutes (\mathbf{X}_{brut}), on doit appliquer la correction de champ plat aux données et créer \mathbf{X}_{CP} . Si cette calibration n'est pas connue, ou si on la croit biaisée, il est recommandé de tout simplement normaliser chaque spectre et ainsi créer la matrice \mathbf{X}_{norm} , tel que le décrit l'équation (3.4). En d'autres mots, on modifie les données pour que la somme des données pour chaque pixel soit de un.

Les données de chaque pixel i sont donc divisées par un scalaire qu'on définira comme d_i . On a donc le vecteur \mathbf{d} contenant les scalaires divisant les pixels de chaque observation de l'ensemble de données.

Dans la méthode, \mathbf{X}_{norm} sera considéré comme des données prétraitées. Tel qu'il le sera démontré dans les prochains paragraphes, à la fois \mathbf{X}_{brut} et \mathbf{X}_{norm} seront utilisés.

4.2 Méthode développée (*MCR-LLM*)

Il sera expliqué dans les prochaines lignes que la méthode développée au cours de la maîtrise est basée sur l'algorithme *MCR-ALS*, mais qui comporte une modification majeure basée sur la maximisation du logvraisemblance (*loglikelihood* en anglais). De ce fait, on nommera l'algorithme *Multivariate Curve Resolution – Loglikelihood Maximization* (*MCR-LLM*).

4.2.1 Initialisation

Les sous-sections 3.1.4 à 3.1.6 ont présenté deux méthodes d'initialisation et les avantages et inconvénients respectifs de celles-ci. Selon les nombreux tests qui ont été effectués pour arriver à la version finale de *MCR-LLM*, plusieurs échecs de convergence ont été dus à l'utilisation d'une initialisation par *SIMPLISMA*, dont les principaux désavantages ont été mis en relief dans la sous-section 3.1.6.

À l'opposé, toujours par les raisons mentionnés dans la sous-section 3.1.6, l'algorithme des k-moyennes augmente grandement le taux de succès d'une analyse par la méthode qui sera expliquée dans les prochains paragraphes. On parle de succès lorsque l'algorithme réussit à extraire des spectres physiquement cohérents et des cartographies de contribution tout aussi cohérentes à convergence. S'il y a présence de zones spatiales de plus petite surface, une initialisation par l'algorithme des k-moyennes augmentera grandement la possibilité d'isoler ces zones dans des composants uniques au fil des itérations.

4.2.2 Cœur de la méthode

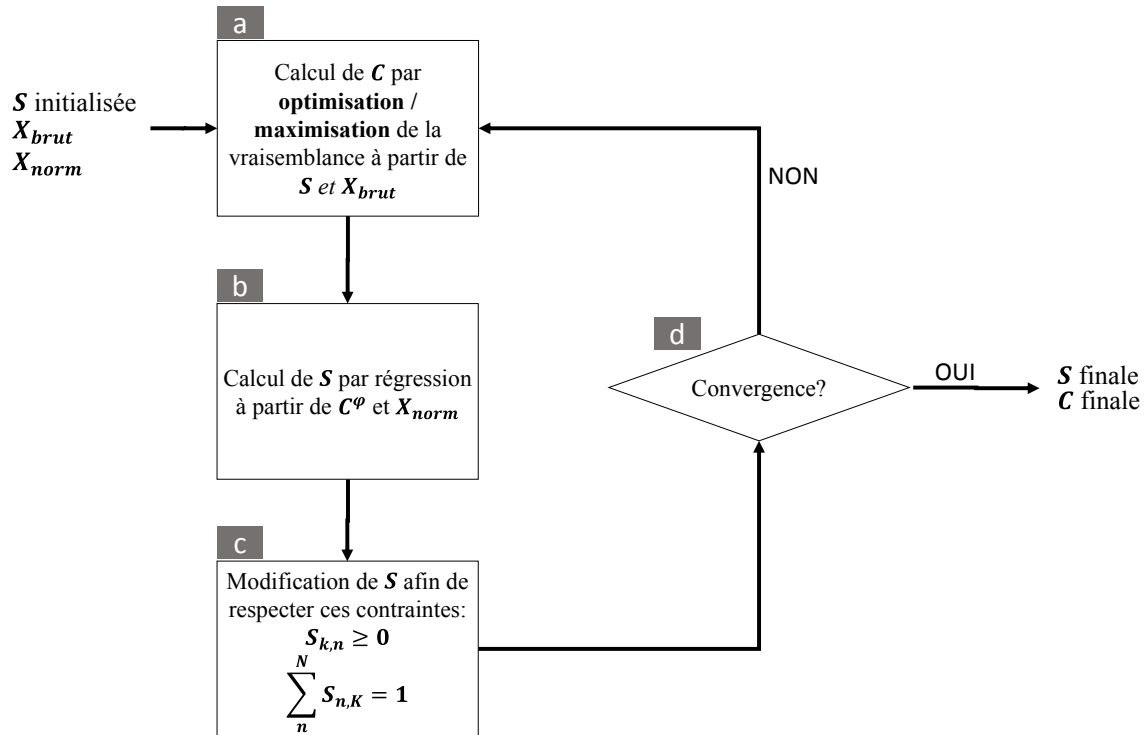


Figure 4.1 Cœur de la méthode proposée

La Figure 4.1 montre le cœur du processus itératif permettant d’analyser les données, d’extraire de réelles contributions et des spectres, tout en considérant le comportement du bruit. La convergence peut être basée sur la différence des valeurs de S ou de C entre l’itération présente et l’itération précédente. Cette différence doit être en-deçà d’un seuil paramétré.

4.2.2.1 Calcul des contributions

L’essentiel du processus itératif se déroule dans la *boîte a*, dans lequel il est écrit que les contributions sont calculées en maximisant la vraisemblance. Cette maximisation est effectuée par optimisation, spécifiquement avec l’algorithme *Active Set* (sous-section 2.2.3). Il est à rappeler qu’originellement, dans l’algorithme *MCR-ALS*, les contributions sont calculées par régression multilinéaire et que, par la suite, les résultats sont modifiés afin de respecter les contraintes.

Ainsi, pour chaque pixel, on cherche les contributions qui permettent de maximiser la vraisemblance. Les contributions permettent de calculer le spectre résultant de ce pixel :

$$\mathbf{W}_{i,n} = \sum_{k=1}^K \mathbf{C}_{i,k} \times \mathbf{S}_{k,n} \times \mathbf{d}_i \quad (4.1)$$

Dans l'équation (4.1), $\mathbf{W}_{i,n}$ correspond à la valeur spectrale prédite pour le pixel i et pour la variable n . K correspond au nombre total de composants à extraire, $\mathbf{C}_{i,k}$ correspond à la contribution pour le pixel i et le composant k . $\mathbf{S}_{k,n}$ correspond à la valeur spectrale pour le composant k et la variable n . Enfin, \mathbf{d}_i correspond au scalaire par lequel toutes les données du pixel i ont été divisées.

Ainsi, dans la *boîte a* de la Figure 4.1, on cherche à trouver les valeurs de \mathbf{C}_i afin de prédire le spectre \mathbf{W}_i qui permet de maximiser la vraisemblance selon les données acquises $\mathbf{X}_{brut,i}$.

Il est à rappeler que la formule pour calculer la vraisemblance diffère selon le comportement du bruit (*sous-section 3.3.1*). Il est donc primordial de connaître le type de distribution qui régit le comportement du bruit et, s'il y a lieu, de connaître les valeurs des paramètres qui régissent cette distribution. Tel que discuté aux sections 2.1 et 2.2, dû à la nature quantique du rayonnement analysé, le bruit suit une loi de Poisson dans la plupart des instruments. La fonction de vraisemblance, l , est donc décrite par l'équation (4.2).

$$l_i(\mathbf{X}_i; \mathbf{W}_i) = \prod_{n=1}^N \frac{\mathbf{W}_{i,n}^{\mathbf{X}_{i,n}} e^{-\mathbf{W}_{i,n}}}{\mathbf{X}_{i,n}!} \quad (4.2)$$

Dans l'équation (4.2), N correspond au nombre de variables dans l'ensemble de données, \mathbf{X}_i contient les données d'une observation i et \mathbf{W}_i , qui est à la fois un spectre, contient les moyennes/variances qui régissent les distributions de Poisson pour chaque variable n .

Les valeurs de \mathbf{C}_i sont soumises à certaines contraintes dans l'optimisation. Premièrement, afin de garder une signification physique, les valeurs de \mathbf{C}_i ne doivent pas être négatives et, deuxièmement, la somme des valeurs de \mathbf{C}_i doit être de un.

La Figure 4.2 illustre à la fois comment l'algorithme *Active Set* trouve itérativement le maximum de la fonction de vraisemblance pour un système à trois composants au fil des itérations, illustre les contraintes qui doivent être respectées et démontre l'allure d'une fonction de vraisemblance.

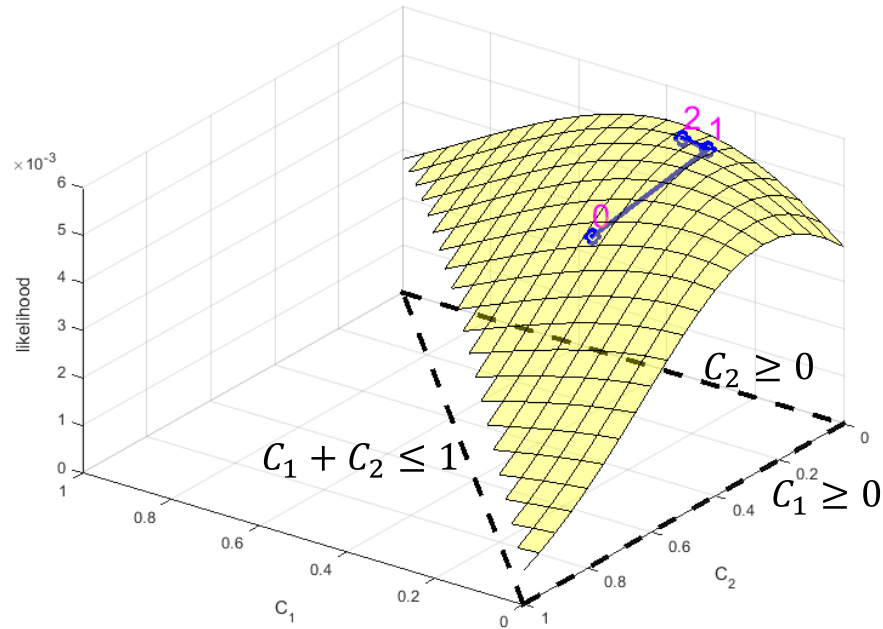


Figure 4.2 Exemple d'optimisation de la fonction de vraisemblance

À noter que dans la Figure 4.2, comme la contribution reliée au 3^e composant est dépendante des deux autres valeurs, on peut définir C_3 de cette façon :

$$C_3 = 1 - C_1 - C_2 \quad (4.3)$$

Comme C_3 n'est plus considérée comme une variable indépendante, la contrainte stipulant que la somme de toutes les contributions doit être de un est donc remplacée par cette contrainte :

$$C_1 + C_2 \leq 1 \quad (4.4)$$

4.2.2.2 Calcul des spectres

Une fois les contributions calculées, le calcul de la matrice spectrale (*boîte b*) se fait toujours par régression multilinéaire à partir de la matrice de contributions, \mathbf{C} , et de la matrice des données normalisées, \mathbf{X}_{norm} .

À travers le développement de *MCR-LLM*, il a été diagnostiqué que l'opération de régression multilinéaire dans le calcul de la matrice spectrale peut causer certains inconvénients, surtout quand les contributions de la plupart des pixels démontrent des mélanges de composants. Le principal inconvénient rencontré dans cette situation est que les spectres résultant s'éloignent des spectres attendus, ce qui fait donc converger l'algorithme vers des résultats fortement erronés. On parle d'ambiguïté rotationnelle au niveau des spectres.

Pour remédier à cet inconvénient, une modification des valeurs de \mathbf{C} a été introduite avant le calcul de régression multilinéaire. Ainsi, toutes les valeurs de \mathbf{C} sont modifiées par un exposant défini par φ (c^φ). (*Figure 4.1 - boîte c*) Généralement paramétré à une valeur entre 1,5 et 2, cet exposant permet d'augmenter le poids des pixels qui correspondent davantage à un composant pur qu'à un mélange de composants. En d'autres mots, les pixels qui contiennent des valeurs de contributions de 1 ou s'approchant de 1 auront un plus grand impact dans le calcul de la matrice spectrale.

On nommera cette matrice des valeurs modifiées \mathbf{C}^* . La régression multilinéaire s'effectue tout à fait normalement à partir de cette nouvelle matrice, soit par l'équation (4.5).

$$\mathbf{S} = (\mathbf{C}^{*T} \mathbf{C}^*)^{-1} \mathbf{C}^{*T} \mathbf{X} \quad (4.5)$$

Une valeur de φ entre 1,5 et 2 permet un juste équilibre. Cette valeur n'est pas trop élevée pour empêcher l'algorithme de converger vers la bonne solution, alors qu'elle permet en même temps d'avantager l'information contenue dans les pixels qui possèdent des contributions s'approchant de 1. Les sous-sections 4.3.2 et 4.3.3 iront davantage en profondeur mathématiquement sur l'utilisation de cette modification et des inconvénients que la modification engendre. Les résultats (*Chapitres 5 et 6*) démontreront l'avantage d'utiliser une valeur de φ entre 1,5 et 2.

La modification des valeurs de C par un exposant φ (Figure 4.1 - boîte c) aide l'algorithme à favoriser la présence de composants purs, soit à réduire l'ambiguïté rotationnelle, tout en considérant, au fil des itérations, les pixels qui contiennent des mélanges de composants.

4.3 Discussions

4.3.1 Maximisation de la fonction de vraisemblance versus minimisation de l'erreur sous contraintes

On peut se demander pourquoi ne pas tout simplement effectuer une minimisation d'erreur au carré sous contraintes pour calculer les contributions.

D'une part, comme la fonction de vraisemblance est basée sur des termes qui se multiplient entre eux, on parle de fonction strictement concave. Cela signifie qu'il n'y a qu'une seule coordonnée à laquelle la fonction est à son maximum et garantit l'unicité de la solution.

D'autre part, l'utilisation d'une minimisation d'erreur absolue peut amener certains problèmes. Par exemple, un pixel a comme données acquises [1 ; 1], le spectre du premier composant a comme valeurs [0,7 ; 1,1] et celui du deuxième composant a comme valeurs [0,8 ; 1,2]. Le Tableau 4.1 montre l'erreur absolue résultante selon différentes combinaisons du composant 1 et du composant 2.

Tableau 4.1 Exemples d'erreurs absolues calculées

| | | | |
|-------------------|------|------|------|
| C_1 | 20% | 50% | 70% |
| C_2 | 80% | 50% | 30% |
| S_1 | 0,78 | 0,75 | 0,73 |
| S_2 | 1,18 | 1,15 | 1,13 |
| $ 1-S_1 + 1-S_2 $ | 0,4 | 0,4 | 0,4 |

Or, qu'on associe le pixel totalement au premier composant, au deuxième composant, ou à n'importe quel mélange de ces deux composants, l'erreur absolue aura dans tous les cas

une valeur de 0,4 (*Tableau 4.1*), ce qui rend totalement impossible le calcul de contributions dans ce cas. Ce petit exemple démontre bien tout l'ambiguïté que peut apporter l'utilisation de l'erreur absolue.

Enfin, si on effectue plutôt une minimisation de l'erreur au carrée, mais cette fois par optimisation sous contraintes, d'autres ambiguïtés surviennent. Bien que la fonction soit concave, le problème ici vient du fait qu'elle ne l'est pas strictement. Peu importe le nombre de variables, en dérivant partiellement la fonction d'erreur au carrée pour toutes les variables et on faisant égaliser toutes ces dérivées partielles par 0, on obtient l'équation (4.6).

$$1 = \sum_{k=1}^{K-1} \tau_k C_k \quad (4.6)$$

L'équation (4.6) prend en compte la contrainte que la somme des contributions doit évaluer 1. Les variables τ_k correspondent à des scalaires qui ne sont pas reliés aux contributions. Le développement complet pour obtenir cette équation linéaire est présenté dans l'Annexe F.

Ainsi, l'équation (4.6) démontre que plusieurs combinaisons de contributions sont possibles pour obtenir l'égalité. Si différentes combinaisons se retrouvent à l'intérieur des contraintes, soit que toutes les contributions doivent avoir des valeurs positives et que la somme des $K - 1$ contributions ne doivent pas dépasser 1, il y a alors ambiguïté.

4.3.2 Les mathématiques sous φ

Pour comprendre l'influence de l'exposant φ pour modifier les valeurs de contributions, il faut tout d'abord considérer l'équation de la régression multilinéaire :

$$\mathbf{S} = (\mathbf{C}^{*T} \mathbf{C}^*)^{-1} \mathbf{C}^{*T} \mathbf{X} \quad (4.7)$$

Dans l'équation (4.7), \mathbf{C}^* correspond à la matrice des contributions dont toutes les contributions ont été modifiées par un exposant φ .

Pour la suite, on décompose l'équation (4.7) en deux parties. La première partie consiste au calcul d'une matrice carrée de dimension K (nombre de composants à extraire) :

$$\mathbf{A} = (\mathbf{C}^{*T} \mathbf{C}^*)^{-1} \quad (4.8)$$

La valeur absolue des nombres de la diagonale de la matrice \mathbf{A} sera toujours plus grande que celle des nombres hors diagonale. Cette propriété vient du fait que la matrice \mathbf{C} ne contient que des nombres positifs et que la somme des valeurs de chaque ligne de la matrice est de 1.

Or, l'exposant φ a un effet sur ces nombres. Plus cet exposant est petit, plus les nombres hors diagonale sont intenses, de façon absolue. En d'autres mots, les nombres hors diagonale tendent vers $-\infty$ avec une valeur de φ qui tend vers 0, alors que ces mêmes nombres tendent vers 0 lorsque la valeur de φ tend vers $+\infty$. À l'opposé, les nombres à l'intérieur de la diagonale tendent vers $+\infty$ avec une valeur de φ qui tend aussi vers $+\infty$. (Figure 4.3)



Figure 4.3 Schématisation de la matrice calculée dans la première partie de la régression multilinéaire

La deuxième partie de la régression multilinéaire résulte en des spectres :

$$\mathbf{B} = \mathbf{C}^{*T} \mathbf{X} \quad (4.9)$$

Dans cette partie de l'équation, les spectres se caractériseront différemment selon la valeur du φ .

- À $\varphi < 1$: Comme les valeurs de contribution tendent à être les mêmes, les différents spectres résultant de (4.9) tendront aussi à être les mêmes.
- À $\varphi > 1$: Plus φ est élevé, plus les spectres se distingueront. Cependant, plus l'exposant est élevé, de moins en moins de pixels sont considérés. Seulement les pixels qui contiennent une valeur de contribution près de 1 seront considérés.

En recombinaison des deux parties de la régression multilinéaire ensemble, on obtient cette équation pour le calcul d'une seule valeur dans la matrice spectrale :

$$S_{k,n} = \sum_{l=1}^K A_{k,l} B_{l,n} \quad (4.10)$$

Dans l'équation (4.10), tel que mentionné auparavant, la matrice \mathbf{A} contient un nombre positif lorsque $k = l$ et, autrement, lorsque $k \neq l$, la matrice contient un nombre négatif. Or, tel que mentionné précédemment, à φ faible, les nombres dans la matrice \mathbf{A} sont intenses.

La matrice \mathbf{A} contient l'information qui permet en quelque sorte de rajuster les spectres contenus dans la matrice \mathbf{B} . Si les nombres hors diagonale sont de faible intensité par rapport aux nombres dans la diagonale, le spectre k résultant de l'équation (4.10) aura donc une allure similaire au spectre k contenu dans la matrice \mathbf{B} . Il est à rappeler que le spectre S_k est normalisé à la suite de son calcul.

Si, au contraire, les nombres hors diagonale de \mathbf{A} ont une intensité notable, des problèmes peuvent survenir. Pour expliquer cette problématique, on prend par exemple une régression multilinéaire dans laquelle on cherche à calculer deux spectres. La matrice \mathbf{B} contient ces deux spectres « préliminaires » :

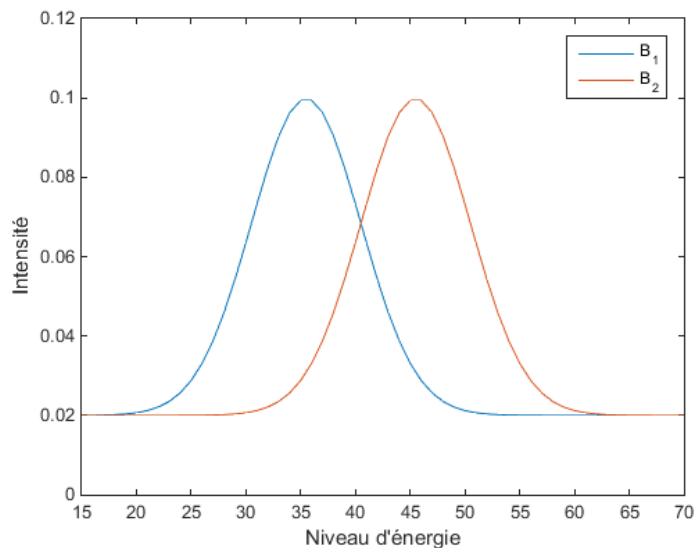


Figure 4.4 Matrice \mathbf{B} simulée

Ces deux spectres simulés contiennent un bruit de fond, car la base de leur pic n'est pas à une intensité de 0. De plus, il y a une grande présence de chevauchement entre les deux spectres.

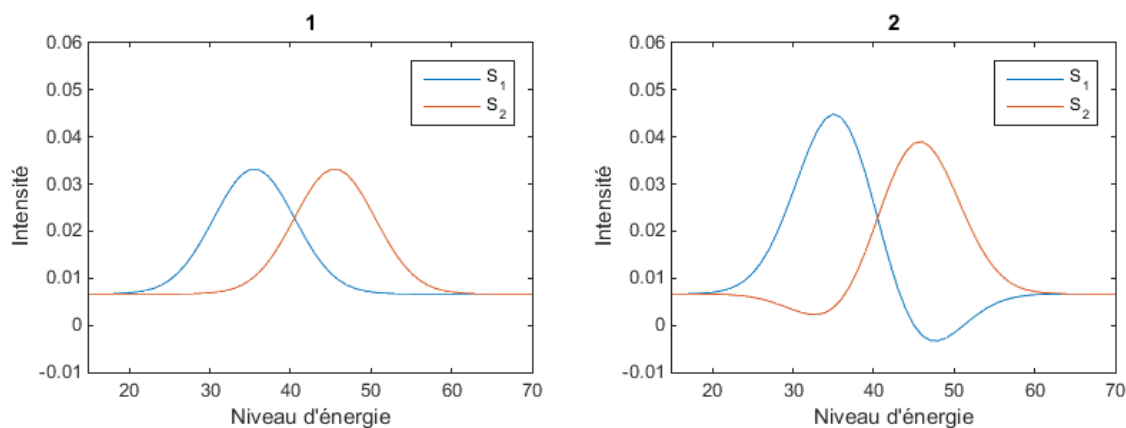


Figure 4.5 Spectres résultant après normalisation avec (1) une matrice \mathbf{A} dont les nombres hors diagonale sont à 0 et avec (2) une matrice \mathbf{A} dont les nombres hors diagonale ont une intensité notable

La Figure 4.5 montre comme 1^{er} graphique des spectres aux allures identiques à ceux simulés et contenus dans la matrice \mathbf{B} (Figure 4.4). Comme, dans ce cas, la matrice \mathbf{A} ne contient aucun nombre négatif (φ très élevé), il n'y a aucun rajustement entre les spectres finaux et ceux inclus dans \mathbf{B} .

Le 2^e graphique de la Figure 4.5 montre très bien la problématique lorsque les valeurs négatives dans \mathbf{A} sont d'une intensité notable (φ faible). Toujours en gardant en tête l'équation (4.10), aux niveaux d'énergie où les pics des différents composants se chevauchent, un des composants prendra de l'ampleur alors que les autres diminueront.

C'est le cas au niveau d'énergie 35 de l'exemple. Il y a présence d'un minimum local pour le composant 2, alors que le composant 1 a pris de l'ampleur comparativement aux spectres contenus dans \mathbf{B} . On nommera ces présences de minima locaux des pics négatifs dans les prochaines lignes. D'ailleurs, au niveau d'énergie 47, le pic négatif se retrouve même sous la barre des 0 pour le 1^{er} composant.

Certains résultats dans les Chapitres 5 et 6 montreront très bien cette problématique qu'est la création de pics négatifs.

4.3.3 Valeur de l'exposant φ optimale

La sous-section 4.3.2 a fait part des mathématiques entourant les modifications des valeurs de contributions dans la régression multilinéaire pour le calcul de la matrice spectrale. La sous-section 4.2.1 a expliqué qu'une initialisation simple, telle que par l'algorithme des k-moyennes, augmente les chances de convergence de l'algorithme *MCR-LLM*.

Tel que mentionné, une bonne initialisation en est une qui calcule des spectres physiquement cohérents, sans aucune discontinuité dans le spectre, et ce même si les spectres peuvent se ressembler entre eux. Ainsi, au fil des itérations de *MCR-LLM*, le but est de conserver la cohérence physique des spectres. Selon la valeur de l'exposant φ utilisée, différentes situations peuvent survenir en calculant de nouveaux spectres à la première itération de *MCR-LLM* et en les comparant aux spectres initiaux :

- φ trop faible : Risques de création de pics négatifs, tels que définis à la sous-section 4.3.2, et de valeurs négatives (remises à 0 par la suite). Conséquemment, les spectres sont grandement modifiés par rapport aux spectres initiaux.
- φ trop élevé : Spectres corrélés seulement aux pixels ayant des contributions s'approchant de 1. Ces spectres sont physiquement cohérents, mais l'algorithme

MCR-LLM ne peut évoluer, car ces spectres seront toujours recalculés à chaque itération.

- φ optimal : Spectres toujours cohérents physiquement. Présence de modifications par rapport aux spectres initiaux.

Le φ optimal doit être quantifié. Un moyen de le quantifier l'exposant optimal passe par la quantification de la modification M entre les spectres calculés à l'itération i et ceux calculés à l'itération $i - 1$:

$$M = \sum_{k=1}^K \sum_{n=1}^N (s_{i-1,k,n} - s_{i,k,n})^2 \quad (4.11)$$

Dans l'équation (4.11), N correspond au nombre de variables dans l'ensemble de données, alors que K correspond au nombre de composants à extraire. Si « $i - 1$ » égale à 0, cela réfère aux spectres calculés à l'initialisation

La Figure 4.6 démontre un exemple de calcul de M par rapport à des spectres à l'itération $i - 1$ considérés comme physiquement cohérents.

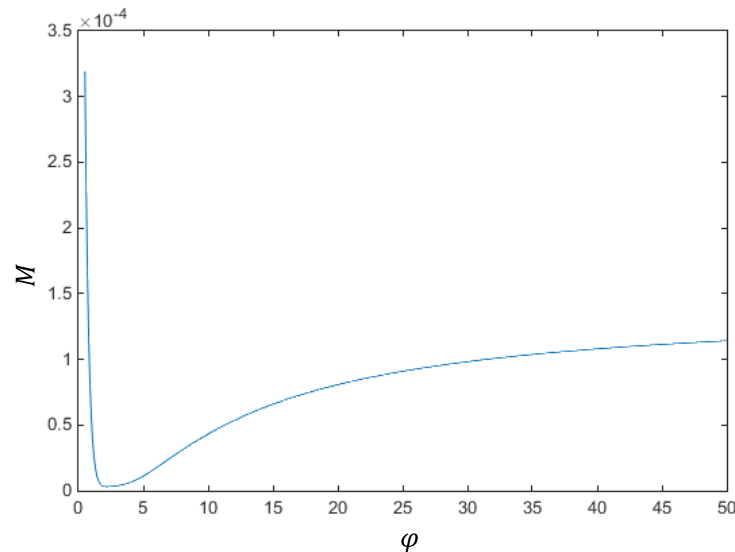


Figure 4.6 Exemple de la quantification de la modification des spectres selon différentes valeurs de φ

La courbe présentée par la Figure 4.6 démontre bien les conséquences d'un φ trop faible ou trop élevé, tel que mentionné précédemment. Ce qui est intéressant est la présence d'un minimum global, qui est d'ailleurs dans les alentours de $\varphi = 2$.

En démarrant l'algorithme *MCR-LLM* avec des spectres physiquement cohérents et en minimisant la modification des spectres entre ceux de l'itération précédente et ceux de l'itération présente, on augmente grandement les probabilités de garder des spectres physiquement cohérents au fil des itérations de *MCR-LLM*.

Or, le calcul des contributions par maximisation de la vraisemblance statistique (*sous-section 4.2.2.1*) est extrêmement sensible aux petites différences entre les spectres. Ainsi, malgré que la modification des spectres soit minimisée au fil des itérations, le calcul très précis des contributions permet de compenser ce fait et permet à l'algorithme *MCR-LLM* de converger petit à petit vers la bonne solution.

La minimisation de M afin de trouver le φ optimal peut être effectuée à chaque itération de *MCR-LLM*. Par la probabilité que la courbe de M par rapport à φ (*Figure 4.6*) puisse contenir des minima locaux, il est suggéré d'effectuer une optimisation/minimisation de M par algorithme génétique [43], ce qui augmente grandement les chances de converger vers le minimum global.

Cette minimisation de M afin de calculer le φ optimal à chaque itération nécessite peu de calcul, mais est optionnelle. L'exposant φ peut tout simplement être fixé pour toutes les itérations. Les Chapitres 5 et 6 présenteront des résultats obtenus avec des φ fixés et d'autres résultats obtenus par des φ calculés par minimisation de M à chaque itération.

4.3.4 *MCR-LLM* par rapport à d'autres algorithmes

Que ce soit dans *MCR-LLM* ou dans l'algorithme d'espérance-maximisation, le but est de maximiser une fonction de vraisemblance. Or, dans le cas de l'espérance-maximisation, le but est de maximiser une fonction de vraisemblance globale prenant en compte tous les composants et tous les pixels/observations, soit l'équation (3.28). En effectuant successivement les étapes d'espérances et de maximisation, la fonction de vraisemblance globale est continuellement maximisée.

Dans le cas de *MCR-LLM*, la maximisation de la fonction de vraisemblance se fait localement, soit dans le calcul des contributions pour chacun des pixels. Si on calcule la fonction de vraisemblance globale à chaque itération lors de *MCR-LLM*, on constate que, contrairement à l'algorithme d'espérance-maximisation, la valeur calculée n'augmente pas nécessairement à chaque itération.

Le but de l'algorithme d'espérance-maximisation est d'attribuer chaque pixel/observation à une ou l'autre des distributions et, par ces attributions, de calculer les paramètres régissant chacune des distributions. Au fil des itérations, les pixels sont de mieux en mieux distribués et, par conséquent, les paramètres régissant les distributions se rapprochent plus de la réalité. Ainsi, la fonction de vraisemblance globale augmentera au fil des itérations.

On définit un pixel contenant un mélange de composants comme étant un pixel dont son spectre représente une certaine combinaison des spectres associés à chacun des composants. La sommation « $\sum_{k=1}^K \mathbf{W}_{mk} \log(\boldsymbol{\pi}_k \times l(\mathbf{X}_m; \boldsymbol{\theta}_k))$ » de l'équation (3.28) désavantage les mélanges de composants, car les valeurs de l ne sont calculées qu'en considérant les spectres de chacun des composants et sont donc, par conséquent, de beaucoup moindre intensité. Aucun spectre résultant d'un mélange de composants n'est considéré. Dans *MCR-LLM*, comme plusieurs pixels reflètent un mélange de composants et que l'équation (3.28) ne favorise pas ce type de situation, la fonction de vraisemblance globale L sera donc plus faible avec *MCR-LLM* qu'avec l'algorithme d'espérance-maximisation.

Ainsi, malgré que la fonction de vraisemblance globale soit plus faible en fin de convergence avec *MCR-LLM*, l'utilisation de cet algorithme est plus approprié, car il maximise localement la fonction de vraisemblance et permet ainsi qu'un pixel représente un mélange de composants.

D'un autre côté, il est important de faire la distinction entre les buts de *MCR-ALS* et de *MCR-LLM*. Dans *MCR-ALS*, les résidus sont minimisés, alors que dans *MCR-LLM*, c'est la vraisemblance statistique locale, au niveau de chaque pixel, qui est maximisée. Ainsi, si on analyse l'évolution des résidus au fil des itérations de *MCR-LLM*, ceux-ci ne diminueront pas nécessairement.

4.4 Résumé du Chapitre 4

La méthode proposée (*MCR-LLM*) se distingue de *MCR-ALS*, principalement par rapport au calcul des contributions. Dans *MCR-ALS*, on effectue un calcul (régression multilinéaire) et on modifie par la suite les résultats pour que ceux-ci respectent les contraintes imposées. Dans l'algorithme proposé (*MCR-LLM*), au contraire, les contraintes sont considérées directement dans le calcul des contributions, grâce à l'optimisation/maximisation de la vraisemblance, ainsi qu'à la modification des valeurs de contributions par un exposant φ lors du calcul de la matrice spectrale par régression multilinéaire.

MCR-ALS standard

MCR-LLM

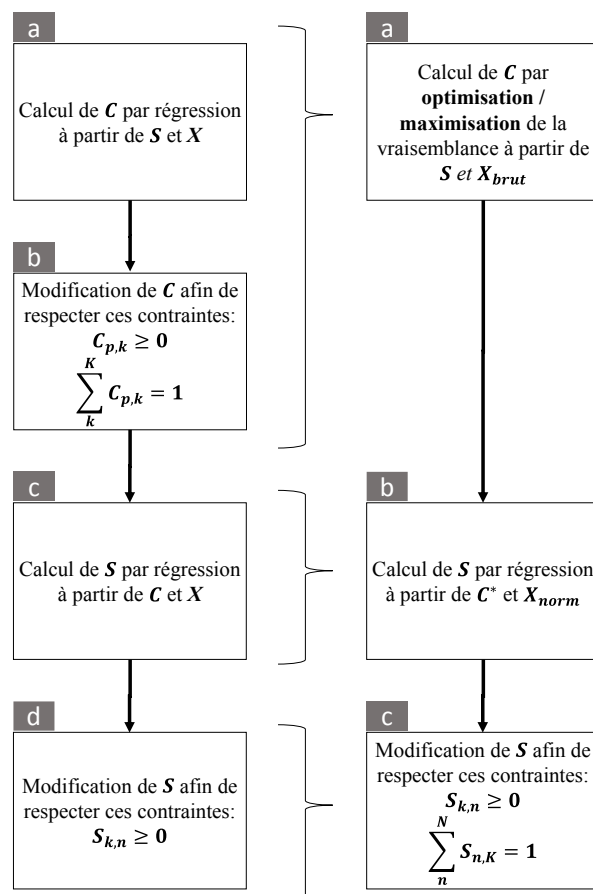


Figure 4.7 Comparaison entre une itération de MCR-ALS standard et une itération de MCR-LLM

Tableau 4.2 Acronyme utilisés dans le Chapitre 4

| Acronyme | Correspondance en anglais | Correspondance en français |
|------------------|---|--|
| <i>MCR-ALS</i> | <i>Multivariate Curve Resolution – Alternating Least Squares</i> | Résolution multivariée de courbes – Moindres carrés alternatifs |
| <i>MCR-LLM</i> | <i>Multivariate Curve Resolution – Loglikelihood Maximization</i> | Résolution multivariée de courbes – Maximisation du logvraisemblance |
| <i>SIMPLISMA</i> | <i>Simple-to-Use Interactive Self-modeling Mixture Analysis</i> | Analyse de mélange interactive facile à utiliser par auto-modélisation |

Tableau 4.3 Variables utilisées dans le Chapitre 4

| Variable | Définition |
|------------------------------|---|
| \mathbf{C} | Matrice des contributions |
| $\mathbf{C}_{i,k}$ | Contribution pour le pixel i et du composant k |
| \mathbf{S} | Matrice spectrale |
| $\mathbf{S}_{k,n}$ | Valeur spectrale pour le composant k et la variable n |
| \mathbf{X}_{norm} | Matrice des données acquises, normalisées pour chaque pixel |
| \mathbf{X}_{brut} | Matrice des données acquises brutes |
| $\mathbf{X}_{brut/norm,i,n}$ | Donnée acquise, brutes ou normalisée, pour le pixel i et la variable n |
| \mathbf{X}_{CP} | Matrice des données acquises avec la correction de champ plat appliquée |
| \mathbf{Q} | Matrice définissant à quels composants pures et à quelles proportions est associée chacun des composants non pures. |
| l | Fonction de vraisemblance |
| φ | Scalaire représentant l'exposant par lequel toutes les données de contributions sont modifiées dans le calcul de la matrice spectrale |
| \mathbf{W} | Matrice contenant le spectre maximisant la vraisemblance pour chaque pixel |
| $\mathbf{W}_{i,n}$ | Valeur spectrale maximisant la vraisemblance pour le pixel i et la variable n |
| \mathbf{P} | Matrice diagonale des poids dans la régression multilinéaire pondérée |
| τ_k | Scalaire relié au composant k |
| \mathbf{d} | Vecteur de scalaires contenant le nombre total de comptes par pixel |
| \mathbf{d}_i | Nombre total de comptes pour le pixel i |
| K | Nombre de composants à extraire |

5. APPLICATION DE LA MÉTHODE DÉVELOPPÉE SUR DES DONNÉES XPS

5.1 Liens entre la méthode développée et les données XPS

5.1.1 Comportement du signal

Les sous-sections 2.1 et 2.2 ont expliqué de quelle façon les photoélectrons sont comptés lors d'une acquisition et surtout comment sont bruitées ces données, c'est-à-dire, selon une loi de Poisson. Par l'utilisation du détecteur par délai de ligne, on approximera pour la suite des explications que le courant d'obscurité est nul, par le fait que celui-ci est extrêmement faible par rapport au signal détecté (*sous-section 2.1.3*).

Dans l'utilisation courante du *Axis Ultra DLD*, une correction de champ plat est appliquée par défaut. Cette correction est basée sur une calibration avec une feuille d'or effectuée au préalable. Afin d'acquérir le nombre réel de photons incidents au microcanal, qui est un nombre entier, il est primordial de désactiver la correction de champ plat, soit la correction qui permet de corriger la différence de sensibilité entre les pixels. Pour des acquisitions standards, cette correction est automatiquement appliquée aux données lors de l'enregistrement (*sous-section 2.2.3*). Si cette option n'est pas désactivée, les données enregistrées consistent ainsi en des nombres avec décimales et les statistiques reliées à la loi de Poisson ne peuvent être appliquées. Conséquemment, la méthode développée (*MCR-LLM*) ne peut être appliquée.

5.1.2 Spectres

La sous-section 2.3 a fait part de la signification physique du bruit de fond dans les spectres *XPS* et des méthodes les plus courantes pour modéliser le bruit de fond. Il a aussi été

démontré, à l'aide du développement inclus dans l'*Annexe C*, que le bruit de fond peut être considéré comme étant additif, selon la contribution de chaque composant.

Le comportement du bruit de fond est la principale préoccupation quant au comportement des spectres. L'*Annexe C* démontre aussi qu'en considérant un bruit de fond qui se comporte de façon linéaire avec l'intensité signal, les spectres seront linéaires, selon la contribution des différents composants. Or, si le matériau analysé s'écarte trop des hypothèses posées afin d'effectuer une approximation, la méthode d'analyse de données proposée souffrira d'imprécisions.

5.2 Ensemble de données des demi-cercles

C'est à partir de l'analyse *XPS* de l'échantillon représentant des demi-cercles que la méthode a été développée.

5.2.1 Description de l'échantillon analysé

Cet échantillon a été développé par photolithographie. Il repose à une gaufre de silicium sur laquelle une couche d'oxyde de titane de 100 nm a été déposée sur la totalité de sa surface. Ensuite, grâce à un photomasque, une couche additionnelle de titane métallique a été déposée de façon sélective en salles blanches.

Plusieurs motifs se retrouvent sur cet échantillon. La Figure 5.1 montre la zone de l'échantillon sélectionnée pour des fins d'analyse.

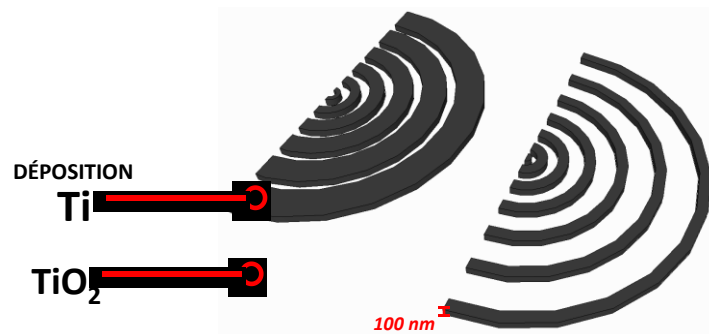


Figure 5.1 Représentation de l'échantillon des demi-cercles

Les motifs créés représentent des demi-cercles dont les largeurs des bandes sont de différentes tailles. Cet échantillon permet donc, entre autres, d'explorer la convolution spatiale créée par le XPS, et par conséquent, la résolution maximale.

Bien qu'il y ait eu à la base une déposition de titane métallique sur un substrat de TiO_2 , les deux différentes phases ne représentent pas au final des composés purs. Ainsi, pour la phase du titane métallique, bien que ce dernier reste en majorité dans la phase correspondante, on retrouvera aussi divers degrés d'oxydation. Il en sera de même pour la phase du TiO_2 , dans laquelle on retrouvera divers degrés d'oxydation, mais en majorité du TiO_2 .

Les pics devraient se retrouver aux énergies de liaison de 453,74 eV et de 459,79 eV pour la phase correspondant au titane métallique, alors que les pics devraient se retrouver aux énergies de liaison de 458,46 eV et de 464,23 eV pour la phase correspondant au TiO_2 .
(Figure 5.2)

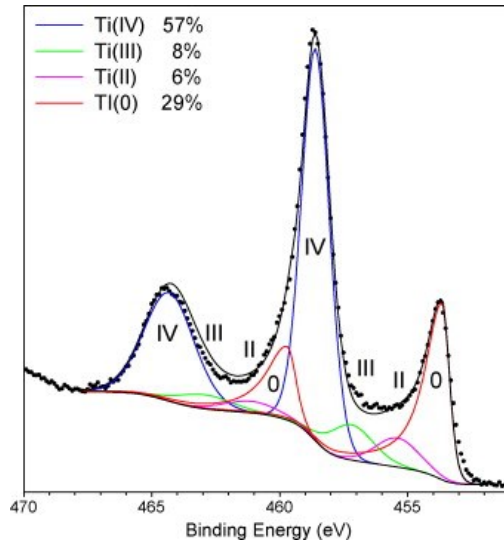


Figure 5.2 Spectres de référence [49]

5.2.2 Paramètres d'acquisition

L'instrument utilisé est le *XPS Axis Ultra DLD* de *Kratos* de l'Université de Sherbrooke. La source de rayons X utilisée est le Mono Al $K\alpha$. Le champ de vision utilisé est le deuxième (*FOV 2*), ce qui signifie que l'image résultante, de 256 par 256 pixels, correspond à un espace spatial de $400 \times 400 \mu\text{m}^2$. L'énergie de passage utilisée est de 80 eV.

La plage d'énergie de liaison analysée varie de 447 eV jusqu'à 465 eV, avec un incrément de 0,1 eV, ce qui crée un total de 181 niveaux d'énergie. Pour chaque énergie de liaison, le temps d'acquisition est de 300 secondes.

Les données ont été acquises en appliquant la correction de champ plat, ce qui résulte donc à des valeurs avec décimales. Ayant en possession la calibration de champ plat qui a été appliquée, il a ensuite été possible de retrouver les données brutes qui correspondent au nombre de photons incidents comptés, donc des nombres entiers.

5.2.3 Analyses PCA

Pour l'analyse PCA, ce sont les données dont la correction de champ plat est appliquée qui ont été utilisées. Le prétraitement effectué sur les données est celui qui est considéré

comme le plus approprié pour des données corrompues par un bruit de Poisson, appelé *Optimal Scaling*. (sous-section 3.1.1)

En analysant les matrices des *scores* résultantes (Figure 5.3), on peut facilement constater visuellement que la segmentation spatiale est bien effectuée dans la première composante. Les autres composantes ne montrent seulement que des artéfacts.

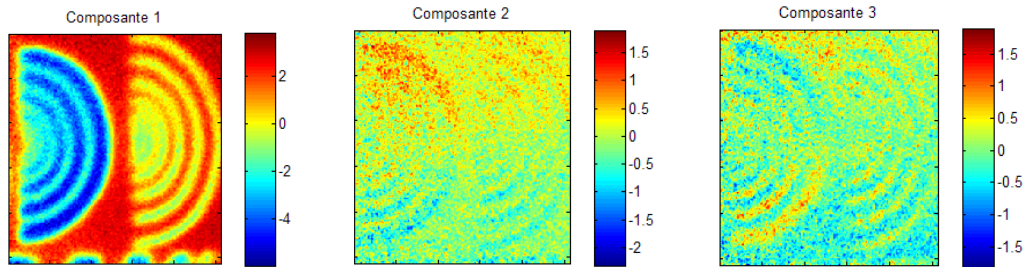


Figure 5.3 Matrice des *scores* de l'ensemble des données des demi-cercles, avec prétraitement *Optimal Scaling*

La courbe des *loadings* correspondant à la première composante présente contient plusieurs pics. Les pics correspondant aux extrema positifs sont associés à la phase de TiO_2 et les pics correspondant aux extrema négatifs sont associés à la phase du titane métallique. Chaque spectre comporte ainsi des éléments caractéristiques de TiO et de TiO_2 .

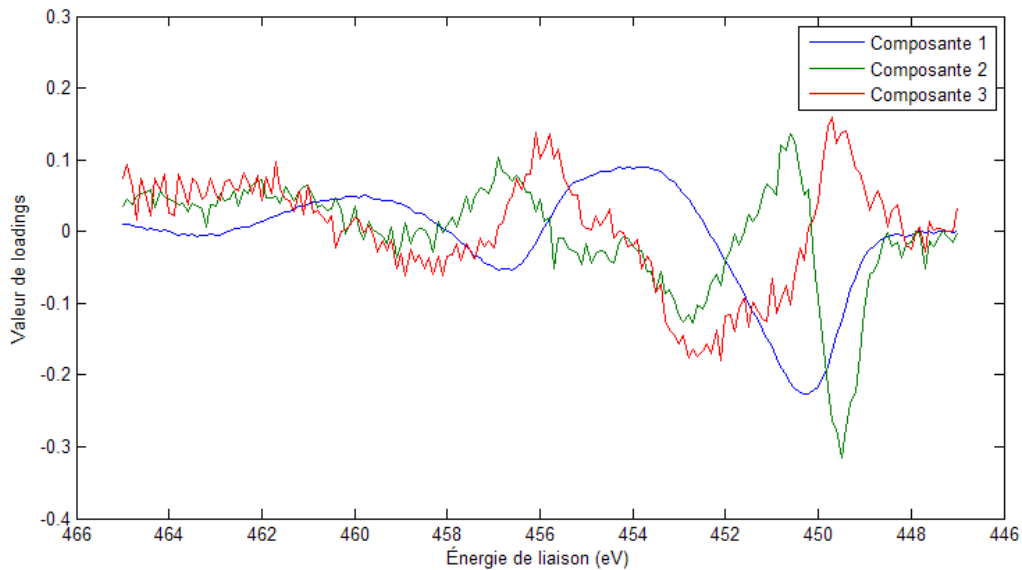


Figure 5.4 *Loadings* de l'ensemble de données des demi-cercles, avec prétraitement *Optimal Scaling*

Il est donc possible de mettre spatialement en évidence les deux composantes grâce à une seule composante de *PCA*. Cependant, on ne peut pas calculer de réelles contributions à partir de *scores*. Les *loadings*, qui ne présentent pas de vrais spectres, restent aussi difficiles à interpréter.

5.4.4 Analyses par *MCR-ALS* et *MCR-LLM*

Dans cette analyse, deux composants ($K = 2$) ont été extraits, car deux espèces sont recherchées. De plus, la contrainte exigeant que la somme des contributions pour chacune des observations soit égale à 1 est appliquée, de même que la contrainte exigeant qu'il ne doit y avoir aucune valeur négative au niveau des spectres et au niveau des contributions. La matrice des données (\mathbf{X}) est normalisée pour chaque observation. À convergence, les résultats obtenus sont les suivants :

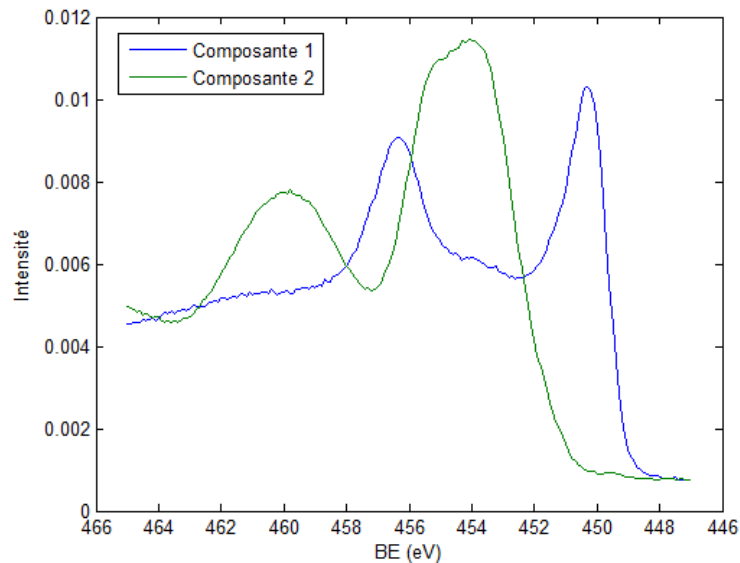


Figure 5.5 Spectres calculés par *MCR-ALS* avec l'ensemble de données des demi-cercles (Ti/TiO₂), acquisition de 300 secondes par niveau d'énergie

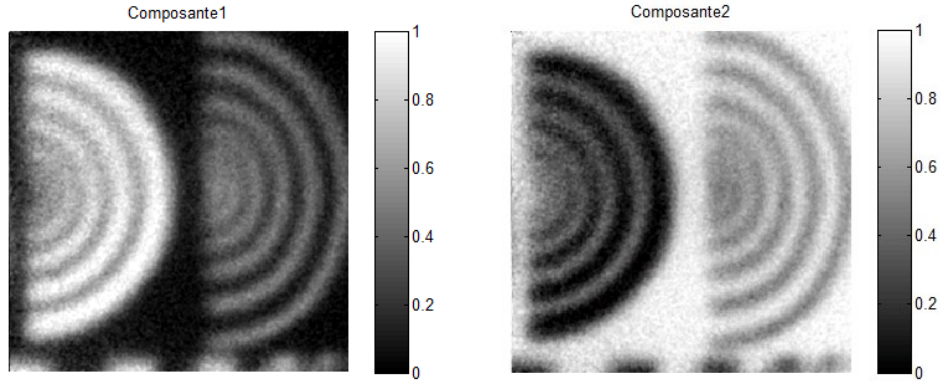


Figure 5.6 Contributions calculées par *MCR-ALS* avec l'ensemble de données des demi-cercles (Ti/TiO₂), acquisition de 300 secondes par niveau d'énergie

La Figure 5.5 montre deux spectres caractéristiques et physiquement cohérents comparativement aux spectres de référence (*Figure 5.2*). Le premier composant démontre les pics caractéristiques du titane métallique, alors que le deuxième composant démontre les pics caractéristiques de l'oxyde de titane (principalement TiO₂). La Figure 5.6 montre que les composants ont correctement été segmentés spatialement.

Les deux prochaines figures montrent les spectres et contributions calculés, mais cette fois-ci à partir de l'algorithme *MCR-LLM*, avec un exposant φ d'une valeur de 1. Tel que dit précédemment, avec seulement deux composants à extraire, la régression multilinéaire standard ($\varphi = 1$) (*Figure 5.7*) peut être utilisée pour révéler les mêmes spectres, à toutes fins pratiques, que ceux obtenus par *MCR-ALS* (*Figure 5.5*).

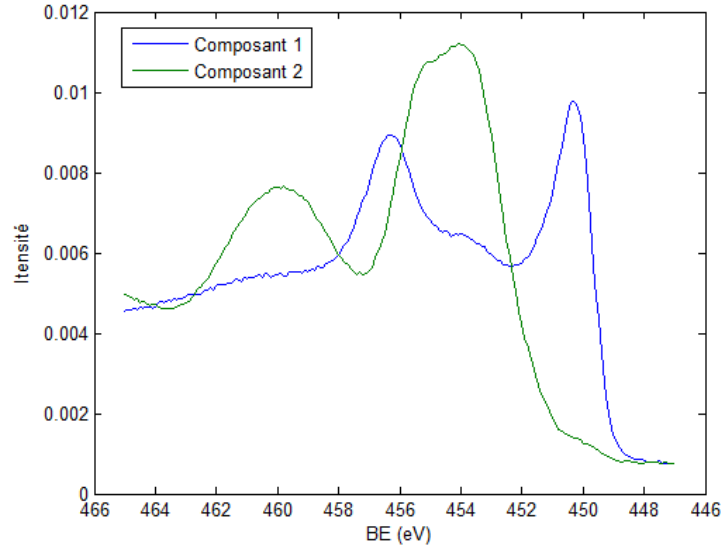


Figure 5.7 Spectres calculés par *MCR-LLM*, $\varphi = 1$, avec l'ensemble de données des demi-cercles (Ti/TiO_2), acquisition de 300 secondes par niveau d'énergie

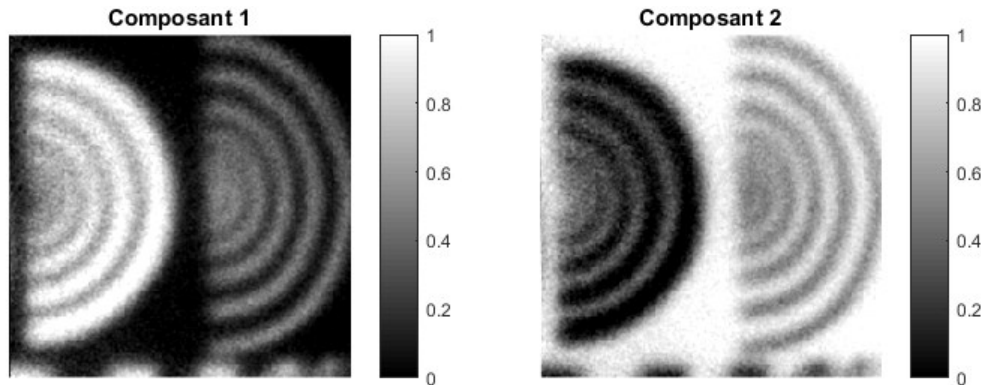


Figure 5.8 Contributions calculées par *MCR-LLM*, $\varphi = 1$, avec l'ensemble de données des demi-cercles (Ti/TiO_2), acquisition de 300 secondes par niveau d'énergie

En comparant les résultats obtenus par *MCR-ALS* standard (*Figure 5.5 et Figure 5.6*) à ceux obtenus par *MCR-LLM* (*Figure 5.7 et Figure 5.8*), on constate que les résultats sont semblables. Cependant, les spectres obtenus par *MCR-LLM* ($\varphi = 1$) présentent un léger biais. Les spectres sont quelque peu moins distincts entre eux par rapport à ceux de *MCR-ALS*. Ce biais est causé par le fait que l'optimisation rapide par *Active Set* peut avoir de la difficulté à calculer correctement les contributions d'un pixel presque pur (qui contient presque juste un seul composant).

Par exemple, on devrait calculer, pour un pixel, des contributions des composants 1 et 2 de, respectivement, 0,98 et 0,02. L'optimisation pourrait plutôt calculer des contributions de 1,00 et de 0. Ainsi, un fil des itérations, ce léger biais mènera à un spectre final quelque peu biaisé.

Ainsi, dans le cas d'une analyse d'un ensemble de données avec de relativement bons ratios signal sur bruit et en n'extrayant que deux composants, *MCR-ALS* standard sera quelque peu plus performant que *MCR-LLM*. Or, la prochaine section démontrera, qu'au contraire, avec plus deux composants à extraire et avec des données très bruitées, l'utilisation de *MCR-ALS* n'est pas recommandable.

5.3 Ensemble de données de trois métaux

5.3.1 Description de l'échantillon

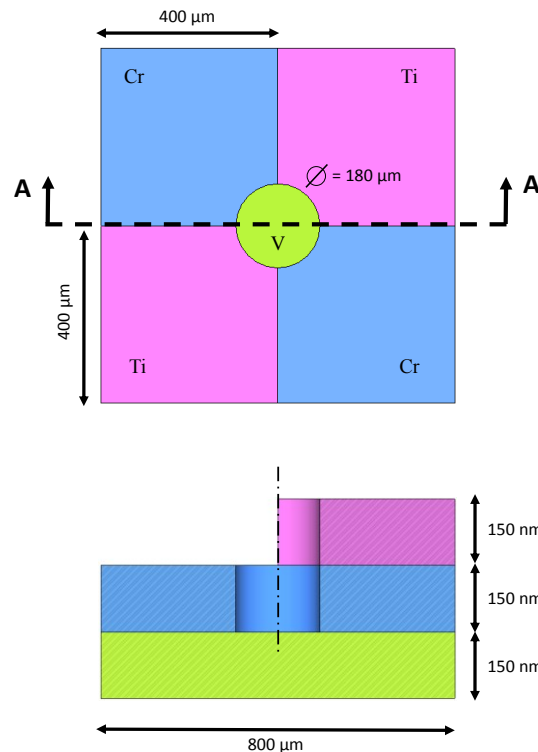


Figure 5.9 Représentation en trois dimensions de l'échantillon V/Cr/Ti

Les motifs présentés par la Figure 5.9 proviennent de l'échantillon créé par photolithographie et dont le procédé de fabrication est expliqué en détail dans l'*Annexe D Fabrication des échantillons*.

Les caractéristiques importantes à noter sur cet échantillon sont, notamment, les matériaux en présence, soit le vanadium, le chrome et le titane selon le motif illustré à la Figure 5.9. Alors que le titane et le chrome occupent un espace important et égal entre eux sur l'échantillon (48 % de la surface pour chacun des deux), le vanadium représente seulement 4 % de la surface. (Figure 5.9) Chaque couche est d'une épaisseur de 150 nm, ce qui est beaucoup plus grand que l'épaisseur caractéristique de ~10 nm sondés par le XPS.

5.3.2 Paramètres d'acquisition

L'instrument utilisé est le *XPS Axis Ultra DLD* de *Kratos* de l'Université de Sherbrooke. La source de rayons X utilisée est le Mono Al K α . Le champ de vision utilisé est le premier (*FOV 1*), ce qui signifie que l'image résultante, de 256 par 256 pixels, correspond à un espace spatial de 800 par 800 μm . L'énergie de passe utilisée est de 40 eV.

La plage d'énergie de liaison analysée varie de 607 eV jusqu'à 430 eV, avec un incrément de 0,33 eV, ce qui crée un total de 537 images par acquisition. Le temps d'acquisition par énergie de liaison a varié de 0,33 seconde à 20 secondes. Les données ont été acquises avec la calibration spatiale désactivée dans le logiciel d'acquisition.

Le nombre de compte moyen pour un pixel pour un niveau d'énergie pour chacun des différents temps d'acquisition est présenté par la Figure 5.10.

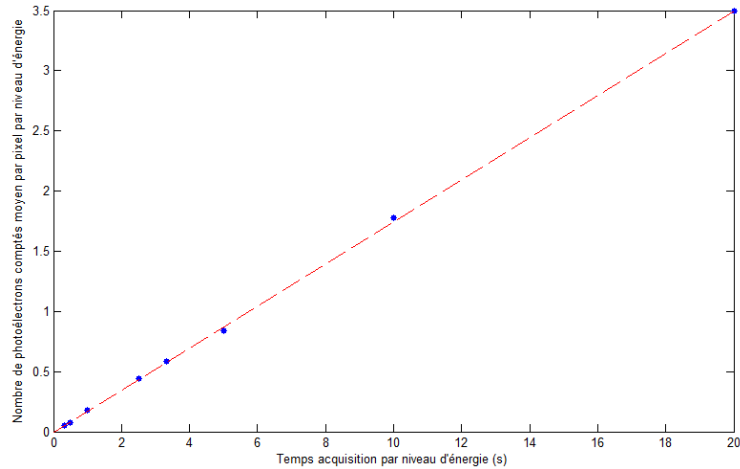


Figure 5.10 Nombre de photoélectrons comptés moyen pour un pixel pour un niveau d'énergie

En analysant la Figure 5.10, tel qu'attendu, on remarque que le nombre de photoélectrons comptés moyen est directement proportionnel au temps d'acquisition par niveau d'énergie. Ainsi, le ratio signal sur bruit des données dépend directement du temps d'acquisition et le détecteur ne présente pas de symptôme de saturation pour l'acquisition obtenue.

5.3.3 Analyses PCA

Cette sous-section démontrera les différents résultats obtenus par *PCA* à partir de différents prétraitements. Afin de comparer les *loadings* obtenus par les différentes analyses *PCA*, les spectres de référence pour le chrome et le vanadium sont présentés par la Figure 5.11. Le spectre de référence de titane a été présenté à la Figure 5.2.

(a)

(b)

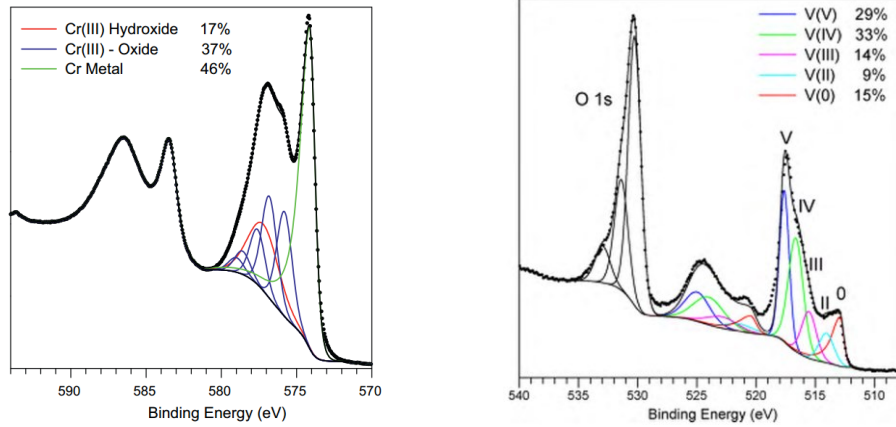


Figure 5.11 Spectres de référence pour (a) le chrome et (b) le vanadium [49], [50]

La Figure 5.12 démontre les matrices de *scores* des trois premiers composants résultant d'une analyse *PCA* avec des données prétraitées par *Optimal Scaling*. Les données utilisées sont celles qui ont été acquises avec un temps d'acquisition de 1 seconde par niveau d'énergie. Aucune calibration spatiale n'a été appliquée sur les données.

Dans la matrice des *scores* associée au premier composant, on remarque donc la variation de sensibilité entre les pixels par les petites raies obliques. Ces petites raies obliques sont aussi visibles dans le troisième composant.

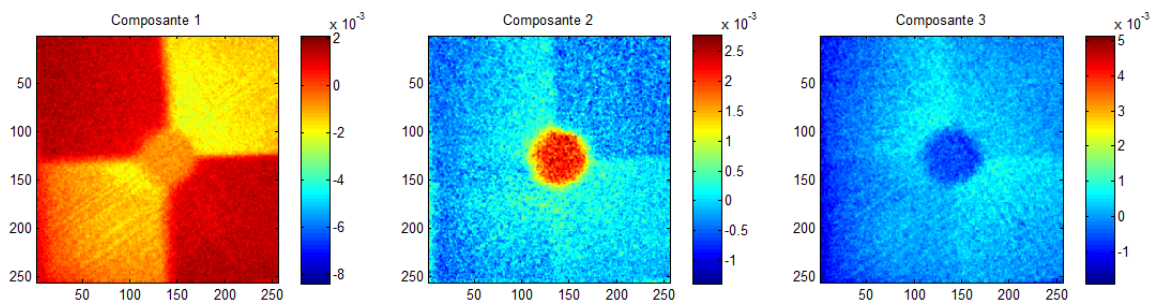


Figure 5.12 Matrice des *scores* obtenues par *PCA* à partir de l'ensemble des données avec trois métaux (Ti/Cr/V), prétraitement *Optimal Scaling*, acquisition de 1 seconde par niveau d'énergie

En analysant les *scores* du premier composant, on remarque le contraste entre le titane (zones jaunes) par rapport au chrome (zones rouges). La matrice des *scores* du deuxième composant met principalement en contraste le vanadium par rapport aux deux autres métaux. Ainsi, il est possible d'effectuer une certaine segmentation spatiale pour un

échantillon comportant trois composés distincts avec seulement les deux premiers composants d'une analyse *PCA*. Cette segmentation doit cependant être supervisée.

Pour effectuer cette segmentation, des pixels sont sélectionnés à partir des matrices de *scores* en ne sélectionnant que les pixels en-dessus ou en-dessous d'une certaine valeur seuil. Les données des pixels extraits sont ensuite moyennées pour chacune des variables afin d'en extraire des spectres. [51]–[53] Cette méthode, par le paramétrage manuel de valeurs seuils, comporte beaucoup d'incertitudes et les résultats peuvent différer grandement d'une personne à l'autre. De plus, il peut être difficile d'extraire des spectres de zones caractéristiques quand l'information se retrouve dans plus d'un composant.

Il est à rappeler qu'un spectre XPS permet de déduire de l'information précise quant à l'élément analysé, soit son degré d'oxydation et son environnement chimique. Les détails qui peuvent être extraits du spectre font l'avantage de l'utilisation du XPS.

Or, les *loadings* calculés ne démontrent pas de spectres physiquement cohérents comparativement aux spectres de référence (*Figure 5.2 et Figure 5.11*). Si le but est d'extraire des contributions et d'associer ces contributions à des spectres, l'information contenue dans les *loadings* est en conséquence inutile. Or, les « spectres » provenant de *loadings* permettent seulement de situer la position énergétique des pics caractéristiques des différents composés en présence. Par l'incohérence physique des spectres, il donc impossible d'utiliser le XPS à son plein potentiel.

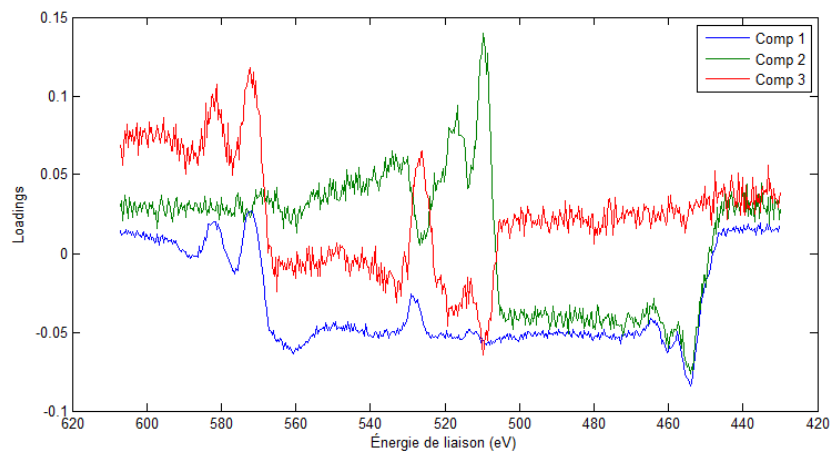


Figure 5.13 *Loadings* obtenus par *PCA* à partir de l'ensemble des données avec trois métaux (Ti/Cr/V), prétraitement *Optimal Scaling*, acquisition de 1 seconde par niveau d'énergie

Les deux prochaines figures montrent les résultats d'une analyse *PCA*, cette fois-ci avec des données centrées-réduites comme prétraitement. Les conclusions sont semblables à l'analyse précédente. Une segmentation spatiale est toujours possible, mais sous supervision.

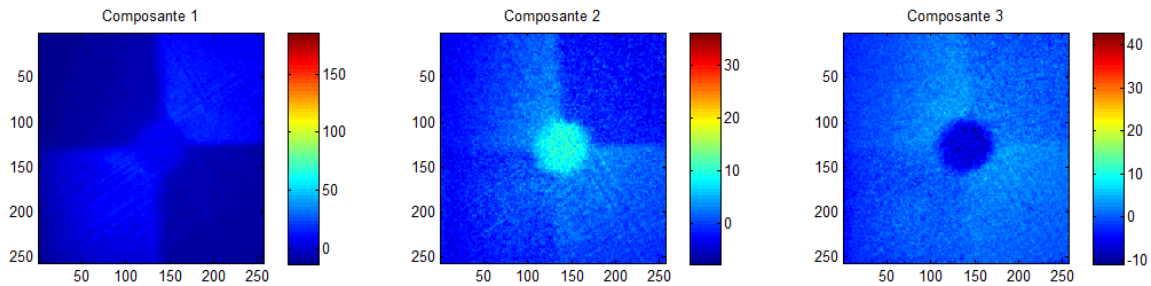


Figure 5.14 Matrice des *scores* obtenues par *PCA* de l'ensemble des données avec trois métaux (Ti/Cr/V), données centrées-réduites, acquisition de 1 seconde par niveau d'énergie

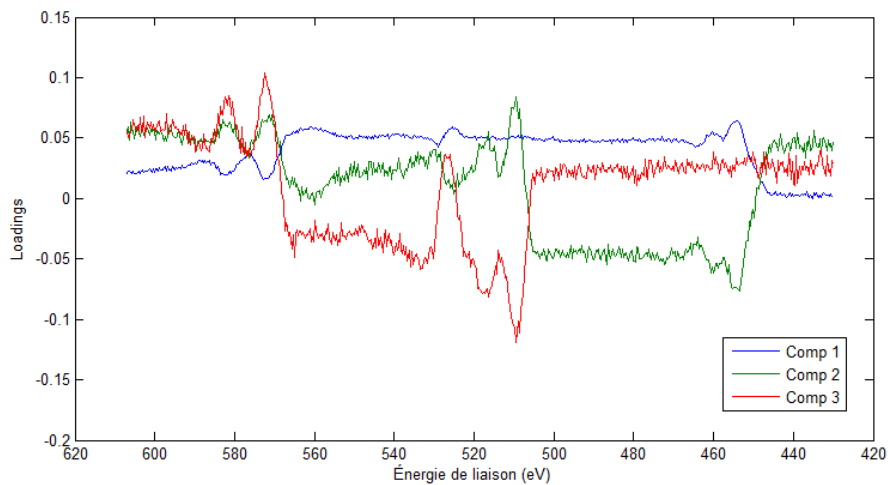


Figure 5.15 *Loadings* obtenus par *PCA* de l'ensemble des données avec trois métaux (Ti/Cr/V), données centrées-réduites, acquisition de 1 seconde par niveau d'énergie

Enfin, les deux prochaines figures démontrent les résultats d'une analyse *PCA*, mais avec un prétraitement plus adapté aux caractéristiques des données. Ainsi, comme aucune calibration spatiale n'a été appliquée aux données, tous les spectres-pixels sont normalisés en premier lieu. Ce calcul fait donc office de calibration spatiale. En deuxième lieu, les données sont centrées et réduites.

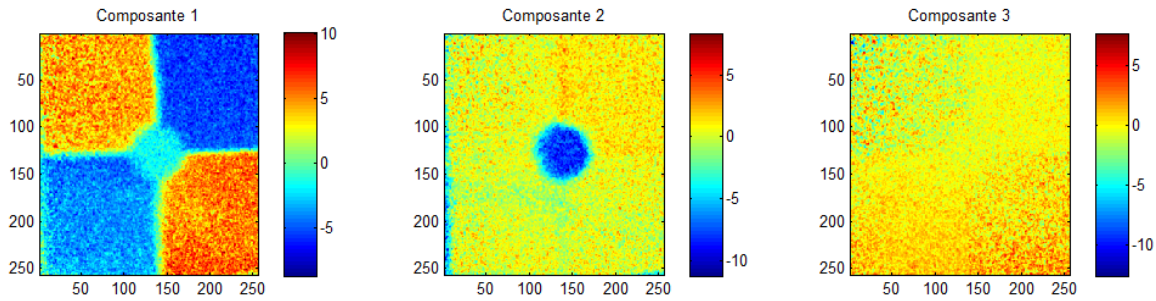


Figure 5.16 Matrice des *scores* obtenues par *PCA* de l'ensemble des données avec trois métaux (Ti/Cr/V), données normalisées au niveau spectral et ensuite centrées-réduites, acquisition de 1 seconde par niveau d'énergie

Contrairement aux deux dernières analyses, les *scores* du troisième composant résultant (Figure 5.16) ne contiennent pas vraiment d'information significative. Cette affirmation peut se confirmer en analysant les *loadings* de ce troisième composant (Figure 5.17) que ne présentent pas vraiment de pics se distinguant de la courbe, qui est d'ailleurs très bruitée.

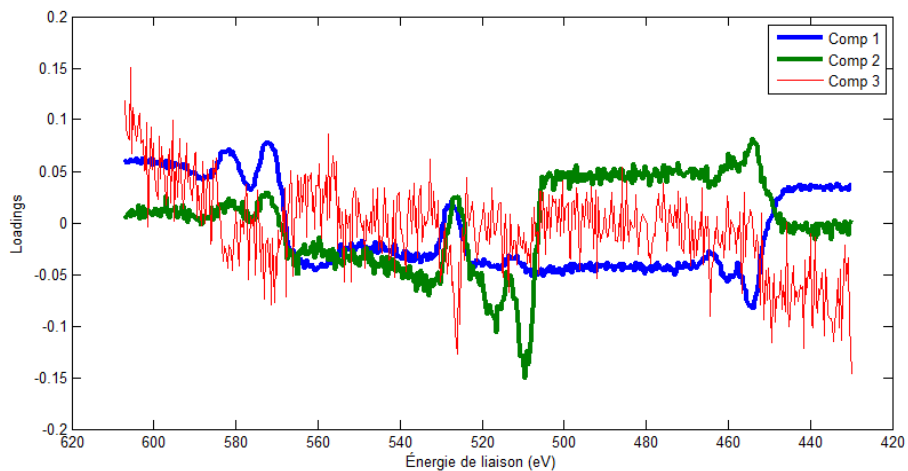


Figure 5.17 *Loadings* obtenus par *PCA* de l'ensemble des données avec trois métaux (Ti/Cr/V), données normalisées au niveau spectral et ensuite centrées-réduites, acquisition de 1 seconde par niveau d'énergie

Dans la dernière analyse *PCA*, l'information pertinente à la segmentation des données n'est contenue que dans les deux premiers composants. Dans les deux premières analyses présentées, de l'information pertinente se trouvait dans le troisième composant, ce qui amplifie la difficulté de l'analyse.

Tel que mentionné précédemment, toute calibration spatiale a été retirée de l'ensemble de données. Or, si une calibration spatiale avait été appliquée et que celle-ci s'avérait adéquate, une analyse avec des données prétraitées par *Optimal Scaling* aurait donc donné de meilleurs résultats que ceux présentés par la Figure 5.12.

5.3.4 Aperçus des résultats à travers les itérations de *MCR-ALS* et *MCR-LLM*

Les prochaines figures présenteront les contributions et les spectres calculées au fil des itérations par l'utilisation de *MCR-ALS* standard et de la méthode proposée (*MCR-LLM*).

Premièrement, avec une acquisition de 5 secondes à chaque énergie de liaison, on obtiendra l'évolution présentée par la Figure 5.18 avec *MCR-LLM* et avec un paramètre de modification des contributions φ paramétré à 2. Tel que mentionné auparavant, l'initialisation utilisée est l'algorithme des k-moyennes.

itération

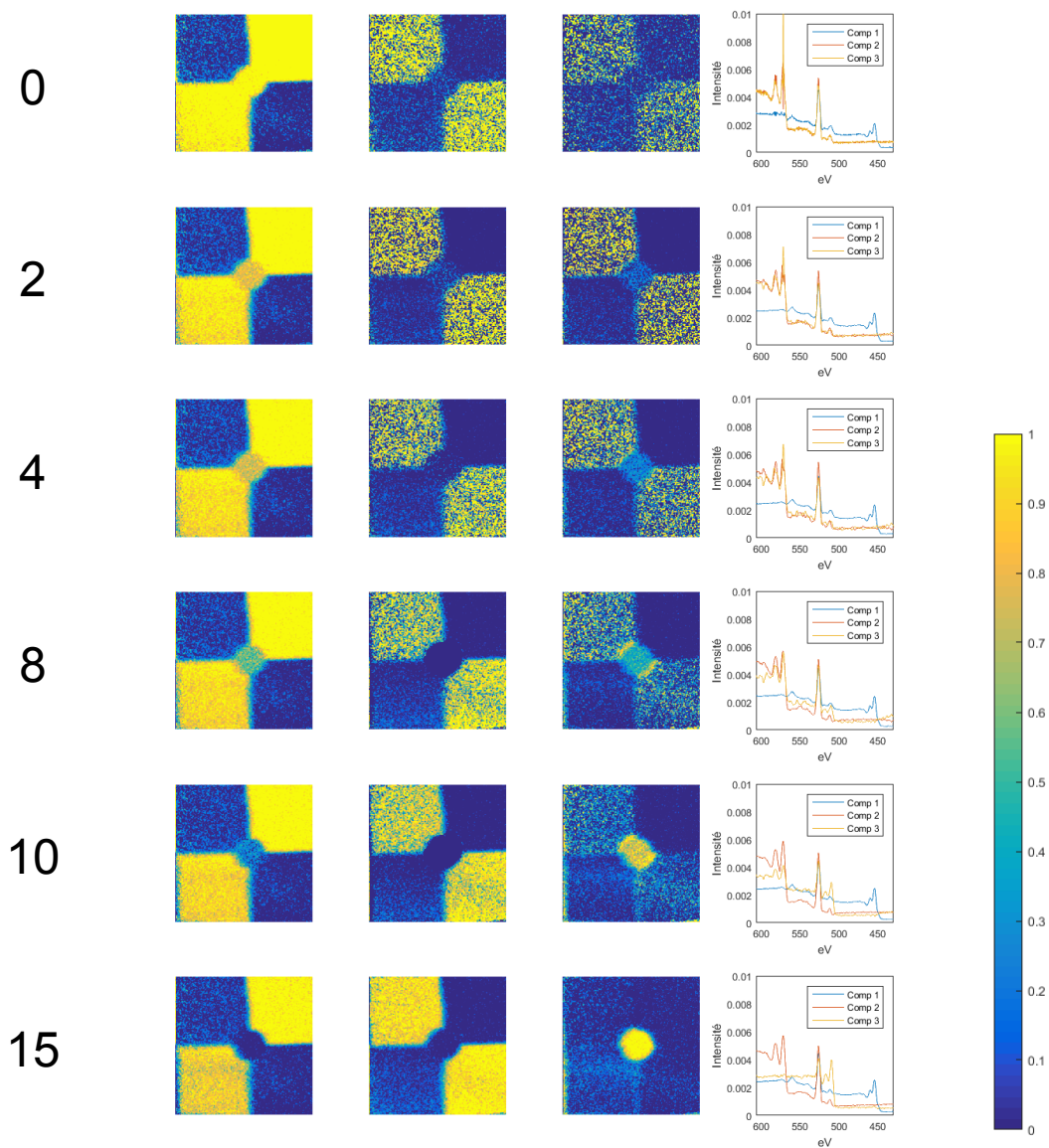


Figure 5.18 Résultats obtenus par *MCR-LLM* à partir de l'ensemble des données avec trois métaux (Ti/Cr/V) au fil des itérations, $\varphi = 2$, acquisition de 5 secondes par niveau d'énergie

En analysant la première ligne de la Figure 5.18, on constate que l'initialisation par l'algorithme des k-moyennes a tendance à inclure deux composés sous un même composant, alors que le troisième composé est dispersé dans deux composants. En analysant les spectres initiaux, on constate que les spectres en premier et du deuxième composant sont presque identiques. C'est au fil des itérations que l'algorithme réussira à

isoler chacun des trois composés et de les associer à trois composants distincts. D’ailleurs, la dernière ligne de la Figure 5.18 démontre bien trois zones spatiales distinctes, ainsi que trois spectres distincts.

La Figure 5.19 montre les résultats de convergence pour les mêmes données que celles utilisées à la Figure 5.18, mais avec *MCR-ALS* standard, en prenant en compte les contraintes normalement imposées à des données d’imagerie spectrale.

itération

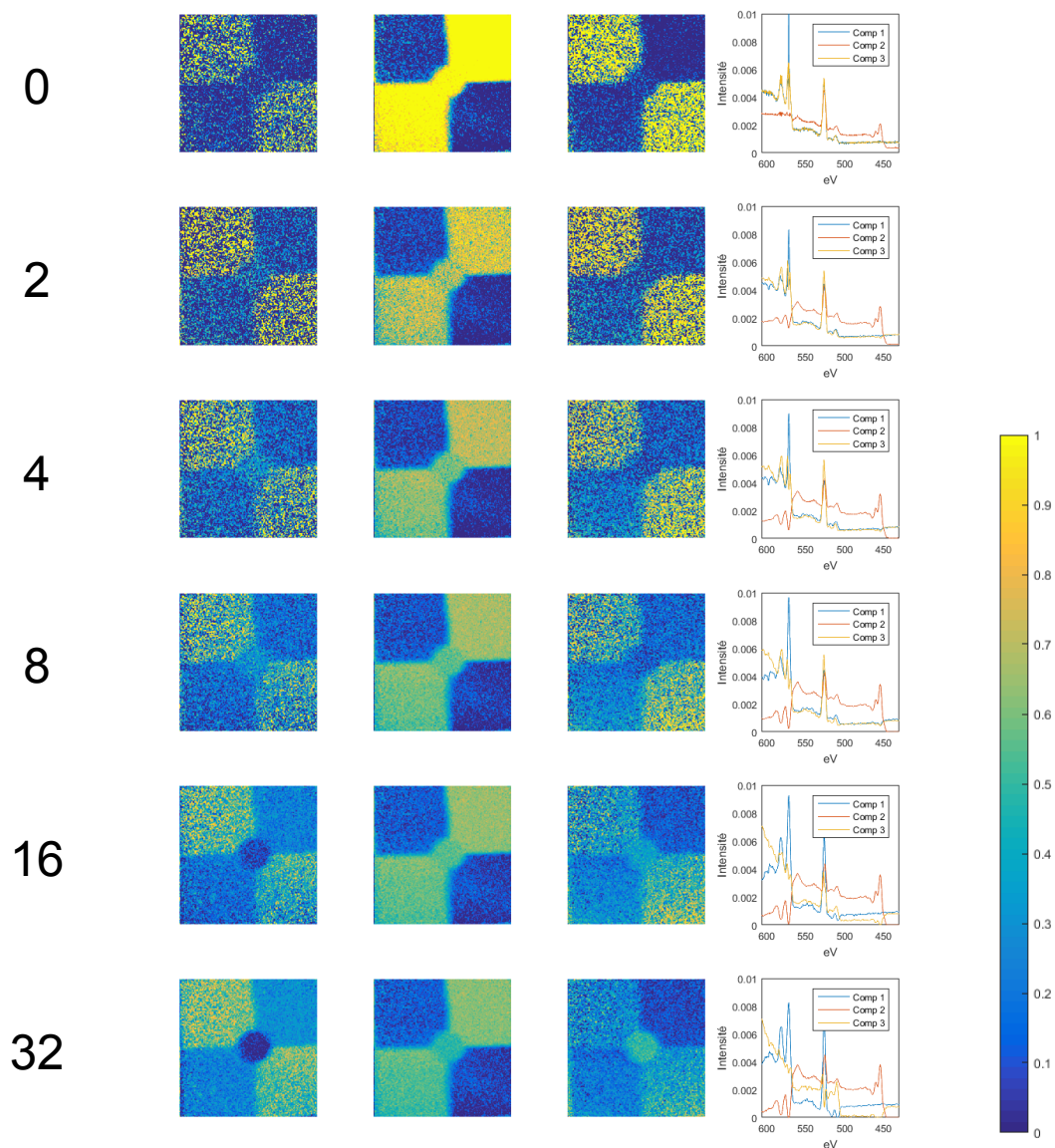


Figure 5.19 Résultats obtenus par *MCR-ALS* standard à partir de l’ensemble des données avec trois métaux (Ti/Cr/V) au fil des itérations, acquisition de 5 secondes par niveau d’énergie

Tel qu'affichée sur la dernière ligne de la Figure 5.19, l'algorithme *MCR-ALS* standard n'est pas en mesure de bien segmenter les trois composés dans trois différents composants. Le troisième composant conserve deux composés, alors que le premier et deuxième composants se partageront essentiellement le même composé.

Il est aussi intéressant d'analyser les résultats en utilisant les modifications proposées par *MCR-LLM* (Chapitre 4), mais en retirant le calcul des contributions par maximisation de la vraisemblance. Cela signifie qu'on modifie les valeurs de contributions par un exposant φ dans le calcul des matrices spectrales, mais qu'on calcule les contributions par simple régression multilinéaire. Afin d'alléger le texte, cette méthode modifiée sera appelée *MCR-ALS+ φ* dans les prochaines lignes. La Figure 5.20 illustre les résultats à convergence de cette méthode pour différentes valeurs de φ .

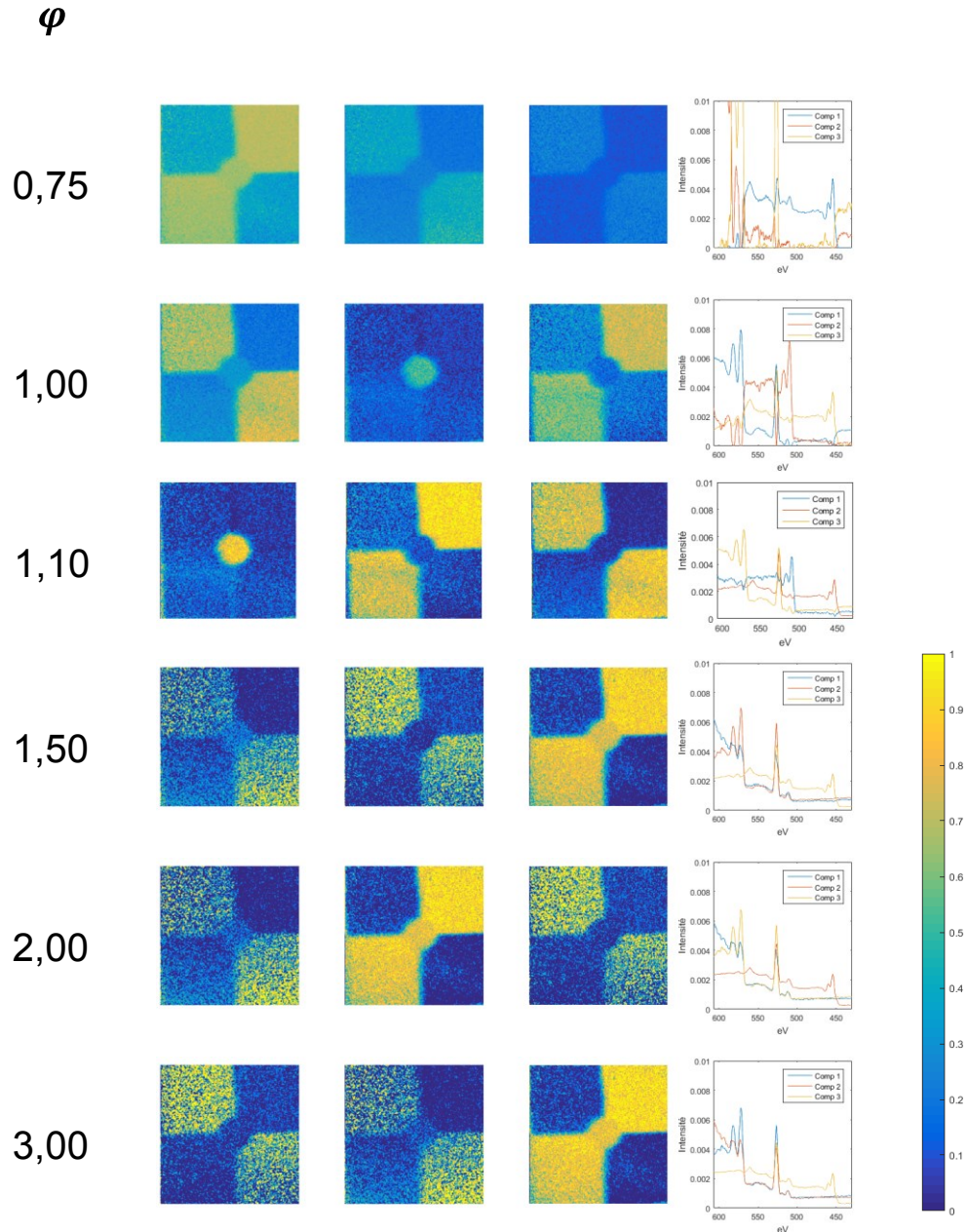


Figure 5.20 Résultats à convergence obtenus par $MCR-ALS+\varphi$ à partir de l'ensemble de données avec trois métaux (Ti/Cr/V), acquisition de 5 secondes par niveau d'énergie

En analysant la Figure 5.20, on constate que les résultats de convergence diffèrent beaucoup d'une valeur de φ à l'autre. À valeur φ s'approchant de 1 (*deuxième et troisième lignes*), les résultats se rapprochent de ceux obtenus avec $MCR-LLM$ (*dernière ligne de la Figure 5.18*), sans toutefois les atteindre complètement. Spatialement, les différentes zones ne sont pas autant clairement définies. Au niveau spectral, les spectres comportent quelques

sections problématiques (Figure 5.21) qui sont principalement dues à la création de pics négatifs, tel qu'expliqué à la sous-section 4.3.3.

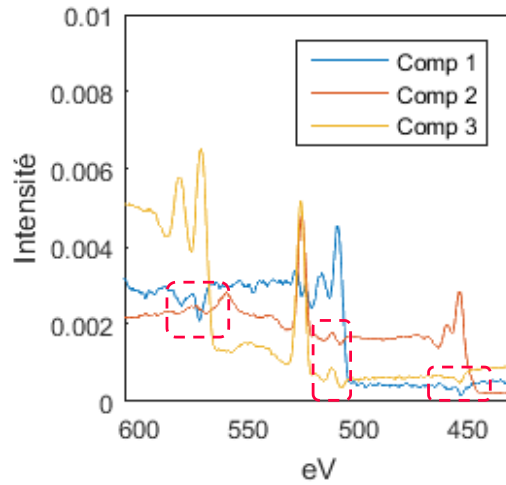


Figure 5.21 Problématiques au niveau des spectres obtenus par $MCR-ALS+\varphi$ à partir de l'ensemble de données avec trois métaux (Ti/Cr/V), $\varphi = 1,1$, acquisition de 5 secondes par niveau d'énergie

En augmentant la valeur de φ , les résultats à convergence restent très semblables aux contributions et spectres calculés à l'initialisation.

Au niveau des contributions, $MCR-ALS$ standard a pour effet de calculer davantage des pixels se rapprochant de substances pures, alors qu'on devrait plutôt calculer un mélange de composants qui représente un constituant. En transformant les contributions avec une valeur φ élevée, on amplifie donc l'erreur causée par $MCR-ALS$. Cette amplification d'erreur empêchera donc les résultats d'évoluer au fil des itérations, tel qu'il est possible de constater sur les dernières lignes de la Figure 5.20 et tel qu'il a été expliqué à la sous-section 4.3.3.

La Figure 5.22 présente cette fois-ci les résultats à convergence obtenus avec une acquisition de seulement une seconde par niveau d'énergie.

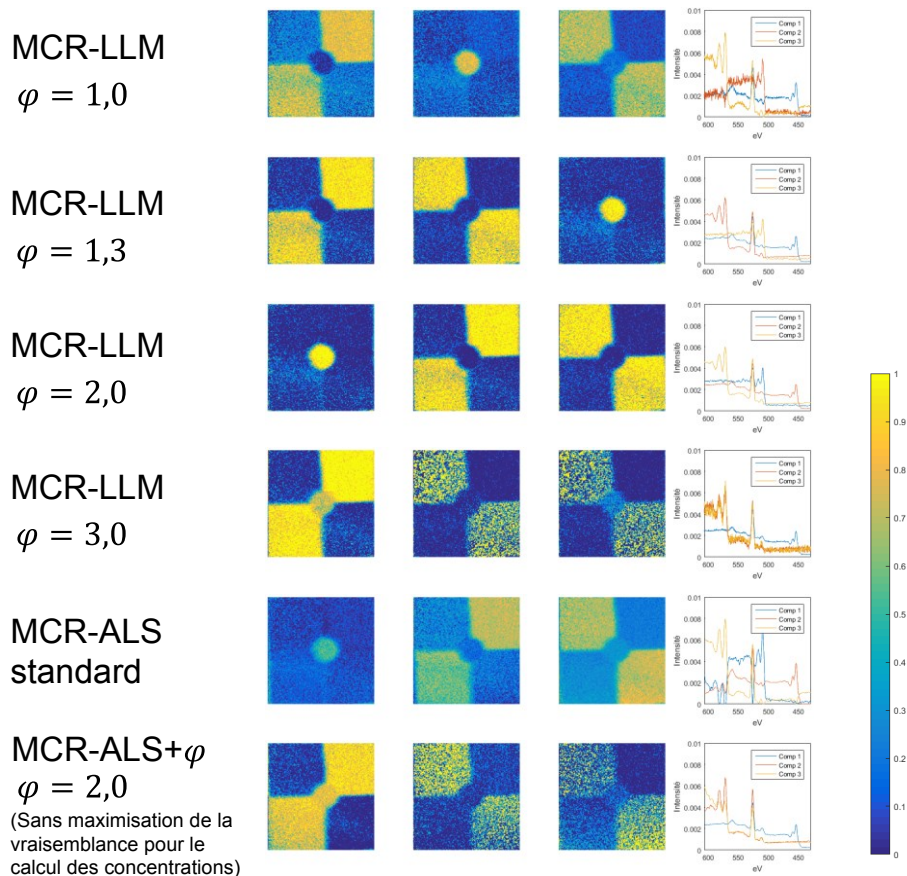


Figure 5.22 Résultats à convergence de différents algorithmes avec une acquisition d'une seconde par énergie de liaison

Tel que présenté par le graphique de la Figure 5.10, le nombre moyen de comptes par pixel et par niveau d'énergie est extrêmement faible avec une acquisition de seulement une seconde par niveau d'énergie. Or, en analysant les résultats présentés par la Figure 5.22, on constate l'avantage de prendre en compte le comportement du bruit (*première ligne*) dans le traitement des données. Par l'utilisation de *MCR-ALS* standard, il est difficile d'extraire trois zones spatiales distinctes et trois spectres cohérents, alors qu'avec la méthode proposée (*MCR-LLM*), on est en mesure de distinguer les trois zones, malgré un ratio du signal sur bruit particulièrement faible.

Or, toujours dans l'analyse de la Figure 5.22, on constate qu'avec une valeur φ trop élevée, l'algorithme *MCR-LLM* n'est pas en mesure de faire évoluer les valeurs à partir de celles de l'initialisation.

Les sous-sections 4.3.2 et 4.3.3 ont démontré et expliqué qu'en modifiant les valeurs de contribution par un exposant φ dans le calcul des spectres, on diminue l'interaction entre les différents composants. Ainsi, comme la valeur de φ est trop élevée, la régression multilinéaire dans le calcul des spectres négligera l'interaction entre les différents composants. L'opération revient, en quelque sorte, à recalculer les spectres associés seulement aux pixels contenant des valeurs de contribution s'approchant de 1.

Ainsi, les spectres ressembleront presque totalement à ceux de la dernière itération. Les résultats à convergence, autant pour les contributions que les spectres, seront donc très semblables à ceux calculés à l'initialisation. L'utilisation d'une valeur de φ entre 1 et 2 permet de considérer une certaine interaction entre les composants, ce qui perturbe suffisamment les donner pour stimuler leur évolution au fil des itérations. Une valeur au-delà de 2 amène le risque d'une stagnation des valeurs à celles obtenus par l'initialisation.

La sous-section 4.3.3 a démontré comment la valeur de φ peut être optimisée à chaque itération, sans que celle-ci ait à être définie. La Figure 5.23 montre les résultats obtenus à convergence avec l'ensemble de données à une seconde d'acquisition par énergie de liaison, mais sans fixer une valeur de φ dans *MCR-LLM*.

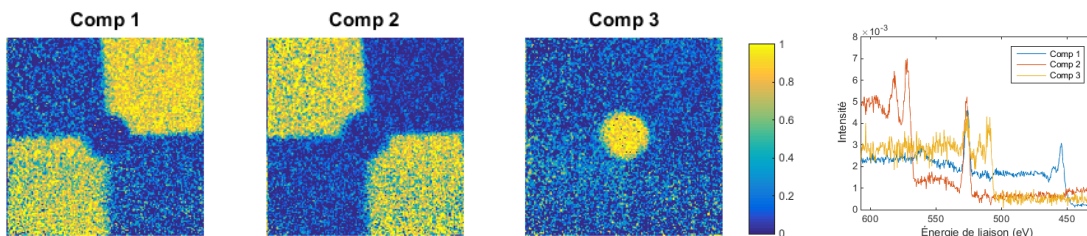


Figure 5.23 Résultats à convergence de MCR-LLM avec φ optimisé à chaque itération avec une acquisition d'une seconde par énergie de liaison

La Figure 5.24 montre quant à elle l'évolution de l'exposant φ à travers les itérations de la même analyse dont les résultats à convergence sont présentés à la Figure 5.23. On peut constater que la valeur de φ est à son maximum aux premières itérations, alors qu'elle diminue par la suite jusqu'à se stabiliser.

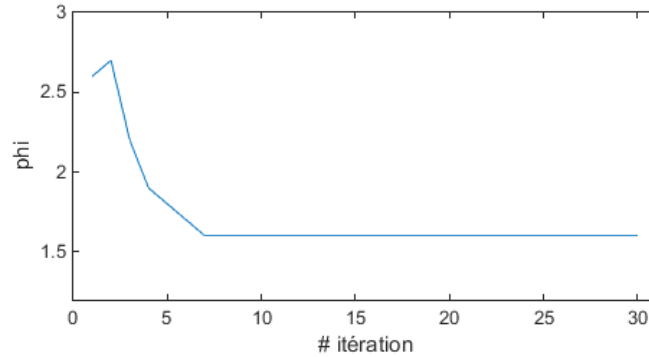


Figure 5.24 Évolution du φ dans l'algorithme de MCR-LLM avec φ optimisé à chaque itération avec une acquisition d'une seconde par énergie de liaison

5.3.5 Indices de performance

Les résultats de l'algorithme *MCR-ALS* standard sont comparés à la méthode proposée (*MCR-LLM*), et ce pour différents temps d'acquisition par énergie de liaison (0.33, 0.5, 1, ...). Dans chacun des essais, le but est d'extraire les spectres et les contributions associées de trois composants. Dans tous les cas, que ce soit avec *MCR-ALS* ou *MCR-LLM*, l'initialisation par l'algorithme des k-moyennes est utilisée.

Deux indices de performance seront étudiés. Le premier indice, I_S , quantifie la capacité de l'algorithme à s'approcher des spectres de référence. Les spectres de référence ont été extraits avec l'ensemble de données avec un temps d'acquisition de 20 secondes par niveau d'énergie. (Figure 5.25) Pour chacun des trois métaux, comme le modèle de contributions est connu, les zones spatiales contenant exclusivement les métaux correspondants sont sélectionnés et à partir de celles-ci, les spectres moyens sont extraits.

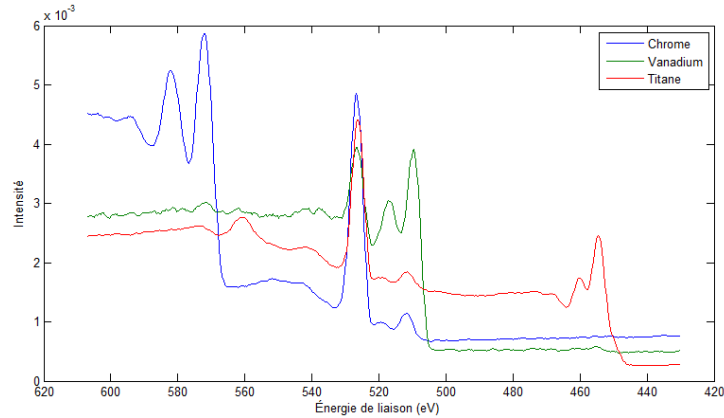


Figure 5.25 Spectres de référence obtenus avec un temps d'éclairement (*dwell time*) de 20 s par niveau d'énergie

L'indice est basé sur une somme d'erreur absolue et se définit par cette équation :

$$I_S = 1 - \frac{\sum_{k=1}^K \sum_{n=1}^N |R_{s_{k,n}} - S_{k,n}|}{2K} \quad (5.1)$$

Dans l'équation (5.1), N correspond au nombre de variables dans l'ensemble de données et K correspond au nombre de composants à extraire. S correspond à la matrice spectrale calculée à la convergence de l'algorithme, alors que RS est la matrice contenant les spectres de référence. L'indice I_S sera donc toujours positif. Ainsi, plus les spectres calculés sont similaires aux spectres de référence, plus l'indice I_S se rapprochera de 1.

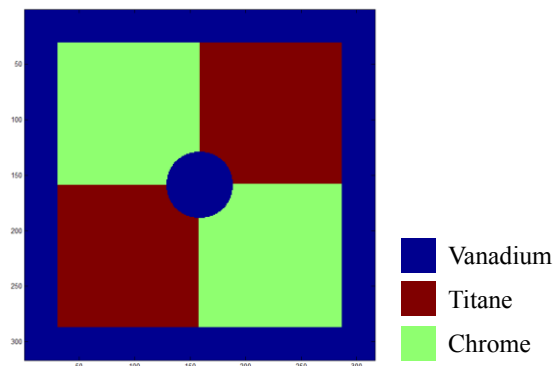


Figure 5.26 Modèle de contributions

Le deuxième indice, I_C est celui relié à la prédiction des contributions et se base sur le modèle de contributions illustré par la Figure 5.26 et qu'on définit par la matrice \mathbf{Rc} . L'indice se calcule de façon similaire à I_S :

$$I_C = 1 - \frac{\sum_{k=1}^K \sum_{m=1}^M |\mathbf{Rc}_{m,k}^* - \mathbf{C}_{m,k}|}{MK} \quad (5.2)$$

Dans l'équation (5.2), M correspond au nombre total de pixels, K au nombre de composants à extraire et \mathbf{C} à la matrice contenant les contributions des M pixels. Comme la zone de l'échantillon analysé se déplace quelque peu d'une acquisition à l'autre, une translation et une rotation du modèle (*Figure 5.26*) sont effectuées afin de maximiser l'indice I_C . Cette matrice transformée est définie comme étant la matrice \mathbf{Rc}^* .

Similairement à I_S , la pire valeur de I_C est de 0, alors qu'une prédiction parfaite donnera une valeur de 1.

5.3.6 Performance entre les différents algorithmes

Les indices de performance I_S , équation (5.1), et I_C , équation (5.2), ont été calculé à partir des résultats à convergence obtenus par les ensembles de données de différents temps d'acquisition avec différents algorithmes. Dans ces différents algorithmes, on retrouve *MCR-ALS* avec φ fixé à 1 (*MCR-ALS* standard) et φ optimisé à chaque itération. Il a été de même pour *MCR-LLM*. À noter que pour toutes les analyses, les mêmes groupes de pixels ont été moyennés entre eux afin d'initialiser les spectres. (*Figure 5.27*)

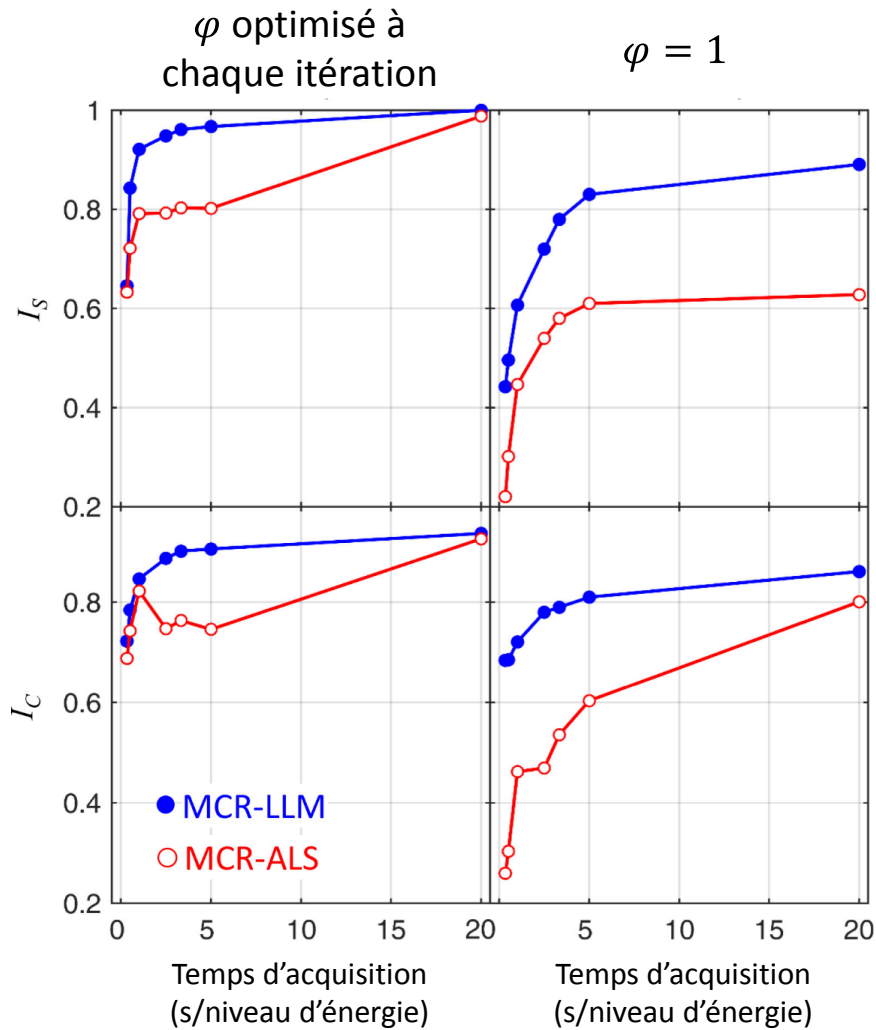


Figure 5.27 Indices de performance calculés à convergence de *MCR-ALS* et *MCR-LLM*, avec $\varphi = 1$ et φ optimisé à chaque itération, avec l'ensemble de données des trois métaux obtenus à différents temps d'acquisition

En analysant la Figure 5.27, on constate que les indices obtenus grâce à *MCR-LLM* sont toujours supérieurs à ceux de *MCR-ALS*. Ceci s'explique par l'incapacité de *MCR-ALS* à isoler correctement les composants à un temps d'acquisition inférieur à 20 secondes par niveau d'énergie. De plus, l'utilisation d'un exposant φ permet d'augmenter la performance de chacun des deux algorithmes.

Pour des temps d'acquisition égaux ou plus longs que 0,5 seconde par niveau d'énergie, *MCR-LLM* génère des spectres très près de ceux de référence, expliquant des valeurs de I_S s'approchant de 1. Malgré le fait que les bons spectres soient calculés, le bruit contenu dans

les résultats réduit la valeur de I_S pour les temps d'acquisition de 0,5 et de 1 seconde par niveau d'énergie. Les résultats sont infructueux pour un temps d'acquisition de 0,33 seconde par niveau d'énergie, ce qui résulte d'une plus basse valeur de I_S .

Au niveau des contributions, une acquisition plus longue que 3,33 secondes par niveau d'énergie devient non nécessaire, par le fait que I_C atteint un plateau avec *MCR-LLM* et φ optimisé.

Dans *MCR-ALS* avec φ optimisé, l'indice I_C a un minimum local à 1 seconde par niveau d'énergie pour ensuite redescendre à des temps d'acquisition supérieurs et pour finalement atteindre la même performance que *MCR-LLM*.

Cet étrange comportement s'explique en comprenant les premières itérations dans l'algorithme. L'initialisation par l'algorithme des k-moyennes créent des valeurs de contributions initiales ségréguées soit à 0, soit à 1. En incluant l'exposant φ dans le calcul de \mathbf{S} à l'effet de coupure dans le calcul de \mathbf{C} , la ségrégation sera renforcée.

Ainsi, la valeur de I_C est ascendante à une acquisition de 0,33 à 1 seconde par niveau d'énergie. À ces temps d'acquisition, les données ont de très basses valeurs et la ségrégation en question est très importante suite à l'initialisation. À des temps d'acquisition au-dessus de 3,33 secondes par niveau d'énergie, grâce à des données de plus haute intensité, l'algorithme est capable de passer outre cette ségrégation initiale et de converger vers la bonne solution. Enfin, entre des temps d'acquisition de 1 et de 3,33 secondes par niveau d'énergie, l'algorithme tergiverse entre converger vers une solution ségréguée ou une solution continue (valeurs continues entre 0 et 1), ce qui mène à un indice I_C plus bas.

5.4 Retour sur le Chapitre 5

Ce chapitre a tout d'abord démontré pourquoi il était possible d'utiliser l'algorithme développé (*MCR-LLM*) avec des données d'imagerie spectrale *XPS*, par le fait que les signaux captés sont bruités par une distribution de Poisson.

Autant avec les données des demi-cercles qu’avec les données des trois métaux, il a été démontré que par analyse *PCA*, malgré des données prétraitées par une normalisation au niveau de chaque observation, il reste difficile d’obtenir de réels composants qui comportent, chacune, un spectre et une cartographie spatiale de contributions.

Dans le cas d’une analyse avec seulement deux composants et avec des données avec ratio signal sur bruit relativement moyen, *MCR-ALS* standard est très efficace.

Ceci n’est pas le cas lorsqu’on veut extraire plus de deux composants d’un ensemble de données, et qu’en plus, les données ont un ratio signal sur bruit relativement faible. C’est ce qui a été démontré avec l’ensemble de données des trois métaux. De cet ensemble de données, il a aussi été démontré qu’en utilisant *MCR-LLM*, un trop grand temps d’acquisition peut s’avérer totalement inutile.

Tableau 5.1 Acronyme utilisés dans le Chapitre 5

| Acronyme | Correspondance en anglais | Correspondance en français |
|----------------|---|--|
| <i>EDX</i> | <i>Energy Dispersive X-Ray</i> | Analyse dispersive à rayons X |
| <i>MCR-ALS</i> | <i>Multivariate Curve Resolution – Alternating Least Squares</i> | Résolution multivariée de courbes – Moindres carrés alternatifs |
| <i>MCR-LLM</i> | <i>Multivariate Curve Resolution – Loglikelihood Maximization</i> | Résolution multivariée de courbes – Maximisation du logvraisemblance |
| <i>EELS</i> | <i>Electron Energy Loss Spectroscopy</i> | Spectroscopie des pertes d’énergie |
| <i>XPS</i> | <i>X-Ray Photoelectronic Spectroscopy</i> | Spectrométrie photoélectronique X |

Tableau 5.2 Variables utilisées dans le Chapitre 5

| Variable | Définition |
|-----------------|---|
| φ | Exposant modifiant les valeurs des contributions dans la régression multilinéaire pour le calcul des spectres |
| K | Nombre de composants à extraire |
| \mathbf{X} | Matrice des données |
| α | Nombre total de composants à calculer dans la méthode préliminaire |
| I_S | Indice de la capacité de l'algorithme à calculer les spectres de référence |
| \mathbf{S} | Matrice spectrale |
| \mathbf{Rs} | Matrice contenant les spectres de référence |
| I_C | Indice de la capacité de l'algorithme à calculer les contributions de référence |
| \mathbf{C} | Matrice des contributions |
| \mathbf{Rc} | Matrice contenant les contributions de référence |
| I_G | Indice de la capacité de l'algorithme à calculer à la fois les contributions et les spectres de référence |

6. APPLICATION DE LA MÉTHODE SUR D'AUTRES TYPES DE DONNÉES

6.1 Analyse dispersive à rayons X

6.1.1 Explications de l'analyse

L'analyse dispersive à rayons X analyse, couramment appelé sous l'acronyme *EDX*, pour *Energy Dispersive X-ray spectroscopy*, permet la caractérisation d'un matériau par l'analyse des rayons X que le matériau émet suite à une excitation par les électrons d'un faisceau électronique dans un microscope électronique en transmission (*MÉT*).

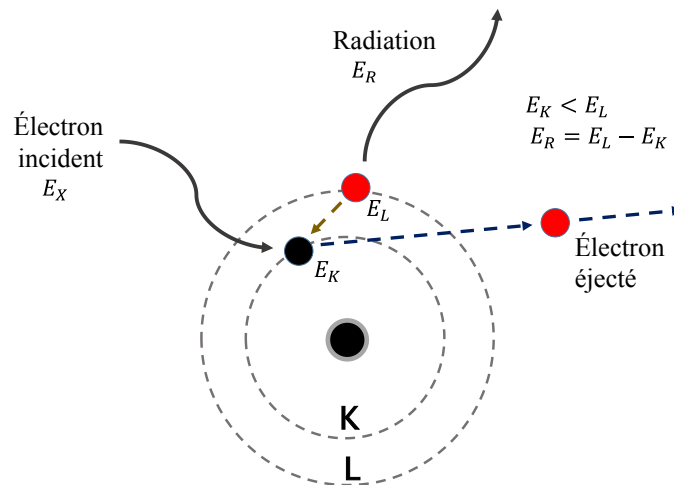


Figure 6.1 Principes physiques de l'EDX

Le processus débute par une ionisation d'un électron de cœur par les électrons incidents. L'atome devient instable et, tout dépendant des probabilités de transition, un électron d'une couche orbitale d'une énergie supérieure transite pour combler la lacune de l'électron de cœur. La différence en énergie est émise sous forme de rayonnement électromagnétique, qui est caractéristique de l'élément en question. (*Figure 6.1*)

L'ensemble de données acquises par analyse dispersive à rayons X (remerciements à Pr. Ray Lapierre pour le partage de ses données) repose sur la caractérisation d'un nanofil dont le cœur est composé d'arséniure de gallium (AsGa) et dont l'enveloppe est composée d'arséniure d'indium (AsIn). La section du nanofil est hexagonale. (Figure 6.2)

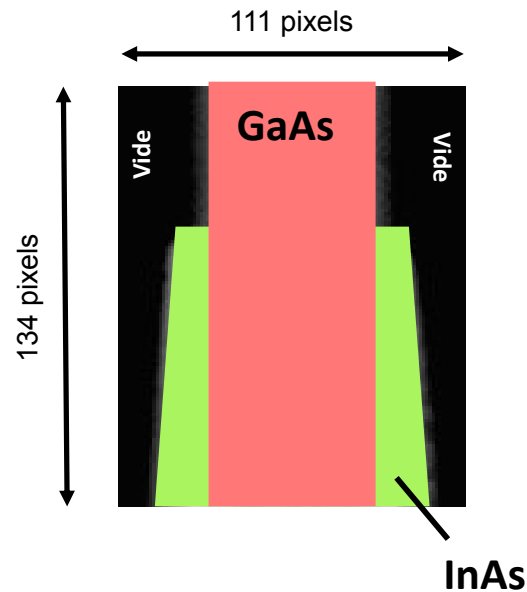


Figure 6.2 Représentation spatiale des sections de l'analyse EDX

La Figure 6.3 montre le modèle calculé des contributions d'InAs et de GaAs. Ce modèle a d'abord été créé par la sommation des comptes pour chaque pixel sur la plage d'énergie caractéristique d'In et la plage d'énergie caractéristique de Ga. Ensuite, par optimisation, en considérant le ratio In/Ga et en considérant la forme hexagonale caractéristique du nanofil, les facteurs de sensibilité de chacun des deux composés, ainsi que l'épaisseur de l'enveloppe pour chaque ligne horizontale de l'image ont été calculés. Ces paramètres connus, il donc possible de calculer le modèle des contributions attendues.

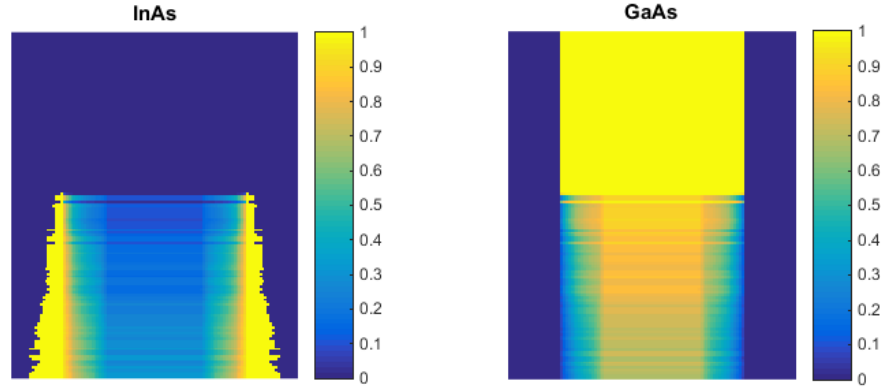


Figure 6.3 Contributions calculées de InAs et de GaAs par intégration par plage d'énergie et par optimisation

Le nombre moyen de photons comptés par pixel est de 52, répartis sur 1024 niveaux d'énergie, ce qui est typique de ce genre de méthode d'analyse. (Figure 6.4)

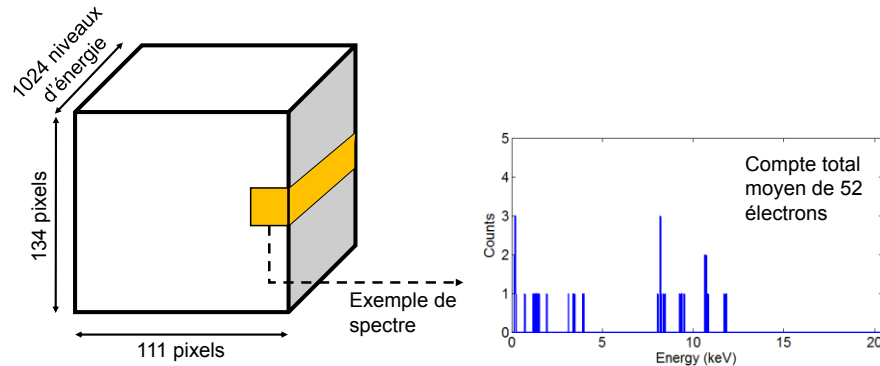


Figure 6.4 Représentation de l'ensemble de données EDX et exemple de spectre

Une très grande partie des comptes est causée par la grille de cuivre sur laquelle repose l'échantillon. En effet, les rayons X provenant du nanofil parviennent à la grille de Cu qui fluoresce également des rayons X caractéristiques du Cu. Ces comptes ne sont donc pas caractéristiques de l'échantillon et ajoutent une difficulté à l'analyse des données. En plus de la faible quantité de photons comptés, la difficulté de cet ensemble de données est le fait que trois zones doivent être extraites, soit le GaAs, le InAs et le vide.

Enfin, on considère le bruit comme poissonnien. La radiation provenant de l'échantillon est bien sûr poissonnienne, alors qu'on considère que le bruit de lecture du capteur n'est

pas significatif sur les résultats. Tel que l'illustre l'exemple de spectre de la Figure 6.4, toutes les données acquises consistent à des nombres entiers.

6.1.2 Comportement de *MCR-ALS* versus *MCR-LLM*

Afin de bien comparer le comportement au fil des itérations entre les deux méthodes, l'initialisation a été la même pour les deux méthodes (algorithme des k-moyennes avec trois composants). La Figure 6.5 illustre l'évolution des contributions et des spectres avec *MCR-ALS* standard. On peut constater que les trois composants ont tendance à se confondre. À la 50^e itération, il est très difficile de distinguer la différence entre les trois spectres, en plus de ne pouvoir rien distinguer dans les images de contribution.

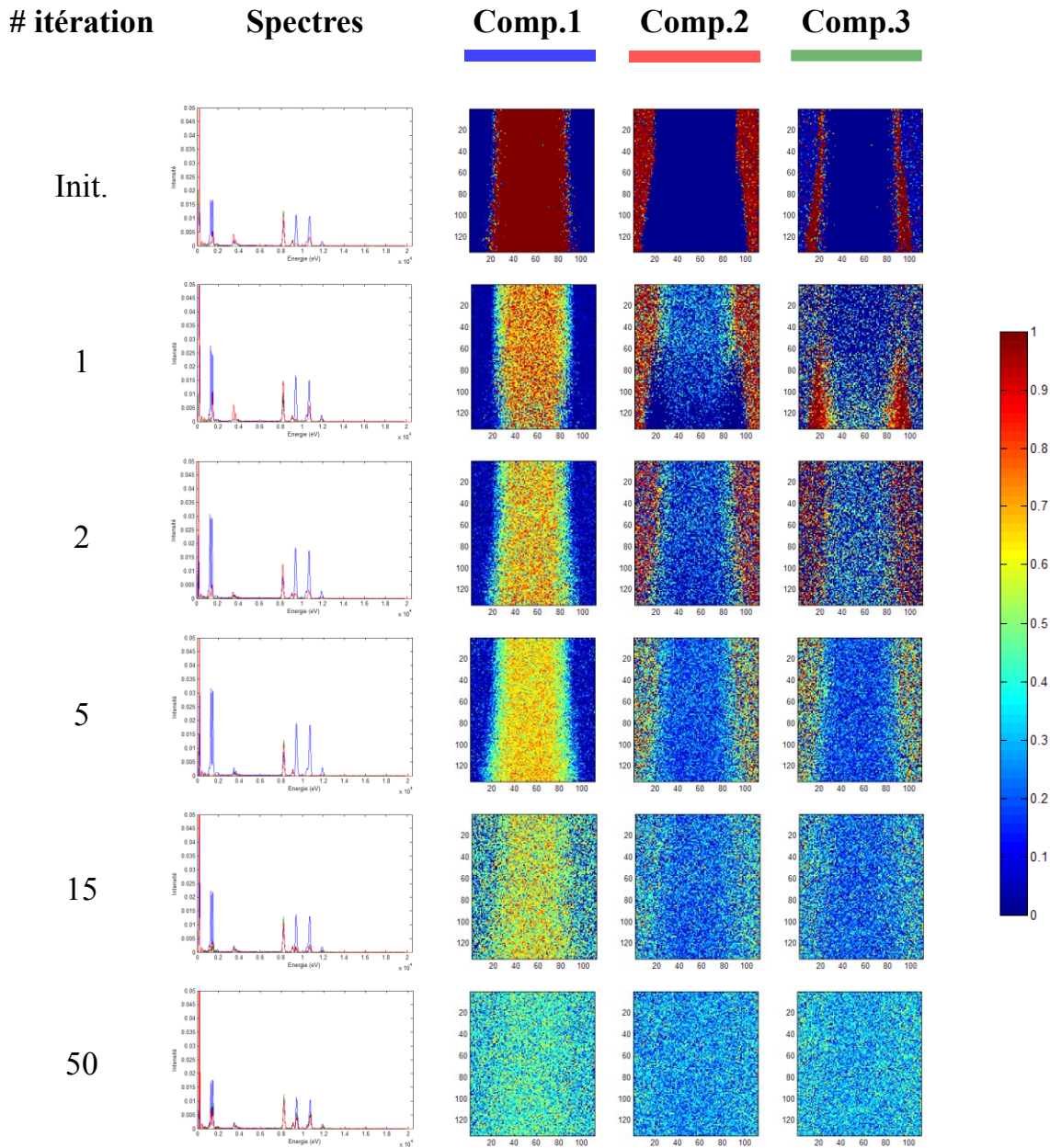


Figure 6.5 Évolution des spectres et des contributions au fil des itérations, avec *MCR-ALS*

La Figure 6.6 illustre l'évolution des spectres et des contributions au fil des itérations avec *MCR-LLM* ($\varphi = 1$).

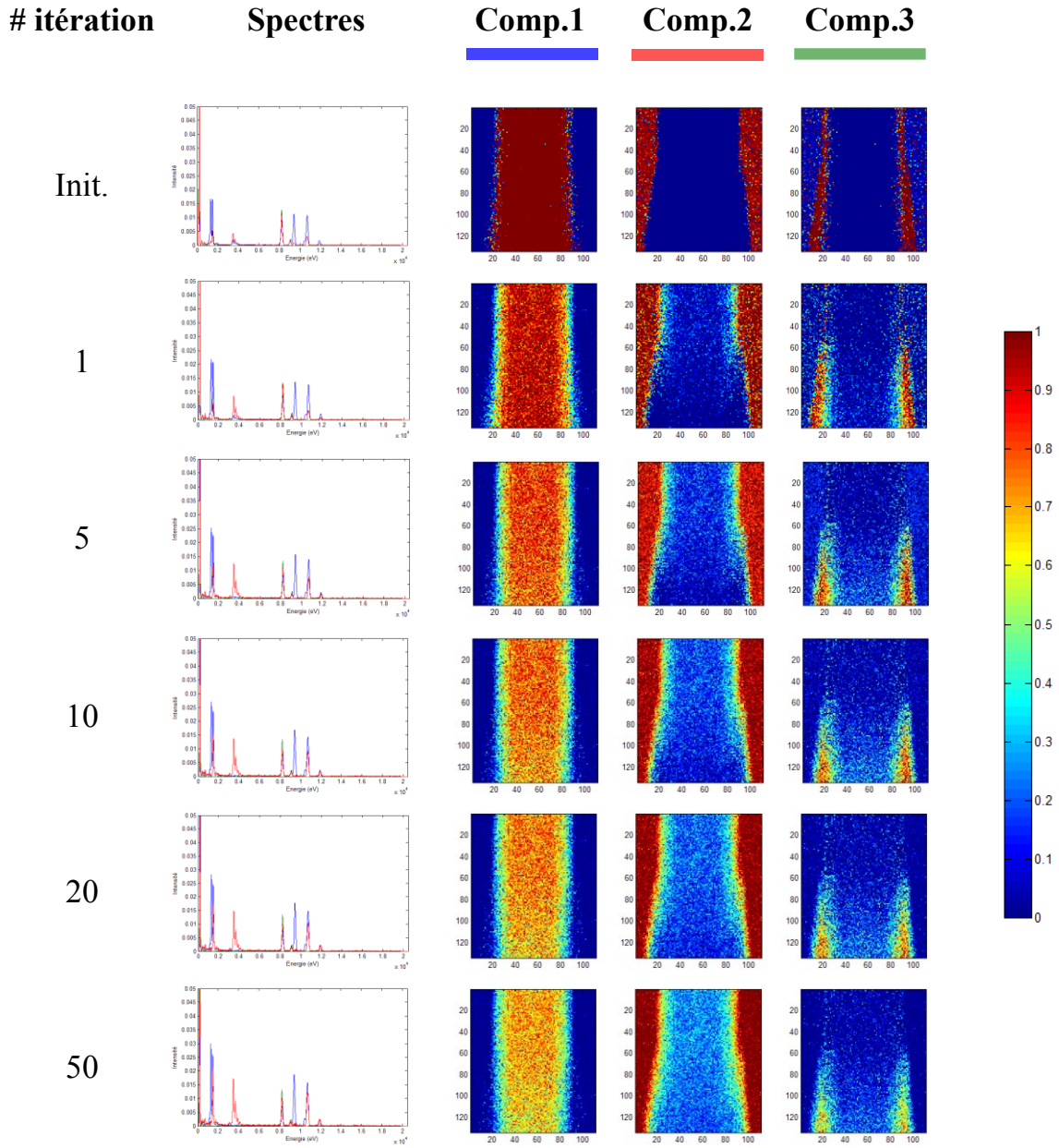


Figure 6.6 Évolution des spectres et des contributions au fil des itérations, avec *MCR-LLM*, $\varphi = 1$

La Figure 6.7 est créée à partir de la dernière ligne de la Figure 6.6, en retirant les contributions associées au composant du vide, soit par l'équation (6.1). Or, les résultats obtenus sont similaires au modèle (*Figure 6.3*).

$$C_{Im,i,j} = \frac{c_{3,l,j}}{c_{1,i,j} + c_{3,i,j}} \quad C_{Ga,i,j} = \frac{c_{1,l,j}}{c_{1,i,j} + c_{3,i,j}} \quad (6.1)$$

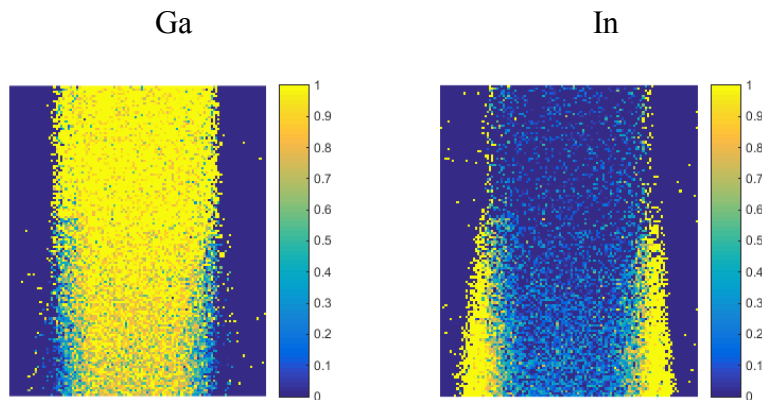


Figure 6.7 Contributions à convergence, avec *MCR-LLM*, $\varphi = 1$, en retirant le composant associé au vide

Si on compare les spectres obtenus à l'initiation et ceux à la 50^e itération de la méthode proposée (*MCR-LLM*), on constate que les spectres de l'initiation semblent représenter un mélange de composants, alors que les spectres de la 50^e itération semblent représenter davantage des composants purs par les pics plus intenses.

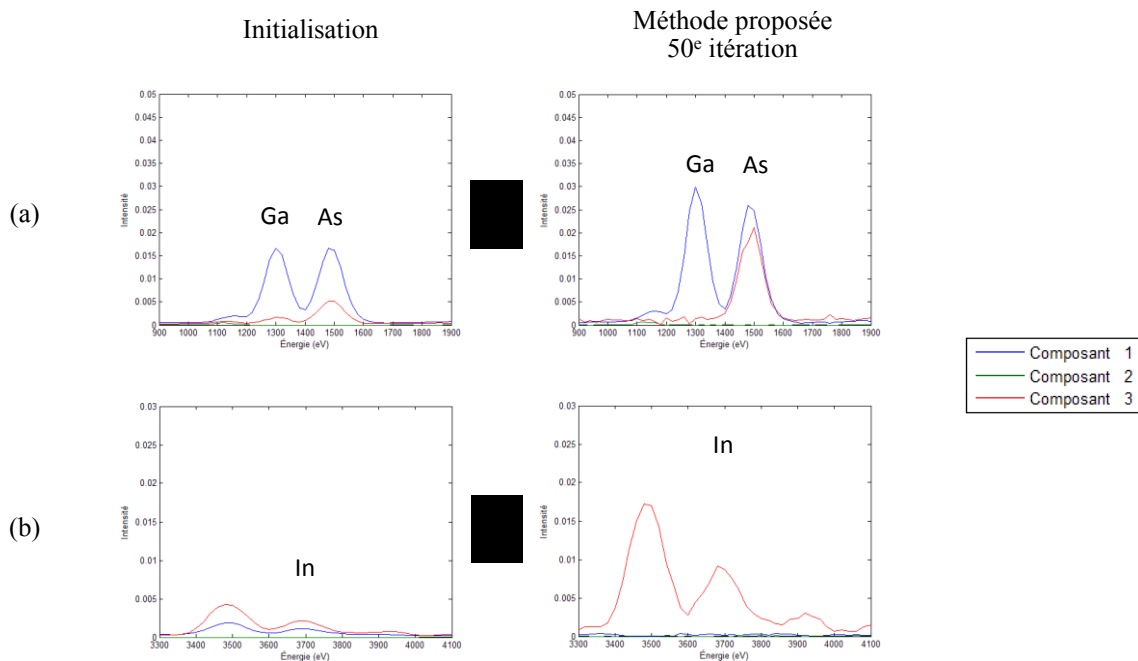


Figure 6.8 Comparaison de sections de spectres entre l'initialisation et la 50^e itération de *MCR-LLM* dans (a) les environs de 1400 eV et dans (b) les environs de 3700 eV

Plus spécifiquement, à l'aide de la Figure 6.8, si on analyse les spectres dans les alentours de 1400 eV (cas *a*), on constate qu'il n'y a qu'un seul pic caractéristique (As) pour le 3^e composant à la 50^e itération, alors que le spectre initial démontre deux pics (Ga et As), de moindre intensité, pour le 1^{er} composant. Similairement, dans les environs de 3700 eV (cas *b*), on constate que, dans les spectres initiaux, le 1^{er} composant possède les mêmes pics, caractéristiques du In, que le 3^e composant, mais à moindre intensité. À la 50^e itération, on constate que ces pics étaient caractéristiques du 3^e composant (InAs), vu la disparition de pics du 1^{er} composant.

Autant au niveau spectral que spatial, les résultats démontrent que la méthode proposée est capable de segmenter les données en trois composants caractéristiques et uniques. *MCR-ALS* standard convergera plutôt vers des composants qui consistent en des mélanges de composants purs.

Les résultats de convergence sont totalement différents entre les deux méthodes. La Figure 6.9 permettra de comprendre pourquoi les résultats diffèrent. On présente ici les résultats de concentrations obtenues lors de la 1^{re} itération pour un pixel en particulier :

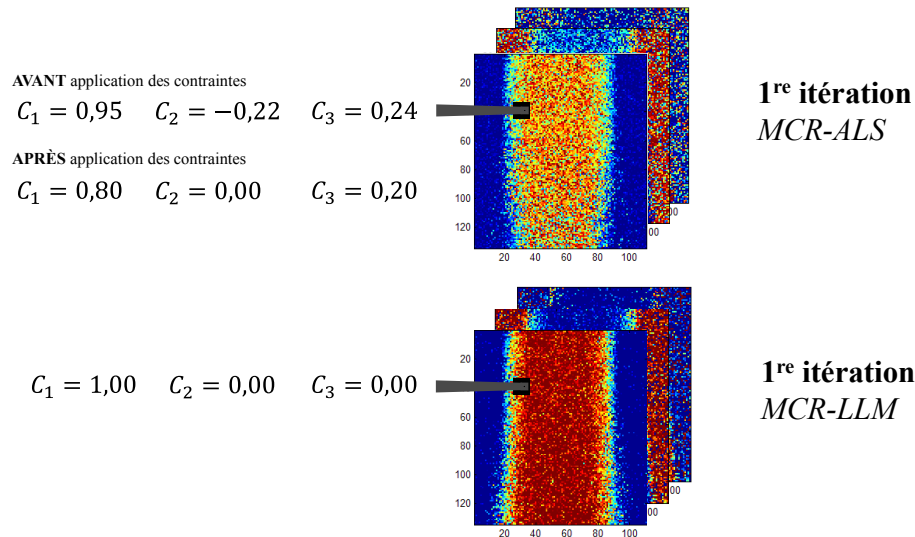


Figure 6.9 Comparaison entre les valeurs de contributions calculées entre les deux méthodes pour un même pixel

Avec *MCR-ALS*, les contributions calculées par régression pour le pixel ciblé à la Figure 6.9 ne respectent pas les contraintes. En modifiant les résultats afin de faire respecter les

contraintes, on constate que les contributions finales calculées diffèrent de beaucoup par rapport aux contributions calculées par la méthode proposée (*MCR-LLM*).

Dans l'exemple de la Figure 6.9, *MCR-LLMS* calcule que seul le 1^{er} composant est en présence sur le pixel, alors qu'après application des contraintes, *MCR-ALS* calcule plutôt une contribution de 80% du 1^{er} composant.

Cet exemple démontre un phénomène très généralisé. Plus les données sont bruitées, plus les contributions calculées s'éloigneront du respect des contraintes. En modifiant les données afin de faire respecter les contraintes, le phénomène démontré par l'exemple de la Figure 6.9 se produira davantage. Les contributions calculées ont tendance à représenter un mélange de composants et non des composants purs.

Ainsi, au fil des itérations, en calculant des contributions représentant des mélanges de composants et en calculant de nouveaux spectres à partir de ces contributions, le processus résultera en des composants qui sont semblables les uns des autres. Ce phénomène est observable à la Figure 6.5, en observant les résultats de la 50^e itération, dans laquelle il n'est plus possible de distinguer spatialement les composants, en plus d'obtenir des spectres très semblables les uns aux autres.

L'utilisation de *MCR-LLM* permet d'éviter ce phénomène. La maximisation de la vraisemblance calcule directement les contributions, tout en respectant les contraintes, et ce, peu importe la qualité des données. Voilà pourquoi l'utilisation de la méthode proposée (*MCR-LLM*) pour analyser l'ensemble de données en question permet une segmentation spatiale et spectrale, contrairement à *MCR-ALS*.

6.1.3 Résultats avec $\varphi > 1$

La Figure 6.6 a démontré que l'algorithme proposé (*MCR-LLM*) a réussi à segmenter l'ensemble de données en trois zones cohérentes avec $\varphi = 1$. La Figure 6.10 compare les contributions obtenues entre $\varphi = 1,0$ et $\varphi = 1,5$, alors que la Figure 6.11 fait la même comparaison, mais au niveau spectral.

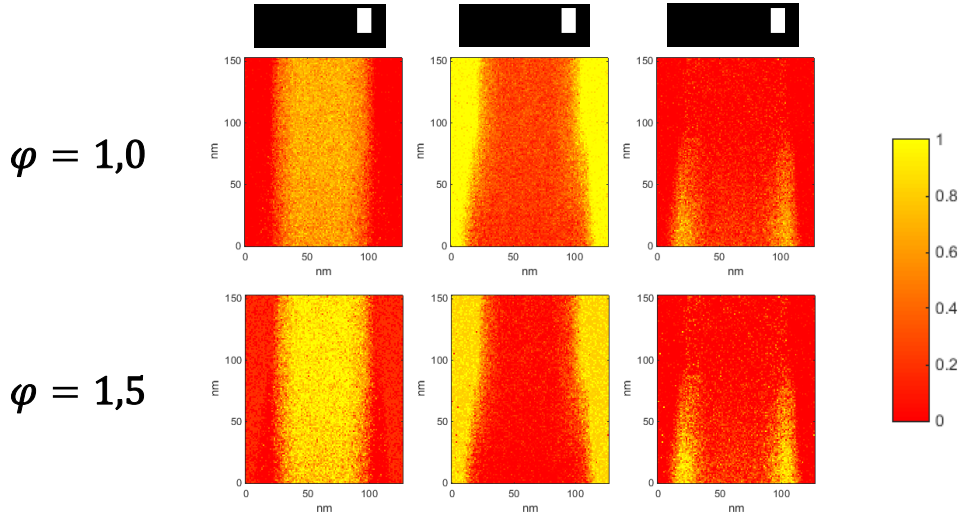


Figure 6.10 Comparaison entre les contributions calculées $MCR-LLM$ $\varphi = 1$, et $MCR-LLM$, $\varphi = 1,5$

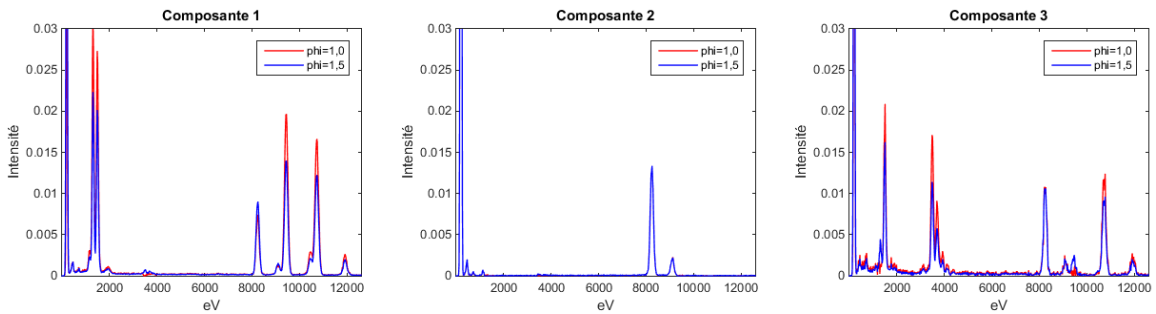


Figure 6.11 Comparaison entre les spectres calculés avec $MCR-LLM$, $\varphi = 1$, et $MCR-LLM$, $\varphi = 1,5$

La Figure 6.10 montre des zones plus intenses et moins bruitées pour les composants 1 et 3 avec $\varphi = 1,5$ par rapport à $\varphi = 1,0$, alors que la zone du composant 2 est plus intense avec $\varphi = 1,0$. Cela signifie que le composant du vide prend plus d'ampleur avec $\varphi = 1,0$. Or, le vide ne comprend aucune information pertinente. Ainsi, si on retire ce composant, par l'équation (6.1), et qu'on ne considère que les contributions des 1^{er} et 3^e composants, les contributions résultantes seront, dans les deux cas, celles présentées par la Figure 6.7.

Cependant, au niveau des spectres (Figure 6.11), dans les composants 1 et 3, les pics caractéristiques de ces deux composants sont plus intenses avec $\varphi = 1,0$. Ce phénomène est relié par le manque d'interaction entre les différents composants, tel qu'expliqué dans les sous-sections 4.3.2 et 4.3.3. À cause du biais causé par la modification des contributions

par l'exposant φ , les spectres calculés ne sont pas totalement purs, c'est-à-dire qu'ils contiennent un peu d'information des autres composants.

Pour remédier à cette situation, on peut réinitialiser l'algorithme à partir des résultats de convergence avec $\varphi = 1,5$, mais cette fois, sans modification des contributions avec un exposant φ (équivalent à $\varphi = 1$). Plus précisément, on initialise avec l'algorithme des k-moyennes et 3 composants ($K = 3$) pour ensuite utiliser l'algorithme *MCR-LLM* avec $\varphi = 1,5$. Une fois ce dernier convergé, on utilise *MCR-ALS* standard avec comme spectres de départ ceux obtenus à convergence de *MCR-LLM*. (Figure 6.12)

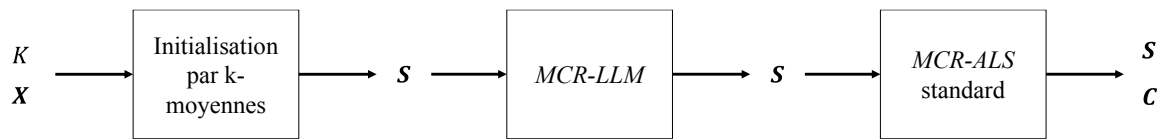


Figure 6.12 Méthodologie de l'utilisation de *MCR-LLM* suivi de *MCR-ALS* standard

Après trois itérations de *MCR-ALS* standard, on atteint de nouveau une convergence. Alors qu'aucune différence notable n'est visible au niveau des contributions, des différences intéressantes au niveau des spectres sont présentes. La Figure 6.13 montre quelques différences au niveau du premier composant :

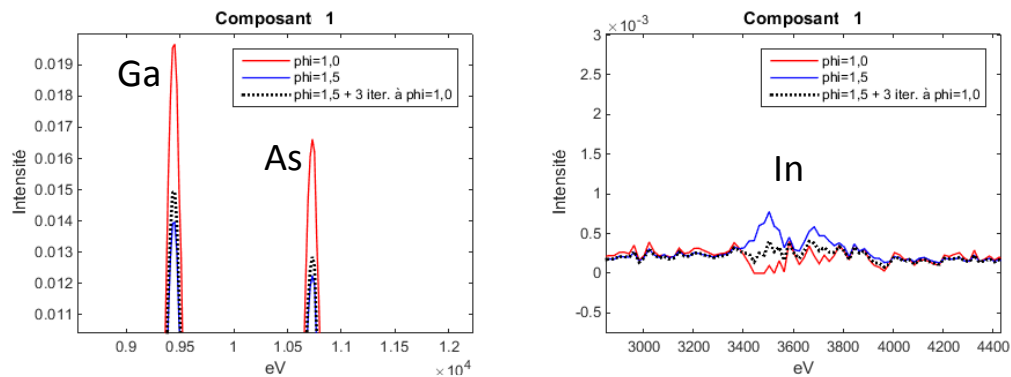


Figure 6.13 Différences ciblées entre les spectres calculés pour le 1^{er} composant

Avec les trois itérations supplémentaires (*lignes pointillées noires de la Figure 6.13*), les pics caractéristiques du premier composant s'intensifient (*graphique de gauche de la Figure 6.13*), alors que les pics non caractéristiques, mais se retrouvant tout de même dans

le 1^{er} composant, diminuent en intensité (*graphique de droite de la Figure 6.13*). Il n'est toutefois pas possible d'atteindre l'intensité des pics qui ont été extraits par l'algorithme sans modification des contributions par un exposant φ (*lignes rouges de la Figure 6.13*).

À noter que bien que la Figure 6.13 ne montre que deux zones du spectre, ceux-ci sont généralisables à tous les composants pour tous les pics.

Cela démontre donc un léger désavantage de l'utilisation de l'exposant φ . Dans le cas de cet ensemble de données, l'analyse a été fructueuse sans utilisation de cet exposant φ , par le fait que, au niveau spectral, le bruit de fond est négligeable et qu'il y a peu de chevauchements de pics. Cependant, dans le cas d'ensembles de données dont lesquelles il y a présence de chevauchement dans les signatures spectrales (*XPS* et *EELS*), la non-utilisation de l'exposant φ ne permettrait pas une analyse fructueuse.

6.1.4 Résultats avec φ variable automatiquement

La sous-section 4.3.3 a fait part d'une méthodologie afin de laisser le paramètre φ libre dans l'algorithme *MCR-LLM*. En appliquant cette méthodologie sur les données en question, les résultats sont très intéressants. La Figure 6.14 montre les valeurs de φ calculées au fil des itérations de *MCR-LLM*.

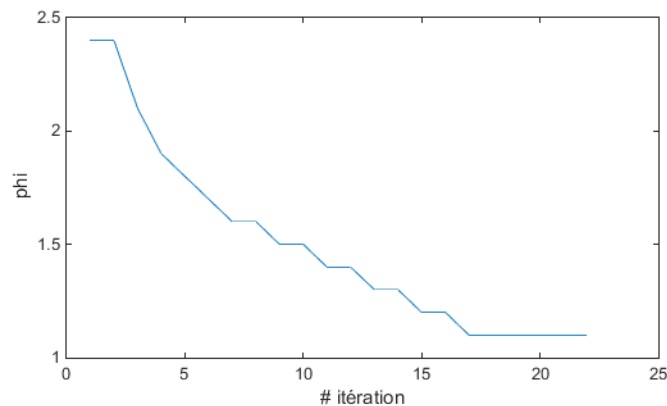


Figure 6.14 φ calculé à travers les itérations

En analysant la Figure 6.14, on constate que la valeur du φ diminue graduellement au fil des itérations. Ce qui est intéressant est que la valeur s'approche de 1 à convergence. Ainsi, automatiquement, on s'assure que les spectres gardent une cohérence physique dans les premières itérations et, près de la convergence, on évite de créer des bords au niveau des spectres, tels que démontrés dans la sous-section 6.1.3.

Les résultats finaux de contribution et des spectres sont donc les mêmes que ceux présentés à la dernière ligne de la Figure 6.6.

6.2 Spectroscopie des pertes d'énergie

L'analyse de données de spectroscopie des pertes d'énergie, ou *Electron Energy Loss Spectroscopy* en anglais (*EELS*), à partir de l'algorithme *MCR-LLM* est un exemple que cet algorithme développé apporte peu par rapport à l'utilisation de *MCR-LLM*, mais sans utilisation de la maximisation de la vraisemblance dans le calcul des contributions.

6.2.1 Explications de l'analyse

La spectroscopie des pertes d'énergie consiste à exposer un matériau à un faisceau d'électrons dont l'énergie cinétique est pratiquement monochromatique. Les électrons ayant traversé l'échantillon peuvent être caractérisés sous trois catégories, selon leur perte d'énergie (*Figure 6.15*) :

- (a) Électrons ayant subi une diffusion élastique avec dispersion : Les électrons ont conservé leur énergie initiale et ont été déviés de leur trajectoire initiale ;
- (b) Électrons ayant subi une diffusion élastique sans dispersion : Les électrons ont conservé leur énergie initiale en plus de conserver leur trajectoire initiale
- (c) Électrons ayant subi une diffusion inélastique avec dispersion : Les électrons n'ont pas conservé leur énergie initiale, ni leur trajectoire initiale

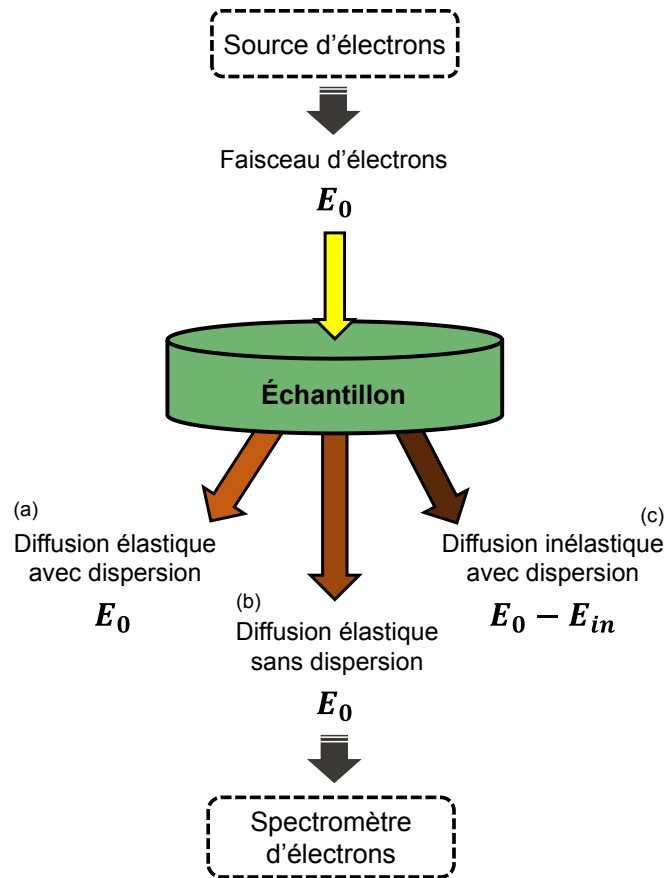


Figure 6.15 Schématisation

La proportion d'électrons subissant une diffusion inélastique, en plus de l'énergie perdue par ces électrons sont dépendants de la composition élémentaire du matériau analysé. Le spectromètre compte le nombre d'électrons à une énergie donnée, ce qui crée une signature spectrale caractéristique de la composition du matériau.

Les spectres *EELS* sont caractérisés par un bruit de fond très important. Les seuils caractéristiques sont souvent minimes par rapport au bruit de fond et sont très difficiles à détecter.

L'ensemble de données étudié (remerciements à Clémence Fauteux-Lefebvre pour le partage de ses données) consiste à imager les clusters de fer contenu dans un nanofilament de carbone de ~ 100 nm de diamètre. Le jeu de données a été acquis en utilisant un microscope électronique en transmission opérant en mode balayage. Le faisceau d'électrons est généré à l'aide d'un canon d'électrons à effet de champs froid (*cold field-*

emission gun) et les électrons ont été accélérés à 300 kV. Au passage du nanofilament, le faisceau d'électrons est analysé à l'aide d'un spectromètre *Gatan Quantum* pour générer un spectre entre 184 eV et 1208 eV sur une zone de 79 X 83 nm, donnant un jeu de données de 243 x 256 x 2048.

La méthode traditionnelle consiste à approximer et soustraire un bruit de fond exponentiel en utilisant le spectre global créé en additionnant les valeurs de tous les pixels pour chaque niveau d'énergie. Ces manipulations servent à repérer sur quelles plages d'énergie se trouvent les pics. [54] Ensuite, pour chaque pixel, le signal est intégré sur une plage d'énergie caractéristique (en l'occurrence, les seuils L23 du Fe, K du C et K de l'O) pour générer des cartes chimiques élémentaires. (*Figure 6.16*)

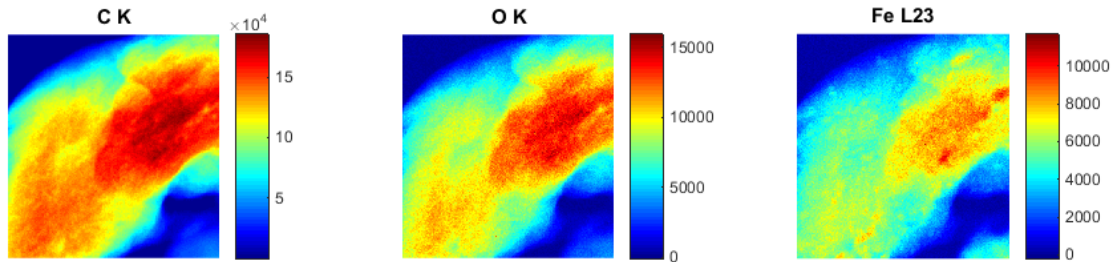


Figure 6.16 Intégration du signal sur des plages d'énergie caractéristiques

Dans ce cas, il est clair que les clusters de Fe sont à peine discernables et sont corrélés spatialement avec l'oxygène, ce qui mène à croire que les clusters sont en fait des oxydes de fer.

Les prochaines sections porteront sur l'application des méthodes développées afin d'augmenter la capacité d'interprétation des différents composants formant le jeu de données EELS.

6.2.2 Résultats

6.2.2.1 Résultats avec trois composants

L'ensemble de données a d'abord été étudié avec trois composants. À noter que la Figure 6.17 montre les résultats obtenus grâce à toutes les modifications incluses dans *MCR-LLM*, la Figure 6.18 montre les résultats obtenus sans utilisation de la maximisation du

logvraisemblance pour le calcul des contributions. Les contributions sont donc calculées comme dans *MCR-ALS* standard, soit par régression multilinéaire. Le 1^{er} composant représente la couche de carbone, le 2^e le vide, alors que le 3^e composant représente l'oxyde de fer.

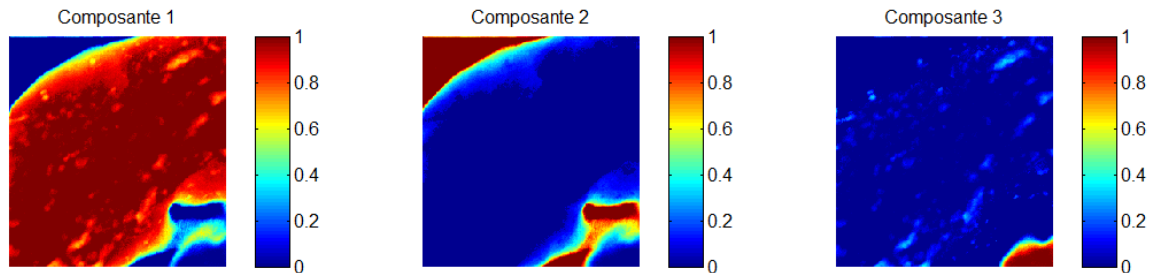


Figure 6.17 Contributions calculée avec les données *EELS*, avec utilisation de toutes les modifications apportées par *MCR-LLM*, avec trois composants et $\varphi = 1,5$

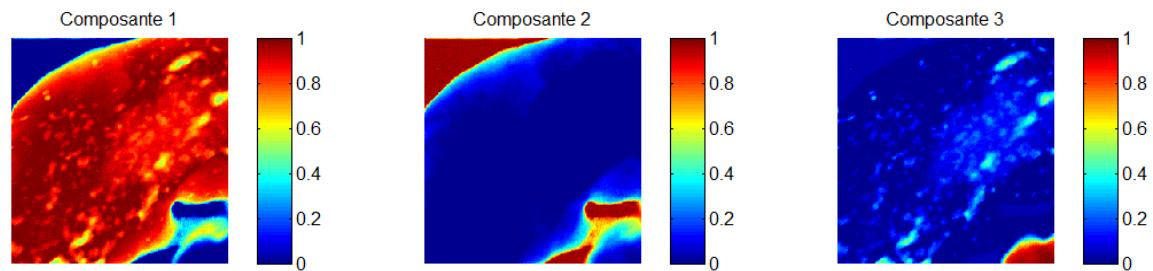


Figure 6.18 Contributions calculée avec les données *EELS*, sans utilisation de la maximisation du logvraisemblance, avec trois composants et $\varphi = 1,5$

Les contributions calculées sont assez similaires entre les deux composants. Il y a tout de même des différences notables. Entre autres, le 1^{er} composant (correspondant au carbone) est beaucoup plus représenté avec les résultats obtenus à partir de l'algorithme *MCR-LLM* (Figure 6.17). Plusieurs pixels qui ont une contribution de 1 dans le 1^{er} composant ont une valeur moindre selon les résultats provenant de l'algorithme sans maximisation (Figure 6.18). En effet, plusieurs pixels se partagent une contribution entre le 1^{er} et le 3^e composant.

Il en est de même pour le deuxième composant. Dans les résultats de l'algorithme *MCR-LLM* (Figure 6.17), des zones sont clairement définies comme faisant parti à 100% dans ce deuxième composant, alors que ces mêmes zones sont quelque peu partagées avec le 3^e

composant en termes de contributions avec l'analyse sans maximisation du logvraisemblance (*Figure 6.18*). En chiffres, la moyenne des contributions calculées dans le coin supérieur gauche de l'image du 2^e composant (zone de 15 X 15 pixels) est de 0,99 avec toutes les modifications apportées par *MCR-LLM*, tandis que cette moyenne se chiffre à 0,95 sans la maximisation du logvraisemblance dans le calcul des contributions.

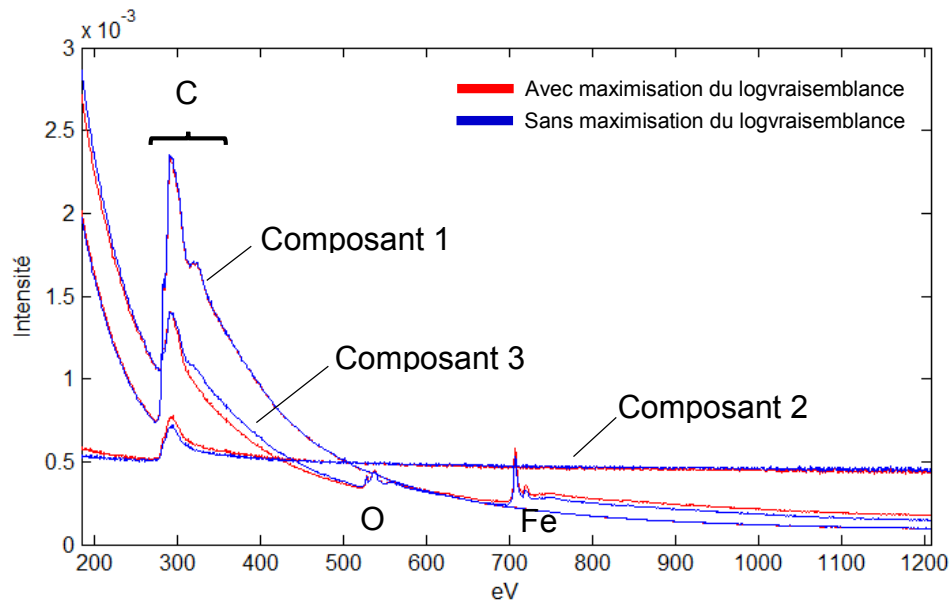


Figure 6.19 Spectres *EELS* calculés pour *MCR-LLM*, mais sans maximisation du logvraisemblance, avec trois composants et $\varphi = 1,5$

Le 2^e composant devrait représenter le vide. Au niveau spectral, cela devrait se traduire par une ligne horizontale (sans aucun pic caractéristique). Cependant, en analysant les spectres obtenus (*Figure 6.19*), peu importe l'algorithme, un pic d'une ampleur considérable se retrouve sur la ligne horizontale. Cela signifie donc que le vide n'est pas parfaitement extrait ou pourrait correspondre au carbone amorphe qui s'est déposé durant l'analyse à l'interface entre le nanofilament et le vide.

De plus, toujours dans l'analyse des spectres calculés (*Figure 6.19*), dans le 3^e composant, un épaulement apparaît sur le seuil du carbone vers 325 eV dans les résultats obtenus à partir de l'algorithme sans maximisation du logvraisemblance, alors qu'il n'apparaît pas dans les résultats de l'algorithme avec maximisation. Ce pic semble plutôt caractéristique du 1^{er} composant et correspond à la convolution du plasmon. Considérant ceci, on peut

considérer que la méthode avec maximisation du logvraisemblance pour calculer les contributions a mieux segmenté les composants qu'avec la méthode utilisant la régression multilinéaire afin de calculer les contributions.

6.2.2.2 Quatre composants

Autant *MCR-ALS* que *MCR-LLM* généreront des résultats totalement différents dépendamment du nombre de composants K à extraire. Ainsi, les prochaines lignes traiteront des résultats obtenus avec quatre composants.

Le deuxième composant issu de l'application de l'algorithme avec trois composants a démontré que l'utilisation de trois composants pour analyser l'ensemble de données en question n'est pas suffisante. Les deux figures ci-dessous présentent donc les cartographies de contributions résultant d'analyse à quatre composants :

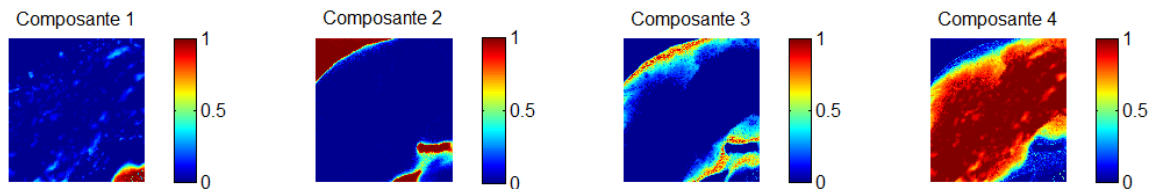


Figure 6.20 Contributions calculée avec les données *EELS*, avec utilisation de toutes les modifications apportées par *MCR-LLM*, avec quatre composants et $\varphi = 1,5$

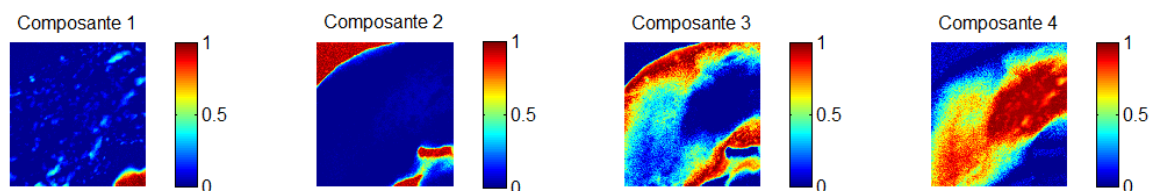


Figure 6.21 Contributions calculée avec les données *EELS*, sans utilisation de la maximisation du logvraisemblance, avec quatre composants et $\varphi = 1,5$

Comme c'était le cas dans l'analyse avec trois composants, l'analyse des contributions obtenues à partir de l'algorithme comprenant la maximisation du logvraisemblance (*Figure 6.20*) démontre qu'il y a beaucoup plus de zones dans laquelle les valeurs de contributions s'approchent de 1 que dans les résultats de contributions obtenus à partir de l'algorithme sans maximisation du logvraisemblance. (*Figure 6.21*)

De plus, toujours dans l'analyse des contributions, il y a une très bonne différence dans les 3^e et 4^e composants entre les deux méthodes. Dans l'algorithme comprenant la maximisation du logvraisemblance (Figure 6.20), le 4^e composant comprend beaucoup de pixels, alors qu'une bonne partie de ces pixels se retrouvent dans le 3^e composant dans l'algorithme sans maximisation du logvraisemblance (Figure 6.21). L'analyse des spectres peut expliquer cette différence. (Figure 6.22)

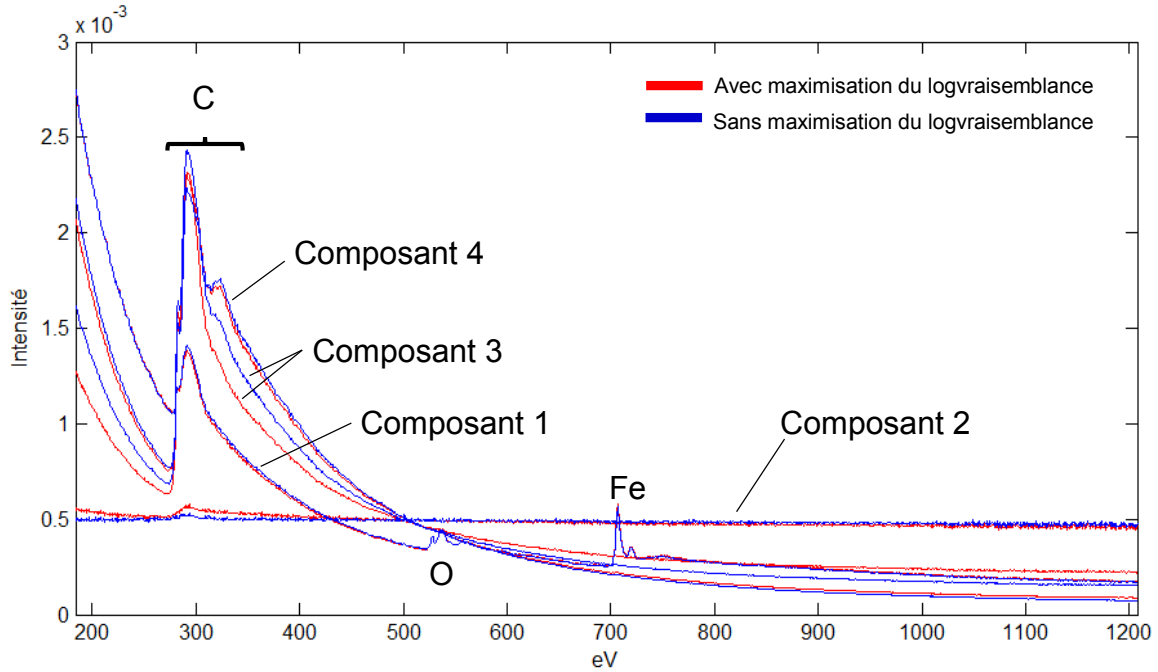


Figure 6.22 Spectres calculés pour *MCR-LLM* et sans maximisation du logvraisemblance, avec trois composants et $\varphi = 1,5$

Le spectre du 3^e composant provenant de l'algorithme sans maximisation contient un petit pic à 325 eV (qui correspond au plasmon), alors que celui-ci est presque absent avec maximisation. Cela explique donc les différences perçues dans les résultats de contributions. Comme un des spectres a une absence de pic, on peut affirmer que l'algorithme avec maximisation du logvraisemblance a été en mesure d'isoler deux signatures spectrales différentes.

Le 3^e composant de l'analyse avec trois composants (Figure 6.19) démontre un spectre avec les seuils de C K (282 eV), O K (532 eV) et du Fe L23 (710 eV). L'analyse avec quatre composants (Figure 6.22) a aussi extrait un spectre avec des pics aux mêmes

positions, mais a aussi extrait un spectre avec un seul pic à 282 eV, qui représente les zones minces du carbone. C'est ce dernier spectre qu'on peut considérer comme s'étant ajouté entre l'analyse à trois composants et celle à quatre composants. Comme le spectre a une signature spectrale unique avec un seul pic caractéristique, on peut considérer qu'il y a bel et bien nécessité d'ajouter un quatrième composant dans l'analyse de l'ensemble de données. On pourrait même affirmer qu'il y aurait une infinité de composants vu que les spectres ne varient pas linéairement selon l'épaisseur du carbone.

Enfin, le spectre du vide comporte toujours un petit pic vers 300 eV (*Figure 6.22*), mais celui-ci est de beaucoup moindre ampleur comparé à ce qui a été extrait avec l'analyse à trois composants (*Figure 6.19*).

Les résultats présentés montrent que la segmentation est légèrement supérieure en calculant les contributions par optimisation du logvraisemblance. Cependant, comme un ensemble de données *EELS* a une grande quantité de variables, une analyse peut donc prendre énormément de temps. On peut compter ce temps en termes d'heures.

6.3 Retour sur le Chapitre 6

Le Chapitre 6 des applications de l'algorithme à des données provenant d'un autre instrument que le *XPS*.

Avec les données provenant d'une analyse dispersive à rayons X (*EDX*), le bruit était considéré comme poissonien. Or, dans ce cas, les données étaient extrêmement bruitées. La comparaison entre le comportement de *MCR-ALS* et la méthode proposée était donc intéressante à étudier dans ce cas. Il a été démontré que *MCR-ALS* était inadéquat à utiliser avec des ensembles de données comportant les mêmes caractéristiques, alors que la méthode proposée (*MCR-LLM*) permet une bonne segmentation spatiale et spectrale des données.

Avec les données provenant de spectroscopie des pertes d'énergie (*EELS*), il a été démontré que vu que les données sont caractérisées par un grand ratio signal sur bruit, l'utilisation de la maximisation du logvraisemblance pour le calcul des contributions peut s'avérer plus ou moins utile. Le grand désavantage à utiliser la maximisation du logvraisemblance est le long temps de calcul à chaque itération, surtout dans ce type d'ensemble de données aux dimensions considérables. Or, les résultats obtenus avec toutes les modifications apportées par *MCR-LLM* sont tout de même supérieures qu'avec seulement la modification au niveau du calcul spectral.

Tableau 6.1 Acronyme utilisés dans le Chapitre 6

| Acronyme | Correspondance en anglais | Correspondance en français |
|----------------|---|--|
| <i>EDX</i> | <i>Energy Dispersive X-Ray</i> | Analyse dispersive à rayons X |
| <i>MCR-ALS</i> | <i>Multivariate Curve Resolution – Alternating Least Squares</i> | Résolution multivariée de courbes – Moindres carrés alternatifs |
| <i>MCR-LLM</i> | <i>Multivariate Curve Resolution – Loglikelihood Maximization</i> | Résolution multivariée de courbes – Maximisation du logvraisemblance |
| <i>EELS</i> | <i>Electron Energy Loss Spectroscopy</i> | Spectroscopie des pertes d'énergie |
| <i>XPS</i> | <i>X-Ray Photoelectronic Spectroscopy</i> | Spectrométrie photoélectronique X |

Tableau 6.2 Variables utilisées dans le Chapitre 6

| Variable | Définition |
|-------------|---|
| φ | Exposant modifiant les valeurs des contributions dans la régression multilinéaire pour le calcul des spectres |
| $C_{k,i,j}$ | Contribution calculée pour le composant k à la position spatiale i et j |

7. CONCLUSION

7.1 Retour sur les objectifs

L'analyse de données spectrales est très complexe et est en soit une spécialité. Si on ajoute une troisième dimension à un ensemble de données provenant d'un microscope, on parle alors de données d'imagerie multispectrale. Déjà qu'un spectre peut être très difficile à analyser, un ensemble de données contenant une multitude de spectres avec de faibles ratios signal sur bruit sème un défi considérable. Il faut aussi considérer que la plupart des techniques d'imagerie spectrale requiert un très long temps d'acquisition. Un temps d'exposition plus long est donc nécessaire afin d'obtenir un ratio signal sur bruit suffisant pour chaque pixel, ce qui n'est pas toujours possible, surtout lorsque l'échantillon est sensible à l'irradiation de la sonde d'analyse.

Pour contrer ce défi, et pour réduire les temps d'acquisition et d'exposition, différentes méthodes ont été développées ces dernières années. Ces techniques ont rendu analysables certains ensembles de données provenant d'instruments de spectroscopie qui n'auraient, sans les méthodes développées, jamais pu être analysés de façon adéquate.

L'objectif global de cette maîtrise était donc d'améliorer les méthodes d'analyse de données d'imagerie multispectrale existantes. Les sous-objectifs se dressaient sous différents points :

- Caractériser les ensembles de données d'imagerie multispectrale générés par divers instruments de spectrométrie
- Développer une méthodologie afin de traiter les données recueillies en tenant compte de la caractéristique du bruit
- Comparer la méthode développée aux méthodologies déjà existantes et généralement utilisées

Plus spécifiquement, les premiers tests et les développements de méthodes ont été développés à partir d'ensembles de données provenant d'image spectrale XPS (*X-Ray Photoelectronic Spectroscopy*). Le Chapitre 2 a donc traité des principales caractéristiques des détecteurs inclus dans le XPS et expliquant le comportement du bruit. Ce chapitre permettait donc de répondre au 1^{er} sous-objectif.

Ensuite, le Chapitre 3 a fait part des techniques de traitements de données et d'algorithmes répandues et sur lesquelles la nouvelle méthode présentée dans cette maîtrise a été basé et qui sont présentés de fond en comble dans le Chapitre 4. Sommairement, l'innovation de cette méthode repose sur le calcul des contributions par maximisation de la vraisemblance statistique. Ces deux chapitres permettaient donc de répondre au 2^e sous-objectif.

Le Chapitre 5 a quant à lui comparé les résultats obtenus par la méthode proposée aux résultats obtenus à partir de méthodes déjà existantes, et ce à partir de données XPS.

Dans le but de démontrer la versatilité de la méthode, le Chapitre 6 a aussi comparé et montré la supériorité des résultats obtenus par la méthode, comparativement à des résultats obtenus de méthodes plus traditionnelles. Ces deux derniers chapitres répondaient donc au 3^e et dernier sous-objectif. Certains bémols sont cependant à tenir compte dans *MCR-LLM*, soit le temps de calcul qui peut être faramineux et l'ambiguïté au niveau des spectres finaux.

L'accomplissement des trois sous-objectifs à travers les Chapitres 2 à 6 permettaient donc d'affirmer que l'objectif global, demandant d'améliorer les méthodes d'analyse de données multispectrales existantes, est accompli.

Les Chapitres 5 et 6 ont présenté des analyses provenant de différents instruments de spectroscopie. Le Tableau 7.1 présente les caractéristiques des données et la nécessité ou non d'utiliser les deux principales modifications proposées à l'algorithme *MCR-ALS* (*sous-section 4.2*). Une case rouge présente un inconvénient lié à l'instrument, alors qu'une case verte présente un avantage.

Tableau 7.1 Caractéristiques des données

| Instrument | Caractéristiques | | Modifications | |
|------------|------------------------|---|---|--|
| | Ratio signal sur bruit | Bruit de fond et chevauchement des pics | Maximisation du logvraisemblance (Calcul de C) | Modification des contributions par l'exposant φ (Calcul de S) |
| XPS | Faible | Moyen | ✓ | ✓ |
| EDX | Faible | Faible | ✓ | ~ |
| EELS | Élevé | Élevé | ~ | ✓ |

Les caractéristiques de spectres provenant de n'importe quel microscope étant grandement documentés, si le ratio signal sur bruit est faible, il est grandement conseillé d'utiliser la maximisation du logvraisemblance dans le calcul des contributions. De plus, si le signal pic sur bruit de fond est faible ou qu'il y a une grande présence de chevauchement entre les pics, il est conseillé d'utiliser la modification des contributions par un exposant φ dans le calcul des spectres. Dans tous les cas, l'utilisation de toutes les modifications, même si cela n'est pas nécessaire, donnera de bons résultats.

7.2 Contributions originales

La principale contribution à la communauté scientifique dans ce projet de maîtrise est la démonstration qu'on peut clairement tirer avantage de la connaissance du comportement du bruit dans l'analyse de données multispectrales, ce qui n'a jamais été fait auparavant.

De plus, l'utilisation de l'algorithme *MCR-ALS* n'a été très que très peu utilisé dans l'analyse de données multispectrales. Les travaux de cette maîtrise ont démontré que le concept de *MCR-ALS* peut être utilisé dans l'analyse de données provenant de différents instruments de spectroscopie, mais en posant certaines hypothèses qui ont été listées dans

ce document. Dans le cas de données d'imagerie spectrale *XPS*, il a été démontré que ces hypothèses créent un biais négligeable.

Dans le cas du *XPS*, les analyses d'ensembles de données d'imagerie spectrale étaient essentiellement effectuées par analyse par composantes principales, ce qui ne fournit pas directement les spectres et les contributions. Plus globalement, la méthode la plus répandue qu'est l'analyse élémentaire par intégration sur une plage d'énergie nécessite des interventions humaines et ne permet pas l'extraction de réels composants, contrairement à *MCR-LLM*.

Enfin, afin de faire rayonner ces découvertes et de faciliter l'accès à ceux qui le désirent, un logiciel gratuit a été développé. Ce logiciel, en constante évolution, permet d'importer plusieurs types de fichiers provenant de différents instruments de spectroscopie, ce que peu de logiciels permettent de faire. Bien sûr, toutes les méthodes développées et présentées dans cette maîtrise ont été implémentées dans le logiciel et sont faciles d'utilisation.

7.3 Travaux futurs

7.3.1 Linéarité des spectres

Pour utiliser *MCR-ALS*, tel que mentionné à plusieurs reprises dans le texte, il faut considérer que le spectre résultant de différentes proportions de composants purs peut se calculer de façon linéaire. Dans le cas de spectres *XPS*, cette considération n'amène pas beaucoup de biais, mais dans les échantillons analysés par *EELS*, ceci était quelque peu problématique.

Cependant, dans le cas d'autres instruments, tels l'*EDX*, l'effet de matrice a une grande importance. Un biais peut donc être causé en calculant un spectre résultant de différents composants purs de façon linéaire dépendamment des proportions que prend chacune des composants purs.

Ainsi, dans le but d'améliorer la méthode proposée dans ce mémoire, certains moyens pourraient être envisagés dans le but d'intégrer les effets de matrice. Pour cela, il faut tout d'abord lister, pour chacun des instruments, les équations à utiliser pour approximer l'effet de matrice aux différents niveaux d'énergie. Ensuite, ces équations doivent être incluses à même l'optimisation pour calculer les contributions pour chacun des pixels.

Intégrer l'effet de matrice dans la méthode proposée permettrait d'augmenter la robustesse de l'algorithme, en plus d'améliorer la qualité des résultats.

7.3.2 Rapidité de l'algorithme

Tel que mentionné dans le texte, le plus gros désavantage de cet algorithme est sa vitesse de calcul. Deux points seraient à considérer pour augmenter la vitesse de cet algorithme.

Premièrement, plusieurs modifications pourraient être effectuées à même le code de l'optimisation *Active Set* afin d'accélérer le processus. À certaines étapes du code, il pourrait y avoir certains calculs qui pourraient diminuer le nombre d'itérations requis. Cependant, toute modification augmente les chances que l'algorithme ne converge pas vers la bonne solution. Ainsi, il devrait y avoir une batterie de tests pour toute modification proposée.

Deuxièmement, tout le développement de la méthode, ainsi que le logiciel rendu disponible à la communauté, a été basé et programmé sur Matlab[®]. Ce logiciel est loin d'être la solution la plus efficace en termes de rapidité de calcul. Une conversion vers un autre langage, tel que C++, pourrait être considérée.

Une option qui pourrait être aussi intéressante serait de convertir le code en Python et de l'installer sur un serveur internet sur lequel Python est installé. Une application de type nuage informatique pourrait ainsi être développée pour l'analyse de données multispectrales.

ANNEXE A PRÉSENTATION DU LOGICIEL DÉVELOPPÉ

Le logiciel développé consiste en une interface graphique basée sur Matlab. Le logiciel se compose en trois sections importantes.

AA.1 Importation de données

L'importation des données s'effectue par le biais de cette fenêtre :

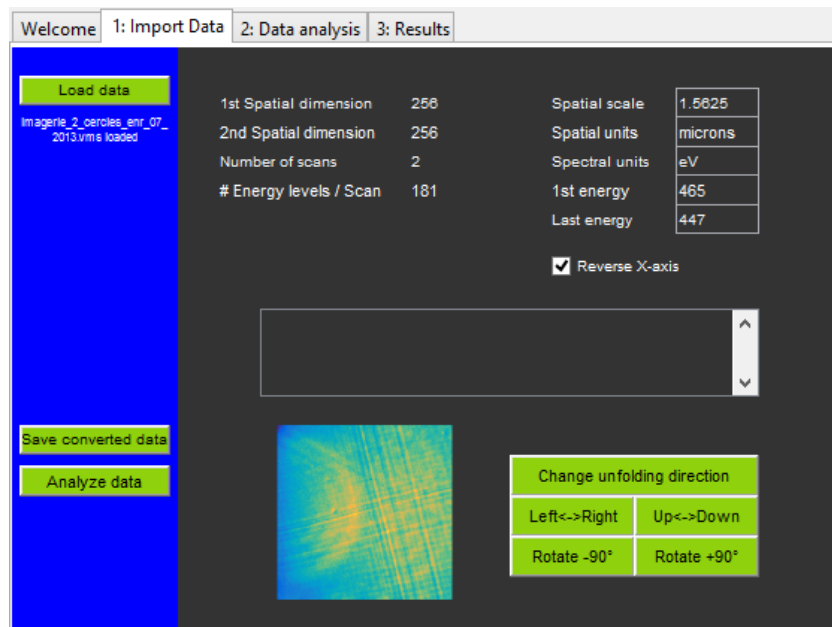


Figure AA.1 Fenêtre d'importation de données dans le logiciel

Dans cette dernière fenêtre, l'utilisateur doit tout d'abord sélectionner un fichier à convertir. Les types de fichiers pris en charge augmenteront au fil du temps. Pour la première version du logiciel, les fichiers ayant l'extension *.vms*, associés à des données

provenant du *XPS*, sont pris en charge ainsi que les fichiers ayant l'extension *.dm3* et *.dm4*, associés au logiciel *Digital Micrograph*.

Une fois le fichier chargé et converti à l'intérieur du logiciel, l'utilisateur peut voir les informations qui ont été extraites. De ces informations, on retrouve toutes les informations spatiales et spectrales. L'utilisateur peut aussi constater les dimensions de l'ensemble de données. Enfin, il peut aussi constater si les données sont spatialement cohérentes. Si cela n'est pas le cas, l'utilisateur a en sa possession tous les outils nécessaires pour retrouver une image cohérente.

Une fois les données chargées et les informations confirmées, l'utilisateur peut facilement transférer les données pour les analyser. L'utilisateur a aussi la possibilité de sauvegarder les données converties dans le format du logiciel.

AA.2 Analyse

La fenêtre d'analyse permet de paramétrer l'algorithme. (*Figure AA.2*)

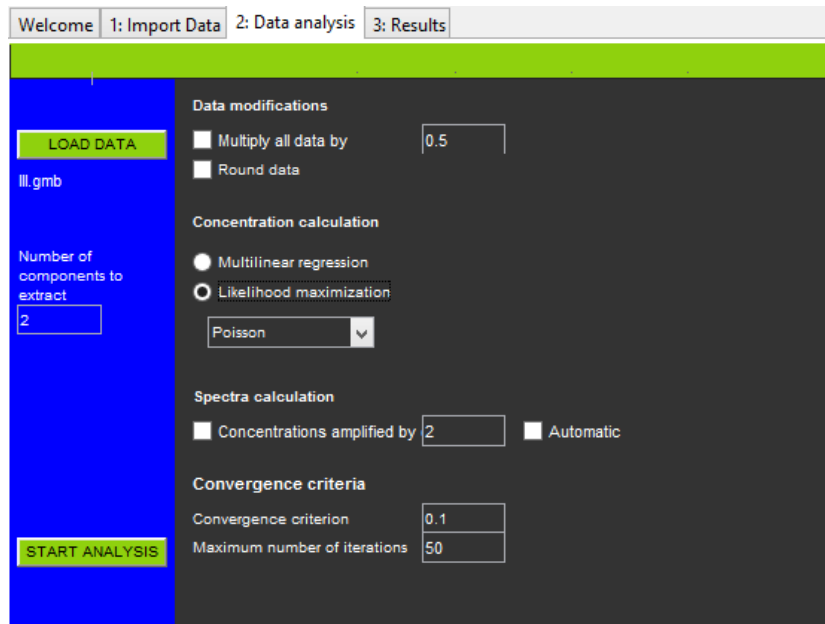


Figure AA.2 Fenêtre d'analyse

Si des données ont précédemment été converties, l'utilisateur n'a pas besoin de charger de nouveau des données. Sinon, il est invité à charger les données qui ont un format spécifique pour le logiciel.

L'utilisateur peut paramétrer tout l'algorithme, que ce soit pour la modification des données, pour l'initialisation des spectres ou le critère de convergence du processus itératif. Le nombre de composants à extraire doit être inscrit dans la barre verticale bleue de gauche.

Une fois les données chargées et le tout paramétré, l'analyse peut être lancée. Pendant le processus itératif, l'utilisateur peut voir les derniers spectres et les dernières cartes de contributions calculées. L'utilisateur peut aussi demander à tout moment l'arrêt des calculs et les reprendre s'il le désire.

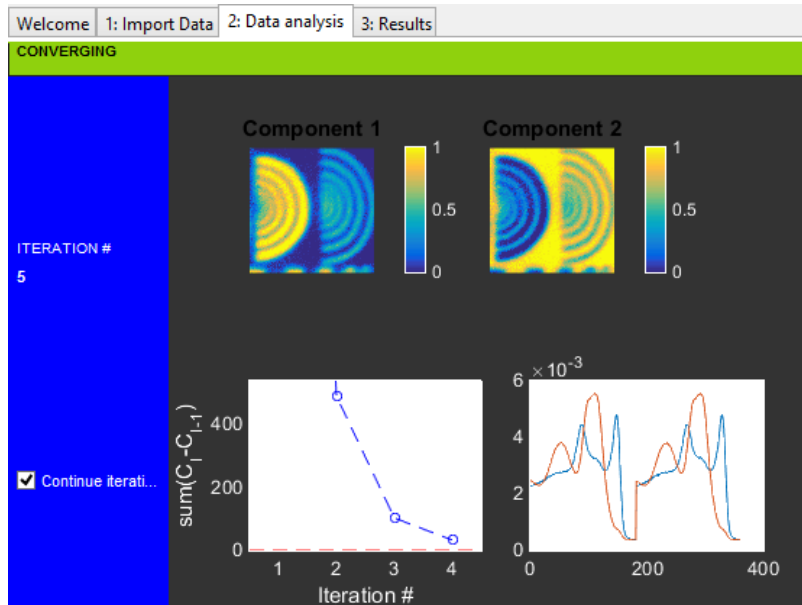


Figure AA.3 Information apparaissant pendant les calculs

AA.3 Visualisation des résultats

La fenêtre de visualisation permet d'analyser autant les spectres et les cartes de contributions extraits. (Figure AA.4)

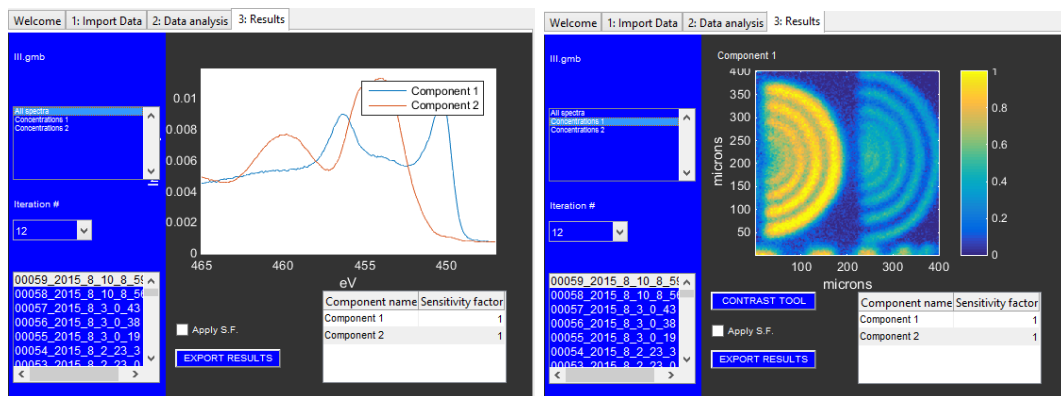


Figure AA.4 Fenêtre de visualisation des résultats

Grâce au tableau du bas, l'utilisateur peut facilement nommer les différents composés calculés et imposer un facteur de sensibilité au niveau des contributions. Grâce au menu du haut à gauche, l'utilisateur peut transiter entre l'analyse des spectres et l'analyse des différentes cartes de contribution. Au niveau des contributions, l'utilisateur a accès à un outil de contraste d'image pour faire ressortir de petits éléments.

Tous les résultats d'analyse sont archivés et peuvent facilement être consultés grâce au menu du bas dans la barre de gauche. De plus, l'utilisateur peut exporter les résultats d'analyse sous format *Matlab* ou sous format *Excel*.

ANNEXE B PSEUDOCODES

AB.1 Analyse en composantes principales (PCA)

Avec l'algorithme *Moindres carrés partiels itératifs non linéaires – Non-linear partial least squares (NIPALS)*

q : Nombre de composantes à extraire

p_h : Vecteur des « loadings » de la composante h

t_h : Vecteur des « scores » de la composante h

X : Matrice des données

X_h : Matrice des données à $h=0$ | Matrice des résidus à $h>0$

s : Critère de convergence

1. $X_1 = X$
2. Pour $h = 1, 2, \dots, q$
 - a. Initialiser t_h en sélectionnant au hasard n'importe quelle ligne de X_h
 - i. Calculer p_h ($p_h = \frac{X_{h-1}' t_h}{t_h' t_h}$)
 - ii. Normaliser p_h à 1 ($p_h = \frac{p_h}{\|p_h\|}$)
 - iii. Calculer un nouveau t_h ($t_h = \frac{X_{h-1} p_h}{p_h' p_h}$)
 - iv. Si $\|nouveau\ t_h\| - \|ancien\ t_h\| < s$, passer à l'étape b, sinon retourner à l'étape i
 - b. Calculer X_h ($X_h = X_{h-1} - t_h p_h^T$)

L'algorithme s'assure que chaque composante est équilibrée grâce à l'étape ii. En effet, cette étape évite que l'information ne soit contenue que dans une seule composante.

AB.2 Résolution multivariée de courbes – Moindres carrés alternatifs (*MCR-ALS*)

En considérant les contraintes de que pour chaque observation, la somme des contributions doit être de 1 et qu'il ne doit y avoir aucune négative autant pour les contributions que pour les spectres.

\mathbf{X} : Matrice des données

K : Nombre de composants à extraire

N : Nombre d'observations

\mathbf{S} : Matrice des spectres

\mathbf{C} : Matrice des contributions

s : Critère de convergence

1. Prétraiter \mathbf{X} (*sous-section 4.1.2*)
2. Initialisation des spectres par SIMPLISMA selon K (*sous-section 3.1.4*)
3. Jusqu'à convergence
 - a. Calculer un nouveau \mathbf{C} ($\mathbf{C} = \mathbf{X}\mathbf{S}(\mathbf{S}\mathbf{S}^T)^{-1}$)
 - b. Modifier toute valeur négative dans \mathbf{C} à 0 [Contrainte]
 - c. Pour $n = 1, 2, \dots, N$
 - i. Calculer $\sum C_n$ ($\sum C_n = \sum_{k=1}^K C_{n,k}$)
 - ii. Pour $k = 1, 2, \dots, K$
 1. Modifier $C_{n,k}$ afin de normaliser le pixel ($C_{n,k} = \frac{C_{n,k}}{\sum C_n}$)
[Contrainte]
 - d. Calculer un nouveau \mathbf{S} ($\mathbf{S} = (\mathbf{C}^T \mathbf{C})^{-1} \mathbf{C}^T \mathbf{X}$)
 - e. Modifier toute valeur négative dans \mathbf{S} à 0 [Contrainte]
 - f. Si $||\text{nouveau } \mathbf{S}|| - ||\text{ancien } \mathbf{S}|| < s$, fin de l'algorithme, sinon retourner à l'étape a

Toutes les lignes marquées de [Contrainte] sont présentes dans le but de calculer des spectres et des contributions. Celles-ci peuvent donc être retirées.

ANNEXE C DÉMONSTRATION DE LA LINÉARITÉ DU BRUIT DE FOND AU XPS

Afin d'effectuer cette démonstration, il faut considérer toutes les hypothèses qui ont été énumérées dans la sous-section 4.5.5 *Approximation du bruit de fond d'un mélange de différents composants*. De plus, tout le développement est basé sur la méthode de Tougaard pour le calcul du bruit de fond. Cette méthode ne produisant que très peu d'erreurs, on la considère comme étant « idéale » pour le développement.

AC.1 Développement et démonstration

On cherche à calculer le spectre résultant d'un mélange de K composants. Les proportions que prennent chaque mélange sont définies par c_k , où k est le numéro du composant. Les spectres de chacune des composants, avec bruit de fond, sont définis par s_k et on définit s'_k les spectres dont toutes les valeurs ont été soustraites par la valeur minimale m_k . Enfin, les spectres dont le bruit de fond (b_k) est retiré sont définis par t_k .

$$s_k(E) = s'_k(E) + m_k \quad (\text{AC.1})$$

$$s'_k(E) = t_k(E) + b_k(E) \quad (\text{AC.2})$$

On cherche donc $s'_{mél}$. On pose l'hypothèse que le spectre sans bruit de fond du mélange des K composants est une somme des spectres sans bruit de fond, pondérés par les contributions :

$$\mathbf{t}_{mél} = \sum_{k=1}^K \mathbf{c}_k \mathbf{t}_k \quad (\text{AC.3})$$

Le bruit de fond est calculé selon la méthode de Tougaard :

$$\mathbf{b}_{tougard}(E) = \int_E^{E_1} W(E' - E) \mathbf{s}'(E') dE' \quad (\text{AC.4})$$

$$E_0 \leq E \leq E_1$$

On définit E_0 comme la plus basse énergie cinétique à laquelle le spectre a été acquis et E_1 comme la plus haute énergie cinétique.

En premier lieu, comme toutes les valeurs de \mathbf{s}' et de \mathbf{t}' sont de zéro à E_1 , on peut affirmer que $\mathbf{s}'_{mél}(E_1) = 0$. Cette valeur étant posée, la valeur de $\mathbf{s}'_{mél}(E_1 - 1)$ doit ensuite être calculée. Pour ce, on doit considérer l'équation permettant de calculer cette valeur :

$$\mathbf{s}'_{mél}(E_1 - 1) = \mathbf{t}_{mél}(E_1 - 1) + \int_{E_1-1}^{E_1} W(E' - (E_1 - 1)) \mathbf{s}'_{mél}(E') dE' \quad (\text{AC.5})$$

En développant l'intégrale de l'équation (AC.5) par la méthode des trapèzes, on obtient une intégrale ayant une valeur de 0. On obtient tout simplement cette équivalence pour $\mathbf{s}'_{mél}(E_1 - 1)$:

$$\mathbf{s}'_{mél}(E_1 - 1) = \mathbf{t}_{mél}(E_1 - 1) \quad (\text{AC.6})$$

On veut calculer $\mathbf{s}'_{mél}(E_1 - 2)$ et on connaît déjà $\mathbf{s}'_{mél}(E_1)$ et $\mathbf{s}'_{mél}(E_1 - 1)$. Le développement décrit par les prochaines lignes permet donc d'atteindre les résultats :

$$\mathbf{s}'_{mél}(E_1 - 2) = \mathbf{b}_{tougard}(E - 2, j') + \mathbf{t}_{mél}(E_1 - 2) \quad (\text{AC.7})$$

$$\mathbf{s}'_{m\acute{e}l}(E_1 - 2) = \int_{E-2}^{E_1} W(E' - (E - 2)) \mathbf{s}'_{m\acute{e}l}(E - 2) dE' + \mathbf{t}_{m\acute{e}l}(E_1 - 2) \quad (\text{AC.8})$$

On approxime l'intégrale par la méthode des trapèzes. À noter que ΔE correspond à différence en énergie entre deux points conjoints dans les spectres acquis.

$$\begin{aligned} \mathbf{s}'_{m\acute{e}l}(E_1 - 2) &= \Delta E \frac{[W(\Delta E) \mathbf{s}'_{m\acute{e}l}(E_1 - 1) - W(0) \mathbf{s}'_{m\acute{e}l}(E_1 - 2)]}{2} \\ &\quad + \Delta E W(\Delta E) \mathbf{s}'_{m\acute{e}l}(E_1 - 1) \\ &\quad + \Delta E \frac{[W(2\Delta E) \mathbf{s}'_{m\acute{e}l}(E_1) - W(\Delta E) \mathbf{s}'_{m\acute{e}l}(E_1 - 1)]}{2} \\ &\quad + \Delta E W(2\Delta E) \mathbf{s}'_{m\acute{e}l}(E_1) + \mathbf{t}_{m\acute{e}l}(E_1 - 2) \end{aligned} \quad (\text{AC.9})$$

Dans l'équation (AC.9), les termes $\mathbf{s}'_{m\acute{e}l}(E_1)$ et $W(0)$ égalent à zéro. On peut donc simplifier l'équation de la façon suivante :

$$\begin{aligned} \mathbf{s}'_{m\acute{e}l}(E_1 - 2) &= \Delta E \frac{W(\Delta E) \mathbf{s}'_{m\acute{e}l}(E_1 - 1)}{2} + \Delta E \frac{-W(\Delta E) \mathbf{s}'_{m\acute{e}l}(E_1 - 1)}{2} \\ &\quad + \Delta E W(\Delta E) \mathbf{s}'_{m\acute{e}l}(E_1 - 1) + \mathbf{t}_{m\acute{e}l}(E_1 - 2) \end{aligned} \quad (\text{AC.10})$$

En simplifiant de nouveau, on obtient cette équation :

$$\mathbf{s}'_{m\acute{e}l}(E_1 - 2) = \Delta E W(\Delta E) \mathbf{s}'_{m\acute{e}l}(E_1 - 1) + \mathbf{t}_{m\acute{e}l}(E_1 - 2) \quad (\text{AC.11})$$

En utilisant l'équation (AC.6), on peut remplacer le terme $\mathbf{s}'_{m\acute{e}l}(E_1 - 1)$ dans l'équation (AC.11) :

$$\mathbf{s}'_{m\acute{e}l}(E_1 - 2) = \Delta E W(\Delta E) \mathbf{t}'_{m\acute{e}l}(E_1 - 1) + \mathbf{t}_{m\acute{e}l}(E_1 - 2) \quad (\text{AC.12})$$

On refait le même développement que dans l'équation (AC.9), soit en résolvant l'intégrale par la méthode des trapèzes, mais cette fois pour $E_1 - 3$:

$$\begin{aligned}
\mathbf{s}'_{m\acute{e}l}(E_1 - 3) &= \Delta E \frac{[W(\Delta E)\mathbf{s}'_{m\acute{e}l}(E_1 - 2) - W(0)\mathbf{s}'_{m\acute{e}l}(E_1 - 3)]}{2} \\
&+ \Delta E W(\Delta E) \mathbf{s}'_{m\acute{e}l}(E_1 - 2) \\
&+ \Delta E \frac{[W(2\Delta E)\mathbf{s}'_{m\acute{e}l}(E_1 - 1) - W(\Delta E)\mathbf{s}'_{m\acute{e}l}(E_1 - 2)]}{2} \\
&+ \Delta E W(2\Delta E) \mathbf{s}'_{m\acute{e}l}(E_1 - 1) \\
&+ \Delta E \frac{[W(3\Delta E)\mathbf{s}'_{m\acute{e}l}(E_1) - W(2\Delta E)\mathbf{s}'_{m\acute{e}l}(E_1 - 1)]}{2} \\
&+ \Delta E W(3\Delta E) \mathbf{s}'_{m\acute{e}l}(E_1) + \mathbf{t}_{m\acute{e}l}(E_1 - 3)
\end{aligned} \tag{AC.13}$$

En simplifiant l'équation (AC.13), on obtient ceci :

$$\begin{aligned}
\mathbf{s}'_{m\acute{e}l}(E_1 - 3) &= \Delta E \frac{W(\Delta E)\mathbf{s}'_{m\acute{e}l}(E_1 - 2)}{2} + \Delta E W(\Delta E) \mathbf{s}'_{m\acute{e}l}(E_1 - 2) \\
&+ \Delta E \frac{[W(2\Delta E)\mathbf{s}'_{m\acute{e}l}(E_1 - 1) - W(\Delta E)\mathbf{s}'_{m\acute{e}l}(E_1 - 2)]}{2} \\
&+ \Delta E W(2\Delta E) \mathbf{s}'_{m\acute{e}l}(E_1 - 1) \\
&+ \Delta E \frac{-W(2\Delta E)\mathbf{s}'_{m\acute{e}l}(E_1 - 1)}{2} + \mathbf{t}_{m\acute{e}l}(E_1 - 3)
\end{aligned} \tag{AC.14}$$

$$\begin{aligned}
\mathbf{s}'_{m\acute{e}l}(E_1 - 3) &= \Delta E W(\Delta E) \mathbf{s}'_{m\acute{e}l}(E_1 - 2) \\
&+ \Delta E W(2\Delta E) \mathbf{s}'_{m\acute{e}l}(E_1 - 1) + \mathbf{t}_{m\acute{e}l}(E_1 - 3)
\end{aligned} \tag{AC.15}$$

Similairement à l'équation (AC.12), on remplace les termes \mathbf{s}' dans l'équation (AC.15) :

$$\begin{aligned}
\mathbf{s}'_{m\acute{e}l}(E_1 - 3) &= \Delta E W(\Delta E)[\Delta E W(\Delta E) \mathbf{t}'_{m\acute{e}l}(E_1 - 1) + \mathbf{t}_{m\acute{e}l}(E_1 \\
&- 2)] + \Delta E W(2\Delta E) \mathbf{t}'_{m\acute{e}l}(E_1 - 1) + \mathbf{t}_{m\acute{e}l}(E_1 - 3)
\end{aligned} \tag{AC.16}$$

$$\begin{aligned}
\mathbf{s}'_{m\acute{e}l}(E_1 - 3) &= [\Delta E W(\Delta E)]^2 \mathbf{t}'_{m\acute{e}l}(E_1 - 1) + \Delta E W(\Delta E) \mathbf{t}_{m\acute{e}l}(E_1 \\
&- 2) + \Delta E W(2\Delta E) \mathbf{t}'_{m\acute{e}l}(E_1 - 1) + \mathbf{t}_{m\acute{e}l}(E_1 - 3)
\end{aligned} \tag{AC.17}$$

Encore là, en analysant l'équation (AC.17), on constate que $\mathbf{s}'_{mél}(E - 3)$ est directement proportionnel aux valeurs de $\mathbf{t}'_{mél}$. En réécrivant l'équation (AC.17) en joignant ΔE , $W(\Delta E)$ et $W(2\Delta E)$ et en les considérant comme des constantes, on obtient cette équation :

$$\mathbf{s}'_{mél}(E_1 - 3) = \alpha_1 \mathbf{t}'_{mél}(E_1 - 1) + \beta \mathbf{t}_{mél}(E_1 - 2) + \alpha_2 \mathbf{t}'_{mél}(E_1 - 1) + \mathbf{t}_{mél}(E_1 - 3) \quad (\text{AC.18})$$

$$\mathbf{s}'_{mél}(E_1 - 3) = \alpha \mathbf{t}'_{mél}(E_1 - 1) + \beta \mathbf{t}_{mél}(E_1 - 2) + \mathbf{t}_{mél}(E_1 - 3) \quad (\text{AC.19})$$

Dans l'équation (AC.18), α_1 , α_2 et β sont des constantes. Dans l'équation (AC.19), α est aussi une constante et regroupe α_1 et α_2 .

À partir de l'équation (AC.19), on peut réécrire l'équation sous forme générale pour n'importe quelle énergie :

$$\mathbf{s}'_{mél}(E_1 - X) = \sum_{x=1}^{X-1} \gamma_x \mathbf{t}'_{mél}(E_1 - x) + \mathbf{t}_{mél}(E_1 - X) \quad (\text{AC.20})$$

Dans l'équation (AC.20), on a deux termes distincts. « $\sum_{x=1}^{X-1} \gamma_x \mathbf{t}'_{mél}(E_1 - x)$ » représente le bruit de fond. Considérant l'équation (AC.3), on peut réécrire l'équation (AC.20) de cette façon :

$$\mathbf{s}'_{mél}(E_1 - X) = \sum_{x=1}^{X-1} \left[\gamma_x \left(\sum_{k=1}^K \mathbf{c}_k \mathbf{t}'_k(E_1 - x) \right) \right] + \sum_{k=1}^K \mathbf{c}_k \mathbf{t}'_k(E_1 - X) \quad (\text{AC.21})$$

Avec quelques manipulations, on obtient ceci :

$$\mathbf{s}'_{mél}(E_1 - X) = \sum_{k=1}^K \mathbf{c}_k \sum_{x=1}^{X-1} [\gamma_x \mathbf{t}'_k(E_1 - x)] + \sum_{k=1}^K \mathbf{c}_k \mathbf{t}'_k(E_1 - X) \quad (\text{AC.22})$$

$$\mathbf{s}'_{m\acute{e}l}(E_1 - X) = \sum_{k=1}^K \mathbf{c}_k \left[\sum_{x=1}^{X-1} [\gamma_x \mathbf{t}'_k(E_1 - x)] + \mathbf{t}'_k(E_1 - X) \right] \quad (\text{AC.23})$$

La sommation « $\sum_{x=1}^{X-1} [\gamma_x \mathbf{t}'_k(E_1 - x)]$ » dans l'équation (AC.23) représente le bruit de fond du composant k pour le niveau d'énergie « $E_1 - X$ ».

$$\mathbf{s}'_{m\acute{e}l}(E_1 - X) = \sum_{k=1}^K \mathbf{c}_k [\mathbf{b}_k(E_1 - X) + \mathbf{t}'_k(E_1 - X)] \quad (\text{AC.24})$$

Considérant ce fait, et considérant la définition posée par l'équation (AC.2), on peut réécrire la dernière équation par cette simple formule :

$$\mathbf{s}'_{m\acute{e}l}(E_1 - X) = \sum_{k=1}^K [\mathbf{c}_k \mathbf{s}'_k(E_1 - X)] \quad (\text{AC.25})$$

En analysant l'équation (AC.25), on constate que $\mathbf{s}'_{m\acute{e}l}$ se calcule de façon directement proportionnelle selon la valeur du spectre, avec bruit de fond, de chacun des composants et selon la proportion que chacun des composants représente dans le mélange.

On pose l'hypothèse que la valeur $m_{m\acute{e}l}$ résultante se définit comme suit :

$$m_{m\acute{e}l} = \sum_{k=1}^K \mathbf{c}_k m_k \quad (\text{AC.26})$$

Grâce à cette hypothèse, et à partir de l'équation (AC.25), $\mathbf{s}'_{m\acute{e}l}$ peut s'approximer à cette équation :

$$\mathbf{s}_{m\acute{e}l}(E_1 - X) = \sum_{k=1}^K [\mathbf{c}_k \mathbf{s}_k(E_1 - X)] \quad (\text{AC.27})$$

L'équation (AC.27) permet donc d'approximer $s_{mél}$ en affirmant que ce spectre se calcule par additivité des spectres de chacun des composants selon la contribution correspondante. En plus de l'hypothèse définie par l'équation (AC.26), il faut considérer que la façon de calculer le bruit de fond, soit par la méthode de Tougaard, ne crée pas d'erreurs.

Il faut aussi considérer que la fonction W se calcule de la même façon pour toutes les composants et tous les mélanges de composants. En pratique, la fonction est la même pour tous les métaux et leurs oxydes. Ainsi, si un échantillon analysé ne contient que ces éléments, il n'y a aucun problème. La fonction diffèrera pour d'autres composés, comme les polymères. Ainsi, si un échantillon contient à la fois des polymères et des métaux, l'approximation de l'équation (AC.27) ne pourra s'appliquer.

AC.2 Exemple numérique

Un script *MATLAB* a été implémenté. La fonction universelle de section efficace utilisée dans ce script est la suivante :

$$W(T) = \frac{BT}{(C + T^2)^2} \quad (\text{AC.28})$$

Dans l'équation (AC.28), B et C sont des constantes référencées qui dépendent de la nature du matériau analysé. Dans le script, elles ont été modifiées manuellement afin de calculer des fonds de spectre significatifs.

On a donc deux spectres dont on trace les fonds de spectre par la méthode qui a été présentée dans la section précédente :

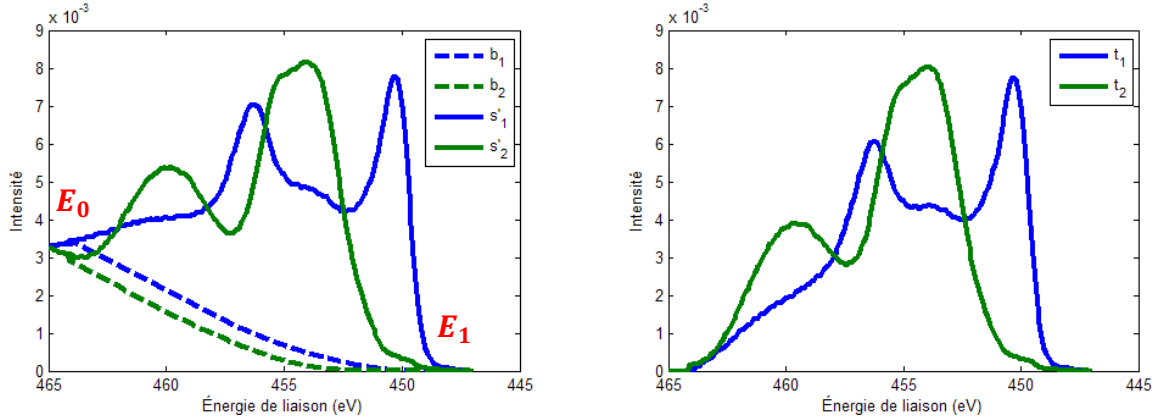


Figure AC.1 Exemples de spectres et bruit de fond calculé selon ceux-ci

On recherche le spectre résultant, avec bruit de fond, d'un mélange des deux composants, avec une proportion égale des deux composants. On connaît donc $t_{mél}$ qui se calcule de cette façon :

$$t_{mél} = 0.5 \times t_1 + 0.5 \times t_2 \quad (\text{AC.29})$$

Pour calculer $s'_{mél}$, deux façons seront comparées. La première est basée sur la résolution de l'intégrale par la méthode des trapèzes, similairement à ce qui est présenté dans l'équation (AC.9). Chaque valeur spectrale est calculée une à la suite de l'autre, en commençant par E_1 et en finissant par E_0 en développant les équations selon la méthode des trapèzes. Ainsi, à une valeur d'énergie X donnée, les valeurs spectrales d'une énergie supérieure à X sont connues. Il est donc possible de calculer le bruit de fond à X , car le calcul du bruit de fond se base sur les valeurs spectrales, avec bruit de fond, aux énergies cinétiques supérieures ($s'_{mél}(E > X)$) et la valeur spectrale, sans bruit de fond, à X ($t'_{mél}(X)$).

La deuxième façon de calculer $s'_{mél}$ est celle qui a été développée. On utilise donc l'équation (AC.25). La Figure AC.2 montre les spectres obtenus par les deux différentes méthodes :

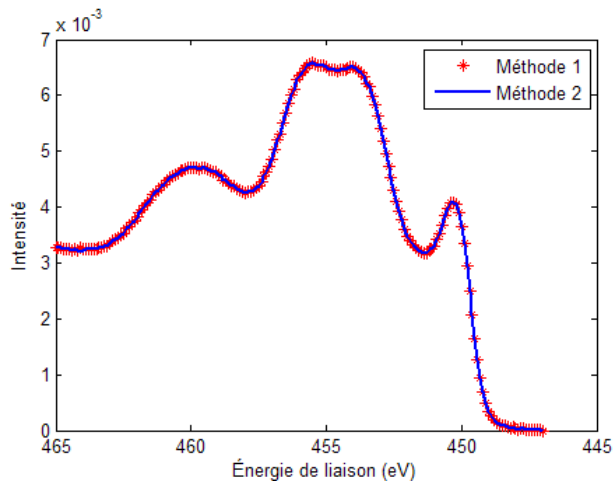


Figure AC.2 Spectres calculés de façon numérique (méthode 1) et par l'équation élaborée (méthode 2)

La Figure AC.2 démontre clairement que le résultat entre les deux méthodes de calcul sont identiques. Le développement qui se conclut par l'équation (AC.25), soit la Méthode 2 dans la Figure AC.2, est donc valide.

ANNEXE D FABRICATION DES ÉCHANTILLONS

AD.1 Conception des masques

Les motifs contenus sur les masques ont été créés grâce au logiciel *Layout Editor*. Tous les motifs consistent en des formes fermées.

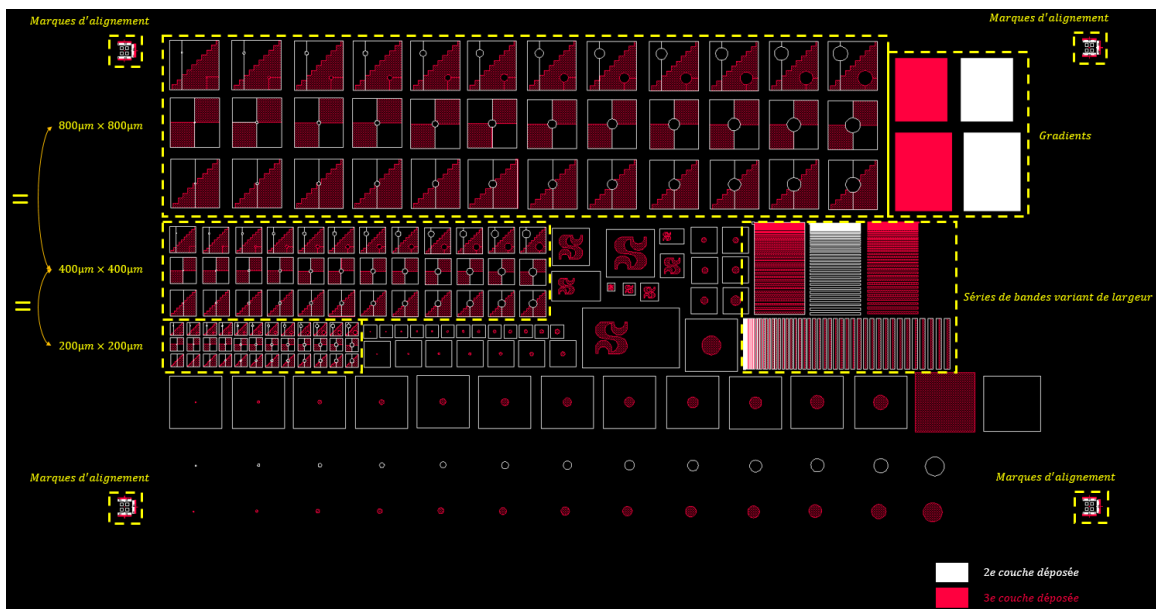


Figure AD.1 Motifs des masques

Différentes zones sont indiquées dans la Figure AD.1. À des fins de précision, à la gauche de l'image, on voit une série de carrés. Cette série s'y trouve trois fois, mais avec trois différentes dimensions de carrés. À la droite, on retrouve une aire de gradient qui consiste à des petits carrés répartis uniformément et dont la quantité diminue au fur et à mesure.

AD.2 Procédé de fabrication

La presque totalité des opérations se sont déroulées dans les salles de blanches du *Laboratoire de nanofabrication et de nanocaractérisation* de l'Université de Sherbrooke. Les étapes qui n'ont pas eu lieu dans ces installations sont clairement indiquées.

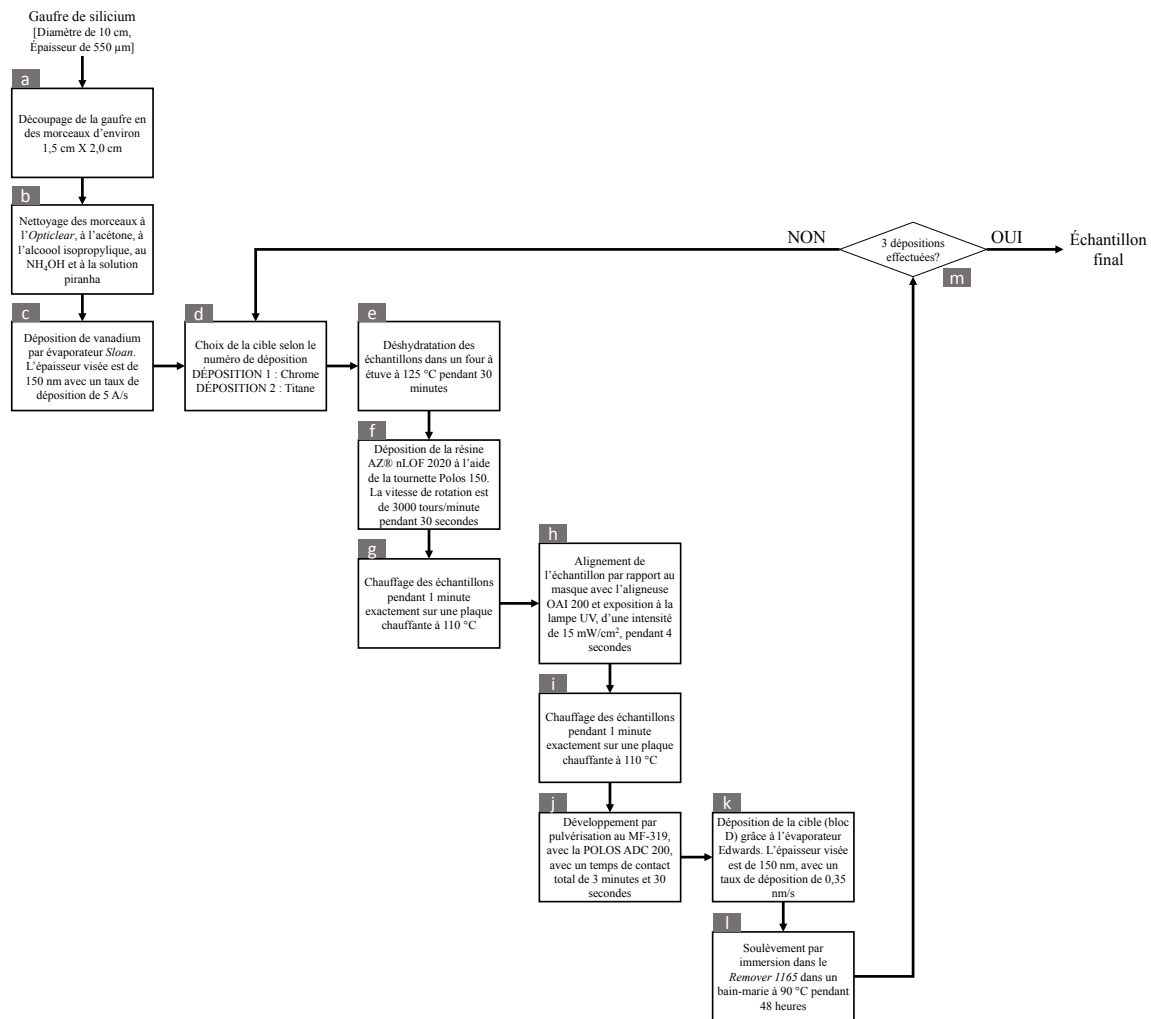


Figure AD.2 Procédé de fabrication des échantillons

AD.2.1 Découpage

Le substrat de base qui est utilisé est une gauffre de silicium, de 550 µm d'épaisseur et de 10 cm de diamètre. La gauffre est découpée en morceaux d'environ 1,5 cm par 2,0 cm à l'aide d'une pointe diamant et d'une pince.

AD.2.2 Nettoyage du substrat

Les échantillons, une fois découpés, passent par une série de nettoyage :

- 5 minutes dans l'*Opticlear* (dégraissant)
- 5 minutes dans l'acétone
- 5 minutes dans l'alcool isopropylique
- Rinçage des échantillons à l'eau
- Séchage au jet d'azote
- Environ une dizaine de secondes dans le NH_4OH
- 15 minutes dans une solution piranha
- Rinçage à l'eau
- Séchage au jet d'azote

AD.2.3 Dépôt du vanadium

La dépôt du vanadium est effectuée par évaporateur *Sloan*. L'épaisseur visée est de 150 nm et le taux de dépôt est de 5 Å/s. Cette opération a été effectuée dans les installations de l'*Infrastructure Matériaux et Dispositifs Quantiques (IMDQ)* de l'Université de Sherbrooke.

AD.2.4 Dépôt de résine

Les étapes suivantes sont les mêmes pour la dépôt du titane et la dépôt du chrome.

Avant l'étalement de la résine, les échantillons sont déshydratés dans un four à étuve à 125 °C pendant 30 minutes.

La résine négative AZ® nLOF 2020. Afin d'étaler la résine, la tournette Polos 150 a été utilisée. La résine est déposée sur l'échantillon à l'aide d'une pipette. La vitesse de rotation ciblée est de 3000 tours par minutes, et ce pendant 30 secondes, ce qui donne une épaisseur de film d'environ 2 µm.

Après la déposition, les échantillons sont déposés pendant une minute sur une plaque chauffante à 110 °C.

AD.2.5 Alignement et exposition

L'aligneuse OAI 200 a été utilisée pour cette étape. Après application du masque au-dessus de l'échantillon et l'alignement de ce dernier, chaque échantillon est exposé à lampe UV pendant 4 secondes. L'intensité de la lampe est de 15 mW/cm².

L'alignement est effectué à l'aide des marques d'alignement. Ces marques sont situées aux quatre coins de l'ensemble des motifs. (*Figure AD.1*)

Après l'exposition, les échantillons sont de nouveau mis sur une plaque chauffante pendant une minute à 110°C.

AD.2.6 Développement

Le développement est effectué grâce à la pulvérisation au MF-319. Le temps de contact avec le développeur est de trois minutes et 30 secondes. Le tout est effectué grâce à la développeuse Polos ADC 200.

AD.2.7 Déposition du titane et du chrome

La déposition du titane et du chrome s'est effectué grâce à un évaporateur Edwards. L'épaisseur visée est de 150 nm, alors que le taux de déposition est de 0,35 nm/s.

AD.2.8 Soulèvement

Après la déposition, les échantillons sont placés dans la solution *Remover 1165* pendant 48 heures, dans un bain-marie dont la température est de 90 °C. Les morceaux de métal flottant à la surface sont ensuite ramassés grâce à une pince.

ANNEXE E CARACTÉRISTIQUES SUPPLÉMENTAIRES DU XPS UTILISÉ

AE.1 Processus de création des photoélectrons

AE.1.1 Sections de l'instrument

Le XPS permet la caractérisation de la surface du matériau grâce à la radiation de ce dernier et par l'analyse des photoélectrons générés par cette radiation. (Figure AE.1) Ainsi, les rayons X sont générés par le tube à rayons X. Ces photons sont ensuite monochromatisés et transmis à l'échantillon, ce qui permet la radiation. Les photoélectrons résultant de ce dernier phénomène passent dans un système de focalisation électronique pour ensuite traverser un spectromètre et être finalement analysés par un détecteur d'électrons. Dans les prochaines sous-sections seront présentées plus en détails les principales parties et pièces qui constituent le XPS.

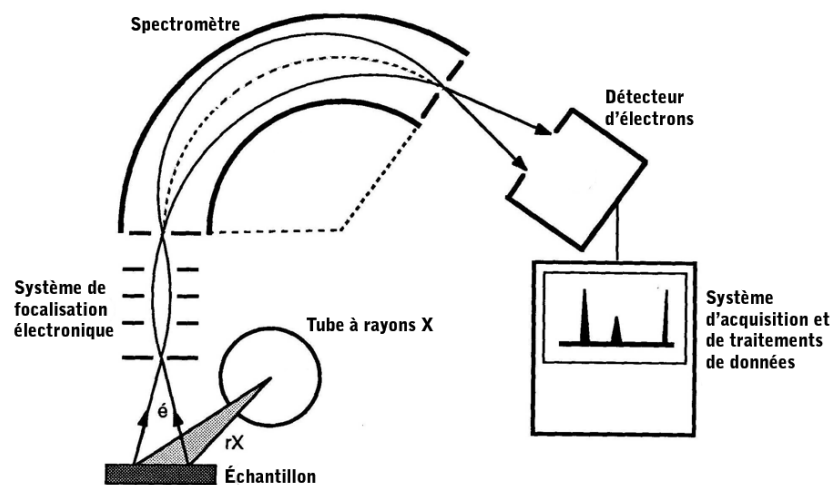


Figure AE.1 Principe global du XPS [55]

AE.1.2 Génération des rayons X

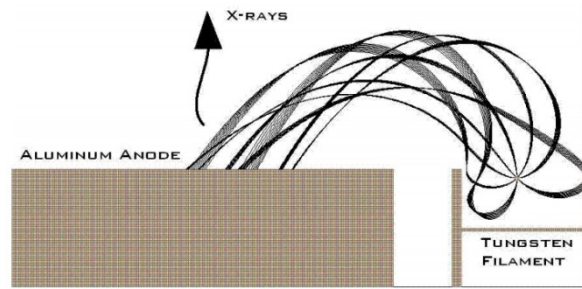


Figure AE.2 Production de rayons X [56]

La Figure AE.2 schématise la production de rayons X dans un tube à rayons X. Le filament de tungstène, qui est mis à la terre, fait office de cathode. Il émet donc des électrons qui sont attirés par l'anode d'aluminium qui est portée à haut voltage, soit à environ 15000 kV. L'excitation des atomes de l'anode émet un rayonnement X. [56]

Les rayons X une fois émis du tube à rayons X, un cristal de quartz permet de filtrer les rayons dans le but d'obtenir une sonde monochromatique. L'alignement de ce cristal monochromateur est très important dans la mesure où la surface analysée doit être irradiée par des rayons X de même énergie. Les rayons incidents sur un cristal sont différemment diffractés selon leur énergie. Cette différence de diffraction peut être calculée par la loi de Bragg, qui met en relation l'angle d'incidence et la longueur d'onde du rayon. Le cristal est courbé selon un certain angle et une fente (diaphragme) sert à sélectionner les rayons X qui constituent la sonde de rayons X. [57] (Figure AE.3)

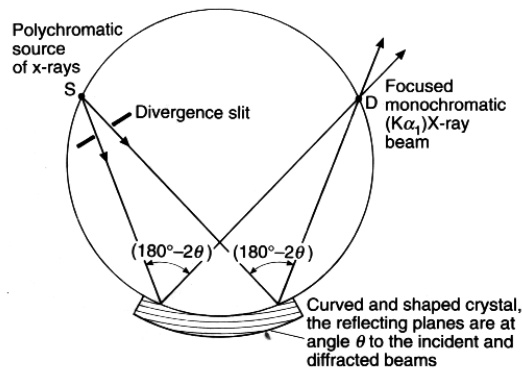


Figure AE.3 Représentation d'un cristal monochromateur [57]

Plusieurs aberrations peuvent être produites à partir de la source de rayons X. On doit tenir compte bien sûr de l'aberration chromatique, en plus de tenir compte de l'aberration de diffraction produit par l'iris à la sortie de l'anode ainsi que les aberrations occasionnées par la diffraction et aux imperfections dues à la surface imparfaite du cristal.

AE.1.3 Système de lentilles

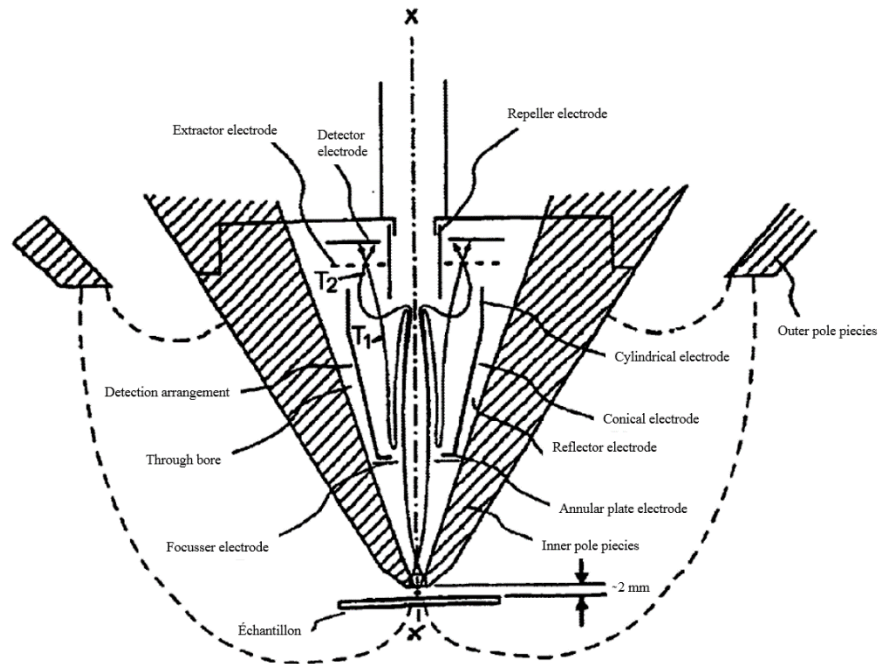


Figure AE.4 Schématisation de la lentille d'immersion magnétique [58]

Les photoélectrons émis sont contraints par un champ magnétique (*voir les lignes pointillées de la Figure AE.4*). Le champ magnétique relie les pièces polaires extérieures et les pièces polaires intérieures. Une fois dans la lentille, les électrons secondaires, qui ont un mouvement circulaire, sont accélérés et concentrés grâce au « Focuser electrode », électrode qui est à un voltage dans les environs de +50 V. [58]

Une fois cette première section passée, les électrons sont dirigés vers l'électrode de répulsion. Cette électrode a un voltage dans les environs de -20 V. Comme les électrons rencontreront un voltage négatif, ceux-ci seront repoussés vers l'électrode de contribution tout en s'éloignant de l'axe longitudinale X-X. Les électrons seront enfin dirigés vers

l'électrode d'extraction. Le haut voltage de cette électrode, soit dans les environs de +150 V, permet d'attirer les photoélectrons. [58]

Enfin, l'électrode de détection représentée dans le haut du schéma de la Figure AE.4 n'est pas présente dans le XPS étudié. Cette absence s'explique par le fait que les photoélectrons sont ensuite dirigés vers des analyseurs dont il sera mention dans les prochaines sous-sections. Dans le XPS utilisé, les photoélectrons entrent à un niveau d'énergie constant. Ainsi, par un jeu de voltage des différentes électrodes de la lentille d'immersion, il est possible de sélectionner les photoélectrons à une énergie cinétique et une énergie de passage données pour que ceux-ci soient quantifiés. [58]

AE.2 Mode spectral versus mode imagerie

AE.2.1 Analyseur hémisphérique concentrique

L'analyseur hémisphérique est la partie de l'instrument où entre en jeu la résolution spectrale. Cet analyseur est utilisé lorsque l'acquisition au XPS se fait en mode spectre. (Figure AE.5)

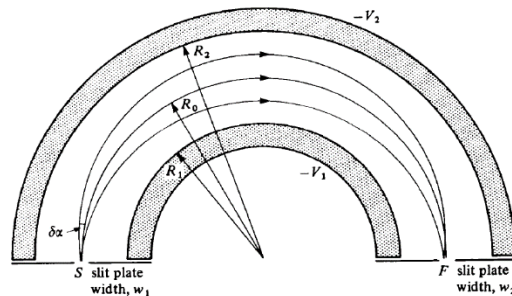


Figure AE.5 Analyseur hémisphérique concentrique [17]

L'équation ci-dessous permet de faire correspondre la trajectoire de l'électron :

$$e(V_2 - V_1) = e\Delta V = E_0 \left(\frac{R_2}{R_1} - \frac{R_1}{R_2} \right) = \frac{1}{k_s} E_0 \quad (\text{AE.1})$$

Dans cette dernière équation et dans la Figure AE.5, les rayons R_1 et R_2 correspondent aux dimensions de l'analyseur et sont bien sûr constants. Pour qu'un électron avec une énergie E_0 puisse passer de l'entrée de l'analyseur à la sortie de l'analyseur (*de S à F sur la Figure AE.5*), on doit donc modifier la différence de potentiel entre les deux hémisphères ($-V_1$ et $-V_2$). Les électrons n'ayant pas l'énergie E_0 dévieront de leur trajectoire et ne seront pas en mesure d'atteindre le point F .

L'équation ci-dessous est une équation généralisée qui permettra de calculer la résolution spectrale de l'analyseur :

$$\frac{\Delta E}{E} = \frac{w_1 + w_2}{2R_0} \quad (\text{AE.2})$$

Dans l'équation précédente, w_1 et w_2 sont les largeurs des fentes respectivement à l'entrée et à la sortie de l'analyseur et R_0 est la distance radiale correspondant au centre de l'analyseur. Ces trois termes étant donc constants, on remplace cette valeur par la constante k_r et on isole ΔE :

$$\Delta E = k_r E \quad (\text{AE.3})$$

ΔE représente la résolution spectrale pouvant être acquise avec une énergie d'électron E . Au niveau spectral, cette résolution correspond au FWHM (*Full-width at half-maximum, Largeur à mi-hauteur*) s'ajoutant aux largeurs naturelles des pics apparaissant dans les spectres XPS. On parle ici de largeurs naturelles de pics comme étant des largeurs calculées à partir de spectres avec une acquisition ne comportant aucun facteur d'erreur. La largeur à mi-hauteur correspond à la distance en x , sur un spectre, entre les deux points qui correspondent à la moitié de l'intensité maximale d'un pic. [59] (*Figure AE.6*)

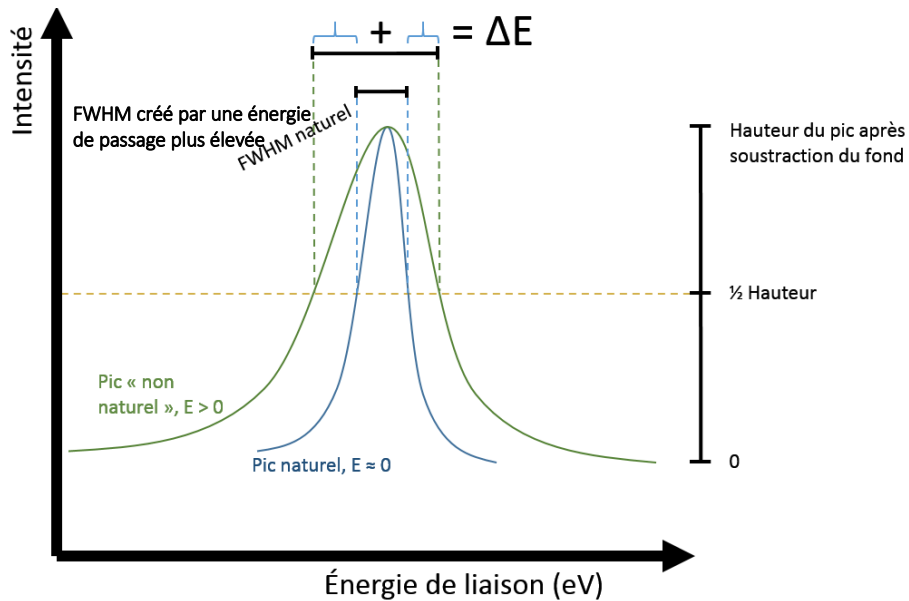


Figure AE.6 Schématisation du FWHM et du ΔE dans les spectres

De toutes ces explications, on peut se demander pourquoi il est nécessaire de ralentir les électrons à une énergie de passage donnée, qui est un paramètre d'acquisition important du XPS. Pour y répondre, on suppose que, par exemple, des électrons qui ont une énergie cinétique de 1500 eV doivent être analysés avec un ΔE de 0.5 V. Or, cette spécification nécessiterait un analyseur hémisphérique de très grande taille (de plus grande taille qu'une pièce de bâtiment standard). C'est pour cela qu'un système de lentilles électromagnétique a été incorporé au système, ce qui permet de réduire considérablement la taille de l'analyseur. Dans l'équation 3.3, E se nomme l'énergie de passage. k_r étant une constante, plus l'énergie de passage (E) sera petite, meilleure sera la résolution des pics (plus petit ΔE). Or, par le fait que l'étendue de la distribution en énergie des électrons soit restreinte, moins d'électrons seront comptés. Les spectres résultant seront donc plus bruités et comporteront donc des pics de moindre intensité. Le système de lentilles électromagnétiques, présenté à la sous-section précédente, permet donc de ralentir les électrons ayant une énergie cinétique désirée dépendamment de l'énergie de passage paramétrée. [17]

En analysant les équations présentées ci-dessus, un analyseur avec une résolution spectrale parfaite aurait des fentes w_1 et w_2 près de 0. Cependant, l'intensité serait très faible et cela

demanderait un très grand temps d'acquisition pour obtenir un bon ratio signal sur bruit. Par le fait que l'énergie de passage des électrons soit fixée et que l'analyseur hémisphérique analyse toujours la même énergie cinétique, on nomme ce mode d'acquisition *Fixed Analyser Transmission* (FAT), ce qui se traduit par *Transmission par analyseur fixe*. [17]

Les paragraphes précédents ont fait part de l'analyseur qui est utilisé en mode spectral. Or, une des innovations majeures du XPS étudié dans ce projet de recherche est l'ajout d'un analyseur utilisé de façon complémentaire en mode imagerie. La prochaine sous-section en fera part.

AE.2.2 Analyseur en miroir sphérique

En mode imagerie, plutôt que dans un analyseur hémisphérique, la déviation des rayons s'effectuera dans un analyseur en miroir sphérique. (*Figure AE.7*) Tel qu'illustré sur la même figure, dans ce mode d'acquisition, deux fentes de l'analyseur hémisphérique concentrique s'ouvrent pour faire passer les rayons dans l'analyseur en miroir sphérique.

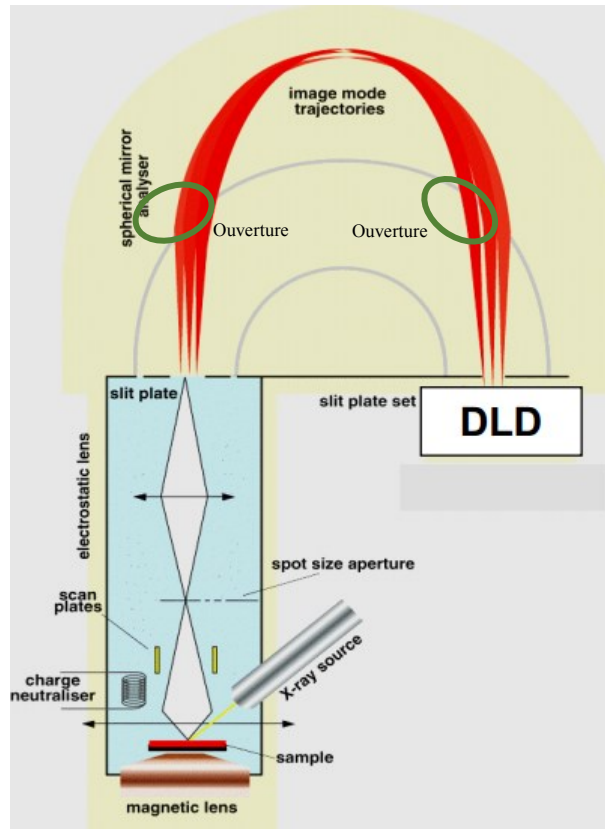


Figure AE.7 Acquisition en mode imagerie de *Kratos Analytical* [60]

Dans l'analyseur en miroir sphérique, les rayons entrent parallèlement entre eux selon leur provenance spatiale sur l'échantillon. Dans l'analyseur en miroir sphérique, chaque rayon se fractionne et se disperse selon son énergie cinétique, phénomène dû à la différence de voltage entre l'hémisphère intérieur et celui extérieur de l'analyseur. Or, grâce à la forme de l'analyseur en miroir sphérique, tous les rayons ayant la même énergie cinétique, peu importe leur provenance spatiale, se focaliseront sur un plan unique. C'est à ce point que se trouve un déflecteur, où une fente ne laissera passer qu'une certaine fraction selon leur énergie cinétique. [61], [62]

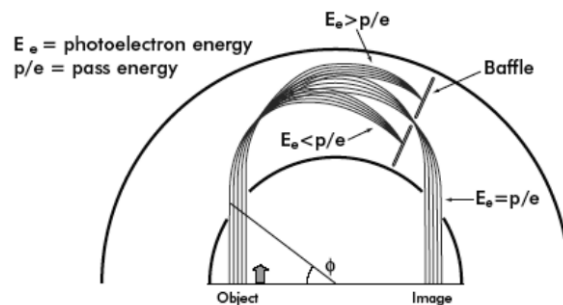


Figure AE.8 Déflecteur dans l'analyseur en miroir sphérique [60]

Passés ce déflecteur, les rayons restants redeviendront parallèles entre eux selon leur provenance spatiale sur l'échantillon. Selon le fabricant du XPS, ce type d'analyseur n'amènerait aucune source d'erreur en termes de diffraction de l'image. La diffraction de l'image se décrit par ce qui est acquis spatialement par l'analyseur ne correspond pas exactement à l'échantillon.

ANNEXE F

Soit cette fonction à minimiser :

$$F = \sum_{n=1}^N \left[\sum_{k=1}^K C_k S_{k,n} \right]^2 \quad (\text{AF.1})$$

La dérivée partielle de F selon une contribution C_l est décrite par l'équation (AF.2). En optimisation, le but est de faire égaliser chaque dérivée partielle à 0.

$$\frac{\delta F}{\delta C_l} = 0 = \sum_{n=1}^N \left[2S_{l,n} \sum_{k=1}^K (C_k S_{k,n}) \right] \quad (\text{AF.2})$$

En additionnant toutes les dérivées partielles pour les K composants, on obtient une somme de 0.

$$\sum_{l=1}^K \frac{\delta F}{\delta C_l} = \sum 0 = 0 = \sum_{l=1}^K \sum_{n=1}^N \left[2S_{l,n} \sum_{k=1}^K C_k S_{k,n} \right] \quad (\text{AF.3})$$

En manipulant l'équation (AF.4), on obtient une sommation pour chacune des N variables.

$$0 = \sum_{n=1}^N \left[2 \sum_{k=1}^K (S_{k,n}) \left(\sum_{k=1}^K C_k S_{k,n} \right) \right] \quad (\text{AF.4})$$

Comme le terme « $2 \sum_{k=1}^K (S_{k,n})$ » n'est pas dépendant des contributions, on le remplace par la constante α_n .

$$0 = \sum_{n=1}^N \left[\alpha_n \left(\sum_{k=1}^K C_k S_{k,n} \right) \right] \quad (\text{AF.5})$$

On insère la constante α_n à l'intérieur de la 2^e sommation.

$$0 = \sum_{n=1}^N \left[\sum_{k=1}^K \alpha_n C_k S_{k,n} \right] \quad (\text{AF.6})$$

On remplace « $\alpha_n S_{k,n}$ » par la constante $\beta_{k,n}$.

$$0 = \sum_{n=1}^N \left[\sum_{k=1}^K \beta_{k,n} C_k \right] \quad (\text{AF.7})$$

On crée la nouvelle constante ε_k en sommant toutes les valeurs $\beta_{k,n}$ associées au composant k .

$$\varepsilon_k = \sum_{n=1}^N \beta_{k,n} \quad (\text{AF.8})$$

Considérant les constantes ε_k , on modifie l'équation (AF.7).

$$0 = \sum_{k=1}^K [\varepsilon_k C_k] \quad (\text{AF.9})$$

On considère cette contrainte :

$$1 = \sum_{k=1}^K C_k \quad (\text{AF.10})$$

On peut donc remplacer C_k par « $1 - \sum_{k=1}^{K-1} C_k$ ».

$$0 = \sum_{k=1}^{K-1} [\varepsilon_k C_k] + \varepsilon_K \left(1 - \sum_{k=1}^{K-1} C_k \right) \quad (\text{AF.11})$$

On modifie l'équation (AF.11) afin de n'avoir qu'une seule sommation.

$$-\varepsilon_K = \sum_{k=1}^{K-1} [(\varepsilon_k - \varepsilon_K) C_k] \quad (\text{AF.12})$$

On divise les deux côtés de l'équation par « $-\varepsilon_K$ ».

$$\frac{-\varepsilon_K}{-\varepsilon_K} = 1 = \sum_{k=1}^{K-1} \left[\frac{(\varepsilon_k - \varepsilon_K)}{-\varepsilon_K} C_k \right] \quad (\text{AF.13})$$

On remplace « $\frac{(\varepsilon_k - \varepsilon_K)}{-\varepsilon_K}$ » par la constante τ_k .

$$1 = \sum_{k=1}^{K-1} [\tau_k C_k] \quad (\text{AF.14})$$

On obtient donc une équation avec $K - 1$ inconnus permettant de trouver les valeurs de contributions minimisant l'erreur au carré.

LISTE DES RÉFÉRENCES

- [1] P. M. Mehl, Y. R. Chen, M. S. Kim, and D. E. Chan, “Development of hyperspectral imaging technique for the detection of apple surface defects and contaminations,” *J. Food Eng.*, vol. 61, no. 1 SPEC., pp. 67–81, 2004.
- [2] S. G. Kong, Y.-R. Chen, I. Kim, and M. S. Kim, “Analysis of hyperspectral fluorescence images for poultry skin tumor inspection,” *Appl. Opt.*, vol. 43, no. 4, pp. 824–833, 2004.
- [3] H. Piao, L. Le Tarte, W. A. Hennessy, and N. Fairley, “Understanding delamination in microelectronic devices using AES, XPS and chemometrics,” *Surf. Interface Anal.*, vol. 39, no. 6, pp. 493–500, 2007.
- [4] Australian Microscopy & Microanalysis Research Facility, “Introduction, Myscope Microanalysis,” 2015. [Online]. Available: <http://www.ammrf.org.au/myscope/analysis/introduction/>.
- [5] J. Farrell, M. Okincha, M. Parmar, and B. Wandell, “Using visible SNR (vSNR) to compare the image quality of pixel binning and digital resizing,” *Library (Lond.)*, vol. 7537, p. 75370C–75370C–9, 2010.
- [6] S. Hajati, J. Walton, and S. Tougaard, “Three-dimensional X-ray photoelectron tomography on the nanoscale: limits of data processing by principal component analysis,” *Microsc. Microanal.*, vol. 19, no. 3, pp. 751–60, 2013.
- [7] P. Cueva, R. Hovden, J. a Mundy, H. L. Xin, and D. a Muller, “Data Processing For Atomic Resolution EELS,” *Arxiv Prepr. arXiv1112.3059*, 2011.
- [8] N. Dobigeon and N. Brun, “Spectral mixture analysis of EELS spectrum-images,” *Ultramicroscopy*, vol. 120, pp. 25–34, 2012.
- [9] S. Pylypenko, K. Artyushkova, and J. E. Fulghum, “Application of XPS spectral subtraction and multivariate analysis for the characterization of Ar⁺ ion beam modified polyimide surfaces,” *Appl. Surf. Sci.*, vol. 256, no. 10, pp. 3204–3210, Jan. 2010.
- [10] G. Schönhense, A. Oelsner, O. Schmidt, G. H. Fecher, V. Mergel, O. Jagutzki, and H. Schmidt-Böcking, “Time-of-flight photoemission electron microscopy – a new way to chemical surface analysis,” *Surf. Sci.*, vol. 480, no. 3, pp. 180–187, Jan. 2001.
- [11] D. M. Photonics, “Microchannel Plates,” pp. 1–8.

- [12] A. Oelsner, O. Schmidt, M. Schicketanz, M. Klais, G. Schönhense, V. Mergel, O. Jagutzki, and H. Schmidt-Böcking, “Microspectroscopy and imaging using a delay line detector in time-of-flight photoemission microscopy,” *Rev. Sci. Instrum.*, vol. 72, no. 10, pp. 3968–3974, 2001.
- [13] U. Vohrer, C. Blomfield, S. Page, and A. Roberts, “Quantitative XPS imaging—new possibilities with the delay-line detector,” *13th Appl. Surf. Anal. Work. AOFA 13*, vol. 252, no. 1, pp. 61–65, 2005.
- [14] B. Schweber, “Understanding the Basics of Low-Noise and Power Amplifiers in Wireless Designs,” 2013.
- [15] SPECS, “2D-Delayline Detector Product Information.”
- [16] I. Mclean, “Signal-to-Noise in Optical Astronomy,” *Report*, pp. 4–19, 2004.
- [17] S. Hofmann, *Auger- and X-Ray Photoelectron Spectroscopy in Materials Science*, vol. 1re. Berlin: Springer, 2013.
- [18] F. Dierks and B. Ag, “Sensitivity and Image Quality of Digital Cameras,” pp. 1–36, 2004.
- [19] S. Tougaard, “QUANTITATIVE ANALYSIS OF THE INELASTIC BACKGROUND IN SURFACE ELECTRON SPECTROSCOPY,” *Surf. Interface Anal.*, vol. 11, no. 9, pp. 453–472, 1988.
- [20] S. Tougaard, “INELASTIC BACKGROUND REMOVAL IN XPS,” *Surf Interface Anal, ECASIA 85, Proc Eur Conf Appl Surf Interface Anal, Oct. 14, 1986 - Oct. 18, 1986*, vol. 9, no. 1–6, p. 135, 1986.
- [21] S. Tougaard, “Universality classes of inelastic electron scattering cross-sections,” *Surf. Interface Anal.*, vol. 25, no. 3, pp. 137–154, 1997.
- [22] J. Végh, “The Shirley-equivalent electron inelastic scattering cross-section function,” *Surf. Sci.*, vol. 563, no. 1–3, pp. 183–190, 2004.
- [23] ASTM, “Standard guide for background subtraction techniques in Auger electron spectroscopy (e995-84),” *ASTM*, vol. 48, pp. 74–115, 1988.
- [24] D. a. Shirley, “High-resolution x-ray photoemission spectrum of the valence bands of gold,” *Phys. Rev. B*, vol. 5, no. 100, pp. 4709–4714, 1972.
- [25] J. Végh, “The Shirley background revised,” *J. Electron Spectros. Relat. Phenomena*, vol. 151, no. 3, pp. 159–164, 2006.
- [26] A. Herrera-Gomez, “The Peak-Shirley Background,” 2012.

- [27] J. Végh, "The Shirley-equivalent electron inelastic scattering cross-section function," *Surf. Sci.*, vol. 563, pp. 183–190, 2004.
- [28] H. Tokutaka, N. Ishihara, K. Nishimori, S. Kishida, and T. Takabuchi, "the Comparison of the Background Removal Methods in XPS Spectra," *Jpn. J. Appl. Phys.*, vol. 29, no. 11, pp. 2512–2513, 1990.
- [29] W. Svante, "Principal Component Analysis," *Chemom. Intell. Lab. Syst.*, vol. 2, pp. 37–52, 1987.
- [30] J. Walton and N. Fairley, "Data scaling for quantitative imaging XPS," *Surf. Interface Anal.*, vol. 41, no. 2, pp. 114–118, 2009.
- [31] M. R. Keenan and P. G. Kotula, "Optimal scaling of TOF-SIMS spectrum-images prior to multivariate statistical analysis," *Appl. Surf. Sci.*, vol. 231–232, pp. 240–244, 2004.
- [32] C. Cordella, "PCA: The Basic Building Block of Chemometrics," *Analytical Chemistry*. 2012.
- [33] M. Wall, A. Rechtsteiner, and L. Rocha, "Singular value decomposition and principal component analysis," *A Pract. Approach to Microarray Data Anal.*, pp. 91–109, 2003.
- [34] W. Windig and C. E. Heckler, "Self-modeling mixture analysis of categorized pyrolysis mass spectral data with the SIMPLISMA approach," *Chemom. Intell. Lab. Syst.*, vol. 14, no. 1–3, pp. 195–207, 1992.
- [35] A. Bogomolov and M. Hachey, "Application of SIMPLISMA purity function for variable selection in multivariate regression analysis: A case study of protein secondary structure determination from infrared spectra," *Spec. Issue WSC-5 2006 Spec. Issue WSC-5 2006*, vol. 88, no. 1, pp. 132–142, 2007.
- [36] A. de Juan, R. Tauler, R. Dyson, C. Marcolli, M. Rault, and M. Maeder, "Spectroscopic imaging and chemometrics: a powerful combination for global and local sample analysis," *TrAC Trends Anal. Chem.*, vol. 23, no. 1, pp. 70–79, 2004.
- [37] A. De Juan, J. Jaumot, and R. Tauler, "Multivariate Curve Resolution (MCR). Solving the mixture analysis problem," *Anal. Methods*, vol. 6, no. 14, pp. 4964–4976, 2014.
- [38] A. de Juan and R. Tauler, "Multivariate Curve Resolution (MCR) from 2000: Progress in concepts and applications," *Crit. Rev. Anal. Chem.*, vol. 36, no. 3–4, pp. 163–176, 2006.

- [39] T. Azzouz and R. Tauler, "Application of multivariate curve resolution alternating least squares (MCR-ALS) to the quantitative analysis of pharmaceutical and agricultural samples," *Talanta*, vol. 74, no. 5, pp. 1201–1210, 2008.
- [40] J. Camacho and A. Ferrer, "Cross-validation in PCA models with the element-wise k-fold (ekf) algorithm: theoretical aspects," *J. Chemom.*, vol. 26, no. 7, pp. 361–373, 2012.
- [41] A. K. Cline and I. S. Dhillon, "Computation of the Singular Value Decomposition," *Handb. Linear Algebr.*, pp. 1–14, 2006.
- [42] A. Kumar, P. K. Jain, and P. M. Pathak, "Curve Reconstruction of Digitized Surface Using K-means Algorithm," *24th DAAAM Int. Symp. Intell. Manuf. Autom. 2013*, vol. 69, no. 0, pp. 544–549, 2014.
- [43] J. Nocedal and S. Wright, "Numerical Optimization," vol. 2. Springer, p. 664, 2006.
- [44] J. F. Bonnans, *Numerical optimization: theoretical and practical aspects*. New York; Heidelberg; Berlin: Springer, 2006.
- [45] J. C. Gilbert and C. A. Sagastizábal, "Numerical Optimization."
- [46] C. W. Therrien, *Discrete Random Signals and Statistical Signal Processing*, 1st ed. Englewood Cliffs: Prentice-Hall, 1992.
- [47] S. Borman, "The Expectation Maximization Algorithm A short tutorial," *Submitt. Publ.*, vol. 25, no. x, pp. 1–9, 2009.
- [48] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *J. R. Stat. Soc.*, vol. 39, no. 1, pp. 1–38, 1977.
- [49] M. C. Biesinger, L. W. M. Lau, A. R. Gerson, and R. S. C. Smart, "Resolving surface chemical states in XPS analysis of first row transition metals, oxides and hydroxides: Sc, Ti, V, Cu and Zn," *Appl. Surf. Sci.*, vol. 257, no. 3, pp. 887–898, 2010.
- [50] M. C. Biesinger, B. P. Payne, A. P. Grosvenor, L. W. M. Lau, A. R. Gerson, and R. S. C. Smart, "Resolving surface chemical states in XPS analysis of first row transition metals, oxides and hydroxides: Cr, Mn, Fe, Co and Ni," *Appl. Surf. Sci.*, vol. 257, no. 7, pp. 2717–2730, 2011.
- [51] J. Walton and N. Fairley, "XPS spectromicroscopy: Exploiting the relationship between images and spectra," *Surf. Interface Anal.*, vol. 40, no. 3–4, pp. 478–481, 2008.

- [52] K. Artyushkova and J. E. Fulghum, “Multivariate image analysis methods applied to XPS imaging data sets,” *Surf. Interface Anal.*, vol. 33, no. 3, pp. 185–195, 2002.
- [53] M.-A. Coulombe, M. Lebeuf, N. Fairley, J. Walton, and G. Soucy, “Cryolitic vein XPS imaging from an industrial grade carbon,” *J. Electron Spectros. Relat. Phenomena*, vol. 185, no. 12, pp. 588–597, 2012.
- [54] R. F. R. F. Egerton and T. Edition, *Electron Energy-Loss Spectroscopy in the Electron Microscope*. 2011.
- [55] IRAMIS, “Spectrométrie de photoélectrons X,” 19-Feb-2013. .
- [56] E. D. Cantando and R. A. Baragiola, “Imagin X-ray photoelectron spectroscopy,” 2000.
- [57] S. Swinnea, “Primary or Incident Beam Monochromator,” 2001. .
- [58] T. Marui, R. K. Badheka, and F. H. Read, “Magntic immersion lense with detection arrangement,” 2003.
- [59] V. S. W. S. I. LTD, *Instruction Manual HA100 Hemispherical Analyser System*. Manchester.
- [60] A. J. Roberts, S. C. Page, and K. Takahashi, “Getting More from XPS Imaging: Multivariate Analysis for Spectromicroscopy,” *J. Surf. Anal.*, vol. 14, no. 2, pp. 160–165, 2007.
- [61] J. E. Fulghum, “Recent developments in high energy and spatial resolution analysis of polymers by XPS,” *J. Electron Spectros. Relat. Phenomena*, vol. 100, no. 1–3, pp. 331–355, 1999.
- [62] J. Walton, “Quantitative surface chemical microscopy by X-ray photoelectron spectroscopy,” *Spectrosc. Eur.*, vol. 19, no. 6, pp. 7–10, 2007.

