



HAL
open science

Cartographie micro-onde et quantification d'incertitudes par inversion-segmentation bayésienne et échantillonnage de Gibbs

Julien Boissy

► **To cite this version:**

Julien Boissy. Cartographie micro-onde et quantification d'incertitudes par inversion-segmentation bayésienne et échantillonnage de Gibbs. Automatique. Université de Bordeaux, 2022. Français. NNT : 2022BORD0216 . tel-03772528

HAL Id: tel-03772528

<https://theses.hal.science/tel-03772528>

Submitted on 8 Sep 2022

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

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



THÈSE PRÉSENTÉE
POUR OBTENIR LE GRADE DE

**DOCTEUR DE
L'UNIVERSITÉ DE BORDEAUX**

ÉCOLE DOCTORALE : SCIENCES PHYSIQUES ET DE L'INGÉNIEUR
SPÉCIALITÉ : AUTOMATIQUE, PRODUCTIQUE, SIGNAL ET IMAGE,
INGÉNIERIE COGNITIVE

Par **Julien Boissy**

**Cartographie micro-onde
et quantification d'incertitudes
par inversion-segmentation bayésienne
et échantillonnage de Gibbs**

Sous la direction de **Jean-François Giovannelli** et **Pierre Minvielle-Larrousse**.

Soutenance prévue le 1^{er} juillet 2022

Membres du jury :

Jérôme Idier	Directeur de recherche	Centrale Nantes	Rapporteur
Marcelo Pereyra	Associate professor	Université Heriot-Watt	Rapporteur
Audrey Giremus	Professeure des universités	Université de Bordeaux	Examineur
Camille Chapdelaine	Ingénieur de recherche	Safran Tech	Examineur
Jean-François Giovannelli	Professeur des universités	Université de Bordeaux	Directeur
Pierre Minvielle-Larrousse	Ingénieur-chercheur	CEA-CESTA	Co-encadrant
Cornelia Vacar	Ingénieure-chercheuse	CEA-CESTA	Invité
Muriel Sesques	Ingénieure-chercheuse	CEA-CESTA	Invité

Titre — Cartographie micro-onde et quantification d’incertitudes par inversion-segmentation bayésienne et échantillonnage de Gibbs

Résumé — On s’intéresse au contrôle électromagnétique d’un objet dans le but de vérifier l’homogénéité de propriétés telles que la perméabilité et la permittivité. Pour réaliser le contrôle, on utilise un moyen expérimental qui mesure des coefficients de réflexion en champ proche. Dans ce contexte, les propriétés électromagnétiques se ramènent à une impédance surfacique homogène par morceaux. Les coefficients de réflexion se déduisent des impédances par un modèle d’observation linéaire dérivé des équations de Maxwell et un terme additif incluant diverses erreurs : mesure, linéarisation et modélisation. Le nombre de données disponibles est petit devant le nombre d’inconnues, ce qui rend l’inversion du modèle d’observation délicate. Ce travail présente une méthode d’inversion bayésienne qui fournit une carte d’impédance homogène par morceaux. On s’appuie sur un modèle *a priori* Gauss-Potts qui repose sur un découpage en région grâce à un jeu d’étiquettes, modélisé par un champ de Potts, et sur l’impédance, modélisée par des champs gaussiens. L’estimation est rendue difficile par le caractère grande dimension et la forme non standard de la distribution *a posteriori*. Pour surmonter cette difficulté, on développe un algorithme de Gibbs afin d’obtenir des échantillons de la loi *a posteriori*. On exploite ces échantillons pour calculer les estimateurs de cartes d’impédances et d’étiquettes, ainsi que les incertitudes. La méthode proposée est validée sur des données simulées.

Mots clés — Inversion-segmentation, statistique bayésienne, problème inverse, impédance surfacique, Gauss-Potts, modèle hiérarchique, MCMC, échantillonneur de Gibbs.

Title — Microwave mapping and uncertainty quantification by Bayesian inversion-segmentation and Gibbs sampling

Abstract — An experimental set-up is dedicated to the microwave control of the homogeneity of electromagnetic properties, such as permeability and permittivity. By moving an antenna close to the material, it is possible to measure near-field reflection coefficients for various locations. One consider that the EM properties can be reduced to a surface impedance with possible discontinuities related to material inhomogeneities. The reflection coefficients are deduced from the impedances by a linearized forward model, derived from Maxwell’s equations. It includes an additive term corresponding to various errors : measurement, linearization and modeling. Since the number of data available is small compared to the number of unknowns, the inversion of the observation model is tricky. Therefore, we have developed a Bayesian inversion method that provides a segmented impedance map. A Gauss-Potts prior relies on a region division by a set of labels, modelled by a Potts field and on the impedance, modelled by Gaussian fields. Estimation is very tricky due to the large dimension and the non-standard form of the posterior. Consequently, we have turned towards a Gibbs algorithm that provides random samples of the *a posteriori* law, that are used to compute the estimate of the impedance map and labels, and to quantify uncertainties. The proposed method is validated on simulated data.

Keywords — Inversion-segmentation, bayesian statistics, inverse problem, surface impedance, Gauss-Potts, hierarchical model, MCMC, Gibbs sampler.

TABLE DES MATIÈRES

Table des matières	8
Notations et acronymes	9
Introduction	11
1 Contexte	13
1.1 Contrôle non destructif en électromagnétisme	14
1.2 Prototype expérimental en champ proche	16
1.2.1 Principe général	16
1.2.2 Banc de mesure	16
1.2.3 Procédure d'acquisition	17
1.3 Contrôle d'homogénéité	19
1.4 Modèle direct	20
1.4.1 Impédance surfacique de matériaux	20
1.4.2 Modèle direct non-linéaire	21
1.4.3 Modèle direct linéarisé	23
1.4.4 Analyse de \mathbf{H} , $\mathbf{H}^\dagger \mathbf{H}$ et $\mathbf{H} \mathbf{H}^\dagger$	24
1.4.5 Analyse complémentaire de la matrice \mathbf{H}	27
1.4.6 Retour au problème de contrôle	28
2 Inversion de modèle linéaire	29
2.1 Introduction	30
2.2 Inversion dans un cadre déterministe	30
2.2.1 Attache aux données et moindres carrés	30
2.2.2 Moindres carrés à norme minimale	31
2.2.3 Régularisation	31
2.3 Inversion dans un cadre probabiliste	32

2.3.1	Maximum de vraisemblance	32
2.3.2	Approche bayésienne	33
2.3.2.1	Maximiseur <i>a posteriori</i>	34
2.3.2.2	Espérance <i>a posteriori</i>	34
2.3.2.3	Conclusion sur l'approche bayésienne	35
2.4	Inversion-segmentation	35
2.4.1	Approches bayésiennes fondées sur le modèle de Potts	35
2.4.2	Analyse complémentaire	37
2.5	Conclusion	38
3	Méthode d'inversion-segmentation	39
3.1	Champ d'impédance par morceaux	40
3.1.1	Mailles, voisinage, projections et troncatures	40
3.1.2	Champ d'étiquette	42
3.1.3	Construction du champ d'impédance	43
3.1.3.1	Champs complets	43
3.1.3.2	Champ par morceaux	45
3.2	Modèle pour le bruit et vraisemblance	49
3.3	Hyperparamètres	50
3.4	Modèle hiérarchique et loi <i>a posteriori</i>	50
3.4.1	Loi jointe	51
3.4.2	Loi <i>a posteriori</i>	51
3.5	Estimateurs	52
3.5.1	Étiquettes	52
3.5.2	Impédance et hyperparamètres	53
3.5.3	Difficultés calculatoires	54
4	Approche numérique	55
4.1	Méthodes Monte Carlo par chaînes de Markov	56
4.1.1	Intégration de Monte Carlo	56
4.1.2	Algorithmes de MCMC	56
4.1.3	Algorithme de Gibbs	58
4.2	Méthode de Gibbs pour l'inversion-segmentation	60
4.2.1	Champs d'impédance	60
4.2.1.1	Difficultés d'échantillonnage	61
4.2.1.2	Échantillonnage efficace en dimension élevée	63
4.2.2	Paramètres d'échelle	65
4.2.2.1	Bruit	66
4.2.2.2	Impédances	66
4.2.3	Champ d'étiquettes	67
4.2.4	Pseudo-code	69

4.3	Calcul des estimateurs	71
4.3.1	Étiquettes	71
4.3.2	Impédance et hyperparamètres	71
5	Résultats numériques	73
5.1	Cas de validation	74
5.1.1	Impédances et étiquettes vraies	74
5.1.2	Données synthétiques	76
5.1.3	Algorithme de Gibbs	77
5.2	Validation de la méthode	78
5.2.1	Paramètre d'échelle	78
5.2.2	Étiquettes	80
5.2.3	Impédance EAP	82
5.2.3.1	Impédances complètes	82
5.2.3.2	Impédance par morceaux	84
5.2.4	Impédances EAP conditionnelle	85
5.3	Conclusion	87
5.4	Application à des données plus réelles	88
6	Conclusion	89
6.1	Synthèse	89
6.2	Contributions	90
6.3	Perspectives	91
A	De Maxwell au système linéaire	93
A.1	Équations de Maxwell	93
A.2	Équations intégrales	94
A.3	Formulation variationnelle	96
A.4	Discrétisation	96
A.5	Calcul de la dérivée	98
B	Impédance complexe	101
C	Matrice de structure	103
D	Matrice bloc	105
D.1	Déterminant de la matrice bloc	105
D.2	Inverse de la matrice bloc	106
E	Échantillonnage des impédances <i>a posteriori</i>	107
F	Résultats complémentaires	111

F.1	Estimation des paramètres d'échelles	111
F.2	Estimation de la carte des étiquettes	112
F.2.1	Résultats obtenus pour d'autres chaînes	112
F.2.2	Utilisation de chaînes plus courtes	113
F.3	Estimation des impédances complètes	113
F.3.1	Résultats obtenus pour d'autres chaînes	113
F.3.2	Utilisation de chaînes plus courtes	113
G	Publication au colloque GRETSI 2019	115
H	Publication dans la revue IEEE Signal Processing Letters	121
	Bibliographie	127

NOTATIONS ET ACRONYMES

Symboles mathématiques

1. $\#$: égal à une constante additive près.
2. \propto : égal à une constante multiplicative strictement positive près.
3. \dagger : transposé conjugué.
4. $-\dagger$: transposé conjugué inverse : $M^{-\dagger} = (M^\dagger)^{-1} = (M^{-1})^\dagger$.
5. \mathcal{N} : loi gaussienne complexe ou réelle.
6. i.i.d. : indépendantes et identiquement distribuées.
7. $\|\bullet\|$: $\|\bullet\|_2$, norme euclidienne.

Densité de probabilité et probabilité

1. $\pi(\mathbf{x})$: $\pi_{\mathbf{X}}(\mathbf{x})$.
2. $\pi(\mathbf{x}|\mathbf{y})$: $\pi_{\mathbf{X}|\mathbf{Y}=\mathbf{y}}(\mathbf{x})$.
3. $\pi(\mathbf{x}, \mathbf{y})$: $\pi_{\mathbf{X}, \mathbf{Y}}(\mathbf{x}, \mathbf{y})$.
4. $\Pr(\mathbf{X} = \mathbf{x}|\mathbf{y})$: $\Pr(\mathbf{X} = \mathbf{x}|\mathbf{Y} = \mathbf{y})$.

Vecteur

1. Pour un vecteur \mathbf{x} de taille N , on note $\mathbf{x}_{-n} = [x_1, x_2, \dots, x_{n-1}, x_{n+1}, \dots, x_N]$ de taille $N - 1$ le vecteur \mathbf{x} sans sa n -ième composante.
2. Pour un vecteur \mathbf{x} de taille N , on note $\mathbf{x}_{1..n} = [x_1, x_2, \dots, x_n]$ le sous vecteur de taille n .
3. On note $\mathbf{1}_N = [1, \dots, 1]^t$ le vecteur de taille N dont toutes les composantes valent 1.
4. On note $\mathbf{1}_n = [0, \dots, 1, \dots, 0]^t$ le vecteur de taille N dont la composante n vaut 1 et les autres 0.

Acronyme

1. MAP : Maximiseur *a posteriori*
2. MMAP : Maximiseur Marginal *a posteriori*
3. EAP : Espérance *a posteriori*
4. CAP : Covariance *a posteriori*

INTRODUCTION

J'ai effectué mon travail de thèse au sein du CEA-CESTA (Centre d'Études Scientifiques et Techniques d'Aquitaine) et du groupe signal et image de l'IMS (Laboratoire de l'Intégration du Matériau au Système), sur le problème du contrôle non destructif.

Le contrôle non destructif (CND) est un outil important dans de nombreux secteurs, par exemple l'aéronautique, l'automobile ou l'industrie de l'énergie. Il permet d'obtenir des informations sur l'état d'une pièce ou d'une structure sans qu'il en résulte des altérations préjudiciables à son utilisation ultérieure [1]. Pour cela, il repose sur l'estimation d'une quantité d'intérêt à partir de mesure. À ce titre, il s'inscrit dans la grande famille des problèmes inverses.

Mon travail porte plus précisément sur le contrôle non destructif en électromagnétisme (EM). L'objectif est de s'assurer que l'objet fabriqué est conforme aux spécifications. Pour y parvenir, on s'intéresse à la reconstruction de propriétés matériaux, permittivité et perméabilité. Cette tâche est rendue possible grâce à deux aspects fondamentaux.

1. Les propriétés matériaux sont modélisées par une impédance surfacique.
2. On s'appuie sur un code EM 3D afin d'en déduire un modèle direct linéaire reliant les mesures aux quantités d'intérêt.

La méthode que l'on présente permet, à partir de mesures, de remonter aux propriétés matériaux et plus précisément à l'impédance surfacique. De plus, on s'intéresse au cas où l'objet mesuré présente d'éventuelles inhomogénéités. Ces inhomogénéités des propriétés électromagnétiques se traduisent par une impédance homogène par morceaux. En plus de l'inversion, il est donc question de reconstruire ces zones au travers d'une segmentation de la surface de l'objet. La méthode proposée se place alors dans la catégorie des méthodes d'inversion-segmentation.

Structure du manuscrit – Afin d'introduire les choses progressivement et de manière cohérente, le manuscrit est partagé en 3 grandes parties.

1. Le Chapitre 1 est dédié au contexte ainsi qu'à la présentation du problème. On y introduit les quantités physiques et le lien qui les unit. Cette partie aboutit à la construction du modèle direct linéaire à bruit additif que l'on inverse par la suite. Le Chapitre 2 présente certaines méthodes classiques dédiées à l'inversion de modèle direct linéaire et aborde le problème joint de l'inversion-segmentation.

2. On présente dans le Chapitre 3 la méthode d'inversion-segmentation proposée, reposant sur une approche bayésienne. On y discute notamment de la modélisation *a priori* des quantités d'intérêt ainsi que de leurs paramètres. La fin de cette partie nous mènera à définir une loi *a posteriori* et des estimateurs, ainsi qu'une manière de quantifier les incertitudes.

3. Une dernière partie regroupe l'aspect résolution numérique. Dans le Chapitre 4, nous présentons les algorithmes Monte Carlo par chaînes de Markov et plus en détail l'algorithme de Gibbs utilisé. Ce dernier nous permet de calculer numériquement les estimateurs choisis et mène au Chapitre 5 dans lequel nous présentons les résultats obtenus à partir de données simulées. Enfin, nous concluons sur les qualités de la méthode, mais également sur ses limites.

Dans le but de rendre le document plus fluide, on regroupe en annexe certains détails, par exemple de complexité calculatoire, et des calculs utiles comme le calcul de déterminant de matrice. On y trouve également des publications au colloque GRETSI 2019 [2] et dans la revue *IEEE Signal Processing Letters* [3]. La dernière partie regroupe les références bibliographiques.

Publications

– J. Boissy, P. Minvielle, M. Sesques et J.-F. Giovannelli, « Cartographie d'impédance par inversion d'un code d'électromagnétisme 3D surfacique », in *Actes du 27^e colloque GRETSI*, Lille, France, août 2019

– J. Boissy, J.-F. Giovannelli et P. Minvielle, « An insight into the Gibbs sampler: Keep the samples or drop them? », *IEEE Signal Processing Letters*, vol. 27, pp. 2069–2073, 2020

Communications

– J. Boissy, « Cartographie d'impédance par inversion d'un code d'électromagnétisme 3D », Rapport de stage de Master, IMS, Talence, France, septembre 2018

– J. Boissy, P. Minvielle, M. Sesques et J.-F. Giovannelli, « Cartographie d'impédance par inversion d'un code Maxwell 3D surfacique », Journées JOSO, Le Barp, France, mars 2019

– J. Boissy, P. Minvielle, M. Sesques et J.-F. Giovannelli, « Inversion pour la cartographie d'impédance surfacique 3D », Journée de l'ED SPI, Talence, France, mars 2020

– J. Boissy, P. Minvielle et J.-F. Giovannelli, « Near-field microwave mapping by Bayesian inversion-segmentation », in *SIAM Conference on Uncertainty Quantification*, Atlanta, Géorgie, États-Unis, avril 2022

CHAPITRE 1

CONTEXTE

Le contrôle non destructif (CND) en électromagnétisme (EM) est un sujet d'importance majeure. Il intervient dans de nombreux domaines, par exemple dans la métallurgie, le nucléaire ou le spatial. Dans notre cas, le but du contrôle est de caractériser les propriétés EM, perméabilité et permittivité, d'un objet et de mettre en évidence d'éventuelles inhomogénéités. Dans ce contexte, les propriétés EM se ramènent à une impédance surfacique homogène par morceaux qui est reliée aux mesures de coefficients de réflexion par un modèle instrument linéaire, dérivé des équations de Maxwell. Notre spécificité tient au fait que les mesures sont réalisées en champ proche par un moyen expérimental développé par le CEA-CESTA. Néanmoins, en l'absence de mesures réelles, nous utiliserons des mesures simulées grâce au modèle instrument.

Sommaire

1.1	Contrôle non destructif en électromagnétisme	14
1.2	Prototype expérimental en champ proche	16
1.2.1	Principe général	16
1.2.2	Banc de mesure	16
1.2.3	Procédure d'acquisition	17
1.3	Contrôle d'homogénéité	19
1.4	Modèle direct	20
1.4.1	Impédance surfacique de matériaux	20
1.4.2	Modèle direct non-linéaire	21
1.4.3	Modèle direct linéarisé	23
1.4.4	Analyse de \mathbf{H} , $\mathbf{H}^\dagger \mathbf{H}$ et $\mathbf{H} \mathbf{H}^\dagger$	24
1.4.5	Analyse complémentaire de la matrice \mathbf{H}	27
1.4.6	Retour au problème de contrôle	28

1.1 Contrôle non destructif en électromagnétisme

Contrôle non destructif – On se place dans le contexte général du contrôle non destructif (CND) qui vise à « *caractériser l'état d'intégrité de structures ou de matériaux, sans les dégrader, soit au cours de la production, soit en cours d'utilisation, soit dans le cadre de maintenances* » (source internet : Wikipédia). Le CND est un élément majeur du contrôle de la qualité des produits. Il se différencie de l'instrumentation de laboratoire et industrielle puisque le but est de détecter des hétérogénéités et anomalies plutôt que de mesurer des paramètres physiques tels que le poids ou les cotes d'une pièce [1]. Présent dans de nombreuses applications, allant de l'industrie automobile à l'aéronautique et l'inspection alimentaire, le CND met en œuvre différents phénomènes physiques et différentes techniques : courants de Foucault, radiographie, ultrasons, thermographie, *etc.*



FIGURE 1.1 – Ligne de transmission (source internet : RF & Microwave Material Measurements: Techniques and Applications).

Contrôle non destructif en EM – Plus spécifiquement, on s'intéresse au contrôle électromagnétique (EM) que l'on retrouve notamment dans les domaines des matériaux agroalimentaires [8], du médical, des absorbants radar, *etc.* Il s'agit de vérifier *via* l'interaction onde-matière que les propriétés EM, *e.g.* perméabilité (μ) et permittivité (ε), de matériaux ou constituants correspondent bien aux valeurs attendues ou spécifiées. Ce sont des nombres complexes qui dépendent de la fréquence. La permittivité exprime le changement d'un matériau pour absorber l'énergie d'un champ électrique incident, la perméabilité est le degré de magnétisation d'un matériau en réponse à un champ magnétique. À noter que l'inspection peut se faire dans un mode de réflexion ou de transmission (voir la Fig. 1.2).

Échantillons homogènes de taille limitée – Les techniques de caractérisation EM [10] vont de la méthode de la ligne de transmission (*cf.* Fig. 1.1), dite de Nicolson-Ross, à la méthode en cavité ou à des approches fondées sur des bancs de mesure en espace libre [10]. Les méthodes classiques en espace libre mesurent la réflectivité d'un échantillon plan et homogène. Ainsi, le Naval Research Laboratory (NRL) a développé ce qui constitue maintenant un standard

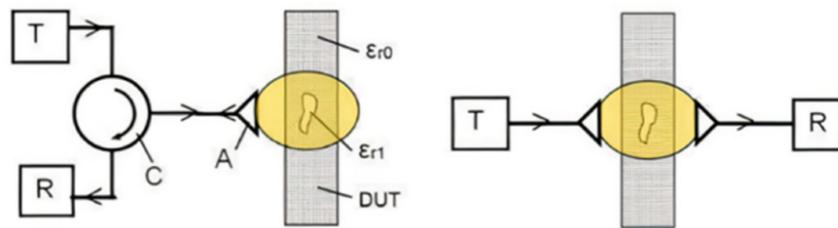


FIGURE 1.2 – Inspection micro-onde en réflexion (à gauche) ou en transmission (à droite) [9].

expérimental, appelé « arche NRL ». Illustré sur la Fig. 1.3, il consiste en un dispositif bistatique apte à mesurer l'échantillon suivant différents angles à l'émission et à la réception. Toutes ces méthodes sont étudiées pour analyser des échantillons homogènes de petite taille, éventuellement positionnés sur un porte-échantillon.

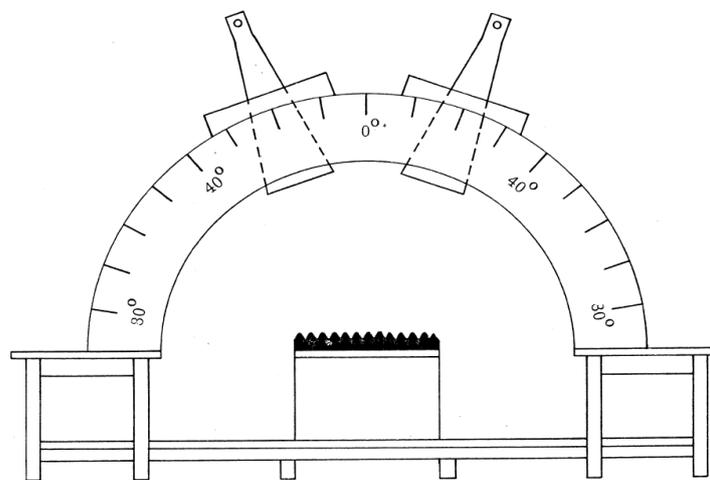


FIGURE 1.3 – Arche bistatique NRL [11] pour le contrôle d'un échantillon plan.

Néanmoins, [12] souligne que ces techniques ne parviennent pas à mettre en évidence les hétérogénéités spatiales qui apparaissent sur des échantillons de grande taille, même si les protocoles de fabrication sont soigneusement contrôlés. Il est important de contrôler les variations des matériaux finaux, puisque des hétérogénéités peuvent avoir des effets considérables sur les performances des composants finaux. Dans ce but, [10] a proposé un nouveau dispositif bistatique qui, adapté à la caractérisation de la perméabilité de bandes d'échantillons de taille plus importante, permet de mettre en évidence d'éventuelles hétérogénéités.

Objet de taille réelle – Le contrôle d'objet de taille réelle consiste à vérifier l'objet dans son ensemble et s'oppose au contrôle d'échantillons ou de maquettes réduites. Les références [13, 14] considèrent la mesure de Surface Équivalente Radar (SER) en espace libre. L'objet est situé à l'intérieur d'une chambre anéchoïque, c'est-à-dire une salle d'expérimentation dont les parois absorbent les ondes électromagnétiques. Cela pose un problème dit de « diffraction inverse » [15], en champ lointain : on cherche à déterminer les propriétés des matériaux, plus précisément les inhomogénéités, à la surface de l'objet à partir de mesures de réflectivité globale, sous différents angles et à différentes fréquences. Si, à géométrie connue, il est possible d'obtenir des cartes de propriétés EM, le processus d'acquisition en espace libre et champ lointain reste long, complexe

à mettre en œuvre et de coût élevé. En outre, la reconstruction proposée dans [13, 14] se limite à des objets axisymétriques.



FIGURE 1.4 – Contrôle micro-onde de matériau composite plan [16].

En rapprochant l'antenne de l'objet, on peut s'affranchir d'une chambre anéchoïque et accélérer le processus d'acquisition. Cela s'apparente à de l'*imagerie micro-onde en champ proche* [16], que l'on retrouve en contrôle non destructif (voir l'illustration de la Fig. 1.4) ainsi que dans d'autres domaines tels que l'imagerie médicale (sein, cerveau, *etc.*), l'imagerie à travers les murs, le contrôle de sécurité (détection d'arme), *etc.* La composante logicielle, incluant le traitement des données et des signaux, y est essentielle. Dans ce contexte, le CEA-CESTA a développé un prototype expérimental pour le contrôle rapide d'objet de taille réelle.

1.2 Prototype expérimental en champ proche

1.2.1 Principe général

Le principe général du prototype expérimental est représenté sur la Fig. 1.5. L'objet sous test, à symétrie de révolution, est positionné verticalement sur son support (colonne en polystyrène, mats en plexiglas, *etc.*). On considère une onde monochromatique émise par une antenne positionnée à quelques longueurs d'onde (quelques centimètres, c'est-à-dire approximativement dans la zone de Fresnel) de l'objet. À la fréquence d'intérêt, l'antenne présente un diagramme de rayonnement directif en champ proche. On peut alors mesurer un « coefficient de réflexion » noté σ , partiellement entaché des interactions multiples antenne-objet.

On donne ci-après quelques détails concernant le banc de mesure ainsi que sur la procédure d'acquisition.

1.2.2 Banc de mesure

Le banc de mesure est représenté sur la photographie de la Fig. 1.6. Sont représentés les 3 axes de translation et les 3 axes de rotation possibles. La précision des positionneurs est inférieure à 1 mdeg pour les platines de rotation et de 4 à 2 μm pour les platines de translation. Le banc de mesure est constitué de

- la cible sous test et son support,

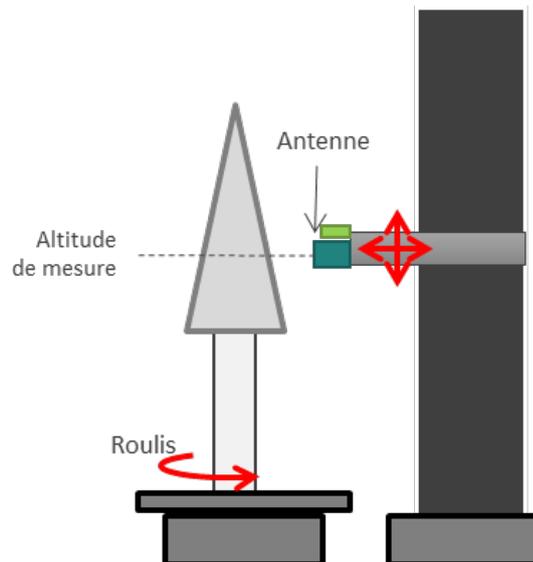


FIGURE 1.5 – Principe du moyen de contrôle en champ proche.

- un analyseur de réseau vectoriel (VNA) embarqué (immobile par rapport à l'antenne),
- quatre platines de translation qui permettent de placer l'antenne à plusieurs altitudes, d'approcher ou d'éloigner l'antenne de l'objet et d'aligner l'axe de rotation de l'objet,
- deux platines de rotation, une placée sous l'objet et une sous l'antenne, qui permettent de modifier l'angle de présentation de l'objet face à l'antenne et de placer l'antenne face à l'objet, mais aussi face à un système de calibration,
- un contrôleur des positionneurs,
- un laser d'alignement.

Antenne « étoile » – L'élément rayonnant utilisé est une antenne en forme d'étoile, voir la Fig. 1.7. Les pétales travaillent par couple opposé et en opposition de phase. Pour améliorer la focalisation ou directivité de la sonde, l'élément rayonnant est placé dans une cavité cylindrique.

1.2.3 Procédure d'acquisition

La procédure d'acquisition consiste en les 4 étapes suivantes. Les acquisitions réalisées sont celles de champ électrique exprimé en Volts.

Étape d'alignement de l'objet à une altitude donnée – Il est nécessaire de maîtriser la distance antenne / objet qui doit rester constante afin de rattraper les erreurs éventuelles de parallaxes. Un algorithme d'alignement permet de minimiser les écarts de distances qui peuvent apparaître en fonction des roulis.

Étape d'acquisition du cylindre étalon (S_{Eta}) – L'antenne est approchée à quelques centimètres du cylindre étalon de diamètre 300 mm et une mesure de coefficient de réflexion est réalisée.

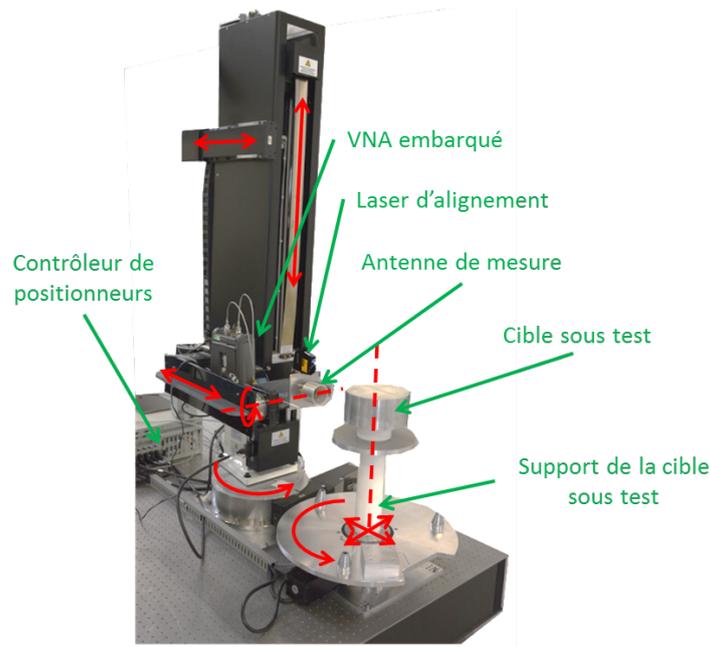


FIGURE 1.6 – Photographie du banc de mesure.



FIGURE 1.7 – Schéma de l'antenne « étoile » en émission et réception.

Étape d'acquisition de l'objet (S_{Obj}) – L'objet sous test est positionné sur un support. L'antenne est positionnée à l'altitude de l'étude. Elle est approchée à quelques centimètres de l'objet et une mesure de coefficient de réflexion est réalisée pour chaque angle de présentation, c'est-à-dire pour différentes rotations de l'objet devant l'antenne. Elle est nécessaire à la soustraction vectorielle qui permet de retirer les contributions dues à l'instrumentation et l'antenne.

Étape d'acquisition de la chambre vide (S_{CV}) – Une antenne est orientée face à un espace libre. Une mesure de coefficient de réflexion est réalisée.

On applique alors la formule de soustraction-calibration suivante pour obtenir le coefficient de réflexion complexe calibré σ :

$$\sigma = \frac{S_{Obj} - S_{CV}}{S_{Eta} - S_{CV}}. \quad (1.1)$$

Cette mesure est effectuée à la fréquence d'intérêt pour M positions d'antenne relativement à l'objet, correspondant à différentes altitudes et à plusieurs angles de présentation. La distance antenne-objet reste constante pendant la procédure. On collecte ainsi M « coefficients de réflexion » mesurés : $\sigma = [\sigma_1, \dots, \sigma_M]^t$, voir la Fig. 1.8.

Remarque 1. Les interactions antennes-objets complexes et multiples n'apparaissent qu'en présence de l'objet. Elles ne sont donc pas supprimées par la soustraction vectorielle. Pour les

réduire, une étape de filtrage spatial de type Réponse Impulsionnelle Finie est mise en œuvre, requérant l'acquisition préalable des coefficients de réflexion dans une bande fréquentielle centrée sur la fréquence d'intérêt.

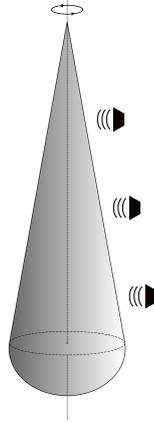


FIGURE 1.8 – Acquisition des mesures en champ proche.

1.3 Contrôle d'homogénéité

On considère qu'il existe des régions ou morceaux de l'objet dans lesquels les propriétés EM sont homogènes (homogénéité intra-régions), mais possiblement différentes entre régions (hétérogénéité inter-régions) et on s'attend à ce qu'elles évoluent lentement à l'intérieur des régions. On parle alors de propriétés EM homogènes par morceaux. À partir des mesures σ obtenues grâce à la procédure d'acquisition décrite précédemment et schématisée par la Fig. 1.9, on cherche à contrôler l'objet, c'est-à-dire à estimer les propriétés électromagnétiques des matériaux qui le constituent et à vérifier leur homogénéité.

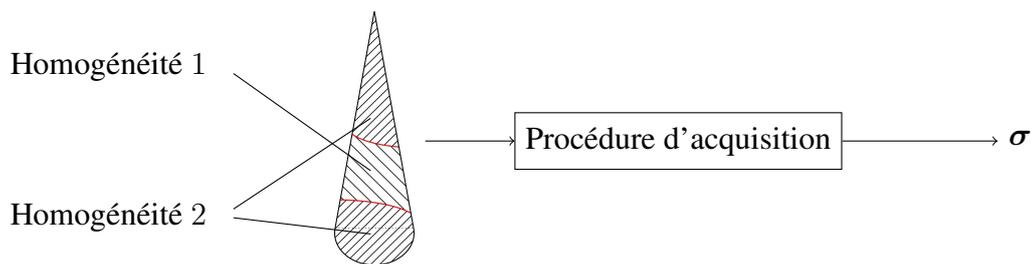


FIGURE 1.9 – Schéma de principe de la procédure d'acquisition des mesures.

Afin de réaliser le contrôle, on s'appuie sur un modèle direct ou modèle d'instrument qui, pour des propriétés électromagnétiques données, est capable de déterminer les coefficients de réflexion que l'on devrait observer. L'objectif est alors de développer un traitement de données pour le contrôle comme représenté sur la Fig. 1.10. La suite du chapitre est dédiée à l'introduction du modèle direct.

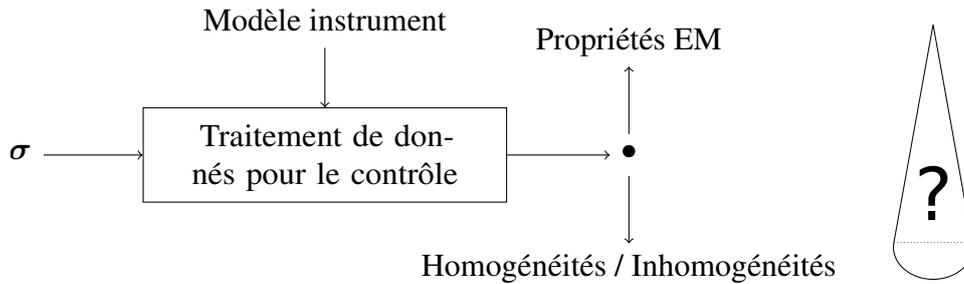


FIGURE 1.10 – Schéma de principe du contrôle d'homogénéité.

1.4 Modèle direct

On considère un objet tridimensionnel situé dans le vide, typiquement un conducteur recouvert de matériaux inhomogènes éclairé par une onde monochromatique émise par une antenne positionnée à proximité, voir la Fig. 1.11.

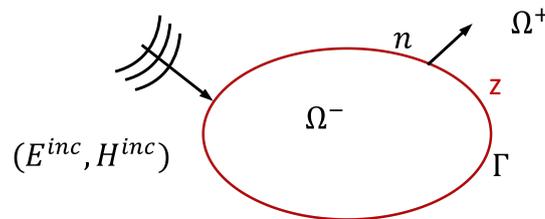


FIGURE 1.11 – Schéma général de la diffraction sur un obstacle.

Le champ diffracté par cet objet est solution des équations de Maxwell dans le régime harmonique. Elles peuvent se mettre sous une forme variationnelle¹. Elles font alors intervenir des équations intégrales définies sur la surface de l'objet et sur les interfaces entre les matériaux qui le composent. Ces équations relient le champ diffracté aux propriétés EM de l'objet. Se référer à l'annexe A pour plus de détails.

1.4.1 Impédance surfacique de matériaux

Naturellement on voudrait considérer les propriétés des matériaux qui recouvrent l'objet, c'est-à-dire la permittivité (ϵ) et la perméabilité (μ), et résoudre les équations de Maxwell sous forme d'intégrale surfacique sur les interfaces entre les matériaux. Or, une discrétisation spatiale appropriée dépend de la longueur d'onde dans le matériau : plus l'indice du milieu est élevé et plus la longueur d'onde est petite. Le nombre de mailles volumiques requis est alors important, conduisant à un gros système linéaire, avec un coût de résolution élevé en temps de calcul et taille de mémoire.

1. Le principe de l'approche variationnelle pour la résolution des équations aux dérivées partielles est de remplacer l'équation par une formulation équivalente, dite, variationnelle, obtenue en intégrant l'équation multipliée par une fonction quelconque, dite test.

Condition aux limites de Leontovich – Il est possible d’obtenir une approximation à moindre coût en remplaçant l’empilement de matériaux par une « condition aux limites d’impédance » à la surface extérieure de l’objet [17], voir la Fig. 1.12. Cela permet de s’affranchir de la résolution des équations de Maxwell dans les matériaux, avec pour seule interface la surface extérieure de l’objet. La discrétisation ne dépend plus que de la longueur d’onde dans le vide. La « condition d’impédance » la plus simple, dite de Leontovich [17, 18], relie par une impédance² scalaire et complexe z les composantes tangentielles des champs électriques et des champs magnétiques. Elle est strictement locale et constitue une bonne approximation pour des matériaux dont les indices sont suffisamment élevés. Les matériaux qui recouvrent l’objet sont alors représentés par l’impédance z à la surface.

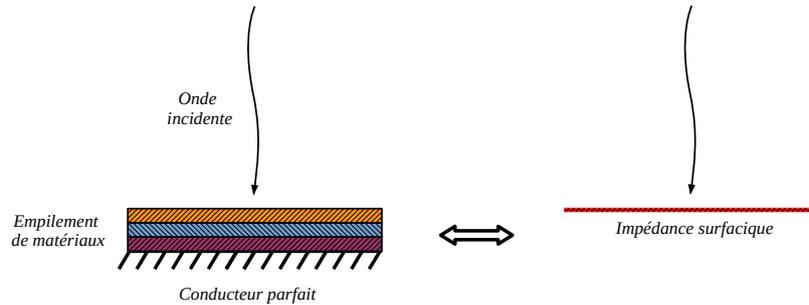


FIGURE 1.12 – Schéma équivalent de l’impédance surfacique.

Remarque 2. Pour déterminer l’impédance surfacique z à partir des propriétés (ϵ, μ) on s’appuie classiquement sur des modèles analytiques [19] adaptés à un matériau multi-couches de type plan infini.

1.4.2 Modèle direct non-linéaire

Afin de traiter le problème numériquement, on considère la surface de l’objet discrétisée sous la forme d’un maillage triangulaires, Fig. 1.13, et l’on note $z \in \mathbb{C}^N$ l’impédance discrétisée. On précise que la hauteur de l’objet est d’environ 60 cm.

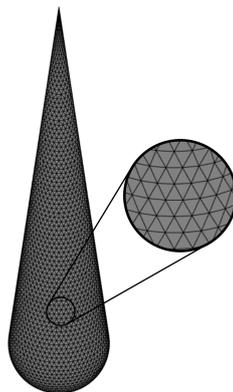


FIGURE 1.13 – Surface du cône-sphère IEEE discrétisée sous la forme de N mailles triangulaires.

2. (Wikipédia) L’impédance est une caractéristique d’un système physique, définie comme le rapport d’une grandeur caractérisant une excitation à laquelle il est soumis à celui d’une grandeur caractérisant sa réponse.

Système d'équations linéaire – À impédance de surface donnée, on sait calculer les coefficients de réflexion en champ proche σ_m à l'aide du code Maxwell 3D, développé au CEA-DAM. Fondé sur la résolution du système de Maxwell sous forme intégrale en régime harmonique, ce code d'électromagnétisme 3D permet de traiter des assemblages de matériaux diélectriques et de conducteurs, et tient compte des invariances géométriques. Ce code massivement parallèle exploite efficacement plusieurs milliers de cœurs du supercalculateur TERA1000 [20]. Les équations intégrales sont résolues par une méthode d'éléments finis de surface sous forme standard de Raviart-Thomas [21]. Lorsque l'empilement de matériaux est modélisé par une impédance surfacique [22], la formulation variationnelle associée à la résolution des équations de Maxwell sous forme intégrale est symétrique, mais non hermitienne [23].

On pourra se référer à la partie A.4 de l'annexe A pour plus de détails sur ce qui suit. Pour la direction d'illumination m , on note $\mathbf{x}_m \in \mathbb{C}^{2N_a}$ les flux de courants électriques et magnétiques à travers les N_a arêtes du maillage 3D. À fréquence et direction d'illumination données, le système discret qui en résulte se met sous la forme :

$$\mathbf{A}(\mathbf{z})\mathbf{x}_m = \mathbf{b}_m \quad (1.2)$$

où $\mathbf{A}(\mathbf{z}) \in \mathbb{C}^{2N_a \times 2N_a}$ est la matrice du système qui dépend de \mathbf{z} , et \mathbf{b}_m le second membre fonction de la direction d'éclairage.

On a $\mathbf{x}_m(\mathbf{z}) = \mathbf{A}(\mathbf{z})^{-1}\mathbf{b}_m$ la solution de l'Eq. (1.2) et on a directement l'intégrale du rayonnement : $\sigma_m = \mathbf{x}_m^t(\mathbf{z})\mathbf{c}$, où \mathbf{c} est l'opérateur d'observation en réception. Considérant toutes les directions, le système s'écrit :

$$\mathbf{A}(\mathbf{z})\mathbf{X} = \mathbf{B} \quad (1.3)$$

avec $\mathbf{X} = [\mathbf{x}_1, \dots, \mathbf{x}_M]$ et $\mathbf{B} = [\mathbf{b}_1, \dots, \mathbf{b}_M]$.

Modèle direct non linéaire – On note $\mathbf{X}(\mathbf{z}) = \mathbf{A}(\mathbf{z})^{-1}\mathbf{B}$ la solution de l'Eq. (1.3) et on obtient alors la relation non-linéaire en \mathbf{z}

$$\boldsymbol{\sigma} = \mathbf{h}(\mathbf{z}) = \mathbf{X}(\mathbf{z})^t \mathbf{c} = \mathbf{B}^t \mathbf{A}(\mathbf{z})^{-t} \mathbf{c}. \quad (1.4)$$

La matrice $\mathbf{A}(\mathbf{z})$ est pleine et la taille du système croît de façon quadratique avec la fréquence. L'emploi d'un solveur direct de type Cholesky est privilégié par rapport à un solveur itératif en raison du grand nombre de seconds membres, ainsi que pour des raisons de précision dues au mauvais conditionnement de la matrice $\mathbf{A}(\mathbf{z})$.

Remarque 3. Lorsque le problème atteint des centaines de milliers d'inconnues, l'emploi d'un solveur direct devient irréaliste, car il exige des ressources mémoire en $\mathcal{O}(N^2)$ et un nombre d'opérations en $\mathcal{O}(N^3)$. On utilise alors la méthode des H-matrices [20], conduisant à un stockage compact ou compressé de la matrice initiale. Ainsi, le coût d'une factorisation de Cholesky d'une H-matrice est en $\mathcal{O}(N \log_2(N))$ contre $\mathcal{O}(N^3)$ pour une matrice dense.

Remarque 4. Pour réaliser le contrôle d'homogénéité, il faut disposer d'un modèle fidèle de l'interaction onde-objet, incluant le rayonnement de l'antenne. Cette modélisation n'est pas parfaite, notamment en raison des interactions complexes évoquées à la Remarque 1, de la représentation imparfaite de l'antenne par un dipôle élémentaire et de l'électronique hyperfréquence associée, de la modélisation éventuellement simplifiée des matériaux, etc. Cependant, on peut s'appuyer sur une hypothèse de faible perturbation que l'on introduit dans la partie suivante.

1.4.3 Modèle direct linéarisé

Dans ce cadre de contrôle, on s'attend à des écarts de propriétés faibles relativement à une référence. Ainsi, en plus de l'impédance z , on considère une impédance de référence notée $z_0 \in \mathbb{C}^N$.

Linéarisation – On considère l'approximation d'ordre 1 autour de cette impédance de référence z_0

$$\mathbf{h}(z) \approx \mathbf{h}(z_0) + \mathbf{H}(z - z_0), \quad (1.5)$$

avec $\mathbf{h}(z_0) = \mathbf{X}(z_0)^t \mathbf{c} = \mathbf{B}^t \mathbf{A}(z_0)^{-t} \mathbf{c}$ et \mathbf{H} la matrice jacobienne ou matrice des dérivées partielles de $\mathbf{h}(z)$, déterminées au point de référence z_0

$$\mathbf{H} = \left. \frac{\partial \mathbf{h}}{\partial z} \right|_{z_0} = \left. \frac{\partial \mathbf{X}}{\partial z} \right|_{z_0}^t \mathbf{c}.$$

Par dérivation de Eq. (1.3) relativement à la n -ième composante de z , on obtient :

$$\mathbf{A}(z) \frac{\partial \mathbf{X}}{\partial z_n} + \frac{\partial \mathbf{A}}{\partial z_n} \mathbf{X}(z) = 0, \quad \text{avec} \quad \frac{\partial \mathbf{A}}{\partial z_n} = \left[\frac{\partial \mathbf{a}_{ij}}{\partial z_n} \right]_{i,j=1 \dots 2N_a},$$

et en z_0 , on a donc le système linéaire :

$$\mathbf{A}(z_0) \left. \frac{\partial \mathbf{X}}{\partial z_n} \right|_{z_0} = - \left. \frac{\partial \mathbf{A}}{\partial z_n} \right|_{z_0} \mathbf{X}(z_0). \quad (1.6)$$

Ainsi, en plus de la résolution standard qui forme $\mathbf{A}(z_0)$ et détermine $\mathbf{X}(z_0)$ puis $\sigma_0 = \mathbf{h}(z_0)$, pour obtenir la matrice jacobienne \mathbf{H} , il faut résoudre les N systèmes supplémentaires, Eq. (1.6), qui diffèrent par leurs seconds membres. On exploite alors le fait que l'impédance intervient dans la formulation variationnelle [23] uniquement sous la forme d'une intégrale simple du type $\int_S z J \cdot J' + z^{-1} M \cdot M'$ ds, voir partie A.5 de l'annexe A, où les courants électriques J et magnétiques M à la surface de l'objet sont les inconnues directes du système de Maxwell tandis que J' et M' sont les fonctions tests de la méthode des éléments finis utilisée pour la résolution, voir partie A.4 de l'annexe A. La dérivation de la matrice \mathbf{A} par rapport à l'impédance conduit ainsi à des matrices $\partial \mathbf{A} / \partial z_m$ très creuses, qui sont stockées de façon optimisée pour effectuer efficacement les produits matrices vecteurs.

Comme mentionné dans la Remarque 4, on considère une représentation imparfaite de l'antenne et, dans notre cas, σ_0 et \mathbf{H} sont calculés à partir d'une modélisation du rayonnement par un dipôle élémentaire.

Modèle direct linéarisé à bruit additif – À partir de l'approximation d'ordre 1 définie par l'Eq. (1.5), on considère un terme d'erreur qui prend en compte les erreurs de modélisation abordée dans la Remarque 4, les erreurs de représentation par une impédance surfacique, les erreurs de linéarisation et les erreurs de mesure. Notre travail se place en amont de la caractérisation du terme d'erreur et pour les besoins de la suite, on considère qu'il est additif et on le note ε , voir la Remarque 5. Finalement, à partir de l'Eq. (1.5), on obtient le modèle direct linéarisé à bruit additif

$$\sigma = \sigma_0 + \mathbf{H}(z - z_0) + \varepsilon, \quad (1.7)$$

sur lequel s'appuie notre méthode de reconstruction de l'impédance présentée dans la suite. Pour notre travail, on considère que le modèle linéaire, Eq. (1.7), est valide pour des variations de l'impédance jusqu'à $\pm 10\%$ de l'impédance de référence et on choisit une référence représentative uniforme $z_0 = z_0 \mathbf{1}_N$, avec $z_0 = 0.686 - i 0.2329$. Cette considération et ce choix sont tirés de [19, 24] (matériau absorbant « école »).

Remarque 5. *Le modèle additif d'erreur, très courant [24–26], est retenu en absence de caractérisation. Un travail de métrologie pourrait être entrepris par la suite pour caractériser les incertitudes de mesure, par exemple l'approche de GUM (Guide pour l'expression de l'incertitude de mesure) ou l'approche de Monte Carlo [27].*

Remarque 6. *L'impédance de référence z_0 est uniforme et l'objet est invariant par rotation. Les coefficients de réflexion σ_0 sont alors constants pour tous les angles à une altitude fixée. L'objet étant discrétisé, on s'attend néanmoins à de petites variations. Ce point est présenté dans la partie 5.1.2 page 76.*

Structure de la matrice \mathbf{H} – Les composantes de la matrice \mathbf{H} sont fonction des directions d'illuminations. Les positions successives de l'antenne sont à distance constante de l'objet. Ainsi, les directions d'illuminations ne sont définies que par leurs altitudes et leurs angles. La matrice \mathbf{H} , représentée sur la Fig. 1.14, est structurée comme suit : la ligne m correspond à la m -ième position d'illumination et la colonne n correspond à la maille d'indice n . Dans notre cas, on considère 25 altitudes régulièrement espacées sur la hauteur de l'objet et 12 angles de rotation régulièrement espacés entre $[-180, 180[$. Au total, il y a donc $M = 300$ positions d'illumination, que l'on représente sur la Fig. 1.15. Les propriétés de la matrice \mathbf{H} sont essentielles pour la suite et dans la partie qui suit, on se propose d'étudier les matrices \mathbf{H} , $\mathbf{H}^\dagger \mathbf{H}$ et $\mathbf{H} \mathbf{H}^\dagger$.

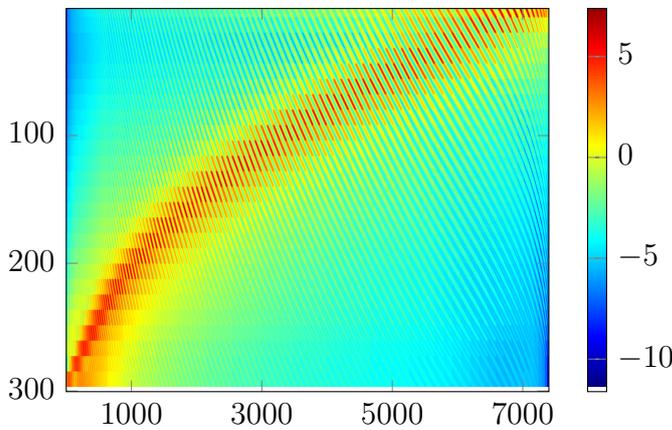


FIGURE 1.14 – Matrice $\log |\mathbf{H}|$ pour $f = 1.5$ GHz.

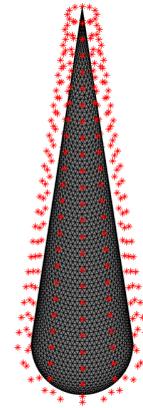


FIGURE 1.15 – Emplacements de l'antenne autour de l'objet.

1.4.4 Analyse de \mathbf{H} , $\mathbf{H}^\dagger \mathbf{H}$ et $\mathbf{H} \mathbf{H}^\dagger$

La matrice \mathbf{H} présentée dans cette section est celle utilisée tout du long de notre travail. Pour l'obtenir, on s'est placé dans le cas représentatif où la fréquence est 1.5 GHz et l'antenne est de type dipôle élémentaire. On propose une analyse de la matrice $\mathbf{H} \in \mathbb{C}^{M \times N}$ et des matrices $\mathbf{H}^\dagger \mathbf{H} \in \mathbb{C}^{N \times N}$ et $\mathbf{H} \mathbf{H}^\dagger \in \mathbb{C}^{M \times M}$, toutes deux symétriques hermitiennes et semi-définies positives.

Matrice \mathbf{H} – On s'intéresse dans un premier temps aux valeurs singulières de la matrice \mathbf{H} que l'on note μ_m , pour $m = 1 \dots M$. Comme \mathbf{H} est une matrice de taille $M \times N$, il existe deux matrices carrées unitaires \mathbf{U} et \mathbf{V} de tailles respectives $M \times M$ et $N \times N$, ainsi qu'une matrice rectangulaire $\mathbf{\Sigma}$ de taille $M \times N$ telles que [28, p. 63] :

$$\mathbf{U}^\dagger \mathbf{H} \mathbf{V} = \mathbf{\Sigma} = \left[\begin{array}{ccc|c} \mu_1 & & & \\ & \mu_2 & & \\ & & \ddots & \\ \bigcirc & & & \mu_M \\ & & & \bigcirc \end{array} \right] \quad (1.8)$$

La matrice $\mathbf{\Sigma}$ est constituée de deux sous matrices. La première est de taille $M \times M$ et contient sur sa diagonale les M valeurs singulières réelles positives de \mathbf{H} . La seconde de taille $M \times (N - M)$ est nulle, voir l'Eq. (1.8). On sait que le rang de la matrice \mathbf{H} est égal au nombre de valeurs singulières non nulles. À partir des valeurs singulières représentées sur la Fig. 1.16, on peut dire que la matrice \mathbf{H} est de rang plein³ et qu'elle est mal conditionnée. Le conditionnement est le ratio entre la plus forte et la plus faible des valeurs propres ou singulières et dans notre cas, il est de l'ordre de 10^5 . Dans le meilleur des cas, il vaut 1.

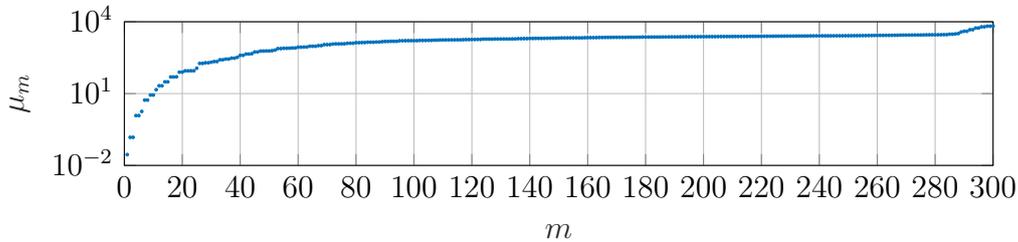


FIGURE 1.16 – Représentation en échelle logarithmique des valeurs singulières de la matrice \mathbf{H} , notées μ_m , pour $m = 1 \dots M$. On note qu'il y a autant de valeurs singulières que de mesures. Le ratio entre la plus forte et la plus faible des valeurs singulières est de l'ordre de 10^5 .

Dans la suite on s'intéresse aux matrices $\mathbf{H}^\dagger \mathbf{H}$ et $\mathbf{H} \mathbf{H}^\dagger$ représentées sur la Fig. 1.17. Puisque le nombre de données est inférieur au nombre d'inconnue, on sait que la matrice $\mathbf{H}^\dagger \mathbf{H}$ est non inversible. Au vu des valeurs singulières de la matrice \mathbf{H} , on s'attend à ce que $\mathbf{H} \mathbf{H}^\dagger$ soit inversible, mais mal conditionnée.

Remarque 7. Comme les matrices $\mathbf{H}^\dagger \mathbf{H}$ et $\mathbf{H} \mathbf{H}^\dagger$ sont symétriques hermitiennes et semi-définies positives, on sait que leurs valeurs propres sont réelles positives [28, p. 62].

Matrice $\mathbf{H} \mathbf{H}^\dagger$ – En s'appuyant sur la décomposition de l'Eq. (1.8), on peut écrire

$$\mathbf{H} \mathbf{H}^\dagger = \mathbf{U} \mathbf{\Sigma} \mathbf{V}^\dagger \mathbf{V} \mathbf{\Sigma}^\dagger \mathbf{U}^\dagger = \mathbf{U} \mathbf{\Sigma} \mathbf{\Sigma}^\dagger \mathbf{U}^\dagger$$

et l'on en déduit la décomposition en valeurs propres de la matrice $\mathbf{H} \mathbf{H}^\dagger$:

$$\mathbf{U}^\dagger (\mathbf{H} \mathbf{H}^\dagger) \mathbf{U} = \mathbf{\Sigma} \mathbf{\Sigma}^\dagger = \left[\begin{array}{ccc|c} \mu_1^2 & & & \\ & \mu_2^2 & & \\ & & \ddots & \\ \bigcirc & & & \mu_M^2 \\ & & & \bigcirc \end{array} \right] \quad (1.9)$$

3. (Wikipedia) Une matrice de taille $M \times N$ est dite de rang plein si son rang est égal à $\min(M, N)$.

La matrice $\mathbf{H}^\dagger \mathbf{H}$ possède un maximum de M valeurs propres non nulles. Par conséquent et comme on s’y attendait, elle n’est pas inversible. L’Eq. (1.10) montre que les valeurs propres non nulles sont égales aux carrés des valeurs singulières de la matrice \mathbf{H} , comme pour la matrice $\mathbf{H}\mathbf{H}^\dagger$. Les valeurs propres non nulles de $\mathbf{H}^\dagger \mathbf{H}$ calculées numériquement sont donc également représentées sur la Fig. 1.18.

1.4.5 Analyse complémentaire de la matrice \mathbf{H}

Chaque donnée σ_m est constituée, en partie, d’une combinaison des éléments du vecteur $\mathbf{z} - \mathbf{z}_0$ pondérée par les coefficients de la m -ième ligne de \mathbf{H} . Pour cette raison, on propose une analyse de ses coefficients. Sur la Fig. 1.19 ci-dessous, on affiche la partie réelle et la partie imaginaire des coefficients d’une ligne de \mathbf{H} . La Fig. 1.19a représente les valeurs à la surface du cône, tandis que la Fig. 1.19b montre la valeur en fonction de l’altitude et de l’angle.

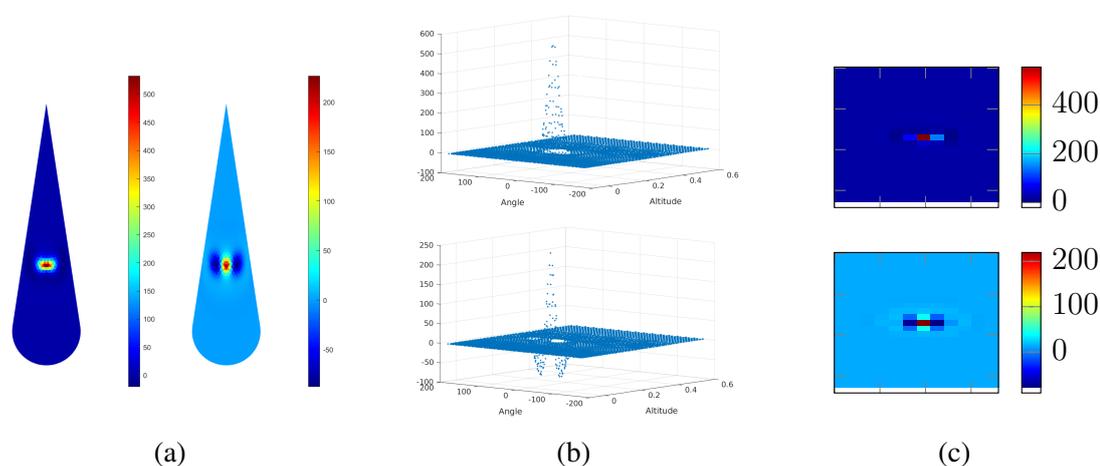


FIGURE 1.19 – Coefficients de la ligne 122 de \mathbf{H} représentés (a) à la surface du cône et (b) en fonction de l’altitude et de l’angle. La ligne 122 correspond à une position centrale. La figure (c) représente sous forme d’hologramme les coefficients de la colonne 2436 de \mathbf{H} , correspondant à une maille proche d’un des emplacements où sont effectuées les mesures. Les hologrammes sont fonction de l’angle en abscisse, et de l’altitude en ordonnée. Les quantités sont représentées en partie réelle à gauche sur (a) et en haut sur (b,c), et en partie imaginaire à droite sur (a) et en bas sur (b,c).

Les coefficients de la matrice \mathbf{H} correspondent à une fonction d’étalement du point et pour l’illustrer, on propose de visualiser les coefficients d’une colonne sur la Fig. 1.19c. Dans notre cas, la fonction d’étalement du point est non invariante et dépend de la position du capteur. Également, on observe sur les Fig. 1.19a et 1.19b que les données dépendent majoritairement de l’impédance locale. La forme de la fonction représentée sur la Fig. 1.19b nous montre que la combinaison pondérée des éléments de $\mathbf{z} - \mathbf{z}_0$ tend à dégrader l’information. Ainsi, reconstruire des informations locales à partir des données obtenues grâce au système d’observation est une tâche difficile. D’autant plus que le nombre de données disponible est très faible devant le nombre d’inconnues.

1.4.6 Retour au problème de contrôle

Dans ce chapitre nous avons présenté le problème du contrôle de propriétés matériaux. Les différentes considérations introduites nous ont mené à résumer ces propriétés par une impédance surfacique, ainsi qu'à un modèle d'observation linéaire à bruit additif donné par l'Eq. (1.7).

Ainsi, le problème du contrôle consiste à inverser le modèle d'observation pour fournir une reconstruction de l'impédance à partir des coefficients de réflexion mesurés. Pour réaliser cela, on dispose des informations suivantes : l'impédance est homogène par morceaux et elle évolue lentement autour d'une référence. Dans la suite, on considère alors conjointement le problème de l'inversion et celui de la reconstruction de région ou morceaux, et l'on parlera d'inversion-segmentation.

Le chapitre suivant traite du problème de l'inversion et de l'inversion-segmentation de modèle linéaire à bruit additif dans un cadre général. Notre méthode d'inversion-segmentation sera présentée dans le Chapitre 3.

CHAPITRE 2

INVERSION DE MODÈLE LINÉAIRE

Dans le chapitre précédent, nous avons montré comment les coefficients de réflexion sont déduits de l'impédance surfacique grâce à un modèle linéaire à bruit additif. On parle de modèle direct. Le problème inverse est celui de la reconstruction de l'impédance à partir des coefficients de réflexion. Deux éléments rendent l'inversion compliquée. D'une part, le nombre de données disponibles est inférieur au nombre d'inconnues et d'autre part, les données sont bruitées. Cependant, on sait que l'impédance surfacique est homogène par morceaux et qu'elle est centrée sur une référence. L'objectif de ce chapitre est de présenter le problème de l'inversion-segmentation. Pour cela, nous présentons les choses en deux temps. Tout d'abord, nous introduisons le problème de l'inversion de modèle linéaire à bruit additif dans un cadre générique et certaines solutions « historiques ». Puis, nous aborderons le problème de l'inversion-segmentation et les méthodes de la littérature.

Sommaire

2.1	Introduction	30
2.2	Inversion dans un cadre déterministe	30
2.2.1	Attache aux données et moindres carrés	30
2.2.2	Moindres carrés à norme minimale	31
2.2.3	Régularisation	31
2.3	Inversion dans un cadre probabiliste	32
2.3.1	Maximum de vraisemblance	32
2.3.2	Approche bayésienne	33
2.3.2.1	Maximiseur <i>a posteriori</i>	34
2.3.2.2	Espérance <i>a posteriori</i>	34
2.3.2.3	Conclusion sur l'approche bayésienne	35
2.4	Inversion-segmentation	35
2.4.1	Approches bayésiennes fondées sur le modèle de Potts	35
2.4.2	Analyse complémentaire	37
2.5	Conclusion	38

2.1 Introduction

Dans de nombreux domaines tels que l'imagerie médicale, la géophysique ou encore l'astronomie, on s'intéresse à la reconstruction d'une quantité d'intérêt \mathbf{x} au vu de mesures \mathbf{y} . Dans notre cas, il s'agit de déterminer l'impédance surfacique à partir de coefficients de réflexion mesurés.

Dans un cadre général, on considère que l'objet est observé au travers d'un opérateur linéaire \mathbf{A} et que les observations sont perturbées par un bruit additif \mathbf{b} . Le modèle s'écrit alors $\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{b}$ et il est représenté sur la Fig. 2.1. Le terme de bruit regroupe plusieurs choses. Cela peut-être des erreurs de mesures introduites par un moyen de mesure imprécis, des perturbations liées à l'environnement dans lequel les mesures sont réalisées, des erreurs dans le modèle qui ne décrit pas totalement la réalité physique du phénomène en jeu, ou encore des erreurs de discrétisation des équations.

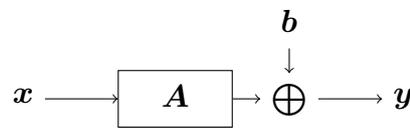


FIGURE 2.1 – Schéma du modèle direct linéaire à bruit additif.

Afin de résoudre le problème inverse associé, il existe plusieurs méthodes. Dans la suite, on présente des méthodes représentatives de deux grandes familles, les méthodes d'inversion dans un cadre déterministe et les méthodes d'inversion dans un cadre probabiliste.

2.2 Inversion dans un cadre déterministe

2.2.1 Attache aux données et moindres carrés

Les méthodes déterministes visent généralement à minimiser, par rapport à \mathbf{x} , un critère nommé terme d'attache aux données qui rend compte d'une « distance » entre la sortie-modèle $\mathbf{A}\mathbf{x}$ et les données \mathbf{y} . Dans de nombreux cas [2, 29, 30], on le choisit convexe, par exemple une norme. Ainsi, on assure l'existence d'un unique minimum et la minimisation peut s'appuyer sur la large littérature dédiée à l'optimisation convexe.

Parmi les critères les plus populaires, on retrouve le critère des moindres carrés

$$\mathcal{J}_{MC}(\mathbf{x}) = \|\mathbf{y} - \mathbf{A}\mathbf{x}\|^2, \quad (2.1)$$

dont le minimiseur, s'il existe, s'écrit $\arg \min_{\mathbf{x}} \mathcal{J}_{MC}(\mathbf{x})$. L'annulation du gradient nous mène à considérer $\mathbf{A}^\dagger \mathbf{A}\mathbf{x} = \mathbf{A}^\dagger \mathbf{y}$ et dans le cas où $\mathbf{A}^\dagger \mathbf{A}$ est inversible, le minimiseur s'écrit

$$\hat{\mathbf{x}} = \arg \min_{\mathbf{x}} \mathcal{J}_{MC}(\mathbf{x}) = (\mathbf{A}^\dagger \mathbf{A})^{-1} \mathbf{A}^\dagger \mathbf{y}. \quad (2.2)$$

Le calcul de la solution (2.2) dépend de l'inversibilité ou du conditionnement de la matrice $\mathbf{A}^\dagger \mathbf{A}$ et l'on se propose de discuter de ces deux cas de figure.

Cas non inversible – Calculer la solution (2.2) n’est pas toujours possible. Il arrive que le nombre de données disponibles soit inférieur au nombre d’inconnues. Le problème est alors sous-déterminé : la matrice $\mathbf{A}^\dagger \mathbf{A}$ est non inversible, et il existe une infinité de solutions cohérentes avec les données qui minimisent et annulent $\mathcal{J}_{MC}(\mathbf{x})$. C’est d’ailleurs notre cas, voir la partie 1.4.4.

Cas inversible – Même si la matrice $\mathbf{A}^\dagger \mathbf{A}$ est inversible, il est possible que celle-ci soit mal conditionnée. Le calcul de la solution des moindres carrés, Eq. (2.2), peut alors entraîner « l’explosion » de la solution [31, p. 37-38]. Afin de servir d’exemple, on considère le problème de la déconvolution d’une image floutée et bruitée dans le cas particulier où la matrice $\mathbf{A}^\dagger \mathbf{A}$ est inversible, mais mal conditionnée. Dans cet exemple, l’information pour certaines composantes fréquentielles est fortement atténuée. La reconstruction au sens des moindres carrés mène alors à une forte amplification de ces mêmes fréquences, entraînant « l’explosion » de la solution [31, p. 91].

2.2.2 Moindres carrés à norme minimale

Dans le cas non inversible, on sait qu’il existe une infinité de solutions qui minimisent et annulent le critère des moindres carrés. Parmi ces solutions, la méthode des moindres carrés à norme minimale consiste à sélectionner celle dont la norme $\|\mathbf{x}\|^2$ est minimale. Elle s’écrit

$$\hat{\mathbf{x}} = \begin{cases} \arg \min_{\mathbf{x}} \|\mathbf{x}\|^2, \\ \mathbf{y} = \mathbf{A}\mathbf{x} \end{cases} = \mathbf{A}^\dagger (\mathbf{A}\mathbf{A}^\dagger)^{-1} \mathbf{y}. \quad (2.3)$$

Son calcul nécessite que $\mathbf{A}\mathbf{A}^\dagger$ soit inversible et correctement conditionnée. Dans notre cas, $\mathbf{H}\mathbf{H}^\dagger$ est inversible, mais mal conditionnée, voir partie 1.4.4. La solution au sens des moindres carrés à norme minimale peut donc être calculée, mais elle est instable numériquement et l’on ne la considèrera pas.

2.2.3 Régularisation

Lorsque de l’information à propos de \mathbf{x} est disponible, on souhaite en tenir compte et une manière de faire cela est d’introduire une pénalité $\mathcal{P}(\mathbf{x})$. Comme pour le terme d’attache aux données, la pénalité peut prendre diverses formes dont celle d’une norme. Souvent, l’information prise en compte assure l’unicité de la solution et limite les phénomènes d’explosion.

Il est courant de retrouver dans la littérature des pénalités quadratiques. Par exemple la régularisation de Tikhonov [32,33], pour laquelle $\mathcal{P}(\mathbf{x}) = \|\mathbf{\Gamma}\mathbf{x}\|^2$. Le choix de la matrice $\mathbf{\Gamma}$ permet alors d’introduire diverses informations, par exemple des informations de textures, une pénalité sur la quantité d’intérêt $\mathcal{P}(\mathbf{x}) = \|\mathbf{x}\|^2$ afin de favoriser des solutions de faible énergie, ou encore une pénalité sur une mesure de rugosité [31, p. 47] qui favorise des solutions douces $\mathcal{P}(\mathbf{x}) = \|\mathbf{D}\mathbf{x}\|^2$, avec \mathbf{D} est une matrice des différences finies. La matrice \mathbf{D} est utilisée dans notre méthode et sera introduite dans le chapitre suivant.

L’ajout d’une pénalité définit un nouveau critère et fait apparaître un hyperparamètre $\lambda \geq 0$

$$\mathcal{J}_{MCP}(\mathbf{x}) = \|\mathbf{y} - \mathbf{A}\mathbf{x}\|^2 + \lambda \mathcal{P}(\mathbf{x}), \quad (2.4)$$

qui règle le compromis entre le terme d'attache aux données et la pénalité. Plus λ est grand, plus la solution est pénalisée et inversement. Le réglage des hyperparamètres est une question importante qui est discutée dans [31, p. 53-56] et sera évoquée dans le chapitre suivant.

Exemple 1. Dans le cas de la régularisation de Tikhonov $\mathcal{P}(\mathbf{x}) = \|\Gamma\mathbf{x}\|^2$, si la solution existe, elle s'écrit

$$\begin{aligned}\hat{\mathbf{x}} &= \arg \min_{\mathbf{x}} \|\mathbf{y} - \mathbf{A}\mathbf{x}\|^2 + \lambda \|\Gamma\mathbf{x}\|^2 \\ &= (\mathbf{A}^\dagger \mathbf{A} + \lambda \Gamma^\dagger \Gamma)^{-1} \mathbf{A}^\dagger \mathbf{y}\end{aligned}\tag{2.5}$$

et nous la retrouvons dans notre méthode comme solution intermédiaire.

Pour certaines applications, l'information disponible sur \mathbf{x} ne peut être prise totalement en compte par une pénalité quadratique. D'autres pénalités sont alors exploitées, par exemple des pénalités du type L_2/L_1 et L_2/L_0 [31, p. 148-152], ou encore une contrainte, par exemple la positivité. Les pénalités L_2/L_1 sont convexes, ce qui assure la convergence des techniques standards de minimisation et fournit une certaine « robustesse » à la solution [31, p. 47]. Parmi les pénalités L_2/L_1 , on peut citer la pénalité de Huber [34] qui favorise des solutions douces, mais qui pénalise moins les fortes variations que la pénalité quadratique et permet ainsi de mieux reconstruire les discontinuités. En début de thèse, j'ai mené un travail préliminaire s'appuyant sur la pénalité de Huber. Ce travail est détaillé en annexe G et a été publié [2].

2.3 Inversion dans un cadre probabiliste

Dans cette partie, on introduit une modélisation stochastique pour le terme de bruit. On choisit alors une loi de probabilité dite *a priori*, notée $f(\mathbf{b})$, qui décrit les connaissances que l'on a sur les erreurs \mathbf{b} . En s'appuyant sur le modèle direct : $\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{b}$ et sur le modèle de bruit, on déduit la loi des données \mathbf{y} sachant \mathbf{x}

$$p(\mathbf{y}|\mathbf{x}) = f(\mathbf{y} - \mathbf{A}\mathbf{x}),\tag{2.6}$$

appelée vraisemblance des inconnues \mathbf{x} attachée aux données \mathbf{y} . On précise que l'écriture (2.6) sous-entend que \mathbf{x} est probabilisé. Le choix de la distribution du bruit est une question délicate, particulièrement lorsque peu d'informations sur celui-ci sont disponibles. À l'inverse, il existe des cas dans lesquels les erreurs sont bien caractérisées et le choix d'une loi pour le bruit est alors plus simple, par exemple, dans le cas d'erreurs systématiques connues.

2.3.1 Maximum de vraisemblance

La méthode du maximum de vraisemblance s'appuie directement sur la loi $p(\mathbf{y}|\mathbf{x})$ et définit l'estimée $\hat{\mathbf{x}}_{MV}$, si elle existe, comme maximiseur de la vraisemblance

$$\hat{\mathbf{x}}_{MV} = \arg \max_{\mathbf{x}} p(\mathbf{y}|\mathbf{x}).\tag{2.7}$$

On retrouve des problématiques similaires à celles abordées pour les moindres carrés : non-inversibilité et mauvais conditionnement. Maximiser $p(\mathbf{y}|\mathbf{x})$ est plus ou moins compliqué selon la loi pour le bruit. Dans l'exemple 2 classique ci-dessous, la vraisemblance est log-concave dérivable et l'expression de l'estimée est obtenue, si elle existe, en annulant le gradient.

Exemple 2. Lorsque le bruit est gaussien : $\mathbf{b} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, la vraisemblance s'écrit

$$p(\mathbf{y}|\mathbf{x}) = \mathcal{N}(\mathbf{y}; \mathbf{A}\mathbf{x} + \boldsymbol{\mu}, \boldsymbol{\Sigma}). \quad (2.8)$$

Le maximiseur de vraisemblance, s'il existe, s'écrit

$$\begin{aligned} \hat{\mathbf{x}}_{MV} &= \arg \max_{\mathbf{x}} p(\mathbf{y}|\mathbf{x}) \\ &= \arg \min_{\mathbf{x}} \|(\mathbf{y} - \boldsymbol{\mu}) - \mathbf{A}\mathbf{x}\|_{\boldsymbol{\Sigma}^{-1}}^2 \\ &= (\mathbf{A}^\dagger \boldsymbol{\Sigma}^{-1} \mathbf{A})^{-1} \mathbf{A}^\dagger \boldsymbol{\Sigma}^{-1} (\mathbf{y} - \boldsymbol{\mu}). \end{aligned} \quad (2.9)$$

La solution des moindres carrés présentée dans la partie précédente, Eq. (2.2), est en réalité un cas particulier du maximiseur de vraisemblance avec $\boldsymbol{\mu} = \mathbf{0}$ et $\boldsymbol{\Sigma} \propto \mathbf{I}$.

Pour des cas plus complexes, des méthodes d'optimisation existent. Par exemple l'algorithme de recuit simulé [35, p. 163-169] qui fournit le maximiseur à convergence, les algorithmes semi-quadratiques ou encore les algorithmes de gradient. Enfin, lorsque le problème est sous déterminé, cela mène à une indétermination et une stratégie reposant sur le maximum de vraisemblance n'est alors pas envisageable, ce qui est notre cas.

2.3.2 Approche bayésienne

Le modèle de vraisemblance décrit la distribution des données en fonction des inconnues. Lorsque l'on dispose d'informations sur les inconnues, il est possible de les modéliser au travers d'une loi *a priori* pour \mathbf{x} que l'on note $p(\mathbf{x})$. Ainsi, à partir du modèle de vraisemblance et de la loi *a priori*, on applique la règle de Bayes pour obtenir la loi *a posteriori*, *i.e.*, la distribution des inconnues compte tenu de l'information *a priori* et des données \mathbf{y} :

$$p(\mathbf{x}|\mathbf{y}) = \frac{p(\mathbf{y}, \mathbf{x})}{p(\mathbf{y})} = \frac{p(\mathbf{y}|\mathbf{x})p(\mathbf{x})}{\int_{\mathbf{x}} p(\mathbf{y}|\mathbf{x})p(\mathbf{x}) d\mathbf{x}}. \quad (2.10)$$

Le dénominateur est la loi marginale pour les données et elle intervient comme une constante multiplicative par rapport aux inconnues. Elle est souvent difficile à obtenir. Néanmoins, des méthodes exploitent le fait qu'il n'est pas nécessaire de la connaître, par exemple des méthodes d'optimisation ou encore certaines méthodes d'intégration de Monte Carlo. Il existe cependant des cadres dans lesquels il est nécessaire de la connaître, par exemple pour de la sélection de modèles [36–38].

À partir de la loi jointe, la construction d'un estimateur peut être réalisée de plusieurs manières. Pour les besoins de la suite, on note $\boldsymbol{\Delta}(\mathbf{y}) = \hat{\mathbf{x}}$ l'estimateur de \mathbf{x} et $\mathcal{L}(\mathbf{x}, \boldsymbol{\delta})$ une fonction de coût. On rappelle que le risque bayésien [39, p. 43] [35, p. 12-13] s'écrit

$$\mathbb{E}_{\mathbf{X}, \mathbf{Y}}[\mathcal{L}(\mathbf{X}, \boldsymbol{\Delta}(\mathbf{Y}))] = \mathbb{E}_{\mathbf{Y}}[\mathbb{E}_{\mathbf{X}|\mathbf{Y}}[\mathcal{L}(\mathbf{X}, \boldsymbol{\Delta}(\mathbf{Y}))]], \quad (2.11)$$

avec $\mathbb{E}_{\mathbf{X}|\mathbf{Y}=\mathbf{y}}[\mathcal{L}(\mathbf{X}, \boldsymbol{\Delta}(\mathbf{y}))]$ le risque *a posteriori*. Afin de minimiser le risque bayésien, on peut minimiser le risque *a posteriori*. Dans la suite, on discute du Maximiseur *a posteriori* (MAP) et de l'Espérance *a posteriori* (EAP).

2.3.2.1 Maximiseur *a posteriori*

Si on considère la fonction de coût

$$\mathcal{L}(\mathbf{x}, \boldsymbol{\delta}) = \begin{cases} 0, & \text{si } \|\mathbf{x} - \boldsymbol{\delta}\| < s \\ 1, & \text{sinon,} \end{cases} \quad (2.12)$$

avec s qui tend vers 0, alors l'estimateur approche le MAP à condition que la distribution *a posteriori* soit quasi-concave [40]. Le MAP est la valeur de \mathbf{x} qui maximise la loi *a posteriori* :

$$\hat{\mathbf{x}}_{MAP} = \arg \max_{\mathbf{x}} p(\mathbf{x}|\mathbf{y}). \quad (2.13)$$

Il peut être calculé en s'appuyant sur les nombreux outils d'optimisation disponibles dans la littérature : descente de gradient, recuit simulé, *etc.* De nombreuses circonstances peuvent rendre cette opération compliquée : loi multimodale, loi discrète, *etc.*

Puisque $\hat{\mathbf{x}}_{MAP}$ maximise la loi *a posteriori*, on peut dire qu'il s'agit d'un estimateur cohérent, qui a du sens. Cependant, il n'est pas toujours un estimateur de choix. Par exemple, dans le cas multimodal, le maximum peut être atteint dans un mode qui ne contient qu'une fraction minime de la masse de probabilité. Le MAP n'est alors pas une bonne représentation de la loi *a posteriori* et, en ce sens, l'EAP peut être un estimateur plus pertinent.

2.3.2.2 Espérance *a posteriori*

Lorsque l'on minimise le risque bayésien pour une fonction de coût quadratique

$$\mathcal{L}(\mathbf{x}, \boldsymbol{\delta}) = \|\mathbf{x} - \boldsymbol{\delta}\|^2, \quad (2.14)$$

l'estimateur optimal est l'EAP :

$$\hat{\mathbf{x}}_{EAP} = \mathbb{E}_{\mathbf{X}|\mathbf{Y}=\mathbf{y}}[\mathbf{X}]. \quad (2.15)$$

Pour obtenir l'EAP, si l'intégration analytique n'est pas possible, par exemple à cause de la grande dimension ou de la forme non standard des lois, on peut s'appuyer sur des méthodes numériques. Parmi les méthodes numériques, on peut citer les méthodes de quadrature qui consistent à discrétiser le domaine en morceaux plus ou moins réguliers puis à calculer une approximation de l'intégrale sur ces morceaux. L'intégrale totale est ensuite obtenue en sommant les résultats de chaque morceau. Simple et rapide à mettre en œuvre pour des cas de faible dimension, cette approche n'est pas envisageable en grande dimension puisque le nombre de points nécessaires pour maintenir une précision croît de manière exponentielle avec la dimension. Une autre classe de méthodes qui permettent de répondre à la difficulté d'intégration est celle des algorithmes Monte Carlo par chaînes de Markov. Ce sont les méthodes que nous utilisons et qui sont discutées plus loin dans le document.

Exemple 3. Dans le cas particulier tout gaussien suivant : $p(\mathbf{x}) = \mathcal{N}(\mathbf{x}; \mathbf{0}, (\gamma_x \mathbf{P}_x)^{-1})$ et $f(\mathbf{b}) = \mathcal{N}(\mathbf{b}; \mathbf{0}, \gamma_b^{-1} \mathbf{I})$, l'espérance *a posteriori* est la solution des moindres carrés pénalisés avec la pénalité $\mathcal{P}(\mathbf{x}) = \|\mathbf{x}\|_{\mathbf{P}_x}^2$ et l'hyperparamètre $\lambda = \gamma_x/\gamma_b$:

$$\hat{\mathbf{x}}_{EAP} = (\mathbf{A}^\dagger \mathbf{A} + \frac{\gamma_x}{\gamma_b} \mathbf{P}_x)^{-1} \mathbf{A}^\dagger \mathbf{y}. \quad (2.16)$$

Si $P_x = \Gamma^\dagger \Gamma$, alors cette solution correspond à la solution (2.5). La matrice P_x est appelée matrice de précision. Sa structure s'appuie sur les relations entre les composantes de x . Lorsque ces relations s'écrivent uniquement à partir de relations entre voisins « proches », la matrice P_x est creuse et x est un champ de Markov [31, p. 169-177] [41]. C'est la modélisation que l'on retrouve dans le chapitre suivant.

L'EAP est un estimateur cohérent, mais il existe des cas dans lesquels il n'est pas un bon estimateur. Prenons l'exemple d'une loi bimodale pour lequel chaque mode concentre une part égale de la masse de probabilité. Dans cette configuration, l'EAP peut se trouver dans une zone de faible probabilité et n'est pas forcément un bon estimateur.

2.3.2.3 Conclusion sur l'approche bayésienne

L'approche bayésienne apporte un cadre solide et rigoureux au problème de l'inversion. Les possibilités de prendre en compte de l'information au travers de loi *a priori* sont riches. La loi *a posteriori* nous renseigne sur l'ensemble des masses de probabilités et synthétise l'information disponible *a posteriori* à propos des inconnus. De là, on peut obtenir une solution ponctuelle : MAP, EAP, *etc*, mais également d'autres quantités comme des incertitudes.

2.4 Inversion-segmentation

L'inversion classique, dont nous venons de parler, fournit une solution au problème de la reconstruction d'une quantité d'intérêt x à partir de données y . Cependant, il arrive que la quantité d'intérêt soit composée de régions et que l'on souhaite en tenir compte. Les cas d'application sont variés, par exemple en imagerie spatiale pour dissocier plusieurs régions de l'image, en imagerie médicale pour détecter des éléments d'intérêt, ou encore en contrôle non destructif afin de vérifier l'intégrité d'une pièce. Ce problème joint de l'inversion et de la segmentation est représenté par le modèle direct de la Fig. 2.2, où ℓ sont des variables cachées décrivant des régions. L'objectif est d'inverser ce modèle direct afin de reconstruire conjointement x et ℓ , et on parle alors d'inversion-segmentation.

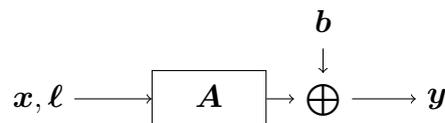


FIGURE 2.2 – Schéma du modèle direct pour l'inversion-segmentation.

Les méthodes présentées précédemment ne permettent pas d'y répondre directement. En effet, bien qu'elles fournissent la plupart du temps des solutions convenables, elles ne réalisent pas de réelle segmentation. Nous discutons dans la suite de ce problème joint en considérant l'aspect segmentation comme partie intégrante du problème.

2.4.1 Approches bayésiennes fondées sur le modèle de Potts

Parmi la littérature sur l'inversion-segmentation, mon analyse s'est principalement portée sur des travaux utilisant une approche bayésienne et un jeu d'étiquettes ℓ , modélisé par un champ

de Potts. Le champ de Potts est une généralisation du modèle de Ising (1925) qui fut proposée par Renfrey Potts dans ses travaux de thèse de doctorat en 1951 [42] et par la suite, démocratisée par Geman et Geman [43] pour le traitement du signal et de l'image. Le champ de Potts est un champ de Markov qui introduit du lien entre éléments (pixels, mailles, *etc*) voisins. Pour une image par exemple, il décrit la dépendance spatiale. Il favorise alors des régions compactes dont la taille moyenne est pilotée par un paramètre appelé coefficient de granularité. Dans la suite, on présente certains travaux desquels s'inspire notre travail. Des similitudes existent entre eux et, pour en faire ressortir les spécificités, on concentre l'analyse sur certains points saillants.

Structure *a priori* et *a posteriori* – La modélisation de l'objet x et des étiquettes ℓ mène à introduire des modèles *a priori* conditionnels $x|\ell$, par exemple gaussien [44, 45], Rayleigh [46], Dirichlet [47] ou encore gaussiens généralisés [48]. La structure de ces modèles est cruciale puisqu'elle détermine, entre autres, la structure de la loi *a posteriori* des étiquettes et de l'objet. Pour l'illustrer, on présente les conséquences de différentes structures pour ℓ puis pour x .

1. La structure de la loi *a posteriori* pour ℓ dépend de la structure *a priori* $x|\ell$.
 - Si $x|\ell$ est séparable [26, 49, 50], la loi *a posteriori* pour ℓ est un champ de Markov.
 - Si $x|\ell$ est markovienne [51, 52], la loi *a posteriori* pour ℓ l'est également.
 - Si $x|\ell$ est non séparable et non markovienne [53], la loi *a posteriori* pour ℓ est non séparable et non markovienne.

2. La structure de la loi *a posteriori* pour x dépend de la vraisemblance et de $x|\ell$.
 - Si la vraisemblance et la loi *a priori* $x|\ell$ sont toutes les deux séparables, la loi *a posteriori* pour x est séparable [47].
 - Si la vraisemblance ou la loi *a priori* $x|\ell$ est non séparable, la loi *a posteriori* pour x ne l'est pas [54].

Ce dernier point peut rendre l'exploitation de la loi *a posteriori* difficile, d'autant plus si elle est de forme non standard [48, 55]. Pour cette raison on préfère considérer des lois *a priori* qui mènent à des lois *a posteriori* de forme standard, par exemple gaussiennes [52, 53, 56].

Une spécificité de [53] est liée à la modélisation de x sachant ℓ . Celle-ci repose sur un modèle déterministe fonction des étiquettes et de plusieurs objets inconnus x_1, x_2, \dots correspondant aux différentes régions. Ces derniers sont alors probabilisés à la place de x , ce qui mène à une hiérarchie différente pour laquelle les étiquettes apparaissent dans le modèle de vraisemblance. La structure de la loi *a posteriori* pour ℓ dépend alors de la structure de la vraisemblance et l'on retrouve les 3 cas de figure cités au point (1.) ci-dessus, avec « vraisemblance » à la place de « $x|\ell$ ». Par ailleurs, c'est cette modélisation que nous adoptons pour notre méthode d'inversion-segmentation.

Estimateurs et algorithmes – On ne sait pas calculer explicitement les estimateurs à partir de la loi *a posteriori* et l'on a recours à des méthodes numériques. Le choix des estimateurs est alors important puisqu'il détermine souvent l'algorithme utilisé ainsi que la qualité des reconstructions.

- Pour les étiquettes ℓ , l'EAP n'a pas de sens et l'on choisit généralement le MAP. Son obtention par des algorithmes tels que l'ICM (Iterated Conditional Modes) [54, 57] ou le recuit simulé [35, p. 163-169] peut être difficile et l'on se contente souvent du MMAP (Maximiseur Marginal *a posteriori*) [44], plus simple à obtenir.

- Pour l'objet x , l'EAP est généralement choisie et il existe deux approches. La première considère l'EAP de la loi *a posteriori* marginalisée par rapport aux étiquettes, ce qui revient à considérer l'espérance d'un mélange de lois pondérées par les probabilités *a posteriori* pour les étiquettes. La deuxième s'appuie sur l'EAP conditionnellement à la configuration des étiquettes la plus probable (MAP ou MMAP), qui ne fait pas de mélange. Obtenir le MAP pour x peut être délicat. [58] propose une approche fondée sur une segmentation puis une régularisation quadratique tirant parti de la segmentation pour tenter d'obtenir le MAP pour ℓ et x .

Parmi les méthodes numériques les plus courantes, on retrouve essentiellement trois classes.

1. Les méthodes MCMC (Monte Carlo par chaînes de Markov) et en particulier l'algorithme de Gibbs qui permet d'exploiter les structures *a posteriori*, séparables ou de forme standard dont nous venons de parler. Par exemple, lorsque la loi *a posteriori* des étiquettes est un champ de Potts, des stratégies reposant sur la propriété markovienne permettent de paralléliser en partie l'échantillonnage [59, 60]. Les algorithmes MCMC peuvent être coûteux en temps de calcul et nécessiter un grand nombre d'itérations.

2. Les méthodes du type ICM (Iterated Conditional Modes) qui sont des algorithmes d'optimisation locaux et qui tirent également parti des structures *a posteriori*, séparables ou de forme standard (optimisation parallélisée pour les étiquettes, *etc*). Ces méthodes, plus rapides dans le cas de [54], peuvent ne fournir qu'un optimum local.

3. Les méthodes bayésiennes variationnelles utilisées dans [25, 61] qui fournissent une approximation de loi *a posteriori* sous la forme de lois paramétriques séparables relativement simples à exploiter. L'obtention de ces lois nécessite des calculs que l'on ne sait faire explicitement que pour des lois de la famille exponentielle. On ajoute qu'il est difficile de quantifier la qualité de l'approximation de la loi *a posteriori*.

2.4.2 Analyse complémentaire

La partie précédente s'est concentrée sur l'analyse de travaux d'inversion-segmentation proches du nôtre, c'est-à-dire des approches bayésiennes reposant sur des étiquettes modélisées par un champ de Potts. Néanmoins, d'autres travaux répondent au même problème avec des méthodes différentes ou répondent à un problème connexe grâce à l'utilisation du champ de Potts. Nous proposons alors d'en présenter une partie.

L'utilisation d'une régularisation adaptée permet également de reconstruire des objets par morceaux. Dans cette veine, on peut citer [62, 63], qui s'appuient sur une modélisation variationnelle du champ de Potts. Celle-ci prend alors la forme d'une pénalité L_0 du type $\|\nabla x\|_0$, appelée « jump term », qui favorise des solutions constantes par morceaux. Cette stratégie d'inversion correspond à celle présentée dans la section 2.2.3, lorsque $\mathcal{P}(x) = \|Dx\|_0$. Néanmoins, l'utilisation d'une pénalité L_0 non convexe mène à des difficultés d'optimisation du critère dues à la présence d'optima locaux. On peut également citer [64], dont le travail repose sur des régularisations quadratiques tronquées L_2/L_0 afin de favoriser des solutions douces par morceaux. Enfin,

dans [65], les auteurs proposent une méthode favorisant des objets constants par morceaux et des morceaux compacts grâce à l'utilisation de pénalités L_1 .

Comme expliqué dans la partie précédente, le champ de Potts modélise des régions. Il trouve donc naturellement sa place dans les problèmes de segmentation [66,67] ou de classification [68]. On cite également [69, p. 169-200] dans lequel le champ de Ising est utilisé pour modéliser l'activation de région 3D.

Par ailleurs, le champ de Potts fait intervenir un paramètre de granularité pilotant la taille moyenne des régions. Des travaux visant à l'estimer existent, mais cette tâche est rendue difficile à cause de la fonction de partition du champ. En effet, cette dernière nécessite de sommer un nombre de termes qui croît de manière exponentielle avec la taille de l'objet. Plusieurs solutions sont proposées dans la littérature pour surmonter cette difficulté, par exemple le précalcul de la fonction de partition pour une fine grille de valeur [53, 70] ou l'échantillonnage du coefficient de granularité par une étape de « ABC Likelihood-Free MH » [71] qui ne nécessite pas le calcul de la fonction de partition.

Enfin, plus récemment, on peut citer la méthode de segmentation/classification proposée par [72, 73] reposant sur l'utilisation de l'équation différentielle d'Allen-Cahn.

2.5 Conclusion

Dans la première partie du chapitre, nous avons présenté des méthodes classiques utilisées pour l'inversion de modèle direct linéaire. Néanmoins, celles-ci ne répondant pas au problème joint de l'inversion et de la segmentation, nous avons présenté différentes approches possibles. L'analyse s'est concentrée sur les méthodes bayésiennes utilisant des étiquettes modélisées par un champ de Potts et sur les conséquences des structures des lois *a priori* sur celles des lois *a posteriori*. Différents estimateurs et méthodes numériques ont également été présentés. Il en ressort que les choix concernant les modèles *a priori* et les estimateurs déterminent souvent les méthodes numériques choisies, ou inversement. Une seconde partie, dédiée à une analyse complémentaire de la bibliographie, présente des travaux reposant sur des méthodes différentes ou traitant de problèmes connexes, par exemple l'estimation du coefficient de granularité.

Dans le chapitre suivant, on présente notre méthode d'inversion-segmentation appliquée à la reconstruction de l'impédance homogène par morceaux.

CHAPITRE 3

MÉTHODE D'INVERSION-SEGMENTATION

Ce chapitre présente la méthode d'inversion-segmentation développée pour répondre au problème du contrôle EM. On rappelle que l'on souhaite inverser le modèle direct défini au Chapitre 1 : $\sigma = \sigma_0 + \mathbf{H}(z - z_0) + \varepsilon$ et fournir une solution homogène par morceaux. La méthode proposée s'appuie sur la modélisation aléatoire d'impédances homogènes et d'un jeu d'étiquettes qui décrit des régions à la surface de l'objet. Les modèles a priori choisis sont fonction d'hyperparamètres que l'on estime, au moins en partie, conjointement aux impédances et étiquettes. La fin du chapitre nous mène à définir le modèle hiérarchique ainsi que la loi a posteriori. On présente également les estimateurs choisis pour reconstruire les inconnues et une manière de quantifier les incertitudes associées.

Sommaire

3.1	Champ d'impédance par morceaux	40
3.1.1	Mailles, voisinage, projections et troncatures	40
3.1.2	Champ d'étiquette	42
3.1.3	Construction du champ d'impédance	43
3.1.3.1	Champs complets	43
3.1.3.2	Champ par morceaux	45
3.2	Modèle pour le bruit et vraisemblance	49
3.3	Hyperparamètres	50
3.4	Modèle hiérarchique et loi a posteriori	50
3.4.1	Loi jointe	51
3.4.2	Loi a posteriori	51
3.5	Estimateurs	52
3.5.1	Étiquettes	52
3.5.2	Impédance et hyperparamètres	53
3.5.3	Difficultés calculatoires	54

3.1 Champ d'impédance par morceaux

On rappelle que l'impédance surfacique modélise les propriétés matériaux, que ces propriétés sont homogènes par morceaux et qu'elles évoluent lentement autour d'une référence. Cela veut dire qu'il existe des régions à la surface de l'objet dans lesquelles l'impédance évolue lentement. On dit que l'impédance est douce par morceaux. Pour modéliser cela, on présente une construction de l'impédance qui s'appuie sur plusieurs impédances homogènes complètes, c'est-à-dire sur l'ensemble du cône, ainsi que sur un jeu d'étiquettes. D'un côté, les différentes impédances modélisent l'aspect homogène pour chaque région, et de l'autre, les étiquettes modélisent les régions et les discontinuités.

3.1.1 Mailles, voisinage, projections et troncatures

On considère un maillage composé de N mailles triangulaires numérotées de manière arbitraire de 1 à N par un indice n . On attribue à chaque maille d'indice n une étiquette notée $\ell_n \in \{1 \dots K\}$ et on note $\ell = [\ell_1 \dots \ell_N]^t$, la carte complète des étiquettes. Les valeurs de 1 à K correspondent aux différentes propriétés EM que l'on cherche à caractériser et dans la suite, on les appelle des classes. Les mailles dont les étiquettes ont la même valeur, par exemple k , forment une région notée $\mathcal{C}_k = \{n \in \mathbb{N}_*^N \text{ tel que } \ell_n = k\}$ pour laquelle les propriétés EM sont homogènes. Attribuer une étiquette à chacune des mailles permet alors le découpage de la surface en région, mais également d'identifier la classe des régions, c'est-à-dire leurs propriétés. De plus, on note $N_k(\ell)$ le nombre d'étiquettes égales à k :

$$N_k(\ell) = \sum_{n=1}^N \delta(\ell_n, k), \text{ avec } \delta(u, v) = \begin{cases} 1 & \text{si } u = v, \\ 0 & \text{sinon.} \end{cases} \quad (3.1)$$

Remarque 8. *Le maillage représente la surface du cône sphère, Fig. 1.13 page 21, mais n'est pas spécifique à cet objet et peut s'appliquer à d'autres formes. Les développements proposés restent valides pour des travaux futurs où l'objet serait différent.*

Pour les besoins de la suite, on introduit pour chaque classe k , les matrices binaires de projection et de troncature correspondantes.

Matrice de projection – On définit

$$\mathbf{P}_k(\ell) = \text{diag}\{\delta(\ell_n, k), n = 1 \dots N\}, \quad (3.2)$$

la matrice diagonale binaire de projection de taille $N \times N$. Son application à un vecteur de taille N force à zéro les composantes d'indices n pour lesquelles $\ell_n \neq k$ et laisse les autres inchangées. Puisque chaque étiquette prend une valeur et une seule dans $\{1 \dots K\}$:

$$\sum_{k=1}^K \mathbf{P}_k(\ell) = \mathbf{I}_N. \quad (3.3)$$

Matrice de troncature – On définit également $\mathbf{T}_k(\ell)$, la matrice binaire de troncature de taille $N_k(\ell) \times N$ qui s'obtient à partir de $\mathbf{P}_k(\ell)$ en enlevant les lignes nulles. Sa transposée $\mathbf{T}_k^t(\ell)$,

de taille $N \times N_k(\ell)$, est une matrice de bourrage. Appliquée à un vecteur de taille $N_k(\ell)$, elle range les composantes dans un vecteur nul de taille N aux indices n pour lesquels $\ell_n = k$. Appliquer la matrice de troncature, puis la matrice de bourrage revient à appliquer la matrice de projection, c'est-à-dire extraire les composantes de la classe considérée et compléter par des zéros aux bons indices. On a donc

$$\mathbf{T}_k^t(\ell)\mathbf{T}_k(\ell) = \mathbf{P}_k(\ell). \quad (3.4)$$

À l'inverse, bourrer par des zéros, puis extraire les composantes de la classe considérée, revient à laisser le vecteur inchangé. On a donc

$$\mathbf{T}_k(\ell)\mathbf{T}_k^t(\ell) = \mathbf{I}_{N_k(\ell)}. \quad (3.5)$$

Pour $k \neq k'$, appliquer la matrice de bourrage $\mathbf{T}_{k'}^t(\ell)$, puis la matrice de troncature $\mathbf{T}_k(\ell)$ revient à bourrer par des zéros, puis extraire des composantes nulles. On a donc

$$\mathbf{T}_k(\ell)\mathbf{T}_{k'}^t(\ell) = \mathbf{0}_{N_k(\ell) \times N_{k'}(\ell)}. \quad (3.6)$$

Exemple 4. Exemple de matrices de troncature et de projection dans le cas où $N = 4$ et où les indices des mailles appartenant à la classe k sont $[2, 3, 4]$:

$$\mathbf{T}_k(\ell) = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$\mathbf{T}_k^t(\ell) = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\mathbf{T}_k^t(\ell)\mathbf{T}_k(\ell) = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = \mathbf{P}_k(\ell)$$

$$\mathbf{T}_k(\ell)\mathbf{T}_k^t(\ell) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \mathbf{I}_{N_k(\ell)}$$

Maille et voisinage – Dans la suite, on introduit des modèles qui s'appuient sur un système de voisinage et on le définit ici. On rappelle que les mailles sont triangulaires. Pour chaque maille, on considère un ensemble de mailles dites voisines et l'on choisit qu'une maille est voisine d'une autre si elle partage une arête avec elle. Pour notre maillage, toutes les mailles possèdent exactement 3 mailles voisines, comme illustré par la Fig. 3.1.

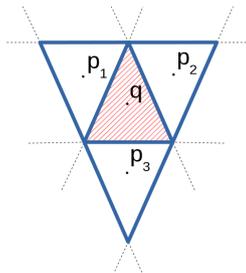


FIGURE 3.1 – Représentation de la maille q et de ses 3 mailles voisines p_1 , p_2 et p_3 .

3.1.2 Champ d'étiquette

Le découpage en région est décrit par une configuration d'étiquettes notée ℓ . On introduit un modèle stochastique pour les étiquettes et on note L_n , l'étiquette aléatoire attachée à la maille d'indice n et $\mathbf{L} = [L_1 \dots L_N]^t$, le jeu d'étiquettes complet. Afin d'introduire du lien entre les étiquettes et de concentrer la masse de probabilité sur un découpage en région, on choisit d'utiliser un champ aléatoire de Potts [74, p. 232]. La probabilité pour une configuration ℓ s'écrit :

$$\Pr(\mathbf{L} = \ell; \beta) = A(\beta)^{-1} \exp \beta \sum_{n \sim n'} \delta(\ell_n, \ell_{n'}), \quad (3.7)$$

avec $\sum_{n \sim n'}$, la somme sur l'ensemble des N_c paires de mailles voisines. La fonction de partition $A(\beta)$ se définit comme

$$A(\beta) = \sum_{\ell} \exp \beta \sum_{n \sim n'} \delta(\ell_n, \ell_{n'}), \quad (3.8)$$

avec \sum_{ℓ} , la somme sur toutes les configurations d'étiquettes possibles. Pour N étiquettes et K classes, il y a K^N configurations. Calculer $A(\beta)$ est alors un problème difficile. Dans ce travail, $N = 7400$ et si $K = 3$, la somme est alors composée de $3^{7400} \approx 10^{3530}$ termes à sommer.

Le paramètre $\beta \in \mathbb{R}^+$ est appelé coefficient de granularité et il pilote la taille moyenne des régions. S'il est grand, on favorise des régions étendues et inversement. On présente sur la Fig. 3.2 des réalisations pour différentes valeurs de β . Lorsque $\beta = 0$, toutes les configurations du champ sont équiprobables et en s'appuyant sur l'Eq. (3.7) et l'Eq. (3.8) on calcule leurs probabilités :

$$\Pr(\mathbf{L} = \ell; \beta = 0) = 1/K^N.$$

Remarque 9. D'après l'Eq. (3.7), les valeurs des étiquettes sont échangeables par classe. Si toutes les étiquettes égales à k échangent leurs valeurs avec toutes celles égales à k' , la probabilité reste inchangée. Cela s'explique par le fait que la somme $\sum_{n \sim n'} \delta(\ell_n, \ell_{n'})$ est égale au nombre de paires de mailles voisines dont les étiquettes ont la même valeur. Ainsi, lorsque les valeurs des étiquettes sont échangées, la valeur de la somme ne change pas.

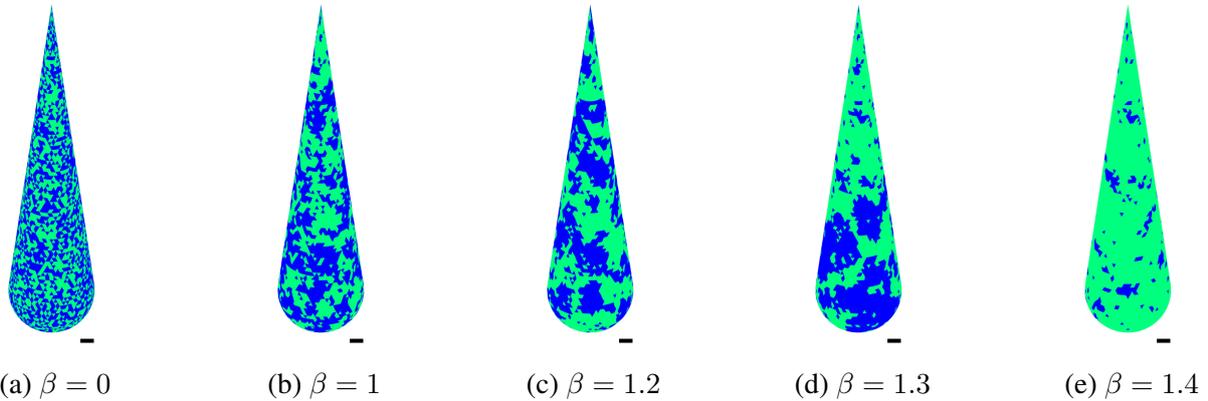


FIGURE 3.2 – Réalisations du champ de Potts, Eq. (3.7), affichées selon l'orientation (-) du cône, pour différentes valeurs de β et $K = 2$. On observe que plus β est grand et plus les régions sont étendues. Les réalisations sont obtenues grâce à un algorithme de Gibbs, décrit plus loin dans le manuscrit, qui échantillonne étiquette par étiquette le champ de Potts *a priori*.

3.1.3 Construction du champ d'impédance

On s'intéresse maintenant à la construction de l'impédance par morceaux. Les morceaux sont définis séparément les uns des autres à partir d'impédances complètes et du jeu d'étiquettes. Dans la suite de cette section, on considère que l'on est à configuration ℓ fixée.

3.1.3.1 Champs complets

Pour construire l'impédance par morceaux, l'idée est de considérer K impédances complètes $\mathbf{z}_k \in \mathbb{C}^N$ que l'on va découper puis rabouter en fonction des étiquettes ℓ , voir la Fig. 3.3. On choisit de les modéliser par des champs gaussiens indépendants dans leur ensemble. On les note \mathbf{Z}_k et leurs distributions s'écrivent [75, p. 77] :

$$\pi(\mathbf{z}_k; \mathbf{m}_k, \mathbf{C}_k) = \pi^{-N} \det \mathbf{C}_k^{-1} \exp - \|\mathbf{z}_k - \mathbf{m}_k\|_{\mathbf{C}_k^{-1}}^2, \quad (3.9)$$

avec $\mathbf{m}_k \in \mathbb{C}^N$ la moyenne et $\mathbf{C}_k \in \mathbb{R}^{N \times N}$ la covariance de la classe k .

Remarque 10. *Le choix d'une covariance réelle pour les champs gaussiens mène à une partie réelle et une partie imaginaire indépendantes et de même covariance $\mathbf{C}_k/2$, voir l'annexe B page 101.*

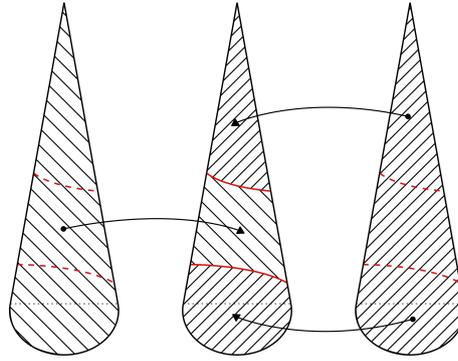


FIGURE 3.3 – Construction de l'impédance par morceaux à partir des impédances complètes et d'une configuration des étiquettes.

Moment d'ordre 1 – Dans ce travail, on sait que l'impédance est attendue proche de la référence \mathbf{z}_0 , voir la section 1.4.3 page 23. Afin de prendre cette information en compte, on choisit la référence comme moyenne pour les K champs gaussiens :

$$\mathbf{m}_k = \mathbf{z}_0. \quad (3.10)$$

De plus, on rappelle que la référence est uniforme $\mathbf{z}_0 = z_0 \mathbf{1}_N$. Les moyennes le sont donc également et peuvent s'écrire $\mathbf{m}_k = z_0 \mathbf{1}_N$. Dans le cas où l'on disposerait d'une linéarisation différente ou d'informations *a priori* supplémentaires, on souhaite que la méthode puisse s'appliquer lorsque les moyennes ne sont pas uniformes. Pour cette raison, on garde l'écriture \mathbf{m}_k dans nos développements.

Remarque 11. *Il serait possible d'introduire une perturbation sur les moyennes de chaque classe. Le modèle perturbé s'écrirait $\mathbf{m}_k = \mathbf{z}_0 + e_k \mathbf{1}_N$, avec e_k que l'on pourrait modéliser comme $e_k \sim \mathcal{N}(0, v_k)$. Ce modèle n'est pas étudié ici, mais constitue une perspective pour de futurs travaux.*

Moment d'ordre 2 – La modélisation des moments d'ordre 2 est déterminante pour la suite. Celle-ci peut être faite en covariance C_k ou en précision C_k^{-1} . Pour les deux possibilités, on peut définir des matrices pleines ou creuses.

- Pour la modélisation en covariance, celle-ci peut être considérée comme une fonction décroissante de la distance. Lorsque cette fonction est à support infini, par exemple de type « gaussienne » ou « exponentielle », la matrice de covariance est pleine. À l'inverse, si la fonction est à support fini, par exemple de type « gaussienne tronquée » ou « triangulaire », la matrice de covariance peut être creuse. Cela dépend de la taille du support.

- Pour la modélisation en précision, c'est la structure des interactions entre les mailles qui pilote le remplissage de la matrice. Si cette structure peut s'écrire par des relations entre voisins « proches », alors la matrice est creuse et il s'agit d'une structure markovienne [31, p. 169-177]. À l'inverse, si la structure s'écrit avec un voisinage étendu, la matrice de précision peut être pleine.

Entre la modélisation en covariance et la modélisation en précision, on choisit la dernière, car elle permet de modéliser plus simplement des champs markoviens. On fait apparaître un facteur d'échelle et une matrice de structure

$$C_k^{-1} = \gamma_k \mathbf{\Pi}_k. \quad (3.11)$$

La distribution des champs gaussiens s'écrit alors

$$\pi(\mathbf{z}_k; \mathbf{m}_k, \gamma_k, \alpha_k) = \pi^{-N} \gamma_k^N \det \mathbf{\Pi}_k \exp -\gamma_k \|\mathbf{z}_k - \mathbf{m}_k\|_{\mathbf{\Pi}_k}^2. \quad (3.12)$$

La suite est dédiée à la modélisation de la matrice de structure $\mathbf{\Pi}_k$. En gardant à l'esprit la connaissance selon laquelle l'impédance évolue lentement à la surface de l'objet, on construit des matrices de structure $\mathbf{\Pi}_k$ qui concentrent les masses de probabilités sur des champs corrélés. Pour cela, on choisit d'utiliser le modèle SAR (« simultaneous autoregressive », [76]) markovien suivant

$$\mathbf{\Pi}_k = \mathbf{D}_k^t \mathbf{D}_k, \quad (3.13)$$

avec \mathbf{D}_k la matrice des différences finies d'ordre 1 pondérée par un coefficient d'interaction $\alpha_k \in \mathbb{R}$. Afin d'introduire de la corrélation positive, on choisira des coefficients d'interaction positifs. Les différences sont calculées entre les valeurs de mailles voisines et on rappelle que chaque maille possède 3 voisines. Ainsi, l'argument de l'exponentielle de la densité de probabilité, Eq. (3.12), s'écrit à partir des relations par paire de mailles voisines comme

$$\begin{aligned} (\mathbf{z}_k - \mathbf{m}_k)^\dagger \mathbf{\Pi}_k (\mathbf{z}_k - \mathbf{m}_k) &= \|\mathbf{D}_k (\mathbf{z}_k - \mathbf{m}_k)\|^2 \\ &= \sum_{n \sim n'} |(z_k(n) - m_k(n)) - \alpha_k (z_k(n') - m_k(n'))|^2 \\ &= \sum_{n \sim n'} |(z_k(n) - \alpha_k z_k(n')) - (m_k(n) - \alpha_k m_k(n'))|^2. \end{aligned} \quad (3.14)$$

Afin que les lois des impédances, Eq. (3.12), soient correctement définies, il suffit que la modélisation choisie mène à des matrices $\mathbf{\Pi}_k$ inversibles. Par construction, la diagonale de $\mathbf{\Pi}_k$ est remplie d'une constante strictement positive fonction de α_k , voir l'exemple 5. Pour trouver les conditions sur α_k qui assurent l'inversibilité de la matrice, on cherche à déterminer pour quelles valeurs de α_k

$$\forall (\mathbf{z}_k - \mathbf{m}_k) \in \mathbb{C}_*^N, \|\mathbf{D}_k (\mathbf{z}_k - \mathbf{m}_k)\|^2 > 0. \quad (3.15)$$

D'après la démonstration développée dans l'annexe C, il est nécessaire que $\alpha_k \in \mathbb{R}_+ \setminus \{1\}$ pour que les matrices $\mathbf{\Pi}_k$ soient inversibles et que la corrélation soit positive. Pour ce travail, on choisit $\alpha_k \in [0, 1[$.

Exemple 5. Exemple d'une matrice de structure $\mathbf{\Pi}_k$ pour $N = 12$ mailles. On note $a = 3(1 + \alpha_k^2)$ et $b = -2\alpha_k$. La matrice $\mathbf{\Pi}_k$ possède 4 éléments non nuls par ligne et colonne. La position d'un b dans la matrice correspond à une paire de mailles voisines. Prenons l'exemple de la Fig. 3.1 page 41, avec la maille d'indice q qui a pour voisine les 3 mailles d'indice p_1 , p_2 et p_3 . On trouve alors sur la ligne q , les b aux indices de colonne p_1 , p_2 et p_3 correspondant aux paires (q, p_1) , (q, p_2) et (q, p_3) . La matrice étant symétrique, on trouve également sur la colonne q les b aux indices de ligne p_1 , p_2 et p_3 .

$$\mathbf{\Pi}_k = \begin{bmatrix} a & b & 0 & b & 0 & 0 & 0 & 0 & 0 & b & 0 & 0 \\ b & a & b & 0 & 0 & b & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & b & a & b & b & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ b & 0 & b & a & 0 & 0 & 0 & 0 & 0 & b & 0 & 0 \\ 0 & 0 & b & 0 & a & b & 0 & b & 0 & 0 & 0 & 0 \\ 0 & b & 0 & 0 & b & a & b & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & b & a & b & 0 & 0 & 0 & b \\ 0 & 0 & 0 & 0 & b & 0 & b & a & 0 & 0 & b & 0 \\ 0 & 0 & 0 & b & 0 & 0 & 0 & 0 & a & b & b & 0 \\ b & 0 & 0 & 0 & 0 & 0 & 0 & 0 & b & a & 0 & b \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & b & b & 0 & a & b \\ 0 & 0 & 0 & 0 & 0 & b & 0 & 0 & b & b & a \end{bmatrix} \quad \begin{aligned} a &= 3(1 + \alpha_k^2) \\ b &= -2\alpha_k \end{aligned}$$

Remarque 12. La structure de voisinage est la même pour toutes les classes. On pourrait donc écrire $\mathbf{\Pi}_k = \mathbf{\Pi}(\alpha_k) = \mathbf{D}^t(\alpha_k)\mathbf{D}(\alpha_k)$, mais on ne le fera pas, car l'écriture $\mathbf{\Pi}_k$ est plus compacte.

Remarque 13. Les mailles sont numérotées de manière arbitraire de 1 à N . Une renumérotation des mailles impliquerait la permutation des lignes et des colonnes des matrices $\mathbf{\Pi}_k$.

Pour la suite, on regroupe les facteurs d'échelle dans le vecteur $\boldsymbol{\gamma} = [\gamma_1 \dots \gamma_K]$ et les coefficients d'interaction dans le vecteur $\boldsymbol{\alpha} = [\alpha_1 \dots \alpha_K]$. De plus, on introduit les notations

$$\bar{\mathbf{z}}_k(\boldsymbol{\ell}) = \mathbf{T}_k(\boldsymbol{\ell})\mathbf{z}_k, \quad (3.16)$$

$$\bar{\mathbf{m}}_k(\boldsymbol{\ell}) = \mathbf{T}_k(\boldsymbol{\ell})\mathbf{m}_k, \quad (3.17)$$

$$\bar{\mathbf{C}}_k(\boldsymbol{\ell}) = \mathbf{T}_k(\boldsymbol{\ell})\mathbf{C}_k\mathbf{T}_k^t(\boldsymbol{\ell}), \quad (3.18)$$

qui seront utiles à la partie suivante. Elles correspondent aux morceaux et blocs des impédances, moyennes et covariances pour lesquelles les étiquettes ont la valeur k .

3.1.3.2 Champ par morceaux

À configuration d'étiquettes $\boldsymbol{\ell}$ donnée, l'impédance par morceaux \mathbf{z} est obtenue en raboutant les impédances complètes tronquées et l'on note sa construction :

$$\mathbf{z} = \sum_{k=1}^K \mathbf{P}_k(\boldsymbol{\ell})\mathbf{z}_k = \sum_{k=1}^K \mathbf{T}_k^t(\boldsymbol{\ell})\mathbf{T}_k(\boldsymbol{\ell})\mathbf{z}_k = \sum_{k=1}^K \mathbf{T}_k^t(\boldsymbol{\ell})\bar{\mathbf{z}}_k(\boldsymbol{\ell}). \quad (3.19)$$

Modélisation alternative de l'impédance par morceaux

Durant nos développements, nous avons envisagé une deuxième modélisation de l'impédance par morceaux. Cette seconde option consistait à s'appuyer sur l'Eq. (3.19) pour déterminer la loi de l'impédance par morceaux conditionnellement aux étiquettes. L'objectif était de ne considérer qu'une seule impédance au lieu des K impédances complètes. Avec ce modèle, nous avons réalisé des développements analogues à ceux que l'on présente dans la suite du manuscrit. Néanmoins, ils se sont avérés infructueux en raison d'un coût calculatoire élevé lors de l'échantillonnage des étiquettes. Ces difficultés sont la conséquence directe du modèle par morceaux conditionnel et de sa fonction de partition pilotée par la configuration d'étiquettes. Malgré le fait que cette voie n'est pas aboutie, nous présentons une partie de nos développements qui montrent comment la loi de \mathbf{Z} conditionnellement à $\mathbf{L} = \ell$ s'exprime à partir des morceaux et blocs $\bar{\mathbf{m}}_k(\ell)$ et $\bar{\mathbf{C}}_k(\ell)$.

Par construction, l'impédance par morceaux \mathbf{Z} est aléatoire et on cherche à définir sa loi afin de s'assurer qu'elle soit douce par morceaux. Conditionnellement à $\mathbf{L} = \ell$, \mathbf{Z} est gaussien, car il est une combinaison linéaire des champs gaussiens \mathbf{Z}_k ; on note $\mathbf{m}_{z|\ell}$ et $\mathbf{C}_{z|\ell}$ sa moyenne et sa covariance. Sa loi se déduit de la loi jointe des champs complets sachant les étiquettes, que l'on note $\pi(\mathbf{z}_1 \dots \mathbf{z}_K | \ell)$. On rappelle que les champs \mathbf{Z}_k sont indépendants dans leur ensemble et l'on ajoute qu'ils sont indépendants des étiquettes. Finalement, la loi jointe s'écrit

$$\pi(\mathbf{z}_1 \dots \mathbf{z}_K | \ell) = \pi(\mathbf{z}_1 \dots \mathbf{z}_K) = \prod_{k=1}^K \pi(\mathbf{z}_k). \quad (3.20)$$

Moment d'ordre 1 – On s'intéresse dans un premier temps au calcul de l'espérance $\mathbf{m}_{z|\ell}$. On l'écrit comme

$$\begin{aligned} \mathbf{m}_{z|\ell} &= \mathbb{E}[\mathbf{Z} | \mathbf{L} = \ell] = \mathbb{E}\left[\sum_{k=1}^K \mathbf{P}_k(\ell) \mathbf{Z}_k\right] \\ &= \sum_{k=1}^K \mathbf{P}_k(\ell) \mathbb{E}[\mathbf{Z}_k] = \sum_{k=1}^K \mathbf{P}_k(\ell) \mathbf{m}_k \\ &= \sum_{k=1}^K \mathbf{T}_k^t(\ell) \mathbf{T}_k(\ell) \mathbf{m}_k = \sum_{k=1}^K \mathbf{T}_k^t(\ell) \bar{\mathbf{m}}_k(\ell). \end{aligned} \quad (3.21)$$

La moyenne du champ conditionnellement aux étiquettes est donc constituée des moyennes découpées puis raboutées des champs \mathbf{Z}_k .

Remarque 14. Dans ce travail, toutes les moyennes $\mathbf{m}_k = z_0 \mathbf{1}_N$ et l'on en déduit $\mathbf{m}_{z|\ell} = z_0 \mathbf{1}_N$. L'impédance par morceaux est donc centrée sur la référence z_0 , ce qui est cohérent de l'information disponible.

Moment d'ordre 2 – On s'intéresse maintenant à la covariance $C_{z|\ell}$ qui s'écrit comme

$$\begin{aligned} C_{z|\ell} &= \text{cov} [\mathbf{Z} \mid \mathbf{L} = \ell] = \text{cov} \left[\sum_{k=1}^K \mathbf{P}_k(\ell) \mathbf{Z}_k \right] \\ &= \text{E} \left[\sum_{k,k'} \mathbf{P}_k(\ell) [\mathbf{Z}_k - \mathbf{m}_k] [\mathbf{Z}_{k'} - \mathbf{m}_{k'}]^\dagger \mathbf{P}_{k'}(\ell) \right] \\ &= \sum_{k,k'} \mathbf{P}_k(\ell) \text{E} \left[(\mathbf{Z}_k - \mathbf{m}_k) (\mathbf{Z}_{k'} - \mathbf{m}_{k'})^\dagger \right] \mathbf{P}_{k'}(\ell). \end{aligned} \quad (3.22)$$

La somme $\sum_{k,k'}$ fait apparaître toutes les combinaisons des champs pour k, k' de 1 à K . En s'appuyant sur le fait qu'ils sont indépendants dans leur ensemble, Eq. (3.20), on écrit l'espérance des combinaisons $k \neq k'$

$$\text{E} \left[(\mathbf{Z}_k - \mathbf{m}_k) (\mathbf{Z}_{k'} - \mathbf{m}_{k'})^\dagger \right] = 0, \quad (3.23)$$

et l'on en déduit la covariance

$$\begin{aligned} C_{z|\ell} &= \sum_{k=1}^K \mathbf{P}_k(\ell) \text{E} \left[(\mathbf{Z}_k - \mathbf{m}_k) (\mathbf{Z}_k - \mathbf{m}_k)^\dagger \right] \mathbf{P}_k(\ell) \\ &= \sum_{k=1}^K \mathbf{P}_k(\ell) \mathbf{C}_k \mathbf{P}_k(\ell) = \sum_{k=1}^K \mathbf{T}_k^t(\ell) \mathbf{T}_k(\ell) \mathbf{C}_k \mathbf{T}_k^t(\ell) \mathbf{T}_k(\ell) \\ &= \sum_{k=1}^K \mathbf{T}_k^t(\ell) \bar{\mathbf{C}}_k(\ell) \mathbf{T}_k(\ell), \end{aligned} \quad (3.24)$$

qui s'exprime à partir des morceaux $\bar{\mathbf{C}}_k(\ell)$ raboutés pour former une matrice bloc diagonale à une permutation près. On montre dans la partie D.2 de l'annexe D, page 105, que l'inverse d'une matrice bloc diagonale à une permutation près est égale à la matrice bloc diagonale des inverses bloc à la même permutation près. La matrice de précision est alors constituée des morceaux $\bar{\mathbf{C}}_k(\ell)^{-1}$ à la même permutation près :

$$\mathbf{C}_{z|\ell}^{-1} = \sum_{k=1}^K \mathbf{T}_k^t(\ell) \bar{\mathbf{C}}_k(\ell)^{-1} \mathbf{T}_k(\ell). \quad (3.25)$$

Remarque 15. De manière générale, l'inverse des covariances tronquées n'est pas égale à la troncature des covariances inverses :

$$[\mathbf{T}_k(\ell) \mathbf{C}_k \mathbf{T}_k^t(\ell)]^{-1} \neq \mathbf{T}_k(\ell) \mathbf{C}_k^{-1} \mathbf{T}_k^t(\ell). \quad (3.26)$$

Dans certains cas particuliers, par exemple lorsque \mathbf{C}_k est diagonale [50, 56], il y a égalité. Dans notre cas, les matrices \mathbf{C}_k sont diagonales si $\alpha_k = 0$. Néanmoins, cela mène à des champs décorrélés et c'est une situation que l'on ne désire pas. Dans la suite, on considère que les matrices \mathbf{C}_k ne sont pas diagonales et l'inégalité (3.26) vraie.

Champ par morceaux – On écrit la loi du champ \mathbf{Z} conditionnellement aux étiquettes $\mathbf{L} = \ell$

$$\pi(\mathbf{z}|\ell; \boldsymbol{\gamma}, \boldsymbol{\alpha}) = \pi^{-N} \det \mathbf{C}_{z|\ell}^{-1} \exp - \left\| \mathbf{z} - \mathbf{m}_{z|\ell} \right\|_{\mathbf{C}_{z|\ell}^{-1}}^2. \quad (3.27)$$

Afin de faire apparaître les covariances tronquées, on s'appuie sur le fait que le déterminant d'une matrice bloc diagonale est égal au produit des déterminants des blocs. Dans notre cas, on montre dans la section D.1 de l'annexe D page 105, que le déterminant d'une matrice bloc diagonale à une permutation près est aussi égal au produit des déterminants des blocs

$$\det \mathbf{C}_{z|\ell} = \prod_{k=1}^K \det \bar{\mathbf{C}}_k(\ell). \quad (3.28)$$

Concernant l'argument de l'exponentielle, on va montrer qu'il s'écrit à partir des impédances, des moyennes et des covariances tronquées. Dans un premier temps, on écrit la norme de manière compacte avec les moments par morceaux $\mathbf{m}_{z|\ell}$ et $\mathbf{C}_{z|\ell}$

$$\|z - \mathbf{m}_{z|\ell}\|_{\mathbf{C}_{z|\ell}^{-1}}^2 = (z - \mathbf{m}_{z|\ell})^\dagger \mathbf{C}_{z|\ell}^{-1} (z - \mathbf{m}_{z|\ell}). \quad (3.29)$$

En s'appuyant sur l'Eq. (3.19) et l'Eq. (3.21), on peut écrire $z - \mathbf{m}_{z|\ell}$ sous la forme d'une somme

$$z - \mathbf{m}_{z|\ell} = \sum_{k=1}^K \mathbf{T}_k^t(\ell) (\bar{z}_k(\ell) - \bar{\mathbf{m}}_k(\ell)). \quad (3.30)$$

À partir de cette somme et de l'Eq. (3.25), on écrit la norme comme :

$$\begin{aligned} \|z - \mathbf{m}_{z|\ell}\|_{\mathbf{C}_{z|\ell}^{-1}}^2 &= (z - \mathbf{m}_{z|\ell})^\dagger \mathbf{C}_{z|\ell}^{-1} (z - \mathbf{m}_{z|\ell}) \\ &= \left(\sum_{k=1}^K \mathbf{T}_k^t(\ell) (\bar{z}_k(\ell) - \bar{\mathbf{m}}_k(\ell)) \right)^\dagger \left(\sum_{k'=1}^K \mathbf{T}_{k'}^t(\ell) \bar{\mathbf{C}}_{k'}(\ell)^{-1} \mathbf{T}_{k'}(\ell) \right) \left(\sum_{k''=1}^K \mathbf{T}_{k''}^t(\ell) (\bar{z}_{k''}(\ell) - \bar{\mathbf{m}}_{k''}(\ell)) \right). \end{aligned}$$

En développant, on fait apparaître des termes croisés k, k', k'' pour lesquels le produit des matrices de troncature et de bourrage est nul lorsqu'elles ne sont pas de la même classe, par exemple $\mathbf{T}_k(\ell) \mathbf{T}_{k'}^t(\ell)$ est nulle lorsque $k \neq k'$, voir l'Eq. (3.6) page 41. De tous les termes, il ne reste que ceux pour lesquels $k = k' = k''$ et l'on en déduit

$$\|z - \mathbf{m}_{z|\ell}\|_{\mathbf{C}_{z|\ell}^{-1}}^2 = \sum_{k=1}^K (\bar{z}_k(\ell) - \bar{\mathbf{m}}_k(\ell))^\dagger \mathbf{T}_k(\ell) \mathbf{T}_k^t(\ell) \bar{\mathbf{C}}_k(\ell)^{-1} \mathbf{T}_k(\ell) \mathbf{T}_k^t(\ell) (\bar{z}_k(\ell) - \bar{\mathbf{m}}_k(\ell)).$$

Au contraire, $\mathbf{T}_k(\ell) \mathbf{T}_k^t(\ell) = \mathbf{I}_{N_k(\ell)}$, Eq. (3.5) page 41, on a donc

$$\begin{aligned} \|z - \mathbf{m}_{z|\ell}\|_{\mathbf{C}_{z|\ell}^{-1}}^2 &= \sum_{k=1}^K (\bar{z}_k(\ell) - \bar{\mathbf{m}}_k(\ell))^\dagger \mathbf{I}_{N_k(\ell)} \bar{\mathbf{C}}_k(\ell)^{-1} \mathbf{I}_{N_k(\ell)} (\bar{z}_k(\ell) - \bar{\mathbf{m}}_k(\ell)) \\ &= \sum_{k=1}^K \|\bar{z}_k(\ell) - \bar{\mathbf{m}}_k(\ell)\|_{\bar{\mathbf{C}}_k(\ell)^{-1}}^2. \end{aligned} \quad (3.31)$$

Finalement, la loi de \mathbf{Z} conditionnellement à $\mathbf{L} = \ell$ s'écrit :

$$\begin{aligned} \pi(z|\ell; \gamma, \alpha) &= \pi^{-N} \det \mathbf{C}_{z|\ell}^{-1} \exp - \|z - \mathbf{m}_{z|\ell}\|_{\mathbf{C}_{z|\ell}^{-1}}^2 \\ &= \pi^{-N} \prod_{k=1}^K \det \bar{\mathbf{C}}_k(\ell)^{-1} \exp - \sum_{k=1}^K \|\bar{z}_k(\ell) - \bar{\mathbf{m}}_k(\ell)\|_{\bar{\mathbf{C}}_k(\ell)^{-1}}^2 \\ &= \prod_{k=1}^K \pi^{-N_k(\ell)} \det \bar{\mathbf{C}}_k(\ell)^{-1} \exp - \|\bar{z}_k(\ell) - \bar{\mathbf{m}}_k(\ell)\|_{\bar{\mathbf{C}}_k(\ell)^{-1}}^2. \end{aligned} \quad (3.32)$$

La loi conditionnelle pour l'impédance par morceaux est bien séparable et chaque facteur du produit d'indice k s'écrit à partir des morceaux $\bar{\mathbf{m}}_k(\ell)$ et blocs $\bar{\mathbf{C}}_k(\ell)$.

Remarque 16. *Si la classe k est vide, on a $N_k(\ell) = 0$ et les vecteurs $\bar{\mathbf{z}}_k(\ell)$, $\bar{\mathbf{m}}_k(\ell)$ et la matrice $\bar{\mathbf{C}}_k(\ell)$ sont vides. Par convention, on choisit $\det \bar{\mathbf{C}}_k(\ell) = 1$ et $\|\bar{\mathbf{z}}_k(\ell) - \bar{\mathbf{m}}_k(\ell)\|_{\bar{\mathbf{C}}_k(\ell)^{-1}} = 0$. Lorsqu'une classe est vide, le facteur correspondant dans le produit de l'Eq. (3.32) vaut alors 1. Lorsque toutes les classes sont vides, sauf une, le produit se ramène alors à un seul facteur qui est celui de la classe non vide. Si la classe non vide est la classe k , on a alors $N_k(\ell) = N$, $\mathbf{m}_{z|\ell} = \mathbf{m}_k$ et $\mathbf{C}_{z|\ell} = \mathbf{C}_k$.*

Comme mentionné au début de la section, le modèle d'impédance par morceaux conditionnellement aux étiquettes n'a pas été retenu puisqu'il mène à un coût calculatoire élevé. Le travail présenté dans la suite s'appuie alors seulement sur les K impédances complètes, la configuration d'étiquettes et l'Eq. (3.19) permettant de constituer l'impédance par morceaux. Ces considérations mènent à une hiérarchie du problème différente. Il s'agit d'un élément clé pour la suite, puisqu'il permet d'éviter les difficultés liées à la fonction de partition.

3.2 Modèle pour le bruit et vraisemblance

Comme mentionné précédemment, les données sont entachées d'erreurs. Ces erreurs, modélisées au travers du terme additif ε , font référence aux erreurs de mesures (partie 1.2 page 16), à l'erreur de l'approximation en impédance (partie 1.4.1 page 20) et aux erreurs dues à la linéarisation (partie 1.4.3 page 23).

Pour le moment le bruit n'est pas caractérisé. En l'absence d'informations, nous travaillons avec le modèle suivant : bruit additif gaussien complexe, de moyenne \mathbf{m}_ε et de covariance \mathbf{C}_ε

$$\varepsilon \sim \mathcal{N}\left(\mathbf{m}_\varepsilon, \mathbf{C}_\varepsilon\right). \quad (3.33)$$

Modéliser le bruit par une loi gaussienne est un choix courant dans de nombreux travaux [26, 58]. Dans notre cas, il s'agit d'un choix pratique qui simplifie les calculs détaillés dans le chapitre suivant. De plus, on pourrait utiliser un mélange de gaussiennes pour approcher certaines lois qui seraient non gaussiennes et bénéficier de nos développements.

Concernant la modélisation de la matrice d'ordre 2, il y a deux options : la modéliser en covariance ou en précision. On choisit la forme précision et on l'écrit au travers d'un facteur d'échelle γ_ε et d'une matrice de structure $\mathbf{\Pi}_\varepsilon$:

$$\mathbf{C}_\varepsilon^{-1} = \gamma_\varepsilon \mathbf{\Pi}_\varepsilon. \quad (3.34)$$

Dans la suite, on considère γ_ε comme un hyperparamètre que l'on estime conjointement au reste. On propose de ne pas choisir de valeur particulière pour la moyenne \mathbf{m}_ε et la matrice $\mathbf{\Pi}_\varepsilon$. Le but est de pouvoir introduire des informations sur le bruit qui viendraient à être connues par la suite, voir la remarque suivante.

Remarque 17. *On peut imaginer qu'à l'issue d'un processus de calibrage l'erreur soit mieux caractérisée et que des « erreurs systématiques » soient connues. On pourrait alors les inclure grâce à \mathbf{m}_ε . À ce jour, on ne dispose pas d'information sur la corrélation des erreurs, mais le modèle permet de l'encoder au travers de la matrice $\mathbf{\Pi}_\varepsilon$.*

En s'appuyant sur la modélisation du bruit, Eq. (3.33), et la construction du champ d'impédance par morceaux, Eq. (3.19), on déduit la loi pour σ sachant une configuration d'étiquettes ℓ et des jeux d'impédance z_k . Pour alléger les notations, on introduit le changement de variable suivant

$$\bar{\sigma}_0 = \sigma_0 - H z_0 + m_\varepsilon. \quad (3.35)$$

Le modèle de vraisemblance s'écrit :

$$f(\sigma | z_{1...K}, \ell, \gamma_\varepsilon) \propto \gamma_\varepsilon^M \det \Pi_\varepsilon \exp -\gamma_\varepsilon \| (\sigma - \bar{\sigma}_0) - H \sum_{k=1}^K P_k(\ell) z_k \|_{\Pi_\varepsilon}^2. \quad (3.36)$$

3.3 Hyperparamètres

De manière générale, les modèles structurent le problème en introduisant de l'information et ce sont les hyperparamètres qui pilotent ces modèles. La démarche bayésienne permet de les inclure dans les inconnues et de les estimer conjointement aux impédances et aux étiquettes [31, p. 195-216]. Cette partie fournit pour chaque hyperparamètre une loi *a priori* qui s'appuie sur les connaissances disponibles, valeurs et incertitudes. Si l'information dont on dispose est précise, par exemple sur sa valeur nominale, alors la loi *a priori* est « piquée » sur cette valeur. À l'inverse, si peu ou aucune information n'est disponible, la loi est diffuse. Dans la suite, on s'intéresse aux paramètres γ et γ_ε , et on considère les paramètres β et α fixés.

Les paramètres d'échelle γ et γ_ε interviennent dans les matrices de précision des impédances et du bruit. Ce sont des facteurs qui pilotent l'écart à la moyenne. On sait qu'ils sont positifs et on évite le cas où ils sont nuls, ce qui mènerait à des lois mal définies. Afin d'encoder cette information, on choisit de les modéliser sous la forme de loi Gamma (Γ) dont la densité de probabilité s'écrit

$$f(\gamma; a, b) = \frac{b^a}{\Gamma(a)} \gamma^{a-1} \exp -b\gamma \quad \mathbb{1}^+(\gamma), \quad (3.37)$$

avec $\mathbb{1}^+$ l'indicatrice de \mathbb{R}^+ . Pour les $\gamma_{1...K}$ et γ_ε , hormis la positivité, aucune autre information n'est disponible. Pour cette raison, on choisit (a, b) proche de $(0, 0)$ de sorte à fournir un *a priori* peu informatif. Finalement, on note

$$\gamma_k \sim \Gamma(a_k, b_k), \quad (3.38)$$

$$\gamma_\varepsilon \sim \Gamma(a_\varepsilon, b_\varepsilon). \quad (3.39)$$

Le choix de lois gamma, pour les facteurs d'échelle, et de lois gaussiennes, pour les impédances et le bruit, mène dans la suite à des lois conjuguées, ce qui simplifie les calculs.

Remarque 18. Si des développements futurs amènent plus d'information concernant les facteurs d'échelle, par exemple sa valeur nominale, la loi gamma pourra en tenir compte en choisissant les paramètres a et b adéquats.

3.4 Modèle hiérarchique et loi *a posteriori*

Les modèles *a priori* choisis introduisent plusieurs dépendances conditionnelles entre certaines variables. Néanmoins, ces définitions ne suffisent pas, car elles ne donnent pas toutes les

informations nécessaires à caractériser les dépendances entre toutes les variables et définir la loi jointe. Le modèle hiérarchique ou réseau bayésien définit clairement toutes les dépendances conditionnelles. Pour cela, il s'appuie sur un graphe qui fait apparaître les inconnues reliées par des arcs dirigés représentant les relations directes. Les dépendances conditionnelles sont définies au travers des règles définies dans [77, p. 19]. Le modèle hiérarchique que l'on considère est représenté sur la Fig. 3.4. Par ailleurs, il offre ainsi une vue d'ensemble synthétique de toutes les variables du problème.

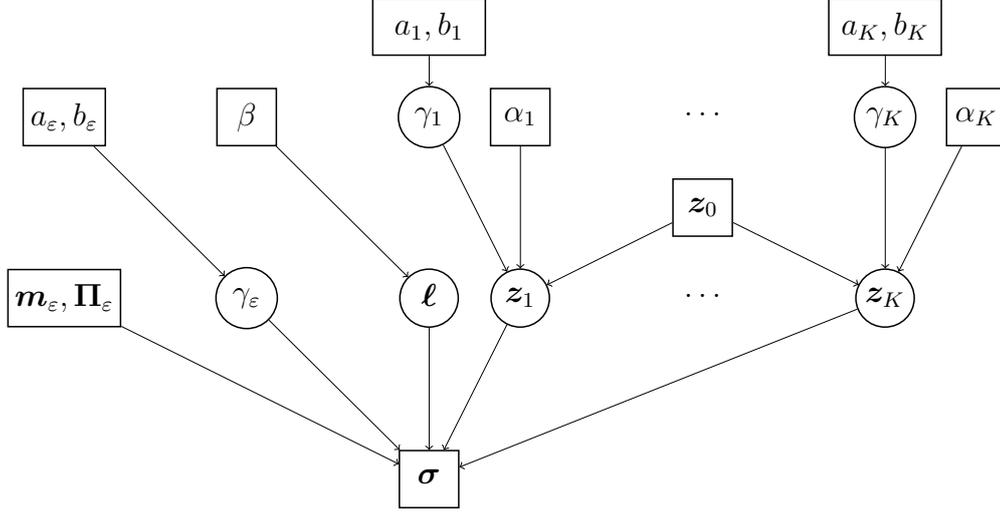


FIGURE 3.4 – Modèle hiérarchique : \square variables observées ou connues *a priori*, \circ variables inconnues estimées.

3.4.1 Loi jointe

La loi jointe est la distribution pour toutes les variables, les données comme les inconnues. Elle se déduit des modèles introduits dans les parties précédentes et du modèle hiérarchique :

$$\pi(\mathbf{z}_{1\dots K}, \ell, \sigma, \gamma_\varepsilon, \gamma) = f(\sigma | \mathbf{z}_{1\dots K}, \ell, \gamma_\varepsilon) f(\gamma_\varepsilon) \pi(\mathbf{z}_{1\dots K} | \gamma) f(\gamma) \Pr(\mathbf{L} = \ell). \quad (3.40)$$

C'est une fonction de forme non standard qui fait intervenir des variables discrètes et continues de grandes dimensions.

3.4.2 Loi a posteriori

À partir de la loi jointe, on obtient la loi *a posteriori* par application de la règle de Bayes. Elle est égale à la loi jointe à une constante multiplicative près et elle s'écrit :

$$\pi(\mathbf{z}_{1\dots K}, \ell, \sigma, \gamma_\varepsilon, \gamma | \sigma) = \frac{\pi(\mathbf{z}_{1\dots K}, \ell, \sigma, \gamma_\varepsilon, \gamma)}{\pi(\sigma)}, \quad (3.41)$$

avec $\pi(\sigma)$ la loi marginale des données qui s'écrit comme la somme

$$\pi(\sigma) = \sum_{\ell} \int_{\mathbf{z}_{1\dots K}, \gamma_\varepsilon, \gamma} \pi(\mathbf{z}_{1\dots K}, \ell, \sigma, \gamma_\varepsilon, \gamma) d\mathbf{z}_{1\dots K} d\gamma_\varepsilon d\gamma. \quad (3.42)$$

Comme abordé dans le chapitre précédent, le calcul de cette marginale n'est pas toujours nécessaire, ce qui est notre cas. Dans la suite du document, on est amené à écrire plusieurs lois qui s'appuient sur la loi *a posteriori*. Afin d'alléger les notations quand cela est possible, on écrit ces lois à une constante multiplicative près. Finalement :

$$\begin{aligned}
\pi(\mathbf{z}_{1\dots K}, \boldsymbol{\ell}, \gamma_\varepsilon, \boldsymbol{\gamma} | \boldsymbol{\sigma}) &\propto f(\boldsymbol{\sigma} | \mathbf{z}_{1\dots K}, \boldsymbol{\ell}, \gamma_\varepsilon) f(\gamma_\varepsilon) \pi(\mathbf{z}_{1\dots K} | \boldsymbol{\gamma}) f(\boldsymbol{\gamma}) \Pr(\mathbf{L} = \boldsymbol{\ell}) \\
&\propto \gamma_\varepsilon^M \exp -\gamma_\varepsilon \| (\boldsymbol{\sigma} - \bar{\boldsymbol{\sigma}}_0) - \mathbf{H} \sum_{k=1}^K \mathbf{P}_k(\boldsymbol{\ell}) \mathbf{z}_k \|^2_{\Pi_\varepsilon} \gamma_\varepsilon^{a_\varepsilon - 1} \exp -b_\varepsilon \gamma_\varepsilon \\
&\quad \exp \beta \sum_{n \sim n'} \delta(\ell_n, \ell_{n'}) \\
&\quad \prod_{k=1}^K \gamma_k^N \exp -\gamma_k \| \mathbf{z}_k - \mathbf{m}_k \|^2_{\Pi_k} \gamma_k^{a_k - 1} \exp -b_k \gamma_k
\end{aligned} \tag{3.43}$$

La loi *a posteriori* synthétise toutes les informations concernant les inconnues contenues par les données et les modèles *a priori*.

Remarque 19. *Comme attendu, l'impédance par morceaux apparaît au travers de la relation $\sum \mathbf{P}_k(\boldsymbol{\ell}) \mathbf{z}_k$. Du point de vue de la méthode elle n'est pas directement probabilisée et le problème de son estimation est substitué par celui de l'estimation des étiquettes et des K impédances.*

3.5 Estimateurs

À partir de la loi *a posteriori*, on introduit plusieurs estimateurs pour les inconnues. Afin de rendre les écritures plus légères, on sépare les inconnues en deux parties : d'un côté les étiquettes, discrètes, et de l'autre les variables continues que l'on place dans le vecteur

$$\boldsymbol{\theta} = [\mathbf{z}_{1\dots K}, \gamma_\varepsilon, \boldsymbol{\gamma}]. \tag{3.44}$$

Dans le chapitre précédent, nous avons discuté de deux estimateurs : le Maximiseur *a posteriori* (MAP) et l'Espérance *a posteriori* (EAP). Dans notre cas, il faut rappeler que les étiquettes sont discrètes et que l'EAP n'a pas de sens. On choisit deux estimateurs différents, un pour $\boldsymbol{\theta}$ et un pour $\boldsymbol{\ell}$. On s'intéresse également aux incertitudes associées.

3.5.1 Étiquettes

Pour les étiquettes, on souhaiterait choisir le MAP puisque cet estimateur fournit la configuration d'étiquettes la plus probable, voir la partie 2.3.2.1 page 34. Cependant, son calcul est difficile à cause de la grande dimension du problème. À la place, on choisit de calculer le Maximiseur Marginal *a posteriori* (MMAP) qui est, séparément pour chaque étiquette, la valeur la plus probable. C'est l'estimateur optimal pour une fonction de coût binaire séparable [53, 78]. Pour $n = 1 \dots N$, il s'écrit

$$\hat{\ell}_n = \arg \max_{\ell_n} \pi(\ell_n | \boldsymbol{\sigma}), \tag{3.45}$$

avec la loi marginale

$$\pi(\ell_n|\sigma) = \sum_{\ell_{-n}} \int_{\theta} \pi(\theta, \ell_n, \ell_{-n}|\sigma) d\theta, \quad (3.46)$$

et $\sum_{\ell_{-n}} \int_{\theta}$ une sommation et une intégration sur toutes les valeurs de θ et toutes les configurations des étiquettes ℓ_{-n} (toutes les étiquettes sauf ℓ_n) possibles. On ne sait pas calculer analytiquement cette quantité. Néanmoins, on présente dans le chapitre suivant une manière d'y parvenir numériquement.

Afin d'obtenir $\hat{\ell}_n$, on peut calculer la probabilité marginale *a posteriori* $\pi(\ell_n|\sigma)$ pour les K valeurs d'étiquette et choisir celle qui maximise la loi. Cette stratégie est efficace, car le nombre de classes que l'on considère n'est pas très élevé. Les incertitudes associées à l'estimation sont alors directement obtenues grâce aux probabilités marginales *a posteriori*.

3.5.2 Impédance et hyperparamètres

Concernant les inconnues θ , on choisit l'EAP et l'on considère deux possibilités permettant d'obtenir des estimées. On propose également une manière de quantifier les incertitudes associées.

1. Dans le cas où les étiquettes sont marginalisées, on définit l'estimateur comme

$$\begin{aligned} \hat{\theta} &= \int_{\theta} \theta \pi(\theta|\sigma) d\theta \\ &= \sum_{\ell} \int_{\theta} \theta \pi(\theta, \ell|\sigma) d\theta \\ &= \sum_{\ell} \Pr(\mathbf{L} = \ell|\sigma) \int_{\theta} \theta \pi(\theta|\ell, \sigma) d\theta, \end{aligned} \quad (3.47)$$

c'est-à-dire comme la moyenne pondérée des espérances conditionnelles. On définit les incertitudes associées par la matrice de covariance *a posteriori*

$$\begin{aligned} \hat{\mathbf{C}}_{\theta} &= \int_{\theta} (\theta - \hat{\theta})(\theta - \hat{\theta})^{\dagger} \pi(\theta|\sigma) d\theta \\ &= \sum_{\ell} \int_{\theta} (\theta - \hat{\theta})(\theta - \hat{\theta})^{\dagger} \pi(\theta, \ell|\sigma) d\theta \\ &= \sum_{\ell} \Pr(\mathbf{L} = \ell|\sigma) \int_{\theta} (\theta - \hat{\theta})(\theta - \hat{\theta})^{\dagger} \pi(\theta|\ell, \sigma) d\theta, \end{aligned} \quad (3.48)$$

qui s'écrit comme la moyenne pondérée des covariances conditionnelles.

2. Dans le cas où la configuration d'étiquettes est fixée à la valeur $\hat{\ell}$, on définit l'estimateur

$$\hat{\theta}(\hat{\ell}) = \int_{\theta} \theta \pi(\theta|\hat{\ell}, \sigma) d\theta, \quad (3.49)$$

et les incertitudes associées par la matrice de covariance *a posteriori*

$$\hat{\mathbf{C}}_{\theta}(\hat{\ell}) = \int_{\theta} (\theta - \hat{\theta}(\hat{\ell}))(\theta - \hat{\theta}(\hat{\ell}))^{\dagger} \pi(\theta|\hat{\ell}, \sigma) d\theta. \quad (3.50)$$

Les matrices \hat{C}_θ et $\hat{C}_\theta(\hat{\ell})$ contiennent toutes les matrices de covariances *a posteriori* des inconnues : z_k, γ , etc, mais également toutes les intercorrélations *a posteriori*. Elles apportent ainsi des informations sur la manière dont les inconnues sont corrélées *a posteriori*.

Remarque 20. L'Eq. (3.47) et (3.48) mènent à mélanger des valeurs issues de toutes les classes, tandis que l'Eq. (3.49) et (3.50) ne font pas de tel mélange.

Remarque 21. Une manière de quantifier les incertitudes, que l'on retrouvera dans le Chapitre 5, est de calculer des intervalles de crédibilité [35, p. 17]. L'intervalle de crédibilité se définit comme un sous-ensemble \mathcal{D} tel que $\Pr(\boldsymbol{\theta} \in \mathcal{D} | \boldsymbol{\sigma}) = 1 - \lambda$, avec $\lambda \in [0, 1]$.

3.5.3 Difficultés calculatoires

Le calcul direct de ces estimateurs est délicat puisque l'on ne sait pas faire les opérations de sommation et d'intégration analytique. La loi *a posteriori* fait intervenir des variables de grande dimension, elle est de forme complexe et combine à la fois des variables continues et discrètes. Pour surmonter ces difficultés, on choisit d'utiliser une méthode numérique du type MCMC (Monte Carlo par chaînes de Markov) qui fournit des échantillons sous la loi *a posteriori* et les exploite afin de calculer numériquement les estimateurs. On aborde ce problème dans le chapitre suivant.

CHAPITRE 4

APPROCHE NUMÉRIQUE

Le calcul des estimateurs définis dans la partie 3.5 nécessite essentiellement l'intégration de la densité a posteriori. Par exemple, l'espérance ou la covariance a posteriori sont de la forme $\mathbf{I}(\varphi) = \int \varphi(\boldsymbol{\theta})\pi(\boldsymbol{\theta}|\boldsymbol{\sigma}) d\boldsymbol{\theta}$, où φ est une fonction test. La loi a posteriori fait intervenir des variables de grandes dimensions, elle est de forme non standard et l'on ne sait pas faire son intégration analytiquement. Afin de surmonter ce problème, on choisit de s'appuyer sur des méthodes numériques. L'utilisation de méthodes d'intégration déterministe, telle que les méthodes de quadrature, est rendue impossible par la grande dimension. Pour cette raison, on se tourne vers des méthodes d'intégration Monte Carlo qui exploitent des réalisations aléatoires : $\boldsymbol{\theta}^{[1]}, \dots, \boldsymbol{\theta}^{[N]}$ de $\pi(\boldsymbol{\theta}|\boldsymbol{\sigma})$ afin d'approximer l'intégrale : $\mathbf{I}(\varphi) \approx \frac{1}{N} \sum_{n=1}^N \varphi(\boldsymbol{\theta}^{[n]})$. Comme on ne sait pas échantillonner directement selon la loi a posteriori, on choisit d'utiliser des méthodes Monte Carlo par chaînes de Markov (MCMC) et, plus particulièrement, l'algorithme de Gibbs.

Sommaire

4.1	Méthodes Monte Carlo par chaînes de Markov	56
4.1.1	Intégration de Monte Carlo	56
4.1.2	Algorithmes de MCMC	56
4.1.3	Algorithme de Gibbs	58
4.2	Méthode de Gibbs pour l'inversion-segmentation	60
4.2.1	Champs d'impédance	60
4.2.1.1	Difficultés d'échantillonnage	61
4.2.1.2	Échantillonnage efficace en dimension élevée	63
4.2.2	Paramètres d'échelle	65
4.2.2.1	Bruit	66
4.2.2.2	Impédances	66
4.2.3	Champ d'étiquettes	67
4.2.4	Pseudo-code	69
4.3	Calcul des estimateurs	71
4.3.1	Étiquettes	71
4.3.2	Impédance et hyperparamètres	71

4.1 Méthodes Monte Carlo par chaînes de Markov

Le calcul intégral est au cœur de nombreux problèmes en statistique bayésienne [79], par exemple : la marginalisation, le calcul d'une fonction de partition ou le calcul de l'espérance. La résolution analytique de tels problèmes est rarement possible : grande dimension, forme non standard, *etc.*, et on les résout numériquement.

Les méthodes de quadrature souffrent de la malédiction de la dimension avec une complexité qui croît de manière exponentielle avec la dimension. Une autre classe de méthode moins sujette à la malédiction de la dimension est celle des méthodes de Monte Carlo et la suite du chapitre leur est dédiée.

4.1.1 Intégration de Monte Carlo

Les méthodes de Monte Carlo approximent des intégrales en s'appuyant sur des réalisations aléatoires. Se référer à [79] pour plus de détails. On considère N réalisations i.i.d. sous $p(\mathbf{x})$ que l'on note $\mathbf{x}_1 \dots \mathbf{x}_N$. En exploitant ces réalisations, on peut approximer $p(\mathbf{x})$ par

$$p_N(\mathbf{x}) = \frac{1}{N} \sum_{n=1}^N \delta(\mathbf{x}_n, \mathbf{x}), \quad (4.1)$$

et le calcul intégral $I(\varphi)$ par la somme $I_N(\varphi)$ qui converge comme suit

$$I_N(\varphi) = \frac{1}{N} \sum_{n=1}^N \varphi(\mathbf{x}_n) \xrightarrow[N \rightarrow \infty]{p.s.} I(\varphi) = \int \varphi(\mathbf{x}) p(\mathbf{x}) \, d\mathbf{x}, \quad (4.2)$$

avec *p.s.* : presque sûrement. L'estimateur $I_N(\varphi)$ est non biaisé et son erreur quadratique moyenne vaut $1/N \text{cov}(\varphi(\mathbf{x}))$.

Remarque 22. *Les méthodes Monte Carlo exploitent une discrétisation non régulière de l'espace, induite par les réalisations aléatoires, pour approximer une intégrale. La différence avec les méthodes de quadrature est que la « grille de discrétisation » est naturellement plus fine dans les zones de fortes probabilités.*

Dans notre problème, on ne sait pas obtenir de réalisations aléatoires de la loi *a posteriori* de manière directe et donc calculer une approximation Monte Carlo telle que définie dans l'Eq. (4.2). Dans la littérature, on trouve plusieurs méthodes d'échantillonnage qui permettent de répondre à ce problème. Par exemple la méthode d'acceptation rejet ou encore l'échantillonnage préférentiel. Malheureusement, ces méthodes deviennent peu efficaces ou difficiles à mettre en place en grande dimension [79], ce qui est notre cas. Pour surmonter ces difficultés, des méthodes plus adaptées appelées Monte Carlo par chaînes de Markov se sont développées durant les dernières décennies.

4.1.2 Algorithmes de MCMC

On considère la situation où l'on ne sait pas échantillonner $p(\mathbf{x})$ directement. Dans la suite, on se propose d'expliquer de manière succincte le fonctionnement des méthodes MCMC. Se reporter à [35, 79] et [80] pour une introduction plus complète.

Principe – Les algorithmes MCMC sont une classe de méthodes qui, en utilisant un mécanisme de chaîne de Markov, fournissent des réalisations aléatoires d’une loi cible $p(\mathbf{x})$. Elles reposent sur un noyau de transition, noté \mathcal{K} , qui rend compte de l’évolution des masses de probabilité associées à un instant i à partir des événements passés. Le processus aléatoire $\mathbf{x}^{[i]}$ est appelé une chaîne de Markov si la probabilité de transition

$$p(\mathbf{x}^{[i]}|\mathbf{x}^{[i-1]}, \mathbf{x}^{[i-2]}, \dots) = \mathcal{K}(\mathbf{x}^{[i]}|\mathbf{x}^{[i-1]}). \quad (4.3)$$

Le plus souvent, on considère que les chaînes de Markov sont *homogènes*, c’est-à-dire que le noyau de transition ne change pas au cours du temps.

Le principe des MCMC est alors le suivant : à partir d’une condition initiale $\mathbf{x}^{[0]} \sim q(\mathbf{x}^{[0]})$, on fournit à chaque itération i une nouvelle réalisation aléatoire notée $\mathbf{x}^{[i]} \sim q(\mathbf{x}^{[i]})$ et la densité de probabilité marginale $q(\mathbf{x}^{[i]})$ est définie par

$$q(\mathbf{x}^{[i]}) = \int_{\mathbf{x}^{[i-1]}} q(\mathbf{x}^{[i-1]}) \mathcal{K}(\mathbf{x}^{[i]}|\mathbf{x}^{[i-1]}) d\mathbf{x}^{[i-1]}. \quad (4.4)$$

Si la chaîne est *ergodique*, c’est-à-dire *irréductible* et *apériodique*, alors elle converge vers une distribution stationnaire ou invariante $p(\mathbf{x})$ [79]. Lorsque la chaîne a convergé, les réalisations aléatoires sont alors distribuées selon la loi stationnaire et on peut les exploiter afin d’approximer des calculs intégraux tels que définis dans la section précédente.

Remarque 23. À la différence des méthodes Monte Carlo, les échantillons ne sont pas indépendants. Pour autant, si la chaîne est ergodique $\mathbf{I}_N^{MCMC}(\varphi) = \frac{1}{N} \sum_{n=1}^N \varphi(\mathbf{x}_n) \xrightarrow[N \rightarrow \infty]{p.s.} \mathbf{I}(\varphi)$ (théorème ergodique [81, p. 182]).

Noyau de transition – Pour que la chaîne converge vers la distribution cible $p(\mathbf{x})$, il est nécessaire de trouver un noyau de transition $\mathcal{K}(\mathbf{x}^{[i]}|\mathbf{x}^{[i-1]})$ adéquat. Dans la littérature, on retrouve divers algorithmes qui, à partir de la densité cible $p(\mathbf{x})$, permettent de produire un tel noyau. L’un des avantages de ces algorithmes est qu’ils ne nécessitent de connaître $p(\mathbf{x})$ qu’à une constante multiplicative près. Parmi les plus connus, on peut citer dans l’ordre chronologique l’algorithme de MH (Metropolis-Hastings) [35], l’algorithme de Gibbs [35], l’algorithme HMC (Hamiltonian Monte Carlo ou Hybride Monte Carlo) [82] ou l’algorithme MALA (Metropolis-Hastings Adjusted Langevin Algorithm) [82].

Temps de chauffe – La chaîne nécessite un certain temps, appelé temps de chauffe, avant d’atteindre la stationnarité. En pratique, cela pose la question suivante : quand pouvons-nous avoir la certitude que nous avons convergé et que le temps de chauffe est atteint ? En effet, il arrive que la chaîne semble stabilisée et que finalement elle se dirige vers un autre centre de masse de probabilité. Dans la littérature, on retrouve plusieurs méthodes [35, p. 459-509] qui tentent de répondre à cette question (critères de convergence, utilisation de plusieurs chaînes, etc).

Concernant notre problème, on choisit de s’appuyer sur l’algorithme de Gibbs et les motivations de ce choix sont détaillées au début de la partie 4.2.

4.1.3 Algorithme de Gibbs

La particularité de l'algorithme de Gibbs est qu'il permet de décomposer le problème complexe de l'échantillonnage d'une loi de grande dimension en sous-problèmes d'échantillonnage de lois de plus petites dimensions. Pour cela, il s'appuie sur la décomposition de la loi jointe en lois conditionnelles. Lorsque cette décomposition fait apparaître des lois de formes standard, on peut tirer parti de méthodes d'échantillonnage efficaces localement afin d'échantillonner globalement une loi jointe de forme non standard. Afin d'illustrer son fonctionnement, on décrit ci-après l'algorithme de Gibbs dans un cas générique.

Cela se présente de la manière suivante dans un contexte multivarié. On souhaite obtenir des réalisations aléatoires de $p(\mathbf{x})$, avec $\mathbf{x} = [x_1, x_2, \dots, x_N]^t$. On suppose que les densités conditionnelles $p(x_n | \mathbf{x}_{-n})$ sont connues, pour $n = 1 \dots N$, et que l'on sait les échantillonner. On se donne la valeur initiale $\mathbf{x}^{[0]}$. L'algorithme de Gibbs est alors représenté par l'Algo. 1 ci-dessous. À chaque itération i , on échantillonne de manière séquentielle, pour $n = 1 \dots N$, les composantes $x_n^{[i]}$ selon la loi conditionnelle correspondante $p(x_n^{[i]} | \mathbf{x}_{1 \dots n-1}^{[i]}, \mathbf{x}_{n+1 \dots N}^{[i-1]})$. Pour l'étape n , les composantes de 1 à $n-1$ ont été échantillonnées à l'itération i et on tient compte de ces nouvelles valeurs, tandis que les composantes de $n+1$ à N n'ont pas encore été échantillonnées et on utilise les réalisations obtenues à l'itération $i-1$.

Après un certain temps de chauffe noté $i_0 \in 1 \dots I$, les échantillons que l'on note $\mathbf{x}^{[i_0 \dots I]}$ sont distribués sous la loi cible et stationnaire $p(\mathbf{x})$.

Algorithme 1 Pseudo-code de l'algorithme de Gibbs pour l'échantillonnage de $p(\mathbf{x})$.

Début

Initialisation : $\mathbf{x}^{[0]}$

POUR $i = 1, 2, \dots, I$ **FAIRE**

- $x_1^{[i]} \sim p(x_1^{[i]} | x_2^{[i-1]}, x_3^{[i-1]}, \dots, x_{N-1}^{[i-1]}, x_N^{[i-1]})$

- $x_2^{[i]} \sim p(x_2^{[i]} | x_1^{[i]}, x_3^{[i-1]}, \dots, x_{N-1}^{[i-1]}, x_N^{[i-1]})$

⋮

- $x_N^{[i]} \sim p(x_N^{[i]} | x_1^{[i]}, x_2^{[i]}, x_3^{[i]}, \dots, x_{N-1}^{[i]})$

FIN POUR

Sortie : $\mathbf{x}^{[1 \dots I]}$

Fin

Lien avec la partie précédente – Dans la partie précédente, on a introduit le fait que d'un instant i à l'autre $i+1$, les masses de probabilité évoluent d'après un noyau de transition. Dans le cas de l'algorithme de Gibbs, le noyau de transition est décomposé et il s'écrit à partir des lois

conditionnelles [35, p. 378]. En reprenant la loi $p(\boldsymbol{x})$, on obtient

$$\mathcal{K}(\boldsymbol{x}^{[i]}|\boldsymbol{x}^{[i-1]}) = \prod_{n=1}^N p(x_n^{[i]}|\boldsymbol{x}_{1\dots n-1}^{[i]}, \boldsymbol{x}_{n+1\dots N}^{[i-1]}). \quad (4.5)$$

Vitesse de convergence – La « vitesse » d’exploration de la loi stationnaire $p(\boldsymbol{x})$, et par extension la vitesse de convergence de la chaîne, est pilotée par la corrélation entre les composantes de \boldsymbol{x} . Plus la corrélation entre les composantes est forte, plus la probabilité d’effectuer de grands déplacements est faible et inversement. Cet aspect est discuté dans [83] ainsi que dans notre analyse théorique [3], voir plus loin. Dans le cas particulier où les composantes de \boldsymbol{x} sont totalement décorréliées, les lois conditionnelles sont égales aux lois marginales et la convergence est immédiate.

Variantes – Parmi les variantes de l’algorithme de Gibbs les plus connues, on peut citer :

- Le Metropolis-Hastings à l’intérieur du Gibbs (MH-within-Gibbs) ou Gibbs hybride [35, p. 392] qui consiste à remplacer une ou plusieurs étapes de l’algorithme de Gibbs par des étapes de Metropolis-Hastings, particulièrement lorsque l’on ne sait pas échantillonner les lois conditionnelles.

- Le Gibbs à balayage aléatoire pour lequel, à chaque itération, une composante sélectionnée aléatoirement est visitée, en opposition au Gibbs à balayage systématique présenté précédemment dans lequel, à chaque itération, chacune des composantes est visitée dans un ordre déterminé. Geman et Geman ont prouvé, dans le cas discret, la convergence vers la distribution jointe pour tous les schémas de balayage qui garantissent que les composantes sont visitées infiniment souvent pour un nombre d’itérations infinies [80, p. 150], [43].

- Le Gibbs à balayage hybride [84] qui représente un « compromis » entre le balayage systématique et le balayage aléatoire. Pour exemple [84] considère des variables latentes en plus de variables d’intérêt et à chaque itération, il échantillonne l’ensemble des variables latentes, ainsi qu’un bloc choisi aléatoirement des variables d’intérêts.

- Le Gibbs « Collapsed » (Collapsed Gibbs) [85] pour lequel une ou plusieurs étapes n’échantillonnent plus les lois conditionnelles associées, mais des lois conditionnelles marginalisées.

- Le Gibbs par bloc (Blocked Gibbs) qui consiste à décomposer le problème non pas en composantes de taille unitaire, mais en bloc. L’idée est de « bloquer » ensemble plusieurs composantes et de les échantillonner en une seule fois.

À noter que des variantes plus récentes se sont développées telles que le « Particle Gibbs » ou le « FUSS-within-Gibbs », présentés dans [86].

Analyse théorique – Nous avons mené un travail analytique sur l’algorithme de Gibbs, publié dans la revue IEEE SPL [3] et disponible en annexe H. Il fournit une analyse non asymptotique de l’algorithme de Gibbs dans le cas simple d’une distribution normale bivariée et caractérise pleinement les estimateurs. Il apporte une compréhension non empirique de l’échantillonneur de Gibbs ainsi que des réponses concernant le temps de chauffe dans ce cas d’école. Dans ce cadre, notre travail détermine la convergence vers la loi stationnaire cible et montre comment la corrélation entre les composantes joue un rôle clé sur la vitesse de convergence.

4.2 Méthode de Gibbs pour l'inversion-segmentation

Pour ce qui est de notre problème, on rappelle que l'on ne sait pas échantillonner directement la loi *a posteriori* $\pi(\mathbf{z}_{1\dots K}, \boldsymbol{\ell}, \gamma_\varepsilon, \boldsymbol{\gamma} | \boldsymbol{\sigma})$. Pour surmonter cette difficulté, on choisit de développer un algorithme de Gibbs. La principale motivation de ce choix est que la décomposition de la loi *a posteriori* en lois conditionnelles *a posteriori* est simple à réaliser et qu'elle fait apparaître des lois de forme standard pour lesquelles des stratégies d'échantillonnage efficaces existent [87].

Lois conditionnelles *a posteriori* – Afin d'obtenir les lois conditionnelles *a posteriori* nécessaires à l'algorithme de Gibbs, on s'appuie sur le fait que les densités sont proportionnelles à la densité jointe, Eq. (3.40). Par exemple, pour le paramètre γ_ε , on a

$$\pi(\gamma_\varepsilon | \boldsymbol{\sigma}, \mathbf{z}_{1\dots K}, \boldsymbol{\ell}, \boldsymbol{\gamma}) = \frac{\pi(\mathbf{z}_{1\dots K}, \boldsymbol{\ell}, \boldsymbol{\sigma}, \gamma_\varepsilon, \boldsymbol{\gamma})}{\pi(\boldsymbol{\sigma}, \mathbf{z}_{1\dots K}, \boldsymbol{\ell}, \boldsymbol{\gamma})}, \quad (4.6)$$

où $\pi(\boldsymbol{\sigma}, \mathbf{z}_{1\dots K}, \boldsymbol{\ell}, \boldsymbol{\gamma})$ peut être considérée comme une constante de normalisation.

La suite de cette partie est consacrée à définir les lois conditionnelles *a posteriori* pour $\boldsymbol{\gamma}$, $\mathbf{z}_{1\dots K}$ et $\boldsymbol{\ell}$, ainsi que les étapes d'échantillonnage associées. On propose également une manière efficace de réaliser les échantillonnages. De plus, on écrira les lois conditionnelles *a posteriori* à une constante multiplicative près (\propto).

Dans le but d'alléger les notations, on utilise l'écriture compacte de l'impédance par morceaux introduite dans le chapitre précédent, Eq. (3.19), que l'on rappelle ici

$$\mathbf{z} = \sum_{k=1}^K \mathbf{P}_k(\boldsymbol{\ell}) \mathbf{z}_k, \quad (4.7)$$

et l'on introduit le changement de variable suivant

$$\bar{\boldsymbol{\sigma}} = \boldsymbol{\sigma} - \bar{\boldsymbol{\sigma}}_0 = \boldsymbol{\sigma} - (\boldsymbol{\sigma}_0 - \mathbf{H} \mathbf{z}_0 + \mathbf{m}_\varepsilon). \quad (4.8)$$

4.2.1 Champs d'impédance

Les choix de lois *a priori* gaussiennes pour les impédances \mathbf{z}_k et pour le bruit ε mènent à des lois conditionnelles *a posteriori* gaussiennes. Pour le montrer, on repart de la loi jointe en ne sélectionnant que les termes faisant apparaître \mathbf{z}_k . On en déduit :

$$\pi(\mathbf{z}_k | \boldsymbol{\ell}, \mathbf{z}_{1\dots K \neq k}, \boldsymbol{\gamma}_k, \gamma_\varepsilon, \boldsymbol{\sigma}) \propto \exp -\gamma_\varepsilon \|\bar{\boldsymbol{\sigma}} - \mathbf{H} \sum_{p=1}^K \mathbf{P}_p(\boldsymbol{\ell}) \mathbf{z}_p\|_{\boldsymbol{\Pi}_\varepsilon}^2 - \gamma_k \|\mathbf{z}_k - \mathbf{m}_k\|_{\boldsymbol{\Pi}_k}^2. \quad (4.9)$$

Afin d'identifier la moyenne et la covariance *a posteriori*, on cherche à mettre l'argument de l'exponentielle, à une constante additive près ($\#$) vis-à-vis de \mathbf{z}_k , sous la forme quadratique suivante

$$\|\mathbf{z}_k - \mathbf{m}_k^*\|_{(\mathbf{C}_k^*)^{-1}}^2 \# \mathbf{z}_k^\dagger (\mathbf{C}_k^*)^{-1} \mathbf{z}_k - 2 \mathbf{z}_k^\dagger (\mathbf{C}_k^*)^{-1} \mathbf{m}_k^*. \quad (4.10)$$

D'après l'Eq. (4.9), on a

$$\begin{aligned}
& -\log \pi(\mathbf{z}_k | \boldsymbol{\ell}, \mathbf{z}_{1 \dots K \neq k}, \gamma_k, \gamma_\varepsilon, \boldsymbol{\sigma}) \\
& \# \gamma_\varepsilon \|\bar{\boldsymbol{\sigma}} - \mathbf{H} \sum_{p=1}^K \mathbf{P}_p(\boldsymbol{\ell}) \mathbf{z}_p\|_{\Pi_\varepsilon}^2 + \gamma_k \|\mathbf{z}_k - \mathbf{m}_k\|_{\Pi_k}^2 \\
& \# \mathbf{z}_k^\dagger \left(\gamma_\varepsilon \mathbf{P}_k(\boldsymbol{\ell}) \mathbf{H}^\dagger \Pi_\varepsilon \mathbf{H} \mathbf{P}_k(\boldsymbol{\ell}) + \gamma_k \Pi_k \right) \mathbf{z}_k \\
& - 2 \mathbf{z}_k^\dagger \left(\gamma_\varepsilon \mathbf{P}_k(\boldsymbol{\ell}) \mathbf{H}^\dagger \Pi_\varepsilon \left(\bar{\boldsymbol{\sigma}} - \mathbf{H} \sum_{p \neq k} \mathbf{P}_p(\boldsymbol{\ell}) \mathbf{z}_p \right) + \gamma_k \Pi_k \mathbf{m}_k \right).
\end{aligned} \tag{4.11}$$

Afin d'alléger les calculs, on introduit l'écart partiel aux données

$$\mathbf{e}_{\sigma_k} = \bar{\boldsymbol{\sigma}} - \mathbf{H} \sum_{p \neq k} \mathbf{P}_p(\boldsymbol{\ell}) \mathbf{z}_p. \tag{4.12}$$

Par identification, entre Eq. (4.10) et Eq. (4.11), on en déduit les lois conditionnelles *a posteriori* pour les impédances surfaciques \mathbf{z}_k

$$\mathbf{z}_k | \boldsymbol{\ell}, \mathbf{z}_{1 \dots K \neq k}, \gamma_k, \gamma_\varepsilon, \boldsymbol{\sigma} \sim \mathcal{N}(\mathbf{m}_k^*, \mathbf{C}_k^*), \tag{4.13}$$

avec la moyenne et la covariance *a posteriori* :

$$\mathbf{m}_k^* = \mathbf{C}_k^* \left(\gamma_\varepsilon \mathbf{P}_k(\boldsymbol{\ell}) \mathbf{H}^\dagger \Pi_\varepsilon \mathbf{e}_{\sigma_k} + \gamma_k \Pi_k \mathbf{m}_k \right), \tag{4.14}$$

$$\mathbf{C}_k^* = \left(\gamma_\varepsilon \mathbf{P}_k(\boldsymbol{\ell}) \mathbf{H}^\dagger \Pi_\varepsilon \mathbf{H} \mathbf{P}_k(\boldsymbol{\ell}) + \gamma_k \Pi_k \right)^{-1}. \tag{4.15}$$

De plus, on note $\mathbf{S}_k^* = \mathbf{C}_k^{*-1}$ la matrice de précision conditionnelle *a posteriori*.

Remarque 24. Plus le nombre de mailles appartenant à la classe k est important, plus la matrice de précision \mathbf{S}_k^* est pleine et inversement. Les deux cas extrêmes sont ceux où aucune maille n'est dans la classe k et où toutes les mailles sont dans la classe k . Dans le premier cas, on retrouve alors la loi *a priori* pour l'impédance \mathbf{z}_k , Eq. (3.9) page 43, et la matrice de précision est très creuse. Dans le second cas, la matrice de précision est pleine.

4.2.1.1 Difficultés d'échantillonnage

Dans cette partie, on s'intéresse de près à l'échantillonnage des impédances \mathbf{z}_k . Cette étape est critique, car elle est en grande dimension et plus coûteuse que les autres. Parmi les méthodes existantes, la plus directe et la plus répandue s'appuie sur la transformation affine d'une gaussienne i.i.d.. Cette transformation utilise la décomposition de Cholesky de la matrice de covariance. Dans notre cas, à partir de \mathbf{C}_k^* ou \mathbf{S}_k^* et \mathbf{m}_k^* , une réalisation peut être obtenue par l'une des deux procédures suivantes :

Procédure en précision :

$$\begin{aligned}
\mathbf{u} & \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_N) \\
\mathbf{S}_k^* & = \mathbf{A}_k^* \mathbf{A}_k^{*\dagger} \\
\mathbf{z}_k & = \mathbf{A}_k^{*-1} \mathbf{u} + \mathbf{m}_k^*
\end{aligned} \tag{4.16}$$

Procédure en covariance :

$$\begin{aligned}
\mathbf{u} & \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_N) \\
\mathbf{C}_k^* & = \mathbf{B}_k^* \mathbf{B}_k^{*\dagger} \\
\mathbf{z}_k & = \mathbf{B}_k^* \mathbf{u} + \mathbf{m}_k^*
\end{aligned} \tag{4.17}$$

En utilisant la structure particulière de la moyenne, Eq. (4.14), on peut rendre la transformation affine moins coûteuse. On prend pour exemple la procédure en précision, Eq. (4.16), et l'on écrit

$$\begin{aligned} \mathbf{z}_k &= \mathbf{A}_k^{*-†} \mathbf{u} + \mathbf{m}_k^* \\ &= \mathbf{A}_k^{*-†} \mathbf{u} + \left(\mathbf{A}_k^* \mathbf{A}_k^{*†} \right)^{-1} \left(\gamma_\varepsilon \mathbf{P}_k(\ell) \mathbf{H}^\dagger \mathbf{\Pi}_\varepsilon \mathbf{e}_{\sigma_k} + \gamma_k \mathbf{\Pi}_k \mathbf{m}_k \right) \\ &= \mathbf{A}_k^{*-†} \left(\mathbf{u} + \mathbf{A}_k^{*-1} \left(\gamma_\varepsilon \mathbf{P}_k(\ell) \mathbf{H}^\dagger \mathbf{\Pi}_\varepsilon \mathbf{e}_{\sigma_k} + \gamma_k \mathbf{\Pi}_k \mathbf{m}_k \right) \right). \end{aligned} \quad (4.18)$$

Le même développement pour la procédure en covariance, Eq. (4.17), s'écrit

$$\mathbf{z}_k = \mathbf{B}_k^* \left(\mathbf{u} + \mathbf{B}_k^{*†} \left(\gamma_\varepsilon \mathbf{P}_k(\ell) \mathbf{H}^\dagger \mathbf{\Pi}_\varepsilon \mathbf{e}_{\sigma_k} + \gamma_k \mathbf{\Pi}_k \mathbf{m}_k \right) \right) \quad (4.19)$$

Remarque 25. La similarité des procédures, Eq. (4.16) et (4.17), et des développements, Eq. (4.18) et (4.19), ne signifie pas pour autant que $\mathbf{B}_k^* = \mathbf{A}_k^{*-†}$. Les deux procédures d'échantillonnage fournissent bien des réalisations de la loi conditionnelle a posteriori, Eq. (4.13), mais elles sont différentes.

Complexité calculatoire – La procédure en covariance, Eq. (4.17), nécessite de calculer l'Eq. (4.15) et donc d'inverser une matrice pour obtenir \mathbf{C}_k^* , de calculer la décomposition de Cholesky de \mathbf{C}_k^* et de faire deux produits matrices triangulaires vecteurs, Eq. (4.19). La procédure en précision, Eq. (4.16), nécessite de calculer la matrice plus ou moins creuse \mathbf{S}_k^* , Eq. (4.15), de calculer la décomposition de Cholesky de \mathbf{S}_k^* et de résoudre deux systèmes triangulaires, Eq. (4.18).

Dans les deux cas, la complexité calculatoire est élevée. D'autant plus que le calcul est répété un grand nombre de fois : pour les K impédances et à chaque itération de l'algorithme de Gibbs. On rappelle la complexité calculatoire d'une inversion : $\mathcal{O}(N^3)$, d'une décomposition de Cholesky : $\mathcal{O}(N^3)$ et d'un produit ou d'une résolution matrice triangulaire vecteur : $\mathcal{O}(N^2/2)$. En covariance, la complexité calculatoire est alors de l'ordre de $\mathcal{O}(2N^3)$ et en précision, la complexité calculatoire est, dans le cas le moins favorable (Remarque 24), de l'ordre de $\mathcal{O}(N^3)$.

D'autres approches – On retrouve dans la littérature d'autres approches pour échantillonner des lois gaussiennes multidimensionnelles. Parmi elles, on peut citer l'algorithme PO (Perturbation Optimisation, [88]), permettant de passer d'un problème d'échantillonnage à un problème d'optimisation. Un élément clé de cet algorithme concerne la méthode d'optimisation mise en œuvre. Dans [88], les auteurs s'appuient sur une méthode du gradient conjugué qui converge en un nombre fini d'itérations, au moins égal à la dimension du système linéaire, menant à une complexité équivalente à la décomposition de Cholesky. Néanmoins, la procédure d'optimisation peut être arrêtée plus tôt sans perte de précision. Ce dernier élément permet de diminuer la complexité calculatoire, mais mène à un algorithme non prouvé convergé. Un algorithme plus récent, appelé RJPO (Reversible Jump Perturbation Optimization, [89]), s'appuie sur le même principe et est prouvé convergé grâce à l'ajout d'une étape d'acceptation-rejet. Outre la diminution de la complexité calculatoire, les algorithmes proposées permettent de réduire considérablement l'occupation mémoire.

L'algorithme RJPO représente un bon candidat pour échantillonner nos lois gaussiennes. Cependant, au vu de la dimension de notre problème ($N = 7400$), l'occupation mémoire n'est pas un élément bloquant. De plus, il ne permet pas d'exploiter la principale particularité de notre

problème qui est le faible nombre de données. Pour cette raison, nous avons développé une méthode d'échantillonnage adaptée, exploitant cette particularité.

4.2.1.2 Échantillonnage efficace en dimension élevée

L'approche que l'on propose s'appuie sur une note d'A. Doucet [90] qui reprend [91]. Elle permet d'obtenir « à moindre coût », comparativement aux stratégies présentées dans la partie précédente, des échantillons de la loi gaussienne conditionnelle *a posteriori*, Eq. (4.13).

Simulation conditionnelle générale telle qu'introduite par A. Doucet – Dans [90], il considère un vecteur gaussien w de moyenne m et de covariance Σ tel que

$$w = \begin{pmatrix} x \\ y \end{pmatrix}, \quad m = \begin{pmatrix} m_x \\ m_y \end{pmatrix} \quad \text{et} \quad \Sigma = \begin{pmatrix} \Sigma_{xx} & \Sigma_{xy} \\ \Sigma_{xy}^\dagger & \Sigma_{yy} \end{pmatrix}, \quad (4.20)$$

avec x prend des valeurs dans \mathbb{R}^{n_x} et y prend des valeurs dans \mathbb{R}^{n_y} . De plus, il montre que sachant $y = y_0$, on a

$$m_{x|y} = m_x + \Sigma_{xy} \Sigma_{yy}^{-1} (y_0 - m_y), \quad (4.21)$$

$$\Sigma_{x|y} = \Sigma_{xx} - \Sigma_{xy} \Sigma_{yy}^{-1} \Sigma_{xy}^\dagger. \quad (4.22)$$

Afin d'obtenir une réalisation de $\bar{x} \sim \mathcal{N}(m_{x|y}, \Sigma_{x|y})$, il propose l'algorithme suivant

-
- ① Échantillonnage de $w = \begin{pmatrix} x \\ y \end{pmatrix} \sim \mathcal{N}(m, \Sigma)$.
 - ② Échantillonnage de $\bar{x} = x + \Sigma_{xy} \Sigma_{yy}^{-1} (y_0 - y)$.
-

Cet algorithme permet de contourner le calcul de la matrice de covariance *a posteriori* $\Sigma_{x|y}$ et de sa décomposition de Cholesky qui, comme on l'a vu dans la partie précédente, peut-être une étape coûteuse en temps de calcul. En contrepartie, l'algorithme nécessite la résolution d'un système de la taille des données n_y . Dans notre cas, le nombre de données est faible et l'échantillonnage de w est peu coûteux, ce qui rend cette stratégie intéressante.

Remarque 26. Cette approche est à la base du filtre de Kalman d'Ensemble (EnKF) qui est appliqué en météorologie ou océanographie à des problèmes de grande dimension [92].

Adaptation à notre problème – Afin d'utiliser l'algorithme présenté ci-dessus, on développe le modèle direct

$$\sigma = \sigma_0 + H \left(\sum_{k=1}^K P_k(\ell) z_k - z_0 \right) + \varepsilon.$$

En s'appuyant sur le changement de variable introduit par l'Eq. (4.8), on obtient

$$\bar{\sigma} = H \sum_{k=1}^K P_k(\ell) z_k + \varepsilon_0,$$

avec $\varepsilon_0 \sim \mathcal{N}(\mathbf{0}, \gamma_\varepsilon^{-1} \mathbf{\Pi}_\varepsilon^{-1})$, puis on sépare la somme pour faire apparaître \mathbf{z}_k

$$\bar{\sigma} = \mathbf{H} \mathbf{P}_k(\ell) \mathbf{z}_k + \varepsilon_0 + \mathbf{H} \sum_{p \neq k} \mathbf{P}_p(\ell) \mathbf{z}_p.$$

En introduisant l'écart partiel aux données $e_{\sigma_k} = \bar{\sigma} - \mathbf{H} \sum_{p \neq k} \mathbf{P}_p(\ell) \mathbf{z}_p$, Eq. (4.12), on aboutit au modèle d'observation

$$e_{\sigma_k} = \mathbf{H} \mathbf{P}_k(\ell) \mathbf{z}_k + \varepsilon_0. \quad (4.23)$$

Remarque 27. L'écart partiel aux données e_{σ_k} est connu sachant $\ell, \mathbf{z}_{1 \dots K \neq k}, \gamma_k, \gamma_\varepsilon$ et σ .

À partir du modèle d'observation, Eq. (4.23) et par analogie à l'Eq. (4.20), on considère le vecteur gaussien \mathbf{w} de moyenne \mathbf{m} et de covariance Σ tel que

$$\mathbf{w} = \begin{pmatrix} \tilde{\mathbf{z}}_k \\ \tilde{\mathbf{e}}_{\sigma_k} \end{pmatrix}, \quad \mathbf{m} = \begin{pmatrix} \mathbf{m}_k \\ \mathbf{H} \mathbf{P}_k(\ell) \mathbf{m}_k \end{pmatrix} \quad (4.24)$$

et $\Sigma = \begin{pmatrix} \mathbf{C}_k & \mathbf{C}_k \mathbf{P}_k(\ell) \mathbf{H}^\dagger \\ \mathbf{H} \mathbf{P}_k(\ell) \mathbf{C}_k & \mathbf{H} \mathbf{P}_k(\ell) \mathbf{C}_k \mathbf{P}_k(\ell) \mathbf{H}^\dagger + \mathbf{C}_\varepsilon \end{pmatrix},$

avec $\mathbf{C}_\varepsilon = \gamma_\varepsilon^{-1} \mathbf{\Pi}_\varepsilon^{-1}$ et $\mathbf{C}_k = \gamma_k^{-1} \mathbf{\Pi}_k^{-1}$. On en déduit la procédure d'échantillonnage suivante :

- ① Échantillonnage du champ *a priori* : $\tilde{\mathbf{z}}_k \sim \mathcal{N}(\mathbf{m}_k, \gamma_k^{-1} \mathbf{\Pi}_k^{-1})$
- ② Échantillonnage de bruit centré : $\tilde{\mathbf{e}}_0 \sim \mathcal{N}(\mathbf{0}, \gamma_\varepsilon^{-1} \mathbf{\Pi}_\varepsilon^{-1})$
- ③ Simulation de l'observation : $\tilde{\mathbf{e}}_{\sigma_k} = \mathbf{H} \mathbf{P}_k(\ell) \tilde{\mathbf{z}}_k + \tilde{\mathbf{e}}_0$
- ④ Échantillonnage du champ conditionnel *a posteriori* :

$$\mathbf{z}_k = \tilde{\mathbf{z}}_k + \gamma_k^{-1} \mathbf{\Pi}_k^{-1} \mathbf{P}_k(\ell) \mathbf{H}^\dagger \left(\gamma_k^{-1} \mathbf{H} \mathbf{P}_k(\ell) \mathbf{\Pi}_k^{-1} \mathbf{P}_k(\ell) \mathbf{H}^\dagger + \gamma_\varepsilon^{-1} \mathbf{\Pi}_\varepsilon^{-1} \right)^{-1} (e_{\sigma_k} - \tilde{\mathbf{e}}_{\sigma_k})$$

Si l'étape ④ de la procédure peut s'écrire en forme covariance, on choisit de garder la forme précision pour considérer des matrices creuses, voir annexe E. L'échantillonnage de $\tilde{\mathbf{z}}_k$ et de $\tilde{\mathbf{e}}_0$ est réalisé grâce à la procédure décrite Eq. (4.16), reposant sur les décompositions de Cholesky des matrices de précision. Afin d'obtenir \mathbf{z}_k de manière plus efficace, on peut factoriser le terme

$$\mathbf{U}_k = \mathbf{H} \mathbf{P}_k(\ell) \mathbf{L}_k^{-\dagger}, \quad (4.25)$$

avec $\mathbf{\Pi}_k = \mathbf{L}_k \mathbf{L}_k^\dagger$ et l'on échantillonne alors l'impédance *a posteriori* par :

$$\mathbf{z}_k = \tilde{\mathbf{z}}_k + \gamma_k^{-1} \mathbf{L}_k^{-\dagger} \mathbf{U}_k^\dagger \left(\gamma_k^{-1} \mathbf{U}_k \mathbf{U}_k^\dagger + \gamma_\varepsilon^{-1} \mathbf{\Pi}_\varepsilon^{-1} \right)^{-1} (e_{\sigma_k} - \tilde{\mathbf{e}}_{\sigma_k}). \quad (4.26)$$

Complexité calculatoire – Les matrices $\mathbf{\Pi}_k$ et $\mathbf{\Pi}_\varepsilon$ étant fixées, leurs décompositions de Cholesky respectives peuvent être précalculées en amont du Gibbs. Afin d'analyser la complexité calculatoire de la procédure d'échantillonnage, on l'étudie en deux parties. Une première liée à l'échantillonnage des lois *a priori* et une seconde liée à l'échantillonnage de l'impédance z_k .

- Les échantillonnages de \tilde{z}_k et de $\tilde{\varepsilon}_0$ s'appuient sur les décompositions de Cholesky de $\mathbf{\Pi}_k$ et de $\mathbf{\Pi}_\varepsilon$. Ces dernières étant précalculées, le coût calculatoire de ces échantillonnages se ramène à celui de la résolution de systèmes triangulaires, voir l'Eq. (4.16).
 - Pour $\tilde{\varepsilon}_0$, cela dépend de la matrice $\mathbf{\Pi}_\varepsilon$ de taille $M \times M$. Si elle est pleine, sa décomposition de Cholesky l'est également et la résolution du système triangulaire a une complexité calculatoire de l'ordre de $\mathcal{O}(M^2/2)$. Inversement, si elle est creuse, la complexité calculatoire est faible. Dans la suite, on considère que la matrice $\mathbf{\Pi}_\varepsilon$ est creuse, menant à un échantillonnage très peu coûteux.
 - Pour \tilde{z}_k , si la matrice $\mathbf{\Pi}_k$ était pleine, sa décomposition de Cholesky le serait également et la complexité calculatoire (de la résolution du système) serait de l'ordre de $\mathcal{O}(N^2/2)$. Grâce à la modélisation markovienne choisie, la matrice $\mathbf{\Pi}_k$ et sa décomposition sont creuses, menant à une complexité bien moindre, voir annexe E, et donc à un échantillonnage peu coûteux.
- L'obtention de z_k nécessite la résolution de deux systèmes, voir l'Eq. (4.26). Pour plus de détails, le lecteur pourra se référer à l'annexe E. Le premier système est plein, mais de la taille des données M et le second est de taille N , mais extrêmement creux. L'étape la plus coûteuse correspond à l'Eq. (4.25) qui est de l'ordre de $\mathcal{O}(MN^2/2)$.

Finalement, la stratégie proposée, Eq. (4.26), fournit un échantillon de l'impédance z_k de manière efficace avec une complexité calculatoire de l'ordre de $\mathcal{O}(MN^2/2)$ ($M = 300$ et $N = 7400$). C'est une amélioration notable comparée à la complexité calculatoire de l'ordre de $\mathcal{O}(N^3)$ des procédures précédentes. Cela est d'autant plus intéressant que l'opération est répétée un grand nombre de fois : pour chaque classe et pour chaque itération de l'algorithme de Gibbs. À noter que cela est vrai seulement parce qu'il est peu coûteux d'échantillonner les lois *a priori*.

Remarque 28. *Dans l'optique de futurs développements, nous nous sommes intéressés à la complexité calculatoire des décompositions de Cholesky des matrices $\mathbf{\Pi}_k$. Ces dernières sont grandement creuses avec seulement 4 éléments non nuls par lignes et colonnes, ce dont tire parti l'algorithme de décomposition. Ces éléments non nuls ne sont pas rassemblés autour de la diagonale, mais dispersés en fonction de la structure du voisinage et de la structure du maillage. En conséquence, la complexité calculatoire est faible, mais difficile à estimer. En pratique, avec $N = 7400$, on observe que le temps de calcul nécessaire à la décomposition de Cholesky de $\mathbf{\Pi}_k$ est négligeable par rapport aux autres opérations. Par ailleurs, cette décomposition est réutilisée ultérieurement dans l'étape (4).*

4.2.2 Paramètres d'échelle

On rappelle que les paramètres d'échelle γ_ε et $\gamma_{1\dots K}$ pilotent la puissance du bruit et des K impédances. Les informations disponibles *a priori* sont encodées par des lois Gamma, voir l'Eq. (3.37) page 50, avec respectivement $\gamma_\varepsilon \sim \Gamma(a_\varepsilon, b_\varepsilon)$ et $\gamma_k \sim \Gamma(a_k, b_k)$. Pour traduire le fait

que l'on ne dispose que de peu d'information, on rappelle que l'on choisit $(a_\varepsilon, b_\varepsilon)$ et (a_k, b_k) proches de $(0, 0)$ afin de définir des lois *a priori* peu informatives. En pratique, on prendra $(a_\varepsilon, b_\varepsilon) = (a_k, b_k) = (10^{-16}, 10^{-16})$.

4.2.2.1 Bruit

À partir de la loi jointe, on obtient la densité conditionnelle *a posteriori*

$$\begin{aligned} \pi(\gamma_\varepsilon | \mathbf{z}_{1\dots K}, \boldsymbol{\ell}, \sigma) &\propto \gamma_\varepsilon^M \exp -\gamma_\varepsilon \|\bar{\boldsymbol{\sigma}} - \mathbf{H}\mathbf{z}\|_{\Pi_\varepsilon}^2 \quad \gamma_\varepsilon^{a_\varepsilon-1} \exp -b_\varepsilon \gamma_\varepsilon \\ &\propto \gamma_\varepsilon^{(a_\varepsilon+M)-1} \exp -\gamma_\varepsilon \left(b_\varepsilon + \|\bar{\boldsymbol{\sigma}} - \mathbf{H}\mathbf{z}\|_{\Pi_\varepsilon}^2 \right), \end{aligned} \quad (4.27)$$

et on retrouve une loi Gamma :

$$\gamma_\varepsilon | \mathbf{z}_{1\dots K}, \boldsymbol{\ell}, \sigma \sim \Gamma(a_\varepsilon^*, b_\varepsilon^*), \text{ avec } \begin{cases} a_\varepsilon^* = a_\varepsilon + M \\ b_\varepsilon^* = b_\varepsilon + \|\bar{\boldsymbol{\sigma}} - \mathbf{H}\mathbf{z}\|_{\Pi_\varepsilon}^2 \end{cases} \quad (4.28)$$

Échantillonnage direct – La loi conditionnelle *a posteriori* est fonction des paramètres a_ε^* et b_ε^* . À l'itération i , on les note $a_\varepsilon^{*[i]}$ et $b_\varepsilon^{*[i]}$, et l'on s'intéresse à leurs évolutions au fil des itérations. Le paramètre a_ε^* ne change pas, puisque l'on a $a_\varepsilon^{*[i]} = a_\varepsilon^{*[i-1]} = \dots = a_\varepsilon + M$, contrairement au paramètre b_ε^* qui est fonction de l'impédance par morceaux $\mathbf{z}^{[i]}$:

$$b_\varepsilon^{*[i]} = b_\varepsilon + \|\bar{\boldsymbol{\sigma}} - \mathbf{H}\mathbf{z}^{[i]}\|_{\Pi_\varepsilon}^2. \quad (4.29)$$

Une autre manière de calculer $b_\varepsilon^{*[i]}$ à l'itération i est de s'appuyer sur $b_\varepsilon^{*[i-1]}$ à l'itération $i-1$ et sur un résidu

$$\begin{aligned} \delta_{b_\varepsilon^{*}}^{[i]} &= b_\varepsilon^{*[i]} - b_\varepsilon^{*[i-1]} \\ &= \|\bar{\boldsymbol{\sigma}} - \mathbf{H}\mathbf{z}^{[i]}\|_{\Pi_\varepsilon}^2 - \|\bar{\boldsymbol{\sigma}} - \mathbf{H}\mathbf{z}^{[i-1]}\|_{\Pi_\varepsilon}^2 \\ &= (\mathbf{z}^{[i]} + \mathbf{z}^{[i-1]})^\dagger \underbrace{\mathbf{H}^\dagger \Pi_\varepsilon \mathbf{H}}_{\text{constante}} \delta_{\mathbf{z}}^{[i]} - 2 \underbrace{\bar{\boldsymbol{\sigma}}^\dagger \Pi_\varepsilon \mathbf{H}}_{\text{constante}} \delta_{\mathbf{z}}^{[i]}, \end{aligned} \quad (4.30)$$

avec $\delta_{\mathbf{z}}^{[i]} = \mathbf{z}^{[i]} - \mathbf{z}^{[i-1]}$ l'écart en impédance entre l'itération i et $i-1$. En exploitant cette forme, on a $b_\varepsilon^{*[i]} = b_\varepsilon^{*[i-1]} + \delta_{b_\varepsilon^{*}}^{[i]}$. L'avantage de ce calcul est qu'il fait apparaître des termes constants pouvant être précalculés. La complexité calculatoire de l'Eq. (4.29) est de l'ordre de $\mathcal{O}(M(N+M))$ tandis que celui du résidu, Eq. (4.30), est de l'ordre de $\mathcal{O}(N^2)$. Dans notre cas, le nombre de données M étant bien inférieur au nombre d'inconnues N , cette stratégie n'est pas efficace et l'on choisit de calculer le paramètre $b_\varepsilon^{*[i]}$ avec l'Eq. (4.29). Néanmoins, si le nombre de données devenait important, la procédure de mise à jour décrite ci-dessus pourrait devenir intéressante.

4.2.2.2 Impédances

Les densités conditionnelles *a posteriori* des paramètres d'échelle γ_k des impédances \mathbf{z}_k , s'écrivent respectivement pour $k = 1 \dots K$:

$$\begin{aligned} \pi(\gamma_k | \mathbf{z}_k) &\propto \gamma_k^N \exp -\gamma_k \|\mathbf{z}_k - \mathbf{m}_k\|_{\Pi_k}^2 \quad \gamma_k^{a_k-1} \exp -b_k \gamma_k \\ &\propto \gamma_k^{(a_k+N)-1} \exp -\gamma_k \left(b_k + \|\mathbf{z}_k - \mathbf{m}_k\|_{\Pi_k}^2 \right). \end{aligned} \quad (4.31)$$

Pour chaque paramètre γ_k , on retrouve une loi Gamma qui fait intervenir la moyenne et la covariance *a priori* : \mathbf{m}_k et $\mathbf{\Pi}_k$ de la classe correspondante. On rappelle que les matrices de structure s'écrivent $\mathbf{\Pi}_k = \mathbf{D}_k^t \mathbf{D}_k$. Les lois conditionnelles *a posteriori* des γ_k s'écrivent alors respectivement :

$$\gamma_k | \mathbf{z}_k \sim \Gamma(a_k^*, b_k^*), \text{ avec } \begin{cases} a_k^* = a_k + N \\ b_k^* = b_k + \|\mathbf{D}_k(\mathbf{z}_k - \mathbf{m}_k)\|^2 \end{cases} \quad (4.32)$$

Échantillonnage direct – Comme dans le cas de γ_ε , le paramètre a_k^* n'évolue pas au cours des itérations. Pour le paramètre b_k^* , une stratégie visant à calculer un résidu comme le montre l'Eq. (4.30) n'apporterait aucun bénéfice concernant le coût calculatoire. Néanmoins, le caractère creux des matrices \mathbf{D}_k rend le calcul des normes $\|\mathbf{D}_k(\mathbf{z}_k - \mathbf{m}_k)\|^2$ peu coûteux ; elles peuvent être calculées à chaque itération.

4.2.3 Champ d'étiquettes

On s'intéresse à la probabilité conditionnelle *a posteriori* des étiquettes ℓ . En partant de la loi jointe Eq. (3.40), on ne considère que les facteurs faisant apparaître les étiquettes. On en déduit la probabilité conditionnelle *a posteriori* que l'on note

$$\Pr(\mathbf{L} = \ell | \mathbf{z}_{1\dots K}, \gamma_\varepsilon, \boldsymbol{\sigma}) \propto \exp \beta \sum_{n \sim n'} \delta(\ell_n, \ell_{n'}) - \gamma_\varepsilon \|\boldsymbol{\sigma} - \mathbf{H} \sum_{p=1}^K \mathbf{P}_p(\ell) \mathbf{z}_p\|_{\mathbf{\Pi}_\varepsilon}^2. \quad (4.33)$$

Dans le terme de droite, on retrouve le terme d'attache aux données issu de la vraisemblance. Cela traduit le fait que la probabilité du champ d'étiquette est directement liée à l'écart entre les données et la sortie modèle. Lorsque les éléments $\mathbf{P}_p(\ell) \mathbf{z}_p$ sont en cohérence avec les données $\boldsymbol{\sigma}$, la probabilité augmente et inversement. Le terme de gauche est issu du modèle de Potts *a priori* pour les étiquettes. On rappelle qu'il introduit du lien entre les étiquettes en définissant des probabilités plus élevées pour des configurations dans lesquelles les étiquettes de mailles voisines ont la même valeur.

Pour une configuration du jeu d'étiquettes, calculer la probabilité associée est simple. Cependant, afin d'échantillonner le champ d'étiquette d'un coup, il faudrait calculer les probabilités des K^N configurations. Cela n'est réalisable que pour des cas en petite dimension et pour un nombre de classes faible. Pour cette raison, on choisit de découper le problème en sous-problèmes de taille unitaire : on exprime la loi conditionnelle *a posteriori* pour chaque étiquette sachant l'ensemble des autres paramètres. Dans notre cas, on s'intéresse à la maille n et à la classe associée à la valeur k , c'est-à-dire à la probabilité conditionnelle *a posteriori* que $L_n = k$. En repartant de l'Eq. (4.33), on déduit

$$\Pr(L_n = k | \ell_{-n}, \mathbf{z}_{1\dots K}, \gamma_\varepsilon, \boldsymbol{\sigma}) \propto \exp \beta \sum_{n' \sim n} \delta(k, \ell_{n'}) - \gamma_\varepsilon \|\boldsymbol{\sigma} - \mathbf{H} \sum_{p=1}^K \mathbf{P}_p(\ell) \mathbf{z}_p\|_{\mathbf{\Pi}_\varepsilon}^2, \quad (4.34)$$

avec $\ell = [\ell_1 \dots \ell_{n-1}, k, \ell_{n+1} \dots \ell_N]^t$.

La somme $\sum_{p=1}^K \mathbf{P}_p(\ell) \mathbf{z}_p$ fait intervenir l'ensemble des étiquettes. Afin d'isoler l'étiquette n dans l'Eq. (4.34), on écrit les matrices de projections comme

$$\mathbf{P}_p(\ell) = \delta(k, p) \mathbf{1}_n \mathbf{1}_n^t + \sum_{q=1, q \neq n}^N \delta(\ell_q, p) \mathbf{1}_q \mathbf{1}_q^t. \quad (4.35)$$

En injectant la nouvelle écriture des matrices de projection dans la construction de l'impédance par morceaux, on obtient :

$$\begin{aligned} \sum_{p=1}^K \mathbf{P}_p(\boldsymbol{\ell}) \mathbf{z}_p &= \sum_{p=1}^K \delta(k, p) \mathbb{1}_n z_p(n) + \sum_{q=1, q \neq n}^N \delta(\ell_q, p) \mathbb{1}_q z_p(q) \\ &= \mathbb{1}_n z_k(n) + \sum_{p=1}^K \sum_{q=1, q \neq n}^N \delta(\ell_q, p) \mathbb{1}_q z_p(q). \end{aligned} \quad (4.36)$$

En notant $\mathbf{z}_{n \rightarrow 0} = \sum_{p=1}^K \sum_{q=1, q \neq n}^N \delta(\ell_q, p) \mathbb{1}_q z_p(q)$ le champ par morceaux dont la $n^{\text{ième}}$ composante est nulle, on a

$$\sum_{p=1}^K \mathbf{P}_p(\boldsymbol{\ell}) \mathbf{z}_p = \mathbb{1}_n z_k(n) + \mathbf{z}_{n \rightarrow 0}. \quad (4.37)$$

Extraite de l'Eq. (4.34), la norme s'écrit à une constante additive près (#) comme

$$\| \bar{\boldsymbol{\sigma}} - \mathbf{H} \sum_{p=1}^K \mathbf{P}_p(\boldsymbol{\ell}) \mathbf{z}_p \|_{\Pi_\varepsilon}^2 \# |z_k(n)|^2 (\mathbf{H}^\dagger \Pi_\varepsilon \mathbf{H})_{n,n} - 2 \operatorname{Re}(z_k(n) (\bar{\boldsymbol{\sigma}} - \mathbf{H} \mathbf{z}_{n \rightarrow 0})^\dagger \Pi_\varepsilon \mathbf{H} \mathbb{1}_n) \quad (4.38)$$

et l'on en déduit la probabilité conditionnelle *a posteriori* à une constante multiplicative près

$$\begin{aligned} \Pr(L_n = k | \boldsymbol{\ell}_{-n}, \mathbf{z}_{1 \dots K}, \gamma_\varepsilon, \boldsymbol{\sigma}) \\ \propto \exp \beta \sum_{n' \sim n} \delta(k, \ell_{n'}) \\ - \gamma_\varepsilon \left(|z_k(n)|^2 (\mathbf{H}^\dagger \Pi_\varepsilon \mathbf{H})_{n,n} - 2 \operatorname{Re}(z_k(n) (\bar{\boldsymbol{\sigma}} - \mathbf{H} \mathbf{z}_{n \rightarrow 0})^\dagger \Pi_\varepsilon \mathbf{H} \mathbb{1}_n) \right). \end{aligned} \quad (4.39)$$

Échantillonnage – Afin d'échantillonner l'étiquette L_n conditionnellement au reste, il est nécessaire de connaître la constante multiplicative de l'Eq. (4.39). Pour des raisons numériques, on s'intéresse plutôt au logarithme à une constante additive près que l'on écrit

$$\begin{aligned} \operatorname{LPr}(L_n = k) &= \beta \sum_{n' \sim n} \delta(k, \ell_{n'}) \\ &- \gamma_\varepsilon \left(|z_k(n)|^2 (\mathbf{H}^\dagger \Pi_\varepsilon \mathbf{H})_{n,n} - 2 \operatorname{Re}(z_k(n) (\bar{\boldsymbol{\sigma}} - \mathbf{H} \mathbf{z}_{n \rightarrow 0})^\dagger \Pi_\varepsilon \mathbf{H} \mathbb{1}_n) \right). \end{aligned} \quad (4.40)$$

Pour toutes les configurations, on la normalise en retranchant la log-probabilité la plus élevée

$$\operatorname{LPr}_{\text{norm}}(L_n = k) = \operatorname{LPr}(L_n = k) - \sup_{k' \in \{1 \dots K\}} \operatorname{LPr}(L_n = k'), \quad (4.41)$$

et l'on obtient de manière exacte la probabilité *a posteriori* pour toutes les étiquettes et toutes les configurations en calculant :

$$p_k = \Pr(L_n = k | \boldsymbol{\ell}_{-n}, \mathbf{z}_{1 \dots K}, \gamma_\varepsilon, \boldsymbol{\sigma}) = \frac{\exp \operatorname{LPr}_{\text{norm}}(L_n = k)}{\sum_{k'=1}^K \exp \operatorname{LPr}_{\text{norm}}(L_n = k')}. \quad (4.42)$$

Finalement, on obtient une réalisation de L_n en échantillonnant la loi discrète $\mathcal{P}_{\mathcal{D}}$ définie par les probabilités p_k pour $k = 1 \dots K$:

$$\ell_n \sim \mathcal{P}_{\mathcal{D}}(p_1, \dots, p_K). \quad (4.43)$$

Calcul efficace de $\text{LPr}(L_n = k)$ – Le calcul de $\text{LPr}(L_n = k)$, Eq. (4.40), fait intervenir deux termes, un terme quadratique et un terme linéaire par rapport à $z_k(n)$. Le terme quadratique est peu coûteux à calculer, puisque $\mathbf{H}^\dagger \mathbf{\Pi}_\varepsilon \mathbf{H}$ peut être précalculé en amont de l'algorithme de Gibbs, tandis que le terme linéaire fait apparaître plusieurs produits. L'objectif de cette partie est de montrer comment le calculer de manière plus efficace.

La stratégie que l'on propose est de maintenir à jour le terme $(\bar{\sigma} - \mathbf{H}z_{n \rightarrow 0})^\dagger \mathbf{\Pi}_\varepsilon \mathbf{H} \mathbf{1}_n$ au fil des itérations sur les étiquettes. Pour cela, on s'appuie sur le « résidu total » $e_\sigma = \bar{\sigma} - \mathbf{H}z$ et l'on écrit

$$\begin{aligned} (\bar{\sigma} - \mathbf{H}z_{n \rightarrow 0})^\dagger \mathbf{\Pi}_\varepsilon \mathbf{H} \mathbf{1}_n &= (\bar{\sigma} - \mathbf{H}(z - \mathbf{1}_n z(n)))^\dagger \mathbf{\Pi}_\varepsilon \mathbf{H} \mathbf{1}_n \\ &= ((\bar{\sigma} - \mathbf{H}z) + \mathbf{H} \mathbf{1}_n z(n))^\dagger \mathbf{\Pi}_\varepsilon \mathbf{H} \mathbf{1}_n \\ &= (e_\sigma + \mathbf{H} \mathbf{1}_n z(n))^\dagger \mathbf{\Pi}_\varepsilon \mathbf{H} \mathbf{1}_n \\ &= e_\sigma^\dagger \mathbf{\Pi}_\varepsilon \mathbf{H} \mathbf{1}_n + z(n)^\dagger \mathbf{1}_n^t \mathbf{H}^\dagger \mathbf{\Pi}_\varepsilon \mathbf{H} \mathbf{1}_n. \end{aligned} \quad (4.44)$$

Remarque 29. Le calcul de l'Eq. (4.44) est peu coûteux puisque $\mathbf{\Pi}_\varepsilon \mathbf{H}$ et $\mathbf{H}^\dagger \mathbf{\Pi}_\varepsilon \mathbf{H}$ peuvent être précalculés en dehors de la boucle du Gibbs.

Dans l'algorithme de Gibbs, après chaque itération $n = 1 \dots N$ sur les étiquettes, la valeur du « résidu total » e_σ change en fonction de la nouvelle valeur de l'étiquette ℓ_n . À l'itération n , on tire une valeur ℓ_n pour L_n et la valeur mise à jour du « résidu total » s'écrit :

$$e_\sigma^{[n+1]} = e_\sigma^{[n]} - \mathbf{H} \mathbf{1}_n (z_{\ell_n}(n) - z(n)). \quad (4.45)$$

Finalement, la log-probabilité conditionnelle *a posteriori* s'écrit

$$\begin{aligned} \text{LPr}(L_n = k) \# \beta \sum_{n' \sim n} \delta(k, \ell_{n'}) \\ - \gamma_\varepsilon \left(|z_k(n)|^2 (\mathbf{H}^\dagger \mathbf{\Pi}_\varepsilon \mathbf{H})_{n,n} - 2 \text{Re}(z_k(n)) (e_\sigma^{[n] \dagger} \mathbf{\Pi}_\varepsilon \mathbf{H} \mathbf{1}_n + z(n)^\dagger (\mathbf{H}^\dagger \mathbf{\Pi}_\varepsilon \mathbf{H})_{n,n}) \right). \end{aligned} \quad (4.46)$$

L'Algo. 2 ci-après fait la synthèse de la méthode d'échantillonnage des étiquettes.

4.2.4 Pseudo-code

L'algorithme de Gibbs par bloc que nous avons développé pour l'inversion-segmentation est synthétisé par l'Algo. 3, où l'on retrouve les échantillonnages des lois conditionnelles *a posteriori* décrits dans les parties précédentes.

Temps de chauffe – Les échantillons sont tirés sous la loi *a posteriori* après un temps de chauffe que l'on considère atteint après l'itération $i_0 \in 1 \dots I$. Les échantillons que l'on exploite sont numérotés de i_0 à I : $\ell^{[i_0 \dots I]}$ et $\theta^{[i_0 \dots I]}$ et on note leur nombre $I_0 = I - i_0$.

Algorithme 2 Pseudo-code de l'algorithme de Gibbs pour les étiquettes à l'itération i .

DébutInitialisation : $\ell^{[i-1]}, \mathbf{z}_{1\dots K}^{[i]}, \gamma_\varepsilon^{[i]}$

- Calculer $\mathbf{z} = \sum_p^K \mathbf{P}_p(\ell^{[i-1]}) \mathbf{z}_p^{[i]}$ Eq. (4.7)

- Calculer $\mathbf{e}_\sigma^{[1]} = \bar{\boldsymbol{\sigma}} - \mathbf{H}\mathbf{z}$

POUR $n = 1 \dots N$ **FAIRE**

- Calculer $(\bar{\boldsymbol{\sigma}} - \mathbf{H}\mathbf{z}_{n \rightarrow 0})^\dagger \mathbf{\Pi}_\varepsilon \mathbf{H} \mathbf{1}_n = \mathbf{e}_\sigma^{[n]\dagger} \mathbf{\Pi}_\varepsilon \mathbf{H} \mathbf{1}_n + z(n)^\dagger \mathbf{1}_n^t \mathbf{H}^\dagger \mathbf{\Pi}_\varepsilon \mathbf{H} \mathbf{1}_n$ Eq. (4.44)

- Calculer $\text{LPr}(L_n = k)$, pour $k = 1 \dots K$ Eq. (4.46)

- Calculer $\text{LPr}_{\text{norm}}(L_n = k)$, pour $k = 1 \dots K$ Eq. (4.41)

- Calculer p_k , pour $k = 1 \dots K$ Eq. (4.42)

- Échantillonner $\ell_n^{[i]} \sim \mathcal{P}_D(p_1, \dots, p_K)$ Eq. (4.43)

- Mise à jour : $\mathbf{e}_\sigma^{[n+1]} = \mathbf{e}_\sigma^{[n]} - \mathbf{H} \mathbf{1}_n (z_{\ell_n^{[i]}}(n) - z(n))$ Eq. (4.45)

FIN POURSortie : $\ell^{[i]}$ **Fin**
Algorithme 3 Pseudo-code de l'algorithme de Gibbs.

DébutInitialisation : $\ell^{[0]}, \mathbf{z}_{1\dots K}^{[0]}$ **POUR** $i = 1, 2, \dots$ **FAIRE**

- Échantillonner $\gamma_\varepsilon^{[i]} \sim \pi(\gamma_\varepsilon^{[i]} | \mathbf{z}_{1\dots K}^{[i-1]}, \ell^{[i-1]}, \boldsymbol{\sigma})$ Eq. (4.28)

POUR $k = 1 \dots K$ **FAIRE**

- Échantillonner $\gamma_k^{[i]} \sim \pi(\gamma_k^{[i]} | \mathbf{z}_k^{[i-1]})$ Eq. (4.32)

- Échantillonner $\mathbf{z}_k^{[i]} \sim \pi(\mathbf{z}_k^{[i]} | \ell^{[i-1]}, \mathbf{z}_{1\dots k-1}^{[i]}, \mathbf{z}_{k+1\dots K}^{[i-1]}, \gamma_k, \gamma_\varepsilon^{[i]}, \boldsymbol{\sigma})$ Eq. (4.13)

FIN POUR

- Échantillonner $\ell^{[i]} \sim \text{Pr}(\mathbf{L} = \ell | \mathbf{z}_{1\dots K}^{[i]}, \gamma_\varepsilon^{[i]}, \boldsymbol{\sigma})$ Algo. 2

FIN POURSortie : $\ell^{[1\dots I]}, \mathbf{z}_{1\dots K}^{[1\dots I]}, \gamma_{1\dots K}^{[1\dots I]}, \gamma_\varepsilon^{[1\dots I]}$ **Fin**

4.3 Calcul des estimateurs

On rappelle que l'on souhaite calculer les estimateurs EAP et MMAP. Pour cela, on introduit leurs approximations numériques qui reposent sur les échantillons fournis par l'algorithme de Gibbs après le temps de chauffe.

4.3.1 Étiquettes

Estimateur – Pour chaque étiquette, on construit une distribution empirique *a posteriori* pour approcher la loi marginale $\pi(\ell_n|\sigma)$ définie par l'Eq. (3.45) à la page 52. Elle s'écrit :

$$\hat{\pi}(L_n = k|\sigma) = \frac{1}{I_0} \sum_{i=i_0}^I \delta(\ell_n^{[i]}, k), \text{ pour } k = 1 \dots K. \quad (4.47)$$

L'estimateur MMAP est déterminé, pour chaque étiquette, par la valeur d'étiquette la plus probable. On note $\hat{\ell} = [\hat{\ell}_1 \dots \hat{\ell}_N]$ le vecteur qui contient toutes les valeurs estimées pour les étiquettes et $\hat{\ell}_n$ la valeur estimée pour une seule étiquette qui s'approxime par

$$\hat{\ell}_n \simeq \arg \max_k \hat{\pi}(L_n = k|\sigma). \quad (4.48)$$

Incertitudes sur $\hat{\ell}_n$ – On propose de quantifier les incertitudes à partir de la distribution empirique *a posteriori* $\hat{\pi}$. Pour une étiquette $\hat{\ell}_n$, plus la probabilité $\hat{\pi}(L_n = \hat{\ell}_n|\sigma)$ est grande et plus la valeur estimée peut être considérée comme fiable et inversement.

Remarque 30. À noter que cette probabilité peut être insuffisante pour caractériser les incertitudes des étiquettes. On s'intéresse alors à la probabilité que l'étiquette prenne la valeur des autres classes. Par exemple pour $K = 3$, les deux configurations ci-dessous conduisent à la même estimation $\hat{\ell}_n = 2$ et à la même probabilité $\hat{\pi}(L_n = \hat{\ell}_n|\sigma)$, pourtant la configuration 2 semble plus incertaine.

$$\begin{bmatrix} \hat{\pi}(L_n = 1|\sigma) \\ \hat{\pi}(L_n = 2|\sigma) \\ \hat{\pi}(L_n = 3|\sigma) \end{bmatrix} : \underbrace{\begin{bmatrix} 0.25 \\ 0.5 \\ 0.25 \end{bmatrix}}_{\text{configuration 1}} \neq \underbrace{\begin{bmatrix} 0.01 \\ 0.5 \\ 0.49 \end{bmatrix}}_{\text{configuration 2}}$$

Incertitudes sur la classe k – On considère également le vecteur

$$\hat{\pi}(\mathbf{L} = k|\sigma) = [\hat{\pi}(L_1 = k|\sigma), \dots, \hat{\pi}(L_N = k|\sigma)],$$

qui approxime les probabilités *a posteriori* que chacune des étiquettes prenne la valeur k .

4.3.2 Impédance et hyperparamètres

Concernant les inconnues θ , nous avons introduit deux estimateurs EAP dans le chapitre précédent.

1. Le premier estimateur est défini par l'Eq. (3.47) et pour l'approximer, on s'appuie directement sur la moyenne empirique des échantillons en sortie de l'algorithme de Gibbs après temps de chauffe

$$\hat{\boldsymbol{\theta}} \simeq \frac{1}{I_0} \sum_{i=i_0}^I \boldsymbol{\theta}^{[i]}. \quad (4.49)$$

Les incertitudes associées peuvent être quantifiées par la matrice de covariance *a posteriori*, Eq. (3.48), que l'on approxime par :

$$\hat{\mathbf{C}}_{\boldsymbol{\theta}} = \frac{1}{I_0} \sum_{i=i_0}^I (\boldsymbol{\theta}^{[i]} - \hat{\boldsymbol{\theta}})(\boldsymbol{\theta}^{[i]} - \hat{\boldsymbol{\theta}})^\dagger. \quad (4.50)$$

2. Le second estimateur, défini par l'Eq. (3.49), est un estimateur conditionnel *a posteriori* sachant le jeu d'étiquettes obtenu au sens du MMAP. Afin de l'approximer, on cherche à s'appuyer sur des échantillons de la loi conditionnelle *a posteriori* $\pi(\boldsymbol{\theta}|\boldsymbol{\sigma}, \boldsymbol{\ell})$. Pour les obtenir, on considère deux possibilités :

- Ⓐ Sélectionner parmi les échantillons disponibles $\boldsymbol{\theta}^{[i]}$, pour $i \in i_0 \dots I$, ceux qui sont tels que la configuration d'étiquettes $\boldsymbol{\ell}^{[i]}$ soit égale à $\hat{\boldsymbol{\ell}}$.
- Ⓑ Réaliser des itérations de l'algorithme de Gibbs avec la configuration d'étiquettes $\hat{\boldsymbol{\ell}}$ « bloquée », ce qui revient à obtenir des réalisations aléatoires de la loi $\pi(\boldsymbol{\theta}|\boldsymbol{\sigma}, \hat{\boldsymbol{\ell}})$. Dans ce cas, il serait nécessaire d'attendre un nouveau temps de chauffe.

L'approche Ⓐ conduit à une variance importante de l'estimateur en raison du faible nombre d'échantillons $\boldsymbol{\ell}^{[i]}$ identiques à $\hat{\boldsymbol{\ell}}$. On lui préfère l'approche Ⓑ même si elle exige des calculs supplémentaires et un algorithme de Gibbs dédié. À partir de ces nouveaux échantillons notés $\boldsymbol{\theta}^{[i]}(\hat{\boldsymbol{\ell}})$, l'EAP conditionnel *a posteriori* s'approxime par

$$\hat{\boldsymbol{\theta}}(\hat{\boldsymbol{\ell}}) \simeq \frac{1}{I'_0} \sum_{i=i'_0}^{I'} \boldsymbol{\theta}^{[i]}(\hat{\boldsymbol{\ell}}), \quad (4.51)$$

avec i'_0 , I'_0 et I' , les nouveaux paramètres algorithmiques associés à cet échantillonneur. Les incertitudes associées peuvent être quantifiées par la matrice de covariance *a posteriori*, Eq. (3.50), qui s'approxime par :

$$\hat{\mathbf{C}}_{\boldsymbol{\theta}}(\hat{\boldsymbol{\ell}}) = \frac{1}{I'_0} \sum_{i=i'_0}^{I'} (\boldsymbol{\theta}^{[i]}(\hat{\boldsymbol{\ell}}) - \hat{\boldsymbol{\theta}}(\hat{\boldsymbol{\ell}}))(\boldsymbol{\theta}^{[i]}(\hat{\boldsymbol{\ell}}) - \hat{\boldsymbol{\theta}}(\hat{\boldsymbol{\ell}}))^\dagger. \quad (4.52)$$

Remarque 31. Dans la Remarque 21, on introduit les intervalles de crédibilité. Ces intervalles peuvent être calculés par comptage en utilisant les réalisations $\boldsymbol{\theta}^{[i]}$ de $\pi(\boldsymbol{\theta}|\boldsymbol{\sigma})$ ou de $\pi(\boldsymbol{\theta}|\boldsymbol{\sigma}, \hat{\boldsymbol{\ell}})$.

CHAPITRE 5

RÉSULTATS NUMÉRIQUES

Ce chapitre est consacré aux résultats numériques obtenus grâce à la méthode proposée. L'objectif est d'analyser sa capacité à reconstruire une impédance surfacique homogène par morceaux. Pour cela, on considère un cas simulé, dit de validation, qui s'appuie sur le modèle d'observation et sur les modèles a priori pour le bruit, les impédances et les étiquettes. La méthode est ensuite appliquée et l'on s'intéresse à la convergence de l'algorithme, à l'erreur d'estimation et à la quantification des incertitudes.

Dans un premier temps, on présente la construction du cas de validation, c'est-à-dire la construction d'impédances et d'étiquettes vraies ainsi que l'obtention de données simulées. On précise également l'initialisation et les paramètres algorithmiques choisis. Dans un second temps, on propose une analyse des résultats afin de mettre en avant les capacités et limitations de la méthode proposée.

Sommaire

5.1	Cas de validation	74
5.1.1	Impédances et étiquettes vraies	74
5.1.2	Données synthétiques	76
5.1.3	Algorithme de Gibbs	77
5.2	Validation de la méthode	78
5.2.1	Paramètre d'échelle	78
5.2.2	Étiquettes	80
5.2.3	Impédance EAP	82
5.2.3.1	Impédances complètes	82
5.2.3.2	Impédance par morceaux	84
5.2.4	Impédances EAP conditionnelle	85
5.3	Conclusion	87
5.4	Application à des données plus réelles	88

5.1 Cas de validation

Afin d'analyser les capacités de la méthode, on s'appuie sur des données simulées à partir du modèle d'observation et des modèles *a priori* pour les impédances, les étiquettes et le bruit. On considère une impédance constituée de $K = 2$ classes et l'on note z_1^* , z_2^* et ℓ^* les valeurs vraies.

5.1.1 Impédances et étiquettes vraies

Étiquette – Pour cette étude, on souhaite que les deux régions soient de tailles similaires, c'est-à-dire que $\bar{N}_1(\ell^*) \approx \bar{N}_2(\ell^*)$, avec

$$\bar{N}_k(\ell) = \frac{1}{N} \sum_{n=1}^N \delta(\ell_n, k), \quad (5.1)$$

la proportion d'étiquettes égales à k . De plus, on veut modéliser des régions de grande taille pour chaque matériau. On rappelle que la taille des régions est pilotée par le coefficient de granularité β et que plusieurs cartes d'étiquette en fonction de β sont présentées sur la Fig. 3.2 page 42. On choisit $\beta = 1.3$ afin d'obtenir la carte ℓ^* correspondant à nos attentes et on la représente sur la Fig. 5.1. Pour quantifier la taille des régions, on considère la proportion de paires de mailles voisines de même valeur d'étiquette

$$\Delta(\ell) = \frac{1}{N_c} \sum_{n \sim n'} \delta(\ell_n, \ell_{n'}), \quad (5.2)$$

avec N_c que l'on rappelle être le nombre de paires de mailles voisines. Cette quantité permet d'évaluer le degré de compacité des régions et est proportionnelle à N_c moins la longueur des contours.

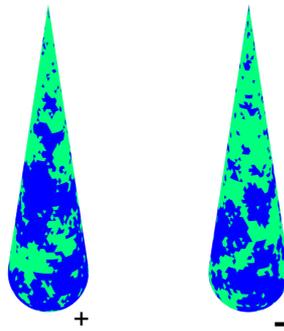


FIGURE 5.1 – Affichage selon l'orientation (+) et (-) de la configuration d'étiquette vraie ℓ^* , avec $\beta = 1.3$ et $K = 2$, obtenue pour 100 000 itérations d'un algorithme de Gibbs. On a $\bar{N}_1(\ell^*) = 0.495$, $\bar{N}_2(\ell^*) = 0.505$ et $\Delta(\ell^*) = 0.862$.

Impédances complètes – L'objet est constitué de deux matériaux représentés par les impédances z_1^* et z_2^* . Afin qu'elles correspondent à des propriétés différentes, on prend des paramètres différents. On rappelle que α règle la corrélation et que plus sa valeur est proche de 1, plus

l'impédance varie lentement et inversement. Le paramètre γ règle l'échelle de variation de l'impédance et plus il est grand, plus les valeurs d'impédance sont concentrées autour de la référence et inversement. Ainsi, on choisit respectivement pour z_1^* et z_2^* les paramètres $\gamma_1^* = 1.10^3$, $\alpha_1 = 0.5$ et $\gamma_2^* = 2.10^4$, $\alpha_2 = 0.999$. Les impédances obtenues sont représentées par la Fig. 5.2 et l'on constate que z_1^* , affiché à gauche, est plus granuleux et varie davantage que z_2^* , affiché à droite.

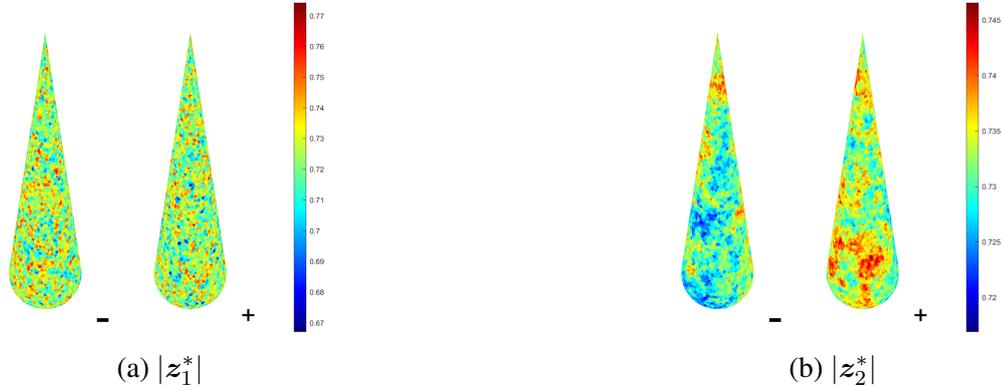


FIGURE 5.2 – Module des impédances (a) z_1^* et (b) z_2^* avec les paramètres $\gamma_1^* = 1.10^3$, $\alpha_1 = 0.5$ et $\gamma_2^* = 2.10^4$, $\alpha_2 = 0.999$. Les deux matériaux sont représentés par deux impédances avec des rugosités et des échelles de variation différentes. Ces réalisations sont obtenues grâce à la procédure présentée par l' Eq. (4.16) à la page 61.

Impédance par morceaux – D'après la construction définie par l'Eq. (3.19), page 45, l'impédance par morceaux vraie z^* s'obtient par $z^* = \sum_{k=1}^K P_k(\ell^*) z_k^*$ et on la représente sur la Fig. 5.3. On constate qu'il y a bien deux comportements différents de l'impédance pour représenter les deux propriétés matériaux. Une première partie, majoritairement dans la partie supérieure, avec une évolution plutôt douce et de petites variations autour de la référence. Une deuxième partie, principalement dans la partie inférieure, plus chahutée et avec des variations plus importantes.

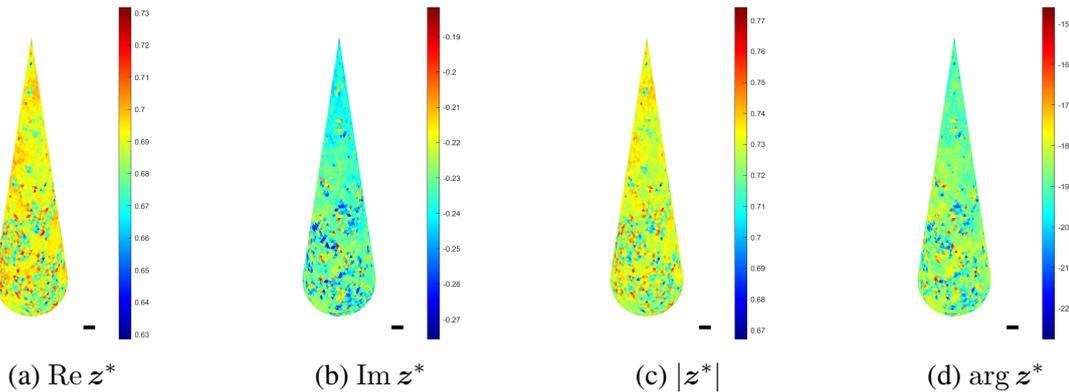


FIGURE 5.3 – Partie réelle, partie imaginaire, module et phase en degré de l'impédance z^* obtenue à partir de z_1^* et z_2^* , représentés sur la Fig. 5.2, et de ℓ^* , représenté sur la Fig. 5.1.

5.1.2 Données synthétiques

La simulation des données repose sur le modèle d'observation $\sigma = \sigma_0 + \mathbf{H}(z^* - z_0) + \varepsilon$, avec ε blanc centré et homogène, c'est-à-dire $\mathbf{m}_\varepsilon = \mathbf{0}$ et $\mathbf{\Pi}_\varepsilon = \mathbf{I}_M$. Pour que le bruit soit de faible puissance, on règle le facteur d'échelle γ_ε à partir du RSB¹ = 65 dB entre la sortie modèle non bruitée $\sigma_0 + \mathbf{H}(z^* - z_0)$ et ε , ce qui nous donne $\gamma_\varepsilon^* = 1.47 \cdot 10^{-2}$.

On affiche les données sous deux formes. La première est un hologramme, Fig. 5.4, où la valeur des coefficients de réflexion est représentée en fonction de l'angle et de l'altitude. La seconde est une représentation en fonction de l'angle et pour toutes les altitudes superposées, Fig. 5.5. On affiche également σ_0 , $\mathbf{H}(z^* - z_0)$ et ε pour illustrer les contributions de z^* et ε , masquées par les fortes variations de σ_0 entre les différentes altitudes (visibles sur les Fig. 5.4 et 5.5).

On observe sur les Fig. 5.4 et 5.5 que les coefficients de référence σ_0 sont constants par altitude, puisque l'objet est invariant par rotation et z_0 est uniforme, voir la Remarque 6 page 24. En pratique, il existe de petites variations en fonction de l'angle dues à la discrétisation de la surface imparfaitement axisymétrique. Les distances antenne-objet ne sont alors pas constantes, entraînant de petites variations de σ_0 (invisibles sur les Fig. 5.4 et 5.5).

La quantité $\mathbf{H}(z^* - z_0)$ représente la différence $\sigma - \sigma_0$ au bruit près. Elle décrit l'évolution des coefficients de réflexion pour des variations de l'impédance autour de la référence. Les représentations des Fig. 5.4 et 5.5 montrent l'impact de ces variations sur les données. Lorsque $z^* - z_0$ est uniforme, la différence $\sigma - \sigma_0$ est constante par altitude. Ici, la différence en impédance n'est pas constante et les données σ diffèrent de σ_0 en angle et en altitude.

Enfin, on observe sur la Fig. 5.5 (à droite) que le bruit est de faible puissance puisque les courbes noires et rouges sont presque superposées.

1. Le Rapport Signal sur Bruit en dB vaut $\text{RSB} = 10 \log_{10} r_\sigma / r_\varepsilon$, avec $r_\sigma = \|\sigma_0 + \mathbf{H}(z^* - z_0)\|^2 / M$ et $r_\varepsilon = 1/\gamma_\varepsilon$.

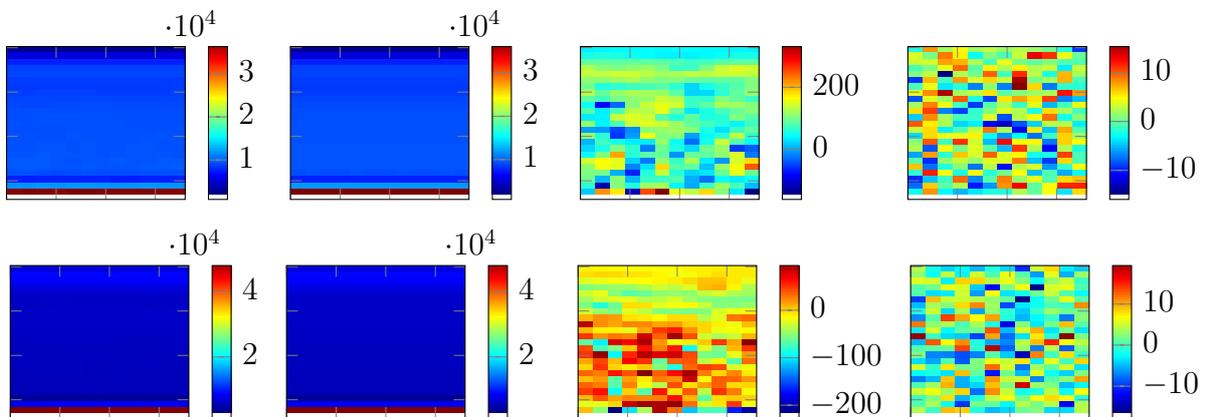


FIGURE 5.4 – Données sous forme d'hologramme en partie réelle en haut et en partie imaginaire en bas. Les hologrammes sont fonction de l'angle en abscisse et de l'altitude en ordonnée. On retrouve de gauche à droite σ , σ_0 , $\mathbf{H}(z^* - z_0)$ et ε .

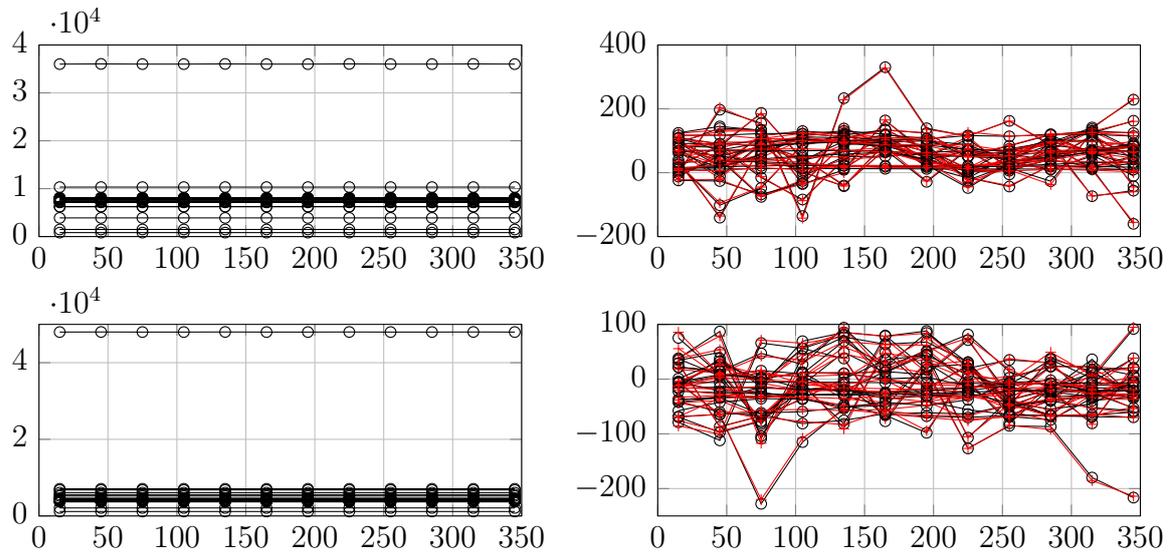


FIGURE 5.5 – Représentation à gauche des coefficients de réflexion de références σ_0 et à droite de $\mathbf{H}(z^* - z_0)$ en noir et de $\mathbf{H}(z^* - z_0) + \varepsilon$ en rouge. La partie réelle est affichée en haut et la partie imaginaire en bas. On les représente pour toutes les altitudes superposées en fonction de l'angle (en abscisse); chaque altitude est tracée par \circ et $+$. Les figures de gauche n'apportent pas d'information supplémentaire sur σ_0 par rapport aux hologrammes de la Fig. 5.4, tandis que les figures de droite montrent la somme des hologrammes de $\mathbf{H}(z^* - z_0)$ et ε .

5.1.3 Algorithme de Gibbs

L'algorithme de Gibbs nécessite l'initialisation d'une partie des inconnues et l'on choisit d'initialiser les étiquettes et les K impédances.

Initialisation des impédances – On initialise les K impédances par l'EAP dans le cas particulier d'une seule classe, ce qui mène à une solution douce sur l'ensemble de l'objet. Pour y parvenir, on utilise l'algorithme de Gibbs avec $K = 1$ et l'on nomme ce cas particulier : Gibbs_{init}. Le principe est le suivant.

- Pour initialiser le Gibbs_{init}, on initialise l'impédance. Une solution serait de prendre la référence z_0 , mais cela mène à $\|\mathbf{D}_k(z_k - m_k)\|^2 = 0$, Eq. (4.32) page 67, et fournit une initialisation, en moyenne, très éloignée de la valeur attendue pour le facteur d'échelle. À la place, on choisit un bruit blanc gaussien centré sur la référence, dont la variance est réglée pour concentrer la masse de probabilité dans les valeurs à $\pm 10\%$ de z_0 , voir la partie 1.4.3 page 23 du Chapitre 1.

- Pour l'unique paramètre d'interaction, on choisit la moyenne $\alpha = (\alpha_1 + \alpha_2)/2 = 0.75$.

- On effectue I_{init} itérations du Gibbs_{init} et l'on calcule une moyenne empirique de l'impédance à partir des $I_{\text{init}}/2$ derniers échantillons, voir la Fig. 5.6b. En pratique, $I_{\text{init}} = 1000$.

Remarque 32. En l'absence d'étiquettes, certains calculs de l'étape ④ de l'échantillonnage des impédances a posteriori, page 64, peuvent être réalisés en dehors de la boucle. En particulier l'opération la plus coûteuse correspondante à l'Eq. (4.25). Cela rend le Gibbs_{init} peu coûteux en temps de calcul, de l'ordre du centième de seconde par itération².

2. Calculs effectués sous Matlab avec un processeur i5 à 6 cœurs cadencés à 4.8 GHz.

Initialisation des étiquettes – On initialise les étiquettes par une réalisation d'un champ blanc équiprobable représenté sur la Fig. 5.6a.

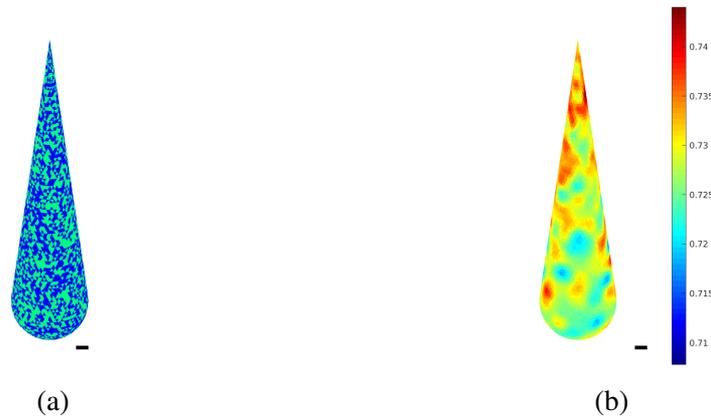


FIGURE 5.6 – Initialisation des étiquettes à gauche et des K impédances à droite.

5.2 Validation de la méthode

Cette partie présente les résultats des estimateurs EAP pour les impédances et les hyperparamètres, et MMAP pour les étiquettes, ainsi que les incertitudes associées. On présente également les reconstructions obtenues pour l'EAP conditionnelle pour les impédances et ses différences avec l'EAP. Les résultats présentés sont ceux d'une unique chaîne (chaîne 1), se reporter à l'annexe F page 111 pour des résultats complémentaires. On rappelle qu'ils sont calculés à partir des réalisations après le temps de chauffe et pour alléger l'écriture, on appellera estimées les approximations obtenues. Enfin, on ajoute que le problème que l'on résout est difficile en raison du faible nombre de données vis-à-vis du nombre d'inconnues (environ 25 fois moins) et des dégradations induites par le modèle d'observation.

Le nombre d'itérations de l'algorithme $I_{\max} = 140\,000$ et le temps de chauffe $i_0 = 120\,000$ sont choisis en s'appuyant sur l'expérience acquise en exploitant plusieurs chaînes. On précise que la méthode est implémentée en Matlab et que le temps de calcul d'une itération est d'environ 0.35 sec^2 . Les 140 000 itérations représentent alors une dizaine d'heures de calculs.

5.2.1 Paramètre d'échelle

On représente sur la Fig. 5.7, l'évolution des paramètres γ_1 , γ_2 et γ_ε . Les figures de gauche montrent une convergence relativement rapide. Autant que l'on puisse le constater, l'algorithme converge au bout de quelques centaines d'itérations. Les mêmes observations sont faites pour d'autres chaînes présentées dans la partie F.1 de l'annexe F.

On donne dans la Table 5.1, les estimées ainsi que les erreurs d'estimation. On constate que la méthode fournit des estimées proches des valeurs vraies et que l'erreur relative est petite pour $\hat{\gamma}_1$ et $\hat{\gamma}_2$, mais plus élevée pour $\hat{\gamma}_\varepsilon$. Puisque le bruit est de faible puissance, celui-ci est noyé dans le signal et son estimation est difficile.

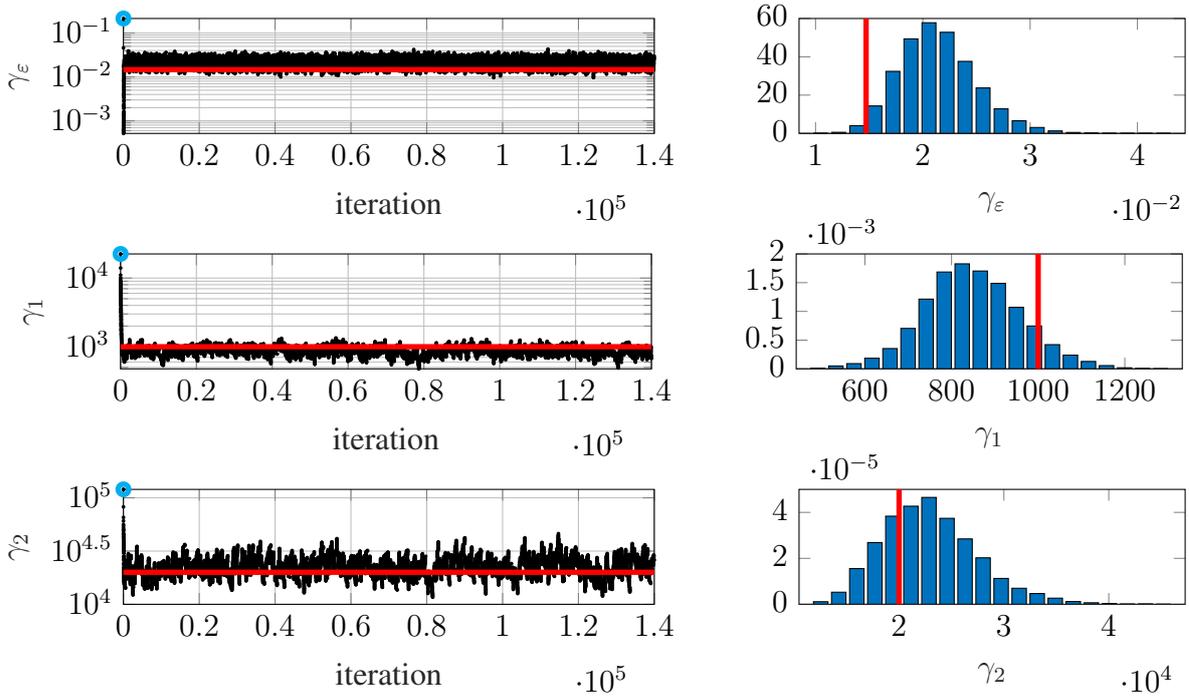


FIGURE 5.7 – Paramètre γ_1 , γ_2 et γ_ϵ sur les $I_{\max} = 140\,000$ itérations à gauche et sur les 70 000 dernières itérations pour les histogrammes à droite (première valeur : \bullet , valeur vraie : —).

Valeur vraie	Valeur estimée	Écart relatif
$\gamma_\epsilon^* = 1.47 \cdot 10^{-2}$	$\hat{\gamma}_\epsilon = 2.15 \cdot 10^{-2}$	$ \gamma_\epsilon^* - \hat{\gamma}_\epsilon /\gamma_\epsilon^* = 45.9\%$
$\gamma_1^* = 1 \cdot 10^3$	$\hat{\gamma}_1 = 0.83 \cdot 10^3$	$ \gamma_1^* - \hat{\gamma}_1 /\gamma_1^* = 16.6\%$
$\gamma_2^* = 2 \cdot 10^4$	$\hat{\gamma}_2 = 2.22 \cdot 10^4$	$ \gamma_2^* - \hat{\gamma}_2 /\gamma_2^* = 10.8\%$

TABLE 5.1 – Estimation et erreur d'estimation pour les paramètres d'échelles γ_1 , γ_2 et γ_ϵ .

Les histogrammes, en lien avec les lois marginales empiriques *a posteriori* $\hat{\pi}(\gamma|\sigma)$, montrent que les masses de probabilités sont relativement concentrées, indiquant des incertitudes limitées. Afin de les quantifier, on fournit les écarts-types et l'on détermine pour chaque paramètre d'échelle un intervalle de crédibilité [35, p. 17] à queues égales $\mathcal{D}_\gamma = [\gamma_m, \gamma_M]$ tel que

$$\Pr(\gamma < \gamma_m|\sigma) = \Pr(\gamma > \gamma_M|\sigma) = \lambda/2, \text{ avec } \lambda = 0.05.$$

On reporte dans la Table 5.2 les intervalles calculés et l'on constate qu'ils contiennent les valeurs vraies γ_1^* et γ_2^* , mais que γ_ϵ^* se situe légèrement en dehors de l'intervalle. Néanmoins, celui-ci reste proche, comme on peut également le voir sur les histogrammes.

Écart-type	Intervalle de crédibilité
$\hat{\sigma}_{\gamma_\epsilon} = 0.34 \cdot 10^{-2}$	$\mathcal{D}_{\gamma_\epsilon} = [\hat{\gamma}_\epsilon - 1.8 \hat{\sigma}_{\gamma_\epsilon}, \hat{\gamma}_\epsilon + 2.2 \hat{\sigma}_{\gamma_\epsilon}] = [1.54 \cdot 10^{-2}, 2.9 \cdot 10^{-2}]$
$\hat{\sigma}_{\gamma_1} = 0.11 \cdot 10^3$	$\mathcal{D}_{\gamma_1} = [\hat{\gamma}_1 - 2.1 \hat{\sigma}_{\gamma_1}, \hat{\gamma}_1 + 2.1 \hat{\sigma}_{\gamma_1}] = [0.61 \cdot 10^3, 1.05 \cdot 10^3]$
$\hat{\sigma}_{\gamma_2} = 0.44 \cdot 10^4$	$\mathcal{D}_{\gamma_2} = [\hat{\gamma}_2 - 1.8 \hat{\sigma}_{\gamma_2}, \hat{\gamma}_2 + 2.2 \hat{\sigma}_{\gamma_2}] = [1.43 \cdot 10^4, 3.18 \cdot 10^4]$

TABLE 5.2 – Quantification des incertitudes associées aux estimées $\hat{\gamma}_1$, $\hat{\gamma}_2$ et $\hat{\gamma}_\epsilon$.

On rappelle que les lois Gamma *a priori* pour ces paramètres sont choisies non informatives avec $(a, b) = (10^{-16}, 10^{-16})$. Les moyennes et écarts-types *a priori* correspondants valent respectivement 1 et 10^8 . On constate alors que les données sont suffisamment informatives pour obtenir des estimées proches des valeurs vraies avec des incertitudes relativement faibles.

5.2.2 Étiquettes

On s'intéresse aux résultats pour les étiquettes. Afin d'évaluer empiriquement la convergence de la chaîne, on s'appuie sur les quantités $\bar{N}_1(\ell)$, $\bar{N}_2(\ell)$ et $\Delta(\ell)$ au fil des itérations, représentées sur la Fig. 5.8.

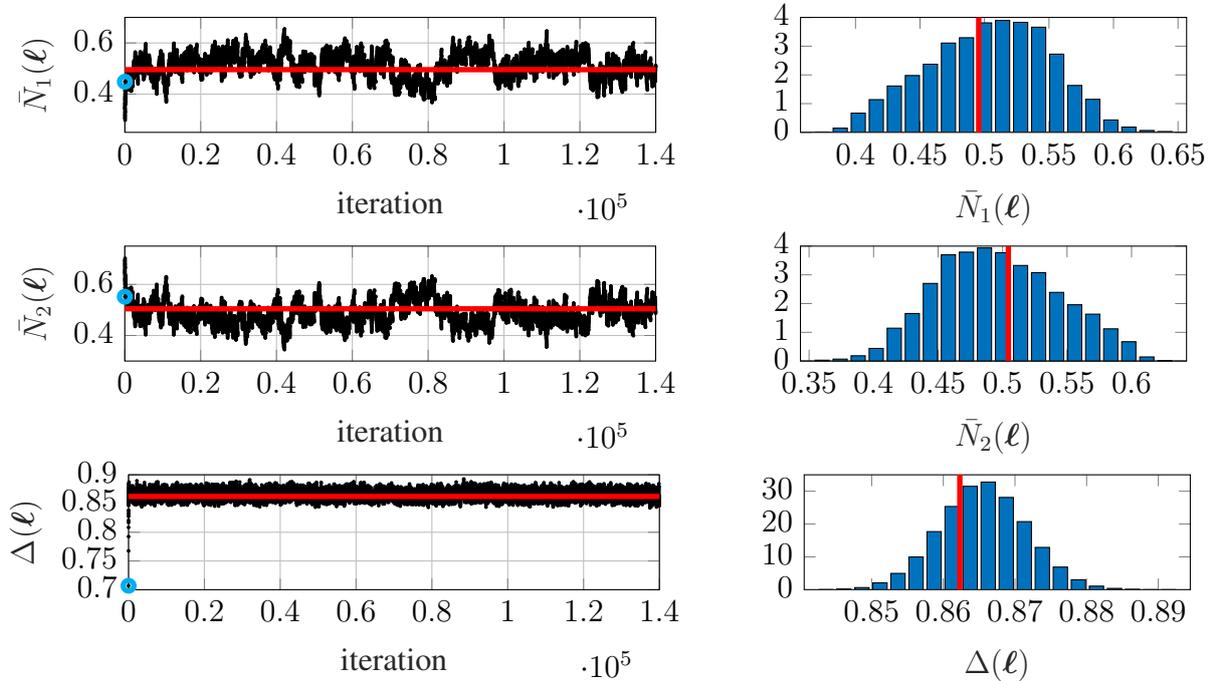


FIGURE 5.8 – Évolution de $\bar{N}_1(\ell)$, $\bar{N}_2(\ell)$ et $\Delta(\ell)$ au fil des itérations. Les histogrammes, à droite, sont produits à partir des 70 000 dernières itérations (première valeur : \bullet , valeur vraie : $-$). Par définition : $\bar{N}_1(\ell) = 1 - \bar{N}_2(\ell)$.

On observe sur les figures de gauche qu'une stabilité survient relativement vite (quelques centaines d'itérations). De plus, les échantillons sur la Fig. 5.9 montrent qu'au-delà d'environ 1 000 itérations, on retrouve visuellement les régions principales. La même observation est faite pour d'autres chaînes. Cette analyse empirique donne une idée du nombre d'itérations nécessaire pour avoir une configuration stable. Néanmoins, déterminer si la chaîne a convergé est une question difficile à laquelle nous essaierons d'apporter des éléments de réponse en fin de section.

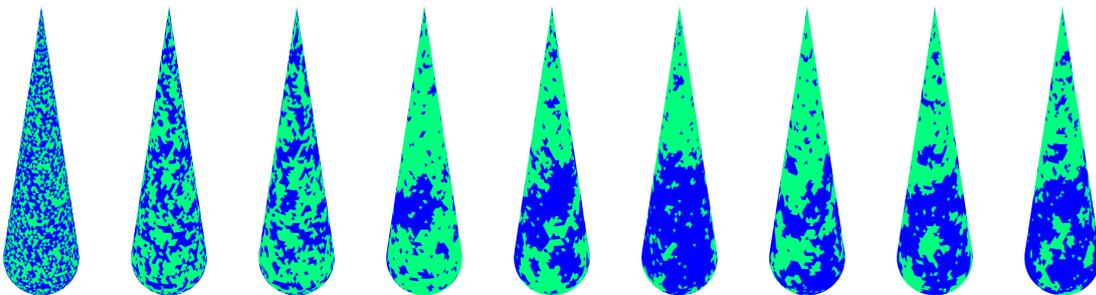


FIGURE 5.9 – Représentation des échantillons $\ell^{[i]}$ pour les itérations, de gauche à droite, $i = 1, 2, 3, 100, 1000, 10000, 45000, 100000, 140000$.

À partir des échantillons après le temps de chauffe, on calcule les distributions marginales *a posteriori* empiriques $\hat{\pi}(L_n = k|\sigma)$ et l'estimée $\hat{\ell}$, que l'on représente sur les Fig. 5.10d et 5.10b. De plus, on définit la carte des erreurs de classification comme $\delta(\ell^*, \hat{\ell}) = [\delta(\ell_n^*, \hat{\ell}_n) \dots \delta(\ell_N^*, \hat{\ell}_N)]^t$ et on la représente sur la Fig. 5.10c.

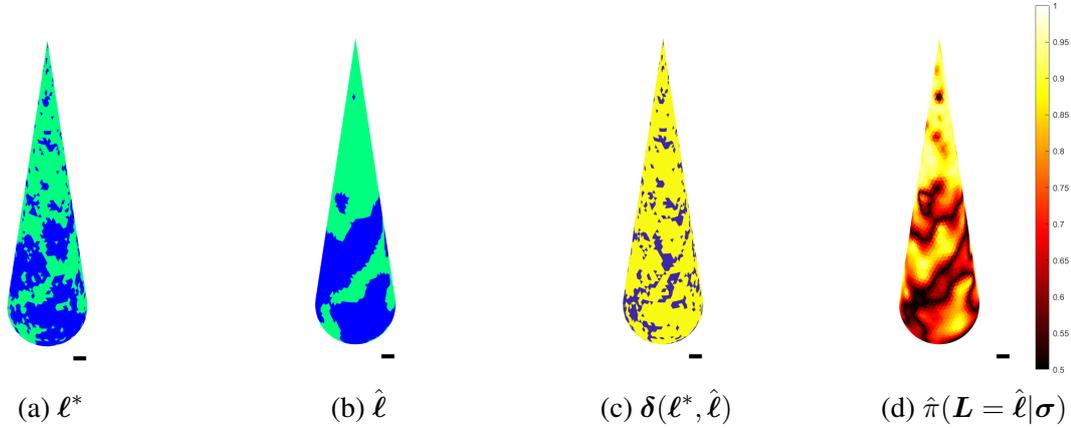


FIGURE 5.10 – (a) Étiquettes vraies ℓ^* . (b) Étiquettes estimées $\hat{\ell}$. (c) Carte des erreurs $\delta(\ell^*, \hat{\ell})$ (80% des mailles sont correctement classées). (d) Carte des probabilités $\hat{\pi}(L = \hat{\ell}|\sigma)$.

L'estimée $\hat{\ell}$ est de relativement bonne qualité, avec 80% des mailles correctement classifiées. De plus, on retrouve les régions principales (supérieure et inférieure). On constate que des erreurs sont naturellement commises aux frontières entre les régions. D'autres erreurs proviennent du fait que les régions de petite taille sont mal reconstruites, ce qui semble normal compte tenu de la difficulté du problème.

La carte des probabilités, Fig. 5.10d, est un résultat très important. Elle fournit pour chaque étiquette la probabilité de la classe sélectionnée et donne ainsi le degré de confiance que l'on peut accorder à la segmentation réalisée. Elle permet de déterminer les zones où la méthode commet probablement des erreurs et offre ainsi, sans connaître la vérité, une idée des erreurs de classification. De plus, on note qu'elle est de relativement bonne qualité puisqu'elle est cohérente de la carte des erreurs Fig. 5.10c. Utiliser conjointement l'estimée et les probabilités permet alors de reconstruire la carte des régions et de savoir si celle-ci est fiable ou non.

Remarque 33. Dans la Remarque 30 page 71, nous avons discuté du fait que pour pleinement quantifier les incertitudes, nous devons également nous intéresser à la probabilité qu'une étiquette prenne la valeur des autres classes. Néanmoins, les incertitudes sont totalement résumées par la Fig. 5.10d, puisque toutes les probabilités somment à 1 et $K = 2$.

Les résultats présentés sont similaires à ceux d'autres chaînes et l'on présente les résultats d'une autre chaîne dans la section F.2.1 de l'annexe F.

Éléments de réponse sur les convergences – Il y a en réalité deux questions de convergence : celle de la chaîne et celle des moyennes empiriques. Le contrôle visuel des échantillons, Fig. 5.9, et des indicateurs tels que $\Delta(\ell)$, Fig. 5.8, laisse penser que la chaîne a convergé après environ 1 000 itérations. Néanmoins, l'exploitation de chaînes plus courtes (30 000 itérations) mène à des résultats différents de ceux présentés, voir section F.2.2 de l'annexe F page 113. Cela peut signifier que le temps de chauffe n'est pas encore atteint ou que les moyennes empiriques n'ont

pas convergé. D'après notre expérience reposant sur l'exploitation de plusieurs chaînes, le temps de chauffe semble atteint au-delà de 100 000 itérations et 20 000 échantillons sont suffisants pour obtenir des estimées de bonne qualité. On précise que cette question mériterait d'être approfondie, par exemple en surveillant la convergence des $\hat{\pi}(L_n = k|\sigma)$.

Remarque 34. *L'exploration de la loi a posteriori par l'algorithme de Gibbs est un point important. La structure de l'échantillonnage pour les étiquettes (étiquette par étiquette) et leur dépendance rendent l'exploration lente. Par ailleurs, le passage entre deux modes, c'est-à-dire entre deux configurations d'étiquettes fortement probables, nécessite la visite de régions de faible probabilité. Cela se traduit par le fait que l'algorithme reste bloqué dans un mode.*

Enfin, sans information a priori en faveur d'une classe ou d'une configuration d'étiquettes privilégiée, nous devrions rencontrer le problème du « label switching » à cause de la symétrie de la loi a posteriori. Néanmoins, dans notre cas les paramètres α_k sont fixés, rendant la loi a posteriori asymétrique.

5.2.3 Impédance EAP

On présente dans cette section les reconstructions des impédances obtenues par le calcul de l'EAP. On rappelle que cet estimateur est défini comme la moyenne des espérances conditionnelles pondérée par les probabilités *a posteriori* des étiquettes, voir l'Eq. (3.47) page 53.

5.2.3.1 Impédances complètes

On s'intéresse aux réalisations des impédances présentées sur la Fig. 5.11. Après environ 1 000 itérations, on observe une stabilité des granularités et des échelles. Cette stabilité était attendue puisque les réalisations des paramètres d'échelles γ_1 , γ_2 et γ_ε sont concentrées.

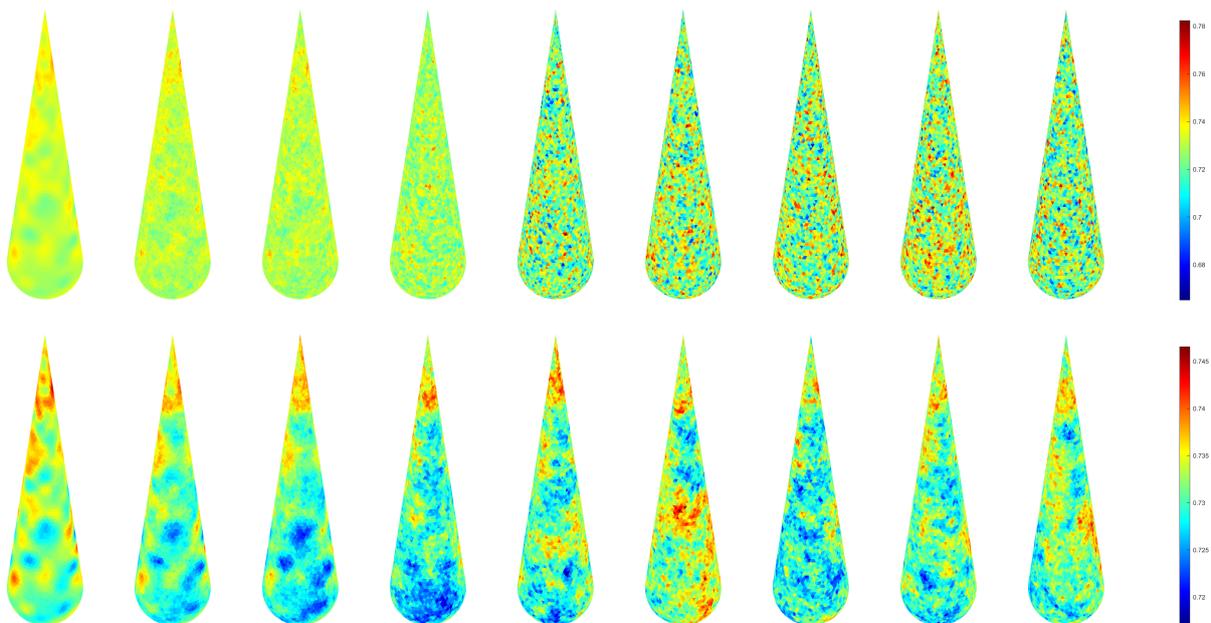


FIGURE 5.11 – Réalisation $|z_1^{[i]}|$ en haut et $|z_2^{[i]}|$ en bas pour les itérations, de gauche à droite, $i = 1, 2, 3, 100, 1000, 10000, 45000, 100000, 140000$.

On présente sur la Fig. 5.12, les estimées \hat{z}_1 et \hat{z}_2 , les erreurs de reconstruction et les incertitudes sous la forme des écarts-types des impédances marginales *a posteriori*, obtenus à partir de la covariance *a posteriori*.

Les estimées sont conformes aux informations *a priori*, c'est-à-dire proches de la référence et douces, et ressemblent localement aux impédances vraies. La partie supérieure de \hat{z}_2 , Fig. 5.12f, est correctement reconstruite, car la probabilité que les étiquettes prennent la valeur 2 est grande, contrairement à la partie inférieure. On constate l'inverse pour \hat{z}_1 , Fig. 5.12b. Par ailleurs, les frontières entre régions sont douces puisque les probabilités que les étiquettes prennent la valeur 1 ou 2 sont proches. Enfin, en raison de la difficulté du problème, les variations rapides ne sont pas reconstruites. Les estimées correspondent alors à des compromis entre les données et les douceurs introduites par les modèles *a priori*.

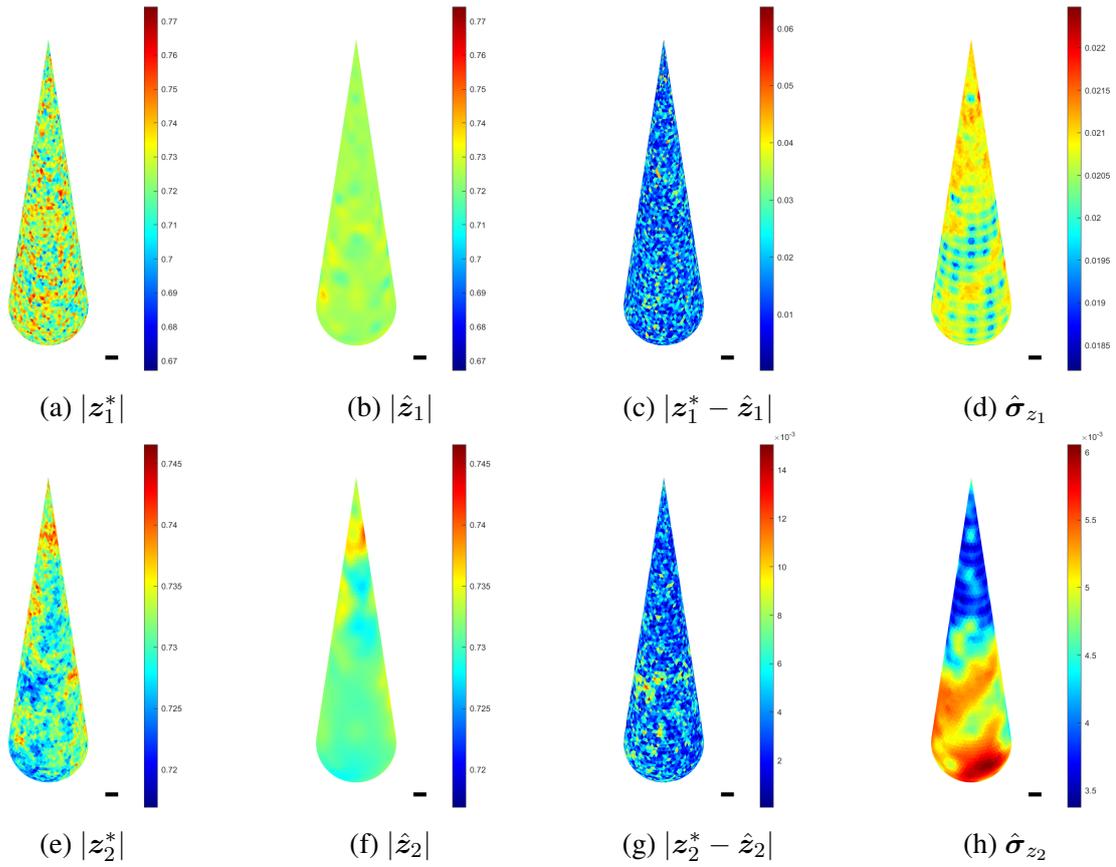


FIGURE 5.12 – (a,e) Impédances vraies. (b,f) Impédances estimées. (c,g) Cartes d'erreurs. (d,h) Écarts-types. Résultats obtenus à partir des 20 000 derniers échantillons. La valeur du module de la référence est $|z_0| = 0.7245$.

Les incertitudes $\hat{\sigma}_{z_1}$ et $\hat{\sigma}_{z_2}$, Fig. 5.12d et 5.12h, sont plus faibles dans les zones où l'impédance est correctement reconstruite, c'est-à-dire dans les zones où la probabilité que les étiquettes prennent la valeur de la classe considérée est forte (partie inférieure pour \hat{z}_1 et supérieure pour \hat{z}_2).

Par ailleurs, on constate que les incertitudes sont localement plus faibles aux endroits où les mesures ont été prises. Cette localité est en lien avec les coefficients de la matrice \mathbf{H} , analysés dans la section 1.4.5 du Chapitre 1 page 27, et la corrélation introduite par les modèles *a priori*. Cela est particulièrement visible sur la partie inférieure de $\hat{\sigma}_{z_1}$, avec les marques rondes

correspondant aux positions de l'antenne, mais moins visible sur $\hat{\sigma}_{z_2}$ puisque les paramètres $\hat{\gamma}_2$ et α_2 réduisent la variabilité de l'impédance, diminuant l'incertitude.

Les incertitudes fournies par la méthode renseignent sur la confiance que l'on peut accorder aux reconstructions. C'est un résultat crucial qui permet de savoir à quel point la méthode peut se tromper. Ici, la méthode prédit que la reconstruction \hat{z}_1 est moins fiable que celle de \hat{z}_2 , ce qui s'avère exact au vu des cartes des erreurs de reconstruction vraies Fig. 5.12c et 5.12g.

Des résultats similaires pour une autre chaîne sont disponibles dans la partie F.3.1 de l'annexe F. On affiche dans la section F.3.2 de l'annexe F, les résultats de chaînes plus courtes (30 000). Les observations que l'on peut faire sont identiques à celles faites pour les étiquettes à la page 81, à savoir que les estimées sont différentes d'une chaîne à l'autre.

5.2.3.2 Impédance par morceaux

Finalement, on reconstruit l'impédance \hat{z} à partir des estimées \hat{z}_1 , \hat{z}_2 et $\hat{\ell}$, et on la représente sur la Fig. 5.13b. On retrouve les morceaux provenant de \hat{z}_1 (partie inférieure) et de \hat{z}_2 (partie supérieure). L'estimée \hat{z} est alors homogène par morceaux et les propriétés sont respectées. Comme mentionné précédemment, on constate sur la Fig. 5.13c, que les erreurs correspondantes à \hat{z}_1 sont plus élevées que celles correspondantes à \hat{z}_2 . Par ailleurs, pour quantifier l'erreur de manière plus globale, on calcule l'écart relatif entre l'estimée et la carte vraie

$$\text{Err}(z^*, \hat{z}) = \frac{\sum_{n=1}^N |z_n^* - \hat{z}_n|}{\sum_{n=1}^N |z_n^*|} = 1.45\%. \quad (5.3)$$

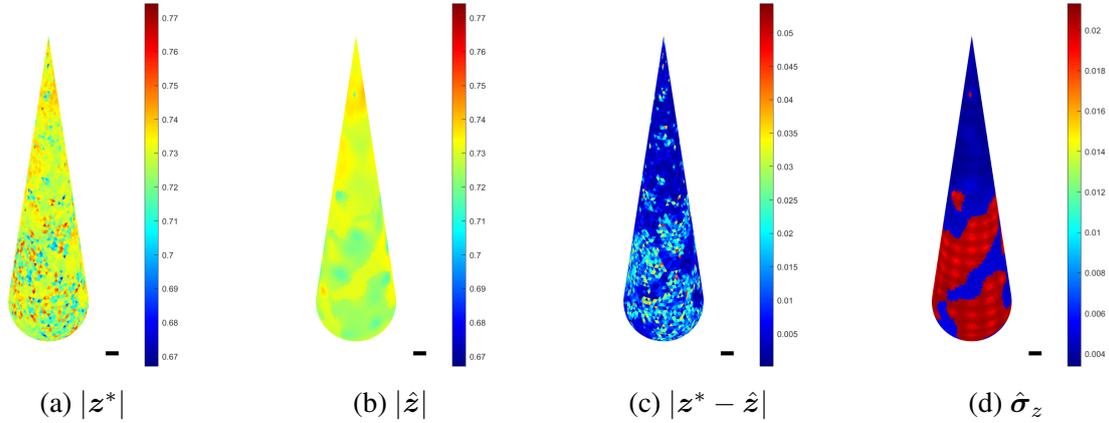


FIGURE 5.13 – (a) Impédance vraie. (b) Impédance estimée. (c) Carte des erreurs. (d) Écarts-types. L'impédance \hat{z} et les écarts-types $\hat{\sigma}_z$ sont reconstruits par $\hat{z} = \sum_{k=1}^K \mathbf{P}_k(\hat{\ell}) \hat{z}_k$ et $\hat{\sigma}_z = \sum_{k=1}^K \mathbf{P}_k(\hat{\ell}) \hat{\sigma}_{z_k}$.

On reconstruit également la carte des incertitudes sous la forme d'écarts-types, Fig. 5.13d, qui quantifie la confiance que l'on peut accorder à la reconstruction. En comparant les Fig. 5.13c et 5.13d, on confirme que la méthode prédit correctement les zones où des erreurs sont commises, ce qui est un résultat crucial en pratique.

Afin de rendre les incertitudes exploitables plus facilement, on fournit des intervalles de crédibilité. L'impédance étant un vecteur complexe, on donne pour chaque maille z_n un intervalle de crédibilité à queues égales pour la partie réelle $\mathcal{D}_{R,n} = [z_r^m, z_r^M]$ et pour la partie

imaginaire $\mathcal{D}_{I,n} = [z_i^m, z_i^M]$ tels que

$$\begin{aligned}\Pr(\operatorname{Re}(z_n) < z_r^m | \sigma) &= \Pr(\operatorname{Re}(z_n) > z_r^M | \sigma) = \lambda/2, \\ \Pr(\operatorname{Im}(z_n) < z_i^m | \sigma) &= \Pr(\operatorname{Im}(z_n) > z_i^M | \sigma) = \lambda/2,\end{aligned}$$

avec $\lambda = 0.05$. Les intervalles obtenus sont d'environ $\operatorname{Re}(\hat{z}) \pm \sqrt{2} \hat{\sigma}_z$ pour la partie réelle et $\operatorname{Im}(\hat{z}) \pm \sqrt{2} \hat{\sigma}_z$ pour la partie imaginaire. On les représente, Fig. 5.14, pour une coupe verticale de l'objet, le long de la génératrice d'angle 135° (correspondant à la génératrice la plus à gauche sur les figures). On affiche également l'impédance vraie et estimée.

On retrouve sur la gauche la partie inférieure (provenant de \hat{z}_1) et sur la droite la partie supérieure (provenant de \hat{z}_2). On retrouve alors également l'incertitude plus forte pour la classe 1 que pour la classe 2. Le résultat clé ici est que l'impédance vraie est en grande partie contenue dans les intervalles de crédibilité. La méthode fournit alors non seulement une estimée mais également un intervalle contenant avec une forte probabilité l'impédance vraie, ce qui en fait un outil pertinent pour le contrôle.

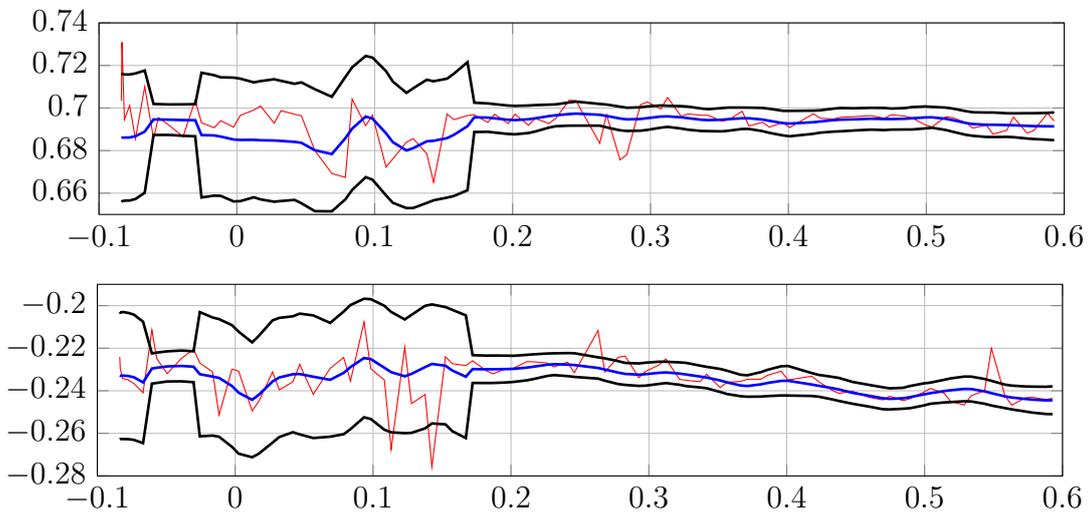


FIGURE 5.14 – Coupe verticale : impédance vraie — , impédance estimée — et intervalle de crédibilité — . Partie réelle en haut et partie imaginaire en bas. L'axe des abscisses représente l'altitude.

5.2.4 Impédances EAP conditionnelle

On présente ici les reconstructions des impédances obtenues par le calcul de l'EAP conditionnellement au MMAP pour les étiquettes. On rappelle que les échantillons nécessaires à ces reconstructions sont obtenus par l'approche (B) présentée page 72. On précise également que les estimées sont calculées à partir de 20 000 réalisations de la loi conditionnelle *a posteriori* et que les paramètres d'échelles γ_ε , γ_1 et γ_2 sont fixés à leurs valeurs estimées précédemment.

Pour ne pas surcharger ce chapitre en résultats, on se contente de représenter l'impédance par morceaux $\hat{z}(\hat{\ell})$, Fig. 5.15b, construite à partir de $\hat{\ell}$ et des estimées conditionnelles $\hat{z}_1(\hat{\ell})$ et $\hat{z}_2(\hat{\ell})$. On représente également les incertitudes sous forme d'écart-type, Fig. 5.15d, et la carte d'erreurs de reconstruction, Fig. 5.15c.

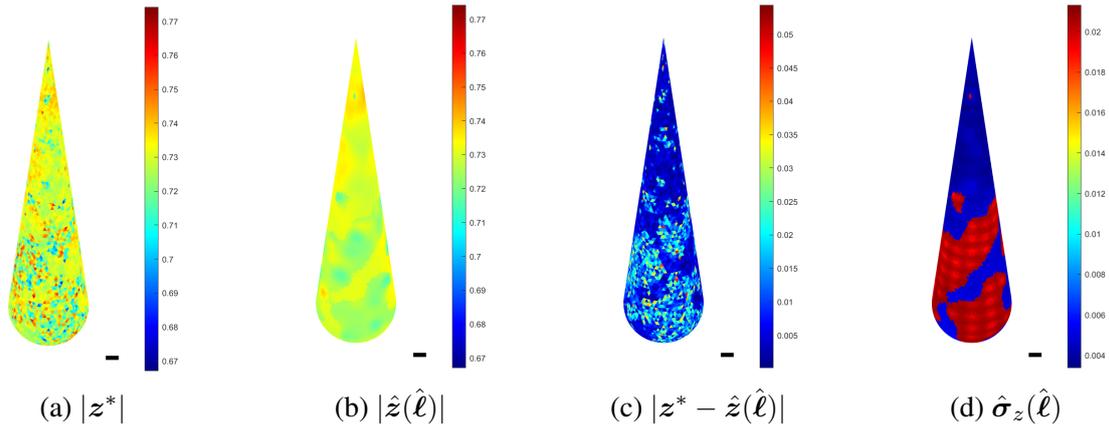


FIGURE 5.15 – (a) Impédance vraie. (b) Impédance estimée. (c) Carte des erreurs. (d) Écart-types. L'impédance $\hat{z}(\hat{\ell})$ et les écart-types $\hat{\sigma}_z(\hat{\ell})$ sont reconstruits par $\hat{z}(\hat{\ell}) = \sum_{k=1}^K P_k(\hat{\ell}) \hat{z}_k(\hat{\ell})$ et $\hat{\sigma}_z(\hat{\ell}) = \sum_{k=1}^K P_k(\hat{\ell}) \hat{\sigma}_{z_k}(\hat{\ell})$.

En comparant les Fig. 5.13b et 5.15b, on constate que les reconstructions de l'impédance par morceaux EAP et EAP conditionnelle sont quasi-similaires. Ce constat est le même pour les incertitudes, Fig. 5.13d et 5.15d, ainsi que pour les intervalles de crédibilité (pour la partie réelle et pour la partie imaginaire). Malgré ces ressemblances, la reconstruction $\hat{z}(\hat{\ell})$ est de qualité légèrement supérieure à la reconstruction \hat{z} au sens de l'écart relatif

$$\text{Err}(z^*, \hat{z}(\hat{\ell})) = \frac{\sum_{n=1}^N |z_n^* - \hat{z}_n(\hat{\ell}_n)|}{\sum_{n=1}^N |z_n^*|} = 1.44\%. \quad (5.4)$$

Cela s'explique par le fait que $\hat{z}(\hat{\ell})$ est plus fidèle à l'impédance vraie que \hat{z} dans les zones où l'incertitude de classification est forte (probabilité *a posteriori* des étiquettes faibles) et possède naturellement des discontinuités plus franches.

Pour l'illustrer, on propose de visualiser sur la Fig. 5.16 une coupe verticale, le long de la génératrice d'angle 135° , des reconstructions EAP et EAP conditionnelle que l'on représente superposées. On trace également la probabilité *a posteriori* pour les étiquettes afin d'identifier facilement les zones de forte et de faible incertitude de classification. Comme on pouvait s'y attendre, les deux reconstructions diffèrent lorsque la probabilité de la classe sélectionnée devient faible. Ce phénomène est particulièrement visible à l'altitude 0.4 et autour des discontinuités. Ces différences viennent du mélange réalisé par l'estimateur EAP. En effet, lorsque la probabilité des étiquettes diminue, la pondération du mélange favorise une reconstruction proche de l'*a priori* en réduisant l'influence des données. À l'inverse, l'EAP conditionnelle ne fait pas de mélange et conserve un compromis entre les données et l'*a priori* homogène à l'intérieur des régions, quelle que soit la probabilité des étiquettes.

Finalement, l'estimée $\hat{z}(\hat{\ell})$ et les incertitudes $\hat{\sigma}_z(\hat{\ell})$ fournissent également une reconstruction homogène par morceaux respectant les propriétés et une quantification de la confiance que l'on peut lui accorder. Ces éléments font de l'EAP conditionnelle un estimateur pertinent pour le contrôle puisqu'il offre les mêmes résultats clé que l'EAP, avec une qualité de reconstruction supérieure plus proche de l'impédance vraie et des discontinuités plus franches.

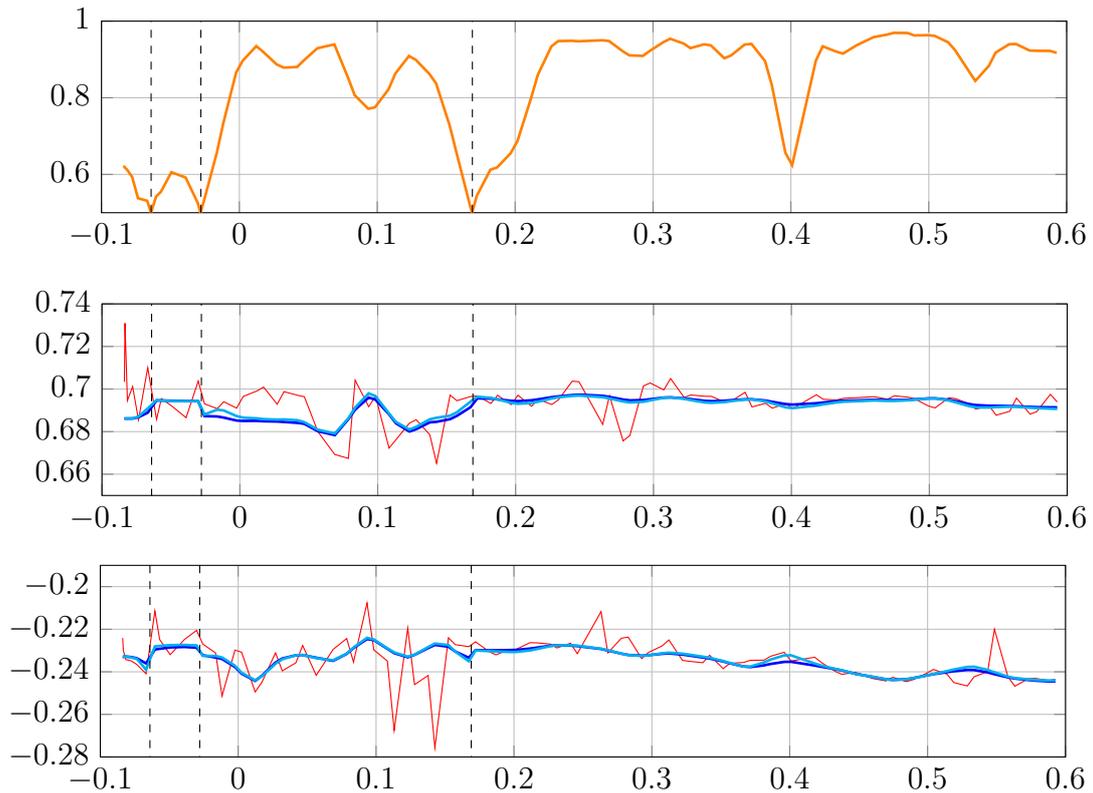


FIGURE 5.16 – Coupe verticale de la probabilité de la classe sélectionnée en haut (comprise entre 0.5 et 1), de la partie réelle au milieu et de la partie imaginaire en bas. L'axe des abscisses représente l'altitude. Impédance vraie —, estimée EAP —, estimée EAP conditionnelle —. Les barres verticales - - - représentent les discontinuités (changements de classe).

5.3 Conclusion

Malgré la difficulté du problème, faible nombre de données vis-à-vis du nombre d'inconnues et dégradation induites par le modèle d'observation, la méthode proposée fournit une reconstruction de l'impédance par morceaux ainsi que la quantification des incertitudes associées. D'un côté, la carte des étiquettes apporte des informations sur la localisation et la forme des régions et de l'autre, les impédances renseignent sur l'homogénéité à l'intérieur des régions. On note cependant que les régions de petite taille et les variations rapides de l'impédance sont mal reconstruites. Les incertitudes (probabilités, écarts-types et intervalles de crédibilité) apportent des informations quant à la confiance que l'on peut accorder aux reconstructions. Il s'agit d'un résultat primordial pour le contrôle puisque, sans connaître la vérité, la méthode est capable de détecter les zones où des erreurs sont probablement commises et de les quantifier. De plus, nous avons constaté la fiabilité de cette capacité grâce à la cohérence entre les erreurs vraies et les incertitudes estimées. L'utilisation conjointe des reconstructions et des incertitudes fournit alors un outil d'aide à la décision pertinent pour identifier si un objet est conforme aux spécifications.

5.4 Application à des données plus réelles

Comme mentionné dans le Chapitre 1, on dispose d'un code d'électromagnétisme (EM) capable, à impédance z fixée, de calculer les coefficients σ correspondants. Dans l'idée d'anticiper les résultats que l'on pourrait obtenir sur des données issues du code EM, on propose de comparer les données obtenues par le code et celles obtenues par le modèle linéarisé pour le patchwork z^* présenté dans ce chapitre. Dans la suite, on adopte les notations suivantes

$$\sigma_{\text{EM}} = \mathbf{h}(z^*), \quad (5.5)$$

où l'on rappelle que \mathbf{h} est la fonction non linéaire en z , et

$$\sigma_{\text{LIN}} = \sigma_0 + \mathbf{H}(z^* - z_0), \quad (5.6)$$

correspondant aux données précédemment présentées, mais sans le terme de bruit.

Pour les comparer, on représente les données sous forme d'hologramme sur la Fig. 5.17. Comme on le constate, la différence entre σ_{EM} et σ_{LIN} est relativement faible, de l'ordre de l'unité. Comparativement, les valeurs du bruit sont de l'ordre de la dizaine. Pour cette raison, on ne s'attend pas à des différences majeures entre les résultats présentés précédemment et ceux que l'on obtiendrait à partir de σ_{EM} .

Remarque 35. *On peut faire remarquer que le patchwork utilisé est relativement proche de la référence z_0 . Il s'agit d'une configuration favorable dans laquelle l'erreur de linéarisation est faible. Pour une impédance autre, qui serait plus éloignée de la référence, des erreurs de linéarisation plus importantes pourraient être attendues.*

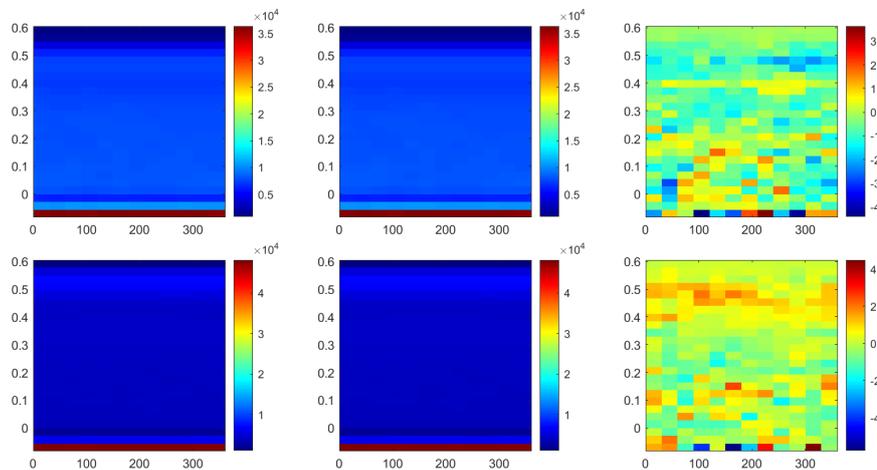


FIGURE 5.17 – Données issues du code EM (sans linéarisation) et du modèle linéaire non bruité. De gauche à droite σ_{EM} , σ_{LIN} et $\sigma_{\text{EM}} - \sigma_{\text{LIN}}$. Partie réelle en haut et partie imaginaire en bas.

CHAPITRE 6

CONCLUSION

6.1 Synthèse

Dans ce travail, nous nous sommes intéressés à un problème de contrôle non destructif (CND) en électromagnétisme (EM). Il s'agit de caractériser l'homogénéité des propriétés matériaux afin de garantir la conformité d'un objet. Pour cela, on s'appuie sur des coefficients de réflexion mesurés en champ proche par un moyen expérimental. Nous avons considéré qu'il existe des régions de l'objet dans lesquelles les propriétés sont homogènes, mais possiblement différentes entre régions. De plus, les propriétés matériaux sont approximées par une impédance surfacique. L'impédance peut être reliée aux coefficients de réflexion par un modèle d'observation linéaire à bruit additif. Le contrôle se ramène alors à l'inversion d'un modèle linéaire pour reconstruire une impédance homogène par morceaux à partir des coefficients de réflexion. Il s'agit d'un problème d'inversion-segmentation. Une des difficultés majeures tient au fait que le nombre de données est très inférieur au nombre d'inconnues.

Une première méthode de reconstruction, reposant sur la minimisation d'un critère pénalisé par un potentiel de Huber [2], a été développée en début de thèse. Celle-ci reconstruit une impédance douce par morceaux avec des discontinuités relativement bien respectées, mais ne réalise pas de segmentation. Afin de répondre au problème joint de l'inversion et de la segmentation, nous avons développé une approche bayésienne.

La modélisation de l'impédance par morceaux repose sur la construction d'un « patchwork » à partir de plusieurs impédances complètes, c'est-à-dire sur toute la surface de l'objet, et d'un jeu d'étiquettes décrivant les régions à la surface. Les impédances sont modélisées *a priori* par des champs aléatoires de Markov gaussiens, correspondant à diverses homogénéités de propriétés matériaux dont les paramètres descriptifs de rugosités sont connus tandis que les paramètres d'échelle de variation sont probabilisés. De plus, le choix de modèles markoviens mène à considérer des matrices creuses, ce qui s'avère bénéfique par la suite. Le jeu d'étiquettes est modélisé *a priori* par un champ de Potts, permettant d'obtenir des régions compactes dont la taille moyenne est pilotée par un paramètre de granularité connu. Introduire des étiquettes augmente le nombre d'inconnues, mais permet de considérer séparément l'aspect « homogène » et l'aspect « par morceaux », ce qui simplifie la modélisation. Concernant le modèle de vraisemblance, il s'appuie sur le modèle d'observation linéaire avec un bruit gaussien de structure connue et de puissance probabilisée. Il en résulte une modélisation hiérarchique que l'on a cherché à exploiter.

Les estimateurs bayésiens (EAP et MMAP) menant aux reconstructions (impédances et étiquettes), ainsi que l'estimation des incertitudes (écarts-types, intervalles de crédibilité et probabilités marginales), nécessitent des calculs d'intégration et de sommation en grande dimension de la loi *a posteriori* que l'on ne peut mener analytiquement. Pour cette raison, nous avons conçu un algorithme de Gibbs efficace et adapté au problème. Il s'appuie sur la décomposition de la loi *a posteriori* en lois de forme standard, en particulier pour les impédances gaussiennes.

L'échantillonnage des impédances gaussiennes présente néanmoins des difficultés liées au fait qu'elles sont en grandes dimensions. Pour cela, nous proposons une stratégie d'échantillonnage, à la base du filtre de Kalman d'Ensemble, qui tire parti du caractère creux des matrices introduites par les modèles *a priori* pour les impédances et du fait que certains calculs sont réalisés dans l'espace des données, qui est de petite dimension. Elle permet ainsi de passer d'un coût calculatoire de l'ordre de $\mathcal{O}(N^3)$, pour des méthodes d'échantillonnage de gaussiennes classiques reposant sur la décomposition de Cholesky, à un coût de l'ordre de $\mathcal{O}(MN^2/2)$, où M (nombre de données) est bien plus petit que N (la dimension des impédances). Cela en fait une stratégie particulièrement adaptée à notre problème. Concernant les étiquettes, il n'est pas possible de réaliser un échantillonnage efficace en damier, courant en traitement d'image. L'échantillonnage est donc réalisé par un balayage systématique optimisé grâce à la mise à jour d'un résidu intermédiaire homogène aux données, réduisant la complexité calculatoire.

L'algorithme de Gibbs nécessite le réglage d'un temps de chauffe délicat à estimer. Pour ce travail, il est réglé de manière empirique en s'appuyant sur l'expérience acquise grâce à l'exploitation de plusieurs chaînes. En complément, nous avons effectué un travail plus théorique sur l'étude non asymptotique de la convergence de l'algorithme de Gibbs en fonction du temps de chauffe dans un cas simple [3].

La méthode proposée fournit une reconstruction de l'impédance par morceaux, ainsi qu'une quantification des incertitudes associées. Elle permet ainsi d'obtenir la carte des régions homogènes et des informations sur l'homogénéité des régions. L'estimation des incertitudes quantifie la confiance que l'on peut accorder à la reconstruction et permet de détecter les zones où celle-ci est probablement erronée. Il s'agit d'un résultat clé. La méthode proposée est alors un outil d'aide à la décision pertinent qui répond au problème du contrôle. Elle nécessitera néanmoins un travail préalable de caractérisation des erreurs de mesure.

6.2 Contributions

Analyse non asymptotique de l'algorithme de Gibbs – Une de nos contributions est l'analyse non asymptotique de l'algorithme de Gibbs publié dans *IEEE Signal Processing Letters* [3]. Cette analyse s'appuie sur un cas d'école gaussien 2D et explicite l'évolution des lois marginales au cours des itérations, permettant une compréhension non empirique du comportement de l'échantillonneur de Gibbs. Ce travail aborde également la question du temps de chauffe et du compromis sur les échantillons à sélectionner pour les estimateurs.

Modélisation hiérarchique adaptée – Une originalité de notre travail vient du fait que l'on considère un modèle *a priori* gaussien corrélé pour l'impédance par morceaux, à l'inverse de beaucoup de travaux en inversion-segmentation. Ce choix mène à un coût calculatoire élevé lors de l'échantillonnage des étiquettes *a posteriori*, à cause du calcul répété de la fonction de partition du champ gaussien. Pour cette raison, nous nous sommes inspirés de la modélisation

originale présentée dans [53], qui propose une structure hiérarchique différente permettant de contourner le problème lié à la fonction de partition. À notre connaissance, nous sommes les seuls avec [53] à exploiter cette hiérarchie particulière. Notre travail se différencie par la manière dont la corrélation est encodée. Dans [53], la corrélation est encodée au travers de la DSP et repose sur le fait que le maillage est une grille régulière, alors que nous avons choisi un modèle de corrélation fondé sur les interactions entre mailles voisines adaptées à notre maillage non régulier, pilotable en longueur et en amplitude.

Simulation efficace des gaussiennes conditionnelles – L'échantillonnage des gaussiennes à l'intérieur de l'algorithme de Gibbs est un point clé puisqu'il représente l'opération la plus coûteuse et qu'il est répété à chaque itération et pour les K classes. Pour cela, nous proposons un algorithme d'échantillonnage particulièrement efficace lorsque le nombre de données disponible est faible. Il permet de passer d'une complexité calculatoire de l'ordre de $\mathcal{O}(N^3)$, pour la méthode de Cholesky classique, à un coût de l'ordre de $\mathcal{O}(MN^2/2)$, où M est le nombre de données et N la dimension des inconnues. Cet algorithme a été inspiré par [90] et n'a, à notre connaissance, jamais été utilisé pour l'inversion-segmentation.

Application à un problème difficile de CND avec quantification des incertitudes – Enfin, on peut mettre en avant le fait que malgré un nombre de données très faible devant le nombre d'inconnues (300 données pour 7400 inconnues) et une dégradation de l'information par le système d'observation, la méthode fournit des résultats de qualité. D'un côté la méthode segmente en régions homogènes et de l'autre, elle reconstruit correctement les propriétés de chaque classe. De plus, la quantification des incertitudes permet d'accorder un degré de crédibilité aux reconstructions en dressant une carte des erreurs attendues, ce qui est un élément crucial dans ce contexte de contrôle.

6.3 Perspectives

De nombreuses perspectives se dégagent du travail présenté dans ce manuscrit. Sans les hiérarchiser, nous en présentons quelques-unes qui nous semblent intéressantes.

Analyse de performance et validation de la méthode – Un travail visant à quantifier les performances en fonction des données nous semble important. Il pourrait concerner l'étude des performances en fonction du niveau de bruit et la robustesse de la méthode aux données aberrantes. Une validation de la méthode devrait également être faite sur des données provenant d'autres sources, par exemple du code Maxwell 3D.

La validation à partir de mesures expérimentales pourrait également être faite, mais cela exigerait un travail préalable de métrologie afin de caractériser les incertitudes de mesure.

Initialisation de l'algorithme de Gibbs – L'initialisation des impédances s'appuie sur les données, à l'inverse de celle des étiquettes. Cette dernière est alors possiblement éloignée de la solution recherchée et pourrait être améliorée grâce à une étape de présegmentation dédiée, par exemple en s'inspirant de [58]. L'initialisation des étiquettes serait alors probablement de

meilleure qualité, ce qui limiterait le nombre d'itérations de l'algorithme de Gibbs et donc le temps de calcul.

Échantillonnage des étiquettes – La loi *a posteriori* pour les étiquettes est non séparable et non markovienne, ce qui rend l'utilisation de stratégies telles que l'échantillonnage en damier impossible. L'échantillonnage des étiquettes est alors un point critique puisque celui-ci est réalisé étiquette par étiquette, ce qui rend l'exploration lente. Une étude visant à comparer différents balayages pourrait déboucher sur des améliorations de performances. L'adaptation de l'algorithme de Swendsen–Wang [93, 94] est également envisageable.

Échantillonnage des impédances – Les performances de la stratégie d'échantillonnage des impédances pourraient être comparées avec d'autres méthodes de simulation de gaussiennes en grande dimension, par exemple l'algorithme RJPO (Reversible Jump Perturbation Optimization [89]) ou Clone MCMC [95]. Par ailleurs, la mise à jour lors des itérations des décompositions de Cholesky (des matrices de covariance *a posteriori*) permettrait d'accélérer l'échantillonnage.

Paramètres supplémentaires – L'estimation des paramètres d'interaction α et du coefficient de granularité β permettrait à la méthode de reconstruire des impédances de rugosités inconnues et d'ajuster automatiquement la taille des régions. Ces estimations sont rendues difficiles à cause des fonctions de partitions. Néanmoins, ces fonctions peuvent être précalculées en amont pour une fine grille de valeur et utilisées par la suite, comme cela est présenté dans [53, 70, 96]. Une autre piste intéressante est également présentée par [97] qui propose une méthode numérique rapide pour estimer les fonctions de partition des champs de Potts symétriques.

Comme mentionné dans la Remarque 34, l'estimation des paramètres d'interaction α conjointement aux autres inconnues mène à une loi *a posteriori* symétrique et au problème de « label switching ». Pour surmonter cette difficulté, des stratégies existent, par exemple l'ajout de contraintes d'identifiabilité, par exemple $\alpha_1 < \alpha_2$, ou l'utilisation d'un algorithme de ré-étiquetage [98].

Estimer le nombre de classes K renseignerait sur le nombre de matériaux et éviterait une surreprésentation d'un même matériau ou que plusieurs matériaux soient confondus au sein d'une même classe. Cette estimation pourrait s'appuyer sur des méthodes de sélection de modèles telles que la méthode de Chib ou l'algorithme RJMCMC (Monte Carlo par chaînes de Markov à sauts réversibles).

Modèles *a priori* – On pourrait tenir compte d'une condition de positivité sur la partie réelle de l'impédance. De même, on pourrait introduire une éventuelle corrélation entre la partie réelle et la partie imaginaire.

La corrélation spatiale introduite par le modèle *a priori* des impédances dépend, entre autres, de la structure de voisinage. L'utilisation d'une structure plus étendue, par exemple d'ordre 2, permettrait de modéliser des corrélations différentes.

Enfin, la méthode permet de considérer des moyennes différentes pour chacune des classes. Ces dernières pourraient alors être probabilisées, introduites dans le modèle hiérarchique et estimées conjointement au reste.

ANNEXE A

DE MAXWELL AU SYSTÈME LINÉAIRE

Dans ce chapitre, on explique comment on passe des équations de Maxwell au système linéaire $A(z)X = B$ et à la dérivée $\partial A/\partial z$, à la base du modèle direct $\sigma = Hz$. Ces développements en *électromagnétisme numérique*, extérieurs à la thèse, sont issus des travaux antérieurs [20]. Ils consistent en la modélisation du problème de diffraction sur un obstacle à l'aide d'équations intégrales et de leur formulation variationnelle.

A.1 Équations de Maxwell

Diffraction sur un obstacle – On considère sur la Fig. A.1 un obstacle Ω^- ouvert borné de \mathbb{R}^3 , de frontière régulière Γ , constitué d'un noyau conducteur recouvert d'un empilement de matériau modélisable par une impédance z . L'obstacle est éclairé à la fréquence f par une onde électromagnétique incidente avec un champ incident, électrique et magnétique, (E^{inc}, H^{inc}) solution des équations de Maxwell dans le vide¹.

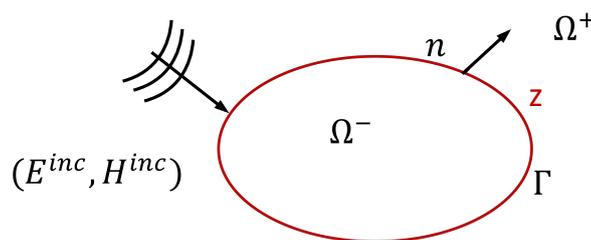


FIGURE A.1 – Schéma général de la diffraction sur un obstacle

Équations de Maxwell en régime harmonique – Le champ électromagnétique total² (E, H) induit par le champ incident (E^{inc}, H^{inc}) en présence de l'obstacle est solution des équations de

1. La variable d'espace x est omise dans la suite, notamment dans l'expression des phaseurs complexes des champs $E(x)$ et $H(x)$.

2. Par souci de lisibilité, on ne fait pas paraître la dépendance temporelle implicite en $e^{+i2\pi ft}$.

Maxwell harmonique dans Ω^+ avec la condition de radiation de Silver-Müller à l'infini (condition d'onde sortante) et la condition de Leontovich [99] à la surface de l'obstacle :

$$\left\{ \begin{array}{ll} \text{rot } E + ik\eta H = 0 & \text{dans } \Omega^+ \\ \text{rot } H - ik\eta^{-1} E = 0 & \text{dans } \Omega^+ \\ \lim_{|x| \rightarrow \infty} |x| \left(\eta (H - H^{inc}) \times \frac{x}{|x|} - (E - E^{inc}) \right) = 0 & \\ n \times (E \times n) - \eta z (n \times H) = 0 & \text{sur } \Gamma \end{array} \right. \quad (\text{A.1})$$

avec $k = 2\pi f/c$ le nombre d'onde, c la vitesse de la lumière, η l'impédance du vide et $z(x)$ l'impédance de surface complexe de l'obstacle à partie réelle non négative. La variation d'impédance $z(x)$ permet de prendre en compte la présence de différents matériaux ainsi qu'une variation d'épaisseur du ou des matériaux à la surface Γ de l'obstacle. La condition sur le signe de la partie réelle de l'impédance est nécessaire pour assurer l'existence et l'unicité de la solution.

Remarque 36. *La condition de Leontovich [99] posée à la surface extérieure de l'objet permet de s'affranchir de la résolution des équations de Maxwell dans l'empilement de matériaux qui le recouvre. Sa validité repose sur plusieurs hypothèses, la principale portant sur les propriétés des matériaux.*

A.2 Équations intégrales

Formules de Stratton-Chu – On introduit les densités de courants électrique et magnétique à la surface de l'obstacle définies comme la trace tangentielle (tournée de $\pm\pi/2$) du champ électrique et magnétique sur Γ :

$$j(x) = n \times H|_{\Gamma}(x), m(x) = E|_{\Gamma}(x) \times n \quad (\text{A.2})$$

Tout champ électromagnétique (E, H) dans Ω^+ est somme d'un champ incident (E^{inc}, H^{inc}) et du champ diffracté :

$$E^d = (E - E^{inc}), H^d = (H - H^{inc}) \quad (\text{A.3})$$

Il vérifie la condition de radiation de Silver-Müller sortante admettant la représentation intégrale suivante connue sous le nom de formules de Stratton-Chu [100] :

$$\left\{ \begin{array}{ll} E(x) = E^{inc}(x) - ik\eta Tj(x) + Km(x) & x \in \Omega^+ \\ H(x) = H^{inc}(x) - i\eta^{-1}kTm(x) - Kj(x) & x \in \Omega^+ \end{array} \right. \quad (\text{A.4})$$

T et K sont les opérateurs intégraux définis par :

$$\left\{ \begin{array}{l} Tj(x) = \int_{\Gamma} G(x, y) j(y) \, d\Gamma(y) + \frac{1}{k^2} \nabla \int_{\Gamma} G(x, y) \text{div}_{\Gamma} j(y) \, d\Gamma(y) \quad x \notin \Gamma \\ Kj(x) = \int_{\Gamma} \nabla_y G(x, y) \times j(y) \, d\Gamma(y) \quad x \notin \Gamma \end{array} \right. \quad (\text{A.5})$$

où le noyau de Green

$$G(x, y) = \frac{\exp(-ik|x-y|)}{4\pi|x-y|} \quad (\text{A.6})$$

est la solution sortante de l'équation de Helmholtz dans \mathbb{R}^3 et $\text{div}_{\Gamma} j$ est la divergence surfacique du courant.

Traces des opérateurs intégraux – Les traces surfaciques des opérateurs intégraux T et K sont connues [100]. La trace tangentielle de $Tj(x)$ est continue à la traversée de Γ , celle de $Kj(x)$ peut être discontinue. Elles sont données en tout point x par :

$$\left\{ \begin{array}{l} (Tj)_{tg}^{\pm}(x) = Tj(x), \\ Tj(x) = \left(\int_{\Gamma} G(x,y)j(y) d\Gamma(y) \right)_{tg} + \frac{1}{k^2} \nabla_{\Gamma} \int_{\Gamma} G(x,y) \operatorname{div}_{\Gamma} j(y) d\Gamma(y) \\ (Kj)_{tg}^{\pm}(x) = \pm \frac{1}{2} n \times j + Kj(x), \\ Kj(x) = \left(\int_{\Gamma} \nabla_y G(x,y) \times j(y) d\Gamma(y) \right)_{tg} \end{array} \right. \quad (\text{A.7})$$

Pour $x \in \Gamma$, l'expression (A.4) devient :

$$\left\{ \begin{array}{l} E_{tg}(x) = E_{tg}^{inc}(x) - ik\eta Tj(x) + \frac{1}{2} n \times m(x) + Km(x) \\ H_{tg}(x) = H_{tg}^{inc}(x) - ik\eta^{-1} Tm(x) - \frac{1}{2} n \times j(x) - Kj(x) \end{array} \right. \quad (\text{A.8})$$

Équations intégrales – En utilisant (A.2), on obtient deux équations intégrales sur les densités de courants électrique et magnétique :

$$\left\{ \begin{array}{l} E_{tg}^{inc}(x) = ik\eta Tj(x) - Km(x) + \frac{1}{2} n \times m(x) \\ -H_{tg}^{inc}(x) = -ik\eta^{-1} Tm(x) - Kj(x) + \frac{1}{2} n \times j(x) \end{array} \right. \quad (\text{A.9})$$

Remarque 37. La première équation de (A.9) est appelée dans la littérature EFIE pour Electric Field Integral Equation, la seconde est appelée MFIE pour Magnetic Field Integral Equation. Ces deux équations ne tiennent pas compte de la condition aux limites sur la frontière Γ de l'obstacle.

Pour un obstacle dont le revêtement peut être approximé par une condition d'impédance, la condition que nous considérons dans cette étude est la condition d'impédance de Leontovich [99] :

$$n \times (E \times n) - \eta z (n \times H) = 0 \quad \text{sur } \Gamma \quad (\text{A.10})$$

qui se réexprime sur les densités de courants comme :

$$n \times m = \eta z j \quad (\text{A.11})$$

ou bien

$$n \times j = -(\eta z)^{-1} m \quad \text{sur } \Gamma. \quad (\text{A.12})$$

En injectant la condition aux limites (A.11) dans l'EFIE et (A.12) dans la MFIE, on obtient un système de deux équations intégrales sur les densités de courants électrique et magnétique qui tient compte de la condition aux limites sur Γ :

$$\left\{ \begin{array}{l} E_{tg}^{inc}(x) = ikT\eta j(x) - Km(x) + \frac{1}{2} z \eta j(x) \\ -\eta H_{tg}^{inc}(x) = -ikTm - K\eta j(x) - \frac{1}{2} z^{-1} m(x) \end{array} \right. \quad (\text{A.13})$$

A.3 Formulation variationnelle

L'originalité du système (A.13) réside dans le choix de conserver les deux densités de courant comme inconnues alors qu'elles sont liées par la condition de Leontovich [99] sur la frontière de l'objet. Cette formulation n'utilise que les opérateurs intégraux classiques T et K qui ne posent pas de problème numérique lorsque l'on utilise les éléments finis classiques de Raviart-Thomas. L'autre avantage est que l'impédance n'est pas appliquée aux opérateurs intégraux T et K et intervient de façon localisée dans le système linéaire discret.

On introduit le produit scalaire $\langle u, v \rangle_\Gamma = \int_\Gamma u \cdot v \, d\Gamma$ et le changement d'inconnue $J = (\eta j)$ que l'on notera par abus de notation de nouvelle j , la formulation variationnelle associée au système (A.13) consiste alors à trouver $(j, m) \in \mathcal{H}^{-1/2}(\text{div}_\Gamma, \Gamma)$ tels que :

$$\begin{aligned} \forall (j', m') \in \mathcal{H}^{-1/2}(\text{div}_\Gamma, \Gamma), \quad & \langle E_{tg}^{inc}, j' \rangle_\Gamma - \langle \eta H_{tg}^{inc}, m' \rangle_\Gamma = \langle ikTj, j' \rangle_\Gamma \\ & + \langle -ikTm, m' \rangle_\Gamma - \langle Km, j' \rangle_\Gamma - \langle Kj, m' \rangle_\Gamma + \frac{1}{2} \langle zj, j' \rangle_\Gamma - \frac{1}{2} \langle z^{-1}m, m' \rangle_\Gamma \end{aligned} \quad (\text{A.14})$$

A.4 Discrétisation

Éléments finis de Raviart-Thomas – Pour résoudre numériquement la formulation variationnelle (A.14), on utilise les éléments finis de Raviart-Thomas [21], conformes pour l'opérateur de divergence surfacique. Ces éléments finis sont les plus simples et les plus utilisés pour discrétiser les équations intégrales surfaciques de l'électromagnétisme. La surface Γ est approchée par une triangulation (ou maillage) $T_h = \cup_{l=1}^N T_l$ constituée de N_a arêtes et N éléments triangulaires. Les courants surfaciques s'expriment comme une combinaison linéaire des fonctions de base associées aux arêtes de T_h :

$$j(x) = \sum_{j=1}^{N_a} \alpha_j \varphi_j(x), \quad m(x) = \sum_{j=1}^{N_a} \beta_j \varphi_j(x) \quad (\text{A.15})$$

Les coefficients (α_j, β_j) dans la base correspondent au flux des courants à travers les arêtes : $\alpha_j = \int_{a_j} j \cdot \nu_j \, dl$, $\beta_j = \int_{a_j} m \cdot \nu_j \, dl$, où ν_j est la normale à l'arête a_j . Le support de la fonction de base φ_j associée à la j -ième arête est la réunion des triangles (T^+, T^-) adjacents à cette arête.

Fonctions test – Pour la formulation variationnelle (A.14), on utilise comme fonctions test les fonctions de base des densités de courant :

$$\left(j' = \varphi_i, \quad i = 1 \dots N_a \right) \quad \text{et} \quad \left(m' = \varphi_i, \quad i = 1 \dots N_a \right), \quad (\text{A.16})$$

et ceci nous donne l'approximation suivante :

$$\begin{aligned} \langle E_{tg}^{inc}, \varphi_i \rangle_\Gamma - \langle \eta H_{tg}^{inc}, \varphi_i \rangle_\Gamma &= \sum_{j=1}^{N_a} ik \langle T \varphi_j, \varphi_i \rangle_\Gamma \alpha_j - \langle K \varphi_j, \varphi_i \rangle_\Gamma \beta_j + \frac{1}{2} \langle z \varphi_j, \varphi_i \rangle_\Gamma \alpha_j \\ &+ \sum_{j=1}^{N_a} -ik \langle T \varphi_j, \varphi_i \rangle_\Gamma \beta_j - \langle K \varphi_j, \varphi_i \rangle_\Gamma \alpha_j - \frac{1}{2} \langle z^{-1} \varphi_j, \varphi_i \rangle_\Gamma \beta_j \end{aligned} \quad (\text{A.17})$$

On note alors :

$$\begin{cases} \mathbf{T}_{i,j} = \langle T\varphi_j, \varphi_i \rangle_\Gamma \\ \mathbf{K}_{i,j} = \langle K\varphi_j, \varphi_i \rangle_\Gamma \\ \mathbf{M}_{i,j}(z) = \langle z\varphi_j, \varphi_i \rangle_\Gamma \end{cases} \quad (\text{A.18})$$

et notamment :

$$\begin{aligned} \mathbf{T}_{i,j} &= \int_{T_h} T\varphi_j(x) \varphi_i(x) \, dx = \int_{T_h} \left(\int_{T_h} G(x,y) \varphi_j(y) \, dy \right) \varphi_i(x) \, dx \\ &\quad + \frac{1}{k^2} \int_{T_h} \nabla_x \left(\int_{T_h} G(x,y) \operatorname{div}_\Gamma \varphi_j(y) \, dy \right) \varphi_i(x) \, dx \end{aligned} \quad (\text{A.19})$$

Et après intégration par parties, on a

$$\mathbf{T}_{i,j} = \iint_{T_h} G(x,y) \{ \varphi_i(x) \cdot \varphi_j(y) - \frac{1}{k^2} \operatorname{div}_\Gamma \varphi_i(x) \operatorname{div}_\Gamma \varphi_j(y) \} \, dy \, dx \quad (\text{A.20})$$

et de même :

$$\begin{aligned} \mathbf{K}_{i,j} &= \int_{T_h} K\varphi_j(x) \varphi_i(x) \, dx \\ &= \int_{T_h} \left(\int_{T_h} \nabla_y G(x,y) \times \varphi_j(y) \, dy \right) \varphi_i(x) \, dx \end{aligned} \quad (\text{A.21})$$

$$\mathbf{M}_{i,j}(z) = \int_{T_h} z(x) \varphi_j(x) \varphi_i(x) \, dx \quad (\text{A.22})$$

Matrice $\mathbf{A}(z)$ – La discrétisation de la formulation variationnelle (A.14) permettant d’obtenir les densités de courants électrique et magnétique donne le système matriciel final :

$$\begin{bmatrix} ik\mathbf{T} + \frac{1}{2}\mathbf{M}(z) & -\mathbf{K} \\ -\mathbf{K} & -ik\mathbf{T} - \frac{1}{2}\mathbf{M}(z^{-1}) \end{bmatrix} \begin{bmatrix} \mathbf{X}_J \\ \mathbf{X}_M \end{bmatrix} = \begin{bmatrix} \mathbf{b}_E \\ \mathbf{b}_H \end{bmatrix} \quad (\text{A.23})$$

avec les coefficients

$$\begin{aligned} \mathbf{X}_J &= (\alpha_1, \dots, \alpha_{N_a})^t, \\ \mathbf{X}_M &= (\beta_1, \dots, \beta_{N_a})^t, \\ \mathbf{b}_E &= (\langle E_{tg}^{inc}, \varphi_1 \rangle_\Gamma, \dots, \langle E_{tg}^{inc}, \varphi_{N_a} \rangle_\Gamma)^t, \\ \mathbf{b}_H &= (\langle -H_{tg}^{inc}, \varphi_1 \rangle_\Gamma, \dots, \langle -H_{tg}^{inc}, \varphi_{N_a} \rangle_\Gamma)^t. \end{aligned}$$

On simplifie les notations du système matriciel, en posant :

$$\mathbf{A}(z) = \begin{bmatrix} ik\mathbf{T} + \frac{1}{2}\mathbf{M}(z) & -\mathbf{K} \\ -\mathbf{K} & -ik\mathbf{T} - \frac{1}{2}\mathbf{M}(z^{-1}) \end{bmatrix}, \mathbf{X} = (\mathbf{X}_J, \mathbf{X}_M)^t, \mathbf{b} = (\mathbf{b}_E, \mathbf{b}_H)^t \quad (\text{A.24})$$

Les densités de courants électrique et magnétique sont ainsi solutions du système matriciel de dimension $(2N_a \times 2N_a)$:

$$\mathbf{A}(z) \mathbf{X} = \mathbf{b} \quad (\text{A.25})$$

La matrice \mathbf{A} dépend de la fréquence k et de l'impédance, z mais est indépendante du champ incident. Pour l'analyse de la diffraction d'un objet, de nombreux angles d'éclairage radar sont à considérer. Ainsi, pour une fréquence donnée le système linéaire à résoudre s'exprime plutôt comme :

$$\mathbf{A}(z) \mathbf{X} = \mathbf{B} \quad (\text{A.26})$$

où \mathbf{B} est de dimension $2N_a \times M$ avec M le nombre d'angles d'éclairage et de polarisations des champs électromagnétiques incidents.

Remarque 38. *Les équations intégrales ont été construites de façon à assurer la symétrie du système linéaire associé. Cette symétrie est exploitée en ne stockant que la partie inférieure de la matrice \mathbf{A} permettant ainsi de réduire le coût du traitement des deux courants j et m . La matrice \mathbf{A} est dense, c'est-à-dire que les $4N_a^2$ coefficients sont différents de zéro. Cependant l'impédance caractérisant le revêtement à la surface de l'objet n'intervient que dans une intégrale simple de la formulation variationnelle. La « sous-matrice » \mathbf{M} associée à la contribution de l'impédance est creuse. C'est un des avantages de la formulation variationnelle retenue, la dérivée de \mathbf{A} par rapport à l'impédance est très simple à exprimer et son coût de calcul est négligeable.*

Discrétisation de z – En pratique, l'impédance z est considérée comme constante par élément de maillage. On introduit alors le vecteur d'impédance complexe $\mathbf{z} = [z_1, z_2, \dots, z_N]^t$ à N composantes. Le système (A.26) s'écrit alors :

$$\mathbf{A}(\mathbf{z}) \mathbf{X} = \mathbf{B} \quad (\text{A.27})$$

A.5 Calcul de la dérivée

On rappelle que la matrice jacobienne \mathbf{H} s'écrit :

$$\mathbf{H} = \left. \frac{\partial \mathbf{h}}{\partial \mathbf{z}} \right|_{z_0} = \left. \frac{\partial \mathbf{X}}{\partial \mathbf{z}} \right|_{z_0}^t \mathbf{c}. \quad (\text{A.28})$$

Afin d'obtenir $\partial \mathbf{X} / \partial \mathbf{z}$, on dérive l'Eq. (A.27) par rapport à \mathbf{z} :

$$\mathbf{A}(\mathbf{z}) \frac{\partial \mathbf{X}}{\partial \mathbf{z}} + \frac{\partial \mathbf{A}}{\partial \mathbf{z}} \mathbf{X}(\mathbf{z}) = \mathbf{0} \quad (\text{A.29})$$

et en \mathbf{z}_0 , la matrice jacobienne se construit à partir de la résolution du système :

$$\mathbf{A}(\mathbf{z}_0) \left. \frac{\partial \mathbf{X}}{\partial \mathbf{z}} \right|_{z=z_0} = - \left. \frac{\partial \mathbf{A}}{\partial \mathbf{z}} \right|_{z=z_0} \mathbf{X}(\mathbf{z}_0) \quad (\text{A.30})$$

On pose $\mathbf{z}^{-1} = [z_1^{-1}, z_2^{-1}, \dots, z_N^{-1}]^t$. On reprend alors le système linéaire (A.24) et on dérive par rapport à l'impédance. Il vient :

$$\frac{\partial \mathbf{A}}{\partial \mathbf{z}} = \frac{1}{2} \frac{\partial}{\partial \mathbf{z}} \begin{bmatrix} \mathbf{M}(\mathbf{z}) & \mathbf{0} \\ \mathbf{0} & -\mathbf{M}(\mathbf{z}^{-1}) \end{bmatrix} \quad (\text{A.31})$$

où l'élément i, j de $\mathbf{M}(z)$ s'écrit

$$\mathbf{M}_{i,j}(z) = \sum_{l=1}^N \int_{T_l} z_l \varphi_j(x) \varphi_i(x) dx = \sum_{l=1}^N z_l \mathbf{m}_{ij}^l, \quad (\text{A.32})$$

et $\mathbf{m}_{ij}^l = \int_{T_l} \varphi_j(x) \varphi_i(x) dx$. φ_i (respectivement φ_j) est la fonction de base attachée à l'arête i (respectivement j) du maillage et z_l est la valeur moyenne de l'impédance sur l'élément T_l calculée au barycentre de l'élément.

Matrice creuse – La dérivée par rapport à la l -ième composante de l'impédance z_l est :

$$\frac{\partial \mathbf{A}}{\partial z_l} = \frac{1}{2} \begin{bmatrix} \mathbf{m}^l & \mathbf{0} \\ \mathbf{0} & z_l^{-2} \mathbf{m}^l \end{bmatrix} \quad (\text{A.33})$$

En pratique, une variation d'impédance sur un élément T_l du maillage donne une matrice dérivée $\partial \mathbf{A} / \partial z_l$ avec seulement dix-huit coefficients non nuls, neuf coefficients (matrice \mathbf{m}^l) affectent la densité du courant électrique, les neuf autres (matrice $z_l^{-2} \mathbf{m}^l$) la densité du courant magnétique.

Conclusion – Le travail présenté dans cette annexe a été réalisé en amont de ma thèse par le CEA. Il a permis d'obtenir une linéarisation des équations de Maxwell vis-à-vis de l'impédance surfacique et de calculer la matrice jacobienne \mathbf{H} essentielle à mon travail.

ANNEXE B

IMPÉDANCE COMPLEXE

On s'intéresse à un champ gaussien complexe de moyenne \mathbf{m} et de covariance \mathbf{C} . Le but est de montrer que lorsque la matrice \mathbf{C} est réelle, la partie réelle et la partie imaginaire sont indépendantes. Pour cela, on note

- $\mathbf{z} = \mathbf{x} + i\mathbf{y}$, le vecteur complexe de \mathbb{C}^N , \mathbf{x} et \mathbf{y} de \mathbb{R}^N les parties réelles et imaginaires.
- $\mathbf{m} = \mathbf{m}_x + i\mathbf{m}_y$, sa moyenne avec $\mathbf{m} \in \mathbb{C}^N$, $\mathbf{m}_x \in \mathbb{R}^N$ et $\mathbf{m}_y \in \mathbb{R}^N$.
- \mathbf{C} , sa matrice de covariance avec $\mathbf{C} \in \mathbb{R}^{N \times N}$.

La densité de probabilité de \mathbf{z} s'écrit

$$f(\mathbf{z}) = \pi^{-N} \det \mathbf{C}^{-1} \exp - \|\mathbf{z} - \mathbf{m}\|_{\mathbf{C}^{-1}}^2. \quad (\text{B.1})$$

Par changement de variable, on peut montrer que

$$f(\mathbf{x}, \mathbf{y}) = \pi^{-N} \det \mathbf{C}^{-1} \exp - \|\bar{\mathbf{x}} + i\bar{\mathbf{y}}\|_{\mathbf{C}^{-1}}^2, \quad (\text{B.2})$$

avec $\bar{\mathbf{x}} = \mathbf{x} - \mathbf{m}_x$ et $\bar{\mathbf{y}} = \mathbf{y} - \mathbf{m}_y$.

En développant la norme on a

$$\begin{aligned} \|\bar{\mathbf{x}} + i\bar{\mathbf{y}}\|_{\mathbf{C}^{-1}}^2 &= (\bar{\mathbf{x}} + i\bar{\mathbf{y}})^\dagger \mathbf{C}^{-1} (\bar{\mathbf{x}} + i\bar{\mathbf{y}}) \\ &= \bar{\mathbf{x}}^\dagger \mathbf{C}^{-1} \bar{\mathbf{x}} + \bar{\mathbf{x}}^\dagger \mathbf{C}^{-1} (i\bar{\mathbf{y}}) + (i\bar{\mathbf{y}})^\dagger \mathbf{C}^{-1} \bar{\mathbf{x}} + (i\bar{\mathbf{y}})^\dagger \mathbf{C}^{-1} (i\bar{\mathbf{y}}) \\ &= \bar{\mathbf{x}}^\dagger \mathbf{C}^{-1} \bar{\mathbf{x}} + i\bar{\mathbf{x}}^\dagger \mathbf{C}^{-1} \bar{\mathbf{y}} - i\bar{\mathbf{y}}^\dagger \mathbf{C}^{-1} \bar{\mathbf{x}} - \bar{\mathbf{y}}^\dagger \mathbf{C}^{-1} \bar{\mathbf{y}} \\ &= \bar{\mathbf{x}}^\dagger \mathbf{C}^{-1} \bar{\mathbf{x}} + \bar{\mathbf{y}}^\dagger \mathbf{C}^{-1} \bar{\mathbf{y}} = \|\bar{\mathbf{x}}\|_{\mathbf{C}^{-1}}^2 + \|\bar{\mathbf{y}}\|_{\mathbf{C}^{-1}}^2. \end{aligned} \quad (\text{B.3})$$

Remarque 39. Comme \mathbf{x} , \mathbf{y} et \mathbf{C} sont réelles, on a

$$\begin{aligned} i\bar{\mathbf{x}}^\dagger \mathbf{C}^{-1} \bar{\mathbf{y}} - i\bar{\mathbf{y}}^\dagger \mathbf{C}^{-1} \bar{\mathbf{x}} &= i(\bar{\mathbf{x}}^\dagger \mathbf{C}^{-1} \bar{\mathbf{y}} - \bar{\mathbf{y}}^\dagger \mathbf{C}^{-1} \bar{\mathbf{x}}) \\ &= i(\bar{\mathbf{x}}^\dagger \mathbf{C}^{-1} \bar{\mathbf{y}} - \bar{\mathbf{x}}^\dagger \mathbf{C}^{-1} \bar{\mathbf{y}}) \\ &= 0. \end{aligned}$$

En introduisant la matrice $\tilde{\mathbf{C}} = \mathbf{C}/2$, la densité devient

$$\begin{aligned}
 f(\mathbf{x}, \mathbf{y}) &= \pi^{-N} \det(2\tilde{\mathbf{C}})^{-1} \exp - \|\bar{\mathbf{x}}\|_{(2\tilde{\mathbf{C}})^{-1}}^2 - \|\bar{\mathbf{y}}\|_{(2\tilde{\mathbf{C}})^{-1}}^2 \\
 &= (2\pi)^{-N} \det(\tilde{\mathbf{C}})^{-1} \exp - \frac{1}{2} \|\bar{\mathbf{x}}\|_{\tilde{\mathbf{C}}^{-1}}^2 - \frac{1}{2} \|\bar{\mathbf{y}}\|_{\tilde{\mathbf{C}}^{-1}}^2 \\
 &= (2\pi)^{-N/2} \det(\tilde{\mathbf{C}})^{-1/2} \exp - \frac{1}{2} \|\bar{\mathbf{x}}\|_{\tilde{\mathbf{C}}^{-1}}^2 \\
 &\quad (2\pi)^{-N/2} \det(\tilde{\mathbf{C}})^{-1/2} \exp - \frac{1}{2} \|\bar{\mathbf{y}}\|_{\tilde{\mathbf{C}}^{-1}}^2.
 \end{aligned} \tag{B.4}$$

Finalement, on peut écrire : $f(\mathbf{x}, \mathbf{y}) = f(\mathbf{x})f(\mathbf{y})$, avec

$$f(\mathbf{x}) = \mathcal{N}(\mathbf{x}; \mathbf{m}_x, \mathbf{C}/2) \tag{B.5}$$

$$f(\mathbf{y}) = \mathcal{N}(\mathbf{y}; \mathbf{m}_y, \mathbf{C}/2). \tag{B.6}$$

Par construction, la partie réelle \mathbf{x} et la partie imaginaire \mathbf{y} du champ \mathbf{z} sont donc indépendantes. Cela est dû au fait que la matrice de covariance \mathbf{C} est réelle. La partie réelle \mathbf{x} et la partie imaginaire \mathbf{y} ont pour moyenne respectivement \mathbf{m}_x et \mathbf{m}_y , tandis qu'elles ont la même covariance qui vaut $\mathbf{C}/2$.

Remarque 40. *Si la matrice de covariance \mathbf{C} était complexe, des termes croisés apparaîtraient, témoignant d'une corrélation entre la partie réelle \mathbf{x} et la partie imaginaire \mathbf{y} .*

ANNEXE C

MATRICE DE STRUCTURE

Cette annexe porte sur les matrices de structure $\mathbf{\Pi}_k$ utilisées dans la modélisation *a priori* des impédances z_k . L'objectif est de trouver les valeurs du paramètre d'interaction α_k qui mènent à des matrices $\mathbf{\Pi}_k$ inversibles. Pour simplifier les notations, on ne fait pas apparaître la moyenne et on ne s'intéresse qu'à une seule classe. On note alors, seulement dans cette annexe, $\mathbf{\Pi} = \mathbf{\Pi}_k$, $\mathbf{D} = \mathbf{D}_k$ et $\alpha = \alpha_k$.

On rappelle que la matrice de structure s'écrit

$$\mathbf{\Pi} = \mathbf{D}^t \mathbf{D}, \quad (\text{C.1})$$

avec \mathbf{D} la matrice des différences finies d'ordre 1 pondérée par un coefficient d'interaction $\alpha \in \mathbb{R}$. L'inversibilité de $\mathbf{\Pi}$ implique la condition suivante sur $\alpha \in \mathbb{R}$:

$$\forall \mathbf{z} \in \mathbb{C}_*^N, \mathbf{z}^\dagger \mathbf{\Pi} \mathbf{z} = \|\mathbf{D} \mathbf{z}\|^2 = \sum_{n \sim n'} |z_n - \alpha z_{n'}|^2 > 0. \quad (\text{C.2})$$

La condition (C.2) n'est pas vérifiée si la norme est nulle, c'est-à-dire si chaque terme de la somme est nul. Dit autrement, il suffit qu'un terme soit non nul pour que la somme soit non nulle. On précise que pour chaque paire (n, n') , il y a deux termes associés : $|z_n - \alpha z_{n'}|^2 + |z_{n'} - \alpha z_n|^2$. Pour trouver les configurations qui ne vérifient pas (C.2), on s'intéresse dans un premier temps au cas où au moins une composante de \mathbf{z} est nulle et, dans un second temps, au cas où toutes les composantes sont non nulles.

Au moins une composante nulle – On se place dans le cas où au moins une composante de \mathbf{z} est nulle. Comme $\mathbf{z} \in \mathbb{C}_*^N$ alors il existe au minimum une composante non nulle. Dans cette configuration, parmi tous les termes de la somme $\sum_{n \sim n'}$, il existe forcément un terme pour lequel une composante est nulle et l'autre est non nulle. Par exemple, $z_{n_0} = 0$ et $z_{n^*} \neq 0$ et on a alors :

$$|z_{n_0} - \alpha z_{n^*}|^2 + |z_{n^*} - \alpha z_{n_0}|^2 = (1 + \alpha^2) |z_{n^*}|^2 > 0. \quad (\text{C.3})$$

On en déduit que $\forall \mathbf{z} \in \mathbb{C}_*^N$, si au moins une composante de \mathbf{z} est nulle, alors il existe au minimum un terme de la somme strictement positif et cela, peu importe la valeur de $\alpha \in \mathbb{R}$.

Tous les éléments non nuls – On se place dans le cas où toutes les composantes de $z \in \mathbb{C}_*^N$ sont non nulles, c'est-à-dire $z_n \in \mathbb{C}^*$ pour $n = 1 \dots N$. Dans un premier temps, on cherche la condition sur α pour annuler un terme n, n' quelconque de la somme :

$$z_n - \alpha z_{n'} = 0 \quad \text{si} \quad \alpha = z_n / z_{n'}.$$

Maintenant, on considère le terme symétrique n', n associé au premier. Pour annuler ce terme, il faut que $\alpha = z_{n'} / z_n$. Pour annuler les deux termes, α doit alors vérifier les deux égalités simultanément :

$$\alpha = \frac{z_n}{z_{n'}} = \frac{z_{n'}}{z_n}. \quad (\text{C.4})$$

Cela nécessite que

$$z_n = \pm z_{n'} \quad \text{avec} \quad \alpha = \pm 1. \quad (\text{C.5})$$

Pour une paire (n, n') , il existe donc deux configurations dans lesquelles les termes associés peuvent s'annuler. On souhaite maintenant étendre cette analyse dans le cas où l'on considèrerait toutes les paires. On rappelle que le maillage représente la surface d'un objet fermé. Il y a donc une certaine circularité.

Pour servir d'exemple, on se place dans un cas circulant 1D de dimensions N où chaque maille n est voisine de la maille précédente $n - 1$ et de la maille suivante $n + 1$. D'après (C.4), pour que tous les termes associés à toutes les paires s'annulent, il faut que l'on ait les égalités suivantes :

$$\alpha = \left(\frac{z_1}{z_2} = \frac{z_2}{z_1} \right) = \left(\frac{z_2}{z_3} = \frac{z_3}{z_2} \right) = \dots = \left(\frac{z_N}{z_1} = \frac{z_1}{z_N} \right). \quad (\text{C.6})$$

Pour que les égalités (C.6) soient vérifiées, il n'y a que deux possibilités :

1. Les composantes sont identiques : $z_1 = z_2 = z_3 = \dots = z_N$ et $\alpha = 1$.
2. Les composantes sont à valeurs alternées : $z_1 = -z_2 = z_3 = \dots = -z_N$ (cela n'est possible que lorsque N est pair) et $\alpha = -1$.

Si l'on revient à notre cas surfacique et au voisinage à 3 voisins, la première possibilité est facile à réaliser tandis que la seconde n'est pas possible. En effet, le maillage ne possède pas la même régularité que celui d'un cas 1D ou même d'un maillage 2D régulier. En conséquence, on retient qu'il faut $\alpha \neq 1$ pour que la condition (C.2) soit vérifiée.

Corrélation positive – Dans ce travail, on souhaite modéliser un champ d'impédance pour lequel la corrélation est positive. Pour cette raison, on ne s'intéressera qu'à des coefficients α positifs. Finalement, il suffit que $\alpha \in \mathbb{R}_+ \setminus \{1\}$ pour que la matrice $\mathbf{\Pi}$ soit inversible et que la corrélation soit positive.

ANNEXE D

MATRICE BLOC

On introduit la matrice carrée de taille $N \times N$ contenant toutes les matrices de troncature des K classes concaténées de 1 à K

$$\mathcal{T} = \begin{bmatrix} \mathbf{T}_1(\ell) \\ \vdots \\ \mathbf{T}_K(\ell) \end{bmatrix}, \quad (\text{D.1})$$

qui est une matrice de permutation¹ puisque chaque étiquette prend une valeur et une seule dans $\{1 \dots K\}$. On a donc $\mathcal{T}\mathcal{T}^t = \mathcal{T}^t\mathcal{T} = \mathbf{I}_N$, soit $\mathcal{T}^t = \mathcal{T}^{-1}$. On introduit également la notation de la matrice bloc diagonale constituée des blocs $\bar{\mathbf{C}}_k(\ell)$ rangés de 1 à K

$$\text{diag}\{\bar{\mathbf{C}}_1(\ell), \dots, \bar{\mathbf{C}}_K(\ell)\} = \begin{bmatrix} \bar{\mathbf{C}}_1(\ell) & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \bar{\mathbf{C}}_K(\ell) \end{bmatrix}. \quad (\text{D.2})$$

On rappelle l'écriture de la matrice bloc diagonale à une permutation près $\mathbf{C}_{z|\ell}$ constituée des blocs $\bar{\mathbf{C}}_k(\ell) = \mathbf{T}_k(\ell)\mathbf{C}_k\mathbf{T}_k^t(\ell)$ comme

$$\begin{aligned} \mathbf{C}_{z|\ell} &= \sum_{k=1}^K \mathbf{T}_k^t(\ell)\bar{\mathbf{C}}_k(\ell)\mathbf{T}_k(\ell) \\ &= \mathbf{T}_1^t(\ell)\bar{\mathbf{C}}_1(\ell)\mathbf{T}_1(\ell) + \dots + \mathbf{T}_K^t(\ell)\bar{\mathbf{C}}_K(\ell)\mathbf{T}_K(\ell) \\ &= \mathcal{T}^t \text{diag}\{\bar{\mathbf{C}}_1(\ell), \dots, \bar{\mathbf{C}}_K(\ell)\} \mathcal{T}. \end{aligned} \quad (\text{D.3})$$

D.1 Déterminant de la matrice bloc

On souhaite montrer que le déterminant de $\mathbf{C}_{z|\ell}$ s'écrit comme le produit des déterminants des blocs $\bar{\mathbf{C}}_k(\ell)$. Pour cela, on reprend la forme matricielle de l'Eq. (D.3)

$$\det \mathbf{C}_{z|\ell} = \det (\mathcal{T}^t \text{diag}\{\bar{\mathbf{C}}_1(\ell), \dots, \bar{\mathbf{C}}_K(\ell)\} \mathcal{T}). \quad (\text{D.4})$$

1. Une matrice de permutation est une matrice carrée qui vérifie les propriétés suivantes : les coefficients sont 0 ou 1 ; il y a un et un seul 1 par ligne ; il y a un et un seul 1 par colonne.

Les matrices étant carrées, le déterminant des produits étant égal au produit des déterminants

$$\det \mathbf{C}_{z|\ell} = \det (\mathcal{T}^t) \det (\text{diag}\{\bar{\mathbf{C}}_1(\ell), \dots, \bar{\mathbf{C}}_K(\ell)\}) \det (\mathcal{T}). \quad (\text{D.5})$$

Le déterminant d'une matrice de permutation vaut 1 si le nombre de permutations est pair et -1 s'il est impair. On a donc

$$\det \mathbf{C}_{z|\ell} = \det (\text{diag}\{\bar{\mathbf{C}}_1(\ell), \dots, \bar{\mathbf{C}}_K(\ell)\}). \quad (\text{D.6})$$

Le déterminant d'une matrice bloc diagonale étant égal au produit des déterminants des blocs, on a

$$\det \mathbf{C}_{z|\ell} = \prod_{k=1}^K \det \bar{\mathbf{C}}_k(\ell). \quad (\text{D.7})$$

On retrouve donc bien que le déterminant d'une matrice bloc diagonale à une permutation près est égale au produit des déterminants des blocs.

D.2 Inverse de la matrice bloc

Dans cette partie, on souhaite montrer que l'inverse de la matrice $\mathbf{C}_{z|\ell}$ peut s'écrire à partir de l'inverse des blocs $\bar{\mathbf{C}}_k(\ell)$. Pour cela, on reprend la forme matricielle de l'Eq. (D.3)

$$\mathbf{C}_{z|\ell}^{-1} = (\mathcal{T}^t \text{diag}\{\bar{\mathbf{C}}_1(\ell), \dots, \bar{\mathbf{C}}_K(\ell)\} \mathcal{T})^{-1}. \quad (\text{D.8})$$

Les matrices de troncature à gauche et à droite et la matrice bloc diagonale sont inversibles. On écrit donc

$$\mathbf{C}_{z|\ell}^{-1} = \mathcal{T}^{-1} \text{diag}\{\bar{\mathbf{C}}_1(\ell), \dots, \bar{\mathbf{C}}_K(\ell)\}^{-1} \mathcal{T}^{-t}. \quad (\text{D.9})$$

D'après la propriété $\mathcal{T}^t = \mathcal{T}^{-1}$, on obtient

$$\mathbf{C}_{z|\ell}^{-1} = \mathcal{T}^t \text{diag}\{\bar{\mathbf{C}}_1(\ell), \dots, \bar{\mathbf{C}}_K(\ell)\}^{-1} \mathcal{T}. \quad (\text{D.10})$$

L'inverse d'une matrice bloc diagonale est égale à la matrice bloc diagonale constituée des blocs inverses. On a donc

$$\mathbf{C}_{z|\ell}^{-1} = \mathcal{T}^t \text{diag}\{\bar{\mathbf{C}}_1(\ell)^{-1}, \dots, \bar{\mathbf{C}}_K(\ell)^{-1}\} \mathcal{T}, \quad (\text{D.11})$$

et en développant le calcul matriciel, on obtient

$$\begin{aligned} \mathbf{C}_{z|\ell}^{-1} &= \mathbf{T}_1^t(\ell) \bar{\mathbf{C}}_1(\ell)^{-1} \mathbf{T}_1(\ell) + \dots + \mathbf{T}_K^t(\ell) \bar{\mathbf{C}}_K(\ell)^{-1} \mathbf{T}_K(\ell) \\ &= \sum_{k=1}^K \mathbf{T}_k^t(\ell) \bar{\mathbf{C}}_k(\ell)^{-1} \mathbf{T}_k(\ell). \end{aligned} \quad (\text{D.12})$$

L'inverse d'une matrice bloc diagonale à une permutation près est également une matrice bloc diagonale à la même permutation près et les blocs de la matrice inverse sont égaux aux inverses des blocs de la matrice initiale.

ANNEXE E

ÉCHANTILLONNAGE DES IMPÉDANCES *A POSTERIORI*

On s'intéresse à la procédure d'échantillonnage des K impédances *a posteriori* proposée dans la partie 4.2.1.2 page 63. Elle s'inspire de la note de A. Doucet [90] et mène à l'échantillonnage de l'impédance tel que défini par l'étape ④ page 64. Cette dernière peut être réalisée à partir des deux formes suivantes où, pour alléger les notations, on ne fait pas apparaître les indices de classe k et les étiquettes ℓ .

$$\text{Forme covariance} : z = \tilde{z} + CPH^\dagger(HPCPH^\dagger + C_\varepsilon)^{-1}(e_\sigma - \tilde{e}_\sigma). \quad (\text{E.1})$$

$$\text{Forme précision} : z = \tilde{z} + \gamma^{-1}\Pi^{-1}PH^\dagger(HP\gamma^{-1}\Pi^{-1}PH^\dagger + C_\varepsilon)^{-1}(e_\sigma - \tilde{e}_\sigma). \quad (\text{E.2})$$

La forme covariance est obtenue en remplaçant $\gamma^{-1}\Pi^{-1}$ par C . L'objectif est de trouver celle qui possède la complexité calculatoire la plus faible. On peut factoriser certains termes en exploitant les décompositions de Cholesky

$$C = L_C L_C^\dagger, \quad (\text{E.3})$$

$$\Pi = L_\Pi L_\Pi^\dagger, \quad (\text{E.4})$$

et en introduisant

$$A_C = HPL_C, \quad (\text{E.5})$$

$$A_\Pi = HPL_\Pi^{-\dagger}. \quad (\text{E.6})$$

Les deux formes, respectivement covariance et précision, deviennent alors :

$$z = \tilde{z} + L_C A_C^\dagger (A_C A_C^\dagger + C_\varepsilon)^{-1} (e_\sigma - \tilde{e}_\sigma), \quad (\text{E.7})$$

$$z = \tilde{z} + \gamma^{-1} L_\Pi^{-\dagger} A_\Pi^\dagger (\gamma^{-1} A_\Pi A_\Pi^\dagger + C_\varepsilon)^{-1} (e_\sigma - \tilde{e}_\sigma). \quad (\text{E.8})$$

Dans la suite, on considère que la matrice de projection $P = I$ afin de simplifier l'analyse des complexités calculatoires.

Complexité calculatoire – Dans les Eq. (E.7) et (E.8), si l'on agence les calculs de droite à gauche, alors on effectue systématiquement des produits matrice vecteur, à l'exception de $A_C A_C^\dagger$

et $\mathbf{A}_\Pi \mathbf{A}_\Pi^\dagger$ qui sont des produits de matrices de petite taille $M \times M$, donc peu coûteux : $\mathcal{O}(M^3)$. Les résolutions des systèmes

$$\begin{aligned} (\mathbf{A}_C \mathbf{A}_C^\dagger + \mathbf{C}_\varepsilon) \mathbf{u}_C &= \mathbf{e}_\sigma - \tilde{\mathbf{e}}_\sigma, \\ (\gamma^{-1} \mathbf{A}_\Pi \mathbf{A}_\Pi^\dagger + \mathbf{C}_\varepsilon) \mathbf{u}_\Pi &= \mathbf{e}_\sigma - \tilde{\mathbf{e}}_\sigma, \end{aligned}$$

sont peu coûteuses : $\mathcal{O}(M^3)$, car ces derniers sont de petite taille M . Les calculs suivants, qui reposent sur les matrices triangulaires \mathbf{L}_C et \mathbf{L}_Π , ont les complexités calculatoires :

$$\begin{aligned} \mathbf{L}_C \mathbf{A}_C^\dagger \mathbf{u}_C &: \mathcal{O}(MN + N^2/2), \\ \gamma^{-1} \mathbf{L}_\Pi^\dagger \mathbf{A}_\Pi^\dagger \mathbf{u}_\Pi &: \mathcal{O}(MN + N^2/2). \end{aligned}$$

L'opération la plus coûteuse correspond aux factorisations des Eq. (E.5) et (E.6). On rappelle que l'on a choisi, pour simplifier, de considérer que la matrice de projection $\mathbf{P} = \mathbf{I}$. Pour obtenir \mathbf{A}_C , il suffit de faire le produit de deux matrices tandis que pour obtenir \mathbf{A}_Π , il faut résoudre M systèmes. Cela mène aux complexités calculatoires suivantes :

$$\mathbf{A}_C = \mathbf{H} \mathbf{L}_C : MN(N + 1)/2 \quad \text{produits} : \mathcal{O}(MN^2/2) \quad (\text{E.9})$$

$$\mathbf{L}_\Pi \mathbf{A}_\Pi^\dagger = \mathbf{H}^\dagger : MN(N - 1)/2 \quad \text{produits et } MN \text{ divisions} : \mathcal{O}(MN^2/2) \quad (\text{E.10})$$

Dans notre cas, la matrice triangulaire \mathbf{L}_Π est grandement creuse tandis que la matrice triangulaire \mathbf{L}_C est pleine. La matrice \mathbf{L}_Π compte environ $3.5 \cdot 10^5$ éléments non nuls contre environ $2.7 \cdot 10^7 \approx N^2/2$ éléments non nuls pour \mathbf{L}_C . De ce fait, le nombre de produits à réaliser pour obtenir \mathbf{A}_Π est bien moindre que pour \mathbf{A}_C , et cela d'un facteur 100.

Constats pratiques – Dans la pratique, on obtient les temps de calcul suivants selon que l'on choisit un encodage « creux » ou « plein » en Matlab. On rappelle que $M = 300$ et $N = 7400$.

	Encodage « plein »	Encodage « creux »
$\mathbf{A}_C = \mathbf{H} \mathbf{L}_C$	~ 0.25 sec	~ 7.5 sec
$\mathbf{L}_\Pi \mathbf{A}_\Pi^\dagger = \mathbf{H}^\dagger$	~ 0.17 sec	~ 0.11 sec

TABLE E.1 – Temps de calcul obtenus sur une machine standard¹ avec Matlab 2020b.

Pour comprendre les temps de calcul, on propose une analyse séparée en fonction de l'encodage.

- Pour l'encodage « plein », le rapport entre les deux temps de calcul est de l'ordre de 1.5. Cela vient du fait que Matlab ne tient pas compte du caractère creux des matrices et que la fonction de résolution exploite la propriété triangulaire, contrairement à la fonction de multiplication. Cette dernière réalise donc deux fois plus d'opérations que la fonction de résolution.

- Pour l'encodage « creux », le rapport entre les deux temps de calcul est de 70. Il s'approche donc du facteur 100 que l'on s'attendait à trouver et s'explique par le fait que Matlab tient compte du caractère creux des matrices.

On constate un écart important entre les temps de calcul obtenus avec l'encodage « plein » et « creux ». Pour aller plus loin, il faudrait étudier cette question en ayant plus d'informations sur les fonctions utilisées par Matlab (bibliothèques BLAS, LAPACK). Pour la suite, on retient que résoudre $\mathbf{L}_\Pi \mathbf{A}_\Pi^\dagger = \mathbf{H}^\dagger$ avec un encodage « creux » est la solution la plus économe en temps de calcul.

1. Calculs effectués sous Matlab avec un processeur i5 à 6 cœurs cadencés à 4.8 GHz.

Conclusion – Afin d'échantillonner les champs d'impédance *a posteriori*, on choisit de s'appuyer sur la forme précision, Eq. (E.6) et Eq. (E.8), qui amène à considérer des matrices très creuses et qui possède la complexité calculatoire la plus faible.

ANNEXE F

RÉSULTATS COMPLÉMENTAIRES

Cette annexe présente des résultats complémentaires à ceux du Chapitre 5.

F.1 Estimation des paramètres d'échelles

On représente l'évolution au fil des itérations des échantillons de γ_1 , γ_2 et γ_ε issus de plusieurs chaînes. On représente sur la Fig. F.1, les échantillons obtenus à partir de la chaîne 2 pour 140 000 itérations, et sur la Fig. F.2, ceux des chaînes 3 et 4 pour 30 000 itérations.

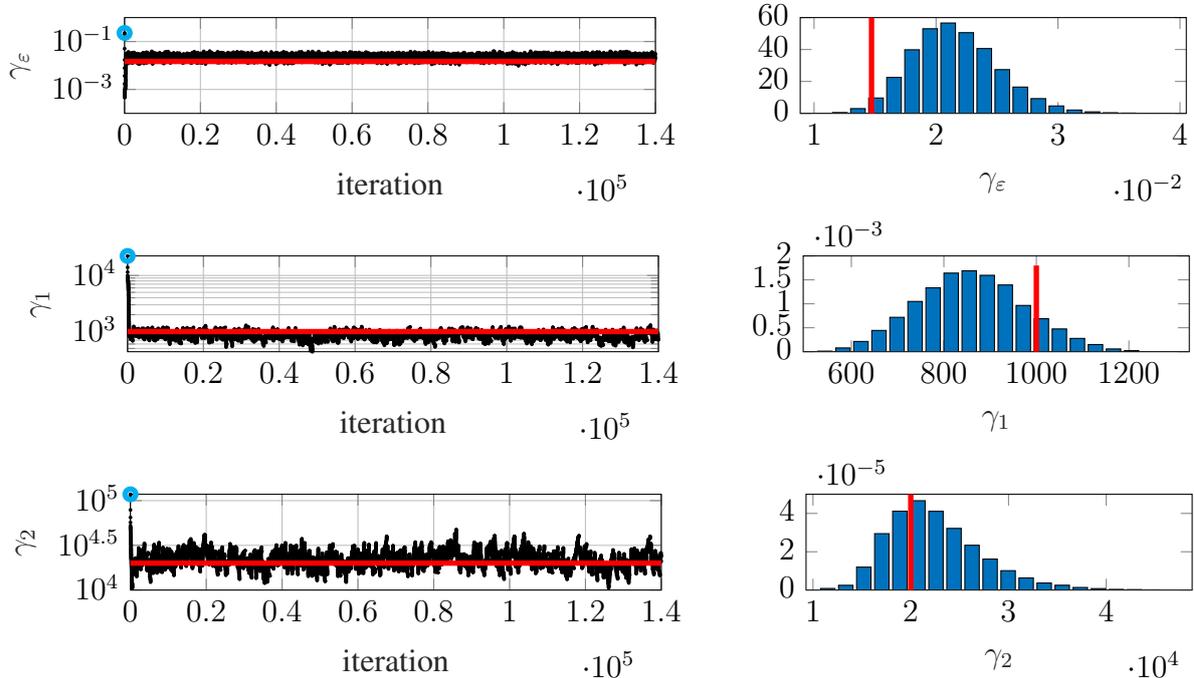


FIGURE F.1 – Réalisations des paramètres d'échelle γ_1 , γ_2 et γ_ε de la chaîne 2 sur 140 000 itérations à gauche et sur les 70 000 dernières itérations pour les histogrammes à droite (première valeur : \bullet , valeur vraie : $-$).

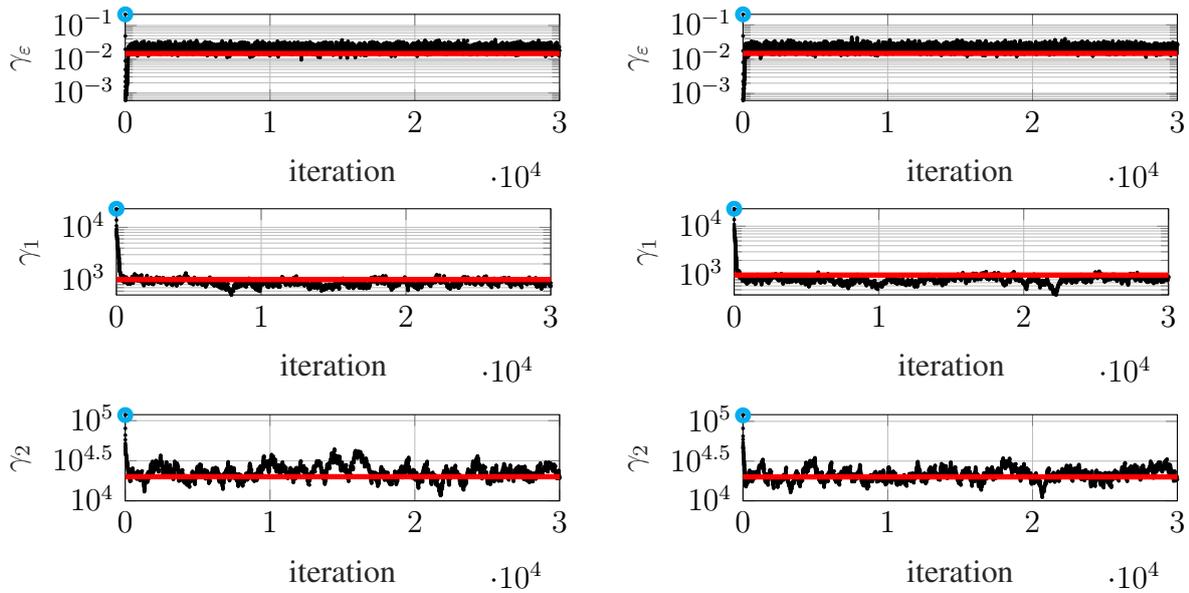


FIGURE F.2 – Réalisations des paramètres d'échelle γ_1 , γ_2 et γ_ε sur 30 000 itérations, pour la chaîne 3 à gauche et pour la chaîne 4 à droite (première valeur : \bullet , valeur vraie : $-$).

F.2 Estimation de la carte des étiquettes

F.2.1 Résultats obtenus pour d'autres chaînes

Dans la section 5.2.2 page 80, on présente les résultats obtenus à partir de la chaîne 1. On les rappelle ici, en complément de ceux obtenus à partir de la chaîne 2, sur la Fig. F.3. On précise également que les résultats de la chaîne 2 sont obtenus dans les mêmes conditions que ceux de la chaîne 1, c'est-à-dire pour le même nombre d'itérations et le même temps de chauffe.

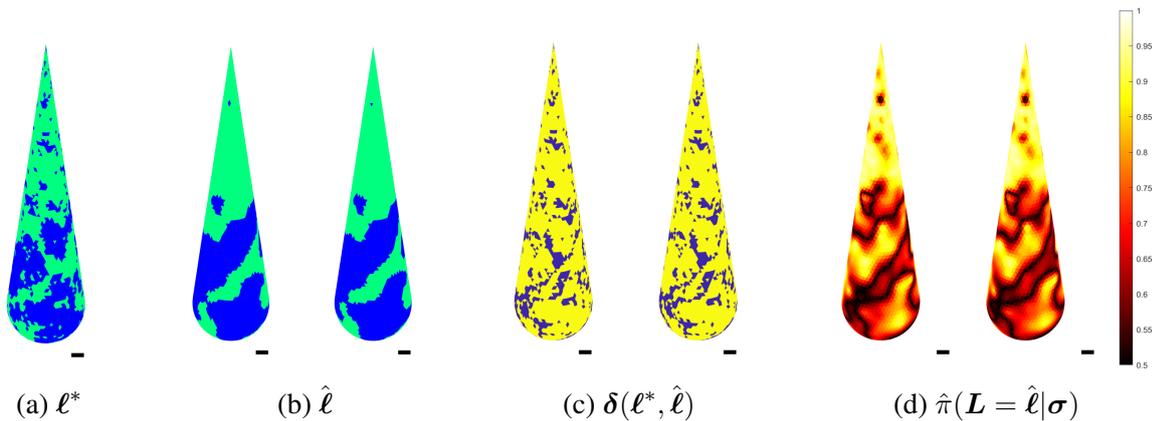


FIGURE F.3 – (a) Étiquettes vraies ℓ^* . (b) Étiquettes estimées $\hat{\ell}$. (c) Carte des erreurs $\delta(\ell^*, \hat{\ell})$ (80% des mailles sont correctement classées). (d) Carte des probabilités $\hat{\pi}(L = \hat{\ell} | \sigma)$. On réalise 140 000 itérations et les estimées sont calculées à partir des 20 000 derniers échantillons. On affiche pour chacune des sous-figures (b), (c) et (d) les quantités correspondantes à la chaîne 1 à gauche et à la chaîne 2 à droite.

F.2.2 Utilisation de chaînes plus courtes

Dans la section 5.2.2 page 81, on mentionne l'utilisation de chaînes plus courtes (30 000 itérations). Les estimées et les probabilités calculées à partir de ces chaînes sont représentées sur la Fig. F.4.

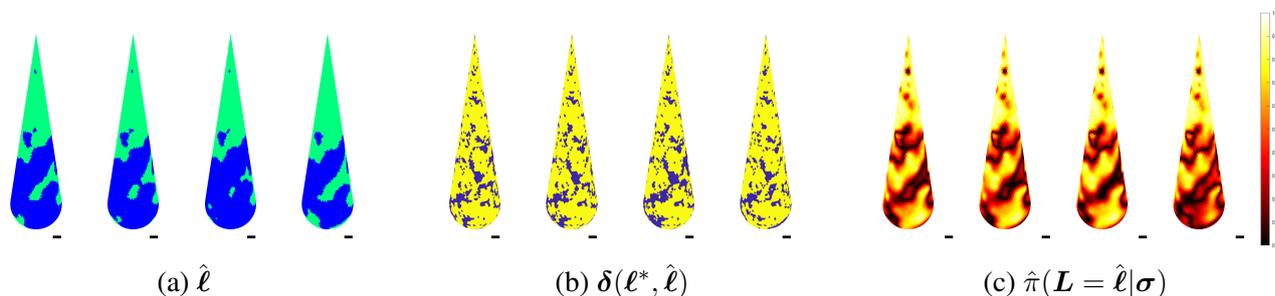


FIGURE F.4 – (a) Étiquettes estimées $\hat{\ell}$. (b) Carte des erreurs $\delta(\ell^*, \hat{\ell})$ (entre 78% et 80% des mailles sont correctement classées). (c) Carte des probabilités $\hat{\pi}(\mathbf{L} = \hat{\ell} | \sigma)$. Résultats issus des chaînes, de gauche à droite, 1, 2, 3 et 4. Les résultats sont obtenus à partir des 20 000 derniers échantillons de chaîne de 30 000 itérations.

F.3 Estimation des impédances complètes

F.3.1 Résultats obtenus pour d'autres chaînes

Dans la partie 5.2.3.1 page 82, on présente les estimées et les incertitudes de la chaîne 1 (140 000 itérations) calculées à partir des 20 000 dernières itérations. On les rappelle sur la Fig. F.5 en plus des résultats de la chaîne 2 obtenus dans les mêmes conditions, c'est-à-dire pour le même nombre d'itérations et le même temps de chauffe. On note que les résultats issus des deux chaînes sont quasi-similaires.

F.3.2 Utilisation de chaînes plus courtes

Les résultats présentés sur la Fig. F.6 sont complémentaires de ceux de la Fig. F.4. Ils sont issus de chaînes longues de 30 000 itérations et les estimées sont calculés à partir des 20 000 derniers échantillons. Comme on le constate, il y a des différences entre les résultats issus de ces chaînes.

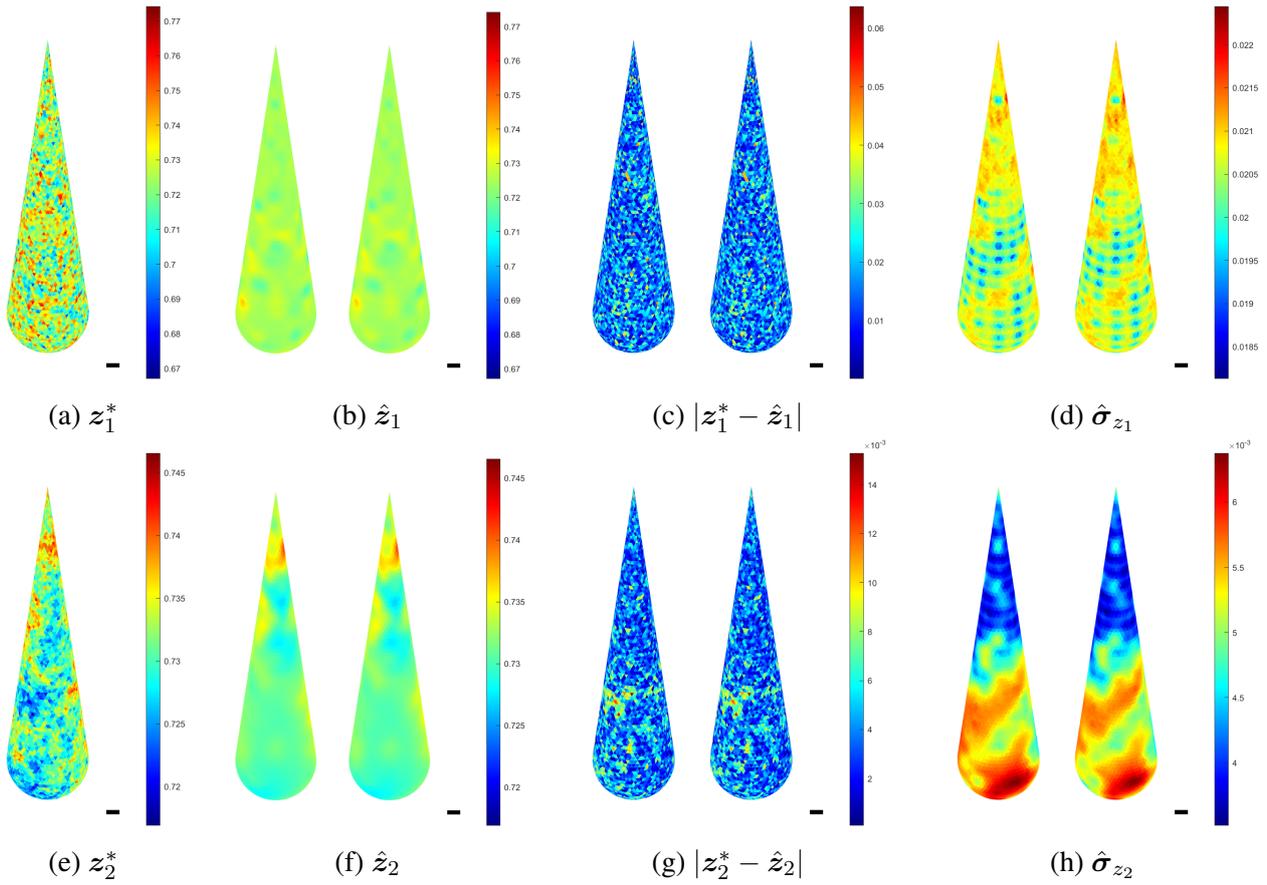


FIGURE F.5 – (a,e) Impédances vraies. (b,f) Impédances estimées. (c,g) Cartes d’erreurs. (d,h) Écart-types. La valeur du module de la référence est $|z_0| = 0.7245$. Pour chaque sous-figure, on représente à gauche la chaîne 1 et à droite la chaîne 2.

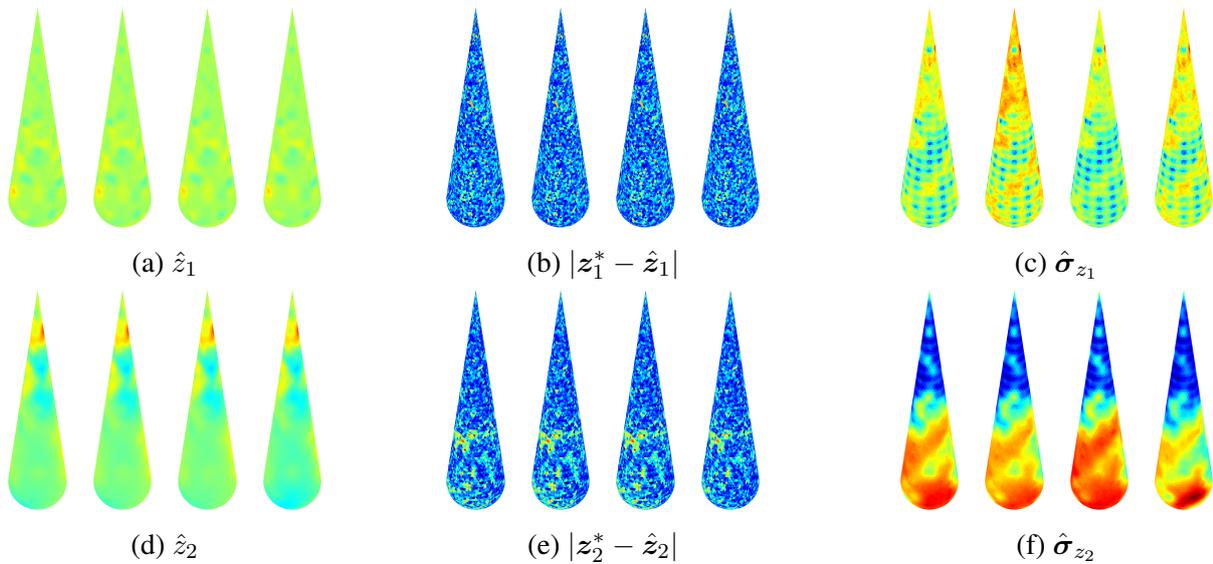


FIGURE F.6 – (a,d) Impédances estimées. (b,e) Cartes d’erreurs. (c,f) Écart-types. Affichage sur chaque sous-figure des résultats issus de la chaîne, de gauche à droite, 1, 2, 3 et 4.

ANNEXE G

**PUBLICATION AU COLLOQUE GRETSI
2019**

Cartographie d'impédance par inversion d'un code d'électromagnétisme 3D surfacique

Julien BOISSY¹⁻², Pierre MINVIELLE¹, Muriel SESQUES¹, Jean-François GIOVANNELLI²

¹CEA, CESTA, DAM, 15 Avenue des Sablières, 33114 Le Barp, France

²IMS (Univ. Bordeaux, CNRS, BINP), 351 Cours de la Libération, 33400 Talence, France

Julien.Boissy@cea.fr, Pierre.Minvielle@cea.fr, Muriel.Sesques@cea.fr, Giova@IMS-Bordeaux.fr

Résumé – Cet article est dédié au problème de la reconstruction d'impédance surfacique d'un objet tridimensionnel. Il se base sur la mesure de coefficients de rétrodiffusion par un radar monostatique. Il s'agit d'un problème qui nécessite la résolution coûteuse des équations de Maxwell. Dans un premier temps on présente comment, en s'appuyant sur la linéarisation des équations de Maxwell sous forme intégrale au sein d'un code de calcul d'électromagnétisme 3D, il est possible de se ramener à un modèle direct. Puis, on propose deux méthodes de reconstruction fondées sur la minimisation de critères pénalisés convexes ; elles sont validées sur des données simulées.

Abstract – This article is dedicated to the problem of reconstructing the surface impedance of a three-dimensional object. It is based on the measurement of backscattering coefficients by a monostatic radar. This problem requires the costly resolution of Maxwell's equations. First we present how it is possible to reduce it to a direct model, based on the linearization of Maxwell's equations in integral form within a 3D electromagnetism numerical code. Then, two reconstruction methods are proposed based on the minimization of convex penalized criteria ; they are validated on simulated data.

1 Introduction

La reconstruction de propriétés matériaux est un enjeu majeur dans de nombreux domaines (imagerie médicale, acoustique, *etc.*). Dans ce contexte, on souhaite remonter aux propriétés électromagnétiques d'un objet sous la forme d'impédance surfacique en exploitant des mesures de coefficients de rétrodiffusion. Pour cela, on s'appuie sur un code d'électromagnétisme 3D résolvant les équations de Maxwell par la méthode des « équations intégrales ». Cependant le problème est en grande dimension, le calcul direct est alors coûteux en temps de calcul. Dans une première partie, on montre comment il est possible de s'appuyer sur un modèle direct linéarisé que l'on désire inverser [1, 2]. Pour autant, le nombre d'inconnues est grandement supérieur au nombre de données observées. Il existe donc une infinité de solutions vérifiant le système. Pour surmonter cette difficulté, on développe dans une seconde partie des méthodes de reconstruction fondées sur les moindres carrés avec pénalités quadratique et Huber [1, 2].

2 Problématique

Le problème revient à déterminer, à partir de mesures de Surface Équivalente Radar (SER) monostatique, émetteur et récepteur co-localisés, les propriétés électromagnétiques locales des matériaux recouvrant un objet de géométrie connue. Le principe de mesure de SER est le suivant : l'objet d'intérêt est placé à l'intérieur d'une chambre anéchoïque (recouverte d'ab-

sorbants radar) où les échos parasites sont réduits. L'illumination en champ lointain (CL) est produite par un radar émettant une onde monochromatique. À la réception, diverses techniques (calibration, substitution de l'objet, filtrage, *etc.*) sont appliquées pour conduire à la mesure calibrée de SER. Finalement, on mesure les coefficients de rétrodiffusion complexes $\{\sigma_1, \dots, \sigma_K\}$ quantifiant l'atténuation de l'onde rétrodiffusée ainsi que son déphasage, pour les conditions d'observation $\{\Psi_1, \dots, \Psi_K\}$ illustrées sur la figure 1. Elles correspondent à différentes directions d'illumination de la cible et à différentes polarisations du champ électromagnétique, HH (horizontale) ou VV (verticale), à l'émission et à la réception.

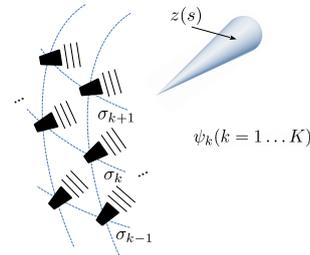


FIGURE 1: Scénario d'observation.

On modélise l'empilement de matériaux par une impédance surfacique sous les *conditions de Leontovich* [3]. Cette approximation n'est valide que si les matériaux recouvrant l'objet sont d'indices suffisamment élevés. L'objectif est de déterminer l'im-

pédance $z(s)$, quantité complexe à la surface de l'objet, à partir des coefficients de rétrodiffusion observés.

3 Modèle direct

3.1 Code Maxwell 3D

À impédance de surface donnée, on sait calculer les coefficients de rétrodiffusion σ_k à l'aide du code ARLENE, développé au CEA-DAM. Fondé sur la résolution du système de Maxwell sous forme intégrale en régime harmonique, ce code d'électromagnétisme 3D permet de traiter des assemblages de matériaux diélectriques et de conducteurs, et tient compte des invariances géométriques. Ce code massivement parallèle exploite efficacement plusieurs milliers de cœurs du supercalculateur TERA1000 [4]. Les équations intégrales sont résolues par une méthode d'éléments finis de surface sous forme standard de Raviart-Thomas [5]. Lorsque l'empilement de matériaux est modélisé par une impédance surfacique [3], la formulation variationnelle associée à la résolution des équations de Maxwell sous forme intégrale est symétrique, mais non hermitienne [6]. On note $z \in \mathbb{C}^L$ l'impédance discrétisée à la surface de l'objet, L étant le nombre de mailles, et $\mathbf{x}_k \in \mathbb{C}^N$ les flux des courants électriques et magnétiques à travers les arêtes du maillage 3D. À fréquence f fixée et direction Ψ_k donnée, le système discret qui en résulte se met sous la forme :

$$\mathbf{A}(z)\mathbf{x}_k = \mathbf{b}_k \quad (1)$$

où $\mathbf{A}(z) \in \mathbb{C}^{N \times N}$ est la matrice du système qui dépend de z , et \mathbf{b}_k le second membre fonction de la direction d'éclairement Ψ_k . Pour $\mathbf{x}_k(z)$ solution de (1), on a directement l'intégrale du rayonnement : $\sigma_k = \mathbf{x}_k^t(z)\mathbf{c}$, où \mathbf{c} est l'opérateur d'observation en réception. Considérant toutes les directions, le système s'écrit :

$$\mathbf{A}(z)\mathbf{X} = \mathbf{B} \quad (2)$$

avec $\mathbf{X} = [\mathbf{x}_1, \dots, \mathbf{x}_K]$ et $\mathbf{B} = [\mathbf{b}_1, \dots, \mathbf{b}_K]$. Pour $\mathbf{X}(z)$ solution de (2), on obtient alors la relation non-linéaire $\boldsymbol{\sigma} = \mathbf{h}(z) = \mathbf{X}(z)^t \mathbf{c}$. La matrice $\mathbf{A}(z)$ est pleine et la taille du système croît de façon quadratique avec la fréquence. L'emploi d'un solveur direct de type Cholesky est privilégié par rapport à un solveur itératif en raison du grand nombre de seconds membres, ainsi que pour des raisons de précision dû au mauvais conditionnement de la matrice $\mathbf{A}(z)$.

Remarque : Lorsque le problème atteint des centaines de milliers d'inconnues, l'emploi d'un solveur direct devient irréaliste car il exige des ressources mémoire en $O(N^2)$ et un nombre d'opérations en $O(N^3)$. On utilise alors la méthode des H-matrices [4], conduisant à un stockage compact ou compressé de la matrice initiale. Ainsi, le coût d'une factorisation de Cholesky d'une H-matrice est en $O(N \log_2(N))$ contre $O(N^3)$ pour une matrice dense.

3.2 Linéarisation intrusive

On considère l'approximation d'ordre 1 autour d'une impédance de surface z_0 : $\mathbf{h}(z) \approx \mathbf{h}(z_0) + \mathbf{H}(z - z_0)$. On a :

$\mathbf{h}(z_0) = \mathbf{X}(z_0)^t \mathbf{c}$, avec $\mathbf{X}(z_0)$ solution du système linéaire $\mathbf{A}(z_0)\mathbf{X}(z_0) = \mathbf{B}$, et \mathbf{H} la matrice jacobienne en z_0 :

$$\mathbf{H} = \left. \frac{\partial \mathbf{h}}{\partial z} \right|_{z=z_0} = \left. \frac{\partial \mathbf{X}}{\partial z} \right|_{z=z_0}^t \mathbf{c}$$

Par dérivation de (2) relativement à la l -ième composante de z , on obtient :

$$\mathbf{A}(z) \frac{\partial \mathbf{X}}{\partial z_l} + \frac{\partial \mathbf{A}}{\partial z_l} \mathbf{X}(z) = 0, \text{ avec } \frac{\partial \mathbf{A}}{\partial z_l} \triangleq \left[\frac{\partial \mathbf{a}_{ij}}{\partial z_l} \right]_{i,j=1 \dots N}$$

En z_0 , on a donc le système linéaire :

$$\mathbf{A}(z_0) \left. \frac{\partial \mathbf{X}}{\partial z_l} \right|_{z=z_0} = - \left. \frac{\partial \mathbf{A}}{\partial z_l} \right|_{z=z_0} \mathbf{X}(z_0) \quad (3)$$

Ainsi, en plus de la résolution standard qui forme $\mathbf{A}(z_0)$ et détermine $\mathbf{X}(z_0)$ puis $\boldsymbol{\sigma}_0$, pour obtenir \mathbf{H} il faut résoudre les L systèmes supplémentaires (3) qui diffèrent par leurs seconds membres. On exploite alors le fait que l'impédance intervient dans la formulation variationnelle [6] uniquement sous la forme d'une intégrale simple du type $\int_S z J \cdot J' + z^{-1} M \cdot M' ds$, où les courants électriques J et magnétiques M à la surface de l'objet sont les inconnues directes du système de Maxwell tandis que J' et M' sont les fonctions tests de la méthode de Galerkin pour la résolution. La dérivation de la matrice \mathbf{A} par rapport à l'impédance conduit ainsi à des matrices $\partial \mathbf{A} / \partial z_k$ très creuses, qui sont stockées de façon optimisée pour effectuer efficacement les produits matrices vecteurs.

Finalement, on obtient le modèle direct linéarisé bruité suivant :

$$\boldsymbol{\sigma} = \boldsymbol{\sigma}_0 + \mathbf{H}(z - z_0) + \boldsymbol{\epsilon} \quad (4)$$

où $\boldsymbol{\epsilon}$ est un terme d'erreur d'approximation et de mesure. En posant $\tilde{\boldsymbol{\sigma}} = \boldsymbol{\sigma} - \boldsymbol{\sigma}_0 + \mathbf{H}z_0$, le modèle devient :

$$\tilde{\boldsymbol{\sigma}} = \mathbf{H}z + \boldsymbol{\epsilon} \quad (5)$$

4 Reconstruction

Dans le but de fournir une solution, on pose le critère des moindres carrés, $J_{LS}(z)$, qui s'appuie sur le modèle direct (5) :

$$J_{LS}(z) = \|\tilde{\boldsymbol{\sigma}} - \mathbf{H}z\|_2^2 \quad (6)$$

Le système est sous-déterminé, il n'existe donc pas de solution unique minimisant le critère. Afin d'orienter la reconstruction en tenant compte des informations disponibles sur l'objet, on introduit une pénalisation $P(z)$. On note μ le paramètre pondérant le terme d'attache aux données (6) et la pénalité. Le critère devient :

$$J_{PLS}(z) = J_{LS}(z) + \mu P(z) \quad (7)$$

et la solution :

$$\hat{z}_{PLS} = \arg \min_z J_{PLS}(z) \quad (8)$$

On considère deux cas. Le premier où l'impédance est régulière sur l'ensemble de la surface et le second où elle est régulière par parties. Pour inclure ces informations, on pénalise le gradient de l'impédance sur la surface :

$$P(z) = \sum_{p \sim q} \varphi(\delta_{pq}) \quad \text{avec } \delta_{pq} = \frac{z_p - z_q}{d_{pq}}$$

et φ un potentiel que l'on définit par la suite. On note d_{pq} la distance entre les barycentres des mailles p et q ($p \sim q$ signifie que les mailles p et q sont voisines). De plus, on note δ le vecteur des δ_{pq} et il peut s'exprimer : $\delta = \mathbf{D}\mathbf{z}$, avec \mathbf{D} la matrice des différences pondérées :

$$\mathbf{D} = \begin{bmatrix} \vdots & \vdots & \vdots \\ \dots & -1/d_{pq} & \dots & 1/d_{pq} & \dots \\ \vdots & \vdots & \vdots \end{bmatrix}$$

Par la suite, on s'intéresse aux pénalités quadratique et Huber. On note φ_Q , φ_H et J_Q , J_H les potentiels et critères correspondant, ainsi que $\hat{\mathbf{z}}_Q$ et $\hat{\mathbf{z}}_H$ les solutions associées.

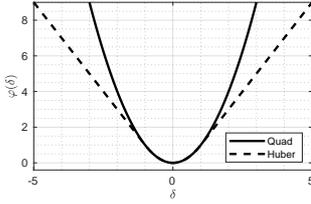


FIGURE 2: Fonctions de pénalité Quadratique et Huber.

4.1 Cas lisse et pénalité quadratique

Afin de favoriser des solutions homogènes et régulières sur l'ensemble la surface, on choisit le potentiel quadratique $\varphi_Q(\delta) = |\delta|^2$, voir la figure 2. La minimisation du critère mène à la solution classique :

$$\hat{\mathbf{z}}_Q = (\mathbf{H}^\dagger \mathbf{H} + \mu \mathbf{D}^\dagger \mathbf{D})^{-1} \mathbf{H}^\dagger \tilde{\sigma} \quad (9)$$

4.2 Cas lisse par parties et pénalité Huber

On cherche à favoriser des solutions régulières, mais également à préserver des discontinuités. De plus, on souhaite que le critère soit convexe afin de s'assurer de l'existence et de l'unicité de la solution et de permettre l'exploitation des nombreuses méthodes de minimisation disponibles. Pour cela, on choisit le potentiel de Huber [2] appartenant à la famille des potentiels L2/L1 (*hyperbole, log cosinus hyperbolique, etc.*), voir la figure 2 :

$$\varphi_H(\delta) = \begin{cases} |\delta|^2 & |\delta| \leq T \\ 2T|\delta| - T^2 & |\delta| \geq T \end{cases} \quad (10)$$

Afin de minimiser le nouveau critère, on choisit de s'appuyer sur des algorithmes semi-quadratiques, dont la particularité consiste en l'introduction de variables auxiliaires, notées \mathbf{a} . En effet, ces algorithmes permettent de passer du problème de minimisation d'un critère non-quadratique, $J_{PLS}(\mathbf{z})$, à la minimisation d'un critère augmenté $J_{SQ}(\mathbf{z}, \mathbf{a})$, tel que :

$$J_H(\mathbf{z}) = \inf_{\mathbf{a}} J_{SQ}(\mathbf{z}, \mathbf{a})$$

où $J_{SQ}(\mathbf{z}, \mathbf{a})$ est quadratique selon \mathbf{z} et $\arg \min_{\mathbf{a}} J_{SQ}(\mathbf{z}, \mathbf{a})$ est séparable et explicite. Le critère augmenté s'écrit [2] :

$$J_{SQ}(\mathbf{z}, \mathbf{a}) = \|\tilde{\sigma} - \mathbf{H}\mathbf{z}\|_2^2 + \mu' \left[\frac{1}{2} \|\mathbf{D}\mathbf{z} - \mathbf{a}\|_2^2 + \sum_{p \sim q} \zeta_\alpha(a_{pq}) \right]$$

avec $\zeta_\alpha(a) = \sup_{\delta} [\alpha \varphi_H(\delta) - \frac{1}{2} |\delta - a|^2]$ et $\mu' = \mu/\alpha$, $\alpha \in]0, 1/2[$. La minimisation selon \mathbf{z} pour un \mathbf{a} fixé conduit à :

$$\hat{\mathbf{z}}(\mathbf{a}) = (\mathbf{H}^\dagger \mathbf{H} + \bar{\mu} \mathbf{D}^\dagger \mathbf{D})^{-1} (\mathbf{H}^\dagger \tilde{\sigma} + \bar{\mu} \mathbf{D}^\dagger \mathbf{a}) \quad (11)$$

où $\bar{\mu} = \mu/2\alpha$. Pour un \mathbf{z} fixé, la minimisation selon \mathbf{a} est explicite et séparable telle que :

$$\hat{\mathbf{a}}(\mathbf{z}) = \delta - \alpha \varphi'_H(\delta) \quad (12)$$

Le minimiseur $\hat{\mathbf{z}}_H$ est obtenu à convergence d'un algorithme itératif minimisant tour à tour le critère suivant \mathbf{a} et \mathbf{z} .

Algorithme Semi-Quadratique

Début

Initialisation $\mathbf{a}^{[0]} = \mathbf{0}$

Tant que $\|\mathbf{z}^{[n]} - \mathbf{z}^{[n-1]}\| > s$ faire

- $\mathbf{z}^{[n]}(\mathbf{a}^{[n-1]}) = \arg \min_{\mathbf{z}} J_{SQ}(\mathbf{z}, \mathbf{a}^{[n-1]})$, Eq. (11)

- $\mathbf{a}^{[n]}(\mathbf{z}^{[n]}) = \arg \min_{\mathbf{a}} J_{SQ}(\mathbf{z}^{[n]}, \mathbf{a})$, Eq. (12)

Fin tant que

Fin

Remarques :

- Le coefficient α permet de régler la vitesse de convergence sans conséquence sur la solution. La convergence est connue pour être plus rapide si $\alpha \rightarrow 1/2$ [2].
- Si $T \rightarrow \infty$, alors la solution $\hat{\mathbf{z}}_H$ coïncide avec $\hat{\mathbf{z}}_Q$.
- Les solutions (9) et (11) sont calculées par résolution des systèmes linéaires correspondant sous Matlab.

5 Résultats

L'objet considéré est une *sphère cône IEEE* [7] de 60 cm de hauteur. Il est illuminé par une onde monochromatique à la fréquence de 1.5 GHz suivant $K = 900$ conditions d'observation définies par la polarisation (HH ou VV), le gisement φ ($0^\circ : 10^\circ : 170^\circ$) et le site θ ($-60^\circ : 5^\circ : 60^\circ$), voir la figure 3.

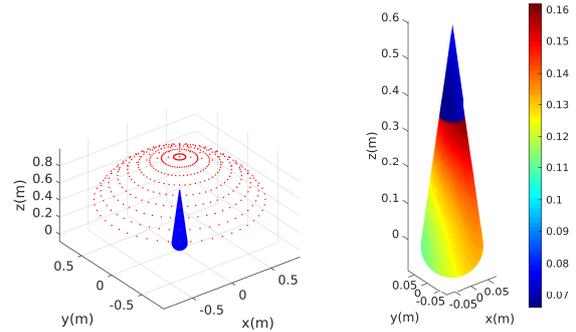


FIGURE 3: Directions d'observation en CL à gauche et carte d'impédance vraie $|z^*|$ à droite.

5.1 Données simulées

Pour simuler les données σ , on considère la surface approchée par un maillage de $L = 7400$ éléments et on prend pour impédance de référence $\mathbf{z}_0 = [0.1, \dots, 0.1]^t$ au centre du domaine d'intérêt. Les impédances considérées sont normalisées

par l'impédance du vide (en Ohms). En exploitant la partie 3.2, on obtient σ_0 et \mathbf{H} utilisés pour la simulation et la reconstruction. On se donne une carte d'impédance vraie z^* lisse par parties, voir la figure 3. Les données sont alors simulées en utilisant le modèle direct (4), où l'on choisit ϵ comme un bruit gaussien complexe circulaire avec un RSB de 20dB entre $\mathbf{H}(z - z_0)$ et ϵ . On représente les données sous forme d'hologrammes en fonction des angles θ et φ sur la figure 4.

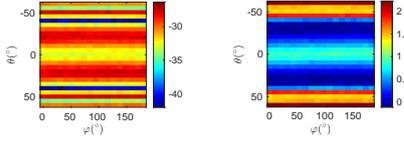


FIGURE 4: Données simulées σ (polarisation HH). Représentation en log du module à gauche et en phase à droite.

5.2 Reconstruction

On représente les reconstructions de z et les cartes d'erreurs absolues sur la figure 5. Dans les deux cas, les hyperparamètres sont réglés manuellement. On prend $\mu = 5.10^{-8}$ pour les deux reconstructions et $T = 0.3$ pour le cas Huber. La solution avec pénalité quadratique \hat{z}_Q fournit une reconstruction régulière sur la totalité de l'objet, respectant le niveau moyen ainsi que les zones de fortes et faibles valeurs d'impédances. Cependant, comme attendu pour cette solution, la discontinuité est mal reconstruite. Cela est dû au choix de l'hyperparamètre μ . Plus il est grand et plus la solution est lisse et inversement, plus il est petit et plus la solution est chahutée. Le potentiel quadratique pénalisant de manière équivalente l'ensemble des différences, il en résulte une incompatibilité entre la valeur du μ permettant une reconstruction de qualité des zones lisses et celle autorisant la reconstruction d'une discontinuité. Pour dépasser cette limitation, on privilégie alors la solution \hat{z}_H fondée sur le potentiel de Huber, voir la figure 5, qui accorde un degré de liberté supplémentaire en autorisant les fortes variations en faible nombre au-dessus d'un seuil T . La solution obtenue est de qualité, non seulement les niveaux d'impédance sont respectés, mais la discontinuité est plus correctement reconstruite.

6 Conclusion et perspectives

La reconstruction de l'impédance surfacique repose sur la linéarisation des équations Maxwell sous forme intégrale au sein d'un code d'électromagnétisme 3D surfacique, ainsi que sur la mise en œuvre de méthodes de régularisation fondées sur les pénalités quadratique et Huber. Les résultats obtenus sont satisfaisants pour des cas lisses et lisses par morceaux. Pour autant, afin de reconstruire au mieux les discontinuités, on envisage des potentiels type L2/L0 [2] malgré le fait que la non-convexité et la non-différentiabilité de ces potentiels amènent à un problème d'optimisation difficile. De plus, on envisage le réglage des hyperparamètres de manière non supervisée au moyen de méthodes bayésiennes.

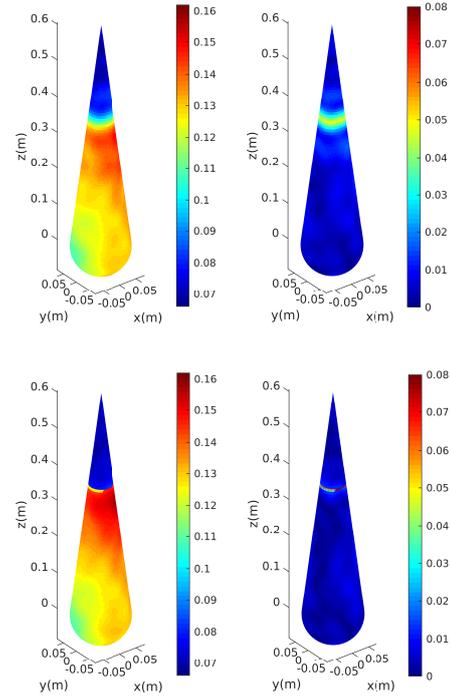


FIGURE 5: Reconstruction $|\hat{z}_Q|$ et carte d'erreur $|\hat{z}_Q - z^*|$ en haut et reconstruction $|\hat{z}_H|$ et carte d'erreur $|\hat{z}_H - z^*|$ en bas.

Références

- [1] A. Mohammad - Djafari. *Problèmes inverses en imagerie et en vision Tome 1 et 2*. IC2 Traité du Signal et Image, Lavoisier, Hermès, Paris, 2009.
- [2] J. Idier. *Approche bayésienne pour les problèmes inverses*. IC2 Traité du Signal et Image, Lavoisier, Hermès, Paris, 2001.
- [3] S. Tretyakov. *Analytical Modeling in Applied Electromagnetics*. Chapitre 3, Artech house 2003.
- [4] C. Augonnet, A. Pujols, M. Sesques. Fast Direct Solver for Integral Equations on Massively Parallel Architecture, *2015 Antennas and Propagation Society International Symposium*. Vancouver, Canada, Juillet 2015.
- [5] P.-A. Raviart, J.M. Thomas. A mixed finite element method for 2nd order elliptic problems, *Mathematical aspects of finite element methods* Springer, Berlin, pp. 292-315, 1977.
- [6] A. Bendali, M'B. Fares, J. Gay. A Boundary-Element Solution of the Leontovitch problem. *IEEE transaction on Antennas and Propagation*, Volume : 47, Issue : 10, pp 1597-1605, 1999.
- [7] A.C. Woo, H.T.G. Wang, M.J. Schuh, M.L. Sanders. EM programmer's notebook-benchmark radar targets for the validation of computational electromagnetics programs. *IEEE Antennas and Propagation Magazine*, Volume : 35, pp 84-89, Février 1993.

ANNEXE H

**PUBLICATION DANS LA REVUE IEEE
SIGNAL PROCESSING LETTERS**

An Insight Into the Gibbs Sampler: Keep the Samples or Drop Them?

Julien Boissy, Jean-François Giovannelli, and Pierre Minvielle 

Abstract—In this letter, we propose an insight into Markov Chain Monte Carlo (MCMC) algorithms and more precisely the Gibbs sampler. From a didactic toy model, based on a normal bivariate distribution, a non-asymptotic analysis is derived and estimators are fully characterized. It provides a worthwhile and non-empirical understanding of the Gibbs sampler behaviour. Issues are investigated, such as the influence of the “burn-in” phase, useful in practice. Especially, the trade-off between discarding samples and integrating them into estimators is studied. On the whole, it leads to an analytical awareness of MCMC sampler.

Index Terms—Bayesian statistics, burn-in, gibbs, MCMC.

I. INTRODUCTION

THE Gibbs sampler is an ubiquitous MCMC algorithm in Bayesian statistics and machine learning [1], [2]; it was introduced in the context of image processing by [3]. Especially applied to hierarchical models, the Gibbs sampler is, as well as the Metropolis-Hastings (MH) sampler of which it is a special case, the workhorse of general MCMC techniques. Hereafter, we provide an analytical insight into the standard Gibbs algorithm and its mechanisms.

We consider a couple of real random variables (X, Y) of joint probability density function $f_{X,Y}$. Let us suppose that though it is difficult to sample from it directly, realisations can easily be obtained from the conditional distributions $f_{X|Y}$ and $f_{Y|X}$. The elementary two-stage Gibbs sampler, see below, is appropriate in such a context [1]. It generates a realisation of the Markov chain $(X^{(n)}, Y^{(n)})_{n \geq 1}$ that has the target distribution $f_{X,Y}$ as invariant distribution. It is typically used to solve intractable integration problems [2], central to Bayesian statistics, such as expectation $E_{f_{X,Y}}[h(X, Y)] = \iint h(x, y) f_{X,Y}(x, y) dx dy$ or marginalization $f_X(x) = \int f_{X,Y}(x, y) dy$.

The algorithm is parametrized by N , the chain run length and N_0 , the so-called “burn-in” period during which samples are assumed not to be samples of the target distribution and discarded in order to reduce starting biases [2]. The initial value of Y is randomly drawn from the density $f_{Y^{(0)}}$. At the end, the Gibbs output is made of the samples: $(x^{(n)}, y^{(n)})_{N_0 \leq n \leq N}$.

Manuscript received October 27, 2020; revised November 11, 2020; accepted November 11, 2020. Date of publication November 16, 2020; date of current version December 7, 2020. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Arash Mohammadi. (Corresponding author: Pierre Minvielle.)

Julien Boissy and Pierre Minvielle are with CEA CESTA DAM, F-33114 Le Barp, France (e-mail: julien.boissy@cea.fr; pierre.minvielle@cea.fr).

Jean-François Giovannelli is with IMS (Univ. Bordeaux, CNRS, Bordeaux INP), F-33400 Talence, France (e-mail: jean-francois.giovannelli@u-bordeaux.fr).

Digital Object Identifier 10.1109/LSP.2020.3038136

Algorithm 1: Two-Stage Gibbs Sampler.

Input: $N, N_0, f_{Y^{(0)}}$

$y^{(0)} \sim f_{Y^{(0)}}(\cdot)$

for $n = 1 \cdots N$

$x^{(n)} \sim f_{X|Y}(\cdot|y^{(n-1)})$

$y^{(n)} \sim f_{Y|X}(\cdot|x^{(n)})$

end for

Output: samples $(x^{(n)}, y^{(n)})_{N_0 \leq n \leq N}$

The determination of the appropriate lengths N and N_0 is known to be a tricky task. How long must the “burn-in” period N_0 be chosen? How many samples $\Delta N \triangleq N - N_0 + 1$ are required to accurately approximate for instance $E_{f_{X,Y}}[h(X, Y)]$? For a fixed N , the higher N_0 is, the “better” the samples are, but the fewer there are. What is the optimal trade-off between the sample length and the sample quality? Conversely, for a fixed N_0 , the higher N is, the more accurate the approximation is, but the longer the Gibbs sampler is.

Similarly to the seminal work by Gelman *et al.* [4] for the Metropolis-Hastings algorithm, its optimal scale factor and its known acceptance rate of 0.234, we intend to examine critical parameters of the Gibbs sampler, such as the “burn-in” length. As extensively done in [4], our exploration makes use of a toy model; it is of great interest to investigate such intricate issues. Although it can be related to basic examples in [5, p. 131], [1, p. 340] and [6], we do believe that we provide an attractive investigation into MCMC algorithms and the “burn-in” issue. We consider for instance the following straightforward MCMC estimate or approximation of $E[Y]$, e.g. the posterior mean in the Bayesian setting; it consists of the empirical mean of the samples $(y^{(n)})_{N_0 \leq n \leq N}$ from the Gibbs output:

$$E[Y] \approx \frac{1}{\Delta N} \sum_{n=N_0}^N y^{(n)}. \quad (1)$$

Our main contribution is therefore to develop closed-form expressions for the bias and the variance of related estimators, from which efficient and meaningful analyses can be performed. It leads to a new study of the influence of the “burn-in” length N_0 . It differs from former works about Gibbs sampling: [7] is based on subsampling and stationarity considerations, [8]–[11] propose improvements where numerical experiments are developed for performance assessment and [12] uses bivariate Gaussian toy examples again for numerical experiments. Even

more recent works, such as [12]–[15], do not develop such a theoretical understanding of “burn-in”.

II. NORMAL BIVARIATE TOYMODEL

We focus on a Gaussian bivariate toy model: $(X, Y) \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, with mean $\boldsymbol{\mu} = [\mu_X \mu_Y]^T$ and covariance matrix

$$\boldsymbol{\Sigma} = \begin{bmatrix} \sigma_X^2 & \rho \cdot \sigma_X \sigma_Y \\ \rho \cdot \sigma_X \sigma_Y & \sigma_Y^2 \end{bmatrix}. \quad (2)$$

Hereinafter, $f_{X,Y}$ represents the target distribution. It is well-known that the conditional laws are normal (see [16, p. 43-44]):

$$X|Y = y \sim \mathcal{N}(\mu_{X|y}, \sigma_{X|y}^2), \quad (3)$$

where $\mu_{X|y} = \mu_X + \rho\sigma_X(y - \mu_Y)/\sigma_Y$ and $\sigma_{X|y}^2 = \sigma_X^2(1 - \rho^2)$, and symmetrically for $Y|X = x$. In addition, the initial distribution is defined by:

$$f_{Y^{(0)}}(y^{(0)}) = \mathcal{N}(y^{(0)}; m_{Y^{(0)}}, \sigma_{Y^{(0)}}^2). \quad (4)$$

The algorithm determines a realisation of the Markov chain: $y^{(0)} \rightarrow x^{(1)} \rightarrow y^{(1)} \rightarrow x^{(2)} \rightarrow y^{(2)} \rightarrow \dots \rightarrow x^{(N)} \rightarrow y^{(N)}$, leading to the successive couples $(x^{(n)}, y^{(n)})_{1 \leq n \leq N}$. It is straightforward to check that $f_{X,Y}$ is a stationary distribution [1]. Next, we determine analytical expressions of the marginal probability densities $f_{X^{(n)}}$ and $f_{Y^{(n)}}$. After, we focus on the MCMC estimator of the expectation, such as $E(Y)$.

A. Elementary AR(1) Model

Let us note $\mathcal{X}^{(n)} \triangleq (X^{(n)} - \mu_X)/\sigma_X$ and $\mathcal{Y}^{(n)} \triangleq (Y^{(n)} - \mu_Y)/\sigma_Y$, it is direct to exhibit an elementary auto-regressive AR(1) model (see toy models [1, p. 340] and [17, p. 9]):

$$\mathcal{Y}^{(n+1)}|\mathcal{Y}^{(n)} = y^{(n)} \sim \mathcal{N}(\rho^2 y^{(n)}, 1 - \rho^4). \quad (5)$$

Noting the expectation as $m_{\mathcal{Y}^{(n)}} \triangleq E[\mathcal{Y}^{(n)}]$ and the variance as $\sigma_{\mathcal{Y}^{(n)}}^2 \triangleq V(\mathcal{Y}^{(n)})$, the recursive expressions are obtained:

$$m_{\mathcal{Y}^{(n+1)}} = \rho^2 m_{\mathcal{Y}^{(n)}}, \quad \sigma_{\mathcal{Y}^{(n+1)}}^2 = 1 + \rho^4(\sigma_{\mathcal{Y}^{(n)}}^2 - 1). \quad (6)$$

B. Closed-Form Expression of Marginal Laws

The expectations and variances of $X^{(n)}$ and $Y^{(n)}$ are noted respectively as $m_{X^{(n)}}$ and $m_{Y^{(n)}}$, $\sigma_{X^{(n)}}^2$ and $\sigma_{Y^{(n)}}^2$. We have:

$$m_{Y^{(n+1)}} - \mu_Y = \sigma_Y m_{\mathcal{Y}^{(n+1)}} = \rho^2(m_{Y^{(n)}} - \mu_Y). \quad (7)$$

The following expectation can then be obtained:

$$m_{X^{(n)}} = \mu_X + \rho^{2n-1} \sigma_X (m_{Y^{(0)}} - \mu_Y)/\sigma_Y \quad (8)$$

$$m_{Y^{(n)}} = \mu_Y + \rho^{2n} (m_{Y^{(0)}} - \mu_Y). \quad (9)$$

Otherwise, $\sigma_{\mathcal{Y}^{(n+1)}}^2 = \rho^4 \sigma_{\mathcal{Y}^{(n)}}^2 + 1 - \rho^4$ directly leads to:

$$\sigma_{Y^{(n+1)}}^2 - \sigma_Y^2 = \rho^4(\sigma_{Y^{(n)}}^2 - \sigma_Y^2). \quad (10)$$

This results in the following variance:

$$\sigma_{X^{(n)}}^2 = \sigma_X^2 + \rho^{4n-2} \sigma_X^2 (\sigma_{Y^{(0)}}^2 - \sigma_Y^2)/\sigma_Y^2 \quad (11)$$

$$\sigma_{Y^{(n)}}^2 = \sigma_Y^2 + \rho^{4n} (\sigma_{Y^{(0)}}^2 - \sigma_Y^2). \quad (12)$$

Similarly, the following covariance can be obtained:

$$\begin{aligned} \text{Cov}(X^{(n)}, Y^{(n)}) &\triangleq E[(X^{(n)} - m_{X^{(n)}})(Y^{(n)} - m_{Y^{(n)}})] \\ &= \rho \sigma_X \sigma_Y \cdot [1 + \rho^{4n+2} (\sigma_{Y^{(0)}}^2/\sigma_Y^2 - 1)]. \end{aligned} \quad (13)$$

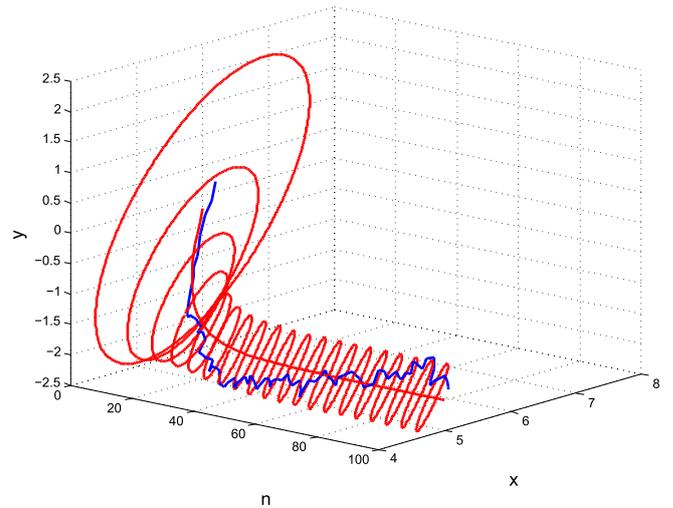


Fig. 1. Marginal distribution convergence to the target distribution ($\rho = 0.95$), as a function of the iteration n . In red, the mean and covariance ellipse every 10 iterations (with 95% confidence). In blue, a realisation of the Gibbs sampler.

Consequently, since all the distributions are normal, Eq. (8) and (11) state the convergence of the marginal normal distributions $f_{X^{(n)}}$ and $f_{Y^{(n)}}$, i.e. their first two moments, respectively towards the invariant target distribution f_X and f_Y , when $n \rightarrow \infty$. It clearly depends on the correlation parameter ρ . The lower $|\rho|$ is, the more decorrelated the components X and Y are and the faster the convergence is. Conversely, the more correlated they are, the slower it converges. It also depends on the initial distribution. Tautologically, if one initiates with the target distribution, the convergence is already reached. Fig. 1 provides an illustration of the convergence of the marginal distributions $f_{X^{(n)}, Y^{(n)}}$ to the target invariant normal distribution, closely linked to the “burn-in” period.

III. PROPERTIES OF THE EXPECTATION ESTIMATOR

We are interested in approximating expectations, such as $\mu_X = E(X)$ or $\mu_Y = E(Y)$. For the latter, we consider the usual MCMC estimate of Eq. (1):

A. Bias Determination

The expectation of \bar{M}_Y is given by:

$$\begin{aligned} E(\bar{M}_Y) &= \frac{1}{\Delta N} \sum_{n=N_0}^N m_{Y^{(n)}} \\ &= \mu_Y + \frac{\sigma_Y}{\Delta N} \sum_{n=N_0}^N m_{\mathcal{Y}^{(n)}} \\ &= \mu_Y + \frac{\sigma_Y}{\Delta N} m_{\mathcal{Y}^{(0)}} \sum_{n=N_0}^N \rho^{2n} \\ &= \mu_Y + \frac{m_{Y^{(0)}} - \mu_Y}{\Delta N} \rho^{2N_0} \frac{1 - \rho^{2\Delta N}}{1 - \rho^2}. \end{aligned} \quad (14)$$

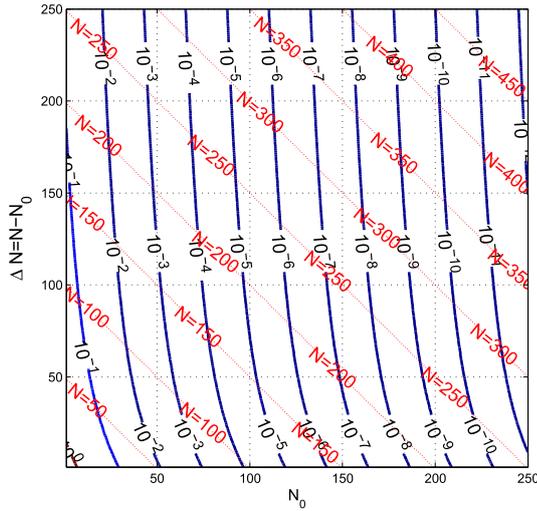


Fig. 2. Bias of \bar{M}_Y ($\rho = 0.95$), function of N_0 and ΔN . The straight isocontours of the sample budget ($N_0 + \Delta N$) are indicated by the red dotted line.

Noting $\alpha(\Delta N, \rho) \triangleq (1 - \rho^{2\Delta N}) / [(1 - \rho^2)\Delta N]$, the bias $B(\bar{M}_Y) \triangleq E(\bar{M}_Y) - \mu_Y$ is then given by:

$$B(\bar{M}_Y) = \alpha(\Delta N, \rho) \cdot \rho^{2N_0} \cdot (m_{Y^{(0)}} - \mu_Y). \quad (15)$$

The bias also depends on the initial mean agreement $m_{Y^{(0)}} - \mu_Y$. Besides, the higher $|\rho|$ is and the more correlated the components are, the stronger the bias is. For a given ρ , it depends on N_0 . A long “burn-in” phase reduces the bias. It also depends on ΔN . One can notice that $\alpha(1, \rho) = 1$ and $\alpha(\Delta N, \rho) \xrightarrow{\Delta N \rightarrow \infty} 0$. Furthermore, consistently with well-known general results [17, p. 21], the bias is $O(\Delta N^{-1})$ for high ΔN (for a given N_0); it will be further shown to be negligible for sufficiently long runs. The bias behaviour is illustrated in Fig. 2. For a given sample budget N (in red), the optimal trade-off between N_0 and ΔN is straightforward. Concerning the bias, it is better to have a long “burn-in” period, discarding many samples and keeping only a few of the latest good ones.

B. Variance Determination

The variance of \bar{M}_Y is given by:

$$\begin{aligned} V(\bar{M}_Y) &= E \left[(\bar{M}_Y - E(\bar{M}_Y))^2 \right] \\ &= \frac{1}{(\Delta N)^2} E \left[\left(\sum_{n=N_0}^{N-1} (Y^{(n)} - m_{Y^{(n)}}) \right)^2 \right] \\ &= \frac{1}{(\Delta N)^2} \sum_{n=N_0}^{N-1} \sum_{p=N_0-n}^{N-n-1} \gamma_Y^{n,p}, \end{aligned} \quad (16)$$

noting the lag- p auto-covariance at time n

$$\gamma_Y^{n,p} \triangleq E \left[(Y^{(n)} - m_{Y^{(n)}})(Y^{(n+p)} - m_{Y^{(n+p)}}) \right]. \quad (17)$$

It can be noticed that:

$$Y^{(n)} - m_{Y^{(n)}} = \sigma_Y \left(\mathcal{Y}^{(n)} - m_{\mathcal{Y}^{(n)}} \right). \quad (18)$$

Remark that $\gamma_Y^{n,p}$ can be rewritten as:

$$\begin{aligned} \gamma_Y^{n,p} &= \sigma_Y^2 E \left[\left(\mathcal{Y}^{(n)} - m_{\mathcal{Y}^{(n)}} \right) \left(\mathcal{Y}^{(n+p)} - m_{\mathcal{Y}^{(n+p)}} \right) \right] \\ &= \sigma_Y^2 \left(E \left[\mathcal{Y}^{(n)} \mathcal{Y}^{(n+p)} \right] - m_{\mathcal{Y}^{(n)}} m_{\mathcal{Y}^{(n+p)}} \right). \end{aligned} \quad (19)$$

For $p \geq 0$, it can be directly shown from Eq. (5) that:

$$E \left[\mathcal{Y}^{(n)} \mathcal{Y}^{(n+p)} \right] = \rho^{2p} E \left[\mathcal{Y}^{(n)} \mathcal{Y}^{(n+p-1)} \right], \quad (20)$$

ending as: $E[\mathcal{Y}^{(n)} \mathcal{Y}^{(n+p)}] = \rho^{2p} E[(\mathcal{Y}^{(n)})^2]$. Then, the auto-covariance recurrence can be written as:

$$\begin{aligned} \gamma_Y^{n,p} &= \sigma_Y^2 \rho^{2p} E \left[(\mathcal{Y}^{(n)})^2 \right] - m_{\mathcal{Y}^{(n)}} m_{\mathcal{Y}^{(n+p)}} \\ &= \sigma_Y^2 \rho^{2p} \left(\rho^{2(p-1)} E \left[(\mathcal{Y}^{(n)})^2 \right] - m_{\mathcal{Y}^{(n)}} m_{\mathcal{Y}^{(n+p-1)}} \right) \\ &= \rho^2 \gamma_Y^{n,p-1}. \end{aligned} \quad (21)$$

For any p , the auto-covariance is asymmetric:

$$\gamma_Y^{n,p} = \rho^{2|p|} \sigma_Y^2 + \rho^{4n+2p} (\sigma_{Y^{(0)}}^2 - \sigma_Y^2) \quad (22)$$

Notice that it tends to be symmetric when $n \rightarrow \infty$.

From Eq. (16), the \bar{M}_Y -variance can now be expressed as:

$$\begin{aligned} V(\bar{M}_Y) &= \frac{1}{(\Delta N)^2} \sum_{n=N_0}^{N-1} \left[\sigma_Y^2 \left(\sum_{p=N_0}^{N-n-1} \rho^{2p} + \sum_{q=1}^{n-N_0} \rho^{2q} \right) \right. \\ &\quad \left. + \rho^{4n} (\sigma_{Y^{(0)}}^2 - \sigma_Y^2) \sum_{p=N_0-n}^{N-n-1} \rho^{2p} \right]. \end{aligned} \quad (23)$$

A few geometric sums latter, it results in:

$$\begin{aligned} V(\bar{M}_Y) &= \delta^2(\Delta N, \rho) \cdot \sigma_Y^2 + \alpha^2(\Delta N, \rho) \\ &\quad \cdot \rho^{4N_0} \cdot (\sigma_{Y^{(0)}}^2 - \sigma_Y^2), \end{aligned} \quad (24)$$

with notations:

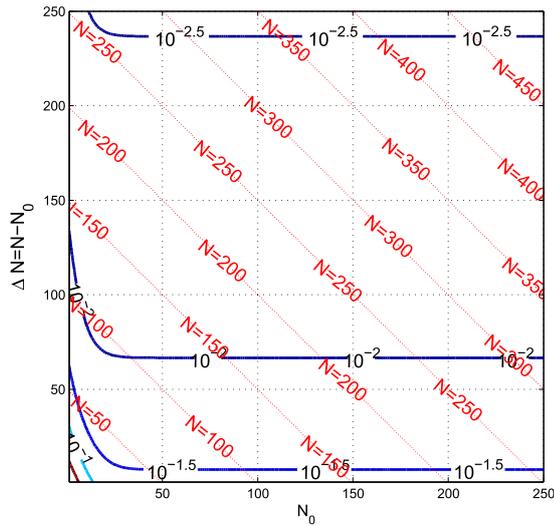
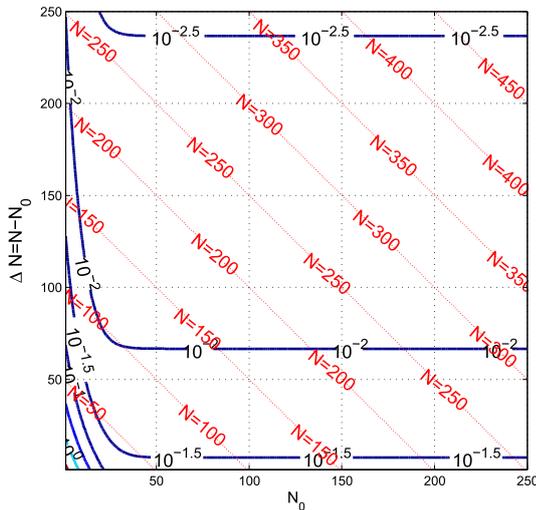
$$\delta^2(\Delta N, \rho) \triangleq \alpha^2(\Delta N, \rho) \beta(\Delta N, \rho) \quad (25)$$

$$\beta(\Delta N, \rho) \triangleq \frac{\Delta N(1 - \rho^4) - 2\rho^2(1 - \rho^{2\Delta N})}{(1 - \rho^{2\Delta N})^2}. \quad (26)$$

The first factor is proportional to the variance σ_Y^2 . The second one is very similar to the bias of Eq. (15), with a dependence on the initial variance agreement $\sigma_{Y^{(0)}}^2 - \sigma_Y^2$. For the special case of $\Delta N = 1$, notice that $\delta^2(1, \rho) = 1$ and $\alpha^2(1, \rho) = 1$. The variance $V(\bar{M}_Y)$ is then determined by N_0 and ρ , its lower bound being σ_Y^2 .

Asymptotically, for high ΔN and a given ρ , it is straightforward to show that: $\delta^2(\Delta N, \rho) = O(\Delta N^{-1})$ and $\alpha^2(\Delta N, \rho) = O(\Delta N^{-2})$. That means that the second term and the bias of Eq. (15) are going to be negligible, compared to the first one. Let us stress that the convergence order $O(\Delta N^{-1})$ of the toy model variance $V(\bar{M}_Y)$ is consistent with Central Limit Theorems for Markov chains [1], [17]. On the other hand, it can be shown that $V(\bar{M}_Y)$ is non-zero when $|\rho| \rightarrow 1$.

The variance behaviour is illustrated in Fig. 3. The convergence order $O(\Delta N^{-1})$ of the variance is confirmed. For a given sample budget N (in red), the optimal trade-off between N_0 and ΔN corresponds to the intersection of the variance level set and the budget line corresponding to N . It turns out to be always around $N_0 = 20$, whatever the value of N is. Concerning the

Fig. 3. Variance of \bar{M}_Y ($\rho = 0.95$).Fig. 4. Mean Squared Error of \bar{M}_Y ($\rho = 0.95$).

variance, it is better to have a short “burn-in” phase, discarding a few tens of samples and keeping all the next ones.

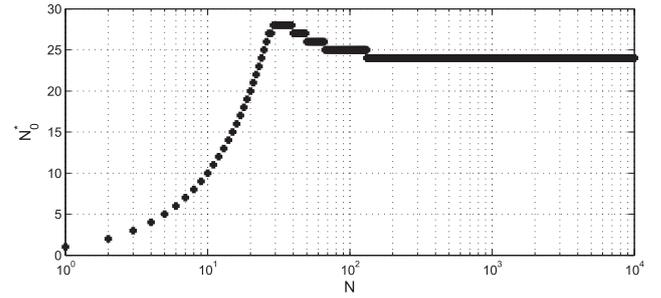
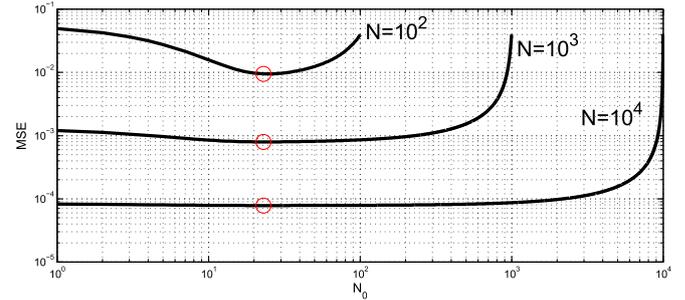
We obtain a closed-form expression of the variance $V(\bar{M}_Y)$ that provides a better comprehension. Let us stress that $V(\bar{M}_Y)$ cannot generally be calculated in closed-form, except for a few small problems; it must be estimated from simulations.

C. Mean Squared Error

The Mean Squared Error (MSE) of \bar{M}_Y is given by:

$$\begin{aligned} \text{MSE}(\bar{M}_Y) &\triangleq V(\bar{M}_Y) + B(\bar{M}_Y)^2 \\ &= \delta^2(\Delta N, \rho) \cdot \sigma_Y^2 + \alpha^2(\Delta N, \rho) \cdot \rho^{4N_0} \\ &\quad \cdot [(\sigma_{Y^{(0)}}^2 - \sigma_Y^2) + (m_{Y^{(0)}} - \mu_Y)^2]. \end{aligned} \quad (27)$$

The second term depends on both the initial mean and variance agreements $(\sigma_{Y^{(0)}}^2 - \sigma_Y^2) + (m_{Y^{(0)}} - \mu_Y)^2$. It must be noticed that: $\text{MSE}(\bar{M}_Y) = \sigma_Y^2 / \Delta N$ if $\rho = 0$. Asymptotically, for high ΔN and a given ρ , it is again straightforward to show that the

Fig. 5. Optimal “burn-in” length N_0^* ($\rho = 0.95$).Fig. 6. MSE function of the “burn-in” length N_0 ($\rho = 0.95$). The optimal values are represented in red.

second term in $O(\Delta N^{-2})$ is negligible, compared to the first one in $O(\Delta N^{-1})$.

The MSE behaviour, illustrated in Fig. 4, is similar to the previous Fig. 3. For a given sample budget N (in red), the trade-off between N_0 and ΔN seems to vary slightly, depending on N . The optimal value N_0^* is represented in Fig. 5, as a function of N . For low N , it is dominated by the bias reduction. It increases linearly up to $N \approx 30$, with the corresponding ΔN being constant and equal to 1. Thus, for low N , it is optimal, according to the MSE, to discard all the samples except the last one. For higher N , the optimal N_0^* slightly decreases and stabilises around 25 for high $N = 10^3$ to 10^6 . In addition, Fig. 6 represents the MSE as a function of N_0 for various values of N . In accordance with [17], the “burn-in” turns out to be useless, i.e. the curve is flat for high N .

IV. CONCLUSION

This didactic investigation ends up with a reinforced understanding of Gibbs sampling and more generally, MCMC algorithms. Our non-asymptotical results are given in the context of a normal bivariate model. The Markov chain estimator is characterised, leading to a non-stationary analysis of the influence of the “burn-in” length. This theoretical work is new, even compared to recent works [11]–[15].

So, keep the samples or drop them? The trade-off is analysed in an original way. Roughly speaking, the answer is to have a long run, in which case the question is useless. Of course, there are limits to the transposition of toy models to real-world problems [17] where there are no closed-form expressions and where the target distribution is obviously not known. And yet, as in [4], it can provide useful guidelines to apply MCMC samplers.

REFERENCES

- [1] C. Robert and G. Casella, *Monte Carlo Statistical Methods*. Berlin, Germany: Springer, 2013.
- [2] C. Andrieu, N. De Freitas, A. Doucet, and M. I. Jordan, "An introduction to MCMC for machine learning," *Mach. Learn.*, vol. 50, nos. 1/2, pp. 5–43, 2003.
- [3] S. Geman and D. Geman, "Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-6, no. 6, pp. 721–741, Nov. 1984.
- [4] A. Gelman *et al.*, "Efficient Metropolis jumping rules," *Bayesian Statist.*, vol. 5, pp. 599–608, 1996.
- [5] J. S. Liu, *Monte Carlo Strategies in Scientific Computing*. Berlin, Germany: Springer, 2008.
- [6] S. Chib and E. Greenberg, "Understanding the Metropolis-Hastings algorithm," *Amer. Statist.*, vol. 49, no. 4, pp. 327–335, 1995.
- [7] S. N. MacEachern and L. M. Berliner, "Subsampling the Gibbs sampler," *Amer. Statist.*, vol. 48, no. 3, pp. 188–190, 1994.
- [8] W. R. Gilks, G. O. Roberts, and E. I. George, "Adaptive direction sampling," *J. Roy. Statist. Soc., Ser. D*, vol. 43, no. 1, pp. 179–189, 1994.
- [9] M.-H. Chen and B. Schmeiser, "Toward black-box sampling: A random-direction interior-point Markov chain approach," *J. Comput. Graph. Statist.*, vol. 7, no. 1, pp. 1–22, 1998.
- [10] W. A. Link and M. J. Eaton, "On thinning of chains in MCMC," *Methods Ecol. Evol.*, vol. 3, no. 1, pp. 112–115, 2012.
- [11] L. Martino, J. Read, and D. Luengo, "Independent doubly adaptive rejection Metropolis sampling within Gibbs sampling," *IEEE Trans. Signal Process.*, vol. 63, no. 12, pp. 3123–3138, Jun. 2015.
- [12] L. Martino, V. Elvira, and G. Camps-Valls, "The recycling Gibbs sampler for efficient learning," *Digit. Signal Process.*, vol. 74, pp. 1–13, 2018.
- [13] C. Fox *et al.*, "Accelerated Gibbs sampling of normal distributions using matrix splittings and polynomials," *Bernoulli*, vol. 23, no. 4B, pp. 3711–3743, 2017.
- [14] A.-C. Barros, F. Caron, J.-F. Giovannelli, and A. Doucet, "Clone MCMC: Parallel high-dimensional Gaussian Gibbs sampling," in *Proc. Adv. Neural Inf. Process. Syst.*, 2017, pp. 5020–5028.
- [15] M. Vono, N. Dobigeon, and P. Chainais, "Split-and-augmented Gibbs sampler—Application to large-scale inference problems," *IEEE Trans. Signal Process.*, vol. 67, no. 6, pp. 1648–1661, Mar. 2019.
- [16] Y. Bar-Shalom and X.-R. Li, "Estimation and tracking: Principles, techniques, and software," *IEEE Antennas Propag. Mag.*, vol. 38, no. 1, pp. 1–62, Feb. 1996.
- [17] S. Brooks, A. Gelman, G. Jones, and X.-L. Meng, *Handbook of Markov Chain Monte Carlo*. Boca Raton, FL, USA: CRC Press, 2011.

BIBLIOGRAPHIE

- [1] J. Dumont-Fillon, *Contrôle non destructif (CND)*, Ed. Techniques de l'Ingénieur, 1996.
- [2] J. Boissy, P. Minvielle, M. Sesques et J.-F. Giovannelli, « Cartographie d'impédance par inversion d'un code d'électromagnétisme 3D surfacique », in *Actes du 27^e colloque GRETSI*, Lille, France, août 2019.
- [3] J. Boissy, J.-F. Giovannelli et P. Minvielle, « An insight into the Gibbs sampler : Keep the samples or drop them ? », *IEEE Signal Processing Letters*, vol. 27, pp. 2069–2073, 2020.
- [4] J. Boissy, « Cartographie d'impédance par inversion d'un code d'électromagnétisme 3D », Rapport de stage de Master, IMS, Talence, France, septembre 2018.
- [5] J. Boissy, P. Minvielle, M. Sesques et J.-F. Giovannelli, « Cartographie d'impédance par inversion d'un code Maxwell 3D surfacique », Journées JOSO, Le Barp, France, mars 2019.
- [6] J. Boissy, P. Minvielle, M. Sesques et J.-F. Giovannelli, « Inversion pour la cartographie d'impédance surfacique 3D », Journée de l'ED SPI, Talence, France, mars 2020.
- [7] J. Boissy, P. Minvielle et J.-F. Giovannelli, « Near-field microwave mapping by Bayesian inversion-segmentation », in *SIAM Conference on Uncertainty Quantification*, Atlanta, Géorgie, États-Unis, avril 2022.
- [8] M. Venkatesh et V. Raghavan, « An Overview of Dielectric Properties Measuring Techniques », *Canadian Biosystems Engineering / Le Genie des biosystems au Canada*, vol. 47, pp. 15–30, janvier 2005.
- [9] W. S. Teng, M. Firdaus Akbar, G. Nihad Jawad, S. Y. Tan et M. I. S. Mohd Sazali, « A Past, Present, and Prospective Review on Microwave Nondestructive Evaluation of Composite Coatings », *Coatings*, vol. 11, N° 8, pp. 913, 2021.
- [10] X. Faget, *Application expérimentale de méthodes inverses avancées pour l'imagerie des propriétés électromagnétiques d'un matériau magnéto-diélectrique*, Thèse de Doctorat, Aix-Marseille, 2018.

- [11] R. E. Hiatt, E. F. Knott et T. Senior, « A study of VHF absorbers and anechoic rooms », Rapport interne, University of Michigan, National Aeronautics and Space Administration, février 1963.
- [12] X. Faget, A. Litman, E. Dieudonné, S. Enoch et N. Malléjac, « Free-space characterization of the permeability of inhomogeneous magneto-dielectric materials », *IEEE Transactions on Microwave Theory and Techniques*, vol. 65, N° 12, pp. 5035–5045, 2017.
- [13] F. Giraud, P. Minvielle et P. Del Moral, « Advanced interacting sequential Monte Carlo sampling for inverse scattering », *Inverse Problems*, vol. 29, N° 9, 2013.
- [14] P. Minvielle, A. Todeschini, F. Caron et P. Del Moral, « Particle MCMC for Bayesian microwave control », in *Journal of Physics : Conference Series*. IOP Publishing, 2014, vol. 542.
- [15] K. Hopcraft et P. Smith, *An introduction to electromagnetic inverse scattering*, Kluwer Academic Publishers, 1992.
- [16] W. Shao et T. McCollough, « Advances in Microwave Near-Field Imaging : Prototypes, Systems, and Applications », *IEEE Microwave Magazine*, vol. 21, N° 5, pp. 94–119, 2020.
- [17] P. Payen, *Modélisation d'un empilement de matériaux dans le domaine fréquentiel par une condition d'impédance d'ordre élevé*, Thèse de Doctorat, Université Paris-Nord, décembre 2020.
- [18] M. Leontovich, « Investigation of propagation of radio waves, part II », *USSR Acad. Sci., Moscow*, 1948.
- [19] E. F. Knott, J. F. Schaeffer et M. T. Tulley, *Radar cross section*, SciTech Publishing, 2004.
- [20] C. Augonnet, A. Pujols et M. Sesques, « Fast direct solver for integral equations on massively parallel architectures », in *2015 USNC-URSI Radio Science Meeting (Joint with AP-S Symposium)*. IEEE, 2015, pp. 149–149.
- [21] P. A. Raviart et J. M. Thomas, « A mixed finite element method for 2-nd order elliptic problems », in *Mathematical Aspects of Finite Element Methods*, I. Galligani et E. Magenes, Eds., Berlin, Heidelberg, 1977, pp. 292–315, Springer Berlin Heidelberg.
- [22] S. Tretyakov, *Analytical Modeling in Applied Electromagnetics*, Artech House electromagnetic analysis series. Artech House, 2003.
- [23] A. Bendali, M. B. Fares et J. Gay, « A boundary-element solution of the Leontovitch problem », *IEEE Transactions on Antennas and Propagation*, vol. 47, N° 10, pp. 1597–1605, 1999.
- [24] P. Pagani, P. Minvielle et M. Sesques, « Uncertainty Quantification for the RCS of a Coated Target using an IBC-based Metamodel », 2020.
- [25] S. Arhab, H. Ayasso, B. Duchêne et A. Mohammad-Djafari, « Optical imaging in a variational Bayesian framework », in *4th International Workshop on New Computational Methods for Inverse Problems*, Cachan, France, mai 2014, p. 6.

- [26] H. Ayasso, B. Duchêne et A. Mohammad-Djafari, « Bayesian estimation with Gauss-Markov-Potts priors in optical diffraction tomography », in *Computational Imaging IX*. International Society for Optics and Photonics, 2011, vol. 7873.
- [27] A. Mohammadi, Y. El Hachimi, S. Marichane et A. Mouad, « Étude comparative de trois approches d'évaluation des incertitudes : GUM, GUM S1 et globale », *Techniques de l'ingénieur*, juin 2019.
- [28] P. Lascaux et R. Théodor, *Analyse numérique matricielle appliquée à l'art de l'ingénieur*, vol. 1. Masson, 1993.
- [29] Y. Song, E.-H. Djermoune, D. Brie et C. Richard, « Séparation et déconvolution conjointes d'images hyperspectrales », in *Actes du 27^e colloque GRETSI*, Lille, France, août 2019.
- [30] E. Chouzenoux, M. Legendre, S. Moussaoui et J. Idier, « Fast constrained least squares spectral unmixing using primal-dual interior point optimization », *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 7, N° 1, pp. 59–69, janvier 2014.
- [31] J. Idier, Ed., *Approche bayésienne pour les problèmes inverses*, Traité IC2, Série traitement du signal et de l'image, Hermès, Paris, 2001.
- [32] A. N. Tikhonov, « On the stability of inverse problems », *Proceedings of the USSR Academy of Sciences*, vol. 39, pp. 195–198, 1943.
- [33] G. H. Golub, P. C. Hansen et D. P. O'Leary, « Tikhonov Regularization and Total Least Squares », *SIAM Journal on Matrix Analysis and Applications*, vol. 21, N° 1, pp. 185–194, 1999.
- [34] T. Martin et J. Idier, « Problème mal posé et reconstruction bayésienne en tomographie d'impédance électrique », in *Actes du 15^e colloque GRETSI*, Juan-les-Pins, France, septembre 1995, pp. 905–908.
- [35] C. Robert et G. Casella, *Monte Carlo Statistical Methods*, Springer Texts in Statistics. Springer New York, 2005.
- [36] B. Harroué, J.-F. Giovannelli et M. Pereyra, « Sélection de modèles en restauration d'image. Approche bayésienne dans le cas gaussien », in *Actes du 27^e colloque GRETSI*, Lille, France, août 2019.
- [37] B. Harroué, J.-F. Giovannelli et M. Pereyra, « Bayesian Model Selection for Unsupervised Image Deconvolution with Structured Gaussian Priors », in *2021 IEEE Statistical Signal Processing Workshop (SSP)*, 2021, pp. 241–245.
- [38] C. Vacar, J.-F. Giovannelli et A.-M. Roman, « Bayesian texture model selection by harmonic mean », in *2012 19th IEEE International Conference on Image Processing*, 2012, pp. 2533–2536.
- [39] F. Samaniego, *A Comparison of the Bayesian and Frequentist Approaches to Estimation*, Springer Series in Statistics. Springer New York, 2010.
- [40] R. Bassett et J. Deride, « Maximum *a posteriori* estimators as a limit of Bayes estimators », *Mathematical Programming*, vol. 174, N° 1-2, pp. 129–144, janvier 2018.

- [41] P. Pérez, « Markov random fields and images », Rapport interne 1196, Irisa, 1998.
- [42] F. Y. Wu, « The Potts model », *Rev. Mod. Phys.*, vol. 54, pp. 235–268, janvier 1982.
- [43] S. Geman et D. Geman, « Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images », *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PAMI-6, N° 6, pp. 721–741, 1984.
- [44] Y. Altmann, N. Dobigeon, S. McLaughlin et J.-Y. Tournet, « Residual Component Analysis of Hyperspectral Images—Application to Joint Nonlinear Unmixing and Nonlinearity Detection », *IEEE Transactions on Image Processing*, vol. 23, N° 5, pp. 2148–2158, 2014.
- [45] A. Mohammad-Djafari, O. Féron et Z. Chama, « Reconstruction of piecewise homogeneous images from partial knowledge of their Fourier transform », in *MaxEnt 04, Garching, Germany*, août 2004.
- [46] M. Pereyra, N. Dobigeon, H. Batatia et J.-Y. Tournet, « Segmentation of Skin Lesions in 2-D and 3-D Ultrasound Images Using a Spatially Coherent Generalized Rayleigh Mixture Model », *IEEE Transactions on Medical Imaging*, vol. 31, N° 8, pp. 1509–1520, 2012.
- [47] O. Eches, J. A. Benediktsson, N. Dobigeon et J.-Y. Tournet, « Adaptive Markov Random Fields for Joint Unmixing and Segmentation of Hyperspectral Images », *IEEE Transactions on Image Processing*, vol. Vol. 22, pp. 5–16, janvier 2013.
- [48] N. Zhao, A. Basarab, D. Kouamé et J.-Y. Tournet, « Joint Segmentation and Deconvolution of Ultrasound Images Using a Hierarchical Bayesian Model Based on Generalized Gaussian Priors », *IEEE Transactions on Image Processing*, vol. vol. 25, N° 8, pp. 3736–3750, août 2016.
- [49] H. Ayasso, B. Duchêne et A. Mohammad-Djafari, « Bayesian Inversion for Optical Diffraction Tomography », *Journal of Modern Optics*, vol. 57, N° 9, pp. 765–776, mai 2010.
- [50] O. Eches, N. Dobigeon et J.-Y. Tournet, « Enhancing Hyperspectral Image Unmixing With Spatial Correlations », *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, N° 11, pp. 4239–4247, 2011.
- [51] A. Mohammad-Djafari, « Gauss-Markov-Potts Priors for Images in Computer Tomography Resulting to Joint Optimal Reconstruction and segmentation », *International Journal of Tomography and Statistics*, vol. 11, N° W09, pp. 76–92, 2008.
- [52] H. Ayasso et A. Mohammad-Djafari, « Joint NDT Image Restoration and Segmentation Using Gauss–Markov–Potts Prior Models and Variational Bayesian Computation », *IEEE Transactions on Image Processing*, vol. 19, N° 9, pp. 2265–2277, 2010.
- [53] C. Vacar et J.-F. Giovannelli, « Unsupervised joint deconvolution and segmentation method for textured images : a Bayesian approach and an advanced sampling algorithm », *EURASIP Journal on Advances in Signal Processing*, mars 2019.
- [54] C. Chapdelaine, A. Mohammad-Djafari, N. Gac et E. Parra-Denis, « A 3D Bayesian Computed Tomography Reconstruction Algorithm with Gauss-Markov-Potts Prior Model and its Application to Real Data », *Fundamenta Informaticae*, vol. 155, N° 4, pp. 373–405, 2017.

- [55] M.-C. Corbineau, D. Kouamé, E. Chouzenoux, J.-Y. Tourneret et J.-C. Pesquet, « Preconditioned P-ULA for Joint Deconvolution-Segmentation of Ultrasound Images », *IEEE Signal Processing Letters*, vol. 26, N° 10, pp. 1456–1460, 2019.
- [56] A. Mohammadpour, O. Féron et A. Mohammad-Djafari, « Bayesian segmentation of hyperspectral images », *AIP Conference Proceedings*, vol. 735, N° 1, pp. 541–548, 2004.
- [57] S. Ouali, J.-B. Courbot, R. Pierron et O. Haerberle, « Segmentation statistique d’images du bois de vigne par champs de Markov triplets », in *Actes du 28^e colloque GRETSI*, Nancy, France, septembre 2022.
- [58] M. Mignotte, « A segmentation-based regularization term for image deconvolution », *IEEE Transactions on Image Processing*, vol. 15, N° 7, pp. 1973–1984, 2006.
- [59] O. Féron, B. Duchêne et A. Mohammad-Djafari, « Microwave imaging of inhomogeneous objects made of a finite number of dielectric and conductive materials from experimental data », *Inverse Problems*, vol. 21, N° 6, pp. S95–S115, décembre 2005.
- [60] O. Féron et A. Mohammad-Djafari, « Image fusion and unsupervised joint segmentation using a HMM and MCMC algorithms », *Journal of Electronic Imaging*, vol. 14, N° 2, pp. 1–12, 2005.
- [61] H. Ayasso, B. Duchêne et A. Mohammad-Djafari, « Optical diffraction tomography within a variational Bayesian framework », *Inverse Problems in Science and Engineering*, vol. 20, N° 1, pp. 59–73, février 2012.
- [62] M. Storath, A. Weinmann, J. Friel et M. Unser, « Joint image reconstruction and segmentation using the Potts model », *Inverse Problems*, vol. 31, N° 2, janvier 2015.
- [63] J. Cui et Z. Qiang, « Potts Model for Limited Angle Tomography », in *2017 2nd International Conference on Image, Vision and Computing (ICIVC)*, 2017, pp. 611–615.
- [64] M. Nikolova, « Thresholding implied by truncated quadratic regularization », *IEEE Transactions on Signal Processing*, vol. 48, N° 12, pp. 3437–3450, 2000.
- [65] M. Ravasi et C. Birnie, « A joint inversion-segmentation approach to assisted seismic interpretation », *Geophysical Journal International*, vol. 228, N° 2, pp. 893–912, septembre 2021.
- [66] R. Morris, X. Descombes et J. Zerubia, « The Ising/Potts model is not well suited to segmentation tasks », in *1996 IEEE Digital Signal Processing Workshop Proceedings*, 1996, pp. 263–266.
- [67] X. Descombes, R. Morris, J. Zerubia et M. Berthod, « Estimation of Markov random field prior parameters using Markov chain Monte Carlo maximum likelihood », *IEEE Transactions on Image Processing*, vol. 8, N° 7, pp. 954–963, 1999.
- [68] C. Tison, J.-M. Nicolas, F. Tupin et H. Maître, « A new statistical model for Markovian classification of urban areas in high-resolution SAR images », *IEEE Transactions on Geoscience and Remote Sensing*, vol. 42, N° 10, pp. 2046–2057, 2004.

- [69] P. Ciuciu, F. Forbes, T. Vincent et L. Chaari, « Joint Detection-Estimation in Functional MRI », in *Regularization and Bayesian Methods for Inverse Problems in Signal and Image Processing*, J.-F. Giovannelli et J. Idier, Eds., Digital signal and image processing series, chapitre 7. Wiley, 2015.
- [70] R. Rosu, J.-F. Giovannelli, A. Giremus et C. Vacar, « Potts model parameter estimation in Bayesian segmentation of piecewise constant images », in *Proceedings of the International Conference on Acoustic, Speech and Signal Processing*, Brisbane, Australie, avril 2015, pp. 4080–4084.
- [71] M. A. Pereyra, N. Dobigeon, H. Batatia et J.-Y. Tourneret, « Estimating the granularity coefficient of a Potts-Markov random field within an MCMC algorithm », *IEEE Transactions on Image Processing*, vol. vol. 22, N° n° 6, pp. 2385–2397, juin 2013.
- [72] J. Budd, Y. van Gennip et J. Latz, « Classification and image processing with a semi-discrete scheme for fidelity forced Allen–Cahn on graphs », *GAMM-Mitteilungen*, vol. 44, N° 1, mars 2021.
- [73] J. Budd, « Joint Reconstruction-Segmentation with Graph PDEs », Présentation dans le cadre du One World Seminar on the Mathematics of Machine Learning. Vidéo disponible à l'adresse : <https://youtu.be/-NkIxrnv8DU?list=PL726PK-VPGxspstscGKcqS21KCI25IJHS9>, novembre 2021.
- [74] J.-M. Marin et C. P. Robert, *Bayesian Core. A Practical Approach to Computational Bayesian Statistics*, Texts in statistics. Springer, Paris, France, 2007.
- [75] K. Miller, *Complex Stochastic Processes : an Introduction to Theory and Application*, Advanced book program. Addison-Wesley Publishing Company, 1974.
- [76] J. M. Ver Hoef, E. M. Hanks et M. B. Hooten, « On the relationship between conditional (CAR) and simultaneous (SAR) autoregressive models », *Spatial Statistics*, vol. 25, pp. 68–85, 2018.
- [77] J. Pearl, *Causality*, Cambridge University Press, 2009.
- [78] A. Doucet, S. Godsill et C. Robert, « Marginal maximum *a posteriori* estimation using Markov chain Monte Carlo », *Statistics and Computing*, vol. 12, pp. 77–84, janvier 2002.
- [79] C. Andrieu, N. de Freitas, A. Doucet et M. I. Jordan, « An Introduction to MCMC for Machine Learning », *Machine Learning*, vol. 50, N° 1, pp. 5–43, janvier 2003.
- [80] D. Gamerman et H. Lopes, *Markov Chain Monte Carlo : Stochastic Simulation for Bayesian Inference, Second Edition*, Chapman & Hall/CRC Texts in Statistical Science. CRC Press, 2006.
- [81] J. Drosbeke, J. Fine et G. Saporta, *Méthodes bayésiennes en statistique*, Technip, 2002.
- [82] M. Girolami et B. Calderhead, « Riemann Manifold Langevin and Hamiltonian Monte Carlo », *Journal of the Royal Statistical Society : Series B (Statistical Methodology)*, vol. 73, N° 2, pp. 123–214, 2011.
- [83] N. Shephard, « Partial non-Gaussian state space », *Biometrika*, vol. 81, N° 1, pp. 115–131, mars 1994.

- [84] G. Backlund, J. P. Hobert, Y. J. Jung et K. Khare, « A Hybrid Scan Gibbs Sampler for Bayesian Models with Latent Variables », *Statistical Science*, vol. 36, N° 3, pp. 379–399, 2021.
- [85] D. A. van Dyk et T. Park, « Partially Collapsed Gibbs Samplers », *Journal of the American Statistical Association*, vol. 103, N° 482, pp. 790–796, 2008.
- [86] D. Luengo, L. Martino, M. Bugallo, V. Elvira et S. Särkkä, « A survey of Monte Carlo methods for parameter estimation », *EURASIP Journal on Advances in Signal Processing*, vol. 2020, mai 2020.
- [87] V. Churchill et A. Gelb, « Synthetic Aperture Radar Image Formation with Uncertainty Quantification », 2020.
- [88] F. Orieux, O. Féron et J.-F. Giovannelli, « Sampling High-Dimensional Gaussian Distributions for General Linear Inverse Problems », *IEEE Signal Processing Letters*, vol. 19, N° 5, pp. 251–254, 2012.
- [89] C. Gilavert, S. Moussaoui et J. Idier, « Efficient Gaussian Sampling for Solving Large-Scale Inverse Problems Using MCMC », *IEEE Transactions on Signal Processing*, vol. 63, N° 1, pp. 70–80, 2015.
- [90] A. Doucet, « A Note on Efficient Conditional Simulation of Gaussian Distributions », Rapport interne disponible à l'adresse : http://www.stats.ox.ac.uk/~doucet/doucet_simulationconditionalgaussian.pdf, avril 2010.
- [91] Y. Hoffman et E. Ribak, « Constrained Realizations of Gaussian Fields : A Simple Algorithm », *The Astrophysical Journal*, vol. 380, octobre 1991.
- [92] G. Evensen, *Data Assimilation : The Ensemble Kalman Filter*, Springer Berlin Heidelberg, 2009.
- [93] R. Swendsen et J.-S. Wang, « Nonuniversal Critical Dynamics in Monte Carlo Simulation », *Physical review letters*, vol. 58, pp. 86–88, février 1987.
- [94] J. Sodjo, A. Giremus, N. Dobigeon et F. Caron, « A Bayesian Nonparametric Model for Unsupervised Joint Segmentation of a Collection of Images », *IEEE Access*, vol. 7, pp. 120176–120188, 2019.
- [95] A.-C. Barros, F. Caron, J.-F. Giovannelli et A. Doucet, « Clone MCMC : Parallel High-Dimensional Gaussian Gibbs Sampling », in *Advances in Neural Information Processing Systems*, I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan et R. Garnett, Eds. 2017, vol. 30, Curran Associates, Inc.
- [96] J.-F. Giovannelli et A. Barros, « Unsupervised segmentation of piecewise constant images from incomplete, distorted and noisy data », in *2016 IEEE Statistical Signal Processing Workshop (SSP)*, 2016, pp. 1–5.
- [97] L. Risser, T. Vincent, F. Forbes, J. Idier et P. Ciuciu, « Min-max extrapolation scheme for fast estimation of 3D Potts field partition functions. Application to the joint detection-estimation of brain activity in fMRI », *Journal of Signal Processing Systems*, vol. 65, N° 3, pp. 325–338, décembre 2011.

- [98] M. Stephens, « Dealing with label switching in mixture models », *Journal of the Royal Statistical Society : Series B (Statistical Methodology)*, vol. 62, N° 4, pp. 795–809, 2000.
- [99] T. B. Senior et J. L. Volakis, *Approximate boundary conditions in electromagnetics*, 41. Institution of Electrical Engineers, 1995.
- [100] D. Colton et R. Kress, *Integral equation methods in scattering theory*, SIAM, 2013.