



HAL
open science

Surface representation by linear forms aimed at a statistical study applied to cerebral imaging

Benjamin Coulaud

► **To cite this version:**

Benjamin Coulaud. Surface representation by linear forms aimed at a statistical study applied to cerebral imaging . Mathématiques [math]. Aix-Marseille Université, 2017. Français. NNT: . tel-01842157

HAL Id: tel-01842157

<https://hal.science/tel-01842157>

Submitted on 18 Jul 2018

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.

AIX-MARSEILLE UNIVERSITÉ
ECOLE DOCTORALE Mathématiques et Informatique (ED 184)

I2M / UMR 7373

Thèse présentée pour obtenir le grade universitaire de docteur

Benjamin COULAUD

Construction d'un cadre statistique consistant pour l'analyse de surfaces au travers de processus généralisés. Application à la classification de surfaces cérébrales extraites d'IRM.

Soutenue le 21/09/2017 devant le jury composé de :

Alain TROUVÉ	ENS Cachan	Rapporteur
Jean-Michel LOUBÈS	Université Paul Sabatier	Rapporteur
Simon MASNOU	Institut Camille Jordan	Examinateur
Stéphanie ALLASSONNIÈRE	Université Paris Descartes	Examinateur
Julien LEFÈVRE	Aix-Marseille Université	Examinateur
Frédéric RICHARD	Aix-Marseille Université	Directeur de thèse

Table des matières

I Cadre consistant pour l'analyse statistique de surfaces représentées par des courants	11
1 Introduction	12
1.1 Motivations	12
1.2 Difficultés liées à l'analyse statistique de surfaces, solutions envisagées	13
1.2.1 Caractère riemannien des surfaces	13
1.2.2 Représentation aléatoire des formes linéaires	13
1.2.3 Traitement de la dimension infinie	14
1.3 Objectifs	14
1.3.1 Contributions	14
1.3.2 Contenus de cette partie	15
2 État de l'art	16
2.1 Représentation des surfaces par des courants	16
2.1.1 Pourquoi utiliser des courants ?	16
2.1.2 Cadre de représentation de Glaunès et Vaillant	16
2.1.3 Cadre de représentation de Durreleman	18
2.1.4 Contributions de notre étude par rapport aux cadres originaux de Glaunès et Durreleman	20
2.2 Espaces de Hilbert à noyau reproduisant	21
2.2.1 Théorie des espaces de Hilbert à noyau reproduisant	21
2.2.2 Exemples de RKHS	24
2.3 Processus stochastiques généralisés	25
2.3.1 Fonctionnelles aléatoires, concept de processus généralisé	25
2.3.2 Moments d'un processus aléatoire généralisé	27
2.3.3 Processus gaussiens	27

3	Modèle d'observation : du modèle continu au modèle discret	29
3.1	Objectifs	29
3.2	Modèle statistique	30
3.2.1	Cadre déterministe de représentation des surfaces	30
3.2.2	Formes linéaires aléatoires	31
3.2.3	Modèle de variation	35
3.3	Discrétisation et modèle d'observation	36
4	Estimation (dans le cadre d'un modèle de bruit blanc)	42
4.1	Objectifs	42
4.2	Approximation	43
4.3	Construction des estimateurs	44
4.3.1	Estimateurs dans le modèle de variation	44
4.3.2	Estimateurs dans le modèle d'observation	45
4.4	Consistance	46
4.4.1	Estimateur du représentant moyen	46
4.4.2	Estimateur de la variance du bruit	50
5	Estimation (dans le cadre d'un modèle général de bruit gaussien)	56
5.1	Objectifs	56
5.2	Formes bilinéaires aléatoires	56
5.2.1	Définitions	57
5.2.2	Structure d'espace de Banach	57
5.3	Construction des estimateurs	58
5.3.1	Notations	58
5.3.2	Estimation du représentant moyen	59
5.3.3	Estimateur de l'opérateur d'autocovariance	59
5.4	Consistance	60
6	Expériences de validation	66
6.1	Objectifs	66
6.2	Simulation d'échantillons	67
6.2.1	Surfaces à creux et bosses	67
6.2.2	Surfaces "Browniennes"	68
6.3	Espace des représentants	70
6.3.1	Espace de champs de vecteurs	70

6.3.2	Espaces d'approximation	71
6.4	Recherche de la dimension optimale de l'espace d'approximation	71
6.4.1	Protocole	71
6.4.2	Résultats	72
6.5	Convergence numérique de l'estimateur du représentant moyen	73
6.5.1	Protocole	73
6.5.2	Résultats	74
7	Discussion	76
7.1	Bilan	76
7.2	Perspectives	77
7.2.1	Perspectives immédiates	77
7.2.2	Perspectives ultérieures	77
II	Application à la classification de surfaces	78
8	Introduction	79
8.1	Motivations et difficultés	79
8.2	Objectifs	79
8.2.1	Contributions	79
8.2.2	Contenus de cette partie	80
9	Structure d'espace métrique sur l'ensemble des surfaces	82
9.1	Distance sur les surfaces	82
9.1.1	Choix des ensembles	82
9.1.2	Injectivité de la fonction Ψ	83
9.1.3	Extension de la distance aux surfaces discrétisées	85
9.2	Propriétés de l'espace métrique des surfaces	87
9.2.1	Question de l'homéomorphisme entre Θ et V^*	87
9.2.2	Question du caractère complet de Θ	89
10	Méthode de classification Bayésienne adaptée aux surfaces	91
10.1	Principe	91
10.1.1	Classification : pose du problème	91
10.1.2	Classification Bayésienne	91
10.2	Algorithme	93

10.2.1 Cas général	93
10.2.2 Cas particulier de l'autocovariance diagonale	93
11 Application à l'imagerie cérébrale	95
11.1 Objectifs	95
11.2 Protocole	95
11.2.1 Échantillon de surfaces	95
11.2.2 Choix des paramètres	97
11.2.3 Procédure de validation croisée	97
11.3 Résultats	98
12 Discussion	99
12.1 Bilan	99
12.2 Perspectives	99
12.2.1 Espace des surfaces	99
12.2.2 Classification	100
A Calcul explicite du représentant d'une surface dans une base de Fourier	107
A.1 Objectifs	107
A.2 Méthode de calcul	107
A.3 Résultats	108
B Éléments sur la notion de surface moyenne	110
B.1 Définitions	110
B.1.1 Dans un cadre déterministe	110
B.1.2 Dans un cadre probabiliste, moyenne théorique et moyenne empirique	111
B.2 Éléments sur l'estimation de la surface moyenne	111
B.2.1 Objectifs et notations	111
B.2.2 Estimation par la méthode du plus proche représentant	112

Remerciements

Il y a quatre ans, Frédéric Richard me proposait de travailler avec lui, dans le cadre d'un stage de fin de Master, sur le sujet "Représentation des surfaces par des formes linéaires sur un espace fonctionnel en vue d'une étude statistique". À cette époque, ma motivation personnelle pour les mathématiques allait décroissant; je ne pensais, pour dire vrai, pas avoir les capacités pour aller plus loin. Cette nouvelle thématique a ravivé ma curiosité; je remercie donc celui qui fut mon responsable de stage et est maintenant mon directeur de thèse car c'est à lui que je dois ce regain de foi en mon avenir dans cette belle discipline.

Faire des mathématiques vouées à l'application fut pour moi un immense changement. J'avais alors une sacrosainte aversion pour les expériences et pour l'utilisation de matlab en général. Un grand merci à Julien Lefèvre et Hamed Rabiei pour leur aide providentielle, sans laquelle les premières implémentations sur machine auraient constitué un obstacle conséquent.

Merci à tous les camarades enseignants-chercheurs que j'ai cotôyés pendant ces quatre années. Votre disponibilité et votre penchant naturel pour l'entraide vous honorent. Un salut tout particulier à Khuyen pour sa bonne humeur constante et sa gentillesse sans pareil. Il faut croire que parfois, sous les mornes carreaux plastifiés d'un sol de laboratoire, peuvent pousser les racines de l'amitié.

J.R.R. Tolkien écrivit : "Il est fort dangereux de sortir de chez soi, on prend la route et, si l'on ne regarde pas où l'on met les pieds, on ne sait pas jusqu'où cela peut nous mener". Cette fausse mise en garde, en réalité une véritable invitation au voyage, vaut pour la pratique des mathématiques : on s'engage sur un sujet avec une vision quelque peu limitée et, au fil des lectures, des difficultés et des rencontres, on peut aller bien plus loin que sa motivation initiale et explorer des idées dont on ne soupçonnait même pas l'existence. Ces découvertes inattendues font la beauté de notre science, comme elles font le charme du monde et de la vie en général.

Cette thèse, je la dédie à la mémoire de Pierre Liardet, ce professeur aussi humain que passionné, dont les cours me laissent le souvenir le plus fort de toutes mes années d'études.

Résumé

Le thème principal de cette thèse est l'analyse statistique de surfaces. Nous introduisons un formalisme pour étendre dans un cadre aléatoire la représentation des surfaces introduite par Glaunès et Vaillant (2005). Ce formalisme repose sur une notion de forme linéaire aléatoire définie sur des espaces de champs vectoriels, qui est inspirée de la théorie des processus linéaires généralisés développée par Itô (1954) et Gel'fand et Vilenkin (1964). A partir de cette représentation, nous mettons en place un modèle probabiliste décrivant la variabilité des surfaces. Par passage du continu au discret, nous prolongeons ce dernier en un modèle d'observation permettant de décrire des données expérimentales. À partir de ce modèle, nous construisons des estimateurs du représentant moyen d'un échantillon de surfaces et de l'autocovariance du bruit. Nous démontrons des résultats de consistance de ces estimateurs. Nous présentons quelques expériences de validation de la méthode d'estimation sur des données simulées. Nous appliquons cette méthode à la classification de surfaces cérébrales issues de l'IRM en nous plaçant dans un cadre bayésien.

Abstract

The main topic of this manuscript is surface statistic analysis. We define a new formalism that extends to a random framework the surface representation introduced by Glaunès and Vaillant (2005). This formalism is based on a notion of random linear form, which is inspired from the theory of generalized random process, developed by Itô (1954) and Gel'fand and Vilenkin (1964). On this representation, we set a probabilistic model that describes the variability of surfaces; by a transition from continuous to discrete, we extend it to an observation model that describes experimental data. We build estimators for the mean representant of a surface sample and for the autocovariance of the noise. We demonstrate that these estimators are consistent. We report some experiments in which we test our estimation method on simulated data. We apply our statistical framework to classification of brain surfaces from MRI, using a Bayesian classification procedure.

Introduction générale

Motivations pour l'analyse statistique de surfaces

Les surfaces sont les objets idéaux pour modéliser les données tridimensionnelles. Ainsi, en raison du développement croissant des méthodes d'imagerie 3-D, ces objets ont pris une place importante dans la littérature mathématique associée au traitement du signal. Dans cette optique, on a actuellement à traiter, dans de nombreux champs d'application, de grandes quantités de données représentées sous formes de surfaces. Ceci est particulièrement vrai en ce qui concerne le domaine de l'imagerie médicale, où se posent, de fait, des problèmes de nature statistique sur des ensembles de surfaces : définir et déterminer des caractères moyens et/ou des modes de variation, exhiber des régularités, établir des groupes. . . De ces problématiques émerge la première motivation de notre étude : construire un cadre statistique rigoureux adapté à l'analyse de surfaces.

Les surfaces sont au cœur du sujet majeur de ce manuscrit ; de ce fait, il est nécessaire de s'accorder de façon précise sur une définition pour cette notion, afin de garantir la cohérence du propos. On prend le parti d'utiliser à ce dessein les sous-variétés différentielles.

Définition 1 (Surface). *On appelle surface toute variété différentielle de dimension 2 plongée dans \mathbb{R}^3 . En d'autres termes, une surface est un sous-ensemble S de \mathbb{R}^3 tel que, pour tout x de S , il existe un voisinage W de x dans \mathbb{R}^3 , un ouvert U de \mathbb{R}^2 et une bijection de classe C^1 reliant U et $S \cap W$.*

En tant que sous-variétés différentielles, les surfaces ne constituent pas un espace vectoriel : on ne dispose pas d'opérateurs de somme interne ni de produit externe ; les calculs de distances, de moyennes, etc. ne font donc a priori pas sens sur ces objets. On ne peut donc pas leur appliquer les méthodes statistiques les plus usuelles, qui reposent en général sur une structure d'espace Euclidien. Il convient alors de développer une méthode innovante, spécifiquement adaptée au cas des surfaces.

Contenus

Première partie : Cadre consistant pour l'analyse statistique de surfaces

Dans la logique des motivations évoquées plus haut, on va tout d'abord mettre en place un cadre de représentation des surfaces par des éléments d'un espace euclidien ; en l'occurrence, on adaptera la représentation par courants, introduite par Glaunès et Vaillant [33, 16], en établissant une correspondance entre surfaces et formes linéaires continues sur un ensemble de champs de vecteurs de l'espace. Ensuite, en utilisant des processus aléatoires généralisés, on confèrera à ce cadre la possibilité de tenir compte de variations aléatoires affectant les représentants des surfaces d'un échantillon ; cela permettra la construction d'un modèle d'observation décrivant ces variations autour d'un représentant moyen.

Dans notre modèle de variation, on écrira, via la méthode du maximum de vraisemblance, des estimateurs pour le représentant moyen d'un échantillon de surfaces ainsi que pour l'opérateur de covariance associé. On démontrera la consistance de ces estimateurs, validant ainsi le caractère pertinent de notre cadre statistique, d'un point de vue théorique. Enfin, on mettra en œuvre ces estimateurs sur des exemples simples de données simulées.

Seconde partie : Application à la classification de surfaces

Dans la seconde partie, l'objectif principal sera d'appliquer le modèle statistique précédemment établi à la résolution numérique de problèmes de classification au sein d'ensembles de surfaces. Dans ce but, on commencera par prolonger le cadre de représentation par des formes linéaires en une structure d'espace métrique sur les surfaces. Cette structure permettra ensuite d'adapter une méthode usuelle de classification Bayésienne au cas particulier des surfaces. Enfin, on effectuera une expérience mettant en application cette procédure, en l'occurrence sur des surfaces cérébrales obtenues à partir de l'IRM.

Première partie

Cadre consistant pour l'analyse statistique de surfaces représentées par des courants

Chapitre 1

Introduction

1.1 Motivations

Ainsi que cela a été mentionné dans l'Introduction générale, les surfaces sont des objets de premier intérêt dans de nombreuses branches des mathématiques : en géométrie, en analyse, mais aussi en traitement du signal, domaine dans lequel ce sont les objets idéaux pour représenter la plupart des primitives extraites d'images tridimensionnelles. De par l'essor actuel des techniques d'imagerie 3-D, on fait ainsi face, dans de nombreux champs d'application, à des quantités de plus en plus grandes de données modélisées par des surfaces. C'est particulièrement vrai dans le domaine de l'imagerie médicale, où l'on dispose de populations de surfaces particulièrement étendues issues des données cliniques. Ces populations correspondent à des observations d'organes internes (coeur, foie, cerveau...) provenant de patients sains ou présentant un même syndrome ou encore issus d'échantillons représentatifs... D'un point de vue diagnostique, il serait intéressant de dresser des classifications de ces populations, les comparer entre elles et établir des caractères moyens associés à un organe sain ou atteint d'une anomalie donnée, calculer modes de variation et indicateurs de dispersion relatifs à diverses données géométriques associées à ces échantillons de surfaces ou encore savoir ranger un individu parmi différentes catégories via l'observation de quelques unes de ses caractéristiques. Compte tenu de cela, il apparaît qu'une analyse statistique de telles données serait une aide majeure quant à la connaissance et au dépistage de certaines maladies et anomalies. Dans cette optique, le but de cette partie est de mettre en place un cadre statistique adapté à l'étude des surfaces.

1.2 Difficultés liées à l'analyse statistique de surfaces, solutions envisagées

1.2.1 Caractère riemannien des surfaces

Les outils usuels de l'analyse statistique, tels que les méthodes d'estimation classiques, paramétriques ou non paramétriques, sont conçus pour être opérants sur des espaces vectoriels. En effet, ils nécessitent de pouvoir définir et calculer des sommes, des différences, des produits par les éléments d'un corps de scalaires ou encore des moyennes... Le cadre idéal pour les mettre en oeuvre étant, bien entendu, un espace normé de type euclidien ou hermitien, sur lequel on dispose d'une distance et d'un produit scalaire. Les échantillons de données expérimentales qui peuvent être traités via ces outils sont ceux que l'on peut modéliser par des éléments de tels espaces ; malheureusement, ce n'est pas le cas des surfaces à première vue. Les surfaces, en effet, sont des variétés riemanniennes différentiables bidimensionnelles de \mathbb{R}^3 ; elles ne sont éléments d'aucun espace vectoriel et, par conséquent, ne sont pas analysables au moyen de méthodes statistiques usuelles. Ainsi, l'analyse statistique de surfaces nécessite le développement de réponses spécifiques en matière de représentation de données ou d'algorithmes statistiques.

Dans la littérature, il existe de nombreux articles liés à notre thématique principale. De ces ressources, il émerge deux grandes catégories de solutions pour pallier notre problème. La première consiste en le développement de méthodes statistiques spécifiquement adaptées aux variétés riemanniennes [27, 5]. La deuxième repose sur le fait de plonger les surfaces dans un espace euclidien via un procédé de représentation de ces dernières par des éléments d'un espace vectoriel puis de mener une étude indirecte sur les représentants. On opte pour la seconde manière. Pour représenter les surfaces dans un espace vectoriel, on adopte le point de vue initié par J.Glaunès et M.Vaillant en 2005 [16, 33, 17] et prolongé, deux ans plus tard, par S.Durreleman [10, 9] : associer aux surfaces discrétisées des courants sur un espace fonctionnel de type Hilbert. On développe ensuite une approche statistique originale dans ce cadre.

Suivant une démarche analogue, étant donné un ensemble de surfaces continument différentiables Θ et un espace de Hilbert V , on représentera une surface $S \in \Theta$ par une forme linéaire s sur V . Sous des hypothèses raisonnables, S est un élément du dual topologique V^* de V . On représente ainsi les surfaces par des éléments d'un espace de Hilbert de dimension infinie. On étendra cette représentation par formes linéaires aux surfaces discrètes, à partir du modèle continu.

1.2.2 Représentation aléatoire des formes linéaires

Pour établir un cadre statistique avec un modèle d'observation, il faudra prendre en compte les variations aléatoires des surfaces en les rendant visibles à travers leurs représentants. Ces représentants étant des courants, *i.e.* des formes linéaires continues sur un espace de Hilbert, il est particulièrement difficile d'en établir une modélisation probabiliste. On surmontera cette difficulté en utilisant des formes linéaires aléatoires,

concept défini à partir d'éléments empruntés à la théorie des processus stochastiques généralisés, telle que développée par Itô [20, 21], Fernique [12] et surtout Gel'fand et Vilenkin [15] ou Yaglom [35]. Sur ce cadre aléatoire, il deviendra possible d'attribuer de manière rigoureuse des lois de probabilité aux représentants des surfaces, construire un modèle d'observation décrivant ces derniers, puis appliquer des outils statistiques usuels, afin de construire des estimateurs pour le représentant moyen ainsi que pour les principaux modes de variation d'un échantillon.

1.2.3 Traitement de la dimension infinie

Comme il a été mentionné plus haut, on représente les surfaces, lisses ou discrétisées, par des éléments d'un espace de dimension infinie. Les représentants dépendent donc d'une infinité de paramètres. Il est donc nécessaire, pour pouvoir effectuer sur ces derniers des calculs numériques, de mettre en place une procédure d'approximation. Pour ce faire, on utilisera un sous-espace vectoriel de dimension finie de l'espace des représentants et on approximera toute forme linéaire par sa projection orthogonale sur ce sous-espace ; on obtiendra ainsi des représentants approximatifs dépendant d'un nombre de paramètres fini et fixé. Ce point constitue une autre originalité par rapport aux travaux de Glaunès, dans lesquels les approximations dépendent de la discrétisation des surfaces.

1.3 Objectifs

1.3.1 Contributions

Nos motivations étant singulières au regard de ce qui existe dans la littérature, on va devoir développer des réponses spécifiques. La première de nos contributions est la mise en place, pour pouvoir modéliser les variations aléatoires des surfaces à travers leurs représentants, d'un cadre probabiliste dans lequel les courants sont vus comme des réalisations de formes linéaires aléatoires, concept que l'on définit à partir de la théorie des processus stochastiques généralisés de Gel'fand et Vilenkin. La deuxième contribution de cette partie est notre modèle d'observation pour les représentants d'un échantillon de surfaces ; ce modèle représente les variations autour d'un élément moyen via un bruit blanc et un terme représentant l'erreur due à l'approximation discrète. La troisième contribution réside dans la construction d'estimateurs pour le représentant moyen et la variance du bruit ainsi que la démonstration de la consistance de ces derniers.

Une autre contribution réside dans le passage du continu au discret. Dans les travaux de Glaunès et Durrleman, la représentation par courants est mise en place directement sur les surfaces discrètes, le cas continu n'est pas traité. À l'inverse, on construit notre cadre de représentation sur des ensembles de surfaces continument différentiables puis on l'étend, via une procédure d'approximation, au cas des surfaces discrètes.

tisées correspondant mieux à la réalité expérimentale. On démontre un théorème garantissant la cohérence de ce passage du continu au discret : le représentant de l'approximation discrète d'une surface continue converge vers le représentant de cette dernière lorsque la discrétisation devient de plus en plus précise. De manière analogue, on élabore le modèle d'observation par passage aux approximations discrétisées sur modèle de variation établi pour les représentants des surfaces continues. De façon similaire, la construction des estimateurs et la démonstration de leur consistance se font également en passant d'un estimateur du maximum de vraisemblance sur le modèle continu à un estimateur empirique calculé par extension au cas discret.

1.3.2 Contenus de cette partie

Cette partie s'organise en quatre chapitres. Dans le premier, on dressera un état de l'art détaillé autour des principales ressources issues de la littérature mathématique utilisées pour construire cette étude, que ces dernières concernent l'analyse de surfaces, les espaces de Hilbert, les processus aléatoires... Dans le deuxième chapitre, on mettra en place un cadre de représentation des surfaces par des formes linéaires sur un espace de Hilbert. Via une modélisation aléatoire de ces formes linéaires, on établira un modèle d'observation décrivant la variabilité autour d'un élément moyen des représentants associés à un échantillon de surfaces. Dans le troisième chapitre, on construira des estimateurs pour le représentant moyen d'un échantillon de surfaces et pour la variance du bruit dans notre modèle d'observation. Enfin, dans le quatrième chapitre, on effectuera des expériences sur données simulées pour attester numériquement de la convergence de notre méthode d'estimation du représentant moyen d'un échantillon de surfaces.

Chapitre 2

État de l'art

2.1 Représentation des surfaces par des courants

2.1.1 Pourquoi utiliser des courants ?

Comme évoqué en Introduction, on prend le parti, pour pallier les difficultés liées au caractère riemannien des surfaces, de plonger ces dernières dans un espace euclidien. À cet effet, on utilise le cadre introduit en 2005 par Glaunès et Vaillant : étant donné un espace de Hilbert W (espace fonctionnel) judicieusement choisi, on attribue à chaque surface un représentant dans le dual topologique W^* de cet espace. Nous allons voir, à travers l'examen précis de ce cadre puis de celui, analogue en de nombreux points, développé par Durrleman, que cette représentation présente deux intérêts majeurs. D'une part, elle permet de faire, via les représentants dans l'espace euclidien W^* , toutes les opérations usuelles qui ne sont pas définies directement sur les surfaces. D'autre part, le courant associé à une surface garde en mémoire l'action sur cette dernière de chaque fonction élément de W et, par conséquent, contient la majeure partie de l'information géométrique qui la caractérise.

2.1.2 Cadre de représentation de Glaunès et Vaillant

L'idée de la représentation des surfaces par des courants introduite en 2005 par Joan Glaunès et Marc Vaillant est initialement vouée à faciliter la résolution de problèmes de recalage de surfaces. Dans [33], l'idée directrice est de dépasser l'absence de distance euclidienne sur les ensembles de surfaces et, à travers la représentation via un espace euclidien de courants, notion empruntée à la théorie des distributions, pouvoir déterminer "combien une surface est proche d'une autre".

Le cadre de représentation des surfaces repose sur l'utilisation d'un espace de Hilbert W dont les éléments sont des 2-formes à support compact sur \mathbb{R}^3 . Une 2-forme différentielle est une application différen-

tielle w qui, à tout élément x de \mathbb{R}^3 , associe une forme bilinéaire et antisymétrique $w(x)$ définie sur \mathbb{R}^3 . Une surface orientée lisse S est représentée par l'application S définie par

$$\begin{aligned} S : W &\longrightarrow \mathbb{R} \\ w &\longrightarrow \int_S w(x)(u_x^1, u_x^2) d\sigma(x), \end{aligned} \quad (2.1)$$

où $d\sigma$ est l'élément d'aire sur S et (u_x^1, u_x^2) une base orthonormale du plan tangent à S au point x . Le "courant" associé à la surface S est ainsi une forme linéaire sur l'espace W des 2-formes. Cette représentation est d'autant plus simple qu'il existe une relation de dualité entre les 2-formes et les champs de vecteurs : si w est une 2-forme différentielle sur \mathbb{R}^3 , il existe un unique champ de vecteurs différentiable \bar{w} sur \mathbb{R}^3 tel que

$$w(x)(\eta, \nu) = \bar{w}(x).(\eta \wedge \nu), \forall x, \eta, \nu \in \mathbb{R}^3, \quad (2.2)$$

où \cdot et \wedge désignent respectivement le produit scalaire et le produit vectoriel dans \mathbb{R}^3 . Pour peu que l'espace W soit construit à partir d'un espace de Hilbert à noyau reproduisant (notion abordée plus loin en Section 2.2), la forme linéaire S donnée par (2.1) sera systématiquement continue. L'espace des représentants sera ainsi le dual topologique W^* de W et, par conséquent, aura une structure d'espace de Hilbert, avec une norme et un produit scalaire. Via cet espace normé des représentants, on pourra ainsi comparer les surfaces : la distance entre deux surfaces étant la distance entre leurs représentants au sens de l'espace W^* . Remarquons que l'utilisation pour l'espace W d'un espace de 2-formes présente un autre intérêt majeur : les représentants des surfaces présentent une sorte de stabilité sous l'action d'un groupe de déformations. Dans cette optique, l'action $\#$ d'une déformation ϕ sur une 2-forme w est définie par

$$\phi\#w(x)(\eta, \nu) = w(\phi(x))((d_x\phi).\eta, (d_x\phi).\nu), \forall x, \eta, \nu \in \mathbb{R}^3, \quad (2.3)$$

où $(d_x\phi)$ désigne la matrice jacobienne de ϕ . Il est démontré, via les règles de changements de coordonnées sur les formes différentielles, que le représentant $\phi.S$ dans W^* de la surface déformée $\phi(S)$ s'écrit

$$\phi.S(w) = S(\phi\#w), \forall w \in W. \quad (2.4)$$

L'application principale de ce cadre est le recalage de surfaces dans un modèle de type "large deformation".

Dans [17], le champ d'applications de la représentation par courants est élargi dans une visée nettement plus proche de la nôtre. Il y est proposé une méthode de construction d'atlas dans un modèle de prototype déformable, à partir d'un échantillon de surfaces images par des difféomorphismes inconnus d'un prototype,

sorte de "surface moyenne", inconnu. Notre étude se rapproche de cet article dans le sens où elle visera également à établir, à partir d'une population de surfaces, un caractère moyen. La différence principale étant que nous ne construirons pas un modèle de prototype déformable mais un modèle d'observation, décrivant les variations des surfaces via un bruit aléatoire affectant leurs représentants. On se fixera le but de trouver le représentant moyen via une méthode d'estimation paramétrique classique. Ce point de vue nous rapproche davantage des travaux de Durrleman, évoqués ci-après.

2.1.3 Cadre de représentation de Durrleman

En 2007, Durrleman utilise un cadre de représentation des surfaces analogue à celui de Glaunès en de nombreux points, proposant également pour sa part une représentation des courbes plongées dans \mathbb{R}^2 . Etant donné un espace W de champs de vecteurs sur \mathbb{R}^3 , le représentant S d'une surface \mathcal{S} est donné par

$$\begin{aligned} S : W &\longrightarrow \mathbb{R} \\ w &\longrightarrow \int_{\mathcal{S}} w(x) \cdot (u_x^1 \wedge u_x^2) d\sigma(x), \end{aligned} \tag{2.5}$$

et le représentant L d'une courbe \mathcal{L} par

$$\begin{aligned} L : W &\longrightarrow \mathbb{R} \\ w &\longrightarrow \int_{\mathcal{L}} w(x) \cdot \tau(x) dx, \end{aligned} \tag{2.6}$$

où $\tau(x)$ désigne la tangente à la courbe \mathcal{L} au point x , les autres notations étant similaires à celles définies dans la Section 2.1.2. Bien que cette définition du représentant d'une surface, utilisée ultérieurement par Glaunès et alii dans [29], soit à première vue différente de celle évoquée précédemment (2.1), elle est équivalente compte tenu de la dualité entre 2-formes et champs vectoriels (2.2). Pour peu que W soit défini à partir d'un espace de Hilbert à noyau reproduisant, les représentants des courbes et des surfaces sont, ici aussi, des éléments du dual topologique de W . D'un point de vue pratique, Durrleman utilise comme noyau reproduisant associé à l'espace W le noyau gaussien K^W de la forme

$$K^W(x, y) = \exp(-\lambda_W^2 \|x - y\|_{\mathbb{R}^3}^2), \quad \forall x, y \in \mathbb{R}^3,$$

où λ_W est le paramètre d'échelle associé à l'espace de champs de vecteurs W . En utilisant la propriété de reproduction et la dualité liée au Théorème de représentation de Riesz-Fréchet, on peut exprimer les courants de l'espace W^* dans la base des "delta-Dirac". Les courants de type "delta-Dirac" (δ_x^α) se définissent comme

suit

$$\begin{aligned}\delta_x^\alpha : W &\longrightarrow \mathbb{R} \\ w &\longrightarrow {}^t \alpha \cdot w(x),\end{aligned}$$

où $\alpha, x \in \mathbb{R}^3$. L'espace W est l'adhérence de l'ensemble des combinaisons linéaires de "delta-Dirac". Dans cette optique, on approche le courant S représentant la surface par le "courant discret"

$$\tilde{S}(w) = \sum_k \delta_{x_k}^{\alpha_k},$$

où $(x_k)_k$ est l'ensemble des barycentres des faces de l'approximation discrète de S et α_k la normale à la surface discrète au point x_k dont la norme vaut l'aire de la face dont x_k est barycentre. La représentation approximée \tilde{S} converge vers le représentant S dans l'espace des courants quand la discrétisation de la surface devient de plus en plus fine. Une représentation analogue approximée peut être définie pour les courbes.

Le produit scalaire entre deux courants "delta-Dirac" est simple à exprimer :

$$\langle \delta_x^\alpha, \delta_y^\beta \rangle_{W^*} = {}^t \alpha \cdot K^W(x, y) \beta,$$

et, par bilinéarité, le produit scalaire entre deux courants $S = \sum \delta_{x_k}^{\alpha_k}$ et $T = \sum \delta_{y_l}^{\beta_l}$ est alors donné par :

$$\langle S, T \rangle_{W^*} = \sum_k \sum_l {}^t \alpha_k \cdot K^W(x_k, y_l) \cdot \beta_l.$$

La structure d'espace euclidien de W^* induit sur les formes géométriques, courbes ou surfaces, la distance :

$$d(S, T) = \|S - T\|_{W^*},$$

où S et T sont deux courbes ou surfaces de représentants respectifs S et T . Ce modèle de représentation permet de "capter" des variations géométriques à une échelle unique qui est déterminée par la constante λ_W . C'est en cela que réside le principal défaut de la méthode : le fait qu'elle ne permet pas l'analyse multi-échelles.

C'est ce cadre de représentation qui est utilisé dans une visée d'application à l'analyse statistique des courbes et surfaces dans [10]. Les auteurs y présentent un modèle intégrant, via les représentants, une variabilité aléatoire entre les surfaces ; cette variabilité est modélisée en utilisant une méthode de génération de vecteurs aléatoires gaussiens sur un espace de Hilbert. En définissant l'action $*$ d'une déformation de façon totalement analogue à l'action \sharp chez Glaunès 2.3 et en prolongeant cette dernière sur les "delta-Dirac",

il est établi sur les courants qui représentent un échantillon de courbes ou de surfaces le modèle de variation

$$T_k = \phi_k * \bar{T} + \epsilon_k, \quad (2.7)$$

où les T_k sont les courants représentant les formes qui constituent l'échantillon, les ϕ_k des déformations et les ϵ_k des termes d'erreur, définis comme des variables aléatoires sur W^* ; le prototype \bar{T} et les ϕ_k sont inconnus tandis que les T_k sont observés. Les variables aléatoires sur l'espace de Hilbert W^* sont définies par la correspondance suivante entre vecteurs de cet espace et variables aléatoires de carré intégrable : à tout vecteur $V = \sum_{k \in \mathbb{N}} v_k e_k \in W^*$, on associe la variable aléatoire $\sum_{k \in \mathbb{N}} v_k \zeta_k$, où les ζ_k sont des variables aléatoires i.i.d. de loi $\mathcal{N}(0, 1)$.

Durrleman et alii donnent un algorithme d'estimation des "paramètres" du modèle que sont le prototype \bar{T} de l'échantillon de formes linéaires et les déformations ϕ_k . La connaissance d'une approximation des fonctions ϕ_k en sortie d'algorithme fait qu'il est possible de "visualiser" la surface correspondant au prototype de l'échantillon. Afin de pouvoir traiter les données "lourdes" que sont les représentants des surfaces, éléments du dual d'un espace fonctionnel, il est également mis en place un algorithme de représentation parcimonieuse des courants via une base formée de "delta-dirac". Cet algorithme approxime à une précision fixée en entrée et ne donne pas en sortie un nombre fixe de composantes pour l'ensemble de l'échantillon.

Ce modèle est prolongé, entre autres, dans [9]. Durrleman, Trouvé et alii appliquent ici un procédé similaire, dérivé de ce qui a été fait pour les courbes, à des objets de type "amas de fibres", ce qui est naturellement idéal pour de l'imagerie 3-D haute résolution. La procédure développée dans ce papier permet d'estimer le prototype d'un échantillon d'amas de fibres et effectuer le problème de recalage sans individualiser les fibres. Chaque fibre représentée, à la manière d'une courbe discrétisée, par un courant, tous les courants sont sommables, qu'ils viennent ou non de fibres extraites du même amas. Cet article accomplit, de fait, un travail beaucoup plus détaillé que [10] sur les propriétés des sommes de courants et les répercussions sur l'objet représenté d'une opération sur le représentant.

2.1.4 Contributions de notre étude par rapport aux cadres originaux de Glaunès et Durrleman

Comme évoqué précédemment, notre étude s'inscrit dans une optique d'analyse statistique de surfaces. On se rapproche ainsi plus des travaux de Durrleman évoqués en 2.1.3 que de ceux de Glaunès évoqués en 2.1.2, quant à l'utilisation du cadre de représentation par courants. On se démarque néanmoins de ces deux auteurs sur plusieurs points. Tout d'abord, on mettra en place une représentation des surfaces basée sur un cadre aléatoire dans lequel les représentants ne sont plus simplement des formes linéaires mais des réalisations de "formes linéaires aléatoires", concept que l'on introduira en se basant sur la notion de pro-

cessus stochastique généralisé (voir Section 2.3). Ensuite, on optera pour une construction dans laquelle les représentants des surfaces s'expriment, du moins de façon approximée, en fonction d'un nombre fini d'éléments, fixé par avance de manière arbitraire et donc constant sur l'échantillon. Ceci constitue un changement par rapport aux cadres existants, où les éléments de W^* s'expriment dans une base dont le nombre de composantes dépend de la précision des discrétisations (ce qui le rend variable au sein de l'échantillon). Ce nombre varie aussi en fonction du paramètre de précision du noyau. On utilisera à cet effet une base d'un sous-espace vectoriel de dimension finie fixée et on identifiera les représentants des surfaces à leur projection sur ce sous-espace. Notre méthode d'approximation présente un autre avantage du fait de son aspect multi-échelle. Enfin, alors que les travaux de Glaunès et Durrleman ne traitent que de la représentation des surfaces discrètes via des courants exprimés dans une base discrétisée de "delta-Dirac", on étudiera en détail dans ce manuscrit le passage du continu au discret. On définit tout d'abord notre représentation par courants pour des surfaces continues puis on l'étend à des approximations discrètes de ces dernières. On traitera le problème de la convergence du représentant de l'approximation discrète vers le représentant de la surface continue initiale.

2.2 Espaces de Hilbert à noyau reproduisant

2.2.1 Théorie des espaces de Hilbert à noyau reproduisant

Comme dans les travaux de Durrleman et Glaunès, on construira l'espace des représentants des surfaces à partir d'un espace de Hilbert à noyau reproduisant dans le but de garantir de bonnes propriétés. La notion d'espace de Hilbert à noyau reproduisant (terme qui sera, par la suite, désigné par le sigle anglophone R.K.H.S.) a été introduite dans les années 1950. N. Aronszajn en est le principal théoricien avec l'article "The Theory of Reproducing Kernels" [4].

Définition 2. Soit F un espace de Hilbert sur \mathbb{R} (respectivement sur \mathbb{C}) dont les éléments sont des fonctions définies sur un sous-ensemble Ω de \mathbb{R}^p et à valeurs dans \mathbb{R} (respectivement dans \mathbb{C}). On notera $\langle \cdot, \cdot \rangle_F$ le produit scalaire de l'espace F et $\|\cdot\|_F$ la norme induite par ce dernier. Une fonction K_0 définie sur Ω^2 et à valeurs dans \mathbb{R} (respectivement dans \mathbb{C}) est appelée un noyau reproduisant si et seulement si elle vérifie les deux propriétés suivantes :

(1) pour tout y de Ω , la fonction f_y définie par

$$\begin{aligned} f_y : \Omega &\longrightarrow \mathbb{R} \text{ (resp. } \mathbb{C}) \\ x &\longrightarrow K_0(x, y) \end{aligned}$$

est un élément de F ;

(2) pour tout y de Ω , pour tout f de F ,

$$f(y) = \langle f, f_y \rangle_F.$$

La deuxième propriété est généralement nommée "reproduction". On dit alors que F est un "espace de Hilbert à noyau reproduisant" (ou R.K.H.S.).

Dans la suite, on donne quelques propriétés élémentaires des R.K.H.S. ; on se bornera, pour des raisons liées aux besoins de notre étude, au cas des R.K.H.S. sur \mathbb{R} et donc des noyaux à valeurs réelles. On conserve, ci-après, les notations données dans la définition : F désigne ainsi un R.K.H.S. sur \mathbb{R} dont les fonctions sont définies sur un sous-ensemble connexe Ω de \mathbb{R}^p et K_0 un noyau reproduisant pour l'espace F .

Tout d'abord, on peut affirmer que tout noyau reproduisant sur un \mathbb{R} -espace de Hilbert est commutatif. En effet, pour tous x et y de Ω , on a par propriété de reproduction :

$$\begin{aligned} K_0(x, y) &= f_y(x) = \langle f_y, f_x \rangle_F \\ &= \langle f_x, f_y \rangle_F \\ &= K_0(y, x). \end{aligned}$$

Une conséquence immédiate est que dans l'espace F de noyau reproduisant K_0 on a, pour tout x de Ω :

$$K_0(x, \cdot) = K_0(\cdot, x) = f_x.$$

Ensuite, on peut aisément démontrer l'unicité du noyau sur un R.K.H.S. ; en effet, si H_0 est un noyau sur F , on a, en définissant pour tout x de Ω la fonction $g_x : y \rightarrow H_0(y, x)$,

$$\begin{aligned} \forall x, y \in \Omega, |K_0(x, y) - H_0(x, y)| &= |K_0(y, x) - H_0(x, y)| \\ &= |f_x(y) - g_y(x)| \\ &= |\langle f_x, g_y \rangle_F - \langle g_y, f_x \rangle_F| \\ &= 0, \end{aligned}$$

d'où le fait que $H_0 = K_0$.

Dans un R.K.H.S. à noyau borné, on peut toujours contrôler les valeurs d'une fonction par la norme hilbertienne, via la propriété suivante, qui sera essentielle au Chapitre 3 pour démontrer la continuité des représentants des surfaces.

Proposition 1 (Contrôle de la borne supérieure par la norme). *Soit F un R.K.H.S. de noyau K_0 . Si K_0 est borné, alors*

il existe $C > 0$ tel que, pour tout élément w de F ,

$$\sup_x |w(x)| \leq C \|w\|_F.$$

Démonstration. Pour tout $w \in F$, pour tout x , on a

$$|w(x)| = |\langle f_x, w \rangle_F| \leq \|w\|_F \|f_x\|_F \leq \|w\|_F |K_0(x, x)|.$$

Le noyau K_0 étant borné, le résultat est démontré. □

Remarque 1. La démonstration de cette proposition exhibe une valeur possible de la constante

$$C = \sup_{x \in \mathbb{D}} |K_0(x, x)| = \sup_{x \in \mathbb{D}} \|K_0(x, \cdot)\|_F.$$

On donne maintenant la propriété fondamentale de caractérisation des R.K.H.S., ainsi que sa preuve.

Théorème 1 (Aronszajn). Soit F un \mathbb{R} -espace de Hilbert dont les éléments sont des fonctions définies sur un sous-ensemble connexe Ω de \mathbb{R}^p . Il existe un noyau reproduisant sur F si et seulement si, pour tout x de Ω , la fonction δ_x , définie par

$$\begin{aligned} \delta_x : F &\longrightarrow \mathbb{R} \\ f &\longrightarrow f(x) \end{aligned}$$

est continue sur F .

Démonstration. Si F admet un noyau reproduisant K_0 , on, pour tout x de Ω , en conséquence de la propriété de reproduction et de l'inégalité de Cauchy-Schwarz,

$$\forall f \in F, |\delta_x(f)| \leq \|f\|_F \|f_x\|_F,$$

d'où la continuité de δ_x sur F . Réciproquement, si δ_x est continue sur F pour tout x de Ω , δ_x admet, en tant que forme linéaire continue sur F , un représentant de Riesz dans F que l'on notera r_x . On pose alors

$$K_0(x, y) = r_x(y), \quad \forall x, y \in \Omega.$$

Il est aisé de vérifier que K_0 est bien un noyau reproduisant sur F . □

2.2.2 Exemples de RKHS

Espace de Hardy

Soit \mathbb{D} le disque-unité ouvert de \mathbb{R}^p et $H^2(\mathbb{D})$ l'espace, dit de Hardy, des fonctions analytiques sur \mathbb{D} à valeurs réelles dont le développement en série entière est de carré sommable, *i.e.*

$$H^2(\mathbb{D}) = \left\{ f : \mathbb{D} \rightarrow \mathbb{R} \mid f(x) = \sum_{k \in \mathbb{N}} a_k x^k, \sum_{k \in \mathbb{N}} a_k^2 < \infty \right\}.$$

L'espace $H^2(\mathbb{D})$ est un R.K.H.S. pour le produit scalaire

$$\langle f, g \rangle_{H^2} = \sum_{k \in \mathbb{N}} a_k b_k,$$

où $f(x) = \sum_{k \in \mathbb{N}} a_k x^k$ et $g(x) = \sum_{k \in \mathbb{N}} b_k x^k$, et le noyau reproduisant suivant, dit "noyau de Szego" :

$$K_{H^2}(x, y) = \sum_{k \in \mathbb{N}} x^k y^k = \frac{1}{1 - xy}.$$

Espace de Sobolev

On pose $\mathcal{D} = [-\pi, \pi]^3$, $m \geq 1$ et, pour tout $k = (k_1, k_2, k_3)$ de \mathbb{N}^3 , $\lambda_k = \frac{1}{(1+|k|^2)^m}$, où $|k| = k_1 + k_2 + k_3$.

Soient $(\phi_k)_{k \in \mathbb{N}^3}$ et $(\psi_k)_{k \in \mathbb{N}^3}$ les suites de fonctions définies par :

$$\begin{aligned} \phi_k : \mathcal{D} &\longrightarrow \mathbb{R} \\ x &\longrightarrow \frac{1}{\sqrt{2\pi}^3} \cos(\langle k, x \rangle_{\mathbb{R}^3}), \forall x \in \mathcal{D}, \\ \psi_k : \mathcal{D} &\longrightarrow \mathbb{R} \\ x &\longrightarrow \frac{1}{\sqrt{2\pi}^3} \sin(\langle k, x \rangle_{\mathbb{R}^3}), \forall x \in \mathcal{D}. \end{aligned}$$

On considère à présent l'ensemble de fonctions :

$$V_0 = \left\{ u : \mathcal{D} \longrightarrow \mathbb{R} \mid \exists (f_k)_{k \in \mathbb{N}^3}, (g_k)_{k \in \mathbb{N}^3} \subset \mathbb{R}, \sum_{k \in \mathbb{N}^3} \frac{|f_k|^2 + |g_k|^2}{\lambda_k} < \infty, u(x) = \sum_{k \in \mathbb{N}^3} (f_k \phi_k(x) + g_k \psi_k(x)), \forall x \in \mathcal{D} \right\}.$$

Remarque 2. On a les faits suivants :

- Si $u \in V_0$, alors $\sum_{k \in \mathbb{N}^3} |f_k| < \infty$ et $\sum_{k \in \mathbb{N}^3} |g_k| < \infty$.
- Les éléments de V_0 sont bien définis ponctuellement.

- Les éléments de V_0 sont des fonctions carré-intégrables.

- Les éléments de V_0 sont des fonctions C^1 .

On définit sur V_0 le produit scalaire

$$\langle u, v \rangle_{V_0} = \sum_{k \in \mathbb{N}^3} \frac{f_k h_k}{\lambda_k} + \sum_{k \in \mathbb{N}^3} \frac{g_k l_k}{\lambda_k},$$

où $u = \sum_{k \in \mathbb{N}^3} f_k \phi_k + \sum_{k \in \mathbb{N}^3} g_k \psi_k$, et $v = \sum_{k \in \mathbb{N}^3} h_k \phi_k + \sum_{k \in \mathbb{N}^3} l_k \psi_k$.

Avec le noyau K_0 , défini par

$$\begin{aligned} K_0 : \mathcal{D}^2 &\longrightarrow \mathbb{R} \\ (x, y) &\longrightarrow \sum_{k \in \mathbb{N}^3} \lambda_k (\phi_k(x) \phi_k(y) + \psi_k(x) \psi_k(y)), \end{aligned}$$

l'espace $(V_0, \langle \cdot, \cdot \rangle_{V_0})$ est un R.K.H.S. ; de plus, le noyau K_0 est borné sur \mathcal{D}^2 .

Cet espace sera utilisé, entre autres, dans le cadre d'expériences au Chapitre 6.

2.3 Processus stochastiques généralisés

2.3.1 Fonctionnelles aléatoires, concept de processus généralisé

Comme cela a été dit en Section 2.1.4, on utilisera, pour donner à l'espace des représentants des surfaces un cadre probabiliste sous-jacent, des processus aléatoires généralisés. Cette notion a été théorisée par Itô [20, 21] et par Gel'fand et Vilenkin [15] dans les années 1970 ; sa construction combine théories des distributions et probabilités. Les processus généralisés sont assez fréquemment utilisés dans la littérature associée au traitement du signal [32, 11, 26, 25].

Variabiles aléatoires, égalité et convergence

Les processus aléatoires généralisés sont définis comme des distributions indexées sur un espace fonctionnel et à valeurs dans un ensemble de variables aléatoires. Il est donc nécessaire d'établir de façon claire des critères d'égalité et de convergence pour les variables aléatoires, pour pouvoir construire une topologie sur l'ensemble des processus aléatoires généralisés.

On considère que deux variables aléatoires X et Y sont égales si et seulement si, pour tous réels a et b ,

$$\mathbb{P}(a \leq X < b, a \leq Y < b) = \mathbb{P}(a \leq X < b).$$

On se donne le critère de convergence faible suivant : une suite de variables aléatoires $(X_n)_{n \in \mathbb{N}}$ converge vers une variable aléatoire X si et seulement si, pour toute fonction bornée f de \mathbb{R}^2 dans \mathbb{R} et pour toute variable aléatoire Z ,

$$\int f(x, y) d\mathbb{Q}_n(x, y) \xrightarrow{n \rightarrow \infty} \int f(x, y) d\mathbb{Q}(x, y),$$

où \mathbb{Q}_n est la loi du couple (X_n, Z) et \mathbb{Q} celle du couple (X, Z) .

Quelques définitions

Définition 3 (fonctionnelle aléatoire). *On considère l'ensemble K des fonctions C^∞ à support compact de \mathbb{R}^q dans \mathbb{R}^p . On définit une fonctionnelle aléatoire Φ sur K par la donnée, pour tout f de K , d'une variable aléatoire que l'on note $\Phi(f)$.*

On dit qu'une fonctionnelle aléatoire Φ est continue si, pour tout entier naturel non nul N , pour tous éléments f_1, \dots, f_N de K , pour toutes suites $(f_1^k)_{k \in \mathbb{N}}, \dots, (f_N^k)_{k \in \mathbb{N}}$ telles que f_1^k, \dots, f_N^k convergent respectivement vers f_1, \dots, f_N dans l'espace K quand k tend vers l'infini, on a :

$$\lim_{k \rightarrow \infty} \Phi(f_i^k) = \Phi(f_i), \forall i \in \{1, \dots, N\},$$

au sens de la convergence des variables aléatoires.

Définition 4 (Processus aléatoire généralisé). *Une fonctionnelle aléatoire continue sur K est appelée fonction aléatoire généralisée. Une fonction aléatoire à valeurs dans \mathbb{R} est appelée processus aléatoire généralisé.*

Soient Φ et Ψ deux processus aléatoires généralisés. On définit leur somme $\Phi + \Psi$ comme le processus aléatoire généralisé qui à toute fonction f de K associe la variable aléatoire $\Phi(f) + \Psi(f)$. Pareillement, pour tout réel λ , on définit le processus aléatoire $\lambda\Phi$ comme le processus aléatoire généralisé qui à tout élément f de K associe la variable aléatoire généralisée $\lambda\Phi(f)$. Ces deux opérations donnent à l'ensemble des processus aléatoires généralisés une structure d'espace vectoriel.

Définition 5 (Linéarité). *Le processus aléatoire généralisé Φ est dit linéaire si pour tous éléments f et g de K , on a*

$$\Phi(\lambda f + g) = \lambda\Phi(f) + \Phi(g),$$

au sens de l'égalité p.s. entre variables aléatoires. C'est ce concept de processus aléatoire linéaire qui servira à concevoir les formes linéaires aléatoires telles que nous les introduirons au Chapitre 3.

2.3.2 Moments d'un processus aléatoire généralisé

Définition 6 (Moyenne). Soit Φ un processus aléatoire généralisé tel que, pour tout élément f de K , la variable aléatoire $\Phi(f)$ admet un moment d'ordre 1. On appelle moyenne du processus Φ l'application m_Φ définie par

$$\begin{aligned} m_\Phi : K &\longrightarrow \mathbb{R} \\ f &\longrightarrow \mathbb{E}[\Phi(f)]. \end{aligned}$$

Remarquons que m_Φ est une forme linéaire sur K .

Définition 7 (Corrélation). Soit Φ un processus aléatoire généralisé tel que, pour tous f et g de K , l'espérance $\mathbb{E}[\Phi(f)\Phi(g)]$ existe. On appelle corrélation du processus Φ l'application B_Φ définie par

$$\begin{aligned} B_\Phi : K \times K &\longrightarrow \mathbb{R} \\ (f, g) &\longrightarrow \mathbb{E}[\Phi(f)\Phi(g)]. \end{aligned}$$

On remarque que B_Φ est une forme bilinéaire semi-définie-positif sur $K \times K$.

Définition 8 (Autocovariance). Si la corrélation d'un processus généralisé Φ existe, on définit son autocovariance comme l'application

$$\begin{aligned} C_\Phi : K \times K &\longrightarrow \mathbb{R} \\ (f, g) &\longrightarrow B_\Phi(f, g) - m_\Phi(f)m_\Phi(g). \end{aligned}$$

Notons que cette définition équivaut à

$$C_\Phi(f, g) = \text{Cov}(\Phi(f), \Phi(g)), \forall f, g \in K.$$

L'autocovariance est également une forme bilinéaire semi-définie-positif sur $K \times K$.

Définition 9 (Moment d'ordre N). Suivant le même principe, le moment d'ordre N d'un processus aléatoire généralisé est défini, pour peu que $\mathbb{E}[\Phi(f_1)\dots\Phi(f_N)]$ existe pour tout entier naturel non nul N et pour tous f_1, \dots, f_N de K , comme l'application

$$\begin{aligned} M_\Phi^N : K^N &\longrightarrow \mathbb{R} \\ (f_1, \dots, f_N) &\longrightarrow \mathbb{E}[\Phi(f_1)\dots\Phi(f_N)]. \end{aligned}$$

2.3.3 Processus gaussiens

Définition 10 (Processus gaussien). Un processus aléatoire généralisé Φ est dit gaussien si, pour tout entier N strictement positif et pour tous éléments f_1, \dots, f_N de K linéairement indépendants, le vecteur aléatoire $(\Phi(f_1), \dots, \Phi(f_N))$ est un vecteur gaussien.

Compte tenu des connaissances générales sur les vecteurs gaussiens, un processus aléatoire généralisé gaussien est univoquement défini par la donnée de ses fonctions de moyenne et de corrélation. Il existe une correspondance entre les processus aléatoires gaussiens et les couples (m, B) formés d'une forme linéaire continue m sur K et d'une forme bilinéaire continue semi-définie-positive B sur K . Cette correspondance est assurée par le théorème suivant.

Théorème 2 (Gel'fand - Vilenkin). *Soit m une forme linéaire continue sur K et soit B une forme bilinéaire continue définie positive sur K . Il existe un processus aléatoire gaussien de moyenne m et de corrélation B si et seulement si la forme bilinéaire C définie par*

$$\begin{aligned} C : K \times K &\longrightarrow \mathbb{R} \\ (f, g) &\longrightarrow B(f, g) - m(f)m(g) \end{aligned}$$

est semi-définie-positive.

Chapitre 3

Modèle d'observation : du modèle continu au modèle discret

3.1 Objectifs

Ainsi que cela a été mentionné en introduction de cette partie, (*c.f.* 1), le problème de l'estimation de l'élément moyen d'un échantillon de surfaces est une thématique particulièrement intéressante, notamment en raison des applications possibles au traitement d'image. Ce problème est, nous l'avons vu précédemment, difficile car les surfaces ne sont pas éléments d'un quelconque espace vectoriel. En conséquence, elles ne peuvent être soumises aux outils et algorithmes usuels de l'analyse statistique, initialement conçus pour opérer sur les espaces euclidiens, sur lesquels on peut définir et calculer sommes, moyennes algébriques, produits scalaires, distances... En Section 2.1.2, nous avons constaté qu'un moyen efficace de contourner cette difficulté est de représenter les surfaces par des formes linéaires sur un espace fonctionnel de type Hilbert, les plongeant ainsi dans une structure d'espace euclidien. Cette représentation, développée entre autres par Glaunès [33, 17] et Durrleman [10, 9], est déjà très utilisée en traitement d'image, 2-D comme 3-D, en particulier pour résoudre des problèmes de recalage sur des courbes ou des surfaces. Dans le présent chapitre, on va l'utiliser pour décrire la variabilité d'échantillons de surfaces autour d'une moyenne à travers un modèle d'observation.

Au vu des ressources existant actuellement dans la littérature, nos visées sont originales ; on va donc devoir développer des réponses spécifiques. Tout d'abord on met en place, pour pouvoir modéliser les variations aléatoires des surfaces d'un échantillon et les visualiser à travers leurs représentants, un cadre probabiliste dans lequel les courants sont vus comme des réalisations de formes linéaires aléatoires. On définit le concept de forme linéaire aléatoire à partir de la théorie des processus stochastiques généralisés de

Gel'fand et Vilenkin, évoquée en Section 2.3. Ce cadre aléatoire sera d'abord utilisé pour décrire les variations des représentants d'un échantillon de surfaces continument différentiables autour d'un élément moyen. Ensuite, on définit une procédure d'approximation des surfaces lisses par des surfaces discrètes formées de facettes triangulaires et on étend la représentation à ces dernières. Ceci conduit à la construction d'un modèle d'observation adapté à l'étude des variations aléatoires de surfaces issues des données expérimentales. On démontrera un théorème établissant la convergence du représentant de l'approximation discrète vers le représentant de la surface lisse originelle. Cela garantit la cohérence de notre modèle construit en allant du continu vers le discret.

3.2 Modèle statistique

3.2.1 Cadre déterministe de représentation des surfaces

Soient \mathcal{D} un cube fermé et borné de \mathbb{R}^3 et Θ un ensemble de surfaces compactes incluses dans \mathcal{D} et continument différentiables. Soit V_0 un sous-espace vectoriel de $\mathcal{L}^2(\mathcal{D}, \mathbb{R}) \cap C^1(\mathcal{D}, \mathbb{R})$ qui soit un R.K.H.S. dont le noyau K_0 est borné sur $\mathcal{D} \times \mathcal{D}$. On considère l'espace

$$V = V_0^3$$

qui est clairement un espace de Hilbert pour le produit scalaire $\langle \cdot, \cdot \rangle_V$ défini par

$$\langle v, w \rangle_V = \sum_{i=1}^3 \langle v_i, w_i \rangle_{V_0},$$

où $\langle \cdot, \cdot \rangle_{V_0}$ est le produit scalaire usuel de l'espace V_0 et v, w sont deux éléments de V dont les composantes dans V_0 sont respectivement v_1, v_2, v_3 et w_1, w_2, w_3 . On désignera par $\|\cdot\|_V$ la norme induite par ce produit scalaire.

On représente toute surface $S \in \Theta$ par la forme linéaire s définie par

$$\begin{aligned} s : V &\longrightarrow \mathbb{R} \\ w &\longrightarrow s(w) \end{aligned}$$

avec

$$s(w) = \int_S \langle w(x), e_1(x) \wedge e_2(x) \rangle_{\mathbb{R}^3} d\sigma(x), \forall w \in V, \quad (3.1)$$

où \wedge est le produit vectoriel sur \mathbb{R}^3 , $\langle \cdot, \cdot \rangle_{\mathbb{R}^3}$ le produit scalaire euclidien de \mathbb{R}^3 , σ l'élément d'aire sur S , et les

vecteurs $e_1(x), e_2(x)$ forment une base orthonormale directe du plan tangent à S au point x .

Sous nos hypothèses, le représentant s est un élément de V^* . Sa continuité est aisée à démontrer; en effet, si w est un élément de V dont les composantes dans V_0 sont w_1, w_2, w_3 , on a

$$\begin{aligned} |s(w)| &= \left| \int_S \langle w(x), e_1(x) \wedge e_2(x) \rangle_{\mathbb{R}^3} d\sigma(x) \right| \\ &\leq \int_S \|e_1(x) \wedge e_2(x)\|_{\mathbb{R}^3} \|w\|_{\infty} d\sigma(x) \\ &\leq \|w\|_{\infty} \mathcal{A}(S), \end{aligned}$$

où \mathcal{A} désigne la mesure d'aire sur les surfaces. Par ailleurs, au vu de la Proposition 1, il existe $C > 0$ tel que

$$\begin{aligned} \|w\|_{\infty}^2 &\leq \sum_{i=1}^3 \|w_i\|_{\infty}^2 \\ &\leq \sum_{i=1}^3 C^2 \|w_i\|_{V_0}^2 \\ &\leq C^2 \|w\|_V^2. \end{aligned}$$

On a donc

$$|s(w)| \leq C \|w\|_V \mathcal{A}(S),$$

ce qui assure la continuité de s sur V .

3.2.2 Formes linéaires aléatoires

Dans cette section, on introduit une notion de forme linéaire aléatoire, inspirée des travaux de Itô [20, 21], Fernique [12] ou encore Gel'fand et Vilenkin [15] sur les processus aléatoires généralisés (*c.f.* Section 2.3). On donne quelques définitions et propriétés essentielles des formes linéaires aléatoires puis on met en place sur ces dernières une structure d'espace vectoriel normé.

Définitions

Soit $\mathcal{E} = (\Omega, \mathcal{F}, \mathbb{P})$ un espace probabilisé. On note $L^2(\mathcal{E})$ l'ensemble des variables aléatoires de carré intégrable définies sur \mathcal{E} . On désigne toujours par V l'espace de champs de vecteurs défini en Section 3.2.1.

Définition 11 (Forme linéaire aléatoire). *On appelle forme linéaire aléatoire sur V une application linéaire continue définie sur V et à valeurs dans $L^2(\mathcal{E})$.*

On note $V_{\mathcal{E}}^*$ l'ensemble des formes linéaires aléatoires sur l'espace V . Les processus généralisés introduits par Itô ou Gel'fand et Vilenkin correspondent à des applications de l'ensemble des fonctions C^{∞} à support compact vers un ensemble de variables aléatoires plus général que $L^2(\mathcal{E})$ (voir Section 2.3). Ici,

les formes linéaires aléatoires sont définies sur l'espace V , produit cartésien de R.K.H.S. à noyau borné. Le choix de limiter l'ensemble de variables aléatoires considéré à $L^2(\mathcal{E})$ permet, d'une part, de définir une norme sur $V_{\mathcal{E}}^*$ et, d'autre part, de garantir l'existence et la continuité de l'espérance et de l'autocovariance définies ci-dessous.

Définition 12 (Espérance et autocovariance). *Soit S une forme linéaire aléatoire sur V . L'espérance $\mathbb{E}^*[S]$ de S est définie par*

$$\begin{aligned}\mathbb{E}^*[S] : V &\longrightarrow \mathbb{R} \\ w &\longrightarrow \mathbb{E}[S(w)]\end{aligned}$$

et son autocovariance γ_S est définie par

$$\begin{aligned}\gamma_S : V \times V &\longrightarrow \mathbb{R} \\ (v, w) &\longrightarrow \text{Cov}(S(v), S(w))\end{aligned}$$

où \mathbb{E} et Cov désignent respectivement l'espérance et la covariance usuelles sur les variables aléatoires de $L^2(\mathcal{E})$.

Remarque 3. *Une forme linéaire aléatoire dont l'espérance est identiquement nulle sera dite centrée.*

Proposition 2. *Pour toute forme linéaire aléatoire S de $V_{\mathcal{E}}^*$, $\mathbb{E}^*[S]$ est une forme linéaire continue sur V et γ_S est une forme bilinéaire continue sur $V \times V$.*

Démonstration. Soit S une forme linéaire aléatoire sur V . Comme S est continue de V dans $\mathcal{L}^2(\mathcal{E})$, il existe $C_S > 0$ tel que

$$\sup_{\|w\|_V \leq 1} \|S(w)\|_{L^2(\mathcal{E})}^2 \leq C_S.$$

Par ailleurs, on a, pour tout w de V ,

$$|\mathbb{E}^*[S](w)| \leq \|S(w)\|_{L^2(\mathcal{E})}.$$

D'où $\sup_{\|w\|_V \leq 1} |\mathbb{E}^*[S](w)| \leq C_S$, ce qui implique la continuité de $\mathbb{E}^*[S]$.

Pour tous v et w dans V ,

$$\begin{aligned}|\gamma_S(v, w)| &= |\text{Cov}(S(v), S(w))| \\ &\leq 2\sqrt{\mathbb{E}[|S(v)|^2]}\sqrt{\mathbb{E}[|S(w)|^2]}.\end{aligned}$$

On a donc $\sup_{\|v\|_V \leq 1, \|w\|_V \leq 1} |\gamma_S(v, w)| \leq 2C_S^2$, d'où $\gamma_S \in V^{**}$ (ensemble des formes bilinéaires continues sur $V \times V$). □

Corollaire 1 (Riesz). *Pour toute forme linéaire aléatoire $S \in V_{\mathcal{E}}^*$, l'espérance $\mathbb{E}^*[S]$ admet un représentant dans V et l'autocovariance γ_S admet un représentant dans V^* .*

Définition 13 (Loi d'une forme linéaire aléatoire). *On appelle loi d'une forme linéaire aléatoire S la donnée, pour tout entier strictement positif N et pour tous éléments v_1, \dots, v_N de V , de la loi du vecteur aléatoire $(S(v_1), \dots, S(v_N))$.*

Définition 14 (Forme linéaire gaussienne). *Une forme linéaire aléatoire S est dite gaussienne si, pour tout entier strictement positif N et pour tous éléments v_1, \dots, v_N de V , le vecteur aléatoire $(S(v_1), \dots, S(v_N))$ est gaussien.*

Proposition 3. *La loi d'une forme linéaire aléatoire gaussienne S est entièrement déterminée par la donnée de $\mathbb{E}^*[S]$ et de γ_S .*

Définition 15 (Indépendance). *Deux formes linéaires aléatoires S et T sont dites indépendantes si, pour tout entier strictement positif N et pour tous éléments v_1, \dots, v_N de V , les vecteurs aléatoires $(S(v_1), \dots, S(v_N))$ et $(T(v_1), \dots, T(v_N))$ sont indépendants.*

Définition 16 (Bruit blanc). *Une forme linéaire aléatoire $B \in V_{\mathcal{E}}^*$ est un bruit blanc si les deux conditions suivantes sont vérifiées :*

(1) *la forme linéaire $\mathbb{E}^*[S]$ est identiquement nulle sur V ;*

(2) *il existe $\sigma > 0$ tel que*

$$\gamma_B(v, w) = \sigma^2 \langle v, w \rangle_V, \forall v, w \in V.$$

Le paramètre σ^2 est appelé variance du bruit blanc B .

Structure d'espace vectoriel

On définit la somme de deux formes linéaires aléatoires ainsi que le produit d'une forme linéaire aléatoire par un nombre réel comme suit, de manière analogue à ce qui est fait pour les processus généralisés en Section 2.3 : pour toutes formes linéaires aléatoires S et T de $V_{\mathcal{E}}^*$ et pour tout réel λ ,

$$\lambda S : V \longrightarrow L^2(\mathcal{E})$$

$$w \longrightarrow \lambda S(w)$$

$$S + T : V \longrightarrow L^2(\mathcal{E})$$

$$w \longrightarrow S(w) + T(w).$$

Le produit d'une forme linéaire aléatoire par un réel et la somme de deux formes linéaires aléatoires sont des éléments de $V_{\mathcal{E}}^*$. Par conséquent l'ensemble $V_{\mathcal{E}}^*$ a une structure de \mathbb{R} -espace vectoriel sur \mathbb{R} et c'est même un espace de Banach, d'après la proposition qui suit.

Proposition 4. L'application $\|\cdot\|_{V_{\mathcal{E}}^*}$ définie par

$$\begin{aligned} \|\cdot\|_{V_{\mathcal{E}}^*} : V_{\mathcal{E}}^* &\longrightarrow \mathbb{R}^+ \\ S &\longrightarrow \sup_{\|w\|_V \leq 1} \|S(w)\|_{L^2(\mathcal{E})}, \end{aligned}$$

où $\|\cdot\|_{L^2(\mathcal{E})}$ désigne la norme usuelle sur l'espace $L^2(\mathcal{E})$, est une norme sur $V_{\mathcal{E}}^*$. L'espace vectoriel normé $(V_{\mathcal{E}}^*, \|\cdot\|_{V_{\mathcal{E}}^*})$ est un espace de Banach.

Démonstration. On démontre tout d'abord que $\|\cdot\|_{L^2(\mathcal{E})}$ est une norme sur $V_{\mathcal{E}}^*$, en utilisant les propriétés de norme de $\|\cdot\|_{L^2(\mathcal{E})}$. On a de façon évidente, pour tout S de $V_{\mathcal{E}}^*$, pour tout λ de \mathbb{R} , $\|S\|_{V_{\mathcal{E}}^*} = 0$ si et seulement si $S = 0$ dans $V_{\mathcal{E}}^*$ et $\|\lambda S\|_{V_{\mathcal{E}}^*} = |\lambda| \|S\|_{V_{\mathcal{E}}^*}$; d'autre part, pour tous S et T de $V_{\mathcal{E}}^*$, on a

$$\begin{aligned} \|S + T\|_{V_{\mathcal{E}}^*} &= \sup_{\|w\|_V \leq 1} \|S(w) + T(w)\|_{L^2(\mathcal{E})} \\ &\leq \sup_{\|w\|_V \leq 1} (\|S(w)\|_{L^2(\mathcal{E})} + \|T(w)\|_{L^2(\mathcal{E})}) \\ &\leq \sup_{\|w\|_V \leq 1} \|S(w)\|_{L^2(\mathcal{E})} + \sup_{\|w\|_V \leq 1} \|T(w)\|_{L^2(\mathcal{E})} \\ &\leq \|S\|_{V_{\mathcal{E}}^*} + \|T\|_{V_{\mathcal{E}}^*}. \end{aligned}$$

$\|\cdot\|_{V_{\mathcal{E}}^*}$ est donc une norme sur $V_{\mathcal{E}}^*$.

Il reste à démontrer que cet espace est complet. Soit (S_N) une suite de Cauchy dans $(V_{\mathcal{E}}^*, \|\cdot\|_{V_{\mathcal{E}}^*})$. On a

$$\forall \epsilon > 0, \exists N_{\epsilon}, \forall m, n > N_{\epsilon}, \sup_{\|w\|_V \leq 1} \|S_n(w) - S_m(w)\|_{L^2(\mathcal{E})} < \epsilon.$$

Pour tout $n \in \mathbb{N}$, on note \tilde{S}_n la restriction de S_n à la boule unité B_V de V . Par ce qui précède, cette suite vérifie un critère de Cauchy uniforme; elle converge donc uniformément vers une fonction \tilde{S} définie sur B_V et à valeurs dans $L^2(\mathcal{E})$. On définit S comme l'application

$$\begin{aligned} S : V &\longrightarrow L^2(\mathcal{E}) \\ w &\longrightarrow \|w\|_V \tilde{S}(\tilde{w}), \end{aligned}$$

où $\tilde{w} = 0$ si $w = 0_V$ et $\tilde{w} = \frac{w}{\|w\|_V}$ sinon.

S est clairement linéaire. On en montre à présent la continuité; soit w un élément de V et soit (w_n) une suite qui

converge vers w dans V . Pour tous entiers m, n , on a

$$\begin{aligned}
\|S(w) - S(w_n)\|_{L^2(\mathcal{E})} &\leq \|S(w) - S_m(w)\|_{L^2(\mathcal{E})} + \|S_m(w) - S_m(w_n)\|_{L^2(\mathcal{E})} \\
&\quad + \|S_m(w_n) - S(w_n)\|_{L^2(\mathcal{E})} \\
&\leq \|w\|_V \|\tilde{S}(\tilde{w}) - \tilde{S}_m(\tilde{w})\|_{L^2(\mathcal{E})} + \|w - w_n\|_{S_m} \|V_{\mathcal{E}}^* \\
&\quad + \|w_n\|_V \|\tilde{S}_m(\tilde{w}_n) - \tilde{S}(\tilde{w}_n)\|_{L^2(\mathcal{E})}.
\end{aligned}$$

On déduit alors, par la convergence uniforme de (\tilde{S}_m) vers \tilde{S} et le critère de Cauchy, l'existence d'une constante positive C telle que pour tout ϵ positif et pour tout n "suffisamment grand",

$$\|S(w) - S(w_n)\|_{L^2(\mathcal{E})} \leq 2\epsilon \|w\|_V + \epsilon C,$$

d'où la continuité de S , qui est donc un élément de $V_{\mathcal{E}}^*$. On obtient alors la convergence de S_n vers S au sens de l'espace $V_{\mathcal{E}}^*$ en conséquence immédiate de la convergence uniforme de \tilde{S}_n vers \tilde{S} . D'où le caractère complet de $V_{\mathcal{E}}^*$. \square

Dans la suite, on utilise cette notion de forme linéaire aléatoire, ainsi que l'espace $V_{\mathcal{E}}^*$ que l'on vient de construire, pour adapter le cadre statistique défini en Section 3.2.1 afin de tenir compte de la variabilité aléatoire des surfaces au sein d'un échantillon et de refléter cette dernière en termes de représentants.

3.2.3 Modèle de variation

On se donne une population de surfaces de Θ , notée (S_1, \dots, S_K) et la suite (s_1, \dots, s_K) de leurs représentants respectifs dans V^* . On décrit la variabilité des surfaces à travers leurs représentants. On considère que, pour tout k de $\{1, \dots, K\}$, s_k est une réalisation d'une forme linéaire aléatoire S_k de $V_{\mathcal{E}}^*$. On suppose que les formes linéaires S_1, \dots, S_K sont mutuellement indépendantes et de même loi, selon le modèle suivant :

$$S_k = \mu + B_k, \forall k \in \{1, \dots, K\}, \quad (3.2)$$

où μ est une forme linéaire de V^* et (B_1, \dots, B_K) est une suite i.i.d. de formes linéaires aléatoires centrées. La forme linéaire μ est le représentant moyen de l'échantillon; ce représentant est vu comme inconnu et à estimer, ainsi que l'opérateur d'autocovariance des $(B_k)_{1 \leq k \leq K}$.

Ce travail d'estimation sera effectué sous différentes conditions. Tout d'abord, au Chapitre 4, on supposera que B_1, \dots, B_K sont des bruits blancs de variance σ^2 . On construira, sous cette hypothèse, des estimateurs de μ et de σ^2 à partir des réalisations de S_1, \dots, S_K . Ensuite, au Chapitre 5, on se placera dans le cas, plus général, où B_1, \dots, B_K sont des formes linéaires gaussiennes i.i.d., centrées et d'autocovariance $\Gamma \in V^{**}$; on estimera alors μ et Γ . On étudiera enfin le cas intermédiaire dans lequel l'opérateur Γ est

supposé diagonal.

Remarquons toutefois que le modèle de variation 3.2 est seulement théorique et ne permet pas de décrire des situations expérimentales où les surfaces sont observées sous forme discrète. Dans la suite, on va le compléter pour obtenir un véritable modèle d'observation.

3.3 Discrétisation et modèle d'observation

Discrétisation des surfaces

En pratique, une surface continument différentiable \mathcal{S}_k du modèle de variation 3.2 est observée sous la forme d'une surface discrète F_k^ρ dépendant d'un paramètre de précision ρ , indépendant de k ainsi que de la taille K de l'échantillon. Cette surface discrète est définie par :

$$F_k^\rho = \bigcup_{m=1}^{M_k^\rho} f_{m,k}^\rho$$

où les $f_{m,k}^\rho$, $m \in \{1, \dots, M_k^\rho\}$ sont des facettes triangulaires. Typiquement, on prendra comme valeur du paramètre ρ , dans les applications, l'aire maximale des facettes triangulaires constituant la surface discrète F_k^ρ . Pour chaque k compris entre 1 et K , on définit le représentant $s_{k,\rho}$ de la surface discrète F_k^ρ de manière analogue à ce qui a été fait pour les surfaces lisses :

$$\begin{aligned} s_{k,\rho} : V &\longrightarrow \mathbb{R} \\ w &\longrightarrow \sum_{s=1}^{M_k^\rho} \int_{f_{m,k}^\rho} \langle N_{m,k}^\rho, w(x) \rangle_{\mathbb{R}^3} dx, \end{aligned} \tag{3.3}$$

où $N_{m,k}^\rho$ désigne le vecteur normal unitaire à la facette $f_{m,k}^\rho$, externe à la surface F_k^ρ .

Il est clair que l'application $s_{k,\rho}$ est linéaire. D'autre part, on a pour tout w de V ,

$$\begin{aligned} |S_{k,\rho}| &\leq \|w\|_\infty \sum_{m=1}^{M_k^\rho} \int_{f_{m,k}^\rho} \|N_{m,k}^\rho\|_{\mathbb{R}^3} dx \\ &\leq \|w\|_\infty \mathcal{A}(F_k^\rho). \end{aligned}$$

Par la Proposition 1, il existe alors une constante positive C telle que, pour tout w de V ,

$$|S_{k,\rho}| \leq C \mathcal{A}(F_k^\rho) \|w\|_V,$$

ce qui assure la continuité de $s_{k,\rho}$, qui est donc un élément de V^* .

Comme dans le cas de l'échantillon de surfaces lisses en Section 3.2.3, on considère ici les formes

linéaires $s_{k,\rho}, k \in \{1, \dots, K\}$ comme des réalisations de formes linéaires aléatoires de $V_{\mathcal{E}}^*$, que l'on notera $S_{k,\rho}, k \in \{1, \dots, K\}$. Dans la suite, on note

$$\forall k \in \{1, \dots, K\}, \epsilon_{k,\rho} = S_k - S_{k,\rho}$$

l'écart entre les représentants d'une surface continue de Θ et sa discrétisation, en d'autres termes l'erreur commise sur le représentant en approximant une surface lisse par sa discrétisation. On notera également

$$\overline{\epsilon_{K,\rho}} = \frac{1}{K} \sum_{k=1}^K \epsilon_{k,\rho},$$

la moyenne des écarts sur l'échantillon de surfaces. $\overline{\epsilon_{K,\rho}}$ correspond à l'erreur moyenne de discrétisation sur l'échantillon complet. De par la structure d'espace vectoriel de $V_{\mathcal{E}}^*$, il va de soi que les $\epsilon_{k,\rho}$ et $\overline{\epsilon_{K,\rho}}$ sont des formes linéaires aléatoires. On suppose, bien entendu, que les $\epsilon_{k,\rho}, k \in \{1, \dots, K\}$ sont mutuellement indépendantes et que, pour tous entiers k et l de $\{1, \dots, K\}$, $\epsilon_{k,\rho}$ et S_l sont indépendantes. Ceci reflète, en outre, le fait que les surfaces sont discrétisées indépendamment les unes des autres et que l'erreur survenant lors de la discrétisation est indépendante de la surface initiale. Le terme d'erreur permet de passer du modèle de variation 3.2 à un modèle d'observation décrivant les représentants d'un échantillon de surfaces discrètes.

Modèle d'observation

En combinant la définition 3.3 et le modèle 3.2 on obtient un modèle d'observation, exprimé en termes de représentants de l'échantillon de surfaces discrètes $(F_k^\rho)_{k \in \{1, \dots, K\}}$,

$$S_{k,\rho} = \mu + B_k + \epsilon_{k,\rho}, k \in \{1, \dots, K\}, \quad (3.4)$$

où les $S_{k,\rho}$ sont observées et μ est inconnue et à estimer. Ce modèle peut, bien entendu, être développé sous chacune des deux hypothèses envisagées quant à la loi de la suite $(B_k)_{1 \leq k \leq K}$ en Section 3.2.3.

Aux Chapitres 4 et 5, on donne un estimateur de la forme linéaire moyenne μ dans ce modèle, qui est asymptotiquement un représentant de la surface moyenne continue. On étendra également aux données discrètes l'estimateur de l'autocovariance du bruit.

Convergence de l'approximation discrète

On établit à présent un théorème garantissant la convergence de l'écart entre le représentant d'une surface lisse et le représentant de son approximation discrète lorsque le paramètre de précision ρ tend vers 0. Ce théorème est fondamental dans le sens où il atteste de la cohérence de notre modèle d'observation

construit par passage du continu à l'approximation discrète. Pour assurer de bonnes propriétés permettant de démontrer cette convergence, il suffit d'avoir les trois hypothèses ci-dessous.

Hypothèse 1. *Il existe une fonction a_1 avec $\lim_{x \rightarrow 0} a_1(x) = 0$ telle que, pour tout ρ positif, pour tout k de $\{1, \dots, K\}$, il existe une partition $(\mathcal{S}_{m,k}^\rho)_{m \in \{1, \dots, M_k^\rho\}}$ de \mathcal{S}_k et un ensemble de difféomorphismes $(H_{m,k}^\rho)_{m \in \{1, \dots, M_k^\rho\}}$ qui mettent en correspondance de manière univoque les $\mathcal{S}_{m,k}^\rho$ avec les facettes $f_{m,k}^\rho$ et vérifient*

$$\mathbb{E} \left[\sup_{m \in \{1, \dots, M_k^\rho\}} (\|H_{m,k}^\rho - I\|_\infty + \|J(H_{m,k}^\rho) - I\|_\infty)^2 \right] \leq a_1(\rho),$$

où I désigne l'identité de \mathbb{R}^3 et $J(H_{m,k}^\rho)$ la matrice jacobienne de $H_{m,k}^\rho$.

Hypothèse 2. *Il existe $C_0 > 0$ tel que, pour tout entier strictement positif K ,*

$$\mathbb{P} \left[\sup_{k \in \{1, \dots, K\}} (\mathcal{A}(\mathcal{S}_k))^2 \leq C_0 \right] = 1,$$

où \mathcal{A} désigne la mesure d'aire.

Hypothèse 3. *Il existe $\rho_0 > 0$ tel que, pour tout $K \in \mathbb{N}^*$, on ait presque sûrement :*

$$\forall \rho \in]0, \rho_0[, \forall k \in \{1, \dots, K\}, \forall w \in V \text{ t.q. } \|w\|_V \leq 1, |\epsilon_{k,\rho}(w)| \leq 1.$$

Théorème 3 (Convergence de l'erreur de discrétisation). *Sous les Hypothèses 1 et 2, on a les convergences :*

$$\lim_{\rho \rightarrow 0} \sup_{\|w\|_V=1} \sup_{k \in \{1, \dots, K\}} \mathbb{E} [(\epsilon_{k,\rho}(w))^2] = 0, \quad (3.5)$$

$$\lim_{\rho \rightarrow 0, K \rightarrow \infty} \sup_{\|w\|_V=1} \mathbb{E} [(\overline{\epsilon_{K,\rho}(w)})^2] = 0. \quad (3.6)$$

Pour démontrer ce théorème, on a besoin de trois lemmes.

Lemme 1. *Il existe une fonction a_2 telle que $\lim_{x \rightarrow 0} a_2(x) = 0$ et, $\forall K \in \mathbb{N}^*, \forall k \in \{1, \dots, K\}$,*

$$\mathbb{E} \left[\sup_{m \in \{1, \dots, M_k^\rho\}} \sup_{u \in \mathcal{S}_{m,k}^\rho} |1 - \text{Det}(J_{H_m^\rho}(u))|^2 \right] \leq a_2(\rho).$$

Démonstration. On utilise le résultat intermédiaire : soit P une fonction polynômiale à p variables et de degré q , soit \mathcal{K} un compact de \mathbb{R}^p ; il existe une constante $C > 0$ telle que, $\forall x \in \mathcal{K}, \forall \epsilon > 0$ avec $x + \epsilon \in \mathcal{K}$,

$$|P(x + \epsilon) - P(x)| \leq C \sum_{k=1}^{q-1} \|\epsilon\|^k.$$

La preuve de ce lemme est conséquence immédiate de ce résultat appliqué au déterminant (qui est une fonction polynômiale des coefficients de la matrice). \square

Lemme 2. *Il existe une fonction a_3 avec $\lim_{x \rightarrow 0} a_3(x) = 0$ telle que*

$$\forall \rho > 0, \forall K \in \mathbb{N}^*, \forall k \in \{1, \dots, K\},$$

$$\mathbb{E} \left[\sup_{m \in \{1, \dots, M_k^\rho\}} \sup_{u \in \mathcal{S}_{m,k}^\rho} \|N_k(u) - N_{m,k}^\rho\|_{\mathbb{R}^3}^2 \right] \leq a_3(\rho),$$

où $N_k(u)$ et $N_{m,k}^\rho$ désignent respectivement la normale externe à la surface \mathcal{S}_k au point u et la normale à la facette $f_{m,k}^\rho$ externe à la surface discrète F_k^ρ .

Démonstration. La preuve repose sur l'égalité

$$\|N_k(u) - N_{m,k}^\rho\|_{\mathbb{R}^3} = \|N_{m,k}^\rho(J(H_{m,k}^\rho) - I)\|_{\mathbb{R}^3},$$

qui permet de déduire

$$\|N_k(u) - N_{m,k}^\rho\|_{\mathbb{R}^3} \leq \|J(H_{m,k}^\rho) - I\|_{\mathcal{L}(\mathbb{R}^3, \mathbb{R}^3)},$$

puis de conclure le résultat. \square

Lemme 3. *Il existe une fonction a_4 avec $\lim_{x \rightarrow 0} a_4(x) = 0$ telle que,*

$$\forall \rho > 0, \forall K \in \mathbb{N}^*, \forall k \in \{1, \dots, K\},$$

$$\mathbb{E} \left[\sup_{m \in \{1, \dots, M_k^\rho\}} \sup_{u \in \mathcal{S}_{m,k}^\rho} \|K_0(H_{m,k}^\rho(u), \cdot) - K_0(u, \cdot)\|_{V_0}^2 \right] \leq a_4(\rho).$$

Démonstration. Le résultat découle immédiatement de la continuité, uniforme sur le compact \mathcal{D} , de la fonction $y \rightarrow K_0(y, \cdot)$ et de l'Hypothèse 1. \square

Démonstration du Théorème 3. Pour tout entier positif K , pour tout k dans $\{1, \dots, K\}$, pour tout élément w de V , on pose

$$I_k^\rho(w) = \sum_{m=1}^{M_k^\rho} \int_{\mathcal{S}_{m,k}^\rho} |\langle N_k(u) - N_{m,k}^\rho, w(u) \rangle_{\mathbb{R}^3}| d\sigma(u),$$

$$\text{et } J_k^\rho(w) = \sum_{m=1}^{M_k^\rho} \left| \int_{f_{m,k}^\rho} \langle N_{m,k}^\rho, w(u) \rangle_{\mathbb{R}^3} du - \int_{\mathcal{S}_{m,k}^\rho} \langle N_{m,k}^\rho, w(u) \rangle_{\mathbb{R}^3} d\sigma(u) \right|.$$

On a alors

$$\mathbb{E} [(\epsilon_{k,\rho}(w))^2] \leq \mathbb{E} [(I_k^\rho(w))^2] + \mathbb{E} [(J_k^\rho(w))^2] + 2\mathbb{E} [I_k^\rho(w)J_k^\rho(w)],$$

et donc

$$\mathbb{E} [(\epsilon_k^\rho(w))^2] \leq \mathbb{E} [(I_k^\rho(w))^2] + \mathbb{E} [(J_k^\rho(w))^2] + 2\sqrt{\mathbb{E} [(I_k^\rho(w))^2] \mathbb{E} [(J_k^\rho(w))^2]}, \quad (3.7)$$

D'une part, comme $\forall u \in \mathcal{D}$, $|w(u)| \leq \|K_0(u, \cdot)\|_{V_0} \|w\|_V$, on a

$$|I_k^\rho(w)| \leq \|w\|_V \sup_{u \in \mathcal{D}} \|K_0(u, \cdot)\|_{V_0} \mathcal{A}(\mathcal{S}_k) \sup_{m \in \{1, \dots, M_k^\rho\}} \sup_{u \in \mathcal{S}_k} \|N_k(u) - N_{m,k}^\rho\|_{\mathbb{R}^3},$$

et, en utilisant le Lemme 2 puis l'Hypothèse 2, on obtient

$$\begin{aligned} \mathbb{E} [(I_k^\rho(w))^2] &\leq \|w\|_V^2 \sup_{u \in \mathcal{D}} \|K_0(u, \cdot)\|_{V_0}^2 \mathbb{E} [\mathcal{A}(\mathcal{S}_k)^2] a_3(\rho) \\ &\leq C_0 \|w\|_V^2 \sup_{u \in \mathcal{D}} \|K_0(u, \cdot)\|_{V_0}^2 a_3(\rho). \end{aligned}$$

Par conséquent, $\exists C_1 > 0, \forall K \in \mathbb{N}^*, \forall k \in \{1, \dots, K\}, \forall w \in V$,

$$\mathbb{E} [(I_k^\rho(w))^2] \leq C_1 a_3(\rho) \|w\|_V^2. \quad (3.8)$$

D'autre part,

$$\begin{aligned} |J_k^\rho(w)| &\leq \sum_{m=1}^{M_k^\rho} \int_{\mathcal{S}_{m,k}^\rho} |\langle N_{m,k}^\rho, w(u) - \text{Det}(J_{H_{m,k}^\rho}) w(H_{m,k}^\rho(u)) \rangle_{\mathbb{R}^3}| d\sigma(u) \\ &\leq \sum_{m=1}^{M_k^\rho} \left(\int_{\mathcal{S}_{m,k}^\rho} |\langle N_{m,k}^\rho, w(u) - w(H_{m,k}^\rho(u)) \rangle_{\mathbb{R}^3}| d\sigma(u) \right. \\ &\quad \left. + \int_{\mathcal{S}_{m,k}^\rho} |\langle N_{m,k}^\rho, w(H_{m,k}^\rho(u)) - \text{Det}(J_{H_{m,k}^\rho}(u)) w(H_{m,k}^\rho(u)) \rangle_{\mathbb{R}^3}| d\sigma(u) \right) \\ &\leq \sum_{m=1}^{M_k^\rho} \left(\int_{\mathcal{S}_{m,k}^\rho} \|w(u) - w(H_{m,k}^\rho(u))\|_{\mathbb{R}^3} d\sigma(u) \right. \\ &\quad \left. + \int_{\mathcal{S}_{m,k}^\rho} \|w(H_{m,k}^\rho(u)) - \text{Det}(J_{H_{m,k}^\rho}(u)) w(H_{m,k}^\rho(u))\|_{\mathbb{R}^3} d\sigma(u) \right). \end{aligned}$$

Par propriété de reproduction, on déduit alors qu'il existe une constante positive C_2 telle que $\forall K \in \mathbb{N}^*, \forall k \in \{1, \dots, K\}, \forall w \in V$,

$$|J_k^\rho(w)| \leq C_2 \sum_{m=1}^{M_k^\rho} (\|w\|_V \int_{\mathcal{S}_{m,k}^\rho} \|K_0(u, \cdot) - K_0(H_{m,k}^\rho(u), \cdot)\|_{V_0} d\sigma(u) + \int_{\mathcal{S}_{m,k}^\rho} |1 - \text{Det}(J_{H_{m,k}^\rho}(u))| d\sigma(u)).$$

On applique à présent les Lemmes 1 et 3, l'Hypothèse 2 ainsi que la Proposition 1 pour conclure qu'il existe $C_3 > 0$ tel que $\forall K \in \mathbb{N}^*, \forall k \in \{1, \dots, K\}, \forall w \in V$,

$$\mathbb{E} [(J_k^\rho(w))^2] \leq C_3 \|w\|_V^2 (a_1(\rho) + a_2(\rho)). \quad (3.9)$$

D'après (3.7), (3.8) et (3.9), on a alors

$$\sup_{\|w\|_V \leq 1} \sup_{K \in \mathbb{N}^*} \sup_{k \in \{1, \dots, K\}} \mathbb{E} [(\epsilon_{k,\rho}(w))^2] \leq C_1 a_3(\rho) + C_3 a_1(\rho) + C_3 a_2(\rho) + 2\sqrt{C_1 C_3 a_3(\rho)(a_1(\rho) + a_2(\rho))},$$

ce qui prouve la convergence (3.5).

On montre maintenant le résultat (3.6). Par construction des modèles d'observation et de variation, la suite $(\epsilon_{k,\rho})_{k \in \{1, \dots, K\}}$ est *i.i.d.* ; ainsi, on a, $\forall w \in V, \forall K \in \mathbb{N}^*, \forall \rho > 0$,

$$\mathbb{E} [(\overline{\epsilon_{K,\rho}}(w))^2] \leq \frac{1}{K^2} \sum_{k=1}^K \mathbb{E} [(\epsilon_{k,\rho}(w))^2] + \frac{1}{K^2} \sum_{k \neq l} \mathbb{E} [\epsilon_{k,\rho}] \mathbb{E} [\epsilon_{l,\rho}].$$

On obtient donc, en utilisant l'inégalité de Cauchy-Schwarz,

$$\sup_{\|w\|_V \leq 1} \mathbb{E} [(\overline{\epsilon_{K,\rho}}(w))^2] \leq \left(\frac{1}{K} + \frac{K-1}{K} \right) \sup_{\|w\|_V \leq 1} \sup_{K \in \mathbb{N}^*} \sup_{k \in \{1, \dots, K\}} \mathbb{E} [(\epsilon_{k,\rho}(w))^2],$$

ce qui prouve la convergence (3.6), vu le résultat (3.5). □

Chapitre 4

Estimation (dans le cadre d'un modèle de bruit blanc)

4.1 Objectifs

Dans ce chapitre, on considère les Modèles (3.2) et (3.4), sous l'hypothèse : $(B_k)_{1 \leq k \leq K}$ est une suite i.i.d. de bruits blancs gaussiens de variance σ^2 .

Dans un premier temps, on construit des estimateurs pour le représentant moyen d'un échantillon de surfaces et pour la variance du bruit. Dans chaque cas, on opère toujours dans une logique de passage du continu au discret : on détermine tout d'abord l'estimateur du maximum de vraisemblance dans le modèle de variation (3.2) puis, en remplaçant les représentants des surfaces lisses par ceux de leurs approximations discrètes, on obtient un estimateur empirique dans le modèle d'observation (3.4). C'est cet estimateur qui sera utilisé en pratique dans la mesure où, dans la réalité expérimentale, on ne dispose que d'approximations discrètes des surfaces.

Comme les représentants des surfaces sont des éléments d'un espace de dimension infinie, il est indispensable, en amont de cette construction, de définir une procédure pour approximer ces derniers par des éléments dépendant d'un nombre fini de paramètres. Dans cette optique, on se donne un sous-espace vectoriel de dimension finie de l'espace des représentants, que l'on nommera "sous-espace d'approximation". On approxime les représentants des surfaces, lisses ou discrètes, par leur projection sur ce sous-espace. Cette procédure présente l'avantage de donner des approximations dépendant toutes du même nombre de paramètres, fixé comme la dimension du sous-espace d'approximation.

Dans un second temps, on démontre la consistance des estimateurs du représentant moyen et de la variance du bruit lorsque la taille de l'échantillon et la dimension du sous-espace d'approximation tendent

vers l'infini et lorsque le paramètre de précision ρ de la discrétisation (défini en Section 3.3) tend vers 0.

4.2 Approximation

Pour estimer le représentant moyen μ et la variance σ^2 du bruit dans les modèles (3.2) et (3.4), on passe par des approximations des éléments de V et V^* dans un sous-espace vectoriel de dimension finie. Cette approximation se construit selon le processus détaillé ci-après.

Soit $(\phi_i)_{i \in I}$ une famille d'éléments de V , totale et linéairement indépendante. Soit $(I_N)_{N \in \mathbb{N}^*}$ une suite croissante de sous-ensembles emboîtés de I dont l'union est égale à I . Pour $N \in \mathbb{N}^*$, on considère les sous-espaces vectoriels de V suivants

$$V_N = \text{Vect}(\{\phi_i | i \in I_N\}),$$

munis du produit scalaire induit par V . Ces sous-espaces, que l'on nommera "espaces d'approximation", sont emboîtés et de réunion dense dans V . Pour $N \in \mathbb{N}^*$, on définit également les sous-espaces \widetilde{V}_N^* de V^* :

$$\widetilde{V}_N^* = \{s \in V^* | v_S \in V_N\},$$

où v_S désigne le représentant de Riesz de la forme linéaire S . On note, pour tout élément $S \in V^*$,

$$S^N = \underset{\nu \in \widetilde{V}_N^*}{\text{argmin}} \|S - \nu\|_{V^*}$$

la projection orthogonale de S sur \widetilde{V}_N^* ; de même, pour tout $v \in V$, on note

$$v^N = \underset{w \in V_N}{\text{argmin}} \|v - w\|_V$$

la projection orthogonale de v sur V_N . On remarque que $v_{S^N} = v_S^N$.

Chaque élément S de V^* peut, par conséquent, être approché par un élément de \widetilde{V}_N^* en utilisant la projection orthogonale v_S^N de son représentant de Riesz v_S sur V_N . On remarque que v_S^N satisfait la relation

$$\langle v_S^N, w \rangle_V = S(w), \forall w \in V_N.$$

La famille $(\phi_i)_{i \in I_N}$ étant linéairement indépendante, on a de manière équivalente $v_S^N = \sum_{i \in I_N} \alpha_i \phi_i$, où

$(\alpha_i)_{i \in I_N}$ est la solution du système linéaire

$$\left(\sum_{j \in I_N} \alpha_j g_{i,j} = \sum_{j \in I_N} S(\phi_j) \right)_{i \in I_N},$$

avec $g_{i,j} = \langle \phi_i, \phi_j \rangle_V$. Pour i et j décrivant I_N , on notera $h_{i,j}^N$ les termes qui satisfont $\sum_{k \in I_N} g_{i,k} h_{k,j}^N = \delta_{i,j}$. On a alors :

$$v_S^N = \sum_{i \in I_N} \alpha_i \phi_i,$$

avec $\alpha_i = \sum_{j \in I_N} h_{i,j}^N S(\phi_j)$.

4.3 Construction des estimateurs

4.3.1 Estimateurs dans le modèle de variation

Ici, on construit des estimateurs pour le représentant moyen μ et le paramètre variance σ^2 du bruit dans le modèle de variation (3.2). Ces estimateurs ne sont, en pratique, pas utilisés car ils reposent sur les surfaces lisses $\mathcal{S}_k, k \in \{1, \dots, K\}$ qui ne sont pas observées. Toutefois, dans la Section 4.3.2, ils serviront de base pour la construction d'estimateurs empiriques, reposant sur les surfaces discrètes $F_k^\rho, k \in \{1, \dots, K\}$. Les estimateurs associés au modèle de variation seront aussi utiles en Section 4.4.

La moyenne μ du modèle de variation (3.2) est estimée à travers son représentant de Riesz v_μ . Soit v_μ^N la projection de v_μ sur V_N construite telle qu'expliquée en Section 4.2. Pour tout k de $\{1, \dots, K\}$, on définit les vecteurs aléatoires :

$$\begin{aligned} \mathbf{S}_k^N &= (S_k(\phi_i))_{i \in I_N}, \\ \eta_k^N &= (B_k(\phi_i))_{i \in I_N}, \\ \text{et } \mu^N &= (\mu(\phi_i))_{i \in I_N}. \end{aligned}$$

On notera respectivement $\mathbf{S}_{k,i}^N, \eta_{k,i}^N, \mu_i^N$, la i -ème composante des vecteurs $\mathbf{S}_k^N, \eta_k^N, \mu^N$.

D'après le modèle de variation (3.2) on a

$$\mathbf{S}_k^N = \mu^N + \eta_k^N, \forall k \in \{1, \dots, K\}. \quad (4.1)$$

Pour $k \in \{1, \dots, K\}$, les vecteurs aléatoires \mathbf{S}_k^N sont *i.i.d.* et suivent une loi gaussienne multivariée de vecteur-espérance μ^N et de matrice de covariance Γ^N dont les termes $\gamma_{i,j}^N$ sont définis par

$$\begin{aligned} \gamma_{i,j}^N &= \text{Cov}(\mathbf{S}_{k,i}^N, \mathbf{S}_{k,j}^N) = \gamma_B(\phi_i, \phi_j) \\ &= \sigma^2 \langle \phi_i, \phi_j \rangle_V. \end{aligned}$$

Dans ce modèle, l'estimateur du maximum de vraisemblance pour le vecteur μ^N est alors

$$\overline{\mathbf{S}}_K^N = \frac{1}{K} \sum_{k=1}^K \mathbf{S}_k^N.$$

On estime ainsi le représentant de Riesz v_μ de μ par

$$\widehat{v}_K^N = \sum_{m \in I_N} \left(\sum_{n \in I_N} h_{m,n}^N \overline{\mathbf{S}}_{k,n}^N \right) \phi_m, \quad (4.2)$$

où les $\overline{\mathbf{S}}_{k,n}^N, n \in I_N$ sont les composantes du vecteur $\overline{\mathbf{S}}_k^N$. De manière équivalente, on écrit \widehat{v}_K^N sous la forme développée :

$$\widehat{v}_K^N = \frac{1}{K} \sum_{k \in \{1, \dots, K\}} \left[\sum_{m \in I_N} \left(\sum_{n \in I_N} h_{m,n}^N S_k(\phi_n) \right) \phi_m \right]. \quad (4.3)$$

On estime alors la forme linéaire μ par la forme linéaire aléatoire $\widehat{\mu}_K^N$ définie par

$$\widehat{\mu}_K^N(w) = \langle \widehat{v}_K^N, w \rangle_V, \forall w \in V. \quad (4.4)$$

Par ailleurs, on a l'estimateur du maximum de vraisemblance pour le paramètre de variance σ^2 du bruit blanc du modèle (3.2)

$$\widehat{\nu}_K^N = \frac{1}{K \text{Card}(I_N)} \sum_{i \in I_N} \left[\sum_{j \in I_N} \left(\sum_{k \in \{1, \dots, K\}} (\mathbf{S}_{k,i}^N - \overline{\mathbf{S}}_{k,i}^N) h_{i,j}^N (\mathbf{S}_{k,j}^N - \overline{\mathbf{S}}_{k,j}^N) \right) \right]. \quad (4.5)$$

Cet estimateur se réécrit sous la forme développée

$$\widehat{\nu}_K^N = \frac{1}{K \text{Card}(I_N)} \sum_{i \in I_N} \left[\sum_{j \in I_N} \left(\sum_{k \in \{1, \dots, K\}} \left(S_k(\phi_i) - \frac{1}{K} \sum_{l=1}^K S_l(\phi_i) \right) h_{i,j}^N \left(S_k(\phi_j) - \frac{1}{K} \sum_{l=1}^K S_l(\phi_j) \right) \right) \right]. \quad (4.6)$$

4.3.2 Estimateurs dans le modèle d'observation

On définit maintenant des estimateurs basés sur les représentants des surfaces discrètes qui, pour leur part, sont observés. Ce sont ces estimateurs qui, en pratique, sont utilisés pour estimer les paramètres μ et σ . Soient les vecteurs aléatoires

$$\mathbf{S}_k^{N,\rho} = (S_{k,\rho}^N(\phi_i))_{i \in I_N},$$

et $\epsilon_k^{N,\rho} = (\epsilon_{k,\rho}(\phi_i))_{i \in I_N}$.

Avec ces notations, le modèle d'observation donné par l'Equation (3.4) en Section 3.3 se réécrit :

$$\mathbf{S}_k^{N,\rho} = \mu^N + \eta_k^N + \epsilon_k^{N,\rho}, \quad k \in \{1, \dots, K\}.$$

On note également $\overline{\mathbf{S}}_K^{N,\rho} = \frac{1}{K} \sum_{k=1}^K \mathbf{S}_k^{N,\rho}$.

Il est naturel de construire les estimateurs dans le modèle d'observation (3.4) en remplaçant $\overline{\mathbf{S}}_K^N$ par la moyenne empirique $\overline{\mathbf{S}}_K^{N,\rho}$ des formes linéaires discrètes dans les estimateurs (4.4) et (4.6) construits dans le cas continu du modèle de variation (3.2). On écrit ainsi l'estimateur du représentant moyen μ dans le modèle d'observation comme la forme linéaire aléatoire

$$\begin{aligned} \widetilde{\mu}_K^{N,\rho} : V &\longrightarrow \mathcal{L}^2(\mathcal{E}) \\ w &\longrightarrow \langle \widetilde{v}_K^{N,\rho}, w \rangle_V \end{aligned} \quad (4.7)$$

où

$$\widetilde{v}_K^{N,\rho} = \frac{1}{K} \sum_{k=1}^K \left[\sum_{m \in I_N} \left(\sum_{n \in I_N} h_{m,n}^N S_k^{N,\rho}(\phi_n) \right) \phi_m \right]. \quad (4.8)$$

Pareillement, on estime la variance σ^2 du bruit dans le modèle d'observation par

$$\widetilde{\nu}_k^{N,\rho} = \frac{1}{K \text{card}(I_N)} \sum_{i \in I_N} \left[\sum_{j \in I_N} \left(\sum_{k=1}^K \left(S_k^{N,\rho}(\phi_i) - \frac{1}{K} \sum_{l=1}^K S_l^{N,\rho}(\phi_i) \right) h_{i,j}^N \left(S_k^{N,\rho}(\phi_j) - \frac{1}{K} \sum_{l=1}^K S_l^{N,\rho}(\phi_j) \right) \right) \right]. \quad (4.9)$$

4.4 Consistance

On énonce à présent les résultats garantissant la consistance des estimateurs établis en Section 4.3.2 lorsque la taille K de l'échantillon et le paramètre de dimension de l'espace d'approximation N tendent vers l'infini et quand le paramètre de précision de la discrétisation ρ tend vers 0. On démontre ces résultats en utilisant comme résultat intermédiaire la consistance des estimateurs définis en Section 4.3.1.

4.4.1 Estimateur du représentant moyen

Théorème 4 (Consistance de l'estimateur du représentant moyen). *La forme linéaire aléatoire $\widetilde{\mu}_k^{N,\rho}$ converge vers le représentant moyen μ dans l'espace $(V_{\mathcal{E}}^*, \|\cdot\|_{V_{\mathcal{E}}^*})$ quand N et K tendent vers $+\infty$ et quand ρ tend vers 0, i.e.*

$$\lim_{\substack{\rho \rightarrow 0 \\ K, N \rightarrow \infty}} \|\widetilde{\mu}_k^{N,\rho} - \mu\|_{V_{\mathcal{E}}^*} = 0,$$

autrement dit

$$\lim_{\substack{\rho \rightarrow 0 \\ K, N \rightarrow \infty}} \sup_{\|w\|_V \leq 1} \sqrt{\mathbb{E} \left[|\widetilde{\mu}_k^{N, \rho}(w) - \mu(w)|^2 \right]}.$$

Corollaire 2. Ce théorème implique, de manière évidente, les résultats suivants.

- La forme linéaire $\mathbb{E}^*[\widetilde{\mu}_k^{N, \rho}]$ converge vers μ quand N et K tendent vers $+\infty$ et quand ρ tend vers 0, au sens de la convergence ponctuelle des fonctions, i.e.

$$\lim_{\substack{\rho \rightarrow 0 \\ K, N \rightarrow \infty}} \mathbb{E}^*[\widetilde{\mu}_k^{N, \rho}](w) = \mu(w), \forall w \in V.$$

- L'erreur quadratique converge vers 0, au sens de la convergence ponctuelle, quand N et K tendent vers $+\infty$ et quand ρ tend vers 0, i.e.

$$\lim_{\substack{\rho \rightarrow 0 \\ K, N \rightarrow \infty}} \mathbb{E} \left[|\widetilde{\mu}_k^{N, \rho}(w) - \mu(w)|^2 \right], \forall w \in V.$$

La démonstration du Théorème 4 repose sur le Théorème 3 et le lemme suivant, assurant la convergence de l'estimateur de μ dans le modèle de variation.

Lemme 4. La forme linéaire aléatoire $\widehat{\mu}_K^N$ converge vers μ dans l'espace $(V_{\mathcal{E}}^*, \|\cdot\|_{V_{\mathcal{E}}^*})$ quand N et K tendent vers $+\infty$, i.e.

$$\lim_{N, K \rightarrow \infty} \|\widehat{\mu}_K^N - \mu\|_{V_{\mathcal{E}}^*} = 0.$$

On démontre le Lemme 4 puis le Théorème 4.

Démonstration du Lemme 4. Tout d'abord, on décompose par inégalité triangulaire

$$\|\widehat{\mu}_K^N - \mu\|_{V_{\mathcal{E}}^*} \leq \|\widehat{\mu}_K^N - \mu^N\|_{V_{\mathcal{E}}^*} + \|\mu^N - \mu\|_V,$$

où μ^N est la projection de μ sur \widetilde{V}_N^* .

Par définition de la projection orthogonale sur un sous-espace, le second terme de cette somme tend vers 0 quand N tend vers $+\infty$; il ne reste donc qu'à démontrer la convergence vers 0 du premier terme quand N et K tendent vers $+\infty$. Pour ce faire, on utilise le Lemme 5 dont la démonstration est faite par la suite.

Lemme 5. Pour tout entier naturel non nul N , on a

$$\sup_{\|w\|_V \leq 1} \mathbb{E} \left[(\widehat{\mu}_1^N(w) - \mu^N(w))^2 \right] \leq \sigma^2.$$

Soit $K \in \mathbb{N}^*$; on a

$$\begin{aligned}\|\widehat{\mu}_K^N - \mu^N\|_{V_{\mathcal{E}}^*} &= \sqrt{\sup_{\|w\|_V \leq 1} \mathbb{E} \left[(\widehat{\mu}_K^N(w) - \mu^N(w))^2 \right]} \\ &= \sqrt{\sup_{\|w\|_V \leq 1} \text{Var} \left(\widehat{\mu}_K^N(w) \right)},\end{aligned}$$

car $\mu^N(w) = \mathbb{E} \left[\widehat{\mu}_K^N(w) \right]$, $\forall w \in V$ (de par les propriétés élémentaires de l'espérance).

Or, comme la suite des (S_k) , $1 \leq k \leq K$ est *i.i.d.*, on a

$$\begin{aligned}\text{Var} \left(\widehat{\mu}_K^N(w) \right) &= \frac{1}{K} \text{Var} \left(\widehat{\mu}_1^N(w) \right) \\ &= \frac{1}{K} \mathbb{E} \left[(\widehat{\mu}_1^N(w) - \mu^N(w))^2 \right]\end{aligned}$$

On a ainsi, par le Lemme 5,

$$\begin{aligned}\|\widehat{\mu}_K^N - \mu^N\|_{V_{\mathcal{E}}^*} &= \sqrt{\frac{1}{K} \sup_{\|w\|_V \leq 1} \mathbb{E} \left[(\widehat{\mu}_1^N(w) - \mu^N(w))^2 \right]} \\ &\leq \sqrt{\frac{\sigma^2}{K}}\end{aligned}$$

D'où le résultat. □

Démonstration du Lemme 5. Soient $n, m \in I$; soit $N' \in \mathbb{N}$ tel que $m, n \in I_{N'}$. Soit $N \in \mathbb{N}$ tel que $N > N'$. On a

$$\begin{aligned}\mathbb{E} \left[(\widehat{\mu}_1^N(\phi_m) - \mu^N(\phi_m))(\widehat{\mu}_1^N(\phi_n) - \mu^N(\phi_n)) \right] &= \mathbb{E} \left[\langle v_1^N - v_\mu^N, \phi_m \rangle_V \langle v_1^N - v_\mu^N, \phi_n \rangle_V \right] \\ &= \mathbb{E} \left[\left\langle \sum_{i \in I_N} \phi_i \sum_{l \in I_N} h_{i,l}^N \eta_{1,l}^N, \phi_m \right\rangle_V \left\langle \sum_{j \in I_N} \phi_j \sum_{p \in I_N} h_{j,p}^N \eta_{1,p}^N, \phi_n \right\rangle_V \right] \\ &= \sigma^2 \sum_{l \in I_N} \sum_{p \in I_N} \left[\left(\sum_{i \in I_N} h_{i,l}^N g_{i,m} \right) \left(\sum_{j \in I_N} h_{j,p}^N g_{j,n} \right) \langle \phi_l, \phi_p \rangle_V \right] \\ &= \sigma^2 \sum_{l \in I_N} \sum_{p \in I_N} \delta_{l,m} \delta_{p,n} \langle \phi_l, \phi_p \rangle_V \\ &= \sigma^2 \langle \phi_m, \phi_n \rangle_V.\end{aligned}$$

Soient maintenant $w \in V$ et $w^N = \sum_{i \in I_N} \alpha_i \phi_i$ sa projection orthogonale sur V^N . Comme $\widehat{\mu}_1^N - \mu^N$ est identiquement

nulle sur le supplémentaire orthogonal de V_N , on a :

$$\begin{aligned}
\mathbb{E} \left[(\widehat{\mu}_1^N(w) - \mu^N(w))^2 \right] &= \mathbb{E} \left[(\widehat{\mu}_1^N(w^N) - \mu^N(w^N))^2 \right] \\
&= \sum_{i \in I_N} \sum_{j \in I_N} \alpha_i \alpha_j \mathbb{E} \left[(\widehat{\mu}_1^N(\phi_i) - \mu^N(\phi_i)) (\widehat{\mu}_1^N(\phi_j) - \mu^N(\phi_j)) \right] \\
&= \sigma^2 \sum_{i \in I_N} \sum_{j \in I_N} \alpha_i \alpha_j \langle \phi_i, \phi_j \rangle_V \\
&= \sigma^2 \|w^N\|_V \\
&\leq \sigma^2 \|w\|_V.
\end{aligned}$$

En passant à la borne supérieure sur la boule-unité de V , on obtient le résultat voulu. \square

Démonstration du Théorème 4. On décompose, toujours par l'inégalité triangulaire,

$$\|\widehat{\mu}_K^{N,\rho} - \mu\|_{V_{\mathcal{E}}^*} \leq \|\widehat{\mu}_K^{N,\rho} - \widehat{\mu}_K^N\|_{V_{\mathcal{E}}^*} + \|\widehat{\mu}_K^N - \mu\|_{V_{\mathcal{E}}^*}.$$

La convergence vers 0 du second terme quand K et N tendent vers $+\infty$ est assurée par le Lemme 4 ; il ne reste donc qu'à démontrer que le premier terme tend vers 0 quand K et N tendent vers $+\infty$ et ρ tend vers 0.

$$\begin{aligned}
\|\widehat{\mu}_K^{N,\rho} - \widehat{\mu}_K^N\|_{V_{\mathcal{E}}^*}^2 &= \sup_{\|w\|_V \leq 1} \mathbb{E} \left[|\langle \widehat{v}_K^{N,\rho} - \widehat{v}_K^N, w \rangle_V|^2 \right] \\
&= \sup_{\|w\|_V \leq 1} \mathbb{E} \left[|\langle \widehat{v}_K^{N,\rho} - \widehat{v}_K^N, w^N \rangle_V|^2 \right]
\end{aligned}$$

On pose $w^N = \sum_{i \in I_N} \alpha_i^N \phi_i$; on a alors

$$\begin{aligned}
&\|\widehat{\mu}_K^{N,\rho} - \widehat{\mu}_K^N\|_{V_{\mathcal{E}}^*}^2 \\
&= \sup_{\|w\|_V \leq 1} \sum_{s \in I_N} \sum_{t \in I_N} \alpha_s^N \alpha_t^N \sum_{j \in I_N} \sum_{m \in I_N} \mathbb{E}[\overline{\epsilon}_K^\rho(\phi_m) \overline{\epsilon}_K^\rho(\phi_j)] \sum_{i \in I_N} \langle \phi_s, \phi_i \rangle_V h_{i,j}^N \sum_{l \in I_N} \langle \phi_t, \phi_l \rangle_V h_{l,m}^N \\
&= \sup_{\|w\|_V \leq 1} \sum_{s \in I_N} \sum_{t \in I_N} \alpha_s^N \alpha_t^N \mathbb{E}[\overline{\epsilon}_K^\rho(\phi_s) \overline{\epsilon}_K^\rho(\phi_t)] \\
&= \sup_{\|w\|_V \leq 1} \mathbb{E} \left[(\overline{\epsilon}_K^\rho(w^N))^2 \right] \\
&\leq \sup_{\|w\|_V \leq 1} \mathbb{E}[(\overline{\epsilon}_K^\rho(w))^2].
\end{aligned}$$

Le résultat est alors garanti par le Théorème 3. \square

4.4.2 Estimateur de la variance du bruit

Théorème 5 (Consistance de l'estimateur de la variance). *L'estimateur $\widetilde{\nu}_K^{N,\rho}$ du paramètre de variance σ^2 dans le modèle d'observation est asymptotiquement sans biais et consistant en moment d'ordre 2; autrement dit, on a les convergences :*

$$\lim_{\substack{\rho \rightarrow 0 \\ N, K \rightarrow \infty}} \mathbb{E} \left[\widetilde{\nu}_K^{N,\rho} \right] = \sigma^2, \quad (4.10)$$

et

$$\lim_{\substack{\rho \rightarrow 0 \\ N, K \rightarrow \infty}} \mathbb{E} \left[(\widetilde{\nu}_K^{N,\rho} - \sigma^2)^2 \right] = 0. \quad (4.11)$$

La preuve de ce théorème repose sur le Théorème 3 et les deux résultats ci-dessous.

Lemme 6. *Soient X_1, X_2, X_3 et X_4 des variables aléatoires Gaussiennes centrées. On a*

$$\text{Cov}(X_1 X_2, X_3 X_4) = \text{Cov}(X_1, X_3) \text{Cov}(X_2, X_4) + \text{Cov}(X_1, X_4) \text{Cov}(X_2, X_3),$$

et en particulier

$$\text{Cov}(X_1^2, X_2^2) = 2\text{Cov}(X_1, X_2)^2.$$

Démonstration. On calcule la fonction génératrice des moments f du vecteur (X_1, X_2, X_3, X_4) .

$$f(x_1, x_2, x_3, x_4) = \exp\left(\frac{1}{2} \sum_{i=1}^4 \text{Cov}(X_i) x_i^2 + \sum_{i=1}^3 \sum_{j=i+1}^4 \text{Cov}(X_i, X_j) x_i x_j\right).$$

On a alors :

$$\begin{aligned} \mathbb{E}[X_1 X_2 X_3 X_4] &= \frac{\partial^4 f}{\partial x_1 \partial x_2 \partial x_3 \partial x_4} (0, 0, 0, 0) \\ &= \text{Cov}(X_1, X_2) \text{Cov}(X_3, X_4) + \text{Cov}(X_1, X_3) \text{Cov}(X_2, X_4) + \text{Cov}(X_1, X_4) \text{Cov}(X_2, X_3). \end{aligned}$$

De plus, comme X_1, X_2, X_3, X_4 sont centrées, on a :

$$\mathbb{E}[X_1 X_2] \mathbb{E}[X_3 X_4] = \text{Cov}(X_1, X_2) \text{Cov}(X_3, X_4).$$

On déduit alors le résultat :

$$\text{Cov}(X_1 X_2, X_3 X_4) = \text{Cov}(X_1, X_3) \text{Cov}(X_2, X_4) + \text{Cov}(X_1, X_4) \text{Cov}(X_2, X_3).$$

□

Lemme 7. *L'estimateur $\widehat{\nu}_K^N$ du paramètre de variance σ^2 dans le modèle de variation est asymptotiquement sans biais*

et consistant en moment d'ordre 2, autrement dit on a les convergences :

$$\lim_{N,K \rightarrow \infty} \mathbb{E} \left[\widehat{\nu}_K^N \right] = \sigma^2, \quad (4.12)$$

et

$$\lim_{N,K \rightarrow \infty} \mathbb{E} \left[(\widehat{\nu}_K^N - \sigma^2)^2 \right] = 0. \quad (4.13)$$

Démonstration. On démontre tout d'abord la convergence (4.12).

$$\begin{aligned} \mathbb{E}[\widehat{\nu}_K^N] &= \mathbb{E} \left[\frac{1}{K \text{Card}(I_N)} \sum_{i \in I_N} \left[\sum_{j \in I_N} \left(\sum_{k \in \{1, \dots, K\}} (S_k(\phi_i) - \frac{1}{K} \sum_{l=1}^K S_l(\phi_i)) h_{i,j}^N (S_k(\phi_j) - \frac{1}{K} \sum_{l=1}^K S_l(\phi_j)) \right) \right] \right] \\ &= \frac{1}{K \text{Card}(I_N)} \sum_{k=1}^K \sum_{i,j \in I_N} h_{i,j}^N \mathbb{E} \left[(\eta_{k,i}^N - \overline{\eta_{k,i}^N}) (\eta_{k,j}^N - \overline{\eta_{k,j}^N}) \right] \\ &= \frac{1}{K \text{Card}(I_N)} \sum_{i,j \in I_N} h_{i,j}^N \sum_{k=1}^K \left(\mathbb{E}[\eta_{k,i}^N \eta_{k,j}^N] - \mathbb{E}[\eta_{k,i}^N \overline{\eta_{k,j}^N}] - \mathbb{E}[\eta_{k,j}^N \overline{\eta_{k,i}^N}] + \mathbb{E}[\overline{\eta_{k,i}^N} \overline{\eta_{k,j}^N}] \right) \end{aligned}$$

On calcule les espérances en utilisant l'indépendance de la suite de vecteurs aléatoires $(\eta_k^N)_{1 \leq k \leq K}$:

$$\begin{aligned} \mathbb{E}[\widehat{\nu}_K^N] &= \frac{1}{K \text{Card}(I_N)} \sum_{i,j \in I_N} h_{i,j}^N \sum_{k=1}^K \left(\sigma^2 \langle \phi_i, \phi_j \rangle_V - \frac{\sigma^2}{K} \langle \phi_i, \phi_j \rangle_V \right) \\ &= \sigma^2 \frac{1 - \frac{1}{K}}{K \text{Card}(I_N)} \sum_{i \in I_N} \left(\sum_{j \in I_N} h_{i,j}^N g_{i,j} \right) \\ &= \sigma^2 \left(1 - \frac{1}{K} \right), \end{aligned}$$

ce qui prouve la convergence (4.12).

On démontre maintenant la convergence (4.13). On commence par décomposer en biais-variance :

$$\mathbb{E} \left[(\widehat{\nu}_K^N - \sigma^2)^2 \right] = \text{Var}(\widehat{\nu}_K^N) + \left(\mathbb{E} \left[\widehat{\nu}_K^N \right] - \sigma^2 \right)^2.$$

Compte tenu de (4.12), $\lim_{N,K \rightarrow \infty} \left(\mathbb{E} \left[\widehat{\nu}_K^N \right] - \sigma^2 \right)^2 = 0$. Il ne reste alors plus qu'à démontrer que

$$\lim_{N,K \rightarrow \infty} \text{Var}(\widehat{\nu}_K^N) = 0$$

$$\text{Var}(\widehat{\nu}_K^N) = \frac{1}{K \text{Card}(I_N)} \sum_{k=1}^K \sum_{l=1}^K \sum_{i,j \in I_N} \sum_{m,n \in I_N} \text{Cov} \left((\eta_{k,i}^N - \overline{\eta_{k,i}^N}) (\eta_{k,j}^N - \overline{\eta_{k,j}^N}), (\eta_{k,m}^N - \overline{\eta_{k,m}^N}) (\eta_{k,n}^N - \overline{\eta_{k,n}^N}) \right).$$

En écrivant $\eta_{k,i}^N - \overline{\eta_{k,i}^N} = \frac{K-1}{K}\eta_{k,i}^N + \frac{1}{K}\sum_{\substack{l=1 \\ l \neq k}}^K$, on développe la covariance :

$$\begin{aligned}
\text{Var}(\widehat{\nu_K^N}) &= \frac{1}{K \text{Card}(I_N)} \sum_{k=1}^K \sum_{l=1}^K \sum_{i,j \in I_N} \sum_{m,n \in I_N} \left[\frac{(K-1)^4}{K^4} \text{Cov}(\eta_{k,i}^N \eta_{k,j}^N, \eta_{l,m}^N \eta_{l,n}^N) \right. \\
&\quad - \frac{(K-1)^3}{K^4} \sum_{\substack{t=1 \\ t \neq l}}^K \text{Cov}(\eta_{k,i}^N \eta_{k,j}^N, \eta_{l,m}^N \eta_{t,n}^N) - \frac{(K-1)^3}{K^4} \sum_{\substack{t=1 \\ t \neq l}}^K \text{Cov}(\eta_{k,i}^N \eta_{k,j}^N, \eta_{t,m}^N \eta_{l,n}^N) \\
&\quad + \frac{(K-1)^2}{K^4} \sum_{\substack{t=1 \\ t \neq l}}^K \sum_{s=1}^K \text{Cov}(\eta_{k,i}^N \eta_{k,j}^N, \eta_{t,m}^N \eta_{s,n}^N) - \frac{(K-1)^3}{K^4} \sum_{\substack{h=1 \\ h \neq k}}^K \text{Cov}(\eta_{k,i}^N \eta_{h,j}^N, \eta_{l,m}^N \eta_{l,n}^N) \\
&\quad + \frac{(K-1)^2}{K^4} \sum_{\substack{h=1 \\ h \neq k}}^K \sum_{\substack{t=1 \\ t \neq l}}^K \text{Cov}(\eta_{k,i}^N \eta_{h,j}^N, \eta_{l,m}^N \eta_{t,n}^N) + \frac{(K-1)^2}{K^4} \sum_{\substack{h=1 \\ h \neq k}}^K \sum_{\substack{t=1 \\ t \neq l}}^K \text{Cov}(\eta_{k,i}^N \eta_{h,j}^N, \eta_{t,m}^N \eta_{l,n}^N) \\
&\quad - \frac{K-1}{K^4} \sum_{\substack{h=1 \\ h \neq l}}^K \sum_{\substack{t=1 \\ t \neq l}}^K \sum_{s=1}^K \text{Cov}(\eta_{k,i}^N \eta_{h,j}^N, \eta_{t,m}^N \eta_{s,n}^N) - \frac{(K-1)^3}{K^4} \sum_{\substack{h=1 \\ h \neq k}}^K \text{Cov}(\eta_{h,i}^N \eta_{k,j}^N, \eta_{l,m}^N \eta_{l,n}^N) \\
&\quad + \frac{(K-1)^2}{K^4} \sum_{\substack{h=1 \\ h \neq k}}^K \sum_{\substack{t=1 \\ t \neq l}}^K \text{Cov}(\eta_{h,i}^N \eta_{k,j}^N, \eta_{l,m}^N \eta_{t,n}^N) + \frac{(K-1)^2}{K^4} \sum_{\substack{h=1 \\ h \neq k}}^K \sum_{\substack{t=1 \\ t \neq l}}^K \text{Cov}(\eta_{h,i}^N \eta_{k,j}^N, \eta_{t,m}^N \eta_{l,n}^N) \\
&\quad - \frac{K-1}{K^4} \sum_{\substack{h=1 \\ h \neq l}}^K \sum_{\substack{t=1 \\ t \neq l}}^K \sum_{\substack{s=1 \\ s \neq l}}^K \text{Cov}(\eta_{h,i}^N \eta_{k,j}^N, \eta_{t,m}^N \eta_{s,n}^N) + \frac{(K-1)^2}{K^4} \sum_{\substack{h=1 \\ h \neq k}}^K \sum_{\substack{p=1 \\ p \neq k}}^K \text{Cov}(\eta_{h,i}^N \eta_{p,j}^N, \eta_{l,m}^N \eta_{l,n}^N) \\
&\quad - \frac{K-1}{K^4} \sum_{\substack{h=1 \\ h \neq k}}^K \sum_{\substack{p=1 \\ p \neq k}}^K \sum_{\substack{t=1 \\ t \neq l}}^K \text{Cov}(\eta_{h,i}^N \eta_{p,j}^N, \eta_{l,m}^N \eta_{t,n}^N) - \frac{K-1}{K^4} \sum_{\substack{h=1 \\ h \neq k}}^K \sum_{\substack{p=1 \\ p \neq k}}^K \sum_{\substack{t=1 \\ t \neq l}}^K \text{Cov}(\eta_{h,i}^N \eta_{p,j}^N, \eta_{t,m}^N \eta_{l,n}^N) \\
&\quad \left. + \frac{1}{K^4} \sum_{\substack{h=1 \\ h \neq k}}^K \sum_{\substack{p=1 \\ p \neq k}}^K \sum_{\substack{t=1 \\ t \neq l}}^K \sum_{\substack{s=1 \\ s \neq l}}^K \text{Cov}(\eta_{h,i}^N \eta_{p,j}^N, \eta_{t,m}^N \eta_{s,n}^N) \right].
\end{aligned}$$

On simplifie chaque terme en calculant les covariances avec le Lemme 6, puis en réduisant les sommes de manière analogue à ce qui a été fait lors de la démonstration du Théorème 4. On a alors

$$\begin{aligned}
\text{Var}(\widehat{\nu_K^N}) &= \frac{(K-1)^2}{K^5 \text{Card}(I_N)} [12(K-1) + 8(K-2) + 2] \\
&= \frac{(K-1)^2}{K^5 \text{Card}(I_N)} [K^2 + 18K - 17],
\end{aligned}$$

ce qui permet de conclure que $\lim_{N,K \rightarrow \infty} \text{Var}(\widehat{\nu_K^N}) = 0$. □

Démonstration du Théorème 5. On démontre tout d'abord la convergence (4.10). Pour ce faire, on commence par décomposer :

$$|\mathbb{E}[\widehat{\nu_{K,\rho}^N} - \sigma^2]| \leq |\mathbb{E}[\widehat{\nu_{K,\rho}^N} - \widehat{\nu_K^N}]| + |\mathbb{E}[\widehat{\nu_K^N} - \sigma^2]|.$$

D'après le Lemme 7, le second terme tend vers 0 quand ρ tend vers 0 et N tend vers $+\infty$; il ne reste ainsi qu'à démontrer la convergence du premier terme. On rappelle les notations :

$$\begin{aligned}\overline{\epsilon}_K^{N,\rho} &= \frac{1}{K} \sum_{k=1}^K \epsilon_k^{N,\rho}, \\ \overline{\eta}_K^N &= \frac{1}{K} \sum_{k=1}^K \eta_k^N.\end{aligned}$$

On a, par les équations (4.6) et (4.9) :

$$\begin{aligned}|\mathbb{E}[\widetilde{\nu}_{K,\rho}^N - \widehat{\nu}_K^N]| &= \left| \frac{1}{\text{Card}(I_N)} \mathbb{E} \left[\frac{1}{K} \sum_{k=1}^K ((\eta_k^N - \overline{\eta}_K^N + \epsilon_k^{N,\rho} - \overline{\epsilon}_K^{N,\rho})^t (\Gamma^N)^{-1} (\eta_k^N - \overline{\eta}_K^N + \epsilon_k^{N,\rho} - \overline{\epsilon}_K^{N,\rho}) \right. \right. \\ &\quad \left. \left. - (\eta_k^N - \overline{\eta}_K^N)^t (\Gamma^N)^{-1} (\eta_k^N - \overline{\eta}_K^N) \right) \right],\end{aligned}$$

où Γ^N est la matrice de Gram du système $(\phi_i)_{i \in I_N}$. On développe et on réduit cette expression; les suites $(\epsilon_k^{N,\rho})_{k \in \mathbb{N}}$ et $(\eta_k^N)_{k \in \mathbb{N}}$ étant mutuellement indépendantes, on obtient :

$$\begin{aligned}|\mathbb{E}[\widetilde{\nu}_{K,\rho}^N - \widehat{\nu}_K^N]| &= \left| \frac{1}{K \text{Card}(I_N)} \sum_{k=1}^K \mathbb{E}[(\epsilon_k^{N,\rho} - \overline{\epsilon}_K^{N,\rho})^t (\Gamma^N)^{-1} (\epsilon_k^{N,\rho} - \overline{\epsilon}_K^{N,\rho})] \right|, \\ &= \left| \frac{1}{\text{Card}(I_N)} \mathbb{E}[(\epsilon_1^{N,\rho} - \overline{\epsilon}_K^{N,\rho})^t (\Gamma^N)^{-1} (\epsilon_1^{N,\rho} - \overline{\epsilon}_K^{N,\rho})] \right|.\end{aligned}$$

La suite $(\epsilon_k^{N,\rho})_{k \in \mathbb{N}}$ étant i.i.d. et la matrice $(\Gamma^N)^{-1}$ étant symétrique, ceci équivaut à :

$$|\mathbb{E}[\widetilde{\nu}_{K,\rho}^N - \widehat{\nu}_K^N]| = \left| \frac{1 - \frac{1}{K}}{\text{Card}(I_N)} \mathbb{E}[(\epsilon_1^{N,\rho})^t (\Gamma^N)^{-1} \epsilon_1^{N,\rho}] \right|.$$

Comme la matrice Γ^N est une matrice de Gram inversible, il existe une matrice $M_N := (m_{i,j}^N)_{i,j \in I_N}$ telle que $(\Gamma^N)^{-1} = M_N^t M_N$. On a alors :

$$|\mathbb{E}[\widetilde{\nu}_{K,\rho}^N - \widehat{\nu}_K^N]| = \left| \frac{1 - \frac{1}{K}}{\text{Card}(I_N)} \mathbb{E}[(M_N \epsilon_1^{N,\rho})^t (M_N \epsilon_1^{N,\rho})] \right|,$$

D'où, par linéarité de $\epsilon_{1,\rho}$:

$$\begin{aligned}|\mathbb{E}[\widetilde{\nu}_{K,\rho}^N - \widehat{\nu}_K^N]| &= \frac{1 - \frac{1}{K}}{\text{Card}(I_N)} \sum_{i \in I_N} \mathbb{E}[(\epsilon_{1,\rho}(\sum_{j \in I_N} m_{i,j}^N \phi_j))^2], \\ &\leq \frac{1 - \frac{1}{K}}{\text{Card}(I_N)} \sum_{i \in I_N} \left\| \sum_{j \in I_N} m_{i,j}^N \phi_j \right\|_{\mathbb{V}}^2 \|\epsilon_{1,\rho}\|_{\mathbb{V}_\mathcal{E}}^2, \\ &\leq \frac{(1 - \frac{1}{K}) \|\epsilon_{1,\rho}\|_{\mathbb{V}_\mathcal{E}^*}^2}{\text{Card}(I_N)} \sum_{j \in I_N} \sum_{n \in I_N} g_{n,j} \sum_{i \in I_N} m_{i,j}^N m_{i,n}^N.\end{aligned}$$

Vu que $M_N^t M_N = (\Gamma^N)^{-1}$, on a, pour tous m, n de I_N , $\sum_{i \in I_N} m_{i,j}^N m_{i,n}^N = h_{j,n}^N$. On a alors :

$$\begin{aligned} |\mathbb{E}[\widetilde{\nu}_{K,\rho}^N - \widehat{\nu}_K^N]| &\leq \frac{(1 - \frac{1}{K}) \|\epsilon_{1,\rho}\|_{V_{\mathcal{E}}^*}^2}{\text{Card}(I_N)} \sum_{j \in I_N} \sum_{n \in I_N} g_{n,j} h_{j,n}^N \\ &= \frac{(1 - \frac{1}{K}) \|\epsilon_{1,\rho}\|_{V_{\mathcal{E}}^*}^2}{\text{Card}(I_N)} \sum_{j \in I_N} \delta_{j,j} \\ &= (1 - \frac{1}{K}) \|\epsilon_{1,\rho}\|_{V_{\mathcal{E}}^*}^2. \end{aligned}$$

Le Théorème 3 permet alors de conclure.

On démontre maintenant la convergence (4.11) ; pour ce faire, on commence par décomposer :

$$\mathbb{E}[(\widetilde{\nu}_{K,\rho}^N - \widehat{\nu}_K^N)^2] = \text{Var}(\widetilde{\nu}_{K,\rho}^N - \widehat{\nu}_K^N) + (\mathbb{E}[\widetilde{\nu}_{K,\rho}^N - \widehat{\nu}_K^N])^2.$$

La convergence vers 0 du second terme quand N et K tendent vers $+\infty$ et ρ tend vers 0 a déjà été établie ; il suffit donc de démontrer la convergence du premier terme. On commence par développer ce dernier en utilisant le caractère i.i.d. des suites $(\epsilon_k^{N,\rho})_{k \in \mathbb{N}}$ et $(\eta_k^N)_{k \in \mathbb{N}}$ ainsi que la symétrie de $(\Gamma^N)^{-1}$.

$$\begin{aligned} \text{Var}(\widetilde{\nu}_{K,\rho}^N - \widehat{\nu}_K^N) &= \frac{1}{(\text{Card}(I_N))^2 K} \text{Var}((\epsilon_1^{N,\rho} - \overline{\epsilon_K^{N,\rho}})^t (\Gamma^N)^{-1} (\epsilon_1^{N,\rho} - \overline{\epsilon_K^{N,\rho}}) \\ &\quad - 2(\epsilon_1^{N,\rho} - \overline{\epsilon_K^{N,\rho}})^t (\Gamma^N)^{-1} (\eta_1^{N,\rho} - \overline{\eta_K^N}), \\ &\leq \frac{1}{(\text{Card}(I_N))^2 K} \mathbb{E}[(\epsilon_1^{N,\rho} - \overline{\epsilon_K^{N,\rho}})^t (\Gamma^N)^{-1} (\epsilon_1^{N,\rho} - \overline{\epsilon_K^{N,\rho}}) \\ &\quad - 2(\epsilon_1^{N,\rho} - \overline{\epsilon_K^{N,\rho}})^t (\Gamma^N)^{-1} (\eta_1^{N,\rho} - \overline{\eta_K^N})]^2], \\ &= \frac{1}{(\text{Card}(I_N))^2 K} \mathbb{E}[(\epsilon_1^{N,\rho} - \overline{\epsilon_K^{N,\rho}})^t (\Gamma^N)^{-1} (\epsilon_1^{N,\rho} - \overline{\epsilon_K^{N,\rho}})]^2 \\ &\quad + \frac{4}{(\text{Card}(I_N))^2 K} \mathbb{E}[(\epsilon_1^{N,\rho} - \overline{\epsilon_K^{N,\rho}})^t (\Gamma^N)^{-1} (\eta_1^{N,\rho} - \overline{\eta_K^N})]^2 \\ &\quad - \frac{4}{(\text{Card}(I_N))^2 K} \mathbb{E}[(\epsilon_1^{N,\rho} - \overline{\epsilon_K^{N,\rho}})^t (\Gamma^N)^{-1} (\epsilon_1^{N,\rho} - \overline{\epsilon_K^{N,\rho}}) (\Gamma^N)^{-1}] \mathbb{E}[\eta_1^{N,\rho} - \overline{\eta_K^N}], \\ &= \frac{1}{(\text{Card}(I_N))^2 K} \mathbb{E}[(\epsilon_1^{N,\rho} - \overline{\epsilon_K^{N,\rho}})^t (\Gamma^N)^{-1} (\epsilon_1^{N,\rho} - \overline{\epsilon_K^{N,\rho}})]^2 \\ &\quad + \frac{4}{(\text{Card}(I_N))^2 K} \mathbb{E}[(\epsilon_1^{N,\rho} - \overline{\epsilon_K^{N,\rho}})^t (\Gamma^N)^{-1} (\eta_1^{N,\rho} - \overline{\eta_K^N})]^2. \end{aligned}$$

On développe à présent les produits matriciels :

$$\begin{aligned}
\text{Var}(\widetilde{\nu}_{K,\rho}^N - \widehat{\nu}_K^N) &\leq \left| \frac{(K-1)^4}{(\text{Card}(I_N))^2 K^5} \mathbb{E}[((\epsilon_1^{N,\rho})^t (\Gamma^N)^{-1} \epsilon_1^{N,\rho})^2] \right| + \left| \frac{2(K-1)^2}{(\text{Card}(I_N))^2 K^5} \mathbb{E}\left[\left(\sum_{k=1}^K (\epsilon_1^{N,\rho})^t (\Gamma^N)^{-1} \epsilon_k^{N,\rho}\right)^2\right] \right| \\
&+ \frac{1}{(\text{Card}(I_N))^2 K^5} \mathbb{E}\left[\left(\sum_{k=1}^K \sum_{l=1}^K (\epsilon_k^{N,\rho})^t (\Gamma^N)^{-1} \epsilon_l^{N,\rho}\right)^2\right] + \left| \frac{4(K-1)^4}{(\text{Card}(I_N))^2 K^5} \mathbb{E}[((\epsilon_1^{N,\rho})^t (\Gamma^N)^{-1} \eta_1^N)^2] \right| \\
&+ \left| \frac{4(K-1)^2}{(\text{Card}(I_N))^2 K^5} \mathbb{E}\left[\left(\sum_{k=1}^K (\epsilon_1^{N,\rho})^t (\Gamma^N)^{-1} \eta_k^N\right)^2\right] \right| + \left| \frac{4(K-1)^2}{(\text{Card}(I_N))^2 K^5} \mathbb{E}\left[\left(\sum_{k=1}^K (\epsilon_k^{N,\rho})^t (\Gamma^N)^{-1} \eta_1^N\right)^2\right] \right| \\
&+ \left| \frac{4}{(\text{Card}(I_N))^2 K^5} \mathbb{E}\left[\left(\sum_{k=1}^K \sum_{l=1}^K (\epsilon_k^{N,\rho})^t (\Gamma^N)^{-1} \eta_l^N\right)^2\right] \right|.
\end{aligned}$$

On réduit chaque terme de la somme en utilisant à nouveau le fait que $(\Gamma^N)^{-1} = (M_N)^t M_N$. On élimine la dépendance en N de façon similaire à ce qui est fait plus haut, dans la démonstration de la convergence (4.10). On obtient :

$$\text{Var}(\widetilde{\nu}_{K,\rho}^N - \widehat{\nu}_K^N) \leq 4 \frac{(K-1)^4}{K^5} \sup_{\|w\|_V=1} \mathbb{E}[(\epsilon_{1,\rho}(w))^4] + 8\sigma^2 \frac{(K-1)^4 + (K-1)^3}{K^5} \sup_{\|w\|_V=1} \mathbb{E}[(\epsilon_{1,\rho}(w))^2].$$

D'après l'Hypothèse 3, on a, pour ρ suffisamment proche de 0 et pour tout w dans la boule-unité de V , $|\epsilon_{1,\rho}| \leq 1$ p.s., ce qui implique que $(\epsilon_{1,\rho})^4 \leq (\epsilon_{1,\rho})^2$ p.s. et donc $\mathbb{E}[(\epsilon_{1,\rho}(w))^4] \leq \mathbb{E}[(\epsilon_{1,\rho}(w))^2]$. On a alors :

$$\text{Var}(\widetilde{\nu}_{K,\rho}^N - \widehat{\nu}_K^N) \leq \left(4 \frac{(K-1)^4}{K^5} + 8\sigma^2 \frac{(K-1)^4 + (K-1)^3}{K^5} \right) \sup_{\|w\|_V=1} \mathbb{E}[(\epsilon_{1,\rho}(w))^2].$$

On conclut à l'aide du Théorème 3. □

Chapitre 5

Estimation (dans le cadre d'un modèle général de bruit gaussien)

5.1 Objectifs

On considère maintenant les Modèles (3.2) et (3.4), sous l'hypothèse plus générale : $(B_k)_{1 \leq k \leq K}$ est une suite i.i.d. de formes linéaires aléatoires gaussiennes d'espérance identiquement nulle et d'autocovariance $\Gamma \in V^{**}$, inconnue et à estimer. Cette hypothèse est plus en accord avec la réalité expérimentale, dans laquelle on ne connaîtra généralement pas, a priori, l'opérateur Γ . L'objectif de ce chapitre est d'estimer le représentant moyen μ ainsi que la forme bilinéaire Γ à partir des observations $S_{1,\rho}, \dots, S_{K,\rho}$. Dans la suite, on conservera toutes les notations des deux chapitres précédents.

Afin de disposer d'un cadre approprié à l'estimation de Γ , qui est un élément de V^{**} , on va dans un premier temps introduire une notion de forme bilinéaire aléatoire. De manière analogue à ce qui a été fait en Section 3.2.2 pour les formes linéaires aléatoires, on utilisera les processus aléatoires généralisés. Dans un deuxième temps, on mettra en place l'estimation de μ et Γ . On procèdera à la manière de ce qui a été fait au Chapitre 4 : construction d'estimateurs dans le modèle continu puis extension au modèle discret. Enfin, dans un troisième temps, on établira la consistance des estimateurs sous les hypothèses courantes sur le bruit.

5.2 Formes bilinéaires aléatoires

Dans cette section, on établit, de manière analogue à ce qui a été fait en Section 3.2.2 sur les formes linéaires aléatoires, une notion de forme bilinéaire aléatoire, basée sur les processus stochastiques généra-

lisés. On donne tout d'abord quelques définitions puis on met en place sur ces objets une structure d'espace de Banach.

5.2.1 Définitions

Dans la suite, soit $\mathcal{E} = (\Omega, \mathcal{F}, \mathbb{P})$ un espace de probabilité et soit V un espace de Hilbert.

Définition 17 (Forme bilinéaire aléatoire). *On appelle forme bilinéaire aléatoire sur V toute application S bilinéaire et continue de $V \times V$ dans $L^2(\mathcal{E})$. On note $V_{\mathcal{E}}^{**}$ l'espace des formes bilinéaires aléatoires sur V .*

Définition 18 (Espérance). *On définit l'espérance d'une forme bilinéaire aléatoire S comme l'application*

$$\begin{aligned} \mathbb{E}^{**}[S] : V \times V &\longrightarrow \mathbb{R} \\ (v, w) &\longrightarrow \mathbb{E}[S(v, w)]. \end{aligned}$$

Proposition 5. *Pour tout S de $V_{\mathcal{E}}^{**}$, l'application $\mathbb{E}^{**}[S]$ est une forme bilinéaire continue sur V .*

Démonstration. Soit $S \in V_{\mathcal{E}}^{**}$. Par continuité de S , il existe une constante positive C_S telle que

$$\forall v, w \in V, \|S(v, w)\|_{L^2(\mathcal{E})} \leq C_S \|v\|_V \|w\|_V.$$

Or $\forall v, w \in V$, on a

$$|\mathbb{E}^{**}[S](v, w)| = |\mathbb{E}[S(v, w)]| \leq \|S(v, w)\|_{L^2(\mathcal{E})}.$$

Donc, $\forall v, w \in V$,

$$|\mathbb{E}^{**}[S](v, w)| \leq C_S \|v\|_V \|w\|_V,$$

d'où la continuité de la forme bilinéaire $\mathbb{E}^{**}[S]$. □

5.2.2 Structure d'espace de Banach

De manière analogue à ce qui a été fait pour les formes linéaires aléatoires en Section 3.2.2, on définit sur $V_{\mathcal{E}}^{**}$ la somme et le produit suivants

$$\begin{aligned} \forall S, T \in V_{\mathcal{E}}^{**}, \forall \lambda \in \mathbb{R}, \\ \lambda S : V \times V &\longrightarrow L^2(\mathcal{E}) \\ (v, w) &\longrightarrow \lambda S(v, w) \\ S + T : V \times V &\longrightarrow L^2(\mathcal{E}) \\ (v, w) &\longrightarrow S(v, w) + T(v, w). \end{aligned}$$

Les applications λS et $S + T$ ainsi définies sont toujours des formes bilinéaires aléatoires. L'espace $V_{\mathcal{E}}^{**}$ est donc un \mathbb{R} -espace vectoriel. C'est un espace de Banach au vu de la proposition ci-dessous.

Proposition 6. *L'application $\|\cdot\|_{V_{\mathcal{E}}^{**}}$ définie par*

$$\begin{aligned} \|\cdot\|_{V_{\mathcal{E}}^{**}} : V_{\mathcal{E}}^{**} &\longrightarrow \mathbb{R}^+ \\ S &\longrightarrow \sup_{\|w\|_{V_{\mathcal{E}}^{**}}} \sup_{\|v\|_{V_{\mathcal{E}}^{**}}} \|S(v, w)\|_{L^2(\mathcal{E})} \end{aligned}$$

*est une norme sur $V_{\mathcal{E}}^{**}$; de plus, l'espace vectoriel normé $(V_{\mathcal{E}}^{**}, \|\cdot\|_{V_{\mathcal{E}}^{**}})$ est un espace de Banach.*

Démonstration. La démonstration est analogue à celle de la Proposition 4 de la Section 3.2.2.

Le fait que $\|\cdot\|_{V_{\mathcal{E}}^{**}}$ est une norme sur $V_{\mathcal{E}}^{**}$ est conséquence immédiate des propriétés de norme de $\|\cdot\|_{L^2(\mathcal{E})}$.

On démontre ensuite que $(V_{\mathcal{E}}^{**}, \|\cdot\|_{V_{\mathcal{E}}^{**}})$ est complet. Soit $(S_n)_{n \in \mathbb{N}}$ une suite d'éléments de $V_{\mathcal{E}}^{**}$ qui soit une suite de Cauchy au sens de la norme $\|\cdot\|_{V_{\mathcal{E}}^{**}}$. Pour tout entier naturel n , on note \tilde{S}_n la restriction de S_n à $B_V \times B_V$ (B_V désignant toujours la boule-unité de V). La suite (\tilde{S}_n) vérifie alors un critère de Cauchy uniforme sur $B_V \times B_V$ et, par conséquent, converge uniformément vers une fonction \tilde{S} , de $B_V \times B_V$ dans $L^2(\mathcal{E})$. On définit alors

$$\begin{aligned} S : V \times V &\longrightarrow L^2(\mathcal{E}) \\ (v, w) &\longrightarrow \|v\|_V \|w\|_V \tilde{S}(\tilde{v}, \tilde{w}), \end{aligned}$$

où \tilde{v} et \tilde{w} sont définis comme dans la démonstration de la Proposition 3. Il est alors simple, comme dans cette dernière, de montrer que S est un élément de $V_{\mathcal{E}}^{**}$ et que S_n converge vers S pour la norme $\|\cdot\|_{V_{\mathcal{E}}^{**}}$. On en conclut que l'espace $(V_{\mathcal{E}}^{**}, \|\cdot\|_{V_{\mathcal{E}}^{**}})$ est complet. □

5.3 Construction des estimateurs

5.3.1 Notations

Pour tout i et pour tout j de I , on note $\gamma_{i,j} = \Gamma(\phi_i, \phi_j)$. On a en particulier :

$$\forall u, v \in V, \Gamma(u, v) = \sum_{i \in I} \sum_{j \in I} \gamma_{i,j} u_i v_j,$$

où $u = \sum_{i \in I} u_i \phi_i$ et $v = \sum_{i \in I} v_i \phi_i$.

Pour tout entier naturel N , on notera Γ^N la forme bilinéaire sur V définie par :

$$\forall i, j \in I, \Gamma^N(\phi_i, \phi_j) = \gamma_{i,j} \text{ si } i \in I_N \text{ et } j \in I_N, \Gamma^N(\phi_i, \phi_j) = 0 \text{ si } i \notin I_N \text{ ou } j \notin I_N.$$

De fait, on a :

$$\forall u, v \in V, \Gamma^N(u, v) = \sum_{i \in I_N} \sum_{j \in I_N} \gamma_{i,j} u_i v_j,$$

où $u = \sum_{i \in I} u_i \phi_i$ et $v = \sum_{i \in I} v_i \phi_i$.

Remarque 4. Γ^N est un élément de V^{**} ; de plus on a

$$\|\Gamma^N\|_{V^{**}} \leq \|\Gamma\|_{V^{**}}.$$

5.3.2 Estimation du représentant moyen

Les estimateurs $\widehat{\mu}_K^N$ et $\widetilde{\mu}_{K,\rho}^N$ de la forme linéaire moyenne μ développés au Chapitre 4 ne dépendent pas de l'autocovariance Γ du bruit. Il est par conséquent naturel de conserver les mêmes estimateurs pour le représentant moyen, dans le modèle de variation comme dans le modèle d'observation.

Dans cette optique, on estime le représentant moyen μ , dans le modèle continu, par la forme linéaire aléatoire $\widehat{\mu}_K^N$ définie par l'Équation (4.4) en Section 4.3.1. Dans le modèle discret, μ est estimé par $\widetilde{\mu}_{K,\rho}^N$, défini par l'Équation (4.7) en Section 4.3.2.

5.3.3 Estimateur de l'opérateur d'autocovariance

Estimateur dans le modèle de variation

Comme cela a été fait pour les estimateurs du Chapitre 4, on estime tout d'abord l'opérateur d'autocovariance Γ dans le cas, théorique, d'un échantillon de surfaces lisses dont les représentants dans V^* varient selon le modèle (3.2). On étendra par la suite cet estimateur au cas d'un échantillon de surfaces discrètes.

On estime la forme bilinéaire $\Gamma \in V^{**}$ par la forme bilinéaire aléatoire $\widehat{\Gamma}_K^N$ définie par

$$\forall v, w \in V, \widehat{\Gamma}_K^N(v, w) = \sum_{i \in I_N} \sum_{j \in I_N} v_i w_j \widehat{\gamma}_{K,i,j}^N,$$

$$\text{avec } \forall i, j \in I, \widehat{\gamma}_{K,i,j}^N = \frac{1}{K} \mathbf{1}_{I_N}(i, j) \sum_{k=1}^K (S_k(\phi_i) - \overline{S}_K(\phi_i)) h_{i,j}^N (S_k(\phi_j) - \overline{S}_K(\phi_j)).$$

Comme il a été dit au Chapitre 4, les surfaces lisses ne sont pas observées en pratique. Pour pouvoir estimer l'opérateur Γ dans le cadre d'une situation expérimentale, il convient ainsi de réécrire notre estimateur en termes de représentants de surfaces discrètes.

Estimation dans le modèle d'observation

On construit, de manière intuitive, un estimateur de Γ dans le modèle d'observation (3.4) à partir de l'estimateur dans le modèle de variation établi plus haut, de manière analogue à ce qui a été fait au Chapitre

4. Dans cette optique, on reprend la formule de l'estimateur $\widehat{\Gamma}_K^N$ et on remplace toutes les occurrences des S_k , $k \in \{1, \dots, K\}$ par les représentants S_k^ρ , $k \in \{1, \dots, K\}$ des approximations discrètes respectivement associées. On estime ainsi Γ dans le modèle (3.4) par la forme bilinéaire aléatoire $\widetilde{\Gamma}_K^{N,\rho}$ définie par

$$\forall v, w \in V, \widetilde{\Gamma}_K^{N,\rho}(v, w) = \sum_{i \in I_N} \sum_{j \in I_N} v_i w_j \widetilde{\gamma}_{K,i,j}^{N,\rho},$$

$$\text{avec } \forall i, j \in I, \widetilde{\gamma}_{K,i,j}^{N,\rho} = \frac{1}{K} \mathbb{1}_{I_N}(i, j) \sum_{k=1}^K (S_k^\rho(\phi_i) - \overline{S}_K^\rho(\phi_i)) h_{i,j}^N (S_k^\rho(\phi_j) - \overline{S}_K^\rho(\phi_j)).$$

5.4 Consistance

Consistance de l'estimateur du représentant moyen

Théorème 6. *Le Théorème 4 et le Lemme 4 se généralisent sous l'hypothèse faite en Section 3.1, i.e. pour $(B_k)_{k \in \{1, \dots, K\}}$ suite i.i.d. de formes linéaires aléatoires Gaussiennes centrées.*

Démonstration. La démonstration est totalement analogue à ce qui a été fait au Chapitre 4, à cela près qu'on remplace le Lemme 5 par le Lemme 8, énoncé ci-dessous. \square

Lemme 8. *Sous les hypothèses faites dans ce chapitre quant à la suite $(B_k)_{k \in \{1, \dots, K\}}$, on a pour tout entier naturel non nul N ,*

$$\sup_{\|w\|_V \leq 1} \mathbb{E} \left[\left(\widehat{\mu}_1^N(w) - \mu^N(w) \right)^2 \right] \leq \|\Gamma\|_{V^{**}}^2.$$

Démonstration. Soient m et n deux éléments de I . Soit $N' \in \mathbb{N}$ tel que $m, n \in I_N$ et soit $N \in \mathbb{N}$ tel que $N > N'$.

En passant aux représentants de Riesz, on a :

$$\begin{aligned} \mathbb{E} \left[\left(\widehat{\mu}_1^N(\phi_m) - \mu^N(\phi_m) \right) \left(\widehat{\mu}_1^N(\phi_n) - \mu^N(\phi_n) \right) \right] &= \mathbb{E} \left[\langle v_1^N - v_\mu^N, \phi_m \rangle_V \langle v_1^N - v_\mu^N, \phi_n \rangle_V \right] \\ &= \mathbb{E} \left[\left\langle \sum_{i \in I_N} \phi_i \sum_{j \in I_N} h_{i,j}^N B_1(\phi_j), \phi_m \right\rangle_V \left\langle \sum_{s \in I_N} \phi_s \sum_{t \in I_N} h_{s,t}^N B_1(\phi_t), \phi_n \right\rangle_V \right] \\ &= \sum_{i \in I_N} \sum_{j \in I_N} \sum_{s \in I_N} \sum_{t \in I_N} (h_{i,j}^N h_{s,t}^N \langle \phi_i, \phi_m \rangle_V \langle \phi_s, \phi_n \rangle_V \mathbb{E} [B_1(\phi_j) B_1(\phi_t)]) \\ &= \sum_{j \in I_N} \sum_{t \in I_N} \left(\mathbb{E} [B_1(\phi_j) B_1(\phi_t)] \left(\sum_{i \in I_N} h_{i,j}^N g_{m,i} \right) \left(\sum_{s \in I_N} h_{s,t}^N g_{n,s} \right) \right) \\ &= \sum_{j \in I_N} \sum_{t \in I_N} (\gamma_{j,t} \delta_{j,m} \delta_{t,n}) \\ &= \gamma_{m,n}. \end{aligned}$$

Soient $w \in V$ et $w^N = \sum_{i \in I_N} \alpha_i^N \phi_i$ sa projection orthogonale sur V_N . On a

$$\begin{aligned}
\mathbb{E} \left[\left(\widehat{\mu}_1^N(w) - \mu^N(w) \right)^2 \right] &= \mathbb{E} \left[\left(\widehat{\mu}_1^N(w^N) - \mu^N(w^N) \right)^2 \right] \\
&= \sum_{i \in I_N} \sum_{j \in I_N} \alpha_i^N \alpha_j^N \mathbb{E} \left[\left(\widehat{\mu}_1^N(\phi_i) - \mu^N(\phi_i) \right) \left(\widehat{\mu}_1^N(\phi_j) - \mu^N(\phi_j) \right) \right] \\
&= \sum_{i \in I_N} \sum_{j \in I_N} \alpha_i^N \alpha_j^N \gamma_{i,j} \\
&= \Gamma^N(w, w) \\
&\leq \|w\|_V^2 \|\Gamma^N\|_{V^{**}}.
\end{aligned}$$

On obtient le résultat en appliquant la Remarque 4 puis en passant à la borne supérieure pour w dans la boule-unité de V . □

Consistance de l'estimateur de l'autocovariance du bruit

On énonce à présent la consistance de l'estimateur $\widetilde{\Gamma}_{K,\rho}^N$ lorsque la taille K de l'échantillon et le paramètre de dimension de l'espace d'approximation N tendent vers l'infini et quand le paramètre de précision de la discrétisation ρ tend vers 0. On démontre ce résultat en utilisant comme résultat intermédiaire la consistance de l'estimateur $\widehat{\Gamma}_K^N$.

Théorème 7. *La forme bilinéaire aléatoire $\widetilde{\Gamma}_{K,\rho}^N$ converge vers Γ dans l'espace $V_{\mathcal{E}}^{**}$ quand K et N tendent vers $+\infty$ et quand ρ tend vers 0; en d'autres termes, on a*

$$\lim_{\substack{\rho \rightarrow 0 \\ K, N \rightarrow \infty}} \|\widetilde{\Gamma}_{K,\rho}^N - \Gamma\|_{V_{\mathcal{E}}^{**}} = 0.$$

Lemme 9. *La forme bilinéaire aléatoire $\widehat{\Gamma}_K^N$ converge vers Γ dans l'espace $V_{\mathcal{E}}^{**}$ quand K et N tendent vers $+\infty$, i.e.*

$$\lim_{K, N \rightarrow \infty} \|\widehat{\Gamma}_K^N - \Gamma\|_{V_{\mathcal{E}}^{**}} = 0.$$

Démonstration. On commence par décomposer par inégalité triangulaire

$$\begin{aligned}
\|\widehat{\Gamma}_K^N - \Gamma\|_{V_{\mathcal{E}}^{**}} &\leq \|\Gamma - \Gamma^N\|_{V_{\mathcal{E}}^{**}} + \|\widehat{\Gamma}_K^N - \Gamma^N\|_{V_{\mathcal{E}}^{**}} \\
&= \|\Gamma - \Gamma^N\|_{V_{\mathcal{E}}^{**}} + \|\widehat{\Gamma}_K^N - \Gamma^N\|_{V_{\mathcal{E}}^{**}}.
\end{aligned}$$

Le premier terme de la somme converge vers 0 quand N et K tendent vers $+\infty$ d'après la Remarque 4 (voir Section 5.3.1).

On montre maintenant que le deuxième terme converge vers 0 quand N et K tendent vers $+\infty$. Soient $u, v \in V$.

On pose $u = \sum_{i \in I} u_i \phi_i$ et $v = \sum_{i \in I} v_i \phi_i$.

$$\begin{aligned}
\mathbb{E} \left[\left| \widehat{\Gamma}_K^N(u, v) - \Gamma(u, v) \right|^2 \right] &= \mathbb{E} \left[\left| \sum_{i \in I} u_i \sum_{j \in I} v_j (\widehat{\gamma}_{K,i,j}^N - \gamma_{i,j}) \right|^2 \right] \\
&= \sum_{i \in I} \sum_{m \in I} \sum_{j \in I} \sum_{n \in I} u_i u_m v_j v_n \mathbb{E} \left[(\widehat{\gamma}_{K,i,j}^N - \gamma_{i,j}) (\widehat{\gamma}_{K,m,n}^N - \gamma_{m,n}) \right] \\
&= \sum_{i \in I} \sum_{m \in I} \sum_{j \in I} \sum_{n \in I} u_i u_m v_j v_n \mathbb{E} \left[\frac{1}{K^2} \sum_{k=1}^K \sum_{l=1}^K ((B_k(\phi_i) - \overline{B}_K(\phi_i))(B_k(\phi_j - \overline{B}_K(\phi_j)) - \gamma_{i,j}) \right. \\
&\quad \left. \times ((B_l(\phi_m) - \overline{B}_K(\phi_m))(B_l(\phi_n - \overline{B}_K(\phi_n)) - \gamma_{m,n})) \right].
\end{aligned}$$

On développe le produit et on utilise la linéarité de l'espérance. On obtient :

$$\begin{aligned}
& \mathbb{E} \left[\left| \widehat{\Gamma}_K^N(u, v) - \Gamma^N(u, v) \right|^2 \right] \\
&= \frac{1}{K^2} \sum_{k=1}^K \sum_{l=1}^K \sum_{i \in I_N} \sum_{m \in I_N} \sum_{j \in I_N} \sum_{n \in I_N} u_i u_m v_j v_n \left[\frac{(K-1)^4}{K^4} \mathbb{E} [B_k(\phi_i) B_k(\phi_j) B_l(\phi_m) B_l(\phi_n)] \right. \\
&\quad - \frac{(K-1)^3}{K^4} \sum_{\substack{t=1 \\ t \neq l}}^K \mathbb{E} [B_k(\phi_i) B_k(\phi_j) B_l(\phi_m) B_t(\phi_n)] - \frac{(K-1)^3}{K^4} \sum_{\substack{t=1 \\ t \neq l}}^K \mathbb{E} [B_k(\phi_i) B_k(\phi_j) B_t(\phi_m) B_l(\phi_n)] \\
&\quad + \frac{1}{K^2} \sum_{\substack{t=1 \\ t \neq l}}^K \sum_{\substack{s=1 \\ s \neq l}}^K \mathbb{E} [B_k(\phi_i) B_k(\phi_j) B_t(\phi_m) B_s(\phi_n)] - \frac{(K-1)^3}{K^4} \sum_{\substack{t=1 \\ t \neq k}}^K \mathbb{E} [B_k(\phi_i) B_t(\phi_j) B_l(\phi_m) B_l(\phi_n)] \\
&\quad - \frac{(K-1)^2}{K^2} \gamma_{m,n} \mathbb{E} [B_k(\phi_i) B_k(\phi_j)] + \frac{(K-1)^2}{K^4} \sum_{\substack{t=1 \\ t \neq k}}^K \sum_{\substack{s=1 \\ s \neq l}}^K \mathbb{E} [B_k(\phi_i) B_t(\phi_j) B_l(\phi_m) B_s(\phi_n)] \\
&\quad + \frac{(K-1)^2}{K^4} \sum_{\substack{t=1 \\ t \neq k}}^K \sum_{\substack{s=1 \\ s \neq l}}^K \mathbb{E} [B_k(\phi_i) B_t(\phi_j) B_s(\phi_m) B_l(\phi_n)] + \frac{(K-1)^2}{K^4} \sum_{\substack{t=1 \\ t \neq k}}^K \sum_{\substack{s=1 \\ s \neq l}}^K \mathbb{E} [B_t(\phi_i) B_k(\phi_j) B_s(\phi_m) B_l(\phi_n)] \\
&\quad - \frac{k-1}{K^4} \sum_{\substack{t=1 \\ t \neq k}}^K \sum_{\substack{s=1 \\ s \neq l}}^K \sum_{\substack{x=1 \\ x \neq l}}^K \mathbb{E} [B_k(\phi_i) B_t(\phi_j) B_s(\phi_m) B_x(\phi_n)] - \frac{(k-1)^3}{K^4} \sum_{\substack{t=1 \\ t \neq k}}^K \mathbb{E} [B_t(\phi_i) B_k(\phi_j) B_l(\phi_m) B_l(\phi_n)] \\
&\quad + \frac{K-1}{K^2} \sum_{\substack{t=1 \\ t \neq k}}^K \gamma_{m,n} \mathbb{E} [B_k(\phi_i) B_t(\phi_j)] - \frac{K-1}{K^4} \sum_{\substack{t=1 \\ t \neq k}}^K \sum_{\substack{s=1 \\ s \neq l}}^K \sum_{\substack{x=1 \\ x \neq l}}^K \mathbb{E} [B_t(\phi_i) B_k(\phi_j) B_s(\phi_m) B_x(\phi_n)] \\
&\quad + \frac{(K-1)^2}{K^4} \sum_{\substack{t=1 \\ t \neq k}}^K \sum_{\substack{s=1 \\ s \neq l}}^K \mathbb{E} [B_t(\phi_i) B_k(\phi_j) B_l(\phi_m) B_s(\phi_n)] + \frac{(K-1)^2}{K^4} \sum_{\substack{t=1 \\ t \neq k}}^K \sum_{\substack{s=1 \\ s \neq l}}^K \mathbb{E} [B_t(\phi_i) B_s(\phi_j) B_l(\phi_m) B_l(\phi_n)] \\
&\quad - \frac{K-1}{K^4} \sum_{\substack{t=1 \\ t \neq k}}^K \sum_{\substack{s=1 \\ s \neq l}}^K \sum_{\substack{x=1 \\ x \neq l}}^K \mathbb{E} [B_t(\phi_i) B_s(\phi_j) B_l(\phi_m) B_x(\phi_n)] + \frac{K-1}{K^4} \sum_{\substack{t=1 \\ t \neq k}}^K \gamma_{m,n} \mathbb{E} [B_t(\phi_i) B_k(\phi_j)] \\
&\quad - \frac{K-1}{K^4} \sum_{\substack{t=1 \\ t \neq k}}^K \sum_{\substack{s=1 \\ s \neq l}}^K \sum_{\substack{x=1 \\ x \neq l}}^K \mathbb{E} [B_t(\phi_i) B_s(\phi_j) B_x(\phi_m) B_l(\phi_n)] - \frac{1}{K^2} \sum_{\substack{t=1 \\ t \neq k}}^K \sum_{\substack{s=1 \\ s \neq l}}^K \gamma_{m,n} \mathbb{E} [B_t(\phi_i) B_s(\phi_j)] \\
&\quad - \frac{(K-1)^2}{K^2} \gamma_{i,j} \mathbb{E} [B_l(\phi_m) B_l(\phi_n)] + \frac{K-1}{K^2} \sum_{\substack{t=1 \\ t \neq l}}^K \gamma_{i,j} \mathbb{E} [B_l(\phi_m) B_t(\phi_n)] \\
&\quad + \frac{K-1}{K^2} \sum_{\substack{t=1 \\ t \neq l}}^K \gamma_{i,j} \mathbb{E} [B_t(\phi_m) B_l(\phi_n)] - \frac{1}{K^2} \sum_{\substack{t=1 \\ t \neq l}}^K \sum_{\substack{s=1 \\ s \neq l}}^K \gamma_{i,j} \mathbb{E} [B_t(\phi_m) B_s(\phi_n)] \\
&\quad + \sum_{\substack{t=1 \\ t \neq k}}^K \sum_{\substack{s=1 \\ s \neq l}}^K \sum_{\substack{x=1 \\ x \neq l}}^K \sum_{\substack{y=1 \\ y \neq l}}^K \mathbb{E} [B_t(\phi_i) B_s(\phi_j) B_x(\phi_m) B_y(\phi_n)] + \gamma_{i,j} \gamma_{m,n}].
\end{aligned}$$

On continue les calculs en utilisant le fait que $(B_k)_{1 \leq k \leq K}$ est une suite i.i.d. de formes linéaires aléatoires gaussiennes d'espérance nulle et d'autocovariance Γ . On applique ensuite le Lemme 6 énoncé en Section 4.4.2 ainsi que la

définition de la matrice Γ^N . On obtient alors :

$$\begin{aligned}
& \mathbb{E} \left[\left| \widehat{\Gamma}_K^N(u, v) - \Gamma^N(u, v) \right|^2 \right] \\
&= \Gamma^N(u, u) \Gamma^N(v, v) \left(\frac{(K-1)^2 + 2(K-1)}{K^3} + \frac{(K-1)^2}{K^4} \right) + (\Gamma^N(u, v))^2 \left(1 - \frac{(K-1)^2}{K^2} + \frac{(K-1)^2 + 2(K-1)}{K^3} \right) \\
&\quad + \frac{(K-1)^2 + (K-2)^2}{K^4} \\
&= \Gamma^N(u, u) \Gamma^N(v, v) \left(\frac{(K-1)^2 + 2(K-1)}{K^3} + \frac{(K-1)^2}{K^4} \right) + (\Gamma^N(u, v))^2 \left(\frac{1-2K}{K^2} + \frac{(K-1)^2 + 2(K-1)}{K^3} \right) \\
&\quad + \frac{(K-1)^2 + (K-2)^2}{K^4}.
\end{aligned}$$

On a ainsi

$$\begin{aligned}
\| \widehat{\Gamma}_K^N - \Gamma^N \|_{V_{\varepsilon}^{**}}^2 &= \sup_{\|u\|_V=1, \|v\|_V=1} \left| \Gamma^N(u, u) \Gamma^N(v, v) \left(\frac{(K-1)^2 + 2(K-1)}{K^3} + \frac{(K-1)^2}{K^4} \right) \right. \\
&\quad \left. + (\Gamma^N(u, v))^2 \left(\frac{1-2K}{K^2} + \frac{(K-1)^2 + 2(K-1)}{K^3} + \frac{(K-1)^2 + (K-2)^2}{K^4} \right) \right| \\
&\leq \left(\sup_{\|u\|_V=1, \|v\|_V=1} |\Gamma^N(u, u) \Gamma^N(v, v)| \right) \left| \frac{(K-1)^2 + 2(K-1)}{K^3} + \frac{(K-1)^2}{K^4} \right| \\
&\quad + \left(\sup_{\|u\|_V=1, \|v\|_V=1} |\Gamma^N(u, v)|^2 \right) \left| \frac{1-2K}{K^2} + \frac{(K-1)^2 + 2(K-1)}{K^3} + \frac{(K-1)^2 + (K-2)^2}{K^4} \right| \\
&\leq \| \Gamma^N \|_{V^{**}} \left(\frac{1-2K}{K^2} + \frac{2(K-1)^2 + 4(K-1)}{K^3} + \frac{2(K-1)^2 + (K-2)^2}{K^4} \right).
\end{aligned}$$

Vu la Remarque 4, on conclut que $\| \widehat{\Gamma}_K^N - \Gamma^N \|_{V_{\varepsilon}^{**}}$ converge vers 0 quand N et K tendent vers $+\infty$, ce qui achève la preuve. \square

Démonstration du Théorème 7. On commence par décomposer :

$$\| \widetilde{\Gamma}_K^{N, \rho} - \Gamma \|_{V_{\varepsilon}^{**}} \leq \| \widehat{\Gamma}_K^N - \Gamma \|_{V_{\varepsilon}^{**}} + \| \widetilde{\Gamma}_K^{N, \rho} - \widehat{\Gamma}_K^N \|_{V_{\varepsilon}^{**}}.$$

Le Lemme 9 donne la convergence vers 0 du premier terme quand N et K tendent vers $+\infty$. Il reste à démontrer que le second terme tend vers 0 quand N et K tendent vers $+\infty$ et ρ tend vers 0.

Pour tous u et v de V , on a :

$$\begin{aligned}
& \mathbb{E} \left[\left| \widetilde{\Gamma}_K^{N,\rho}(u, v) - \widehat{\Gamma}_K^N(u, v) \right|^2 \right] \\
&= \sum_{i \in I_N} \sum_{m \in I_N} \sum_{j \in I_N} \sum_{n \in I_N} u_i u_m v_j v_n \mathbb{E} \left[\left(\widetilde{\gamma}_K^{N,\rho}{}_{i,j} - \widehat{\gamma}_K^N{}_{i,j} \right) \left(\widetilde{\gamma}_K^{N,\rho}{}_{m,n} - \widehat{\gamma}_K^N{}_{m,n} \right) \right] \\
&= \frac{1}{K^2} \sum_{k=1}^K \sum_{l=1}^K \sum_{i \in I_N} \sum_{m \in I_N} \sum_{j \in I_N} \sum_{n \in I_N} u_i u_m v_j v_n \mathbb{E} \left[(B_k(\phi_i) \epsilon_k^\rho(\phi_j) - \frac{1}{K} \sum_{t=1}^K (B_k(\phi_i) \epsilon_t^\rho(\phi_j)) - \frac{1}{K} \sum_{t=1}^K (B_t(\phi_i) \epsilon_k^\rho(\phi_j)) \right. \\
&\quad + \frac{1}{K^2} \sum_{t=1}^K \sum_{s=1}^K (B_t(\phi_i) \epsilon_s^\rho(\phi_j)) + \epsilon_k^\rho(\phi_i) B_k(\phi_j) - \frac{1}{K} \sum_{t=1}^K (\epsilon_k^\rho(\phi_i) B_t(\phi_j)) + \epsilon_k^\rho(\phi_i) \epsilon_k^\rho(\phi_j) - \frac{1}{K} \sum_{t=1}^K (\epsilon_k^\rho(\phi_i) \epsilon_t^\rho(\phi_j)) \\
&\quad - \frac{1}{K} \sum_{t=1}^K (\epsilon_t^\rho(\phi_i) B_k(\phi_j)) + \frac{1}{K^2} \sum_{t=1}^K \sum_{s=1}^K (\epsilon_t^\rho(\phi_i) B_s(\phi_j)) - \frac{1}{K} \sum_{t=1}^K (\epsilon_t^\rho(\phi_i) \epsilon_k^\rho(\phi_j)) + \frac{1}{K^2} \sum_{t=1}^K \sum_{s=1}^K (\epsilon_t^\rho(\phi_i) \epsilon_s^\rho(\phi_j)) \\
&\quad \times (B_l(\phi_m) \epsilon_l^\rho(\phi_n) - \frac{1}{K} \sum_{t=1}^K (B_l(\phi_m) \epsilon_t^\rho(\phi_n)) - \frac{1}{K} \sum_{t=1}^K (B_t(\phi_m) \epsilon_l^\rho(\phi_n)) + \frac{1}{K^2} \sum_{t=1}^K \sum_{s=1}^K (B_t(\phi_m) \epsilon_s^\rho(\phi_n)) \\
&\quad + \epsilon_l^\rho(\phi_m) B_l(\phi_n) - \frac{1}{K} \sum_{t=1}^K (\epsilon_l^\rho(\phi_m) B_t(\phi_n)) + \epsilon_l^\rho(\phi_m) \epsilon_l^\rho(\phi_n) - \frac{1}{K} \sum_{t=1}^K (\epsilon_l^\rho(\phi_m) \epsilon_t^\rho(\phi_n)) - \frac{1}{K} \sum_{t=1}^K (\epsilon_t^\rho(\phi_m) B_l(\phi_n)) \\
&\quad \left. + \frac{1}{K^2} \sum_{t=1}^K \sum_{s=1}^K (\epsilon_t^\rho(\phi_m) B_s(\phi_n)) - \frac{1}{K} \sum_{t=1}^K (\epsilon_t^\rho(\phi_m) \epsilon_l^\rho(\phi_n)) + \frac{1}{K^2} \sum_{t=1}^K \sum_{s=1}^K (\epsilon_t^\rho(\phi_m) \epsilon_s^\rho(\phi_n)) \right].
\end{aligned}$$

On développe le produit puis on utilise la linéarité de l'espérance pour séparer les termes et recomposer les vecteurs u et v à l'intérieur de chaque espérance. On passe ensuite à la borne supérieure en valeur absolue sur u et v dans la sphère-unité de V . On obtiendrait ainsi une situation assez analogue à celle de la démonstration précédente, mais avec 144 termes. Il est simple de montrer que ces termes convergent tous vers 0 quand ρ tend vers 0 et quand N et K tendent vers $+\infty$, en utilisant le Lemme 6, le Théorème 4 ainsi que l'Hypothèse 3. \square

Chapitre 6

Expériences de validation

6.1 Objectifs

Le but de ce chapitre est de tester, sur des données simulées, la convergence numérique de l'estimateur du représentant moyen d'un échantillon de surfaces défini en Section 4.3.2. Dans cette optique, on appliquera cet estimateur à des échantillons de surfaces discrètes dont le représentant moyen théorique est connu, pour diverses valeurs des paramètres K (taille de l'échantillon) et ρ (paramètre de précision de la discrétisation). On effectuera tous les calculs avec une valeur du paramètre N (dimension du sous-espace d'approximation) fixée de manière optimale (de façon à réduire considérablement la dimension de l'espace d'étude tout en limitant l'erreur due à la projection sur le sous-espace). Pour chaque valeur du couple (K, ρ) considérée, on calculera l'erreur d'estimation et on étudiera la convergence de cette erreur lorsque ρ tend vers 0 et K tend vers l'infini. L'intérêt principal de cette expérience réside dans le fait qu'elle constitue une évaluation de notre cadre statistique : on cherche à montrer que ce cadre, consistant sur le plan théorique vu le Théorème 4, est efficace d'un point de vue numérique.

Dans un premier temps, on met en place des algorithmes permettant de simuler des échantillons de surfaces présentant des variations aléatoires, dont le représentant moyen théorique est connu. Ensuite, on construit un exemple concret d'espace des représentants ainsi que les sous-espaces d'approximation associés. Dans un troisième temps, on met en place une expérience pour régler de manière optimale le paramètre de dimension de l'espace d'approximation. Enfin, on effectue l'étude proprement dite de la convergence de l'estimateur du représentant moyen.

6.2 Simulation d'échantillons

On donne à présent deux méthodes permettant de générer aléatoirement des échantillons de surfaces dont le représentant moyen est connu. Ce sont de tels échantillons qui sont utilisés dans les expériences d'évaluation de l'estimateur du représentant moyen rapportées plus loin en Section 6.5.

6.2.1 Surfaces à creux et bosses

Soit O l'origine de l'espace \mathbb{R}^3 et soit \mathcal{D} le cube $[-\pi, \pi]^3$. Soit \mathcal{S} la demi-sphère supérieure de centre O et de rayon 2; notons S son représentant dans l'espace des formes linéaires. Soient $K \in \mathbb{N}^*$ et $\rho > 0$. On donne un algorithme permettant de générer un échantillon $(F_1^\rho, \dots, F_K^\rho)$ de K surfaces discrétisées avec précision ρ dont le représentant moyen est S .

Pour k variant de 1 à K :

- On génère tout d'abord F_K^ρ comme une approximation de \mathcal{S} , discrétisée avec précision ρ .
- On génère un nombre entier aléatoire M_k de loi $\mathcal{U}(\{3, \dots, 9\})$.
- On détermine, en les sélectionnant aléatoirement parmi les sommets de F_k^ρ , M_k points qui constitueront les "centres de déformation". On les note $A_1^k, \dots, A_{M_k}^k$.
- On génère une réalisation $(\alpha_1^k, \dots, \alpha_{M_k}^k)$ d'une suite de M_k variables aléatoires i.i.d. de loi $\mathcal{U}([-0.2, 0.2])$.
- Pour tout $x \in \{1, \dots, M_k\}$, on définit le point B_x^k par

$$\overrightarrow{OB_x^k} = \overrightarrow{OA_x^k} + \frac{\overrightarrow{OA_x^k}}{\|\overrightarrow{OA_x^k}\|_{\mathbb{R}^3}} \times \alpha_x^k.$$

- Pour tout $x \in \{1, \dots, M_k\}$, on définit \mathcal{S}_x^k comme la demi-sphère dont le centre est A_x^k et dont le rayon orthogonal au plan diamétral est $[A_x^k, B_x^k]$.
- Pour x variant de 1 à M_k , on remplace les points de F_K^ρ situés dans la boule de centre A_x^k et de rayon $|\alpha_x^k|$ par leur projection sur \mathcal{S}_x^k suivant un axe orthogonal au plan diamétral de \mathcal{S}_x^k .

Cet algorithme permet d'obtenir des surfaces correspondant à des versions discrètes de demi-sphères bruitées autour de centres localisés aléatoirement; les figures ci-après en constituent des exemples. Des échantillons ainsi constitués présentent l'avantage d'être aisés à simuler; en revanche, ils manqueront inévitablement de variabilité. Les figures ci-dessous constituent deux exemples de surfaces construites selon ce procédé.

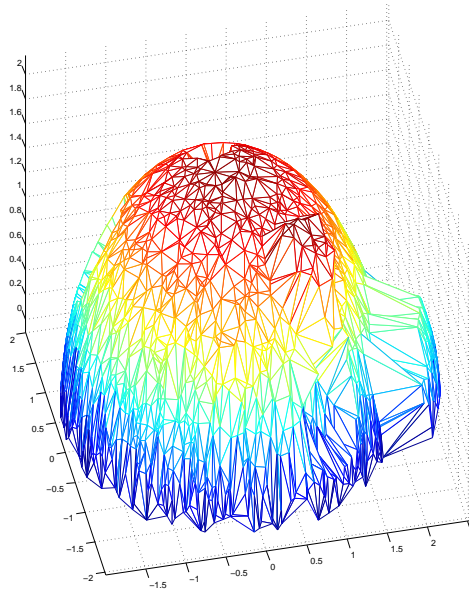


FIGURE 6.1

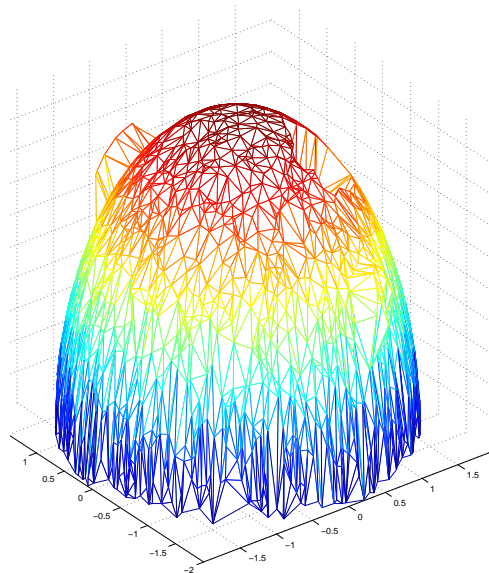


FIGURE 6.2

6.2.2 Surfaces "Browniennes"

Avec les mêmes notations qu'en Section 6.2.1, on donne à présent une méthode permettant de simuler des échantillons de surfaces bruitées, toujours de représentant moyen S , présentant une plus grande variabilité. On opte à cet effet pour des versions discrètes de la sphère S altérées par un bruit Gaussien. On

génère un tel échantillon via l'algorithme suivant.

Pour k variant de 1 à K :

- On construit tout d'abord F_K^ρ comme une approximation de \mathcal{S} , discrétisée avec précision ρ .
- On altère F_K^ρ en utilisant la formule donnée dans [14] pour la sphère bruitée par mouvement brownien tridimensionnel, avec quelques changements afin de rendre nos surfaces simulées plus ressemblantes à des données expérimentales. On choisit dans cette optique de travailler avec un indice de Hurst fixé à 0.9 afin de ne pas obtenir une surface trop irrégulière; par ailleurs, on remplace les variables aléatoires Gaussiennes standard par des variables de loi $\mathcal{N}(0, 0.35)$, afin que les surfaces générées aient un minimum de similarité en termes de forme.

Les surfaces construites via cet algorithme auront des variations nettement plus équiréparties que celles de la section ci-dessus. Les figures ci-après sont des exemples de telles surfaces.

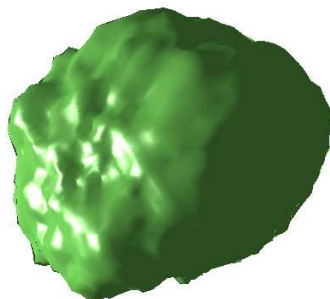


FIGURE 6.3

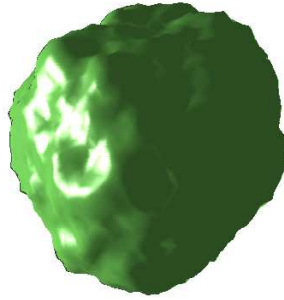


FIGURE 6.4

6.3 Espace des représentants

6.3.1 Espace de champs de vecteurs

Dans cette section, on définit l'espace des représentants qui sera utilisé dans les expériences qui suivent. Pour ce faire, on commence par construire un espace de champs de vecteurs V ayant toutes les propriétés requises en Section 3.2.1.

On considère l'espace de Sobolev V_0 défini en Section 2.2.2, qui est un R.K.H.S. à noyau borné sur \mathcal{D} , et on garde toutes les notations associées. On note $V = (V_0)^3$. On définit sur l'espace V le produit scalaire $\langle \cdot, \cdot \rangle_V$ par

$$\forall f, g \in V, \langle f, g \rangle_V = \sum_{i=1}^3 \langle f_i, g_i \rangle_{V_0},$$

où $f = (f_1, f_2, f_3)$ et $g = (g_1, g_2, g_3)$ avec $f_1, f_2, f_3, g_1, g_2, g_3 \in V_0$.

Soit $\|\cdot\|_V$ la norme sur V induite par le produit scalaire $\langle \cdot, \cdot \rangle_V$. L'espace $(V, \|\cdot\|_V)$ est un espace de Hilbert, avec toutes les propriétés requises pour l'espace du même nom en Section 3.2.1. On représentera les surfaces, selon le modèle défini au Chapitre 3, par des formes linéaires de l'espace V^* , muni de la topologie usuelle d'espace dual.

6.3.2 Espaces d'approximation

On précise maintenant le procédé de construction, dans le cas de l'espace V défini ci-dessus, des espaces d'approximation définis en Section 4.2. Toujours avec les notations de la Section 2.2.2, on pose

$$\forall d \in \{1, 2, 3\}, \forall n \in \mathbb{N}, \phi_n^d = (\delta_{1,d}, \delta_{2,d}, \delta_{3,d})\phi_n,$$

$$\psi_n^d = (\delta_{1,d}, \delta_{2,d}, \delta_{3,d})\psi_n.$$

Le système Φ défini par

$$\Phi = \{\phi_n^d \mid (n, d) \in \mathbb{N}^3 \times \{1, 2, 3\}\} \cup \{\psi_n^d \mid (n, d) \in \mathbb{N}^3 \times \{1, 2, 3\}\}$$

constitue un système libre et total de l'espace V . Pour tout entier naturel non nul N , on note

$$\Phi^N = \{\phi_n^d \mid (n, d) \in \{0, 1, \dots, N\}^3 \times \{1, 2, 3\}\} \cup \{\psi_n^d \mid (n, d) \in \{0, 1, \dots, N\}^3 \times \{1, 2, 3\}\},$$

et on définit les sous-espaces vectoriels de V

$$V_N = \text{Vect}(\Phi^N), N \in \mathbb{N}^*,$$

sur lesquels on met la topologie et le produit scalaire induits par ceux de V . Ces sous-espaces de dimension finie sont, comme requis en Section 4.2, emboîtés et de réunion dense dans V . C'est cette suite $(V_N)_{N \in \mathbb{N}^*}$ qui induit les sous-espaces d'approximation (\widetilde{V}_N^*) , $N \in \mathbb{N}^*$ comme cela a été explicité en Section 4.2.

6.4 Recherche de la dimension optimale de l'espace d'approximation

6.4.1 Protocole

On cherche maintenant à optimiser la valeur du paramètre N donnant la dimension du sous-espace d'approximation sur lesquels les représentants des surfaces seront projetés. On veut fixer cette valeur de manière à, d'une part, réduire au maximum le nombre de composantes à calculer lors de l'estimation et, d'autre part, limiter l'erreur d'approximation commise en projetant les représentants sur \widetilde{V}_N^* . On cherche à construire l'espace d'approximation (et par conséquent la base de test) avec une vérité de terrain : représenter les surfaces de façon approximée, avec une erreur d'estimation liée à l'approximation inférieure à 1% en termes de norme dans V^* . En reprenant la démonstration du Théorème 4, cette erreur s'écrit $\|\mu - \mu^N\|_{V^*}$,

où μ est le représentant moyen théorique de l'échantillon. Il s'agit donc de déterminer N tel que

$$\frac{\|\mu - \mu^N\|_{V^*}}{\|\mu\|_{V^*}} \leq \frac{1}{100}. \quad (6.1)$$

Remarquons que, cependant, le rapport intervenant dans l'inégalité (6.1) est impossible à calculer en pratique car l'espace V est de dimension infinie. On peut pallier ce problème en prenant un entier N_0 induisant un espace V_{N_0} de très grande dimension et en assimilant ce dernier à l'espace V . Typiquement, on choisit pour les expériences menées tout au long de ce chapitre $N_0 = 100$, soit un espace V_{N_0} de dimension 3×10^6 . Ce choix est essentiellement motivé par les performances de la machine à disposition pour mener ces tests. Dans les faits, on calculera ainsi pour chaque valeur de N , en lieu et place du quotient (6.1), le rapport

$$\frac{\|\mu^{N_0} - \mu^N\|_{V^*}}{\|\mu^{N_0}\|_{V^*}}.$$

Comme tous les échantillons de surfaces simulés et étudiés dans ce chapitre sont construits en appliquant un bruit centré aux sommets d'une sphère ou d'une demi-sphère discrètes, on choisit μ comme le représentant dans l'espace V^* d'une sphère discrète \mathcal{S} , triangulée avec précision $\rho = 0.01$.

On effectue donc l'algorithme suivant :

- on simule la sphère \mathcal{S} ;
- pour tout $N \in \{1, \dots, N_0\}$, on calcule μ^N via les formules données en Annexe A.3 ;
- on calcule ensuite les rapports $\frac{\|\mu^{N_0} - \mu^N\|_{V^*}}{\|\mu^{N_0}\|_{V^*}}$, $N \in \{1, \dots, N_0\}$;
- On choisit comme valeur optimale de N la plus petite valeur telle que le rapport $\frac{\|\mu^{N_0} - \mu^N\|_{V^*}}{\|\mu^{N_0}\|_{V^*}}$ soit inférieur à 1%.

6.4.2 Résultats

On a appliqué l'algorithme décrit en Section 6.4.1. Les valeurs pour le rapport considéré dans l'inéquation (6.1) sont en partie récapitulées dans le tableau ci-dessous.

N	2	3	4	6	7	8	9	10	11	12
$\frac{\ \mu^{N_0} - \mu^N\ _{V^*}}{\ \mu^{N_0}\ _{V^*}}$	0.270	0.114	0.058	0.027	0.019	0.015	0.013	0.011	0.009	0.008

TABLE 6.1 – Erreur relative d'approximation.

La plus petite valeur de N telle que le rapport $\frac{\|\mu^{N_0} - \mu^N\|_{V^*}}{\|\mu^{N_0}\|_{V^*}}$ soit plus petit que 0.01 est $N = 11$. Ainsi, lors

du test de l'estimateur du représentant moyen sur des échantillons générés suivant l'algorithme de la Section 6.2.1, on utilisera le sous-espace \widetilde{V}_{11}^* comme sous-espace d'approximation. Pour des raisons de faisabilité des calculs numériques, on assimilera toujours V_{100} à V . Le procédé d'approximation par projection sur le sous-espace d'approximation permet donc de ne calculer qu'une proportion de $(\frac{11}{100})^3$ soit à peu près 1.3% des paramètres, tout en conservant une précision de 99%, en termes de norme dans l'espace des représentants.

6.5 Convergence numérique de l'estimateur du représentant moyen

6.5.1 Protocole

On réalise maintenant l'étude proprement dite de la convergence numérique de l'estimateur du représentant moyen d'un échantillon de surfaces discrètes. On va, dans cette optique, pour différentes valeurs des paramètres K (taille de l'échantillon) et ρ (aire maximale des triangles composant la discrétisation), évaluer de façon approchée, l'erreur moyenne d'estimation en termes de norme dans l'espace des formes linéaires aléatoires.

L'espace V étant toujours celui défini en Section 6.3 et la valeur de N étant choisie de façon appropriée via le test de la Section 6.4, soient $\mathcal{K} = \{10, 20, 50, 200\}$ et $\mathcal{R} = \{1, 0.2, 0.1, 0.05\}$. Pour chaque couple $(K, \rho) \in \mathcal{K} \times \mathcal{R}$, on génère, via l'un des deux algorithmes présentés en Section 6.2, 10 échantillons de K surfaces discrétisées avec précision ρ . On note indifféremment dans chaque cas $(F_{1,E}^\rho, \dots, F_{K,E}^\rho)_{1 \leq E \leq 10}$ ces échantillons de surfaces et \mathcal{S} la surface-modèle théorique associée à l'algorithme choisi, demi-sphère dans le premier cas, sphère dans le second. Soit N la valeur du paramètre de dimension de l'espace d'approximation choisie comme optimal en Section 6.4 et soit $N_0 = 100$. On assimile toujours le sous-espace V_{N_0} à l'espace V .

On commence par évaluer le représentant de Riesz v_μ de la forme linéaire μ associée à \mathcal{S} (dans les faits, on calcule son approximation $v_\mu^{N_0}$ par projection sur le sous-espace V_{N_0}). Ensuite, pour tout K de \mathcal{K} , pour tout ρ de \mathcal{R} , pour tout E de $\{1, \dots, 10\}$, on calcule, au moyen des formules explicitées en Annexe A, les représentants de Riesz approximatifs respectifs $v_{1,\rho}^{N,E}, \dots, v_{K,\rho}^{N,E}$ des formes linéaires $v_{1,\rho}^{N,E}, \dots, v_{K,\rho}^{N,E}$ associées à l'échantillon $(F_{1,E}^\rho, \dots, F_{K,E}^\rho)_{1 \leq E \leq 10}$. Enfin, pour E variant de 1 à 10, on calcule via la formule (4.7) la valeur estimée $\widetilde{v_{k,\rho}^{N,E}}$ du représentant de Riesz de la forme linéaire moyenne associée à l'échantillon $(F_{1,E}^\rho, \dots, F_{K,E}^\rho)_{1 \leq E \leq 10}$.

On conclut en évaluant l'erreur relative d'estimation en termes de norme dans l'espace $V_{\mathcal{E}}^*$ via la formule

$$\frac{\frac{1}{10} \sum_{E=1}^{10} \|\widetilde{v_{k,\rho}^{N,E}} - v_\mu^{N_0}\|_{V_{N_0}}}{\|v_\mu^{N_0}\|_{V_{N_0}}}. \quad (6.2)$$

6.5.2 Résultats

Résultats pour le premier échantillon

On applique le protocole décrit en Section 6.5.1 à des échantillons de surfaces générés via l'algorithme de la Section 6.2.1. Le tableau ci-dessous donne l'erreur relative d'estimation calculée via la formule (6.2), pour les différentes valeurs de la taille d'échantillon $K \in \mathcal{K}$ et de la précision $\rho \in \mathcal{R}$.

$K \backslash \rho$	1	0.5	0.1	0.05
10	0.094	0.079	0.019	0.022
20	0.083	0.094	0.029	0.022
50	0.118	0.078	0.018	0.014
200	0.096	0.055	0.020	0.014

TABLE 6.2 – Erreur relative en norme (premier échantillon).

Les résultats confirment que l'erreur relative d'estimation, en termes de norme dans l'espace des représentants, converge effectivement vers 0 quand le paramètre de précision ρ tend vers 0 et quand la taille d'échantillon K tend vers $+\infty$. Cela va dans le sens du résultat de consistance démontré en Section 4.4 et semble témoigner d'une convergence assez rapide d'un point de vue numérique. Notons toutefois que c'est essentiellement ρ qui semble influencer le résultat. On remarque que l'erreur est toujours supérieure à 0.01. Cet état de fait est somme toute, logique. On a, en effet, vu en Section 6.4.2 que, pour la valeur de N choisie, l'erreur due à la projection sur le sous-espace d'approximation V_N est à peu près de 0.01 ; cette valeur constitue ainsi l'erreur minimale pour l'estimateur de la forme linéaire moyenne.

Résultats pour le second échantillon

En appliquant le protocole de la Section 6.5.1 à des échantillons de surfaces générés via l'algorithme donné en Section 6.2.2, on a obtenu les résultats récapitulés dans le tableau ci-dessous.

$K \backslash \rho$	1	0.5	0.1	0.05
10	34.151	3.675	0.768	0.191
20	19.308	2.183	0.417	0.110
50	13.561	2.001	0.434	0.057
200	7.489	1.560	0.298	0.025

TABLE 6.3 – Erreur relative en norme (second échantillon).

Comme pour l'échantillon précédent, les résultats tendent à confirmer la convergence démontrée en Section 4.4. Dans le cas présent, la convergence est plus lente ; on observe en outre que, pour K petit ou ρ grand, l'estimation est très mauvaise. Ceci s'explique entre autres par le fait, déjà mentionné en Section 6.2, que les échantillons construits avec la seconde méthode présentent une plus grande variabilité d'un point de

vue statistique. L'erreur semble dépendre à la fois de ρ et de K de manière significative. On observe notamment via les deux premières colonnes du tableau que, pour de grandes valeurs de ρ , la taille d'échantillon compense quelque peu la faiblesse de la précision. Pour $K = 200$ et $\rho = 0.05$, on retrouve une estimation presque aussi précise que dans le cas précédent. On pourrait conjecturer que, en améliorant encore les valeurs de K et de ρ , on obtiendrait une erreur d'estimation encore plus basse, avec une valeur limite valant toujours 0.01 pour les raisons évoquées précédemment.

À l'instar de ceux donnés plus haut, ces résultats tendent à confirmer que la convergence de l'erreur commise en estimant le représentant moyen est effective d'un point de vue numérique.

Chapitre 7

Discussion

7.1 Bilan

Dans cette partie, on a développé un cadre adapté à l'analyse statistique de surfaces représentées par des formes linéaires continues sur un espace de champs de vecteurs de type produit cartésien d'espaces de Hilbert à noyau reproduisant. Ce cadre consiste en une extension probabiliste, basée sur des processus stochastiques généralisés, de la représentation par courants introduite initialement par Glaunès et Vaillant. Les surfaces étant ainsi représentés par des éléments d'un espace de Hilbert de dimension infinie, on a mis en oeuvre une procédure d'approximation par projection orthogonale sur un sous-espace vectoriel de dimension finie fixée ; tous les calculs se font dans cet espace d'approximation.

Considérant les représentants des surfaces comme des réalisations de formes linéaires aléatoires, on a donné un modèle décrivant la variabilité des représentants d'un échantillon de surfaces continues autour d'un représentant moyen, vu comme le représentant d'une hypothétique "surface moyenne" de l'échantillon considéré. Ce modèle est ensuite étendu en un modèle d'observation exprimé en termes de représentants d'approximations discrètes (sous formes de surfaces triangulées) des surfaces de l'échantillon considéré ; chaque représentant diffère alors du représentant moyen théorique par un terme de bruit gaussien et un terme correspondant à l'erreur de discrétisation.

On a construit, tout d'abord dans le modèle de variation puis, en passant au cas discret, dans le modèle d'observation, des estimateurs pour le représentant moyen d'un échantillon de surfaces ainsi que pour l'opérateur d'autocovariance du bruit associé. On a énoncé et démontré la consistance de ces estimateurs, au sens de la convergence uniforme des formes linéaires et bilinéaires.

On a effectué des expériences de validation consistant à appliquer notre procédure d'estimation à des échantillons simulés dont le représentant moyen est connu. Ces tests viennent compléter les résultats théoriques, en confirmant que la consistance des estimateurs est effective d'un point de vue numérique.

7.2 Perspectives

7.2.1 Perspectives immédiates

On dispose à présent d'un cadre statistique adapté à l'analyse de surfaces. La suite de ce manuscrit est consacrée à la mise en application de ce dernier. Dans la seconde partie, on s'intéressera à la résolution de problèmes de classification de surfaces, via des méthodes de classification Bayésienne.

7.2.2 Perspectives ultérieures

Dans les travaux à venir, on adaptera, d'une part, notre modèle statistique à des situations plus générales. Par exemple, on pourra reprendre le modèle d'observation donné par l'Équation (3.4) en faisant des hypothèses moins fortes quant à la loi du bruit. On pourra aussi étendre ce modèle pour prendre en compte les variations géométriques des surfaces.

$$S_{k,\rho} = \psi_k \cdot \mu + B_k + \epsilon_{k,\rho}, \forall k \in \llbracket 1, K \rrbracket, \quad (7.1)$$

où $\psi_k \cdot \mu$ représente l'action d'une déformation ψ_k sur le représentant moyen μ . Ce représentant moyen correspond alors à un prototype déformable. De tels modèles probabilistes existent pour les images et les objets euclidiens, mais ne sont pas encore bien établis pour des objets tels que les surfaces ou leurs représentations par des courants. Dans ce nouveau cadre, on projette d'étendre notre méthode d'estimation au calcul du prototype. La littérature est très vaste dans ce domaine ; de fait, le travail à accomplir consiste essentiellement en l'adaptation spécifique à notre cadre de représentation des surfaces de méthodes existantes, telles que celles répertoriées par Allasonnière et alii dans [2]. La littérature propose de nombreuses méthodes pour la résolution de tels problèmes [1, 19, 24, 31]. Le gros du travail à accomplir est de rendre ces modèles compatibles avec notre cadre statistique reposant sur un espace de formes linéaires aléatoires.

On s'intéressera, d'autre part, à la question de la récupération de la surface moyenne d'un échantillon à partir du représentant moyen tel que défini en Section 4.3.2. Pour ce faire, il sera nécessaire de résoudre le problème de l'inversion de la fonction qui à une surface compacte associe son représentant dans l'espace des formes linéaires. Ce problème étant a priori non linéaire et mal posé, intuition que nous confirmons plus loin, en deuxième partie au Chapitre 9, la reconstruction d'une surface moyenne s'avèrera une tâche particulièrement difficile nécessitant l'élaboration de procédures nouvelles et spécifiques. On explore en Annexe B quelques pistes pour définir rigoureusement une notion de surface moyenne d'un échantillon.

Deuxième partie

Application à la classification de surfaces

Chapitre 8

Introduction

8.1 Motivations et difficultés

La classification de surfaces constitue une thématique majeure, aux nombreuses applications à l'imagerie 3-D, notamment dans les domaines de la biologie et de la médecine. Classer les organes en groupes associés à différents phénotypes, et ce en fonction de critères géométriques, tels que forme, aire, etc., constitue un puissant outil de diagnostic.

Dans cette partie, on adapte une procédure de classification existante au cadre statistique que l'on a développé en première partie pour construire un algorithme de classification supervisée adapté aux surfaces. On prend le parti d'utiliser une méthode de classification Bayésienne, telle que décrite dans [8], de manière adaptée aux surfaces. Les procédures usuelles de classification Bayésienne nécessitant de disposer d'une distance, un travail préliminaire essentiel est de mettre en place une structure d'espace métrique sur l'ensemble des surfaces.

8.2 Objectifs

8.2.1 Contributions

Dans cette partie, on accomplit deux apports principaux : l'adaptation d'un algorithme de classification Bayésienne en une forme spécifique aux surfaces et la construction d'une distance sur les surfaces.

Distance sur les surfaces

Dans le cadre de représentation développé au Chapitre 3, un espace de surfaces continûment différentiables Θ est mis en correspondance avec un espace de Hilbert de formes linéaires V^* . Dans la suite, on

note Ψ l'application qui à toute surface de Θ associe son représentant dans V^* . La distance euclidienne sur V^* induit sur Θ l'application :

$$\begin{aligned} \mathcal{D}_\Theta : \Theta \times \Theta &\longrightarrow \mathbb{R} \\ (\mathcal{S}, \mathcal{T}) &\longrightarrow \|\Psi(\mathcal{S}) - \Psi(\mathcal{T})\|_{V^*}. \end{aligned} \tag{8.1}$$

Cette fonction est-elle une distance sur Θ ? D'une part, comme $\|\cdot\|_{V^*}$ est une norme sur V^* , \mathcal{D}_Θ est positive et symétrique et vérifie l'inégalité triangulaire. Il reste à déterminer si

$$\mathcal{D}_\Theta(\mathcal{S}, \mathcal{T}) = 0 \implies \mathcal{S} = \mathcal{T}.$$

Vu que $\|\cdot\|_{V^*}$ est une norme sur V^* , cela est vrai pour peu que

$$\Psi(\mathcal{S}) = \Psi(\mathcal{T}) \implies \mathcal{S} = \mathcal{T},$$

c'est à dire lorsque l'application Ψ est injective. Au vu de l'étude menée en première partie, cette injectivité n'est a priori pas triviale ; par conséquent, une étude précise de la fonction Ψ va s'avérer nécessaire.

Classification Bayésienne sur les surfaces

La classification Bayésienne, telle que définie usuellement dans la littérature, permet de classer des éléments d'un espace métrique parmi des catégories prédéfinies. Une des innovations effectuées dans cette partie est la construction d'une méthode de classification Bayésienne spécifique aux surfaces représentées par courants. C'est la définition d'une structure d'espace métrique sur les surfaces qui rend possible une telle adaptation. On se fixe l'objectif de mettre en place une procédure de classification à deux catégories, qui va permettre de discriminer des populations de surfaces, typiquement des organes ou des sections d'organes, selon un critère binaire, par exemple mâle ou femelle, moitié droite ou moitié gauche, sain ou malade.

8.2.2 Contenus de cette partie

Dans le premier chapitre de cette partie, on met en place une structure d'espace métrique sur l'ensemble Θ des surfaces lisses. Tout d'abord, on établit que, sous certaines conditions, l'application \mathcal{D}_Θ est une distance sur Θ en démontrant l'injectivité de Ψ . Ensuite, on étudie sommairement les propriétés de l'espace métrique $(\Theta, \mathcal{D}_\Theta)$. Enfin, on traite le problème de l'extension de la distance aux surfaces discrètes.

Dans le deuxième chapitre, on décrit un algorithme de classification Bayésienne adapté aux surfaces.

Le troisième chapitre est consacré aux expériences. on y teste l'algorithme de classification, sur un échantillon constitué de deux populations de surfaces issues de l'imagerie cérébrale. Une analyse du taux d'erreur

de généralisation permettra d'évaluer l'efficacité de la méthode.

Chapitre 9

Structure d'espace métrique sur l'ensemble des surfaces

9.1 Distance sur les surfaces

9.1.1 Choix des ensembles

Ainsi que mentionné plus haut, l'application \mathcal{D}_Θ définie via l'Équation (8.1), est une distance sur l'ensemble des surfaces si et seulement si l'application Ψ est injective. Cette propriété dépendant clairement du choix des ensembles Θ et V , il est nécessaire de définir ces derniers de façon plus précise qu'au Chapitre 3, où les hypothèses les concernant sont, somme toute, très générales. Dans la suite de la deuxième partie, on considère que les ensembles Θ et V^* sont définis comme ci-dessous.

Ensemble de surfaces

On construit Θ comme un ensemble de surfaces orientées. Soit \mathcal{D} le cube $[-\pi, \pi]^3$. Les éléments de Θ autres que l'ensemble vide (que l'on rajoute à l'ensemble Θ pour des raisons pratiques) sont les couples $(\mathcal{S}, N_{\mathcal{S}})$, où :

- \mathcal{S} est une surface (non-orientée) incluse dans \mathcal{D} , C^1 , de mesure non nulle et de barycentre $O = (0, 0, 0)$. ;
- $N_{\mathcal{S}}$ est une application C^1 définie sur \mathcal{S} telle que, pour tout point s de \mathcal{S} , le vecteur $N_{\mathcal{S}}(s)$ soit normal à \mathcal{S} et de norme 1.

Remarque 5. Pour une surface non-orientée \mathcal{S} donnée, il y a deux champs normaux $N_{\mathcal{S}}$ possibles. Ces champs sont

mutuellement opposés. Cette propriété est assurée par le fait que N_S doit être de classe C^1 .

Remarque 6 (Égalité entre surfaces orientées). On considère deux surfaces orientées, non vides, (S, N_S) et (T, N_T) comme égales lorsque les surfaces S et T sont égales. On notera l'égalité entre éléments de Θ avec le symbole \equiv .

Espace de champs de vecteurs

Pour avoir de bonnes propriétés (voir ci-dessous en Section 9.1.2), on choisit V_0 comme un R.K.H.S. à noyau borné dont les éléments sont des fonctions définies sur \mathcal{D} et qui contiennent l'espace de Schwarz sur \mathcal{D} , noté $S(\mathcal{D})$. Les espaces de Sobolev considérés en Section 2.2.2, dont un exemple est donné en Section 6.3, vérifient cette hypothèse.

9.1.2 Injectivité de la fonction Ψ

À présent, on étudie, sous les hypothèses sur Θ et V faites ci-dessus, l'injectivité de la fonction Ψ .

Soient $\mathbb{S} = (S, N_S)$ et $\mathbb{T} = (T, N_T)$ deux surfaces orientées de l'ensemble Θ , supposées distinctes, telles que $\Psi(\mathbb{S}) = \Psi(\mathbb{T})$. Il existe alors une forme linéaire s de V^* qui représente à la fois les surfaces orientées S et T . Par la définition du représentant d'une surface, donnée en Section 3.2.1 par l'Équation (3.1), s s'écrit à la fois :

$$s(w) = \int_S \langle w(s), N_S(s) \rangle_{\mathbb{R}^3} d\sigma_S(s), \forall w \in V, \quad (9.1)$$

et

$$s(w) = \int_T \langle w(s), N_T(s) \rangle_{\mathbb{R}^3} d\sigma_T(s), \forall w \in V. \quad (9.2)$$

On va chercher si ces deux définitions de la forme linéaire s sont compatibles ou contradictoires.

On commence par écarter le cas trivial dans lequel une des deux surfaces considérées serait de mesure nulle (*i.e.* vide vu la définition de l'ensemble Θ). Si S est vide, alors, vu l'Équation (9.1), s est la forme linéaire nulle. Comme \mathbb{S} et \mathbb{T} sont distinctes, alors T est de mesure non nulle ; par conséquent, en posant $w|_T = N_T$ dans l'Équation (9.2), on obtient $s(w) = \mathcal{A}(T) \neq 0$ et s n'est pas la forme linéaire nulle. On obtient donc une contradiction, d'où le fait que S n'est pas vide. Par symétrie, T n'est pas vide. Les surfaces S et T sont donc toutes deux de mesure non nulle.

Étant donné que $\mathbb{S} \neq \mathbb{T}$, on a nécessairement l'un des deux cas suivants :

- . $\mathcal{S} = \mathcal{T}$ et $N_{\mathcal{S}} = -N_{\mathcal{T}}$,
- . $\mathcal{S} \neq \mathcal{T}$.

Dans le premier cas, on aurait $\Psi(\mathcal{S}) = -\Psi(\mathcal{T})$, ce qui constitue une contradiction avec l'hypothèse initiale. Par conséquent, le second cas constitue la seule possibilité.

Vu que $\mathcal{S} \neq \mathcal{T}$, on a forcément $\mathcal{S} \not\subset \mathcal{T}$ ou $\mathcal{T} \not\subset \mathcal{S}$. On peut supposer, quitte à permuter, que $\mathcal{T} \not\subset \mathcal{S}$. Dans ce cas, on a alors $\mathcal{T} \setminus (\mathcal{S} \cap \mathcal{T}) \neq \emptyset$. Comme on travaille sur un ensemble de surfaces continues, l'ensemble $\mathcal{T} \setminus (\mathcal{S} \cap \mathcal{T})$ est de mesure non nulle. On considère maintenant un champ de vecteurs w_1 de V qui vérifie :

$$\begin{aligned} \forall t \in \mathcal{S}, w_1(t) &= (0, 0, 0) \\ \text{et } \forall t \in \mathcal{T} \setminus (\mathcal{S} \cap \mathcal{T}), w_1(t) &= x_1(t) \cdot N_{\mathcal{T}}(t), \end{aligned}$$

où x_1 est une fonction C^1 et bornée sur \mathcal{D} , à valeurs dans \mathbb{R}^+ , non nulle sur au moins une partie de $\mathcal{T} \setminus (\mathcal{S} \cap \mathcal{T})$ de mesure non nulle. Une telle fonction existe dans V , car V_0 contient l'espace de Schwartz sur \mathcal{D} .

On calcule maintenant l'image de w_1 par la forme linéaire s de deux manières différentes. L'Équation (9.1) donne

$$\begin{aligned} s(w_1) &= \int_{\mathcal{S}} \langle w_1(s), N_{\mathcal{S}}(s) \rangle_{\mathbb{R}^3} d\sigma_{\mathcal{S}}(s) \\ &= \int_{\mathcal{S}} \langle (0, 0, 0), N_{\mathcal{S}}(s) \rangle_{\mathbb{R}^3} d\sigma_{\mathcal{S}}(s) \\ &= 0. \end{aligned}$$

D'autre part, l'Équation (9.2) donne :

$$\begin{aligned} s(w_1) &= \int_{\mathcal{T}} \langle w_1(s), N_{\mathcal{T}}(s) \rangle_{\mathbb{R}^3} d\sigma_{\mathcal{T}}(s) \\ &= \int_{\mathcal{T} \setminus (\mathcal{S} \cap \mathcal{T})} \langle w_1(s), N_{\mathcal{T}}(s) \rangle_{\mathbb{R}^3} d\sigma_{\mathcal{T}}(s) + \int_{\mathcal{S} \cap \mathcal{T}} \langle w_1(s), N_{\mathcal{T}}(s) \rangle_{\mathbb{R}^3} d\sigma_{\mathcal{T}}(s) \\ &= \int_{\mathcal{T} \setminus (\mathcal{S} \cap \mathcal{T})} x_1(s) \|N_{\mathcal{T}}\|_{\mathbb{R}^3} d\mathcal{T}(s) + \int_{\mathcal{S} \cap \mathcal{T}} \langle (0, 0, 0), N_{\mathcal{T}}(s) \rangle_{\mathbb{R}^3} d\sigma_{\mathcal{T}}(s) \\ &= \int_{\mathcal{T} \setminus (\mathcal{S} \cap \mathcal{T})} x_1(s) d\mathcal{T}(s). \end{aligned}$$

Comme la fonction x_1 est à valeurs positives et non nulle sur une partie de $\mathcal{T} \setminus (\mathcal{S} \cap \mathcal{T})$ de mesure non nulle, la dernière intégrale est nécessairement non nulle, ce qui vient contredire le résultat du calcul qui précède. On ne peut lever cette contradiction qu'en niant l'hypothèse $\mathcal{S} \neq \mathcal{T}$. On a ainsi montré que $\mathcal{S} = \mathcal{T}$. Il ne reste qu'à étudier l'orientation des normales $N_{\mathcal{S}}$ et $N_{\mathcal{T}}$.

D'après la Remarque 5, il n'y a que deux possibilités : $N_{\mathcal{S}} = N_{\mathcal{T}}$ ou $N_{\mathcal{S}} = -N_{\mathcal{T}}$. Si $N_{\mathcal{S}} = -N_{\mathcal{T}}$, alors on se donne une fonction $w_2 \in V$ telle que $w_2(t) = N_{\mathcal{S}}(t), \forall t \in N_{\mathcal{S}}$.

On a alors, par l'Équation (9.1),

$$\begin{aligned}
s(w_2) &= \int_{\mathcal{S}} \langle w_2(s), N_{\mathcal{S}}(s) \rangle_{\mathbb{R}^3} d\sigma_{\mathcal{S}}(s) \\
&= \int_{\mathcal{S}} \|N_{\mathcal{S}}(s)\|_{\mathbb{R}^3}^2 d\sigma_{\mathcal{S}}(s) \\
&= \int_{\mathcal{S}} d\sigma_{\mathcal{S}}(s) \\
&= \mathcal{A}(\mathcal{S})
\end{aligned}$$

$$s(w_2) > 0;$$

mais, par l'Équation (9.2),

$$\begin{aligned}
s(w_2) &= \int_{\mathcal{T}} \langle w_2(s), N_{\mathcal{T}}(s) \rangle_{\mathbb{R}^3} d\sigma_{\mathcal{T}}(s) \\
&= \int_{\mathcal{T}} \langle N_{\mathcal{S}}(s), N_{\mathcal{T}}(s) \rangle_{\mathbb{R}^3} d\sigma_{\mathcal{T}}(s) \\
&= \int_{\mathcal{T}} -\|N_{\mathcal{T}}(s)\|_{\mathbb{R}^3}^2 d\sigma_{\mathcal{T}}(s) \\
&= \int_{\mathcal{T}} -d\sigma_{\mathcal{T}}(s) \\
&= -\mathcal{A}(\mathcal{T})
\end{aligned}$$

$$s(w_2) < 0.$$

On obtient une nouvelle contradiction, qui pousse à conclure que, nécessairement, $N_{\mathcal{S}} = N_{\mathcal{T}}$. Il vient ainsi que $\mathcal{S} \equiv \mathcal{T}$.

On a montré la relation

$$\forall \mathcal{S}, \mathcal{T} \in \Theta, \Psi(\mathcal{S}) = \Psi(\mathcal{T}) \implies \mathcal{S} \equiv \mathcal{T},$$

ce qui prouve que l'application Ψ est injective. On a par conséquent le résultat ci-dessous.

Théorème 8. *Sous les hypothèses faites en Section 9.1.1 sur V et Θ , l'application \mathcal{D}_{Θ} définie en Section 8.2.1 par l'Équation (8.1) est une distance sur Θ .*

9.1.3 Extension de la distance aux surfaces discrétisées

Notations

Afin de pouvoir utiliser dans un cadre pratique la structure d'espace métrique que l'on vient d'établir, il convient de l'étendre rigoureusement au cas des surfaces discrétisées. Pour ce faire, on se donne tout

d'abord une notion d'approximation discrète pour les surfaces orientées, analogue à celle construite en Section 3.3 pour les surfaces non-orientées.

Définition 19. Soit $\mathbb{S} \in \Theta$, une surface orientée, soit ρ un réel strictement positif; on définit une approximation discrète, avec précision ρ , \mathbb{S}_ρ de \mathbb{S} comme détaillé ci-dessous.

- Si \mathbb{S} est l'ensemble vide, alors \mathbb{S}_ρ est également l'ensemble vide.
- Si $\mathbb{S} \equiv (\mathcal{S}, N_{\mathcal{S}})$, alors \mathbb{S}_ρ est un couple $(\mathcal{S}_\rho, N_{\mathcal{S},\rho})$, tel que :
 - . \mathcal{S}_ρ est une approximation discrète, de la surface non-orientée \mathcal{S} , ainsi que définie en Section 3.3, vérifiant notamment les Hypothèses 1 et 2;
 - . $N_{\mathcal{S},\rho}$ est une application définie sur la surface discrétisée \mathcal{S}_ρ qui à chaque point se trouvant sur une arête associe le vecteur nul et à chaque point se trouvant à l'intérieur d'une facette associe un vecteur normal à cette dernière, dont le sens est déterminé de façon à ce que tous les vecteurs normaux engendrés par les fonctions $N_{\mathcal{S}}$ et $N_{\mathcal{S},\rho}$ vérifient le Lemme 2 de la Section 3.3.

Dans la suite, on note $\widetilde{\Theta}_\rho$ l'ensemble des surfaces discrétisées avec un paramètre de précision dans l'intervalle $]0, \rho]$. Soit $\widetilde{\Psi}_\rho$ l'application qui à toute surface discrétisée de $\widetilde{\Theta}_\rho$ associe son représentant dans V^* , tel que construit via l'Equation (3.3) donnée en Section 3.3.

Distance discrète

On définit, sur l'ensemble $\widetilde{\Theta}_\rho$ des surfaces discrètes, l'application $\widetilde{\mathcal{D}}_{\Theta_\rho}$ par

$$\begin{aligned} \widetilde{\Theta}_\rho \times \widetilde{\Theta}_\rho &\longrightarrow \mathbb{R} \\ (\mathcal{S}, \mathcal{T}) &\longrightarrow \|\widetilde{\Psi}_\rho(\mathcal{S}) - \widetilde{\Psi}_\rho(\mathcal{T})\|_{V^*}. \end{aligned}$$

On peut démontrer que $\widetilde{\Psi}_\rho$ est injective et, par conséquent, que $\widetilde{\mathcal{D}}_{\Theta_\rho}$ est une distance sur l'ensemble $\widetilde{\Theta}_\rho$ des surfaces discrétisées. La méthode employée étant similaire à celle utilisée en Section 9.1.2 pour montrer l'injectivité de la fonction Ψ , on n'écrira pas les détails de la preuve : Considérant deux surfaces discrètes \mathbb{S}_ρ et \mathbb{T}_ρ de $\widetilde{\Theta}_\rho$, on montrerait, comme dans le cas continu, que

$$\Psi_\rho(\mathbb{S}_\rho) = \Psi_\rho(\mathbb{T}_\rho) \implies \mathbb{S}_\rho \equiv \mathbb{T}_\rho.$$

La principale différence résiderait dans la formulation du choix des champs de vecteurs w_1 et w_2 , que l'on adapterait de fait au modèle discret. Dans le cas présent, w_1 serait choisi tel que

- . w_1 soit de classe C^1 sur \mathcal{D} ,

. w_1 soit identiquement nul sur \mathcal{S}_ρ ,

. $w_1|_{\mathcal{T}_\rho \setminus \mathcal{S}_\rho \cap \mathcal{T}_\rho} = x_1 \cdot N_{\mathcal{T}, \rho}$, où x_1 est une fonction réelle sur \mathcal{T}_ρ strictement positive,

et w_2 serait choisi tel que

. w_2 soit de classe C^1 sur \mathcal{D} ,

. $w_2|_{\mathcal{S}_\rho} = x_2 \cdot N_{\mathcal{S}, \rho}$, où x_2 est une fonction strictement positive sur \mathcal{S}_ρ .

Remarque 7. Dans la suite, pour des raisons de simplicité, on écrira indifféremment, par abus de langage, $\mathcal{D}_\Theta(\mathcal{S}, \mathcal{T})$ la distance entre deux surfaces \mathcal{S} et \mathcal{T} sans se soucier du caractère lisse ou discrétisé de ces dernières.

9.2 Propriétés de l'espace métrique des surfaces

9.2.1 Question de l'homéomorphisme entre Θ et V^*

On cherche maintenant déterminer les propriétés topologiques de l'espace métrique $(\Theta, \mathcal{D}_\Theta)$. À cet effet, on commence par chercher si Ψ est un homéomorphisme, ce qui impliquerait des propriétés analogues pour les espaces Θ et V^* sur le plan topologique.

La définition de l'application \mathcal{D}_Θ implique clairement la continuité de la fonction Ψ pour les topologies d'espace métrique sur Θ et V^* . On s'intéresse à présent à la question de la surjectivité de Ψ , propriété qui, vu la Section 9.1.2, impliquerait la bijectivité de cette application. Hélas, d'un point de vue intuitif, cela semble faux. En effet, il paraît peu crédible que toutes les formes linéaires de l'espace V^* soient engendrées par une surface, dans la mesure où, le cube \mathcal{D} sur lequel sont définis les champs de vecteurs de l'espace V étant tridimensionnel, il est possible de construire des formes linéaires sur V engendrées par des volumes inclus dans \mathcal{D} . On cherche donc un contre-exemple, sous la forme d'un élément f de V^* qui ne puisse pas s'écrire comme

$$f(w) = \int_{\mathcal{S}} \langle w(s), N_{\mathcal{S}} \rangle_{\mathbb{R}^3} d\sigma_{\mathcal{S}}(s), \forall w \in V,$$

où \mathcal{S} est une surface de l'ensemble Θ . Pour construire une telle forme linéaire, on utilise l'idée d'une forme linéaire engendrée par un volume mentionnée plus haut; on étudie, dans cette optique, la fonction f définie par :

$$f : V \longrightarrow \mathbb{R}$$

$$w \longrightarrow \int_0^1 \int_{\mathbb{S}_r} \langle w(s), N_{\mathbb{S}_r} \rangle_{\mathbb{R}^3} \sigma_{\mathbb{S}_r}(s) dr$$

où \mathbb{S}_r désigne la sphère de centre $(0, 0, 0)$ et de rayon $r \in [0, 1]$, et $N_{\mathbb{S}_r}$ est la normale unitaire externe à la sphère \mathbb{S}_r ; on pose par défaut N_0 comme un vecteur quelconque de norme 1.

On commence par démontrer que f est un élément de V^* . La linéarité de f sur V est évidente. De plus, on a

$$\begin{aligned} \forall w \in V, |f(w)| &= \left| \int_0^1 \int_{\mathbb{S}_r} \langle w(s), N_{\mathbb{S}_r} \rangle_{\mathbb{R}^3} s \sigma_{\mathbb{S}_r}(s) dr \right| \\ &\leq \int_0^1 \left| \int_{\mathbb{S}_r} \langle w(s), N_{\mathbb{S}_r} \rangle_{\mathbb{R}^3} s \sigma_{\mathbb{S}_r}(s) \right| dr. \end{aligned}$$

Ainsi, vu la Section 3.2.1, il existe une constante strictement positive C telle que

$$\forall w \in V, |f(w)| \leq C \int_0^1 \mathcal{A}(\mathbb{S}_r) dr \|w\|_V.$$

L'intégrale $\int_0^1 \mathcal{A}(\mathbb{S}_r) dr$ étant le volume de la boule-unité \mathcal{B} de \mathbb{R}^3 , elle vaut 1 ; on a donc

$$\forall w \in V, |f(w)| \leq C \|w\|_V,$$

où C est une constante positive. Par conséquent, la forme linéaire f est continue sur V , d'où le fait que f est bien un élément de V^* .

On peut montrer que f n'a pas d'antécédent par Ψ , ce qui garantit le résultat ci-dessous.

Proposition 7. *Sous les hypothèses faites en Section 9.1.1 sur V et Θ , l'application Ψ n'est pas surjective.*

Démonstration. On a vu, plus haut, que f est un élément de V^* . On va, à présent, démontrer qu'il n'existe pas de surface \mathcal{S} de Θ telle que $f = \Psi(\mathcal{S})$. Dans le cas où une telle surface existerait, on pourrait alors écrire f sous la forme

$$\forall w \in V, f(w) = \int_{\mathcal{S}} \langle w(s), N_{\mathcal{S}}(s) \rangle_{\mathbb{R}^3} d\sigma_{\mathcal{S}}(s). \quad (9.3)$$

On considère alors un champ de vecteurs w_0 de V qui vérifie les conditions suivantes :

- Le support de w_0 est inclus dans \mathcal{B} ;
- $\forall s \in \mathcal{S}, w_0(s) = 0$;
- il existe une fonction x_0 , de classe C^1 sur \mathcal{B} , nulle sur $\mathcal{B} \cap \mathcal{S}$ et sur \mathbb{S}_1 , et strictement positive partout ailleurs dans \mathcal{B} . Une telle fonction w_0 existe dans V de par le fait que V_0 contient l'espace de Schwartz sur \mathcal{D} .

Étant donné l'Équation (9.3), on devrait alors avoir :

$$\begin{aligned} f(w_0) &= \int_{\mathcal{S}} \langle w_0(s), N_{\mathcal{S}}(s) \rangle_{\mathbb{R}^3} d\sigma_{\mathcal{S}}(s) \\ &= \int_{\mathcal{S}} \langle (0, 0, 0), N_{\mathcal{S}}(s) \rangle_{\mathbb{R}^3} d\sigma_{\mathcal{S}}(s) \\ &= 0. \end{aligned}$$

Cependant, par définition de f , on a :

$$f(w_0) = \int_0^1 \int_{\mathbb{S}_r} \langle w_0(s), N_{\mathbb{S}_r} \rangle_{\mathbb{R}^3} \sigma_{\mathbb{S}_r}(s) dr.$$

Vu la définition du champ de vecteurs w_0 , le produit $\langle w_0(s), N_{\mathbb{S}_r} \rangle_{\mathbb{R}^3}$ est positif sur \mathcal{B} et positif sur une partie de \mathcal{B} de mesure non nulle. Par conséquent, la valeur $f(w_0)$ est non-nulle, ce qui amène une contradiction avec le résultat précédent. Par conséquent, f ne peut s'exprimer sous la forme donnée par l'Équation (9.3); il n'existe donc pas de surface \mathcal{S} de Θ telle que $f = \Psi(\mathcal{S})$. L'application Ψ n'est donc pas surjective. \square

Remarque 8. *La démonstration ci-dessus est indépendante de l'ensemble Θ et du cube \mathcal{D} choisis, de l'instant que \mathcal{D} est un compact de \mathbb{R}^3 et Θ un ensemble de surfaces continues et fermées de \mathcal{D} . Il est par conséquent manifestement clair que, dans notre cadre statistique, l'application Ψ ne sera jamais surjective. Il existe cependant une piste susceptible de permettre d'obtenir la surjectivité : on pourrait considérer un espace Θ plus complexe, incluant les unions au plus dénombrables de surfaces, en s'inspirant du cadre défini par Durrleman pour les unions de courbes dans [9].*

9.2.2 Question du caractère complet de Θ

Dans la section ci-dessus, on n'a pas exhibé d'isomorphisme entre Θ et V^* ; ces deux espaces métriques n'ont donc pas nécessairement les mêmes propriétés. De ce fait, la question du caractère complet de $(\Theta, \mathcal{D}_\Theta)$ n'est pas triviale. En l'état actuel de nos travaux, on ne connaît pas encore la réponse à ce problème; on dispose seulement de la Proposition suivante, établissant une condition nécessaire et suffisante sur l'application Ψ pour que l'espace métrique des surfaces soit complet.

Proposition 8. *Les deux assertions suivantes sont équivalentes :*

- (i) *L'espace métrique $(\Theta, \mathcal{D}_\Theta)$ est complet.*
- (ii) *L'image de Θ par la fonction Ψ est un fermé de l'espace V^* .*

Démonstration. D'une part, si (ii) n'est pas vraie, alors il existe une suite $(s_n)_{n \in \mathbb{N}}$ d'éléments de $\Psi(\Theta)$ qui converge vers un élément s de $V^* \setminus \Psi(\Theta)$. Par définition de la distance \mathcal{D}_Θ , la suite $(\mathcal{S}_n)_{n \in \mathbb{N}}$ des antécédents respectifs par Ψ des $s_n, n \in \mathbb{N}$ est alors une suite de Cauchy. Si elle venait à converger vers une limite $\mathcal{S} \in \Theta$, alors, par continuité de Ψ , la suite $(s_n)_{n \in \mathbb{N}}$ convergerait vers $\Psi(\mathcal{S})$; par unicité de la limite, on aurait alors $s = \Psi(\mathcal{S})$ et donc que $s \in \Psi(\Theta)$, ce qui constitue une contradiction. Donc la suite $(\mathcal{S}_n)_{n \in \mathbb{N}}$ ne peut converger, d'où le fait que $(\Theta, \mathcal{D}_\Theta)$ n'est pas complet.

Réciproquement, si (ii) est vraie, alors, par définition de la distance \mathcal{D}_Θ , pour toute suite de Cauchy $(\mathcal{S}_n)_{n \in \mathbb{N}}$ de Θ , la suite $(\Psi(\mathcal{S}_n))_{n \in \mathbb{N}}$ est également une suite de Cauchy dans V^* ; cet espace étant complet, $(\Psi(\mathcal{S}_n))_{n \in \mathbb{N}}$ admet alors une limite $s \in V^*$; (ii) assure ainsi que $s \in \Psi(\Theta)$, d'où l'existence d'un élément \mathcal{S} de Θ tel que $\Psi(\mathcal{S}) = s$. On a

alors, pour tout entier naturel n ,

$$\mathcal{D}_\Theta(\mathcal{S}, \mathcal{S}) = \|\Psi(\mathcal{S}_n) - \Psi(\mathcal{S})\|_{V^*} = \|\Psi(\mathcal{S}_n) - s\|_{V^*},$$

d'où le fait que la suite $(\mathcal{S}_n)_{n \in \mathcal{N}}$ converge vers \mathcal{S} . L'espace métrique $(\Theta, \mathcal{D}_\Theta)$ est donc complet, d'où l'équivalence. □

Pour répondre à notre interrogation quant au caractère complet de l'espace métrique $(\Theta, \mathcal{D}_\Theta)$, il ne reste ainsi qu'à explorer la question de la véracité de la condition (ii). À cet effet, on pourra par exemple chercher si Ψ est une application fermée. Comme cela a été évoqué plus haut, cette partie du travail reste à faire et la question reste, à ce stade, en suspens.

Chapitre 10

Méthode de classification Bayésienne adaptée aux surfaces

10.1 Principe

10.1.1 Classification : pose du problème

Ce chapitre est consacré à la construction d'un algorithme de classification opérationnel sur les surfaces, dans le cadre de l'espace métrique $(\Theta, \mathcal{D}_\Theta)$ défini précédemment au Chapitre 9. Compte tenu des motivations de notre étude, évoquées entre autres en Introduction de la présente partie, on se limite au cas d'une classification à deux catégories : étant donné deux parties de Θ , disjointes et non vides, notées Θ_1 et Θ_2 , on cherche un critère permettant de décider si une surface \mathcal{S} de $\Theta_1 \cup \Theta_2$ est élément de Θ_1 ou de Θ_2 . Dans les applications concrètes, les sous-ensembles Θ_1 et Θ_2 correspondent typiquement à deux populations d'organes similaires présentant des phénotypes opposés l'un par rapport à l'autre.

Pour mettre en place un critère de décision, on se base sur un échantillon de surfaces que l'on sait décrire en termes d'appartenance aux ensembles Θ_1 et Θ_2 ; on fait donc l'hypothèse que l'on dispose de deux populations de surfaces $(\mathcal{S}_1^1, \dots, \mathcal{S}_{K_1}^1) \subset \Theta_1$ et $(\mathcal{S}_1^2, \dots, \mathcal{S}_{K_2}^2) \subset \Theta_2$. Pour toute surface $\mathcal{S} \in \Theta_1 \cup \Theta_2$, on veut pouvoir, à partir de ces données, calculer un paramètre dont la valeur constituera un indice significatif quant à l'appartenance de \mathcal{S} à Θ_1 ou Θ_2 .

10.1.2 Classification Bayésienne

Pour résoudre le problème présenté ci-dessus, on adapte la méthode de classification Bayésienne à deux catégories, détaillée par Duda et Hart dans [8]. On construit un critère de décision analogue à celui qui y est

explicité, de façon transposée à notre cadre d'espace métrique sur les surfaces.

On utilise des notations similaires à celle de la section 10.1.1. Soient S le représentant de \mathcal{S} et $S_1^1, \dots, S_{K_1}^1, S_1^2, \dots, S_{K_2}^2$ les représentants respectifs de $\mathcal{S}_1^1, \dots, \mathcal{S}_{K_1}^1, \mathcal{S}_1^2, \dots, \mathcal{S}_{K_2}^2$ dans l'espace V . Dans les applications, on travaillera sur les projections de ces formes linéaires sur l'espace d'approximation \widetilde{V}_N .

Lois Conditionnelles

On se place dans un modèle Gaussien multivarié tel que celui construit dans la première partie; dans ce cadre, il est facile d'exprimer les lois conditionnelles de S^N sachant les événements $\{S \in \Theta_i\}, i \in \{1, 2\}$. Comme dans le modèle de variation (3.2) décrit en Section 3.2.3, ces lois s'expriment sous la forme

$$S^N = \mu_i^N + B_i^N, i \in \{1, 2\}, \quad (10.1)$$

où $\mu_i^N, i \in \{1, 2\}$ sont des formes linéaires de \widetilde{V}_N^* et $B_i^N, i \in \{1, 2\}$ des formes linéaires aléatoires Gaussiennes centrées d'autocovariance $\Gamma_i^N, i \in \{1, 2\}$. La forme μ_1^N (respectivement μ_2^N) et l'opérateur Γ_1^N (respectivement Γ_2^N) n'étant a priori pas connus, on les identifie à l'espérance et à l'opérateur d'autocovariance estimés via l'échantillon des réalisations $S_1^1(\omega), \dots, S_{K_1}^1(\omega)$ (respectivement via l'échantillon $S_1^2(\omega), \dots, S_{K_2}^2(\omega)$).

Compte tenu de ces approximations, les lois conditionnelles respectives sachant $\{S \in \Theta_i\}, i \in \{1, 2\}$ du vecteur des coordonnées de S^N , sont données par les densités $f_i^N, i \in \{1, 2\}$ définies sur $\mathbb{R}^{|I_N|}$ par

$$f_i^N(x) = \frac{1}{\sqrt{2\pi \text{Det}(\Gamma_i^N)^{|I_N|}}} \exp((x - \mu_i)^t \Gamma_i^{-1} (x - \mu_i)). \quad (10.2)$$

Critère de décision

Les approximations des deux lois conditionnelles permettent de se donner un critère de décision basé sur la correspondance de l'observation $S^N(\omega)$ avec l'une ou l'autre de ces dernières. On décide donc que $S \in \Theta_{i^*}$, où i^* est défini par

$$i^* = \underset{i \in \{1, 2\}}{\text{argmax}} \mathbb{P}(S = S(w) |_{S \in \Theta_i}).$$

Étant donné les densités exprimées plus haut via l'équation (10.2), i^* s'écrit alors

$$i^* = \underset{i \in \{1, 2\}}{\text{argmin}} \left[\frac{|I_N|}{2} \log(\text{Det}(\Gamma_i^N)) + \frac{1}{2} (s^N - \mu_i)^t \Gamma_i^{-1} (s^N - \mu_i) \right],$$

où s^N désigne le vecteur des coordonnées de la réalisation $S^N(\omega)$.

De fait, on a $i^* = 1$ si

$$\frac{|I_N|}{2} \log\left(\frac{\text{Det}(\Gamma_1^N)}{\text{Det}(\Gamma_2^N)}\right) + \frac{1}{2} ((s^N - \mu_1)^t \Gamma_1^{-1} (s^N - \mu_1) - (s^N - \mu_2)^t \Gamma_2^{-1} (s^N - \mu_2)) < 0,$$

et $i^* = 2$ sinon.

10.2 Algorithme

10.2.1 Cas général

On dispose d'un échantillon de surfaces $(S_1^1, \dots, S_{K_1}^1, S_1^2, \dots, S_{K_2}^2)$ tel que $(S_1^1, \dots, S_{K_1}^1) \subset \Theta_1$ et $(S_1^2, \dots, S_{K_2}^2) \subset \Theta_2$. Soit $S \in \Theta_1 \cup \Theta_2$. On décide si $S \in \Theta_1$ ou $S \in \Theta_2$ via la procédure ci-dessous, se déroulant en deux phases.

Phase d'apprentissage

- À partir de l'échantillon $(S_1^1, \dots, S_{K_1}^1)$, on estime l'espérance μ_1^N et l'autocovariance Γ_1^N de la loi conditionnelle sachant l'événement $\{S \in \Theta_1\}$ du représentant approximé S^N de S .
- À partir de l'échantillon $(S_1^2, \dots, S_{K_2}^2)$, on estime l'espérance μ_2^N et l'autocovariance Γ_2^N de la loi conditionnelle sachant l'événement $\{S \in \Theta_2\}$ du représentant approximé S^N de S .

Phase de test

On calcule le paramètre de décision C défini par

$$C = \frac{|I_N|}{2} \log \left(\frac{\text{Det}(\Gamma_1^N)}{\text{Det}(\Gamma_2^N)} \right) + \frac{1}{2} \left((s^N - \mu_1)^t \Gamma_1^{-1} (s^N - \mu_1) - (s^N - \mu_2)^t \Gamma_2^{-1} (s^N - \mu_2) \right).$$

Si $C < 0$, alors on décide que S est un élément de Θ_1 ; Dans le cas contraire, on décide que S est un élément de Θ_2 .

10.2.2 Cas particulier de l'autocovariance diagonale

Il est possible de simplifier le problème de classification en approximant les opérateurs d'autocovariance associés aux lois conditionnelles par des matrices diagonales. Cette approximation revient à négliger la corrélation entre les coordonnées des représentants des surfaces. Dans un tel cas, en notant, pour tout i dans $\{1, 2\}$ et pour tout j dans I_N , $\sigma_{j,i}$ l'écart-type de $S(\phi_j)$ sachant l'événement $\{S \in \Theta_i\}$, on écrit alors la

valeur du paramètre de décision C de façon plus simple :

$$C = \frac{|I_N|}{2} \log \left(\frac{\prod_{j=1}^{|I_N|} \sigma_{j,1}}{\prod_{j=1}^{|I_N|} \sigma_{j,2}} \right) + \frac{1}{2} \sum_{j=1}^{|I_N|} ((s_j^N - \mu_1(\phi_j))^2 \sigma_{j,1}^{-2} - (s_j^N - \mu_2(\phi_j))^2 \sigma_{j,2}^{-2}),$$

où les $s_j^N, j \in I_N$ sont les coordonnées du vecteur s^N .

Chapitre 11

Application à l'imagerie cérébrale

11.1 Objectifs

Dans ce chapitre, on fait le compte rendu d'une expérience permettant d'appliquer la méthode de classification Bayésienne adaptée aux surfaces que l'on a détaillée plus haut, en Section 10.2. Pour ce faire, on applique cette dernière à un échantillon de données réelles, constitué de surfaces issues de l'imagerie cérébrale.

Dans un premier temps, on explicite le protocole expérimental : on décrit précisément l'échantillon de surfaces employé puis on détaille les modalités d'application de l'algorithme (valeurs des paramètres, hypothèses sur les opérateurs, etc.). Dans un second temps, on donne et on analyse les résultats obtenus.

11.2 Protocole

11.2.1 Échantillon de surfaces

L'échantillon de surfaces utilisé pour cette expérience est construit à partir des 152 surfaces cérébrales, tridimensionnelles et discrétisées, de la base de données ICBM, décrite dans [13]; les individus étudiés sont des adultes des deux sexes, dont l'âge varie entre 19 et 44 ans. L'acquisition de ces données a été réalisée par IRM 3-D (induction : 1.5 Tesla, résolution : $1.5mm \times 1.5mm \times 1.5mm$). La segmentation et le maillage sont effectués avec le logiciel Brainvisa et les méthodes présentées dans [22].

Chacune des surfaces de la base de données ICBM été partagée en deux moitiés, correspondant respectivement à l'un des deux hémisphères du cerveau représenté. Ainsi, on a 304 surfaces, 152 correspondant à des hémisphères droits et 152 à des hémisphères gauches. On dispose donc d'un échantillon $(S_1^1, \dots, S_{152}^1, S_1^2, \dots, S_{152}^2)$, formé des deux populations S_1^1, \dots, S_{152}^1 (hémisphères droits) et S_1^2, \dots, S_{152}^2 (hé-

mispères gauches). Cet échantillon a ainsi toutes les caractéristiques pour être utilisé dans le cadre de notre expérience : il est composé de deux populations géométriquement distinctes et clairement identifiées, ce qui permettra d'évaluer de façon exacte l'erreur de généralisation commise en classant les surfaces via l'algorithme de la Section 10.2.

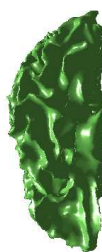


FIGURE 11.1 – Exemple d'hémisphère droit.

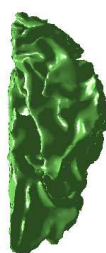


FIGURE 11.2 – Exemple d'hémisphère gauche.

11.2.2 Choix des paramètres

Autocovariance

L'algorithme développé en Section 10.2 est donné pour des opérateurs d'autocovariance Σ_1 et Σ_2 quelconques. Pour cette expérience, on fait l'Hypothèse simplificatrice que ces deux opérateurs sont diagonaux et on utilise donc le critère donné en Section 10.2.2. Cette approximation permet de réduire le nombre de composantes à estimer : l'opérateur d'autocovariance étant une matrice carrée d'ordre $3N^3$, on calculera ainsi $3N^3$ coefficients au lieu de $9N^6$. Cela présente deux avantages majeurs : d'une part la baisse de la dimension permet de passer outre le fait que l'on a peu d'observations ; d'autre part on réduit ainsi la complexité de la procédure, ce qui induira une baisse du temps de calcul d'autant plus importante que l'estimation de l'autocovariance sera, comme expliqué plus bas, faite plusieurs fois au cours de cette expérience.

Dimension de l'espace d'approximation

Pour fixer la valeur de l'entier N , induisant la dimension du sous-espace d'approximation, pour rappel $3N^3$, il convient de tenir compte du nombre de surfaces disponibles dans les données. Le nombre de paramètres mis en jeu ne devant pas être supérieur au cardinal de l'échantillon utilisé, on doit, de fait, avoir

$$3N^3 \leq 152,$$

ce qui donne un paramètre de dimension $N = 3$ au maximum, soit un sous-espace d'approximation de dimension 81, ce qui est très limité en comparaison de ce que l'on a utilisé dans les expériences du Chapitre 6.

11.2.3 Procédure de validation croisée

Pour chaque surface $S_k^i, i \in \{1, 2\}, k \in \{1, \dots, K_i\}$, de l'échantillon, on effectue les étapes suivantes :

- on applique la phase d'apprentissage telle que définie en Section 10.2.1 à l'échantillon privé de S_k^i ;
- avec toutes les valeurs liées à l'apprentissage calculées à l'étape précédente, on applique la phase de test définie en Section 10.2.1 sur la surface S_k^i , comme si on ignorait la valeur de i , et on note enfin k_i^* la valeur de k ainsi estimée.

Enfin, pour chacun des deux échantillons $\{S_1^1, \dots, S_{152}^1\}$ et $\{S_1^2, \dots, S_{152}^2\}$, on calcule la fréquence de réussite de la procédure de classification, c'est à dire la fréquence de réalisation de l'événement " $k_i^* = k$ " lors de la deuxième étape.

11.3 Résultats

On effectue l'expérience avec le protocole et l'échantillon de surfaces cérébrales décrits plus haut en Section 11.2. La fréquence de réussite obtenue à l'issue de l'ensemble des tests est de 88.1% pour la population $\{\mathcal{S}_1^1, \dots, \mathcal{S}_{152}^1\}$ (surfaces droites) et de 90.1% pour la population $\{\mathcal{S}_1^2, \dots, \mathcal{S}_{152}^2\}$ (surfaces gauches). Cette précision, avoisinant en moyenne 90%, est légèrement inférieure aux résultats actuellement existants dans la littérature, notamment ceux publiés dans [7].

La question des moyens d'améliorer les résultats en termes d'erreur de généralisation émerge ainsi tout naturellement. Une première possibilité serait d'augmenter la dimension de l'espace d'approximation, et d'utiliser une méthode de régularisation de type Tykhonov pour compenser le fait que l'opérateur d'autocovariance n'est pas inversible. Il faut cependant demeurer prudent quant à l'augmentation de la dimension dans la mesure où l'erreur de généralisation tend à augmenter avec la complexité. On pourrait aussi faire de la sélection de variables : générer l'opérateur d'autocovariance pour une grande valeur de N et en extraire un nombre de composantes principales adapté à la taille de l'échantillon. Enfin, une troisième possibilité serait de changer la base de travail : on pourrait choisir une base induisant moins de corrélation entre les coefficients et donc réduisant l'erreur due à l'approximation de l'autocovariance par un opérateur diagonal.

Chapitre 12

Discussion

12.1 Bilan

Dans cette partie, on a, d'une part, construit une structure d'espace métrique sur les surfaces à partir du cadre de représentation par courants de la première partie. Pour ce faire, on a montré que, sous des hypothèses raisonnables, la fonction Ψ , associant à chaque surface son représentant, est injective et ainsi que la distance entre les représentants induit une distance sur les surfaces. Du fait de la non surjectivité de Ψ , il n'a pas été possible de transférer à l'espace métrique Θ des surfaces toutes les propriétés de l'ensemble des représentants. Entre autres, la question du caractère complet de Θ reste ouverte.

D'autre part, on a mis en place, dans notre cadre statistique, une procédure de classification des surfaces reposant sur un critère de décision Bayésien. On l'a appliquée à des échantillons de données issues de l'imagerie cérébrale. Pour l'heure, nos résultats en termes de précision sont légèrement en deçà de ceux de la littérature [7].

12.2 Perspectives

12.2.1 Espace des surfaces

Dans les travaux à venir, on poursuivra l'étude des propriétés de l'espace métrique des surfaces construit au Chapitre 9; dans cette optique, une première tâche sera de répondre à la question laissée en suspens en Section 9.2.2 quant au caractère complet de cet ensemble. D'autre part, on tentera d'adapter le cadre d'étude pour obtenir une application Ψ qui soit une bijection. À cet effet, on envisage de considérer un espace Θ constitué non plus seulement de surfaces mais aussi de réunion de surfaces pondérées. À cet effet, on utilisera les résultats existant dans la littérature pour les réunions de courbes; on pourra notamment se

référer aux travaux de Durrleman [9] et Gori [18].

12.2.2 Classification

En ce qui concerne les applications à la classification des surfaces, on travaillera suivant deux axes principaux. D'une part, on mettra en place d'autres expériences : on pourra par exemple appliquer la procédure de classification Bayésienne donnée en Section 10.2 à la résolution du problème de la reconnaissance des catégories "homme" et "femme" au sein de l'échantillon constitué des 152 surfaces cérébrales de la base de données ICBM. Cette expérience est elle aussi déjà présente dans la littérature [7, 34], ce qui nous donnera la possibilité d'évaluer nos résultats de façon relative. D'autre part, on adaptera l'algorithme afin de diminuer l'erreur de généralisation. A cet effet, on envisagera, entre autres, de faire de la sélection de variables sur les composantes de l'opérateur d'autocovariance. On expérimentera également sur d'autres bases de l'espace des champs de vecteurs, afin d'en sélectionner une induisant une plus faible corrélation entre les dimensions.

Conclusion générale

Bilan général

Cadre statistique opérationnel sur les surfaces

Dans la première partie, on a mis en place un cadre statistique permettant d'agir sur les ensembles de surfaces. Ce cadre est construit via une représentation des surfaces par des formes linéaires sur un espace de champs de vecteurs, ce qui induit un plongement de ces objets dans un espace de Hilbert. Cette structure Euclidienne permet de définir un modèle de variation basé sur l'utilisation de processus linéaires généralisés. On a ensuite construit un modèle d'observation par passage du continu au discret.

Dans ce modèle d'observation, on a construit des estimateurs, consistants en termes d'erreur quadratique, pour le représentant moyen d'un échantillon de surfaces et pour la variance du bruit associé. Une approximation des représentants des surfaces, via une projection sur un sous-espace vectoriel de dimension finie, assure la faisabilité des calculs.

Application à la classification

Dans la deuxième partie, on a montré que la représentation par formes linéaires induit une structure d'espace métrique sur l'ensemble des surfaces. On utilise cette structure pour construire une méthode de classification Bayésienne adaptée aux surfaces. Appliquée à deux échantillons de données réelles, dans le cadre d'un problème de classement de surfaces cérébrales en deux catégories (hémisphère droit/hémisphère gauche), cette méthode donne des résultats dans la norme basse de ceux de la littérature actuelle, au vu de la précision obtenue dans [7].

Perspectives

Modèle de prototype déformable

Comme cela a été précédemment évoqué, notamment en Section 7.2.2, il existe dans la littérature de nombreux travaux proposant l'étude d'un modèle de type prototype déformable défini sur les surfaces, en vue d'applications à l'imagerie médicale et à l'anatomie computationnelle [1, 6]. Une suite logique au travail effectué dans ce manuscrit serait l'écriture d'un tel modèle dans le cadre statistique construit en première partie puis le développement d'une procédure, spécifique à ce cadre, d'estimation conjointe des déformations et du prototype ; à cet effet, on adaptera aux surfaces représentées par des courants les méthodes SAEM existant actuellement pour le traitement d'images bidimensionnelles. Une question majeure qui émerge est celle de la stabilité du cadre de représentation lors de l'application d'une déformation à une surface : étant donné une surface \mathcal{S} et une déformation ϕ , a-t-on toujours des équations explicites sur le représentant de $\phi(\mathcal{S})$?

Analyse en composantes principales (ACP)

On envisage d'adapter à notre cadre statistique sur les surfaces représentées par des formes linéaires des techniques d'analyse en composantes principales. En particulier, on se dirigera vers l'ACP fonctionnelle, particulièrement adaptée à l'espace des représentants dont les éléments sont des formes linéaires. Dans la littérature, il existe de nombreuses ressources sur cette thématique [3, 30, 28]. La principale contribution envisagée serait une extension d'un cadre d'ACP fonctionnelle à un espace de formes linéaires aléatoires ; pour ce faire, on s'appuiera sur les travaux de Meidan concernant les processus généralisés [23].

Problématique du choix de la base de l'espace de champs de vecteurs

Pour toutes les expériences détaillées dans ce manuscrit, que ces dernières concernent l'estimation ou la classification, on a exprimé les champs de vecteurs dans la base de Fourier ; pour construire une famille génératrice de l'espace d'approximation de dimension N , on s'est systématiquement contenté d'une sélection par défaut des N premiers vecteurs de cette base, c'est à dire des N plus basses fréquences. Cette absence d'adaptation de la base à la situation est une faiblesse de notre travail, à laquelle on va remédier à l'avenir, selon deux axes. D'une part, on pourra envisager l'utilisation d'une base d'ondelettes, plus appropriée dans certains cas, car reflétant mieux l'information locale que la base de Fourier ; de plus, un tel choix serait susceptible d'introduire une plus grande parcimonie sur la matrice de précision. D'autre part, on mettra en place une procédure de sélection des composantes portant le plus d'information pour un échantillon donné ; on espère ainsi diminuer l'erreur, surtout dans les situations de classification, où on

dispose d'une base de cardinal relativement réduit, comme cela a été constaté au Chapitre 11.

Bibliographie

- [1] S. Allasonnière, Y. Amit, and A. Trouvé. Towards a coherent statistical framework for dense deformable template estimation. *Journal of the Royal Statistical Society : Series B (Statistical Methodology)*, 69(1) :3–29, 2007.
- [2] S. Allasonnière, J. Bigot, J. Glaunès, F. Maire, and F. Richard. Statistical models for deformable templates in image and shape analysis. *Annales Mathématiques de l’Institut Blaise Pascal*, 20(1) :1–35, 2013.
- [3] G.I. Allen. Sparse and functional principal components analysis. *arXiv preprint arXiv :1309.2895*, 2013.
- [4] N. Aronszajn. Theory of reproducing kernels. *Transactions of the American Mathematical Society*, 68(3) :337–404, 1950.
- [5] R. Bhattacharya and V. Patrangenaru. Statistics on manifolds and landmarks based image analysis : a non-parametric theory with applications. *Journal of Statistical Planning and Inference*, 145(1) :1–22, 2014.
- [6] N. Charon and A. Trouvé. The varifold representation of non-oriented shapes for diffeomorphic registration. *SIAM Journal on Imaging sciences*, 6(4) :2547–2580, 2013.
- [7] E. Duchesnay, A. Cachia, A. Roche, D. Rivière, D. Cointepas, D. Papadopoulos-Orfanos, M. Zilbovicius, J. Martinot, J. Régis, and J. Mangin. Classification based on cortical folding patterns. *IEEE Transactions on Medical Imaging*, 26(4) :553–565, 2007.
- [8] R. Duda, P. Hart, and D. Stork. *Pattern Classification*, chapter 2. Wiley New York, 1973.
- [9] S. Durrleman, P. Fillard, X. Pennec, A. Trouvé, and N. Ayache. Registration, atlas estimation and variability analysis of white matter fiber bundles modeled as currents. *Neuroimage*, 55(3) :1073–109, 2009.
- [10] S. Durrleman, X. Pennec, A. Trouvé, and N. Ayache. Statistical models of sets of curves and surfaces based on currents. *Medical Image Analysis*, 13(5) :793–808, 2008.
- [11] J. Fageot and M. Unser. Scaling limits of solutions of spde driven by Lévy white noises. *arXiv preprint arXiv :1610.06711*, 2016.

- [12] X. Fernique. Processus linéaires, processus généralisés. In *Annales de l'institut Fourier*, volume 17, pages 1–92, 1967.
- [13] V. Fonov, A.C. Evans, K. Botteron, C.R. Almlı, R.C. McKinstry, and D.L. Collins. Unbiased average age-appropriate atlases for pediatric studies. *Neuroimage*, 54(1) :313–327, 2011.
- [14] Z. Gelbaum and M. Titus. Simulation of fractional Brownian surfaces via spectral synthesis on manifolds. *IEEE Transactions on Image Processing*, 23(10) :4383–4388, 2014.
- [15] I. Gelfand and N.Y. Vilenkin. *Generalized Functions*, volume 4. Academic Press, 1964.
- [16] J.A. Glaunès. *Transport par Difféomorphismes de Points, de Mesures et de Courants pour la Comparaison des Formes et l'Anatomie Numérique*. PhD thesis, Université Paris 13, 2005.
- [17] J.A. Glaunès and S. Joshi. Template estimation from unlabeled point set data and surfaces for computational anatomy. In X. Pennec and S. Joshi, editors, *Proceedings of the International Workshop on the Mathematical Foundations of Computational Anatomy*, pages 29–39. MFCA, 2006.
- [18] P. Gori, O. Colliot, L. Marrakchi-Kacem, Y. Worbe, F. Fallani, M. Chavez, C. Pouponl, A. Hartmann, N. Ayache, and S. Durrleman. Parsimonious approximation of streamline trajectories in white matter fiber bundles. *IEEE Transactions on Medical Imaging*, 35(12) :2609–2619, 2016.
- [19] P. Gori, O. Colliot, Y. Worbe, L. Marrakchi-Kacem, S. Lecomte, C. Poupon, A. Hartmann, N. Ayache, and S. Durrleman. Bayesian atlas estimation for the variability analysis of shape complexes. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 267–274, 2013.
- [20] K. Itô. Stationary random distributions. *Memoirs of the College of Science, University of Kyoto. Series A : Mathematics*, 28(3) :209–223, 1954.
- [21] K. Itô. Isotropic random current. In *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability*, pages 125–13, 1956.
- [22] J-F. Mangin, D. Riviere, A. Cachia, E. Duchesnay, Y. Cointepas, D. Papadopoulos-Orfanos, D-L. Collins, A. Evans, and J. Régis. Object-based morphometry of the cerebral cortex. *IEEE Transactions on Medical Imaging*, 23(8) :968–982, 2004.
- [23] R. Meidan. On the connection between ordinary and generalized stochastic processes. *Journal of Mathematical Analysis and Applications*, 76(1) :124–133, 1980.
- [24] M. Miller and L. Younes. Bayesian template estimation in computational anatomy. *Neuroimage*, 42(1) :252–261, 2008.
- [25] H. Omer. *Modèles de déformation de processus stochastiques généralisés. Application à l'estimation des non-stationnarités dans les signaux audio*. PhD thesis, Université d'Aix Marseille, 2015.

- [26] H. Omer and B. Torrèsani. Time-frequency and time-scale analysis of deformed stationary processes, with application to non-stationary sound modeling. *Applied and Computational Harmonic Analysis*, 2015.
- [27] X. Pennec. Intrinsic statistics on riemannian manifolds : Basic tools for geometric measurements. *Journal of Mathematical Imaging and Vision*, 25(1) :127–154, 2006.
- [28] H-L. Shang. A survey of functional principal component analysis. *AStA Advances in Statistical Analysis*, 98(2) :121–142, 2014.
- [29] F.M. Tilotta, J.A. Glaunès, F. Richard, and Y. Rozenholc. A local technique based on vectorized surfaces for craniofacial reconstruction. *Forensic Science International*, 200(1) :50–59, 2010.
- [30] Ngoc Mai Tran. *An Introduction to Theoretical Properties of Functional Principal Component Analysis*. PhD thesis, Department of Mathematics and Statistics, The University of Melbourne, 2008.
- [31] A. Trouvé and L. Younes. Local geometry of deformable templates. *SIAM Journal on Mathematical Analysis*, 37(1) :17–59, 2005.
- [32] M. Unser and P. Tafti. *An introduction to sparse stochastic processes*. Cambridge University Press, 2014.
- [33] M. Vaillant and J.A. Glaunès. Surface matching via currents. In G. Christensen and M. Sonka, editors, *Information Processing in Medical Imaging*, volume 3565. Springer Berlin / Heidelberg, 2005.
- [34] C. Wachinger, P. Golland, W. Kremen, B. Fischl, and M. Reuter. Brainprint : a discriminative characterization of brain morphology. *Neuroimage*, 109 :232–248, 2015.
- [35] A.M. Yaglom. *Correlation Theory of Stationary and Related Random Functions : Supplementary Notes and References*. Springer Science & Business Media, 2012.

Annexe A

Calcul explicite du représentant d'une surface dans une base de Fourier

A.1 Objectifs

Cette annexe est consacrée aux calculs associés aux expériences du Chapitre 6. On note $\mathcal{D} = [-\pi, \pi]^3$ et V l'espace de champs de vecteurs défini en Section 6.3. On souhaite calculer explicitement, pour toute surface discrète S de représentant S dans V^* , les valeurs $S^N(w)$, $w \in V$, $N \in \mathbb{N}^*$ ainsi que les projections orthogonales v_S^N , $N \in \mathbb{N}^*$ du représentant de Riesz de S sur les espaces d'approximation, à partir de la seule donnée des triangles constituant S . Vu la linéarité de S et la formule de calcul du représentant de Riesz établie en Section 4.2, il suffit pour cela de calculer, pour tout triangle f inclus dans \mathcal{D} , les valeurs des intégrales

$$\int_f \langle \phi_{n,d}(x), N_f \rangle_{\mathbb{R}^3} dx, \quad (n, d) \in \mathbb{N} \times \{1, 2, 3\},$$

et

$$\int_f \langle \psi_{n,d}(x), N_f \rangle_{\mathbb{R}^3} dx, \quad (n, d) \in \mathbb{N} \times \{1, 2, 3\},$$

où N_f désigne un vecteur normal unitaire au triangle f .

A.2 Méthode de calcul

Soit f un triangle inclus dans le cube \mathcal{D} et soient $A = (x_A, y_A, z_A)$, $B = (x_B, y_B, z_B)$ et $C = (x_C, y_C, z_C)$ ses sommets. On note $(N_{f,1}, N_{f,2}, N_{f,3})$ les coordonnées du vecteur normal externe N_f . Pour simplifier ultérieurement les écritures, on définit, pour tout $d \in \{1, 2, 3\}$, la constante

$$\tilde{C}_d = \frac{N_d}{(2\pi)^3} |(x_B - x_A + y_B - y_A + z_B - z_A)(x_C - x_A + y_C - y_A + z_C - z_A)|.$$

On commence les calculs en effectuant un changement de variables via une paramétrisation du triangle f et en transformant le cosinus et le sinus en exponentielles :

$$\begin{aligned}\int_f \langle \phi_{n,d}(x), N_f \rangle_{\mathbb{R}^3} dx &= \widetilde{C}_d \int_0^1 \int_0^{1-s} \cos(\langle a + (b-a)s + (c-a)t, n \rangle_{\mathbb{R}^3}) dt ds, \\ &= \operatorname{Re} \left(\widetilde{C}_d e^{i\langle a, n \rangle_{\mathbb{R}^3}} \int_0^1 e^{is\langle b-a, n \rangle_{\mathbb{R}^3}} \int_0^{1-s} e^{it\langle c-a, n \rangle_{\mathbb{R}^3}} dt ds \right), \\ \text{et } \int_f \langle \psi_{n,d}(x), N_f \rangle_{\mathbb{R}^3} dx &= \widetilde{C}_d \int_0^1 \int_0^{1-s} \sin(\langle a + (b-a)s + (c-a)t, n \rangle_{\mathbb{R}^3}) dt ds, \\ &= \operatorname{Im} \left(\widetilde{C}_d e^{i\langle a, n \rangle_{\mathbb{R}^3}} \int_0^1 e^{is\langle b-a, n \rangle_{\mathbb{R}^3}} \int_0^{1-s} e^{it\langle c-a, n \rangle_{\mathbb{R}^3}} dt ds \right).\end{aligned}$$

On calcule alors ces deux intégrales en distinguant les cas selon les valeurs de $\langle b-a, n \rangle_{\mathbb{R}^3}$, $\langle c-a, n \rangle_{\mathbb{R}^3}$ et $\langle c-b, n \rangle_{\mathbb{R}^3}$. On procède, lorsqu'il n'est pas possible de trouver directement une primitive, par intégrations par parties successives.

A.3 Résultats

- Si $\langle b-a, n \rangle_{\mathbb{R}^3} = 0$ et $\langle c-a, n \rangle_{\mathbb{R}^3} = 0$,

$$\begin{aligned}\int_f \langle \phi_{n,d}(x), N_f \rangle_{\mathbb{R}^3} dx &= \frac{\widetilde{C}_d \cos(\langle a, n \rangle_{\mathbb{R}^3})}{2}, \\ \text{et } \int_f \langle \psi_{n,d}(x), N_f \rangle_{\mathbb{R}^3} dx &= \frac{\widetilde{C}_d \sin(\langle a, n \rangle_{\mathbb{R}^3})}{2}.\end{aligned}$$

- Si $\langle b-a, n \rangle_{\mathbb{R}^3} = 0$ et $\langle c-a, n \rangle_{\mathbb{R}^3} \neq 0$,

$$\begin{aligned}\int_f \langle \phi_{n,d}(x), N_f \rangle_{\mathbb{R}^3} dx &= \widetilde{C}_d \left[\frac{-\sin(\langle a, n \rangle_{\mathbb{R}^3})}{\langle c-a, n \rangle_{\mathbb{R}^3}} + \frac{\cos(\langle a, n \rangle_{\mathbb{R}^3})}{\langle c-a, n \rangle_{\mathbb{R}^3}^2} + \frac{-\cos(\langle c, n \rangle_{\mathbb{R}^3})}{\langle c-a, n \rangle_{\mathbb{R}^3}^2} \right], \\ \text{et } \int_f \langle \psi_{n,d}(x), N_f \rangle_{\mathbb{R}^3} dx &= \widetilde{C}_d \left[\frac{\cos(\langle a, n \rangle_{\mathbb{R}^3})}{\langle c-a, n \rangle_{\mathbb{R}^3}} + \frac{\sin(\langle a, n \rangle_{\mathbb{R}^3})}{\langle c-a, n \rangle_{\mathbb{R}^3}^2} + \frac{-\sin(\langle c, n \rangle_{\mathbb{R}^3})}{\langle c-a, n \rangle_{\mathbb{R}^3}^2} \right].\end{aligned}$$

- Si $\langle b-a, n \rangle_{\mathbb{R}^3} \neq 0$ et $\langle c-a, n \rangle_{\mathbb{R}^3} = 0$,

$$\begin{aligned}\int_f \langle \phi_{n,d}(x), N_f \rangle_{\mathbb{R}^3} dx &= \widetilde{C}_d \left[\frac{-\sin(\langle a, n \rangle_{\mathbb{R}^3})}{\langle b-a, n \rangle_{\mathbb{R}^3}} + \frac{\cos(\langle a, n \rangle_{\mathbb{R}^3})}{\langle b-a, n \rangle_{\mathbb{R}^3}^2} + \frac{-\cos(\langle b, n \rangle_{\mathbb{R}^3})}{\langle b-a, n \rangle_{\mathbb{R}^3}^2} \right], \\ \text{et } \int_f \langle \psi_{n,d}(x), N_f \rangle_{\mathbb{R}^3} dx &= \widetilde{C}_d \left[\frac{\cos(\langle a, n \rangle_{\mathbb{R}^3})}{\langle b-a, n \rangle_{\mathbb{R}^3}} + \frac{\sin(\langle a, n \rangle_{\mathbb{R}^3})}{\langle b-a, n \rangle_{\mathbb{R}^3}^2} + \frac{-\sin(\langle b, n \rangle_{\mathbb{R}^3})}{\langle b-a, n \rangle_{\mathbb{R}^3}^2} \right].\end{aligned}$$

- Si $\langle b-a, n \rangle_{\mathbb{R}^3} \neq 0$, $\langle c-a, n \rangle_{\mathbb{R}^3} \neq 0$ et $\langle b-c, n \rangle_{\mathbb{R}^3} = 0$,

$$\begin{aligned}\int_f \langle \phi_{n,d}(x), N_f \rangle_{\mathbb{R}^3} dx &= \widetilde{C}_d \left[\frac{\sin(\langle c, n \rangle_{\mathbb{R}^3})}{\langle c-a, n \rangle_{\mathbb{R}^3}} + \frac{\cos(\langle b, n \rangle_{\mathbb{R}^3})}{\langle c-a, n \rangle_{\mathbb{R}^3} \langle b-a, n \rangle_{\mathbb{R}^3}} + \frac{-\cos(\langle a, n \rangle_{\mathbb{R}^3})}{\langle c-a, n \rangle_{\mathbb{R}^3} \langle b-a, n \rangle_{\mathbb{R}^3}} \right], \\ \text{et } \int_f \langle \psi_{n,d}(x), N_f \rangle_{\mathbb{R}^3} dx &= \widetilde{C}_d \left[\frac{-\cos(\langle c, n \rangle_{\mathbb{R}^3})}{\langle c-a, n \rangle_{\mathbb{R}^3}} + \frac{\sin(\langle c, n \rangle_{\mathbb{R}^3})}{\langle c-a, n \rangle_{\mathbb{R}^3} \langle b-a, n \rangle_{\mathbb{R}^3}} + \frac{-\sin(\langle a, n \rangle_{\mathbb{R}^3})}{\langle c-a, n \rangle_{\mathbb{R}^3} \langle b-a, n \rangle_{\mathbb{R}^3}} \right].\end{aligned}$$

- Si $\langle b - a, n \rangle_{\mathbb{R}^3} \neq 0$, $\langle c - a, n \rangle_{\mathbb{R}^3} \neq 0$ et $\langle b - c, n \rangle_{\mathbb{R}^3} \neq 0$,

$$\int_f \langle \phi_{n,d}(x), N_f \rangle_{\mathbb{R}^3} dx = \widetilde{C}_d \left[\frac{\cos(\langle b, n \rangle_{\mathbb{R}^3})}{\langle c - b, n \rangle_{\mathbb{R}^3} \langle b - a, n \rangle_{\mathbb{R}^3}} + \frac{\cos(\langle c, n \rangle_{\mathbb{R}^3})}{\langle b - c, n \rangle_{\mathbb{R}^3} \langle c - a, n \rangle_{\mathbb{R}^3}} + \frac{-\cos(\langle a, n \rangle_{\mathbb{R}^3})}{\langle b - a, n \rangle_{\mathbb{R}^3} \langle c - a, n \rangle_{\mathbb{R}^3}} \right],$$

$$\text{et } \int_f \langle \psi_{n,d}(x), N_f \rangle_{\mathbb{R}^3} dx = \widetilde{C}_d \left[\frac{\sin(\langle b, n \rangle_{\mathbb{R}^3})}{\langle c - b, n \rangle_{\mathbb{R}^3} \langle b - a, n \rangle_{\mathbb{R}^3}} + \frac{\sin(\langle c, n \rangle_{\mathbb{R}^3})}{\langle b - c, n \rangle_{\mathbb{R}^3} \langle c - a, n \rangle_{\mathbb{R}^3}} + \frac{-\sin(\langle a, n \rangle_{\mathbb{R}^3})}{\langle b - a, n \rangle_{\mathbb{R}^3} \langle c - a, n \rangle_{\mathbb{R}^3}} \right].$$

Annexe B

Éléments sur la notion de surface moyenne

B.1 Définitions

B.1.1 Dans un cadre déterministe

Comme cela a été évoqué précédemment, la possibilité de définir et estimer l'élément moyen d'un ensemble de surfaces pourrait s'avérer d'un grand intérêt pratique dans le domaine de l'imagerie 3-D, en particulier dans les applications à la biologie et à la médecine : recherche et visualisation d'un caractère moyen au sein d'une population d'organes, recherche d'une forme-modèle associée à un phénotype donné, construction d'un registre de formes permettant d'indexer les différents phénotypes au sein d'une population. . . La moyenne n'ayant usuellement de sens que dans le cadre d'un espace vectoriel, il convient de se donner une définition originale pour adapter cette notion aux ensembles de surfaces en tenant compte de la spécificité de ces objets. On va, à cet effet, utiliser une inversion de la fonction Ψ définie en Section 8.2.1.

Définition 20 (Moyenne d'un ensemble de surfaces). *Soient S_1, \dots, S_K des surfaces et s_1, \dots, s_K les représentants associés dans l'espace V^* , tel que définis en Section 3.2.1. Soit $\overline{s_K}$ la moyenne algébrique des éléments s_1, \dots, s_K au sens de l'espace vectoriel V^* . On appelle moyenne des surfaces S_1, \dots, S_K toute surface $\overline{S_K}$ telle que $\Psi(\overline{S_K}) = \overline{s_K}$.*

Remarque 9. *Dans la configuration donnée en Section 9.1, la fonction Ψ est injective mais non surjective. De ce fait, la moyenne d'un échantillon de surfaces est unique si elle existe, mais n'existe pas nécessairement.*

Remarque 10. *On définirait de façon analogue, au vu de la section 9.1.3, la moyenne d'un ensemble de surfaces discrètes. La fonction $\widetilde{\Psi}_\rho$ étant elle aussi injective et non surjective, cette moyenne serait elle aussi unique, mais pas nécessairement définie.*

Remarque 11. *Dans le cas général, il n'est pas trivial de reconstruire la surface moyenne d'un échantillon de surfaces à partir du représentant moyen ; cela nécessitera généralement la résolution d'un problème inverse mal posé et a priori impossible à traiter via les méthodes usuelles, définies pour inverser des applications reliant deux espaces vectoriels.*

B.1.2 Dans un cadre probabiliste, moyenne théorique et moyenne empirique

On se place maintenant dans le cas où les représentants s_1, \dots, s_K sont vus comme des réalisations de formes linéaires aléatoires S_1, \dots, S_K de $V_{\mathcal{E}}^*$, décrites par le modèle de variation défini en Section 3.2.3, via l'Équation (3.2), à titre de rappel :

$$S_k = \mu + B_k, \forall k \in \{1, \dots, K\},$$

où $\mu \in V^*$ et $B_k, 1 \leq k \leq K$, sont des formes linéaires aléatoires gaussiennes centrées.

Définition 21 (Moyenne empirique). *On appelle moyenne empirique de la suite (S_1, \dots, S_K) toute moyenne algébrique de cette dernière, telle que donnée par la Définition 20, c'est à dire tout antécédent par la fonction Ψ de la moyenne dans V^* de la suite $(s_k = S_k(\omega))_{1 \leq k \leq K}$.*

Définition 22 (Moyenne théorique). *On appelle moyenne théorique de la suite (S_1, \dots, S_K) tout antécédent de μ par la fonction Ψ .*

Remarque 12. *Il va de soi que moyenne théorique et moyenne empirique n'existent pas nécessairement mais sont uniques en cas d'existence.*

Remarque 13. *On définirait de façon analogue moyenne empirique et moyenne théorique d'un échantillon de surfaces discrétisées, dont les représentants sont décrits par le modèle d'observation défini en Section 3.3 via l'Équation (3.4).*

B.2 Éléments sur l'estimation de la surface moyenne

B.2.1 Objectifs et notations

L'objectif, dans ce chapitre, est de développer une procédure pour estimer la moyenne d'un échantillon de surfaces, à partir du représentant moyen dans V^* , calculé de façon approximée via l'estimateur construit en Section 4.7. Dans la suite, on se donne un ensemble de surfaces de Θ , noté (S_1, \dots, S_K) , dont les représentants dans V^* sont décrits par le modèle de variation (3.2) donné en Section 3.2.3; on suppose que cet échantillon admet une surface moyenne théorique notée \bar{S} .

Dans la suite, on suppose que l'on connaît un échantillon $(F_1^\rho, \dots, F_K^\rho)$, formé d'approximations discrètes respectives des surfaces lisses (S_1, \dots, S_K) , ainsi que les représentants $S_{1,\rho}, \dots, S_{K,\rho}$ associés qui, de fait, suivent le modèle d'observation (3.4). On note $\widetilde{\mu_{K,\rho}}$ le représentant moyen estimé pour cet échantillon discret, calculable à partir des $S_{1,\rho}, \dots, S_{K,\rho}$ et donc également connu. On va, plus loin, construire une méthode pour construire une approximation de la moyenne \bar{S} à partir de ces données.

B.2.2 Estimation par la méthode du plus proche représentant

Algorithme

La méthode que nous introduisons pour estimer la surface moyenne de l'échantillon ne passe pas par l'inversion de la fonction Ψ . Elle est basée sur l'approximation de cette surface moyenne par l'élément de l'échantillon de surfaces discrètes dont le représentant dans V^* est le plus proche du représentant moyen estimé, au sens de la distance euclidienne dans l'espace V^* .

- on détermine l'indice k^* défini par

$$k^* = \operatorname{argmin}_{1 \leq k \leq K} \|\widetilde{\mu}_{K,\rho} - S_k\|_{V^*};$$

- On approxime la surface moyenne \bar{S} de l'échantillon $(S_k)_{1 \leq k \leq K}$ par la surface discrète

$$\widetilde{S}_{K,\rho} := F_{k^*,\rho}.$$

Avantages et inconvénients

Le principal point fort de cette méthode réside dans sa rapidité d'exécution et sa simplicité, essentiellement dues au fait d'éviter le problème de l'inversion de la fonction Ψ . Dans une optique d'application, il existe un avantage inhérent au fait de modéliser la moyenne par un élément de l'échantillon : si on prend le parti d'interpréter la moyenne comme un élément représentatif de la population étudiée, il peut être intéressant de la visualiser via un individu existant plutôt qu'un élément artificiel car reconstruit.

Le revers de la médaille du caractère somme toute fruste de la procédure est un manque de précision, potentiellement assez important : on ne dispose en effet que des surfaces de l'échantillon pour approximer la moyenne, ce qui peut entraîner une erreur considérable si l'échantillon n'admet aucun élément proche de la moyenne, au sens de la distance entre surfaces. De ce fait, il pourrait être intéressant, à l'avenir, de mettre en place une méthode de reconstruction de la surface moyenne, en résolvant le problème inverse

$$\Psi(S) = s,$$

où s est connu et S inconnu.