HAL
open science

Inférence non paramétrique pour les modèles Gibbsiens de processus ponctuels spatiaux

Nadia Morsli

To cite this version:

Nadia Morsli. Inférence non paramétrique pour les modèles Gibbsiens de processus ponctuels spatiaux. Statistiques [math.ST]. Université de Grenoble, 2014. Français. NNT : 2014GRENM055 . tel-01553115

HAL Id: tel-01553115

https://theses.hal.science/tel-01553115

Submitted on 3 Jul 2017

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.

UNIVERSITÉ DE GRENOBLE

THÈSE

Pour obtenir le grade de

DOCTEUR DE L'UNIVERSITÉ DE GRENOBLE

Spécialité : **Mathématiques Appliquées**

Arrêté ministériel : 7 août 2006

Présentée par

Nadia MORSLI

Thèse dirigée par **Jean-François COEURJOLLY**

préparée au sein du **Laboratoire Jean Kuntzmann**
et de **Ecole Doctorale Mathématiques, Sciences et Technologies de l'Information, Informatique**

Inférence non paramétrique pour les modèles gibbsiens de processus ponctuels spatiaux

Thèse soutenue publiquement le **28 novembre 2014**,
devant le jury composé de :

Mme Patricia REYNAUD-BOURET
Directrice de CNRS, Université de Nice, Rapporteur
Mr Rachid SENOUSSI
Directeur de recherche, Université d'Avignon, Rapporteur
Mme Sana LOUHICHI
Professeur, Université de Grenoble, Présidente
Mr Frédéric LAVANCIER
Maître de conférence, Université de Nantes, Examinateur
Mr Jean-François COEURJOLLY
Maître de conférence, Université de Grenoble, Directeur de thèse

En faisant scintiller notre lumière, nous offrons aux autres la possibilité d'en faire autant.
Nelson Mandela

Remerciements

Cette thèse n'aurait jamais pu voir le jour sans le soutien d'un grand nombre de personnes que je tiens à remercier pour leurs encouragements et leurs valeurs humaines.

Je ne remercierai jamais assez mes enseignants du Magister de l'Université Djillali Liabès à Sidi Bel Abbès (Algérie), pour tout ce qu'ils ont fait pour moi au cours de mes années d'études pour leurs soutiens et leurs encouragements. J'ai pris un très grand plaisir à travailler avec eux et j'ai été extrêmement sensible à leurs qualités humaines.

Je remercie Monsieur Jean-François Coeurjolly d'avoir accepté d'être mon directeur de thèse et d'avoir cru en moi et de la confiance qu'il m'a accordée. Je voudrais exprimer mes remerciements les plus vifs et les plus sincères à Monsieur Rachid Senoussi, Directeur de recherche à l'Université d'Avignon, rapporteur de mon travail de thèse de sa lecture minutieuse et de juger mon manuscrit de thèse et je lui suis reconnaissante pour sa participation constructive lors de la préparation de ma soutenance. J'adresse également mes remerciements à Madame Patricia Reynaud-Bouret, Directrice de CNRS à l'Université de Nice, rapporteur de mon travail de thèse, pour m'avoir fait l'honneur d'être présente à ma soutenance. Je tiens à remercier Madame Sana Louichi, Professeur à l'Université de Grenoble, pour m'avoir fait l'honneur d'être la présidente de mon jury, pour ses remarques et commentaires très utiles lors de la préparation de la soutenance ainsi pour ses encouragements continus et conseils pertinents. Je souhaite témoigner toute ma gratitude à Monsieur Frédéric Lavancier, Maître de conférence à l'Université de Nantes pour sa participation, en tant qu'examinateur à ma soutenance de thèse et pour ses questions intéressantes lors de la préparation de la soutenance.

Je tiens à exprimer ma profonde reconnaissance à Eric Bonnetier, Stéphane Labbé et Olivier Gaudoin pour m'avoir offert les meilleures conditions de travail possibles pour achever mon travail de thèse et pour tout ce qu'ils font pour le bien-être des doctorants. Pour leurs générosités, leurs écoutes et leurs hautes compétences humaines ont été des qualités inestimables.

Je souhaite également remercier le personnel administratif du laboratoire LJK pour leur gentillesse, leur présence, leur écoute, leur soutien, ainsi que pour les instants festifs lors des différentes réjouissances du laboratoire.

Je tiens à remercier également chaque membre du laboratoire LJK, permanents ou doctorants. Rencontrer au hasard des couloirs, ou lors des séminaires, ils ont toujours

eu la gentillesse de me demander les nouvelles de l'avancement de mon travail de thèse, ou tout simplement de prendre de mes nouvelles.

Un grand merci également à l'Ecole des Pupilles de l'Air à Montbonnot, Ministère de la Défense pour avoir financé ma thèse.

Mes remerciements vont également aux membres de la bibliothèque de l'institut Joseph Fourier pour leur amabilité et disponibilité. Je tiens à remercier mes amis qui m'ont soutenu pendant cette thèse, même quelques minutes avant le début de ma soutenance ils aidaient à préparer le pot.

Une douce pensée à tous mes anciens collègues aux universités de Sidi Bel Abbès. J'ai eu la chance de rencontrer ces personnes formidables, sur qui j'ai pu compter et qui m'ont soutenus. Je tiens à les remercier.

Je tiens à exprimer toute mon affection à ma famille, qui m'a soutenue dans tous les moments. Pour toute sa volonté de vouloir me pousser le plus loin possible dans mes études.

Résumé

Parmi les modèles permettant d'introduire de l'interaction entre les points, nous trouvons très large famille des modèles gibbsiens de processus ponctuels spatiaux issus de la physique statistique, permettant de modéliser à la fois des motifs répulsifs ou attractifs. Dans cette thèse, nous nous intéressons à l'inférence semi-paramétrique de ces modèles caractérisés par l'intensité conditionnelle de Papangelou. Deux contextes sont étudiés. Dans le premier, nous décrivons une procédure d'estimation du terme d'interaction du premier ordre (qui peut être aussi appelé l'intensité de Poisson) de l'intensité conditionnelle de Papangelou. L'idée sur laquelle l'estimation est basée permet, sous l'hypothèse d'une portée finie, de négliger les termes d'interaction d'ordre supérieur quelle que soit leur nature. La consistance forte et la normalité asymptotique de l'estimateur sont prouvées. Une étude par simulations illustre la performance de l'estimateur sur une fenêtre d'observation finie. Dans le second, nous nous focalisons sur la classe la plus connue et utilisée ; le processus ponctuel à interaction par paires. Nous construisons une nouvelle méthode d'estimation de la fonction d'interaction de paires dans l'esprit des estimations non paramétriques par lissage à partir d'une réalisation du processus ponctuel spatial à interaction par paires. Deux cas sont étudiés : le cas stationnaire et le cas isotrope. Ces estimateurs exploitent à nouveau la propriété de portée finie des processus ponctuels et intègrent l'estimation du paramètre de l'intensité de Poisson vue dans le premier thème. Nous présentons les propriétés asymptotiques telles que la consistance forte ponctuelle, la consistance forte globale avec différentes vitesses de consistance, le comportement de l'erreur quadratique moyenne et la normalité asymptotique de ces estimateurs.

Mots clés : Processus ponctuel spatial ; estimation non paramétrique ; formule de Georgii-Nguyen-Zessin ; processus de Gibbs ; propriétés asymptotiques.

Abstract

Among models allowing to introduce interaction between points, we find the large class of Gibbs models coming from statistical physics. Such models can produce repulsive as well as attractive point pattern. In this thesis, we are interested in the semi-parametric inference of such models characterized by the Papangelou conditional intensity. Two frameworks are considered. First, we describe a procédure which intends to estimate the first-order interaction term (also called Poisson intensity) of the Papangelou conditional intensity. Under the assumption of finite range of the process, the idea upon which the procedure is based allows us to neglect higher-order interaction terms. We study the stong consistency and the asymptotic normality and conduct a simulation study which highlights the efficiency of the method for finite observation window. Second, we focus on the main class of Gibbs models which is the class of pairwise interaction point processes. We construct a kernel-based estimator of the pairwise interaction function. Two cases are studied : the stationary case and the isotropic case. The estimators, we propose, exploit the finite range property and the estimator of the Poisson intensity defined in the first part. We present asymptotic properties, namely the strong consistency, the behavior of the mean squared error and the asymptotic normality.

Keywords : Spatial point process ; non parametric estimation ; Georgii-Nguyen-Zessin formula ; Gibbs process ; asymptotic properties.

Liste des travaux

Publications

J.-F. Coeurjolly and N. Morsli. Poisson intensity parameter estimation for stationary Gibbs point processes. Spatial Statistics, 4 :45–56, 2013.

Communications

1-N.Morsli, J.-F. Coeurjolly. Semi-parametric estimation of the Poisson intensity parameter for stationary Gibbs point processes. In 9th French-Danish Workshop "Spatial Statistics and Image Analysis in Biology", SSIAB. Avignon, France, mai 2012.

2-N.Morsli, J.-F. Coeurjolly. Estimation de l'interaction du premier ordre d'un processus ponctuel spatial d'interaction de paires de portée finie. 45e journée de Statistique de la SFds, Toulouse,France, 2013.

3-N.Morsli. Estimation semi-paramétrique de l'intensité de Poisson d'un processus ponctuel à interaction par paires stationaire et isotrope. 5èmes Rencontres des Jeunes Statisticiens, Aussois, 2013.

Table des matières

Notations **1**

Abréviations **3**

Introduction **5**

1 Processus ponctuels spatiaux **19**
1.1 Définitions et notations 20
1.2 Processus ponctuel de Poisson 23
1.3 Processus de Cox 24
1.3.1 Processus ponctuel de Neymann-Scott 24
1.3.2 Processus ponctuel de Cox log-Gaussien 25
1.4 Processus ponctuel de Gibbs 28
1.4.1 Processus ponctuel de Gibbs dans une région bornée 28
1.4.2 Intensité conditionnelle de Papangelou 28
1.4.3 Processus ponctuels de Gibbs stationnaires sur $\mathbb{R}^d$ 30
1.4.4 Exemples des processus ponctuels de Gibbs dans $\mathbb{R}^d$ 31
1.5 Quelques théorèmes limites 35
1.5.1 Ergodicité et théorème ergodique 35
1.5.2 Théorème central limite pour une fonctionnelle d'un champ aléatoire markovien 36

2 Estimation du paramètre d'intensité de Poisson d'un processus ponctuel de Gibbs stationnaire de portée finie **39**
2.1 Principales hypothèses et discussion 40
2.2 Présentation de l'estimateur de l'intensité de Poisson 42

2.3 Résultats asymptotiques 42
2.3.1 Preuve de la Proposition 2.3.1 44
2.3.2 Preuve du Théorème 2.3.2 45
2.4 Etudes par simulation 49

3 Estimation non paramétrique de fonction d'interaction de paires d'un processus ponctuel à interaction par paires stationnaire 57
3.1 Hypothèses et présentation de l'estimateur 58
3.2 Comportements asymptotiques de la moyenne et de la variance de l'estimateur $\widehat{H}_n(t)$ 60
3.2.1 Preuve de la Proposition 3.2.1 62
3.2.2 Preuve du Théorème 3.2.2 63
3.3 Consistance ponctuelle 69
3.3.1 Preuve de la Proposition 3.3.1 70
3.4 Inégalités de Kahane-Khintchine 71
3.5 Consistance forte ponctuelle 73
3.5.1 Preuve de la Proposition 3.5.2 74
3.6 Consistance forte globale 76
3.6.1 Preuve du Théorème 3.6.2 78
3.6.2 Preuve du Théorème 3.6.1 81
3.7 Normalité asymptotique 81
3.7.1 Preuve du Théorème 3.7.4 85
3.7.2 Preuve du Corollaire 3.7.5 89
3.7.3 Preuve du Théorème 3.7.3 89

4 Estimation non paramétrique de fonction d'interaction de paires d'un processus ponctuel à interaction par paires stationnaire et isotrope 91
4.1 Hypothèses et présentation de l'estimateur 92
4.2 Comportements asymptotiques de la moyenne et de la variance de l'estimateur $\widehat{R}_n(r)$ 94
4.2.1 Preuve de la Proposition 4.2.1 95
4.2.2 Preuve du Théorème 4.2.2 96
4.3 Consistance ponctuelle 100
4.4 Consistance forte ponctuelle 101
4.5 Consistance forte globale 102
4.5.1 Preuve de la Proposition 4.4.1 104
4.5.2 Preuve du Théorème 4.5.1 105
4.6 Normalité asymptotique 108

5 Estimation non paramétrique de fonction d'interaction de paires d'un processus ponctuel à interaction par paires pour les champs mélangeants 111
5.1 Coefficients de mélange 112
5.2 Résultats asymptotiques de l'estimateur $\widehat{G}_n(t)$ 113
5.2.1 Consistance forte ponctuelle 114
5.2.2 Consistance forte globale 115
5.2.3 Normalité asymptotique 116
5.3 Résultats asymptotiques de l'estimateur $\widehat{\Phi}_n(r)$ 117
5.3.1 Consistance forte ponctuelle 118
5.3.2 Consistance forte globale 119
5.3.3 Normalité asymptotique 120
5.4 Preuves 121
5.4.1 Preuve de la Proposition 5.2.2 121
5.4.2 Preuve du Théorème 5.3.2 124
5.4.3 Preuve du Théorème 5.2.4 124

Conclusion et Perspectives 127

Bibliographie 131

Notations

$\mathring{A}$	Intérieur d'un ensemble A.
$\sum^{\neq}$	Somme de points distincts.
Λ^c	Complémentaire de l'ensemble Λ à l'intérieur de $\mathbb{R}^d$.
$B(u,r)$	Boule fermée dans $\mathbb{R}^d$ de centre u et de rayon r.
$B^0(u,r)$	Boule ouverte dans $\mathbb{R}^d$ de centre u et de rayon r.
∂A	Frontière d'un ensemble A.
$\mathbb{1}$	Fonction indicatrice.
σ_d	Aire de surface de boule d'unité $d-$ dimensionnelle.
$\mathbb{S}^{d-1}$	Sphère unité $(d-1)$ dimentionnelle.
o	Origine dans $\mathbb{R}^d$, i.e. $o = (0, ..., 0) \in \mathbb{R}^d$.
$d(u, A)$	Distance entre le point u et l'ensemble A.
$\sharp$	Cardinal d'un ensemble A.
$\mathbb{L}^p$	Pour $p \geq 1$, l'espace des v.a.r X telles que $\mathrm{E}\,\lvert X\rvert^p < \infty$.
$\lvert\lvert X\rvert\rvert_p$	Pour $p \geq 1$, la norme usuelle sur l'espace $\mathbb{L}^p$.
$\mathbb{L}^\infty$	L'espace de variable bornée p.s.
$\lvert\lvert.\rvert\rvert_\infty$	Norme infinie (ou norme sup).

$\mathbb{L}_\psi$	L'espace d'Orlicz associé à une fonctionelle convexe ψ.
$\|\| \cdot \|\|_\psi$	Norme de Luxembourg définie sur L_ψ.
$\mathcal{B}^d$	$\sigma-$ Algébre des boréliens de $\mathbb{R}^d$.
$\mathcal{B}_b^d$	Classe des ensembles des Boréliens bornées de $\mathbb{R}^d$.
$\sigma(X)$	Tribu engendrée par la variable aléatoire X.
τ_x	Opérateur de translation par le vecteur $x \in \mathbb{R}^d$.
Ω	Espace des configurations localement finies.
$\ominus$	Soustraction de Minkowski.

Abréviations

p.s.	Presque sûrement.
p	Probabilité.
v.a.r	Variable aléatoire réelle.
EQM	Erreur Quadratique Moyenne.
GNZ	Formule de Georgii-Nguyen-Zessin.
i.i.d	indépendantes et identiquement distribuées.
PP	Processus Ponctuel.
PPP	Processus Ponctuel de Poisson.
TLC	Théorème Central Limite.
Fig.	Figure.
ssi	si et seulement si.
resp.	respectivement.

Soient $(a_n)_{n\geq 1}$ et $(b_n)_{n\geq 1}$ deux suites réelles.

$a_n = o(b_n), n \to \infty$	si $\lim_{n\to\infty} \frac{\lvert a_n\rvert}{\lvert b_n\rvert} = 0$ p.s.
$a_n = \mathcal{O}(b_n), n \to \infty$	$\exists c > 0$, tel que $\limsup\limits_{n\to\infty} \frac{\lvert a_n\rvert}{\lvert b_n\rvert} \leq c$ p.s.

Introduction

Un processus ponctuel spatial défini sur $\mathbb{R}^d$ (ou un sous-espace) est un nombre aléatoire de points distribués de manière aléatoire sur cet espace. Les processus ponctuels spatiaux étudient des phénomènes observés sur un ensemble spatial. Ce type de processus est utilisé dans des disciplines variées (Diggle [44]), en écologie et en foresterie (répartition d'espèces végétales ; Matérn [92]), en épidémiologie spatiale (positions ponctuelles de malades ; Lawson [86]), en sciences des matériaux (modèles de porosité ; Baddeley et Turner [119]), en sismologie et géophysique (épicentres et éventuellement intensités de tremblements de terre), en astrophysique (positions d'étoiles d'une nébuleuse ; Neyman et Scott [100]). Stoyan et al. [121] traitent des sujets variés dans ces branches.

Dans cette thèse, nous nous consacrerons sur la classe des processus ponctuels de Gibbs définis sur $\mathbb{R}^d$. Les processus ponctuels de Gibbs sont fondés sur la densité de Gibbs utilisée en physique statistique. Ces processus ponctuels sont définis par une densité qui peut être écrite sous une forme énergétique et incorporant des potentiels d'interaction. Ces interactions sont assimilables à des propriétés de voisinage telles que définies dans un champ de Markov dans un cadre stochastique. Les champs de Markov sont les plus utilisés pour le traitement des images (Geman et Geman [56]) dans la mesure où ils permettent de décrire les interactions entre les pixels de façon simple et facilement implémentables sur un ordinateur. Le premier objectif que l'on se fixe dans ces modèles est l'estimation du paramètre d'intensité de Poisson, de l'intensité conditionnelle de Papangelou dans un cadre assez large où l'on suppose que la portée est finie, et ce indépendamment des termes d'interaction d'ordre supérieur. L'idée sur laquelle l'estimation est basée permet de négliger les termes d'interaction d'ordre supérieur quelle que soit leur nature. Dans le second objectif, on s'intéresse au

modèle le plus utilisé, spécifiquement adapté à la modélisation des interactions entre les points grâce à une fonction potentielle associée. On s'intéresse à l'estimation de la partie fonctionnelle de la fonction d'interaction de paires pour les processus ponctuels à interaction par paires caractérisé par son intensité conditionelle de Papangelou, obtenue avec les méthodes non paramétriques. Deux cas sont étudiés : le premier cas, on suppose que le processus ponctuel à interaction par paires est stationnaire et dans le second cas, on suppose que le processus ponctuel à interaction par paires est sationnaire et isotrope.

D'un point de vue historique, il existe dans la littérature d'autres méthodes pour estimer la fonction de potentiel dans un cadre paramétrique : maximisation d'approximation de la vraisemblance (Ogata et Tanemura [103], Ogata et Tanemura [104], Penttinen [108]), la maximisation d'une pseudo-vraisemblance (Besag et al. [11]), Jensen et Møller [78], Guyon [61], Billiot [14]) et des méthodes ad hoc (Strauss [123], Ripley [112], Hanisch et Stoyan [68], Diggle et Gratton [46], Fiksel [53] et Takacs [124], Billiot et Goulard [16]). Toutes ces méthodes sont résumées dans Billiot [14].

Par la suite, nous rappelons la définition de l'intensité conditionnelle de Papangelou et nous décrivons des méthodes pour estimer la fonction de potentiel dans un cadre paramétrique qui ne mettent en jeu que le calcul de l'intensité conditionnelle de Papangelou.

Une densité f est dite héréditaire si pour toute configuration finie de points $\mathbf{x}$,

$$f(\mathbf{x}) > 0 \text{ et pour tout } \mathbf{y} \subseteq \mathbf{x} \Longrightarrow f(\mathbf{y}) > 0.$$

Cette condition, vérifiée pour les densités usuelles, traduit que toute sous configuration d'une configuration de densité > 0 est elle-même de densité > 0.
Si f est héréditaire, l'intensité conditionnelle de Papangelou (voir Papangelou [105]) de $u \notin \mathbf{x}$ conditionnelle à $\mathbf{x}$ est définie par

$$\lambda(u, \mathbf{x}) = \frac{f(\mathbf{x} \cup \{u\})}{f(\mathbf{x})}.$$

λ présente l'avantage par rapport à f d'éliminer la constante de normalisation, c'est pourquoi son usage sera central. L'intensité de Papangelou $\lambda(u, \mathbf{x})$ peut être interprétée comme étant la densité de probabilité qu'il y ait un point de $\mathbf{X}$ en u, sachant que la réalisation de $\mathbf{X}$ est $\mathbf{x}$ ailleurs.

Estimation paramétrique

Cette partie vise à décrire quelques méthodes d'estimation paramétrique des modèles gibbsiens ; θ désigne un paramètre vectoriel à estimer, $\theta \in \mathbb{R}^p$. Cette partie est loin d'être exhaustive et nous renvoyons le lecteur à Møller et Waagepetersen [96] pour des discussions plus complètes incluant des développements autour des processus marqués, fibrés,... Par souci de simplicité ici, nous nous focalisons sur l'inférence de processus stationnaires.

Evidemment la méthode la plus standard lorsqu'il s'agit d'inférer un modèle paramétrique est la méthode du maximum de vraisemblance. Cette méthode souffre cependant de plusieurs lacunes importantes : premièrement elle est extrêmement lourde à mettre en oeuvre car elle nécessite de fixer une configuration extérieure, requiert le calcul de la fonction de partition pour toute valeur du paramètre θ, constante inconnue en règle générale ce qui implique de produire des approximations de la fonction de vraisemblance. Enfin, bien qu'il s'agisse de la méthode la plus standard, la connaissance du comportement asymptotique de l'estimateur du maximum de vraisemblance est très faible. Seul Mase [91] traite ce problème mais les résultats ne couvrent que peu de modèles.

Un processus de Gibbs héréditaire dans un borné ou dans $\mathbb{R}^d$ peut être décrit par l'intensité conditionnelle de Papangelou. Par la suite, nous décrivons des méthodes alternatives au maximum de vraisemblance qui ne mettent en jeu que le calcul de l'intensité conditionnelle de Papangelou et qui donc évitent le calcul de la fonction de partition. Nous supposons que $\mathbf{X}$ est observé sur Λ_n où $(\Lambda_n)_{n\geq 1}$ est une suite de domaines bornés dans $\mathbb{R}^d$ convergeant vers $\mathbb{R}^d$, typiquement une suite de cubes mais nous ne voulons pas rentrer dans ces considérations ici. Par la suite, l'intensité conditionnelle de Papangelou sera notée $\lambda(u, \mathbf{X}) = \lambda(u, \mathbf{X}; \theta)$.

Maximum de pseudo-vraisemblance

S'inspirant de la technique de pseudo-vraisemblance introduite par Besag [10] pour les champs latticiels, Jensen et Møller [78] ont proposé de construire le contraste basé sur le produit de densités conditionnelles calculées dans des domaines infinitésimaux. Cette fonction de contraste, appelée log-pseudo-vraisemblance s'écrit

$$\mathrm{LPL}_{\Lambda_n}(\mathbf{X}; \theta) = \sum_{u \in \mathbf{X}_{\Lambda_n}} \log \lambda(u, \mathbf{X} \setminus u; \theta) - \int_{\Lambda_n} \lambda(u, \mathbf{X}; \theta) \mathrm{d}u. \tag{0.0.1}$$

Jensen et Møller ont montré la consistance de l'estimateur du maximum de pseudo-vraisemblance lorsque λ est log-linéaire en θ (i.e. lorsque le modèle appartient à une

famille exponentielle). La normalité asymptotique est étudiée par Jensen et Künsch [77] pour quelques processus d'interaction de paires au sein de cette classe. Les résultats ont ensuite été étendus par Billiot et al. [15] pour des modèles issus d'une famille exponentielle satisfaisant une hypothèse de locale stabilité et une hypothèse de portée finie, puis par Coeurjolly et Drouilhet [28] pour des modèles non linéaires et superstables au sens de Ruelle avec portée finie. Enfin Dereudre et Lavancier [43] proposent une extension de la méthodologie et des résultats asymptotiques pour des modèles gibbsiens non héréditaires.

La popularité de la méthode du maximum de pseudo-vraisemblance a été sans aucun doute mise en lumière par le papier de Baddeley et Turner [6]. Les auteurs montrent qu'en réalisant une approximation de (0.0.1), on peut montrer que le gradient de cette approximation correspond au score d'une régression poissonienne permettant une implémentation logicielle extrêmement rapide, simple et efficace en utilisant (par exemple en `R`) la fonction `glm` estimant les modèles linéaires généralisés.

Méthode de Takacs-Fiksel

Introduite dans [125, 51] conjointement par Takacs et Fiksel, cette méthode étend la méthode de la pseudo-vraisemblance. Elle s'appuie entièrement sur la formule de Georgii-Nguyen-Zessin. Si $\theta \in \mathbb{R}^p, p \geq 1$, on se fixe $h_1, \ldots, h_q : \mathbb{R}^d \times \Omega \to \mathbb{R}$, $q \geq p$ fonctions mesurables et on construit ensuite la fonction de contraste

$$\mathrm{TF}(\mathbf{X};\theta) = \sum_{i=1}^{q} \left(\sum_{u \in \mathbf{X}_{\Lambda_n}} h_i(u, \mathbf{X} \setminus u) - \int_{\Lambda_n} h_i(u, \mathbf{X}) \lambda(u, \mathbf{X}; \theta) \mathrm{d}u \right)^2 .$$

L'idée de la méthode est que lorsque θ correspond à la vraie valeur du paramètre à estimer, le terme sous le carré a pour espérance zéro grâce à la formule de Georgii-Nguyen-Zessin (1.4.1). Si on choisit $q = p$ et $(h_1(u, \mathbf{X}), \ldots, h_p(u, \mathbf{X}))^\top = \nabla \log \lambda(u, \mathbf{X}; \theta)$, on peut montrer que la méthode de Takacs-Fiksel se réduit à celle du maximum de pseudo-vraisemblance. Les résultats asymptotiques les plus complets sont établis par Coeurjolly et al.[27] : les problèmes de consistance, de normalité asymptotique et les problèmes d'identifiabilité y sont discutés.

L'intérêt de cette méthode réside clairement dans sa flexibilité quant au choix des fonctions tests h_i. Par exemple, des choix adéquats permettent de simplifier le calcul du terme somme ou intégrale. D'autres choix permettent même d'estimer des modèles d'ensembles aléatoires marqués comme le modèle Quermass (extension gibbsienne du modèle booléen) pour lesquels la position des points n'est pas connue. Nous renvoyons le lecteur intéressé à l'article de Dereudre et al. [42].

Régression logistique

Un récent travail de Baddeley et ses coauteurs [3] s'attache à étudier l'approximation de la fonction de pseudo-vraisemblance permettant une implémentation logicielle efficace (voir ci-dessus). En particulier, les auteurs montrent que dans beaucoup de situations cette approximation peut engendrer un biais non négligeable. Comme alternative, ils considèrent la fonction de contraste suivante :

$$\mathrm{LRL}_{\Lambda_n}(\mathbf{X};\theta) = \sum_{u\in\mathbf{X}_{\Lambda_n}} \log \frac{\lambda(u,\mathbf{X}\setminus u;\theta)}{\lambda(u,\mathbf{X}\setminus u;\theta)+\rho} + \sum_{u\in\mathbf{D}_{\Lambda_n}} \log \frac{\rho}{\lambda(u,\mathbf{X};\theta)+\rho}$$

où $\mathbf{D}$ est un processus ponctuel défini sur $\mathbb{R}^d$, indépendant de $\mathbf{X}$ et d'intensité ρ. Lorsque $\rho \to \infty$, ce critère s'apparente à celui de la pseudo-vraisemblance. L'intérêt de l'estimateur issu du maximum de la fonction LRL, dont les propriétés asymptotiques sont également établis par Baddeley et al. [3] est que le score de cette fonction de contraste (qui ne nécessite aucune approximation) correspond au score d'une régression logistique, pouvant donc lui aussi être calculé de manière rapide et efficace en utilisant à nouveau en `R` la fonction `glm`.

Approche variationnelle

En s'inspirant des travaux d'Almeida et Gidas [2] pour des champs markoviens, Baddeley et Dereudre démontrent dans [4] l'équation suivante pour toute fonction $h : \mathbb{R}^d \times \Omega \to \mathbb{R}$ (suffisamment régulière et à support compact)

$$\mathrm{E}\left(\sum_{u\in\mathbf{X}_{\Lambda_n}} h(u,\mathbf{X}\setminus u)\mathrm{div}\log\lambda(u,\mathbf{X}\setminus u;\theta^\star)\right) = \mathrm{E}\left(\sum_{u\in\mathbf{X}_{\Lambda_n}} \mathrm{div}h(u,\mathbf{X})\right)$$

où $\theta^\star$ est le vrai paramètre vectoriel à estimer et où $\mathrm{div}h(u,\mathbf{x}) = \sum_{i=1}^d \frac{\partial h}{\partial u_i}(u,\mathbf{x})$ pour $\mathbf{x}\in\Omega$ et $u = (u_1,\ldots,u_d)^\top$. Basée sur cette équation appelée par les auteurs équation variationnelle, les auteurs montrent que si l'intensité conditionnelle de Papangelou est log-linéaire en θ, il est possible de construire toute une classe des estimateurs de $\theta^\star$ qui ne nécessite aucune discrétisation d'intégrales, ni aucune procédure d'optimisation puisque l'estimateur en question consiste simplement en la solution d'un système d'équations linéaires. Nous renvoyons le lecteur intéressé pour plus de détails à l'article de Baddeley et Dereudre [4] où en particulier sont établies des propriétés de consistance et de normalité asymptotique. Une lacune de cet estimateur est qu'il ne permet pas d'estimer le premier terme d'interaction (appelé aussi terme d'interaction de singletons). Le présent manuscrit répond en partie à cette lacune car nous proposons justement dans le Chapitre 2, un estimateur de ce terme indépendamment des termes d'interaction d'ordre supérieur. Les résultats du Chapitre 2 donc complètent en quelque sorte le travail de Baddeley et Dereudre [4].

Estimation non paramétrique

A notre connaissance, une seule tentative pour résoudre le problème d'estimation non-paramétrique de la fonction d'interaction de paires à base de noyaux a été faite par Diggle et al. [45], malgré le fait que les méthodes proposées soient performantes sur des échantillons simulés, cependant l'article ne contient aucun résultat théorique. Les auteurs considèrent le modèle de densité f défini par l'intermédiaire de la fonction d'interaction Φ définie par :

$$f(s_1, ..., s_n) = \frac{Z}{n!} \prod_{i<j} \Phi(||s_i - s_j||)$$

où Z est une constante de normalisation.
Leur méthode utilise une équation intégrale sur $\mathbb{R}^2$ appellée équation de Ornstein-Zernike :

$$c(r) = h(r) - \lambda(c * h)(r) \tag{0.0.2}$$

où $h(r) = g(r) - 1$ et $*$ désigne le produit de convolution sur $\mathbb{R}^2$. Ainsi qu'une approximation de la physique statistique due à Percus et Yevick (1969) :

$$\Phi(r) \backsimeq \frac{g(r)}{g(r) - c(r)}$$

et c solution de l'équation (0.0.2), appelé la fonction de corrélation directe.

Les auteurs estiment la fonction de correlation de paires g par $\widehat{g}$ obtenue en lissant les différences de l'estimateur de la fonction K Ripley (Ripley [112]). La résolution de l'équation intégrale pourra s'opérer de diverses manières, la plus performante selon Diggle et al. (1987) utilise la transformation de Fourier noté f^T pour une fonction f isotrope. L'équation (0.0.2) s'écrit alors

$$c^T(r) = h^T(r) - \lambda c^T(r) h^T(r).$$

Notre méthode d'estimation ne repose pas contrairement à la précédente sur une approximation, son originalité réside dans l'utilisation combinée de deux techniques d'estimation, l'une paramétrique et l'autre non paramétrique, sous des hypothèses aussi générales que possible. Notons aussi que la composante commune à notre travail est que nous supposons ne disposer alors que d'une seule réalisation du processus ponctuel. Cette restriction importante impose des hypothèses implicites, supplémentaires de stationnarité, d'isotropie pour rendre possible l'analyse de la structure spatiale. Nous présentons maintenant la synthèse des divers travaux en grandes lignes de mon travail de thèse.

Chapitre 1 : Processus ponctuels spatiaux

Le premier chapitre présente quelques éléments de base de la statistique des processus ponctuels spatiaux. Dans ce contexte, de nombreux travaux sont disponibles dans la littérature à titre d'exemple, nous pouvons citer les livres de Daley et Vere-Jones [34], Stoyan et al. [121], van Lieshout [87], Møller et Waagepetersen [95]. Après quelques définitions générales nous présentons quelques modèles standard avant de nous focaliser plus précisément ensuite sur la classe de processus ponctuels de Gibbs. Nous rappelons les éléments importants quant à ces processus dont nous nous servons dans les chapitres suivants. Le chapitre se termine par une collection de théorèmes limites utilisés dans le reste de ce document.

Chapitre 2 : Estimation du paramètre d'intensité de Poisson d'un processus ponctuel de Gibbs stationnaire de portée finie

Le deuxième chapitre se résume de la manière suivante. A partir d'une réalisation d'un processus ponctuel de Gibbs stationnaire observée sur un domaine borné de $\mathbb{R}^d$ ($d \geq 1$), nous supposons que l'intensité conditionnelle de Papangelou se décompose comme

$$\lambda_\beta(u, \mathbf{x}) = \beta\, \widetilde{\lambda}(u, \mathbf{x}), \tag{0.0.3}$$

où $\beta > 0$ est le paramètre de l'intensité de Poisson et $\widetilde{\lambda}$ est une fonction représentant le terme d'interaction d'ordre supérieur non spécifié. Dans un tel modèle, nous décrirons la procédure d'estimation de l'intensité de Poisson indépendamment de la fonction $\widetilde{\lambda}$. L'idée maîtresse sur laquelle l'estimation est basée permet, de négliger les termes d'interaction d'ordre supérieur.

Nous supposerons seulement que le modèle défini en (0.0.3) admet une portée d'interaction finie R, i.e.

$$\lambda_\beta(u, \mathbf{x}) = \lambda_\beta(u, \mathbf{x}_{B(u,R)})$$

pour tout $u \in \mathbb{R}^d$, $\mathbf{x} \in \Omega$ (Ω : espace des configurations localement finies). Nous soulignons également que dans ce qui suit nous ne supposons pas que la portée d'interaction R est connue, mais nous supposons simplement qu'une borne supérieure notée $\widetilde{R}$ est connue.

Supposons que nous observons la réalisation d'un processus ponctuel de Gibbs stationnaire $\mathbf{X}$ dans un domaine borné de $\mathbb{R}^d$ noté Λ_n et supposons que la suite $(\Lambda_n)_{n\geq 1}$ tend vers $\mathbb{R}^d$ quand n tend vers l'infini. Les points situés près du bord de la

zone d'étude posent problème car une partie de boule dans laquelle nous comptons les voisins se trouve hors de la zone d'étude. C'est un problème très général, connu sous le nom d'effet de bord, qui se rencontre dans beaucoup de méthodes d'analyse de la structure spatiale. Plusieurs auteurs ont fait le bilan des méthodes de correction des effets de bord (voir par exemple Ripley [111], Biber [13], Kelly et Ripley [82], Møller et Waagepetersen [95] et Tukey [130]). Ce problème peut par exemple se résoudre par l'introduction du $\widetilde{R}$-intérieur de Λ_n défini par

$$\Lambda_{n,\widetilde{R}} = \{u \in \Lambda_n : B(u, \widetilde{R}) \subseteq \Lambda_n\}.$$

En utilisant la formule de Georgii-Nguyen-Zessin (1.4.1) appliquée à la fonction mesurable non négative h définie par : pour tout $u \in \mathbb{R}^d$, $\mathbf{x} \in \Omega$ et $\widetilde{R} \geq R$,

$$h(u, \mathbf{x}) = \mathbb{1}(\mathbf{x} \cap B(u, \widetilde{R}) = \emptyset) \tag{0.0.4}$$

nous obtenons un estimateur naturel du paramètre $\beta^\star$ le véritable paramètre, noté $\hat{\beta}_n$ est défini par

$$\widehat{\beta}_n = \widehat{\beta}_n(\mathbf{X}; \widetilde{R}) = \frac{N_{\Lambda_{n,\widetilde{R}}}(\mathbf{X}; \widetilde{R})}{V_{\Lambda_{n,\widetilde{R}}}(\mathbf{X}; \widetilde{R})} \tag{0.0.5}$$

où

$$N_\Lambda(\mathbf{X}; \widetilde{R}) = \sum_{u \in \mathbf{X}_\Lambda} h(u, \mathbf{X} \backslash u) \quad \text{et} \quad V_\Lambda(\mathbf{X}; \widetilde{R}) = \int_\Lambda h(u, \mathbf{X}) \mathrm{d}u.$$

La variable $N_\Lambda(\mathbf{X}; \widetilde{R})$ correspond au nombre de points dans $\mathbf{X}_\Lambda$ qui sont séparés des autres points dans $\mathbf{X}$ par une distance supérieure à $\widetilde{R}$, tandis que la variable $V_\Lambda(\mathbf{X}; \widetilde{R})$ représente le volume restreint de Λ après avoir enlever un disque de rayon $\widetilde{R}$ centré en chaque point de $\mathbf{X}$.

Le théorème ergodique de Nguyen et Zessin [101] est un théorème fondamental qui fournit le résultat de la consistance forte de l'estimateur du paramètre de l'intensité de Poisson dès que la suite $(\Lambda_{n,\widetilde{R}})_{n \geq 0}$ tend vers $\mathbb{R}^d$.

Bolthausen [17] semble avoir obtenu l'ensemble minimal de conditions nécessaires sur le coefficient de mélange fort pour un champ stationnaire. Takahata [126] a un résultat similaire où l'hypothèse de stationnarité a été assouplie. Malheureusement, il est souvent difficile d'obtenir les limites nécessaires sur les coefficients de mélange. Certains résultats sont connus pour les champs de Gibbs dans la région d'unicité Dobrushin et ceci est exploité dans Dobrushin [47] et Georgii [57], voir aussi par exemple Guyon [62] et Heinrich [69]. En dehors de la région d'unicité Dobrushin on sait très

peu. Souvent, il y a une décomposition ergodique de la mesure, mais il ne semble y avoir aucune limite générale sur les coefficients de mélange des composantes ergodiques. L'exemple du modèle d'Ising au point critique montre que les coefficients de mélange ne sont pas en général sommable. Pour notre étude de la normalité asymptotique de l'estimateur (0.0.5), au lieu d'utiliser les limites sur les coefficients de mélange, nous exploitons une propriété qui ressemble à celle d'une séquence de centrage conditionnel (ou la différence de martingale) qui est valable pour toutes les valeurs des paramètres avec une hypothèse d'ergodicité par application le théorème central limite pour des champs markoviens établi par Jensen et Künsch [77]. L'énorme avantage de ce résultat est qu'il ne nécessite pas l'unicité de mesure de Gibbs, condition requise pour calculer les coefficients de mélange d'un modèle gibbsien et qui a pour conséquence de contraindre l'espace des paramètres (typiquement pour un modèle de Strauss β doit être suffisamment grand).

Nous soulignons encore un aspect inattendu de ce travail : l'estimateur du paramètre d'intensité de Poisson et l'estimateur de la variance asymptotique sont indépendants des termes d'interaction d'ordre supérieur. Nous avons fait une étude sur des données simulées à l'aide des modèles de Gibbs classiquement définis par leur intensité conditionnelle de Papangelou et vérifiant les conditions citées pour le modèle (0.0.3). Cette étude a montré le bon comportement de l'estimateur proposé.

Chapitre 3 : Estimation non paramétrique de fonction d'interaction de paires d'un processus ponctuel à interaction par paires stationnaire

Dans le troisième chapitre, nous avons formalisé la loi du processus ponctuel à interaction par paires stationnaire $\mathbf{X}$ de $\mathbb{R}^d$ par son intensité conditionnelle de Papangelou (des résultats précisant cette fonction sont établis dans le premier chapitre) notée $\lambda_{\beta^\star}$ et définie pour tout $u \in \mathbb{R}^d$, $\mathbf{x} \in \Omega$, avec Ω l'espace des configurations localement finies par

$$\lambda_{\beta^\star}(u, \mathbf{x}) = \beta^\star \exp\left(-\sum_{v \in \mathbf{x} \setminus u} g(v-u)\right) \tag{0.0.6}$$

avec $\beta^\star > 0$ est le vrai paramètre d'intensité de Poisson, g représente le potentiel d'interaction défini sur Ω à valeur dans $\mathbb{R}$, invariante par translation et $G = \exp(-g)$ représente la fonction d'interaction de paires. Supposons que les interactions entre points ont une portée finie R, dans le sens où l'intensité conditionnelle $\lambda(u, \mathbf{x})$ ne

dépend que de points de $\mathbf{x}$ qui se trouvent à une distance inférieure à R de l'emplacement u. Dans ce modèle semi-paramétrique (0.0.6), l'estimation de $\beta^\star$ a représenté la première étape de notre procédure dans le deuxième chapitre .

Notre objectif est consacré à l'estimation de la fonction d'interaction de paires $G = \exp(-g)$ du processus ponctuel à interaction par paires stationnaire sous des hypothèses aussi générales que possible. La connaissance de cette fonction ou le potentielle g est un outil très utilisé par les statisticiens notamment en physique statistique, où des systèmes constitués d'un très grand nombre de particules sont modélisés. La statistique fonctionnelle a occupé une place de choix dans la recherche au cours de ces dernières années, nous la trouvons dans des domaines reconnus qui ont eu le temps de faire leur preuve, comme les branches de la science économique, l'analyse d'image, la météorologie et la géostatistique. Parmi les ouvrages de référence, nous avons la bibliographie de Ramsay et Silverman [79], Bosq [21], Biau [12], Hallin et al. [67] et Tran et Yakowitz [129].

Nous supposons que nous observons la réalisation du processus ponctuel à interaction par paires stationnaire $\mathbf{X}$ dans un domaine borné de $\mathbb{R}^d$ noté Λ_n où la suite $(\Lambda_n)_{n\geq 1}$ tend vers $\mathbb{R}^d$ quand n tend vers l'infini. Nous prenons en compte les problèmes d'effet de bord, par l'introduction du $2R$-intérieur noté $\Lambda_{n,R}$ de Λ_n. Nous allons introduire un estimateur de la fonction d'interaction de paires $G = \exp(-g)$ par la méthode de lissage, cette méthode nous permet de donner un cadre général à l'estimation de la fonction d'interaction de paires. Cette technique et des exemples d'application ont été débattus par Baddeley et al. [8] et par Gaetan et Guyon [50] pour les champs des résidus lissés. Connaissant le support $T = \{t \in \mathbb{R}^d; g(t) > 0, \text{ pour } ||t|| < R\}$ de la fonction $G = \exp(-g)$ et en se donnant $t \in T$, avec la convention $c/0 = 0$, pour tout réel c, nous définissons l'estimateur de $\beta^\star G$ par,

$$\widehat{G}_n(t) = \frac{\widehat{H}_n(t)}{\widehat{L}_n(t)}, \tag{0.0.7}$$

où

$$\widehat{H}_n(t) = \frac{1}{b_n^d|\Lambda_{n,R}|} \sum_{\substack{u,v\in\mathbf{X} \\ v-u\in B(o,R)}}^{\neq} \mathbb{1}_{\Lambda_{n,R}}(u)h(u,\mathbf{X}\setminus\{u,v\})h(v,\mathbf{X}\setminus\{u,v\})K\left(\frac{v-u-t}{b_n}\right) \tag{0.0.8}$$

est un estimateur de type noyau estimant $\beta^{\star 2}\exp(-g(t))\bar{F}(o,t)$ avec $\bar{F}(o,t)$.

$$\widehat{L}_n(t) = \frac{1}{|\Lambda_{n,R}|} \sum_{u\in\mathbf{X}_{\Lambda_{n,R}}} h(u,\mathbf{X}\setminus\{u\})h(t-u,\mathbf{X}\setminus\{u\}) \tag{0.0.9}$$

est un estimateur de $\beta^\star \bar{F}(o,t)$, où $\bar{F}(o,t) = \mathbb{P}(\mathbf{X} \cap B(o, \widetilde{R}) = \emptyset, \mathbf{X} \cap B(t, \widetilde{R}) = \emptyset)$. L'intérêt principal de cette méthode est qu'elle conduit à une famille d'estimateurs, selon les fonctions tests h choisies dans (0.0.4). Cette flexibilité permet de construire des estimateurs adaptés à notre étude. K étant un noyau défini sur $\mathbb{R}^d$, borné, intégrable et d'intégrale 1. $(b_n)_{n\geq 1}$ une suite réelle appelée paramètre de lissage, tendant vers zéro à l'infini.

Nous présentons d'abord le modèle, sous des hypothèses reliées aux caractéristiques du processus ponctuel à interaction par paires stationnaire. Nous étudions les comportements asymptotiques de la moyenne et de la variance de l'estimateur à base de noyaux $\widehat{H}_n(t)$. Les calculs du biais et la variance sont purement analytique et reposent essentiellement sur la formule de Georgii-Nguyen-Zessin (GNZ) et un développement de Taylor, qui détermineront l'ordre du biais asymptotique en fonction du paramètre de lissage b_n et préciseront également la vitesse de convergence de l'erreur ponctuelle en moyenne quadratique de l'estimateur $\widehat{H}_n(t)$. La consistance forte de l'estimateur $\widehat{L}_n(t)$ est une conséquence assez directe du théorème ergodique classique pour les processus ponctuels spatiaux obtenus par Nguyen et Zessin [101]. L'utilisation de ces deux types de convergence, nous permet d'établir la consistance de l'estimateur (0.0.7). Ensuite, nous juxtaposons des résultats sur les inégalités de Kahane-Khintchine où apparaissent les espaces d'Orlicz associés aux fonctions de Young exponentielles établies par El Machkouri [88]. Pour évaluer la précision de notre estimateur non paramétrique (0.0.7), nous utiliserons dans notre étude deux critères de consistance forte ponctuelle et globale avec des vitesses de convergence, sous certaines hypothèses d'appartenance aux espaces d'Orlicz induits par des fonctions de Young. Nos résultats portent aussi grâce à des cadres particuliers les plus importants des espaces d'Orlicz ; les champs aléatoires bornés et les champs aléatoires p-intégrables. Pour ce dernier, nous utilisons l'inégalité de type Marcinkiewiez-Zygmund pour les moments d'ordre p des sommes partielles des champs aléatoires établie par Dedecker [37]. L'exploitation de résultats obtenus dans le Chapitre 3 permet de conclure l'efficacité de l'estimateur $\widehat{G}_n(t)$, l'accent est mis sur des critères de type projectif (i.e conditions impliquant les espérances conditionnelles) qui sont faciles à vérifier pour des classes importantes de champs aléatoires, comme un champ aléatoire indépendant ou pour un champ qui satisfait l'hypothèse de type centrage conditionnel ou un champ aléatoire mélangeant.

Ces critères de type projectif ont récemment inspiré de nombreux travaux dans l'étude des théorèmes limites centrales des processus stochastiques en une dimension. Voir par exemple Dedecker [36] et Wu [133]. A priori, l'extension des résultats dans une dimension à dimensions élevée n'est pas trivial et induit des changements suffisament profonds nécessitant des traitements adaptés, comme le résume Hallin et al. [66]. Cependant, dans le contexte multidimensionnel, nous utilisons un théorème

central limite proposé par Wang et Woodroofe ([132], Théorème 2) pour les tableaux triangulaires stationnaires de champs aléatoires m-dépendant. Ce résultat améliore le théorème de la limite centrale établi par Heinrich [70]. Ceci nous permet d'étudier la normalité asymptotique de l'estimateur $\widehat{G}_n(t)$.

Chapitre 4 : Estimation non paramétrique de fonction d'interaction de paires d'un processus ponctuel à interaction par paires stationnaire et isotrope

Les processus ponctuels d'interaction de paires stationnaires et isotropes introduits dans la littérature statistique par Ripley [112] comme exemple de processus ponctuels de Markov, ils se sont montrés très utiles dans la pratique par leur capacité à modéliser un large éventail de processus spatiaux. Voir par exemple Diggle et Gratton [46], Ogata et Tanemura [103], Penttinen [108]). Nous adaptons un estimateur au modèle à interaction par paires stationnaire et isotrope, nous allons esquisser quelques preuves et souligner les principales différences apportées dans le cas ou le modèle à interaction par paires est uniquement stationnaire étudié dans le chapitre précédent. Nous considèrons le modèle d'interaction de paires stationnaire et isotrope caractérisé par son intensité conditionnelle de Papangelou

$$\lambda_{\beta^\star}(u,\mathbf{x}) = \beta^\star \cdot \exp\left(-\sum_{v\in\mathbf{x}\backslash u}\gamma(||v-u||)\right),\ \forall u\in\mathbb{R}^d, \mathbf{x}\in\Omega \tag{0.0.10}$$

où $\beta^\star > 0$ est le vrai paramètre d'interaction du premier ordre. γ est le potentiel d'interaction de paires, défini sur $\mathbb{R}^+$ à valeurs dans $\mathbb{R}$. Nous supposons qu'il admet une portée d'interaction finie R, i.e.

$$\gamma(||v-u||) = 0 \quad \text{pour } ||v-u|| > R,\ \forall u\in\mathbb{R}^d, v\in\mathbf{x}\backslash u\in\Omega$$

et nous supposons que

$$\gamma(r) > 0, \quad \text{si} \quad r\in(0,R] \quad \text{pour} \quad R>0.$$

Cette dernière condition implique que les réalisations typiques seront plus ou moins régulières par rapport à un dispositif complètement aléatoire. En particulier, des potentiels négatifs (ou partiellement négatifs) peuvent conduire à des distributions irrégulières ou instables (Kelly et Ripley [81] ; Gates et Westcott [55]).

L'évaluation de l'estimateur de la fonction d'interaction de paires $\Phi = \exp(-\gamma)$, exige l'observation du processus ponctuel à interaction par paires stationnaire et isotrope noté $\mathbf{X}$ dans un domaine borné de $\mathbb{R}^d$ noté Λ_n. Supposons que la suite $(\Lambda_n)_{n\geq 1}$

tend vers $\mathbb{R}^d$ quand n tend vers l'infini. Alors l'estimateur de la fonction inconnu $\beta^\star\Phi$, corrigeant les effets de bord, est donné pour $r \in (0, R]$ par

$$\widehat{\Phi}_n(r) = \frac{\widehat{R}_n(r)}{\widehat{N}_n(r)} \tag{0.0.11}$$

où
(0.0.12)

$$\widehat{R}_n(r) = \frac{1}{b_n|\Lambda_{n,R}|\sigma_d} \sum_{\substack{u,v\in\mathbf{X} \\ ||v-u||\leq R}}^{\neq} \frac{\mathbb{1}_{\Lambda_{n,R}}(u)}{||v-u||^{d-1}} h(u, \mathbf{X}\backslash\{u,v\})h(v, \mathbf{X}\backslash\{u,v\})K_1\left(\frac{||v-u||-r}{b_n}\right)$$

est un estimateur de $\beta^{\star 2}\Phi(r)J(r)$, où $J(r) = \int_{\mathbb{S}^{d-1}} \bar{F}(o, rv)\mathrm{d}v$ et $\sigma_d = \frac{2\pi^{d/2}}{\Gamma(d/2)}$ représente la mesure de la sphère unité de $\mathbb{R}^d$.

$$\widehat{N}_n(r) = \frac{1}{|\Lambda_{n,R}|} \sum_{u\in\mathbf{X}} \mathbb{1}_{\Lambda_{n,R}}(u)h(u, \mathbf{X}\setminus\{u\})h^\star(u, \mathbf{X}\setminus\{u\}) \tag{0.0.13}$$

estimateur de $\beta^\star J(r)$. Où h une fonction mesurable non négative définie pour tout $u \in \mathbb{R}^d$, $\mathbf{x} \in \Omega$ par

$$h(u, \mathbf{x}) = \mathbb{1}\left(\inf_{v\in\mathbf{x}} ||v-u|| > R\right) = \mathbb{1}\left(d(u, \mathbf{x}) > R\right)$$

et

$$h^\star(u, \mathbf{x}) = \int_{\mathbb{S}^{d-1}} h(rv - u, \mathbf{x})\mathrm{d}v.$$

Des conditions standards sont imposées sur le noyau univarié K_1 et la fenêtre b_n, nous imposons aussi certaines hypothèses de régularité dans ce chapitre.

Dans un premier temps, nous définissons précisément la classe de processus que nous considérons. Ensuite, nous définissons un estimateur non-paramétrique d'une nouvelle caractéristique du processus ponctuel à interaction par paires stationnaire et isotrope et nous présentons le problème et son contexte.

Nous étudions les comportements asymptotiques de la moyenne et de la variance de l'estimateur $\widehat{R}_n(r)$. Les calculs du biais et variance reposent essentiellement sur la formule de Georgii-Nguyen-Zessin (GNZ) et un développement de Taylor, qui détermineront l'ordre du biais asymptotique en fonction du paramètre de lissage b_n et préciserons la vitesse de convergence de l'erreur ponctuelle en moyenne quadratique de notre estimateur non paramétrique par lissage $\widehat{R}_n(r)$. La consistance de l'estimateur $\widehat{N}_n(r)$ suit le schéma adapté à l'estimateur empirique ; le théorème ergodique obtenus par Nguyen et Zessin [101], ce qui nous permet de conclure la consistance de l'estimateur (0.0.11).

Nous appliquons des inégalités de Kahane-Khintchine au travers de l'étude de la consistance forte ponctuelle et la consistance forte globale de notre estimateur non paramétrique par lissage, avec des vitesses de convergence. Enfin, il est possible de montrer la normalité asymptotique de l'estimateur (0.0.11), via le théorème central limite proposé par Wang et Woodroofe ([132], Théorème 2).

Chapitre 5 : Estimation non paramétrique de fonction d'interaction de paires d'un processus ponctuel à interaction par paires pour les champs mélangeants

L'objectif dans ce chapitre est de reprendre les propriétés asymptotiques étudiées dans le Chapitre 3 et Chapitre 4 dans le cadre des champs mélangeants. Nous avons ainsi considéré les problèmes d'estimation fonctionnelle sous des hypothèses de mélange. De nombreux résultats dans la littérature sont basés sur ces types de conditions, nous renvoyons à Bosq [19], Delecroix [39], Heinrich et al. [72] et Peligrad [107]. Dans ce cadre Györfi et al. [63] et Bosq [20] ont donc étudié l'estimation d'une densité, de la fonction de régression, de la fonction de hasard. Hall et Hart [64], Robinson [114] et Giraitis et al. [60] ont étudié l'estimation d'une fonction de régression. D'autres part, Doukhan et Louhichi [48] ont introduit une nouvelle notion de dépendance faible qui peut être vérifiée pour des processus plus généraux.

Nous commençons par indiquer les hypothèses de dépendance usuelles. Nous avons choisi, dans ce chapitre, de modéliser la dépendance en considérant des variables ϕ-mélangeantes (ou α-mélangeantes). Il se révèlera que la classe de processus ponctuels de Gibbs sous la condition de l'unicité de Dobrushin (Dobrushin [47] ; Georgii [58] ; Guyon [62]), satisfait ces conditions de mélange. Nous établissons la convergence presque sûre ponctuelle des estimateurs $\widehat{G}_n(t)$ défini en (0.0.7) et $\widehat{\Phi}_n(r)$ défini en (0.0.11). Nous étudions également la convergence presque sûre uniforme de ces estimateurs. La recherche des telles propriétés passe par l'utilisation des inégalités de mélange de Serfling (voir Serfling [118] ou McLeish [93]). La procédure fait intervenir des inégalités exponentielles de type Kahane-Khintchine dans des espaces d'Orlicz induits par des fonctions de Young exponentielles établie par El Machkouri [88] ; elle fournit de nouveau des critères pour les champs aléatoires mélangeants. La force de ces critères est qu'elle s'applique à une grande classe de processus ponctuels. Pour prouver nos résultats sur la normalité asymptotique, nous utiliserons le théorème central limite de Coeurjolly et Lavancier [29], ainsi qu'un contrôle des coefficients de mélange. Nous terminons cette thèse par une conclusion générale et quelques perspectives de recherche.

Chapitre

1 Processus ponctuels spatiaux

Résumé

Ce chapitre est consacré à l'introduction des bases théoriques des processus ponctuels et qui sont aussi à la base de la géométrie stochastique. Quelques modèles courants de la littérature sont présentés, nous rappelons quelques résultats utilisés tout au long de cette thèse et de nombreuses notations utiles dans la suite du chapitre.

1.1 Définitions et notations

Soit S un fermé de $\mathbb{R}^d, (d \geq 1), \mathcal{B}$ (resp. $\mathcal{B}(S), \mathcal{B}_b(S)$) désigne l'ensemble des boréliens de $\mathbb{R}^d$ (resp. des boréliens de S, des boréliens bornés de S). Les sous ensembles de $\mathbb{R}^d$ notés Λ sont toujours supposés des boréliens bornés. La notation $|.|$ sera utilisée sans ambiguité pour différents objets. Pour un ensemble dénombrable $J, |J|$ représente son cardinal, pour un domaine borné Λ, $|\Lambda|$ désigne sa mesure de Lebesgue. On désigne par τ_x la translation par le vecteur $-x \in \mathbb{R}^d$. Notons que nous nous focalisons ici sur les processus ponctuels spatiaux dans $S \subseteq \mathbb{R}^d$. L'extension aux processus ponctuels marqués ou sur des espaces plus complexes comme par exemple sur le cercle ou la sphère s'effectue sans difficulté.

Configuration

Soit S un fermé de $\mathbb{R}^d$, on appelle configuration de points tout ensemble de points $\mathbf{x} = \{x_1, ..., x_n, ...\}$ où $x_i \in S$ sont des points issus d'une expérience aléatoire. On dit qu'une configuration est localement finie si elle possède au plus un nombre fini de points dans n'importe quel borélien borné de S.

Processus ponctuel

Suivant Daley et Vere-Jones [34] (cf. aussi Lieshout [87]), un processus ponctuel sur S est une application $\mathbf{X}$ d'un espace de probabilité dans l'ensemble des configurations localement finies telle que pour tout borélien borné A, le nombre de points $N(A) = N_{\mathbf{X}}(A) = n(\mathbf{x}_A)$ de points de $\mathbf{X}$ tombant dans A est une variable aléatoire.

Caractérisation de la distribution d'un processus ponctuel

La distribution fini-dimensionnelle d'un processus ponctuel $\mathbf{X}$ est la donnée pour tout $m \geq 1$ et tout m-uplet $(A_1, A_2, ..., A_m)$ de $\mathcal{B}_b(S)$ des distributions sur $\mathbb{N}^m$ de $(N(A_1), N(A_2), ..., N(A_m))$.

Si S n'est pas compacte, la loi d'un processus ponctuel est définie de la même façon, les configurations $\mathbf{x}$ (éventuellement infinies) étant finies sur tout borélien borné.

Processus ponctuel stationnaire

Un processus ponctuel sur $\mathbb{R}^d$ est stationnaire si, pour tout $u \in \mathbb{R}^d$, la loi du processus ponctuel translaté $\mathbf{X}_u = \mathbf{X} + u$ est la même que celle de $\mathbf{X}$.

Processus ponctuel isotropique

Un processus ponctuel sur $\mathbb{R}^d$ est isotropique si la loi de $\rho\mathbf{X}$, $\mathbf{X}$ après ρ-rotation, est identique à celle de $\mathbf{X}$ pour toute rotation ρ.

Moments d'un processus ponctuel

Si les moments d'ordre 1 et d'ordre deux sont à la base de l'étude des processus à valeurs réelles, ici les notions adaptées aux PP sont les mesures de moment d'ordre p, $p \geq 1$: le moment d'ordre p d'un PP est la mesure sur $(S, \mathcal{B}(S))^p$ définie sur les produits $B_1 \times ... \times B_p$ par :

$$\mu^{(p)}(B_1 \times ... \times B_p) = \mathrm{E}\left[N(B_1) \cdots N(B_p)\right].$$

Moment d'ordre 1 et intensité ρ d'un processus ponctuel

La mesure d'intensité de $\mathbf{X}$ est la mesure de moment d'ordre 1 :

$$\mu^{(1)}(B) = E \sum_{\xi \in \mathbf{X}} \mathbb{1}(\xi \in B).$$

Généralement, $\mu^{(1)}(\mathrm{d}\xi)$, la probabilité qu'il y ait un point de $\mathbf{X}$ dans le volume infinitésimal $\mathrm{d}\xi$ autour de ξ, est modélisée à partir d'une densité d'intensité $\rho(\xi)$, $\mu^{(1)}(\mathrm{d}\xi) = \rho(\xi)\mathrm{d}\xi$. Si $\mathbf{X}$ est stationnaire, $\mu^{(1)}$ est invariante par translation, donc proportionnelle à la mesure de Lebesgue ν :

$$\mu^{(1)}(B) = \rho\nu(B).$$

Le paramètre ρ est l'intensité de $\mathbf{X}$: c'est le nombre moyen de points de $\mathbf{X}$ par volume unité.

Moment factoriel

La covariance de variables de comptage va s'exprimer à partir de la mesure de moment d'ordre 2, $\mu^{(2)}(B_1 \times B_2) = E[N(B_1)N(B_2)]$. Cependant, notons $\sum_{\xi,\eta \in \mathbf{X}}^{\neq}$ la somme étendue aux sites $\xi \neq \eta$ distincts de $\mathbf{X}$, la décomposition,

$$\begin{aligned}
\mu^{(2)}(B_1 \times B_2) &= \mathrm{E}\left[\sum_{\xi \in \mathbf{X}} \mathbb{1}(\xi \in B_1) \times \sum_{\eta \in \mathbf{X}} \mathbb{1}(\eta \in B_2)\right] \\
&= \mathrm{E}\left[\sum_{\xi,\eta \in \mathbf{X}} \mathbb{1}((\xi,\eta) \in (B_1, B_2))\right] \\
&= \mathrm{E}\left[\sum_{\xi \in \mathbf{X}} \mathbb{1}(\xi \in B_1 \cap B_2)\right] + \mathrm{E}\left[\sum_{\xi,\eta \in \mathbf{X}}^{\neq} \mathbb{1}((\xi,\eta) \in (B_1, B_2))\right],
\end{aligned}$$

montre que $\mu^{(2)}$ a une composante mesure sur $\mathcal{B}(S)$ et une autre mesure sur le produit $\mathcal{B}(S)\times\mathcal{B}(S)$. Cette singularité disparaît en considérant la mesure de moment factoriel $\alpha^{(2)}$ d'ordre 2 définie sur les événements $B_1 \times B_2$ par :

$$\alpha^{(2)}(B_1 \times B_2) = E\Big[\sum_{\xi,\eta\in\mathbf{X}}^{\neq} \mathbb{1}((\xi,\eta) \in (B_1, B_2))\Big].$$

Si $u \neq v$, $\alpha^{(2)}(\mathrm{d}u, \mathrm{d}v)$ s'interprète comme la probabilité que $\mathbf{X}$ possède un point dans le volume $\mathrm{d}u$ (en u) et un point dans le volume $\mathrm{d}v$ (en v). Si $\alpha^{(2)}$ est absolument continue par rapport à la mesure de Lebesgue sur $(S, \mathcal{B}(S))^2$, sa densité $\alpha^{(2)}(u, v)$ est la densité d'intensité d'ordre deux de $\mathbf{X}$.

Corrélation de paires pondérée

L'une des questions centrales dans l'étude des processus ponctuels est de savoir si les points de la réalisation d'un processus ponctuel ont une tendance à s'attirer ou au contraire à se repousser ou enfin à ne pas s'influencer. Cette dernière hypothèse d'indépendance spatiale, que nous noterons CSR (Complete Spatial Randomness), traduit que les points se répartissent dans S indépendamment les uns des autres, mais pas nécessairement uniformément. Les processus poncuels de Poisson (PPP) couvrent exactement cette classe de processus, les PPP stationnaires correspondant à la situation où la densité ρ est constante. Sans faire d'hypothèse de stationnarité, Baddeley et al. [1] définissent une fonction $g(\xi, \eta)$, dite fonction de corrélation de paires pondérée, qui est un bon indicateur au second ordre de la dépendance spatiale : si $\rho(\xi)$ et $\rho(\eta) > 0$,

$$g(\xi, \eta) = \frac{\rho^{(2)}(\xi, \eta)}{\rho(\xi)\rho(\eta)}$$

ce qui conduit à l'interprétation :

- $g(\xi, \eta) = 1$ si les points apparaissent indépendamment (hypothèse CSR).
- $g(\xi, \eta) > 1$ traduit une attraction entre les points (covariance de paire positive).
- $g(\xi, \eta) < 1$ traduit une répulsion entre les points (covariance de paire négative).

On dira que le processus ponctuel $\mathbf{X}$ est stationnaire au second ordre pour la corrélation pondérée si : $\forall \xi, \eta \in S : g(\xi, \eta) = g(\xi - \eta)$. On verra que des PP (par exemple des PP de Cox log-Gaussien) peuvent être stationnaires au second ordre pour la corrélation pondérée sans être stationnaire au premier ordre. Cette corrélation g permettra de construire des tests d'indépendance spatiale sans faire l'hypothèse que l'intensité du PP est stationnaire.

Distance au plus proche voisin (ppv) d'un processus ponctuel

La distance $d(\xi, \mathbf{X}\backslash\xi)$ d'un point $\xi \in \mathbf{X}$ au ppv de $\mathbf{X}$ doit être distinguée de la distance $d(u, \mathbf{X})$ d'un point courant $u \in \mathbb{R}^d$ (qui n'appartient pas nécessairement à $\mathbf{X}$) au ppv de $\mathbf{X}$. Une notion importante lorsqu'on s'intéresse à des processus ponctuels spatiaux est la fonction de vide F_u, qui est la distribution de cette deuxième distance définie par

$$F_u(R) = \mathbb{P}(d(u, \mathbf{X}) \leq R) = \mathbb{P}(\mathbf{X} \cap B(u, R) \neq \emptyset), \ R > 0.$$

Dans le cas d'un processus stationnaire, F_u ne dépend pas de u et on note la fonction de vide $F_u = F$. La fonction $1 - F_u(R) = \mathbb{P}(d(u, \mathbf{X}) > R) = \mathbb{P}(\mathbf{X} \cap B(u, R) = \emptyset)$ est la probabilité que $\mathbf{X}$ n'ait aucun voisin à la distance R de u.

Classes des processus ponctuels

On distingue trois grandes classes parmis les processus ponctuels les plus utilisées dans la littérature :

- Les PP de Poisson (PPP) modélisent une répartition spatiale "au hasard". Ils sont uniquement caractérisés par leur intensité $\rho(\cdot)$, qui n'est pas nécessairement stationnaire et isotrope.
- Les PP de Cox sont des PPP conditionnels à un environnement aléatoire. Ils modélisent des répartitions spatiales moins régulières présentant par exemple des agrégats, comme pour le PP de Neyman-Scott.
- Les PP de Gibbs sont définis à partir d'une spécification conditionnelle. Ils sont utiles pour modéliser des répartitions spatiales plus régulières que celle d'un PPP, par exemple les répartitions à noyaux durs où chaque point préserve un espace vital autour de lui.

D'autres classes existent par exemple les processus ponctuels obtenus par suppression de points, par superposition de processus ponctuels, voir Møller et Waagepetersen [95], ou encore la classe des processus déterminantaux [85].

1.2 Processus ponctuel de Poisson

Le processus ponctuel de Poisson joue un rôle fondamental dans l'étude des processus ponctuels. En effet, il sert de modèle de référence, mathématiquement étudiable, pour lequel il n'existe aucune interaction entre les points. De nombreux modèles, comme les processus ponctuels de Gibbs, par exemple, sont alors construits à partir du processus ponctuel de Poisson. Dans cette section, nous nous intéressons au processus de Poisson spécifié par une fonction d'intensité.

Soit $\rho(\cdot)$ une fonction non négative localement intégrable sur $S \subseteq \mathbb{R}^d$, un processus ponctuel de Poisson d'intensité $\rho(\cdot)$ (on note PPP(ρ)) est défini par :

- Pour tout $A \in \mathcal{B}_b(S)$ de mesure $0 < \int_A \rho(\xi)\mathrm{d}\xi < \infty$, $N(A)$ suit une loi de Poisson de paramètre $\int_A \rho(\xi)\mathrm{d}\xi$.
- Conditionnellement à $N(S)$, les points de $\mathbf{x} \cap A$ sont distribués de façon i.i.d avec une densité proportionnelle à $\rho(\xi), \xi \in A$:

$$\mathbb{P}(N(A) = n) = \exp\left(- \int_A \rho(\xi)\mathrm{d}\xi \right) \frac{\left(\int_A \rho(\xi)\mathrm{d}\xi \right)^n}{n!}.$$

La fonction $\rho(\cdot)$ caractérise entièrement la distribution de $\mathbf{X}$.
Le PPP est stationnaire et isotrope d'intensité ρ, si l'intensité $\rho(\cdot)$ est constante ; à l'indépendance dans la répartition s'ajoute alors l'uniformité de la répartition spatiale.

1.3 Processus de Cox

Soit $Z = \{Z(\xi) : \xi \in S\}$ un champ aléatoire non négatif sur S localement intégrable presque sûrement, pour tout borélien B borné, $\int_B Z(\xi)\mathrm{d}\xi < \infty$. Si la distribution conditionnelle de $\mathbf{X}$ sachant Z est un processus de Poisson sur S avec fonction d'intensité Z, alors $\mathbf{X}$ est appelé un processus Cox gouverné par Z.
La fonction d'intensité est définie par

$$\rho(\xi) = \mathrm{E}\, Z(\xi)$$

et la fonction de corrélation de paire est définie par

$$g(\xi, \eta) = \frac{\mathrm{E}[Z(\xi)Z(\eta)]}{\rho(\xi)\rho(\eta)}.$$

Par ailleurs les probabilités de vide sont définies par

$$v(B) = \mathrm{E} \exp\left(- \int_B Z(\xi)\mathrm{d}\xi \right)$$

pour tout borné $B \subseteq S$. Ci-après nous décrivons deux exemples de processus ponctuels de Cox.

1.3.1 Processus ponctuel de Neymann-Scott

Pour le processus de Neyman-Scott [100], soit C un processus de Poisson stationnaire dans $\mathbb{R}^d$ avec d'intensité $\kappa > 0$. Conditionnellement à C, soit $\mathbf{X}_c$, $c \in C$, un processus de Poisson indépendant dans $\mathbb{R}^d$, où $\mathbf{X}_c$ admet une fonction d'intensité

$$\rho_c(\xi) = \alpha k(\xi - c)$$

où $\alpha > 0$ est un paramètre et k est un noyau (i.e.pour tout $c \in \mathbb{R}^d$, $\xi \to k(\xi - c)$ est une fonction de densité). Alors le processus $\mathbf{X} = \cup_{c \in C} \mathbf{X}_c$ est un processus de Neymann-Scott. Les points $c \in C$ sont appelés parents et ne sont pas observés. Les points de $\mathbf{X}_c$ sont appellés aussi les enfants du parent $c \in C$. Le processus $\mathbf{X}$ est en effet un processus de Cox dans $\mathbb{R}^d$ gouverné par

$$Z(\xi) = \sum_{c \in C} \alpha k(\xi - c).$$

Lorsque k est le noyau gaussien, $\mathbf{X}$ est appelé le processus de Thomas. La figure 1.1 propose quelques réalisations de ce processus pour différentes valeurs des paramètres (l'écart-type des noyaux gaussiens) et κ.

1.3.2 Processus ponctuel de Cox log-Gaussien

Soit $\mathbf{X}$ un processus de Cox dans $\mathbb{R}^d$ gouverné par $Z = \exp Y$, où Y est un champ aléatoire gaussien. Alors, $\mathbf{X}$ est appelé processus de Cox log Gaussien (PCLG). Introduit par Møller et al. [94], ce modèle admet les propriétés suivantes. Soient m et c la fonction de moyenne et de la fonction de covariance du champ gaussien en Y. La fonction d'intensité et la fonction de corrélation de paire de $\mathbf{X}$ sont définies par

$$\rho(\xi) = \exp\left(m(\xi) + c(\xi, \xi)/2\right), \quad g(\xi, \eta) = \exp(c(\xi, \eta)).$$

La figure 1.2 propose quelques réalisations d'un PCLG pour une fonction de corrélation exponentielle

$$C(u, u) = \sigma^2 \exp\left(-\frac{||u||}{2}\right).$$

Figure 1.1 – Quatre réalisations d'un processus ponctuel de Thomas sur $[0,1]^2$ pour différents paramètres

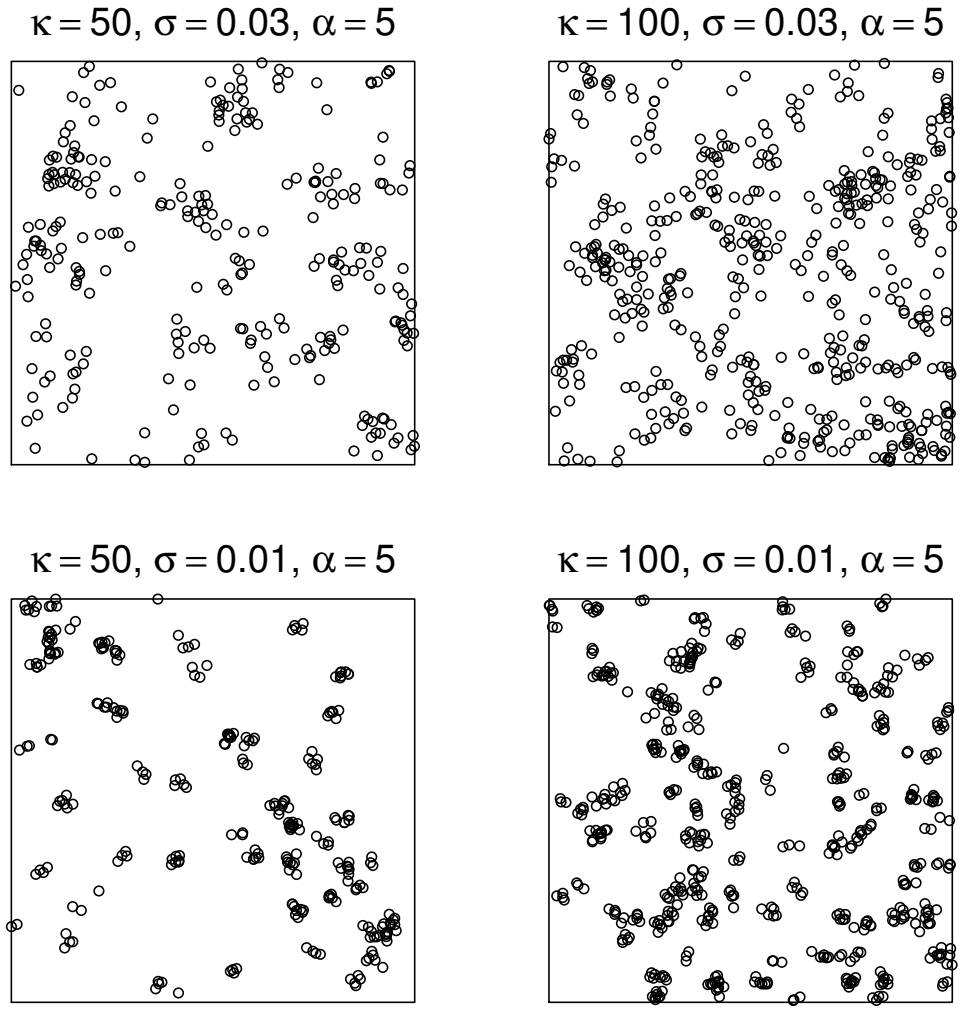
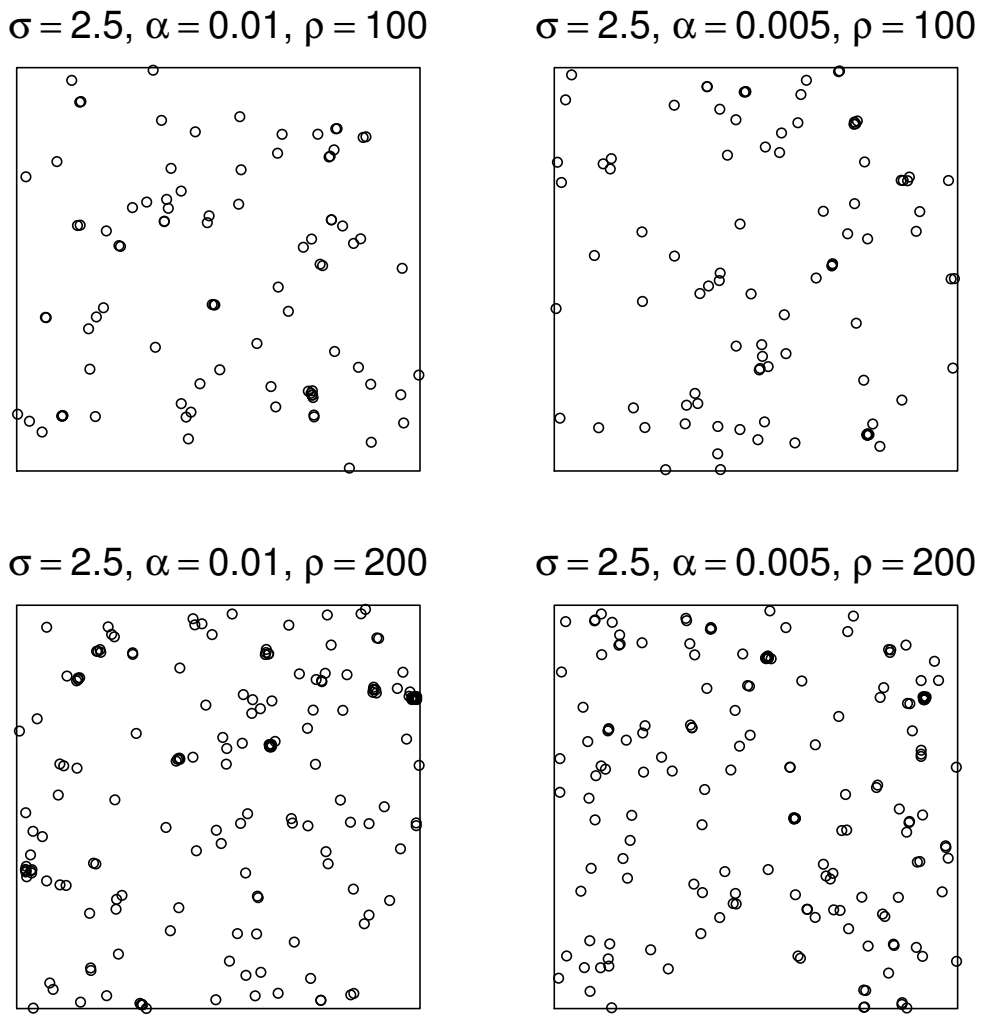


Figure 1.2 – Quatre réalisations d'un PCLG avec fonction de corrélation exponentielle sur $[0,1]^2$ pour différents paramètres. La moyenne m du champ gaussien est fixée de telle sorte que $\rho = \exp(m + \frac{\sigma^2}{2})$.



1.4 Processus ponctuel de Gibbs

1.4.1 Processus ponctuel de Gibbs dans une région bornée

On supposera dans ce paragraphe que S est une partie bornée de $\mathbb{R}^d$. Une façon de modéliser un processus ponctuel est de définir sa densité de probabilité f relativement à celle d'un PP de Poisson homogène de densité 1. Cette approche permet en particulier de définir les PP de Gibbs. Généralement, la densité f d'un PP n'est connue qu'à un facteur multiplicatif près, $f(\mathbf{x}) = cg(\mathbf{x})$, où g est analytiquement connue. Ceci est sans conséquence quant à la simulation d'un PP par une méthode MCMC, méthode qui requiert la seule connaissance de g. Par contre, pour l'estimation par maximum de vraisemblance du modèle $f_\theta(\mathbf{x}) = c(\theta)g_\theta(\mathbf{x}), c(\theta)$, qui est une constante analytiquement incalculable, devra être approchée par la méthode MCMC. Les processus ponctuels de Gibbs viennent de la physique statistique et caractérisent l'état d'équilibre des particules à l'intérieur de sytèmes physiques fermés. Sa densité est donnée par

$$f(\mathbf{x}) = \frac{\exp\left(-V(\mathbf{x})\right)}{Z}$$

où Z est une fonction de partition normalisante, en règle générale incalculable. V une fonction d'énergie, sera souvent dérivée d'un potentiel d'interaction γ :

$$V(\mathbf{x}) = \sum_{\psi \subset \mathbf{x}} \gamma(\psi).$$

1.4.2 Intensité conditionnelle de Papangelou

Une densité f est dite héréditaire si, pour toute configuration finie de points $\mathbf{x} \in N_f$, où $N_f = \{x \subseteq \mathbb{R}^d : n(\mathbf{x}) < \infty\}$,

$$f(\mathbf{x}) > 0 \text{ et pour tout } \mathbf{y} \subseteq \mathbf{x} \Longrightarrow f(\mathbf{y}) > 0.$$

Cette condition, signifie que si une configuration peut se produire, alors toutes les sous-configurations qu'elle contient peuvent se produire aussi.
Si f est héréditaire, l'intensité conditionnelle de Papangelou (voir Papangelou [105]) de $u \notin \mathbf{x}$ conditionnelle à $\mathbf{x}$ est définie par

$$\lambda(u, \mathbf{x}) = \frac{f(\mathbf{x} \cup \{u\})}{f(\mathbf{x})}.$$

λ présente l'avantage par rapport à f d'éliminer la constante de normalisation, c'est pourquoi son usage sera central. L'intensité conditionnelle de Papangelou $\lambda(u, \mathbf{x})$ peut être interprétée comme étant la densité de probabilité qu'il y ait un point de $\mathbf{X}$ en

u, sachant que la réalisation de $\mathbf{X}$ est $\mathbf{x}$ ailleurs. Par ailleurs, l'intensité conditionnelle de Papangelou est reliée à l'intensité du processus par la formule (Møller et Waagepetersen [95]) :

$$\mathrm{E}[\lambda(u, \mathbf{X})] = \rho(u).$$

Pour un PPP(ρ), $\lambda(u, \mathbf{x}) = \rho$ ne dépend pas de $\mathbf{x}$ pour les PP markoviens, l'intensité conditionnelle $\lambda(u, \mathbf{x}) = \lambda(u, \mathbf{x} \cap \partial u)$ ne dépendra que de la configuration $\mathbf{x}$ au voisinage ∂u de u.

Nous disons souvent que $\mathbf{X}$ est

- attractif si
$$\lambda(u, \mathbf{x}) \leq \lambda(u, \mathbf{y}) \quad \text{quand} \quad \mathbf{x} \subset \mathbf{y}$$
- répulsif si
$$\lambda(u, \mathbf{x}) \geq \lambda(u, \mathbf{y}) \quad \text{quand} \quad \mathbf{x} \subset \mathbf{y}.$$

Existence d'un processus ponctuel de Gibbs dans un borné

L'existence des mesures de Gibbs fut démontrée dans les années 70. Selon le type d'interaction, des problèmes liés à la stabilité du processus peuvent apparaître. Dans le cas d'un processus ponctuel attractif, il est important que la densité du processus ponctuel reste intégrable par rapport à un processus de Poisson de référence pour assurer la stabilité du processus. Par ailleurs des critères de stabilité nécessaires pour obtenir la convergence des algorithmes de simulation par les chaînes de Markov ont été proposés par Ruelle [117] et van Lieshout [87]. Ils peuvent être regroupés dans les conditions suivantes :

Définition 1.4.1. *On suppose que $\gamma^\star : S \longrightarrow [0, \infty)$ est une fonction telle que $c^\star = \int_S \gamma^\star(u)du$ est finie. Un processus ponctuel avec une densité non-normalisée $f(\cdot)$ par rapport à une mesure de référence est*

- *localement stable si*
$$f(\mathbf{x} \cup \{u\}) \leq \gamma^\star(u) f(\mathbf{x})$$
pour tout $\mathbf{x} \in N_f$ et $u \in S \backslash \mathbf{x}$.
- *stable au sens de Ruelle (ou $\gamma^\star$- stabilité de Ruelle) si*
$$f(\mathbf{x}) \leq \alpha \prod_{u \in \mathbf{x}} \gamma^\star(u)$$
pour une certaine constante positive α et tout $\mathbf{x} \in N_f$ (l'espace des configurations finies).

La stabilité locale est satisfaite par de nombreux modèles de processus ponctuels (Geyer [59], Kendall et Møller [83]). Un exemple, où la stabilité au sens de Ruelle mais pas la stabilité locale est satisfaite, est le modèle de Lennard-Jones (voir la Section 1.4.4). Ce modèle sera présenté à nouveau plus tard lorsque l'on discute des modèles dans $\mathbb{R}^d$.

1.4.3 Processus ponctuels de Gibbs stationnaires sur $\mathbb{R}^d$

Nous nous focalisons ici sur la définition d'un processus ponctuel de Gibbs dans $\mathbb{R}^d$. Les approches précédentes sur l'existence des processus ponctuels de Gibbs sur des régions bornées sont inapplicables sur $\mathbb{R}^d$. Il y a plusieurs façons de définir un PP de Gibbs stationnaire sur $\mathbb{R}^d$, voir Georgii [57], Møller et Waagepetersen [95], Nguyen et Zessin [102] à ce sujet. La définition de ces PP est importante et primordiale pour évaluer les propriétés asymptotiques des estimateurs d'un modèle du processus ponctuel observé dans une fenêtre Λ destinée à tendre vers $\mathbb{R}^d$. Soit $\Lambda \in \mathcal{B}_b(\mathbb{R}^d)$ et $V_\Lambda(\cdot) : \Omega \longrightarrow \mathbb{R} \cup \{+\infty\}$ est une fonction mesurable, on dit que $(V_\Lambda(\cdot))_{\Lambda \in \mathcal{B}_b(\mathbb{R}^d)}$ est une famille de potentiels compatibles si pour tout $\Lambda \subset \Lambda'$, il existe une fonction $\psi_{\Lambda,\Lambda'} : \Omega \longrightarrow \mathbb{R} \cup \{+\infty\}$, telle que : $\forall \mathbf{x} \in \Omega, V_{\Lambda'}(\mathbf{x}) = V_\Lambda(\mathbf{x}) + \psi_{\Lambda,\Lambda'}(\mathbf{x}_{\Lambda^c})$.
Etant donnée cette famille de potentiel, on peut alors définir un processus ponctuel de Gibbs de la façon suivante.

Définition 1.4.2. *Une mesure de probabilité $\mathbb{P}$ sur Ω est une mesure de Gibbs pour la famille de potentils compatibles $(V_\Lambda(\cdot))_{\Lambda \in \mathcal{B}_b(\mathbb{R}^d)}$. Si pour tout $\Lambda \in \mathcal{B}_b(\mathbb{R}^d)$ et pour $\mathbb{P}$-presque toute configuration $\mathbf{x}_{\Lambda^c}$ sur le complémentaire de Λ, la loi de $\mathbb{P}$ sachant $\mathbf{x}_{\Lambda^c}$ admet la densité conditionnelle suivante par rapport au processus de Poisson d'intensité* 1

$$f_\Lambda(\mathbf{x}_\Lambda \mid \mathbf{x}_{\Lambda^c}) = Z_\Lambda^{-1}(\mathbf{x}_{\Lambda^c}) \exp(-V_\Lambda(\mathbf{x}))$$

où $Z_\Lambda(\mathbf{x}_{\Lambda^c})$ est une fonction de normalisation.

Définition 1.4.3. *L'énergie locale pour insérer un point u dans une configuration $\mathbf{x}$ du processus est définie, pour tout Λ contenant u par : $V(u|\mathbf{x}) = V_\Lambda(\mathbf{x} \cup \{u\}) - V_\Lambda(\mathbf{x})$.*

Notons que l'hypothèse de compatibilité de la famille $(V_\Lambda(\cdot))_{\Lambda \in \mathcal{B}_b(\mathbb{R}^d)}$ implique que l'énergie locale est indépendante de Λ. Étant donnée $\mathbf{x}$ une configuration localement finie, nous pouvons étendre la définition d'intensité conditionnelle de Papangelou, voir Møller et Waagepetersen [95] en posant

$$\lambda(u, \mathbf{x}) = \exp(-V(u|\mathbf{x})).$$

Elle représente la probabilité conditionnelle d'observer un point u dans un volume infinitésimale du autour de u, sachant une configuration $\mathbf{x}$.
La définition suivante contient des éléments de caractérisation d'un processus ponctuel de Gibbs largement utilisé dans toute la suite de ce document.

Définition 1.4.4. *On dit que l'intensité conditionnelle de Papangelou d'un processus ponctuel de Gibbs stationnaire*
1) a une portée finie R si pour tout $u \in \mathbb{R}^d$, $\mathbf{x} \in \Omega$,

$$\lambda(u, \mathbf{x}) = \lambda(u, \mathbf{x}_{B(u,R)}).$$

2) est localement stable si pour tout $u \in \mathbb{R}^d$, $\mathbf{x} \in \Omega$, il existe $K > 0$,

$$\lambda(u, \mathbf{x}) \leq K.$$

L'existence d'une mesure de Gibbs sur $\mathbb{R}^d$ répondant aux lois conditionnelles dites spécifications locales doivent généralement satisfaire certaines conditions de mesurabilité et de compatibilité. Nous renvoyons le lecteur intéressé à Ruelle [117], Preston [109], Bertin et al. [9], Dereudre [40] et Dereudre et al. [41] pour le développement technique et mathématique du problème de l'existence. Rappelons deux résultats issus de cette littérature.

1. Ruelle [117] et Preston [109] montrent que si γ est un potentiel d'interaction de singletons et de paires invariant par translation, $\gamma(x_i) = \alpha, \gamma(\{x_i, x_j\}) = g(x_i - x_j)$, l'existence d'une mesure de Gibbs est satisfaite pour une large classe de fonctions g : il suffit par exemple qu'existent deux fonctions g_1 et g_2 positives et décroissantes, g_1 sur $[0, a_1[$ et g_2 sur $[a_2, \infty[$, avec $0 < a_1 < a_2 < \infty$, telles que $g(t) \geq g_1(t)$ pour $t \leq a_1$ et $g(t) \geq -g_2(t)$ pour $t \geq a_2$ et vérifiant

$$\int_0^{a_1} g_1(t)\mathrm{d}t = \infty \quad \text{et} \quad \int_{a_2}^{\infty} g_2(t)\mathrm{d}t < \infty.$$

2. Bertin et al. [9] ont montré un résultat plus faible que le précédent résultat mais qui a l'avantage important de pouvoir être appliqué et vérifié sur de nombreux modèles. En substance, les auteurs montrent qu'il existe une mesure de Gibbs pour tout modèle héréditaire satisfaisant les propriétés de locale stabilité et de portée finie.

Notons que les résultats précédents parlent d'existence et non d'unicité. Si la mesure de Gibbs est unique elle est nécessairement ergodique. Dans le cas contraire, toute mesure de Gibbs s'obtient comme mélange de mesures ergodiques (Georgii [57]).

1.4.4 Exemples des processus ponctuels de Gibbs dans $\mathbb{R}^d$

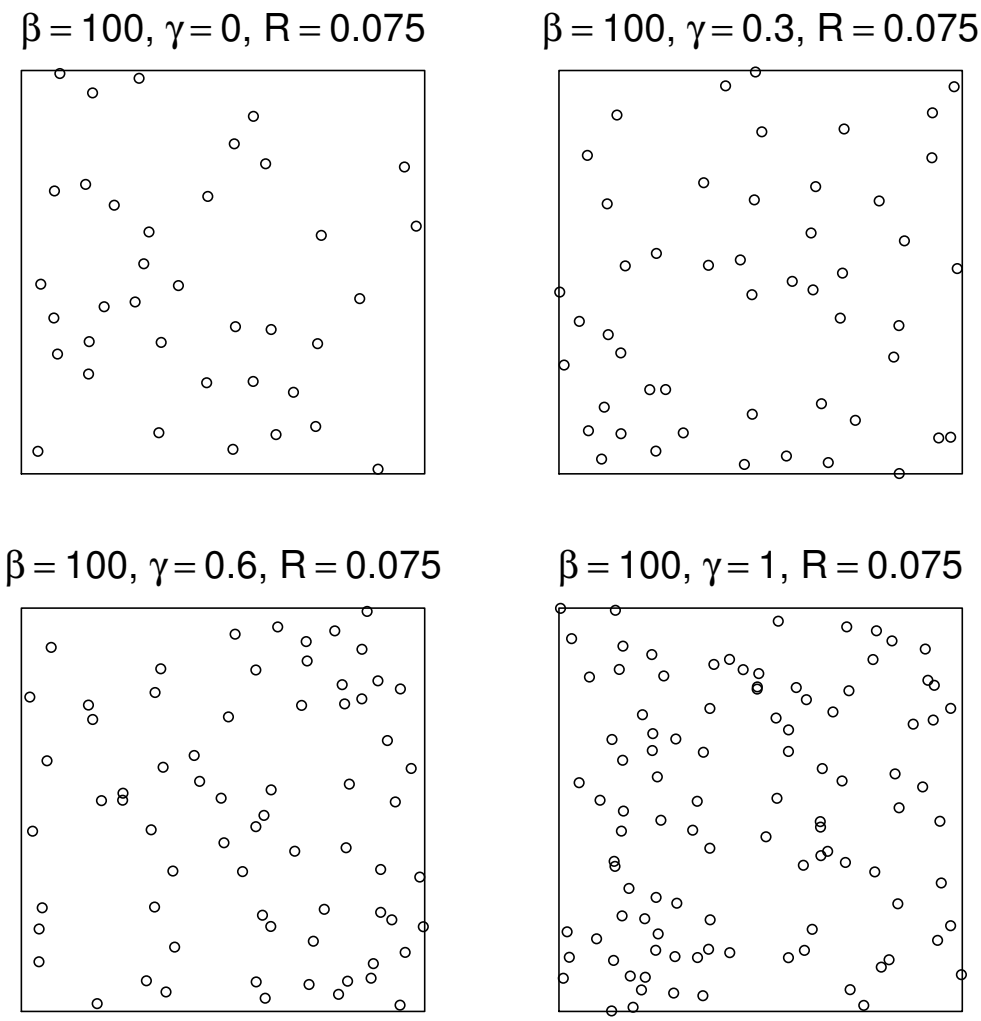
Nous présentons dans ce paragraphe des modèles classiques de processus ponctuels de Gibbs définis par leur intensité conditionnelle de Papangelou. Plus de détails sur les modèles (i)-(iv) et (vi)-(vii) peuvent être trouvés dans les travaux de Møller et Waagepetersen [95] et Geyer [59] pour l'exemple (v). Soient $u \in \mathbb{R}^d, \mathbf{x} \in \Omega$ et $R > 0$.

(i) Processus ponctuel de Strauss.

$$\lambda(u, \mathbf{x}) = \beta\gamma^{n_{[0,R]}(u,\mathbf{x}\setminus u)}$$

où $\beta > 0, \gamma \in [0,1]$ et $n_{[0,R]}(u, \mathbf{x}) = \sum_{v \in \mathbf{x}} \mathbb{1}(\|v - u\| \leq R)$ représente le nombre de R-voisins de u dans $\mathbf{x}$. La figure 1.3 propose quelques réalisations d'un tel processus pour différents paramètres.

Figure 1.3 – Réalisations de processus ponctuel de Strauss dans $[-1, 1]^2$ pour différents paramètres.



(ii) Processus ponctuel de Strauss avec hard-core. Soit $0 < \delta < R$, l'intensité conditionnelle de Papangelou est

$$\lambda(u, \mathbf{x}) = \begin{cases} 0 & \text{si } d(u, \mathbf{x} \setminus u) \leq \delta \\ \beta\gamma^{n_{]\delta,R]}(u,\mathbf{x}\setminus u)} & \text{par ailleurs} \end{cases}$$

où $\beta > 0, \gamma \in [0, 1]$.

(iii) Processus ponctuel de Strauss par morceaux.

$$\lambda(u, \mathbf{x}) = \beta \prod_{j=1}^{p} \gamma_j^{n_{[R_{j-1},R_j]}(u,\mathbf{x}\setminus u)}$$

où $\beta > 0, \gamma_j \in [0, 1], n_{[R_{j-1},R_j]}(u, \mathbf{x}) = \sum_{v\in\mathbf{x}} \mathbb{1}(\|v - u\| \in [R_{j-1}, R_j])$ et $R_0 = 0 < R_1 < \ldots < R_p = R < \infty$.

(iv) Processus ponctuel de triplets.

$$\lambda(u, \mathbf{x}) = \beta\gamma^{s_{[0,R]}(\mathbf{x}\cup\{u\})-s_{[0,R]}(\mathbf{x}\setminus u)}$$

où $\beta > 0, \gamma \in [0, 1]$ et $s_{[0,R]}(\mathbf{x})$ est le nombre de triplets non-ordonnés dont les points sont R-voisins.

(v) Processus ponctuel de saturation de Geyer avec un seuil de saturation s.

$$\lambda(u, \mathbf{x}) = \beta\gamma^{t(\mathbf{x}\cup\{u\})-t(\mathbf{x}\setminus u)}$$

où $\beta > 0, \gamma \in [0, 1], s \geq 1$ et $t(\mathbf{x}) = \sum_{v\in\mathbf{x}} \min(s, n_{[0,R/2]}(v, \mathbf{x} \setminus v))$.

(vi) Modèle Lennard-Jones avec portée finie. Cela correspond au processus ponctuel à interaction par paires défini par

$$\lambda(u, \mathbf{x}) = \beta \prod_{v\in\mathbf{x}\setminus u} g(\|v - u\|)$$

où $r > 0$, $\log g(r) = \left(\theta^6 r^{-6} - \theta^{12} r^{-12}\right) \mathbb{1}_{]0,R]}(r)$, $\theta > 0$ et $\beta > 0$.

(vii) Processus ponctuel à interaction d'aire.

$$\lambda(u, \mathbf{x}) = \beta\gamma^{A(\mathbf{x}\cup\{u\})-A(\mathbf{x}\setminus u)}$$

où $\beta > 0, \gamma > 0$ et $A(\mathbf{x}) = \left| \cup_{v\in\mathbf{x}} B(v, R/2) \right|$.

Notons que d'après les travaux de Ruelle [117], le modèle Lennard-Jones, existe même si $R = +\infty$. L'existence des autres modèles est assurée par le fait que l'intensité conditionnelle de Papangelou est localement stable et satisfait une hypothèse de portée finie.

Formule de Georgii-Nguyen-Zessin

La formule suivante est appellée formule de Georgii-Nguyen-Zessin (GNZ). Elle constitue un outil important à la base de plusieurs méthodes d'estimation et de validation de processus ponctuels de Gibbs. Elle est aussi un moyen de caractérisation d'une mesure de Gibbs. Voir Papangelou [106] et Zessin [135] pour les commentaires historiques, nous citons Georgii [57] ou Nguyen et Zessin [102] pour une présentation générale. Pour toute fonction non négative $h : \mathbb{R}^d \times \Omega \longrightarrow \mathbb{R}$,

$$\mathrm{E} \sum_{u \in \mathbf{X}} h(u, \mathbf{X} \setminus u) = \mathrm{E} \int_{\mathbb{R}^d} h(u, \mathbf{X}) \lambda(u, \mathbf{X}) \mathrm{d}u. \tag{1.4.1}$$

Innovations

Le concept d'innovations des processus ponctuels spatiaux est proposé par Baddeley et al. [8] et est un ingrédient-clé dans ce document. Il découle de la formule GNZ (1.4.1) d'un processus ponctuel de Gibbs dont l'intensité conditionnelle de Papangelou est $\lambda(u, \mathbf{x})$. Pour un choix d'une fonction mesurable h de $\mathbb{R}^d \times \Omega$ dans $\mathbb{R}$, nous définissons la $h-$innovation sur un domaine borné Λ par la variable aléatoire centrée définie par

$$I_\Lambda(\mathbf{x}, h) = \sum_{u \in \mathbf{x}_\Lambda} h(u, \mathbf{x} \setminus u) - \int_\Lambda h(u, \mathbf{x}) \lambda(u, \mathbf{x}) \mathrm{d}u. \tag{1.4.2}$$

Les processus d'innovations qui sont abordés dans le présent chapitre ainsi que le suivant sont de premier ordre. Nous pouvons également définir les processus d'innovation d'ordre supérieur pour les processus spatiaux basés sur la formule (GNZ) itérée suivante : pour tout $n \in \mathbb{N}$ et pour toute fonction mesurable non négative h définie dans $(\mathbb{R}^d)^n \times \Omega$ à valeur dans $\mathbb{R}$

(1.4.3)

$$\mathrm{E} \sum_{u_1, \ldots, u_n \in \mathbf{X}}^{\neq} h(u_1, \ldots, u_n, \mathbf{X} \setminus \{u_1, \ldots, u_n\}) = \int \ldots \int \mathrm{E}\, h(u_1, \ldots, u_n, \mathbf{X}) \lambda(u_1, \ldots, u_n, \mathbf{X}) \mathrm{d}u_1 \ldots \mathrm{d}u_n$$

où $\lambda(u_1, \ldots, u_n, \mathbf{x})$ est appelée intensité conditionnelle de Papangelou d'ordre n. Elle est définie (de façon non unique) par

$$\lambda(u_1, \ldots, u_n, \mathbf{x}) = \lambda(u_1, \mathbf{x}) \lambda(u_2, \mathbf{x} \cup \{u_1\}) \ldots \lambda(u_n, \mathbf{x} \cup \{u_1, \ldots, u_{n-1}\}).$$

Cette formule sera également abondamment utilisée dans la suite de ce document. Plusieurs propriétés sur les processus d'innovation sont établies dans Baddeley et al.[7], Baddeley et al. [5]. Les résultats asymptotiques sont établis dans [29] et un calcul exact de leur covariance ainsi qu'une estimation de cette covariance est proposé dans [30].

Lemme 1.4.5. *(Baddeley et al. [5]) On suppose que $h(u;\mathbf{x}_\Lambda)$ une fonction mesurable non négative telle que $h(u,\mathbf{X}_\Lambda) = h(u;\mathbf{X}_\Lambda \cap B(u,R))$ pour tout $u \in \Lambda$. Alors pour tout ensemble borélien $B \subseteq \Lambda$, $I_B(\mathbf{x},h)$ dépend uniquement de $\mathbf{X}_{B\cup\partial B}$ et pour tout ensemble borélien $C \subseteq \mathbb{R}^d$ tel que $C \supseteq \partial B$,*

$$\mathrm{E}[I_B(\mathbf{x},h) \mid \mathbf{X}_C] = 0.$$

1.5 Quelques théorèmes limites

Dans cette section, nous rassemblons deux résultats théoriques largement utilisés par la suite pour établir des résultats asymptotiques.

1.5.1 Ergodicité et théorème ergodique

Soient $\mathbb{R}^S$ l'espace des états de $\mathbf{X}$ et E sa tribu borélienne, on dira qu'un événement $A \in E$ est invariant par translation si pour tout $i \in S$, $\tau_i(A) = A$ où τ_i est la i-translation sur $\mathbb{R}^S$ ainsi définie : pour tout $i \in S$ et $w \in \mathbb{R}^S, (\tau_i(w))_j = w_{j-i}$. Les événements invariants forment une sous-tribu $\mathcal{I} \subseteq E$.
Un processus stationnaire $\mathbf{X}$ est caractérisé par le fait que sa loi $\mathbb{P}$ est invariante par translation : pour tout événement $A = \{w : X_{s_1}(w) \in B_1, X_{s_2}(w) \in B_2, ..., X_{s_n}(w) \in B_n\}$ générateur de la tribu borélienne, on a :

$$\forall i \in S : \mathbb{P}(A) = \mathbb{P}(\tau_i(A)).$$

L'ergodicité est une propriété qui renforce la notion de stationnarité et qui permet d'obtenir la convergence presque sûre (p.s.) d'une moyenne empirique spatiale lorsque le domaine d'observation "tend vers l'infini".

Un processus ponctuel stationnaire $\mathbf{X}$ est ergodique si pour tout événement invariant $A \in \mathcal{I}, \mathbb{P}(A) = 0$ ou $\mathbb{P}(A) = 1$, où $\mathcal{I}$ est la tribu des invariants.

Si $\mathbf{X}$ est ergodique, la tribu $\mathcal{I}$ des invariants est triviale. L'interprétation heuristique est que si un événement invariant A est de probabilité strictement positif, alors il est de probabilité 1, l'ensemble des translatés de A engendrant l'espace de toutes les trajectoires.

Un autre résultat utilisé dans cette thèse, en dimension $d \geq 2$, la géométrie des dommaines interviendra dans la formulation du théorème ergodique suivant.

Théorème 1.5.1. *(Nguyen et Zessin [101]) Soit $\{F_\Lambda, \Lambda \in \mathcal{B}_b\}$ une famille de fonctions mesurables, que l'on suppose covariante pour tout $x \in \mathbb{R}^d$, i.e.*

$$F_{\tau_x(\Lambda)}(\tau_x(\mathbf{x})) = F_\Lambda(\mathbf{x}) \quad \textit{pour presque tout} \quad \mathbf{x}$$

et additive i.e. pour tout $\Lambda, \Lambda' \in \mathcal{B}_b$ disjoints

$$F_{\Lambda \cup \Lambda'} = F_\Lambda + F_{\Lambda'} \quad p.s. \quad .$$

Supposons qu'il existe une variable aléatoire non négative et intégrable Y tel que $|F_\Lambda| \leq Y$ p.s. pour tout convexe $\Lambda \subset \Lambda_0$. Alors

$$\lim_{n\to\infty} \frac{1}{|\Lambda_n|} F_{\Lambda_n} = \frac{1}{|\Lambda_0|} E[F_{\Lambda_0}/\mathcal{I}] \quad p.s.$$

pour toute suite de domaines réguliers $(\Lambda_n)_{n\geq 1}$ convergeant vers $\mathbb{R}^d$, où $\mathcal{I}$ la sous-σ-algèbre des ensembles invariants.

Si, de plus, le processus est ergodique la limite coïncide avec l'espérance de la variable F_{Λ_0}.

1.5.2 Théorème central limite pour une fonctionnelle d'un champ aléatoire markovien

Dans le cas des variables aléatoires, il existe quelques théorèmes qui traitent le cas de sommes de variables dépendantes, par exemple le théorème de la limite centrale m-dépendante dont la référence de Heinrich [70], le théorème de la limite centrale pour les processus mélangeants, voir par exemple Bolthausen [17] et le théorème de la limite centrale avec des conditions de type centrage conditionnel, on pourra par exemple consulter Jensen et Künsch [77]. Dans ce travail de thèse, nous nous intéressons au comportement asymptotique de suites de champs stationnaires de variables aléatoires rélles. Un champ aléatoire réel $\mathbf{X}$ est une famille $(X_i)_{i\in\mathbb{Z}^d}$ de variables aléatoires réelles indexées par le réseau $\mathbb{Z}^d$ et définies sur un espace probabilisé. Dans le théorème ci-dessous proposé par Jensen et Künsch [77], les auteurs ont remplacé les coefficients de mélange par une hypothèse d'ergodicité et une hypothèse de type centrage conditionnel.

Théorème 1.5.2. *Soit $(X_i)_{i\in\mathbb{Z}^d}$, un champ aléatoire stationnaire et ergodique avec $X_i \in S$ pour un espace mesurable S et soit $f : S^{I_0} \longrightarrow \mathbb{R}^k$, où $I_0 = \{j \in \mathbb{Z}^d; |j| \leq 1\}$. Supposons $Z_j = f(X_{i+j}, i \in I_0), j \in \mathbb{Z}^d$ et $\Delta = \mathrm{E}[\sum_{|j|\leq 1} Z_0^* Z_j]$, où $*$ désigne le vecteur transposé. Pour une région $I_n \subset \mathbb{Z}^d$, soit $S_n = \sum_{i\in I_n} Z_i$. Si*

i) $\mathrm{E}\left[Z_i | X_j, j \neq i\right] = 0, i \in \mathbb{Z}^d$,
ii) $\mathrm{E}\,|Z_i|^3 < \infty$,
iii) I_n *croît vers* $\mathbb{Z}^d$ *avec* $\frac{|\partial I_n|}{|I_n|} \to 0$,
alors

$$|I_n|^{-1/2} S_n \xrightarrow{d} \mathcal{N}_k(0, \Delta) \quad \textit{quand} \quad n \to \infty$$

où $\xrightarrow{d}$ *désigne la convergence en distribution.*

Précisons que ce résultat a été étendu de différentes façons pour des champs non stationnaires par Comets et Janzura [33] (en ajoutant une condition de moment d'ordre 4) et par Coeurjolly et Lavancier [29] pour les tableaux triangulaires, non nécessairement stationnaires.

Chapitre 2

Estimation du paramètre d'intensité de Poisson d'un processus ponctuel de Gibbs stationnaire de portée finie

Résumé

L'approche développée dans ce chapitre a donné lieu à l'article (Coeurjolly et Morsli [31]), intitulé "Poisson intensity parameter estimation for stationary Gibbs point processes of finite interaction range." A partir d'une réalisation d'un processus ponctuel stationnaire de Gibbs observée sur un domaine borné de $\mathbb{R}^d$ ($d \geq 1$), nous décrivons la procédure d'estimation du paramètre d'intensité de Poisson d'une intensité conditionnelle de Papangelou d'un processus ponctuel de Gibbs stationnaire. Sous l'hypothèse d'une portée finie, l'idée sur laquelle l'estimation est basée permet de négliger les termes d'interaction d'ordre supérieur de l'intensité conditionnelle de Papangelou. Dans la Section 2.1, nous précisons les principales hypothèses sur notre modèle. Ensuite, nous définissons précisément l'estimateur du paramètre de l'intensité de Poisson dans la Section 2.2. Les principaux résultats asymptotiques et les preuves techniques sont présentés dans la Section 2.3. Une étude par simulation est menée dans la Section 2.4, où nous illustrons l'efficacité de notre estimateur.

2.1 Principales hypothèses et discussion

Nous partons du fait que nous supposons disposer de la réalisation d'un modèle de Gibbs stationnaire $\mathbf{X}$ de $\mathbb{R}^d$ défini et caractérisé par l'intensité conditionnelle de Papangelou, voir Møller et Waagepetersen [95]. Dans le présent chapitre, nous supposons que l'intensité conditionnelle de Papangelou se décompose comme

$$\lambda_\beta(u, \mathbf{x}) = \beta\, \widetilde{\lambda}(u, \mathbf{x}) \tag{2.1.1}$$

où $\beta > 0$ est le terme d'interaction du premier ordre, appelé le paramètre de l'intensité de Poisson et $\widetilde{\lambda}$ est une fonction de $\mathbb{R}^d \times \Omega$ dans $\mathbb{R}_+$, regroupe les termes d'interaction d'ordre supérieur. L'idée maîtresse dans ce travail est de décrire la procédure d'estimation de la partie paramétrique, c'est-à-dire l'estimation du paramètre de l'intensité de Poisson β indépendamment de la fonction $\widetilde{\lambda}$. Nous avons besoin des deux hypothèses suivantes pour notre étude.

- L'intensité conditionnelle de Papangelou a une portée finie R, i.e.

$$\lambda_\beta(u, \mathbf{x}) = \lambda_\beta(u, \mathbf{x}_{B(u,R)}) \tag{2.1.2}$$

  pour tout $u \in \mathbb{R}^d$, $\mathbf{x} \in \Omega, \beta > 0$.
- La fonction $\widetilde{\lambda}$ satisfait pour tout $u \in \mathbb{R}^d$

$$\widetilde{\lambda}(u, \emptyset) = 1. \tag{2.1.3}$$

Si le processus ponctuel correspond à un processus ponctuel de Poisson stationnaire $\widetilde{\lambda}(u, \mathbf{x}) = 1$, $\beta = \lambda_\beta(u, \mathbf{x})$.

Nous n'exigeons pas dans cette étude que le modèle du processus ponctuel de Gibbs soit localement stable ; hypothèse souvent envisagée lorsqu'il s'agit de tels modèles, voir par exemple Billiot et al. [15]. Nous soulignons aussi que dans ce qui suit, nous ne supposons pas que la portée R est connue. Au lieu de cela, nous supposons simplement qu'un borne supérieure notée $\widetilde{R}$ est connue.

Discutons les hypothèses concernant les exemples présentés dans la Section 1.4.4. Les exemples (i)-(vii) satisfont clairement les hypothèses (2.1.1) et (2.1.2) avec une portée finie R. Par ailleurs les exemples (i)-(vi) satisfont l'hypothèse (2.1.3), alors que l'exemple (vii) est un contre-exemple puisque pour ce dernier

$$\widetilde{\lambda}(u, \emptyset) = \gamma^{\pi(R/2)^2} \neq 1$$

pour tout $\gamma > 0$. La condition (2.1.3) peut donc être considérée comme un caractère d'identifiabilité.

Par la suite, nous notons $\beta^\star$ le véritable paramètre à estimer (supposé un réel positif) et $P_{\beta^\star}$ la mesure de probabilité de $\mathbf{X}$ i.e. $\mathbf{X} \sim \mathbb{P}_{\beta^\star}$. Notre estimation semi-paramétrique est basée sur une application de la formule Georgii-Nguyen-Zessin (1.4.1), pour le choix de la fonction mesurable non négative h définie pour tout $u \in \mathbb{R}^d$, $\mathbf{x} \in \Omega$ et $\widetilde{R} \geq R$ par

$$h(u, \mathbf{x}) = \mathbb{1}(\mathbf{x} \cap B(u, \widetilde{R}) = \emptyset). \tag{2.1.4}$$

L'idée est formulée par le résultat suivant. Soit Λ une région bornée de $\mathbb{R}^d$ et soient $N_\Lambda(\mathbf{X}; \widetilde{R})$ et $V_\Lambda(\mathbf{X}; \widetilde{R})$ les variables aléatoires suivantes :

$$N_\Lambda(\mathbf{X}; \widetilde{R}) = \sum_{u \in \mathbf{X}_\Lambda} h(u, \mathbf{X} \backslash u) \quad \text{et} \quad V_\Lambda(\mathbf{X}; \widetilde{R}) = \int_\Lambda h(u, \mathbf{X}) \mathrm{d}u.$$

La variable $N_\Lambda(\mathbf{X}; \widetilde{R})$ correspond au nombre de points dans $\mathbf{X}_\Lambda$ qui sont séparés des autres points dans $\mathbf{X}$ par une distance supérieure à $\widetilde{R}$, tandis que la variable $V_\Lambda(\mathbf{X}; \widetilde{R})$ représente le volume restreint de Λ après avoir enlever un disque de rayon $\widetilde{R}$ centré en chaque point de $\mathbf{X}$.

Proposition 2.1.1. *Soit $\mathbf{X}$ un processus ponctuel de Gibbs stationnaire d'intensité conditionnelle de Papangelou définie en (2.1.1), de portée finie (2.1.2) et sous la condition (2.1.3). Alors pour tout $\widetilde{R}$ tel que $\widetilde{R} \geq R$,*

$$\mathrm{E}\, N_\Lambda(\mathbf{X}; \widetilde{R}) = \beta^\star \,\mathrm{E}\, V_\Lambda(\mathbf{X}; \widetilde{R}) \tag{2.1.5}$$

où E représente l'espérance par rapport à la mesure de $\mathbb{P}_{\beta^*}$.

Démonstration. En utilisant la formule GNZ (1.4.1) appliquée à la fonction h définie en (2.1.4), nous obtenons

$$\mathrm{E}\, N_\Lambda(\mathbf{X}; \widetilde{R}) = \beta^\star \,\mathrm{E} \int_\Lambda \mathbb{1}(\mathbf{X} \cap B(u, \widetilde{R}) = \emptyset) \widetilde{\lambda}(u, \mathbf{X}) \mathrm{d}u.$$

Maintenant, en utilisant la propriété de portée finie (2.1.2), l'hypothèse (2.1.3) et le fait que $\widetilde{R} \geq R$, nous concluons la preuve avec

$$\begin{aligned}
\mathrm{E}\, N_\Lambda(\mathbf{X}; \widetilde{R}) &= \beta^\star \,\mathrm{E} \int_\Lambda \mathbb{1}(\mathbf{X} \cap B(u, \widetilde{R}) = \emptyset) \widetilde{\lambda}(u, \mathbf{X} \cap B(u, R)) \mathrm{d}u \\
&= \beta^\star \,\mathrm{E} \int_\Lambda \mathbb{1}(\mathbf{X} \cap B(u, \widetilde{R}) = \emptyset) \widetilde{\lambda}(u, \emptyset) \mathrm{d}u \\
&= \beta^\star \,\mathrm{E}\, V_\Lambda(\mathbf{X}; \widetilde{R}).
\end{aligned}$$

□

En utilisant un théorème ergodique classique pour les processus ponctuels spatiaux obtenus par Nguyen et Zessin [101] et rappelé dans le premier chapitre, nous pouvons nous attendre à ce que lorsque le domaine d'observation est large $N_\Lambda(\mathbf{X};\widetilde{R}) \simeq \mathrm{E}\, N_\Lambda(\mathbf{X};\widetilde{R})$ et $V_\Lambda(\mathbf{X};\widetilde{R}) \simeq \mathrm{E}\, V_\Lambda(\mathbf{X};\widetilde{R})$. Avec cela et la Proposition 2.1.1, nous suggérons un estimateur naturel du paramètre $\beta^\star$ et nous présentons ses propriétés asymptotiques dans la section suivante.

2.2 Présentation de l'estimateur de l'intensité de Poisson

Supposons que nous observons la réalisation d'un processus ponctuel de Gibbs stationnaire $\mathbf{X}$ dans un domaine borné de $\mathbb{R}^d$ noté Λ_n et supposons que la suite $(\Lambda_n)_{n\geq 1}$ tend vers $\mathbb{R}^d$ quand n tend vers l'infini. Les points situés près du bord de la zone d'étude posent problème car une partie du cercle dans laquelle nous comptons les voisins se trouve hors de la zone d'étude. C'est un problème très général, connu sous le nom d'effet de bord, qui se rencontre dans beaucoup de méthodes d'analyse de la structure spatiale. Plusieurs auteurs ont fait le bilan des méthodes de correction des effets de bord (voir par exemple Ripley [111], Biber [13], Kelly et Ripley [82], Møller et Waagepetersen [95] et Tukey [130]). Ce problème peut par exemple se résoudre par l'introduction du $\widetilde{R}$-intérieur de Λ_n défini par

$$\Lambda_{n,\widetilde{R}} = \{u \in \Lambda_n : B(u,\widetilde{R}) \subseteq \Lambda_n\}.$$

Nous nous intéressons au comportement asymptotique de suites de champs stationnaires de variables aléatoires rélles. Soit alors Λ_i, $i = (i_1, ..., i_d)$ un cube $\{z \in \mathbb{R}^d; \widetilde{R}(i_j - \frac{1}{2}) \leq z_j \leq \widetilde{R}(i_j + \frac{1}{2}), j = 1, ..., d\}$, pour tout $\widetilde{R} \geq R$. Nous supposons que Λ_n, se décompose comme $\Lambda_n = \cup_{i\in I_n}\Lambda_i$, I_n des sous ensembles de $\mathbb{Z}^d$ et que le processus est observé dans $\Lambda_{n,\widetilde{R}} = \{z \in \Lambda_n \mid ||z-t|| \leq \widetilde{R},\ \text{pour tout } t \in \Lambda_n\} = \cup_{i\in\widetilde{I}_n}\Lambda_i$, avec $\widetilde{I}_n = \{i \in I_n \mid |i-j| \leq 1,\ \text{pour tout } j \in I_n\}$, où $|j| = \max\{|j_1|, ..., |j_d|\}$. Nous supposerons que $\widetilde{I}_n$ croît vers $\mathbb{Z}^d$ et $\frac{|\partial\widetilde{I}_n|}{|\widetilde{I}_n|}$ tend vers 0 quand n tend vers l'infini, où $\partial\widetilde{I}_n = \{i \in \widetilde{I}_n \mid \exists j \notin \widetilde{I}_n : |j-i| = 1\}$. Il résulte de la Proposition 2.1.1 qu'un estimateur naturel du paramètre $\beta^\star$, noté $\hat{\beta}_n$ et défini par

$$\widehat{\beta}_n = \widehat{\beta}_n(\mathbf{X};\widetilde{R}) = \frac{N_{\Lambda_{n,\widetilde{R}}}(\mathbf{X};\widetilde{R})}{V_{\Lambda_{n,\widetilde{R}}}(\mathbf{X};\widetilde{R})}. \tag{2.2.1}$$

2.3 Résultats asymptotiques

Nous explicitons nos résultats asymptotiques de l'estimateur $\widehat{\beta}_n$ défini en (2.2.1) du paramètre $\beta^\star$.

Proposition 2.3.1. *Soit* $\mathbf{X}$ *un processus ponctuel de Gibbs stationnaire, d'intensité conditionnelle de Papangelou définie en* (2.1.1)*. Sous les conditions* (2.1.2) *et* (2.1.3)*. Alors pour tout* $\widetilde{R}$ *tel que* $\widetilde{R} \geq R$*, l'estimateur* $\widehat{\beta}_n$ *défini en* (2.2.1) *du paramètre* $\beta^\star$ *est fortement consistant.*

Théorème 2.3.2. *Soit* $\mathbf{X}$ *un processus ponctuel de Gibbs stationnaire, d'intensité conditionnelle de Papangelou définie en* (2.1.1) *satisfaisant les conditions* (2.1.2) *et* (2.1.3)*. Alors pour tout* $\widetilde{R}$ *tel que* $\widetilde{R} \geq R$*,*
(i) si $\mathbf{X}$ *est ergodique, la convergence en distribution suivante est satisfaite*

$$|\Lambda_{n,\widetilde{R}}|^{1/2}(\widehat{\beta}_n - \beta^\star) \xrightarrow{d} \mathcal{N}(0, \sigma^2)$$

où

$$\sigma^2 = \sigma^2(\beta^\star, \widetilde{R}) = \frac{\beta^\star}{1 - F(\widetilde{R})} + \frac{\beta^{\star 2} \int_{B(o,\widetilde{R})} \bar{F}(o, v)\mathrm{d}v}{(1 - F(\widetilde{R}))^2}$$

avec

$$F(\widetilde{R}) = \mathbb{P}_{\beta^\star}(\mathbf{X} \cap B(o, \widetilde{R}) \neq \emptyset)$$

et

$$\bar{F}(o, v) = \mathbb{P}_{\beta^\star}(\mathbf{X} \cap B(o, \widetilde{R}) = \emptyset, \mathbf{X} \cap B(v, \widetilde{R}) = \emptyset).$$

(ii) La convergence en distribution est satisfaite

$$|\Lambda_{n,\widetilde{R}}|^{1/2} \frac{(\widehat{\beta}_n - \beta^\star)}{\widehat{\sigma}_n} \xrightarrow{d} \mathcal{N}(0, 1) \tag{2.3.1}$$

où $\widehat{\sigma}_n^2 = \widehat{\sigma}_n^2(\mathbf{X}; \widetilde{R})$ *est l'estimateur fortement consistant de* σ^2 *défini par*

$$\widehat{\sigma}_n^2 = |\Lambda_{n,\widetilde{R}}| \left(\frac{\widehat{\beta}_n}{V_{\Lambda_{n,\widetilde{R}}}(\mathbf{X}; \widetilde{R})} + \frac{\widehat{\beta}_n^2 W_{\Lambda_{n,\widetilde{R}}}(\mathbf{X}; \widetilde{R})}{V^2_{\Lambda_{n,\widetilde{R}}}(\mathbf{X}; \widetilde{R})} \right) \tag{2.3.2}$$

avec

$$W_{\Lambda_{n,\widetilde{R}}} = \int_{\Lambda_{n,\widetilde{R}}} \int_{B(u,\widetilde{R}) \cap \Lambda_{n,\widetilde{R}}} \mathbb{1}(\mathbf{X} \cap B(u, \widetilde{R}) = \emptyset) \mathbb{1}(\mathbf{X} \cap B(v, \widetilde{R}) = \emptyset) \mathrm{d}u\mathrm{d}v.$$

L'intérêt du résultat (ii) du précédent théorème est double. Premièrement il est valable même si la mesure de Gibbs n'est pas ergodique. Il peut donc s'appliquer lorsque la mesure de Gibbs n'est pas unique. Deuxièmement, il permet de construire un intervalle de confiance asymptotique du paramètre $\beta^\star$. Nous renvoyons le lecteur à la Section 2.4 pour une étude par simulation de cet estimateur.

2.3.1 Preuve de la Proposition 2.3.1

Démonstration. Il existe au moins une mesure de Gibbs stationnaire. Si cette mesure est unique, elle est ergodique. Dans le cas contraire, elle peut être décomposée par un mélange de mesures ergodiques (voir Georgii [58], Théorème 14.10). Par conséquent, nous pouvons supposer, pour cette preuve, que $\mathbb{P}_{\beta^\star}$ est ergodique. Nous appliquons alors un théorème ergodique pour les processus ponctuels spatiaux obtenu par Nguyen et Zessin [101] annoncé dans le chapitre précédent, dans la Section 1.5.1. Pour cela, nous devons vérifier que :
(I) $N_\Lambda(\mathbf{X};\widetilde{R})$ et $V_\Lambda(\mathbf{X};\widetilde{R})$ sont des variables additives.
(II) Pour tout $y \in \mathbb{R}^d$

$$N_{\tau_y\Lambda}(\tau_y\mathbf{X};\widetilde{R}) = N_\Lambda(\mathbf{X};\widetilde{R}) \quad \text{et} \quad V_{\tau_y\Lambda}(\tau_y\mathbf{X};\widetilde{R}) = V_\Lambda(\mathbf{X};\widetilde{R}).$$

(III) Pour tout convexe Λ_0 de $\mathbb{R}^d$ de volume fini,

$$\mathrm{E}\,|N_{\Lambda_0}(\mathbf{X};\widetilde{R})| < \infty \quad \text{et} \quad \mathrm{E}\,|V_{\Lambda_0}(\mathbf{X};\widetilde{R})| < \infty.$$

Nous démontrons la condition (I). La condition de l'additivité résulte du fait que $N_\Lambda(\mathbf{X};\widetilde{R})$ (resp. $V_\Lambda(\mathbf{X};\widetilde{R})$) est une mesure aléatoire. En ce qui concerne la condition (II), soient Λ un borélien borné de $\mathbb{R}^d$ et y un point de $\mathbb{R}^d$, nous avons donc

$$\begin{aligned}
N_{\tau_y\Lambda}(\tau_y\mathbf{X};\widetilde{R}) &= \sum_{u\in\tau_y\mathbf{X}} \mathbb{1}_{\tau_y\Lambda}(u)\mathbb{1}((\tau_y\mathbf{X}\setminus u)\cap B(u,\widetilde{R}) = \emptyset)\\
&= \sum_{u\in\tau_y\mathbf{X}} \mathbb{1}_{\Lambda}(u+y)\mathbb{1}((\tau_y\mathbf{X}\setminus u)\cap B(u,\widetilde{R}) = \emptyset)\\
&= \sum_{v\in\mathbf{X}} \mathbb{1}_{\Lambda}(v)\mathbb{1}((\tau_y\mathbf{X}\setminus (v-y))\cap B(v-y,\widetilde{R}) = \emptyset)\\
&= \sum_{v\in\mathbf{X}} \mathbb{1}_{\Lambda}(v)\mathbb{1}((\mathbf{X}\setminus v)\cap B(v,\widetilde{R}) = \emptyset)\\
&= N_\Lambda(\mathbf{X};\widetilde{R}).
\end{aligned}$$

Par ailleurs

$$\begin{aligned}
V_{\tau_y(\Lambda)}(\tau_y(\mathbf{X});\widetilde{R}) &= \int \mathbb{1}_{\tau_y\Lambda}(u)\mathbb{1}((\tau_y\mathbf{X})\cap B(u,\widetilde{R}) = \emptyset)\mathrm{d}u\\
&= \int \mathbb{1}_{\Lambda}(u+y)\mathbb{1}((\tau_y\mathbf{X})\cap B(u,\widetilde{R}) = \emptyset)\mathrm{d}u\\
&= \int \mathbb{1}_{\Lambda}(v)\mathbb{1}((\tau_y\mathbf{X})\cap B(v-y,\widetilde{R}) = \emptyset)\mathrm{d}v\\
&= \int_\Lambda \mathbb{1}(\mathbf{X}\cap B(v,\widetilde{R}) = \emptyset)\mathrm{d}v\\
&= V_\Lambda(\mathbf{X};\widetilde{R}).
\end{aligned}$$

Nous démontrons maintenant la condition (III). Soit Λ_0 un convexe de $\mathbb{R}^d$ de volume fini. Par la stationnarité du processus ponctuel $\mathbf{X}$, nous avons donc

$$\mathrm{E}\left|V_{\Lambda_0}(\mathbf{X};\widetilde{R})\right| = \mathrm{E}\, V_{\Lambda_0}(\mathbf{X};\widetilde{R}) = \mathrm{E}\int_{\Lambda_0} \mathbb{1}(\mathbf{X}\cap B(u,\widetilde{R})=\emptyset)\mathrm{d}u = |\Lambda_0|(1-F(\widetilde{R})) < \infty.$$

D'après la Proposition 2.1.1, il résulte également

$$\mathrm{E}\left|N_{\Lambda_0}(\mathbf{X};\widetilde{R})\right| = \mathrm{E}\, N_{\Lambda_0}(\mathbf{X};\widetilde{R}) = \beta^\star\, \mathrm{E}\, V_{\Lambda_0}(\mathbf{X};\widetilde{R}) = \beta^\star|\Lambda_0|(1-F(\widetilde{R})) < \infty.$$

Quand $n\to\infty$, nous concluons par l'ergodicité du processus ponctuel $\mathbf{X}$, les convergences presque sûres suivantes :

$$\begin{aligned} |\Lambda_{n,\widetilde{R}}|^{-1} N_{\Lambda_{n,\widetilde{R}}}(\mathbf{X};\widetilde{R}) &\xrightarrow{p.s} \frac{1}{|\Lambda_0|}\,\mathrm{E}\, N_{\Lambda_0}(\mathbf{X};\widetilde{R}) = \beta^\star(1-F(\widetilde{R})) \\ |\Lambda_{n,\widetilde{R}}|^{-1} V_{\Lambda_{n,\widetilde{R}}}(\mathbf{X};\widetilde{R}) &\xrightarrow{p.s} \frac{1}{|\Lambda_0|}\,\mathrm{E}\, V_{\Lambda_0}(\mathbf{X};\widetilde{R}) = 1-F(\widetilde{R}). \end{aligned} \tag{2.3.3}$$

Puisque la fonction $(y,z)\mapsto \frac{y}{z}$ est continue en tout point de l'ensemble ouvert $\{(y,z)\in\mathbb{R}^2 : z\neq 0\}$, nous déduisons le résultat attendu. □

2.3.2 Preuve du Théorème 2.3.2

Démonstration. (i) De la définition du processus de l'innovation (1.4.2), nous avons

$$\begin{aligned} |\Lambda_{n,\widetilde{R}}|^{1/2}(\widehat{\beta}_n-\beta^\star) &= |\Lambda_{n,\widetilde{R}}|^{1/2}\frac{N_{\Lambda_{n,\widetilde{R}}}(\mathbf{X};\widetilde{R})-\beta^\star V_{\Lambda_{n,\widetilde{R}}}(\mathbf{X};\widetilde{R})}{V_{\Lambda_{n,\widetilde{R}}}(\mathbf{X};\widetilde{R})} \\ &= \frac{|\Lambda_{n,\widetilde{R}}|^{-1/2} I_{\Lambda_{n,\widetilde{R}}}(\mathbf{X};\widetilde{R})}{|\Lambda_{n,\widetilde{R}}|^{-1} V_{\Lambda_{n,\widetilde{R}}}(\mathbf{X};\widetilde{R})}. \end{aligned} \tag{2.3.4}$$

Le dénominateur de l'équation (2.3.4) converge presque sûrement vers $1-F(\widetilde{R})$ d'après (2.3.3). Le numérateur de l'équation (2.3.4) est un processus d'innovation défini en (1.4.2). Puisque $\mathbb{P}^*_\beta$ est ergodique, nous prouvons que $|\Lambda_{n,\widetilde{R}}|^{-1/2} I_{\Lambda_{n,\widetilde{R}}}(\mathbf{X};\widetilde{R})$ tend vers une distribution gaussienne de moyenne nulle. Nous appliquons le Théorème 2.1 de Jensen et Künsch [77], bien adapté à l'étude des mesures de Gibbs cité dans la Section 1.5.2, pour obtenir un théorème de la limite centrale pour $|\Lambda_{n,\widetilde{R}}|^{-1/2} I_{\Lambda_{n,\widetilde{R}}}(\mathbf{X};\widetilde{R})$. Posons $X_i = X_{\Lambda_i}$, $i\in\mathbb{Z}^d$ et nous vérifions les deux conditions fondamentales pour $i\in\mathbb{Z}^d$:
(I) $\mathrm{E}\left[I_{\Lambda_i}(\mathbf{X};\widetilde{R}) \mid X_{\Lambda_i^c}\right] = 0$.
(II) $\mathrm{E}\left|I_{\Lambda_i}(\mathbf{X};\widetilde{R})\right|^3 < \infty$.

Vérifions la condition (I). De la stationnarité du processus ponctuel $\mathbf{X}$, il suffit de prouver que : pour tout Λ un domaine borné de $\mathbb{R}^d$

$$\mathrm{E}\left[I_\Lambda(\mathbf{X};\widetilde{R}) \mid \mathbf{X}_{\Lambda^c}\right] = 0.$$

Pour le choix de la fonction mesurable non négative h définie en (2.1.4) qui satisfait la condition de portée finie, comme $I_\Lambda(\mathbf{X};\widetilde{R})$ dépend de $\mathbf{X}_{\Lambda\cup\partial\Lambda}$ et comme $\partial\Lambda \subseteq \Lambda^c$. Alors d'après le Lemme 1.4.5, nous avons le résultat voulu.

Vérifions maitenant la condition (II). En utilisant notament l'inégalité suivante $\mathrm{E}\,|Z-Y|^p \leq 2^{p-1}(\mathrm{E}\,|Z|^p + \mathrm{E}\,|Y|^p)$, nous obtenons pour tout Λ un domaine borné de $\mathbb{R}^d$

$$\mathrm{E}\,|I_\Lambda(\mathbf{X};\widetilde{R})|^3 \leq 4\left(\mathrm{E}\,N_\Lambda^3(\mathbf{X};\widetilde{R}) + \beta^{\star 3}\,\mathrm{E}\,V_\Lambda^3(\mathbf{X};\widetilde{R})\right). \tag{2.3.5}$$

D'une part, nous connaissons les caractéristiques de la fonction indicatrice, nous avons pour tout $m \geq 1$, $h^m(u,\mathbf{x}) = h(u,\mathbf{x})$,

$$\begin{aligned}\mathrm{E}\,N_\Lambda^3(\mathbf{X};\widetilde{R}) =& \mathrm{E}\,N_\Lambda(\mathbf{X};\widetilde{R}) + 3\,\mathrm{E}\sum_{u,v\in X_\Lambda}^{\neq} h(u,\mathbf{X}\setminus u)h(v,\mathbf{X}\setminus v)\\ &+ \mathrm{E}\sum_{u,v,w\in X_\Lambda}^{\neq} h(u,\mathbf{X}\setminus u)h(v,\mathbf{X}\setminus v)h(w,\mathbf{X}\setminus w).\end{aligned} \tag{2.3.6}$$

Soient T_1, T_2 et T_3 les trois termes de la partie droite de l'équation (2.3.6). D'après la preuve de la Proposition 2.3.1, $\mathrm{E}\,T_1 \leq \beta^\star|\Lambda|$. Maintenant, pour tous points $u, v \in \mathbb{R}^d$ et $\mathbf{x} \in \Omega$, nous rappelons la difinition de l'intensité conditionnelle de Papangelou du second ordre par

$$\lambda_{\beta^\star}(u,v,\mathbf{x}) = \lambda_{\beta^\star}(u,\mathbf{x})\lambda_{\beta^\star}(v,\mathbf{x}\cup\{u\}) = \lambda_{\beta^\star}(v,\mathbf{x})\lambda_{\beta^\star}(u,\mathbf{x}\cup\{v\}).$$

En utilisant les versions itérées de la formule GNZ (1.4.3) définie dans la Section 1.4.4 et la condition (2.1.3), nous avons les calculs suivants

$$\begin{aligned}T_2 &= 3\,\mathrm{E}\int_{\Lambda^2} h(u,\mathbf{X}\cup\{v\})h(v,\mathbf{X}\cup\{u\})\lambda_{\beta^\star}(u,v,\mathbf{X})\mathrm{d}u\mathrm{d}v\\ &= 3\beta^{\star 2}\,\mathrm{E}\int_{\Lambda^2} h(u,\mathbf{X}\cup\{v\})h(v,\mathbf{X}\cup\{u\})\widetilde{\lambda}(u,\mathbf{X})\widetilde{\lambda}(v,\mathbf{X}\cup\{u\})\mathrm{d}u\mathrm{d}v\\ &= 3\beta^{\star 2}\,\mathrm{E}\int_{\Lambda^2} \mathbb{1}((\mathbf{X}\cup\{v\})\cap B(u,\widetilde{R}) = \emptyset)\mathbb{1}((\mathbf{X}\cup\{u\})\cap B(v,\widetilde{R}) = \emptyset)\mathrm{d}u\mathrm{d}v\\ &\leq 3\beta^{\star 2}|\Lambda|^2.\end{aligned}$$

En utilisant le même raisonnement ci-dessus pour T_3, nous obtenons $T_3 \leq \beta^{\star 3}|\Lambda|^3$. D'autre part, nous vérifions également que

$$\begin{aligned}\mathrm{E}\, V_\Lambda^3(\mathbf{X};\widetilde{R}) &= \mathrm{E}\int_\Lambda\int_\Lambda\int_\Lambda h(u,\mathbf{X})h(v,\mathbf{X})h(w,\mathbf{X})\mathrm{d}u\mathrm{d}v\mathrm{d}w \\ &\leq |\Lambda|^3.\end{aligned}$$

Nous avons finalement obtenu

$$\mathrm{E}\,|I_\Lambda(\mathbf{X};\widetilde{R})|^3 \leq 4\beta^\star|\Lambda| + 12\beta^{\star 2}|\Lambda|^2 + 8\beta^{\star 3}|\Lambda|^3 < \infty.$$

Nous constatons que dans la preuve du Théorème 2.1 dans [77] que

$$\mathrm{Var}\left(|\Lambda_{n,\widetilde{R}}|^{-1/2} I_{\Lambda_{n,\widetilde{R}}}(\mathbf{X};\widetilde{R})\right) \longrightarrow \sigma^2 = \mathrm{E}\left[I_{\Lambda_0}(\mathbf{X};\beta^\star)\sum_{j\in B(o,1)} I_{\Lambda_j}(\mathbf{X};\beta^\star)\right]$$

sous la condition que $\mathrm{E}\left[I_{\Lambda_0}(\mathbf{X};\beta^\star)\sum_{j\in B(o,1)} I_{\Lambda_j}(\mathbf{X};\beta^\star)\right]$ est positive.

Pour démontrer cette dernière condition, nous pouvons évoquer une méthode qui simplifie le comportement asymptotique de la variance du processus de l'innovation $I_{\Lambda_{n,\widetilde{R}}}(\mathbf{X};\widetilde{R})$ par le lemme suivant en utilisant le résultat de [30].

Lemme 2.3.3. *Supposons que l'intensité conditionnelle de Papangelou du processus ponctuel* $\mathbf{X}$ *satisfait* (2.1.1) *et la condition de portée finie* (2.1.2). *Soit* $g : \mathbb{R}^d \times \Omega \to \mathbb{R}$ *une fonction mesurable non négative satisfaisant pour tout* $u \in \mathbb{R}^d$ *et* $\mathbf{x} \in \Omega$

$$g(u,\mathbf{x}) = g(o,\tau_u\mathbf{x}) \quad \textit{et} \quad g(u,\mathbf{x}) = g(u,\mathbf{x}_{B(u,\widetilde{R})})$$

avec $\widetilde{R} \geq R$. *Nous supposons également que les conditions d'intégrabilité suivantes sont remplies*

$$\mathrm{E}\int_{B(o,R)} g(o,\mathbf{X})g(v,\mathbf{X})\Big|\lambda_{\beta^\star}(o,\mathbf{X})\lambda_{\beta^\star}(v,\mathbf{X}) - \lambda_{\beta^\star}(o,v,\mathbf{X})\Big|\mathrm{d}v < \infty$$

$$\mathrm{E}\int_{B(o,\widetilde{R})} \Big|\big(g(o,\mathbf{X}\cup\{v\}) - g(o,\mathbf{X})\big)\big(g(v,\mathbf{X}\cup\{o\}) - g(v,\mathbf{X})\big)\Big|\lambda_{\beta^\star}(o,v,\mathbf{X})\mathrm{d}v < \infty.$$

Alors, quand $n \to \infty$

$$|\Lambda_n|^{-1}\,\mathrm{Var}\, I_{\Lambda_n}(\mathbf{X},g) \longrightarrow A_1 + A_2 + A_3$$

où

$$\begin{aligned}A_1 &= \mathrm{E}[g^2(o,\mathbf{X})\lambda_{\beta^\star}(o,\mathbf{X})], \\ A_2 &= \mathrm{E}\int_{B(o,R)} g(o,\mathbf{X})g(v,\mathbf{X})\big(\lambda_{\beta^\star}(o,\mathbf{X})\lambda_{\beta^\star}(v,\mathbf{X}) - \lambda_{\beta^\star}(o,v,\mathbf{X})\big)\mathrm{d}v, \\ A_3 &= \mathrm{E}\int_{B(o,\widetilde{R})} \big(g(o,\mathbf{X}\cup\{v\}) - g(o,\mathbf{X})\big)\big(g(v,\mathbf{X}\cup\{o\}) - g(v,\mathbf{X})\big)\lambda_{\beta^\star}(o,v,\mathbf{X})\mathrm{d}v.\end{aligned}$$

Appliquant le Lemme 2.3.3, avec la fonction $g = h$ définie par (2.1.4). Soient

$$g_0(\mathbf{X}) = \int_{B(o,R)} h(o,\mathbf{X})h(v,\mathbf{X})\Big|\lambda_{\beta^\star}(o,\mathbf{X})\lambda_{\beta^\star}(v,\mathbf{X}) - \lambda_{\beta^\star}(o,v,\mathbf{X})\Big|\mathrm{d}v$$
$$g_1(\mathbf{X}) = \int_{B(o,\widetilde{R})} \Big|\big(h(o,\mathbf{X}\cup\{v\}) - h(o,\mathbf{X})\big)\big(h(v,\mathbf{X}\cup\{o\}) - h(v,\mathbf{X})\big)\Big|\lambda_{\beta^\star}(o,v,\mathbf{X})\mathrm{d}v.$$

Ces deux variables sont bornées, donc admettent des moments de tout ordre. Nous pouvons maintenant déterminer les termes A_1, A_2 et A_3.
Tout d'abord, il est clair que $A_1 = \beta^\star(1 - F(\widetilde{R}))$. Deuxièmement, notons que d'après la condition de portée finie (2.1.2), nous avons pour tout point $v \in B^c(o,R)$ et $\mathbf{x} \in \Omega$

$$\lambda_{\beta^\star}(o,\mathbf{x})\lambda_{\beta^\star}(v,\mathbf{x}) - \lambda_{\beta^\star}(o,v,\mathbf{x}) = 0.$$

Puisque $B(o,R) \subseteq B(o,\widetilde{R})$, nous avons

$$\mathrm{E}\int_{B(o,\widetilde{R})\setminus B(o,R)} h(o,\mathbf{X})h(v,\mathbf{X})\big(\lambda_{\beta^\star}(o,\mathbf{X})\lambda_{\beta^\star}(v,\mathbf{X}) - \lambda_{\beta^\star}(o,v,\mathbf{X})\big)dv = 0$$

ce qui signifie que nous pouvons remplacer R par $\widetilde{R}$ dans le terme A_2. Alors, nous déduisons que

$$A_2 = \mathrm{E}\int_{B(o,\widetilde{R})} \mathbb{1}(\mathbf{X}\cap B(o,\widetilde{R}) = \emptyset)\mathbb{1}(\mathbf{X}\cap B(v,\widetilde{R}) = \emptyset)\big(\beta^{\star 2} - \lambda_{\beta^\star}(o,v,\mathbf{X})\big)\mathrm{d}v$$

et

$$\begin{aligned} A_3 &= \mathrm{E}\int_{B(o,\widetilde{R})} \big(\mathbb{1}((\mathbf{X}\cup\{v\})\cap B(o,\widetilde{R}) = \emptyset) - \mathbb{1}(\mathbf{X}\cap B(o,\widetilde{R}) = \emptyset)\big) \\ &\quad\times \big(\mathbb{1}((\mathbf{X}\cup\{o\})\cap B(v,\widetilde{R}) = \emptyset) - \mathbb{1}(\mathbf{X}\cap B(v,\widetilde{R}) = \emptyset)\big)\lambda_{\beta^\star}(o,v,\mathbf{X})\mathrm{d}v \\ &= \mathrm{E}\int_{B(o,\widetilde{R})} \big(\mathbb{1}(\mathbf{X}\cap B(o,\widetilde{R}) = \emptyset) - \mathbb{1}(\mathbf{X}\cap B(v,\widetilde{R}) = \emptyset)\big)\lambda_{\beta^\star}(o,v,\mathbf{X})\mathrm{d}v. \end{aligned}$$

Ce qui conduit à

$$\begin{aligned} A_2 + A_3 &= \beta^{\star 2}\,\mathrm{E}\int_{B(o,\widetilde{R})} \mathbb{1}(\mathbf{X}\cap B(o,\widetilde{R}) = \emptyset)\mathbb{1}(\mathbf{X}\cap B(v,\widetilde{R}) = \emptyset)\mathrm{d}v \\ &= \beta^{\star 2}\int_{B(o,\widetilde{R})} \bar{F}(o,v)\mathrm{d}v. \end{aligned}$$

Nous déduisons que

$$|\Lambda_{n,\widetilde{R}}|^{-1}\,\mathrm{Var}\, I_{\Lambda_{n,\widetilde{R}}}(\mathbf{X},h) \longrightarrow \beta^\star(1 - F(\widetilde{R})) + \beta^{\star 2}\int_{B(o,\widetilde{R})} \bar{F}(o,v)\mathrm{d}v.$$

(ii) Nous vérifions que l'estimateur $\widehat{\sigma}_n^2$ défini en (2.3.2) tend presque sûrement vers σ^2 quand $n \to \infty$. D'après (2.3.2) et (2.3.3), il suffit de prouver que

$$|\Lambda_{n,\widetilde{R}}|^{-1} W_{\Lambda_{n,\widetilde{R}}} \xrightarrow{p.s} \int_{B(0,\widetilde{R})} \bar{F}(o,v)\mathrm{d}v.$$

Nous décomposons $W_{\Lambda_{n,\widetilde{R}}}$ comme $W^{(1)}_{\Lambda_{n,\widetilde{R}}} + W^{(2)}_{\Lambda_{n,\widetilde{R}}}$ où

$$W^{(1)}_{\Lambda_{n,\widetilde{R}}} = \int_{\Lambda_{n,\widetilde{R}_{\ominus\widetilde{R}}}} \int_{B(u,\widetilde{R})} \mathbb{1}(\mathbf{X} \cap B(u,\widetilde{R}) = \emptyset)\mathbb{1}(\mathbf{X} \cap B(v,\widetilde{R}) = \emptyset)\mathrm{d}u\mathrm{d}v$$

$$W^{(2)}_{\Lambda_{n,\widetilde{R}}} = \int_{\Lambda_{n,\widetilde{R}} \setminus \Lambda_{n,\widetilde{R}_{\ominus\widetilde{R}}}} \int_{B(u,\widetilde{R}) \cap \Lambda_{n,\widetilde{R}}} \mathbb{1}(\mathbf{X} \cap B(u,\widetilde{R}) = \emptyset)\mathbb{1}(\mathbf{X} \cap B(v,\widetilde{R}) = \emptyset)\mathrm{d}u\mathrm{d}v$$

avec $\Lambda_{n,\widetilde{R}_{\ominus\widetilde{R}}} = \{u \in \Lambda_{n,\widetilde{R}} : B(u,\widetilde{R}) \subseteq \Lambda_{n,\widetilde{R}}\}$. En ce qui concerne le premier terme, nous pouvons appliquer le théorème ergodique de Nguyen et Zessin [101] qui prouve que presque sûrement

$$|\Lambda_{n,\widetilde{R}_{\ominus\widetilde{R}}}|^{-1} W^{(1)}_{\Lambda_{n,\widetilde{R}}} \longrightarrow \mathrm{E} \int_{B(0,\widetilde{R})} \mathbb{1}(\mathbf{X} \cap B(o,\widetilde{R}) = \emptyset)\mathbb{1}(\mathbf{X} \cap B(v,\widetilde{R}) = \emptyset)\mathrm{d}v$$

sous l'hypothèse de stationnarité du processus ponctuel de Gibbs $\mathbf{X}$. Étant donné que $|\Lambda_{n,\widetilde{R}}| \sim |\Lambda_{n,\widetilde{R}_{\ominus\widetilde{R}}}|$ quand $n \to \infty$, la preuve est effectuée en vérifiant que

$$|\Lambda_{n,\widetilde{R}}|^{-1} W^{(2)}_{\Lambda_{n,\widetilde{R}}} \longrightarrow 0$$

qui découle du fait que

$$|\Lambda_{n,\widetilde{R}}|^{-1} W^{(2)}_{\Lambda_{n,\widetilde{R}}} \leq \frac{|\Lambda_{n,\widetilde{R}} \setminus \Lambda_{n,\widetilde{R}_{\ominus\widetilde{R}}}|}{|\Lambda_{n,\widetilde{R}}|} |B(u,\widetilde{R})| = o(1).$$

Le résultat de normalité asymptotique du Théorème 2.3.2 résulte du théorème de Slutsky, dans le cas ou $\mathbb{P}_{\beta^\star}$ est ergodique. Dans le cas où la mesure $\mathbb{P}_{\beta^\star}$ n'est pas ergodique elle peut être décomposée comme un mélange de mesures ergodiques par lequel le résultat (ii) reste vrai et puisque tout mélange de gaussiennes standard reste une gaussienne standard, le résultat (ii) reste vrai même si $\mathbb{P}_{\beta^\star}$ n'est pas ergodique. □

2.4 Etudes par simulation

Dans cette section, nous étudions l'efficacité de l'estimation de l'intensité de Poisson dans une étude par simulation. Nous considérons les modèles (i) - (v) décrits dans la Section 1.4.4.

Nous fixons $R = 0.05$ et $\beta^\star = 200$ pour tous les exemples, sauf pour le modèle `g2` où nous fixons $\beta^\star = 50$. Nous considérons

- Processus ponctuel de Strauss où les modèles `s1` et `s2` ont respectivement $\gamma = 0.2$ et $\gamma = 0.8$.
- Processus ponctuel de Strauss avec hard-core où $\delta = R/2$ et où les modèles `shc1` et `shc2` ont respectivement $\gamma = 0.2$ et $\gamma = 0.8$.
- Processus ponctuel de Strauss par morceaux où $p = 3$, $(R_1, R_2, R_3) = (R/3, 2R/3, R)$ et où les modèles `ps1` et `ps2` ont respectivement $(\gamma_1, \gamma_2, \gamma_3) = (0.8, 0.5, 0.2)$ et $(\gamma_1, \gamma_2, \gamma_3) = (0.2, 0.8, 0.2)$.
- Processus ponctuel de triplets où les modèles `t1` et `t2` ont respectivement $\gamma = 0.2$ et $\gamma = 0.8$.
- Processus ponctuel de saturation de Geyer avec un seuil de saturation $s = 1$ où les modèles `g1` et `g2` ont respectivement $\gamma = 0.5$ et $\gamma = 1.5$.

L'estimation de l'intensité de Poisson que nous proposons dépend du paramètre $\widetilde{R}$ et devrait fonctionner pour n'importe quelle valeur supérieure à R. Pour illustrer ceci, le paramètre $\beta^\star$ est estimé pour différentes valeurs de $\widetilde{R}$. Nous considérons $\widetilde{R} = R \times p$, où p est un paramètre variant de 0.9 à 1.2. Chaque sous-échantillon a été généré sur le domaine $[0, L]^2$ et les estimateurs sont calculés sur $[\widetilde{R}, L - \widetilde{R}]^2$. Pour illustrer la convergence quand la fenêtre grandit, les simulations sont effectuées avec $L = 1$ et $L = 2$. Les simulations sont effectuées en utilisant le package `R spatstat` de Baddeley et Turner [7]. Les résultats empiriques sont basés sur 500 estimations sont présentés dans le tableau 2.1 et dans les figures ci-dessous. Les résultats sont très satisfaisants. Nous observons que pour des valeurs de $\widetilde{R} \geq R$ (i.e. pour des valeurs égales ou plus grande que la portée finie), les estimateurs sont presque sans biais tandis qu'un biais non négligeable est observé lorsque $\widetilde{R} < R$. Le biais est négatif lorsque le processus ponctuel est répulsif (i.e. quand $\lambda_{\beta^\star}(u, \mathbf{x}) \geq 1$) et positif lorsque le processus ponctuel est attractif (i.e.quand $\lambda_{\beta^\star}(u, \mathbf{x}) \leq 1$) en particulier pour le modèle `g2`. Le signe du biais lorsque $\widetilde{R} \geq R$ peut être expliqué par le fait que $V_\Lambda(\mathbf{X}, \widetilde{R})$ est sous-estimé pour les processus ponctuels attractifs et surestimé pour les processus ponctuels répulsifs. Plus $\widetilde{R}$ est proche de R et plus les écarts-types sont faibles. Le nombre de points sans $\widetilde{R}$-proches voisins est toujours plus bas quand $\widetilde{R}$ grandit, ce qui explique cette augmentation de la variance. Enfin, nous pouvons également souligner que les résultats asymptotiques sont partiellement contrôlés. Lorsque la fenêtre varie de $[0, 1]^2$ à $[0, 2]^2$, les écarts-types des estimations sont divisés par 2 ce qui s'avère être la racine carrée des rapports des aires des domaines.

Le Tableau 2.1 et les figures 2.1, 2.2 et 2.3, soulignent en particulier les risques de choisir une valeur erronée pour le paramètre de portée finie R. D'un point de vue pratique, si ce paramètre n'est pas défini par l'utilisateur, il doit être estimé. Lorsque l'intensité conditionnelle Papangelou est entièrement modélisée, les paramètres peuvent

être estimés en utilisant par exemple le maximum de vraisemblance ou de la pseudo-vraisemblance (voir Møller et Waagepetersen [95]) et les paramètres irréguliers (tels que la portée finie R) peuvent être estimés à l'aide de la pseudo-vraisemblance [6]. Une telle approche ne peut pas être utilisée ici car nous ne voulons pas modéliser les termes d'interaction d'ordre supérieur. Nous avons remarqué que pour un modèle donné (répulsif) notre estimateur augmente avec la valeur de $\widetilde{R}$ jusqu'à une certaine valeur, où le régime à tendance à changer. Par conséquent, nous proposons d'estimer le paramètre de portée finie R, comme suit : Sélection d'une grille de valeurs raisonnable pour $\widetilde{R}$ (nous avons choisi $[0.02, 0.08]$ dans notre exemple). Calculer $\widehat{\beta}$ pour chaque valeur de la grille. Nous considérons une régression linéaire par morceaux de $\widehat{\beta}$ en fonction de $\widetilde{R}$ et nous définissons $\widehat{R}$ comme le point de rupture de la pente (pour cette étape, nous avons utilisé le package `R segmented` [97]). Les résultats empiriques pour les modèles `s1` et `s2` sont présentés dans le tableau 2.2. En utilisant la stratégie décrite ci-dessus, nous parvenons à estimer R de manière satisfaisante. Les estimateurs obtenus du paramètre $\beta^\star$ semblent être toujours sans biais. Nous notons que l'écart type est plus grand que dans le cas où R est exactement connu (colonne $p = 1$ dans le tableau 2.1). Toutefois, la perte d'efficacité est tout à fait raisonnable. Enfin, nous étudions les propriétés de l'estimateur de la variance asymptotique de $\widehat{\beta}$ donnée par (2.3.2) et le résultat de normalité asymptotique (2.3.1). Le Tableau 2.3 se concentre sur les modèles `s1` et `s2` et affiche les taux de couverture empiriques (c'est à dire le pourcentage d'intervalles de confiance asymptotiques de niveau 95% contenant la vraie valeur du paramètre $\beta^\star = 200$). Comme prévu, le taux de couverture théorique est respecté quelle que soit la valeur de $\widetilde{R} \geq R$, ce qui est loin d'être le cas lorsque $\widetilde{R} = 90\%R$, bien que la Proposition 2.3.1 ne couvre pas ce cas, il est intéressant de noter que les résultats empiriques sont encore tout à fait satisfaisants si l'on remplace R par sa valeur estimée.

		Moyennes (écarts-types.) pour $\widetilde{R} = R \times p$			
Modèle	$\overline{n}$	$p = 0.9$	$p = 1$	$p = 1.1$	$p = 1.2$
`s1` $L = 1$	99	174.0 (26.2)	203.6 (34.2)	204.3 (38.3)	205.5 (44.4)
$L = 2$	393	172.6 (12.7)	200.5 (16.4)	200.3 (18.5)	200.6 (20.8)
`s2` $L = 1$	156	192.0 (31.3)	203.3 (41.8)	203.9 (48.1)	204.7 (55.8)
$L = 2$	622	190.7 (16.0)	200.7 (19.5)	200.3 (22.8)	200.4 (26.5)
`shc1` $L = 1$	94	174.3 (23.2)	201.9 (30.3)	203.3 (36.8)	205.0 (41.9)
$L = 2$	379	174.1 (12.1)	201.6 (15.7)	201.9 (17.8)	202.1 (21.1)
`shc2` $L = 1$	130	194.3 (30.2)	205.4 (36.7)	207.0 (42.8)	209.1 (51.6)
$L = 2$	514	190.2 (15.2)	198.5 (17.8)	199.5 (21.1)	200.6 (24.4)
`ps1` $L = 1$	111	172.2 (25.1)	202.6 (33.2)	203.5 (38.7)	206.5 (47.2)
$L = 2$	445	171.5 (12.4)	201.6 (16.3)	201.7 (18.6)	202.4 (21.6)
`ps2` $L = 1$	134	191.1 (30.7)	201.7 (37.7)	206.2 (45.2)	208.8 (54.2)
$L = 2$	535	192.8 (14.5)	201.7 (17.5)	202.1 (20.6)	202.0 (24.4)
`t1` $L = 1$	159	181.6 (34.6)	204.5 (40.6)	205.2 (48.7)	205.8 (58.8)
$L = 2$	634	180.5 (16.7)	201.6 (19.5)	202.1 (22.9)	202.8 (26.7)
`t2` $L = 1$	178	188.7 (30.7)	200.7 (35.0)	202.4 (39.9)	205.1 (50.2)
$L = 2$	727	189.3 (17.5)	200.0 (20.2)	199.8 (24.5)	199.8 (30.8)
`g1` $L = 1$	110	174.8 (26.6)	201.4 (34.4)	201.5 (40.2)	203.6 (46.0)
$L = 2$	440	175.0 (13.3)	200.7 (17.4)	201.3 (19.9)	201.6 (23.1)
`g2` $L = 1$	70	54.5 (10.4)	50.0 (10.4)	50.2 (11.4)	50.5 (12.1)
$L = 2$	276	54.4 (5.4)	50.2 (5.4)	50.2 (5.8)	50.3 (6.3)

Table 2.1 – Moyennes et écarts-types des estimateurs du paramètre d'intensité de Poisson $\beta^\star$ pour différents paramètres $\widetilde{R}$ basés sur 500 estimations de différents modèles générés sur la fenêtre $[0, L]^2$ et estimés sur la fenêtre $[\widetilde{R}, L - \widetilde{R}]^2$ pour $L = 1, 2$. La deuxième colonne, $\overline{n}$, indique le nombre moyen de points (sur les 500 estimations) tombant dans $[0, L]^2$, i.e. l'estimation Monte-Carlo de $\mathrm{E}\, n(\mathbf{X} \cap [0, L]^2)$.

	$\widehat{R}$	$\widehat{\beta}$
	Moyennes (écarts-types.)	Moyennes (écarts-types.)
`s1` $L = 1$	0.052 (0.004)	197.7 (38.8)
$L = 2$	0.051 (0.003)	198.9 (22.1)
`s2` $L = 1$	0.051 (0.004)	201.7 (38.9)
$L = 2$	0.051 (0.003)	198.4 (21.9)

Table 2.2 – Moyenne et écarts-types des estimateurs du paramètre de portée finie et les estimateurs du paramètre d'intensité de Poisson associés (i.e. $\widehat{\beta}(\widehat{R})$), basés sur 500 estimations des modèles de Strauss observés sur la fenêtre $[0, L]^2$.

		$\widetilde{R} = p \times R$				$\widehat{R}$
		$p = 0.9$	$p = 1$	$p = 1.1$	$p = 1.2$	
`s1`	$L = 1$	77.4%	94.8%	95.0%	94.2%	89.2%
	$L = 2$	40.0%	94.8%	93.8%	95.4%	90.4%
`s2`	$L = 1$	82.6%	93.2%	94.6%	93.8%	93%
	$L = 2$	69.8%	95.0%	93.6%	92.8%	90.8%

Table 2.3 – Taux de couverture empiriques (i.e. le pourcentage des intervalles de confiance à 95% sur la vraie valeur du paramètre $\beta^\star = 200$) basés sur 500 estimations de modèles de Strauss observés sur la fenêtre $[0, L]^2$.

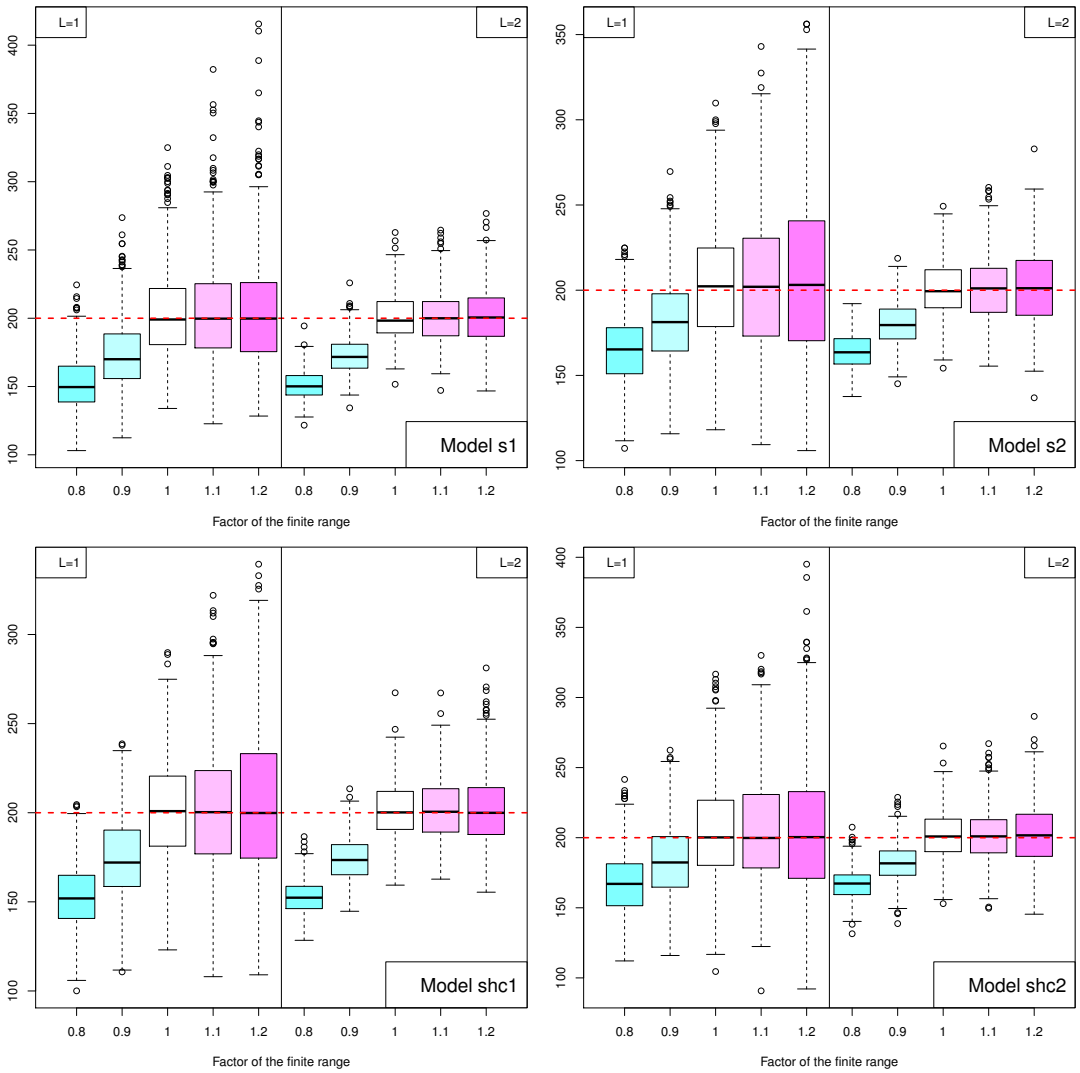


Figure 2.1 – Boxplots des estimateurs du paramètre d'intensité de Poisson pour différents paramètres bassés sur 500 estimations des modèles `s1`, `s2`, `shc1`, `shc2` observés sur la fenêtre $[0, L]^2$ et estimés sur la fenêtre $[\widetilde{R}, L - \widetilde{R}]^2$ pour $L = 1, 2$ et pour $\widetilde{R} = p.\, R$.

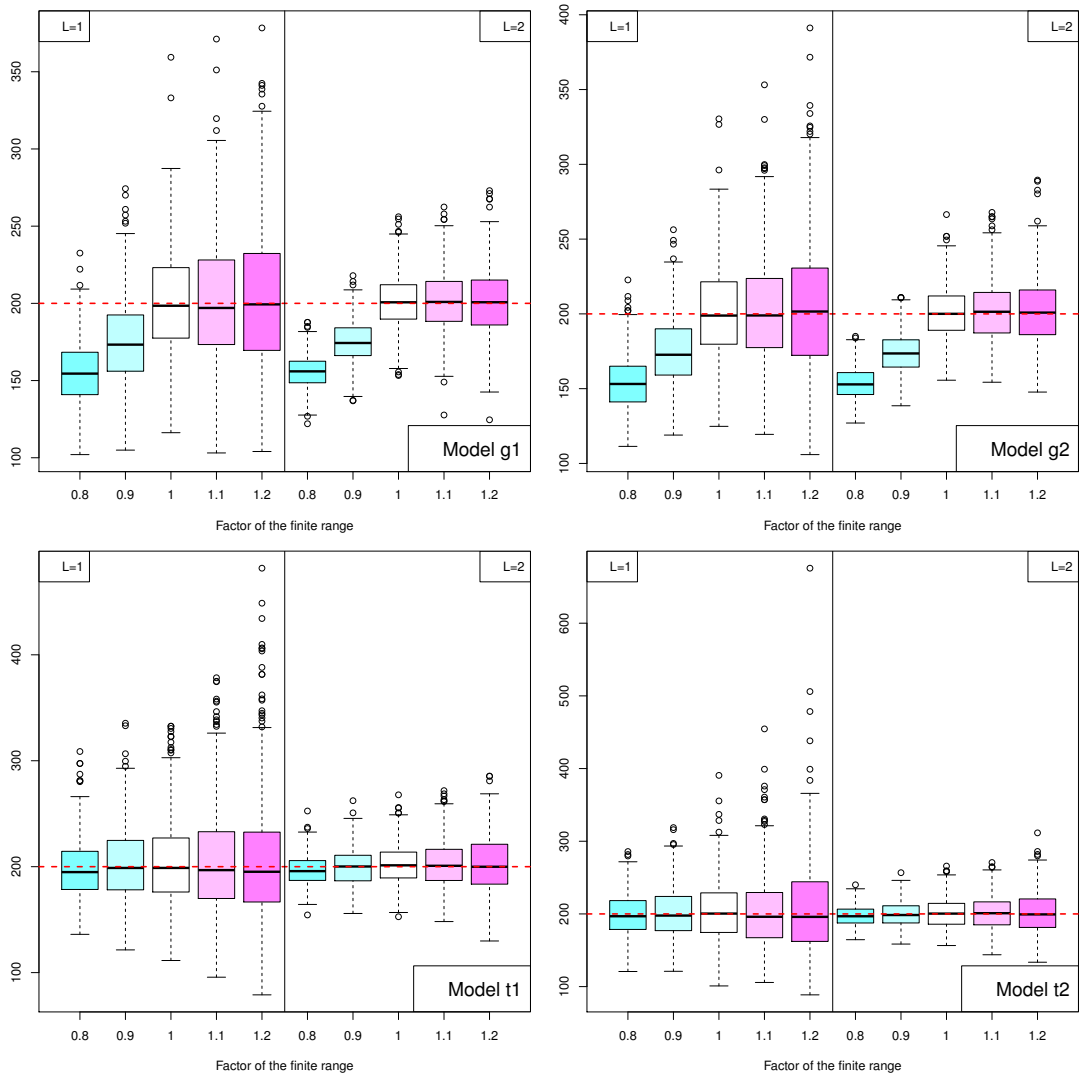


Figure 2.2 – Boxplots des estimateurs du paramètre d'intensité de Poisson pour différents paramètres bassés sur 500 estimations des modèles `g1`, `g2`, `t1`, `t2` observés sur la fenêtre $[0, L]^2$ et estimatés sur la fenêtre $[\widetilde{R}, L - \widetilde{R}]^2$ pour $L = 1, 2$ et pour $\widetilde{R} = p.\, R$.

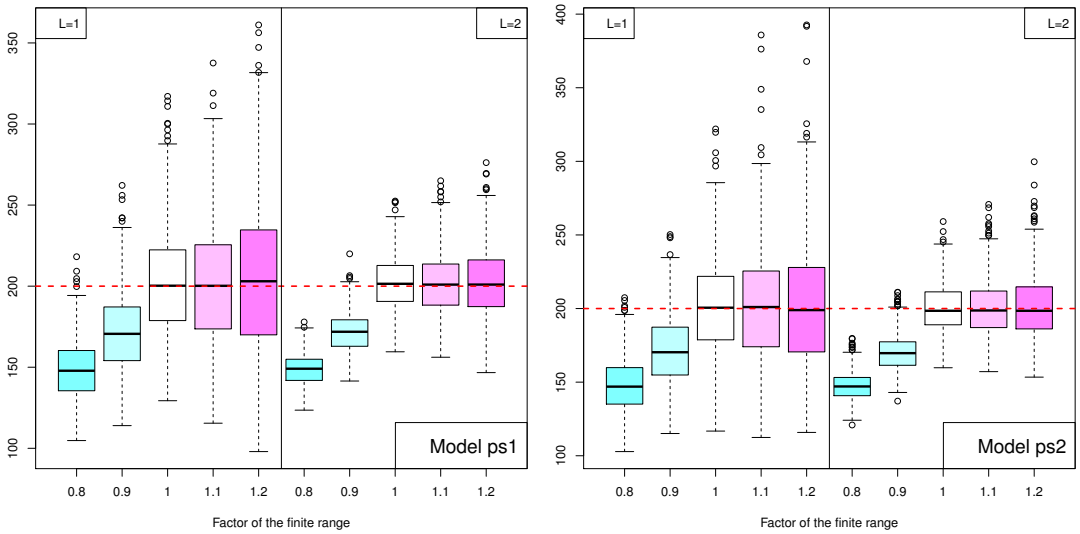


Figure 2.3 – Boxplots des estimateurs du paramètre d'intensité de Poisson pour différents paramètres bassés sur 500 estimations des modèles `ps1`, `ps2` observés sur la fenêtre $[0, L]^2$ et estimés sur la fenêtre $[\widetilde{R}, L - \widetilde{R}]^2$ pour $L = 1, 2$ et pour $\widetilde{R} = p.\,R$.

Chapitre 3

Estimation non paramétrique de fonction d'interaction de paires d'un processus ponctuel à interaction par paires stationnaire

Résumé

Nous nous intéresserons dans ce chapitre au processus ponctuel à interaction par paires. Nous supposons que le processus ponctuel étudié est stationnaire, une propriété qui facilite grandement l'estimation de ses paramètres. Nous construisons une nouvelle méthode d'estimation non paramétrique de la fonction d'interaction de paires à partir d'une réalisation du processus ponctuel spatial à interaction par paires stationnaire. Nous présentons dans un premier temps le modèle que nous étudions, nous donnons des hypothèses techniques sur ce modèle et nous définissons ensuite l'expression de l'estimateur que nous considérons dans la Section 3.1. Dans la Section 3.2, nous étudions les comportements asymptotiques de la moyenne et de la variance de l'estimateur non paramétrique par lissage qui nous intéresse. Dans la Section 3.3, nos résultats portent sur la consistance ponctuelle de nos estimateurs. La Section 3.4 est consacrée à l'étude des fonctions de Young et nous juxtaposons des résultats sur les inégalités de Kahane-Khintchine où apparaissent les espaces d'Orlicz associés aux fonctions de Young. Dans la Section 3.5 et la Section 3.6, nos résultats portent sur la consistance forte ponctuelle et globale des estimateurs obtenus avec des différentes vitesses de consistance, sous certaines hypothèses d'appartenance aux espaces d'Orlicz induits par des fonctions de Young. Nos résultats portent aussi grâce à des cadres particuliers les plus importants des espaces d'Orlicz ; les champs aléatoires p-intégrables et les champs aléatoires bornés. Nous évoquerons ensuite la normalité asymptotique de l'estimateur de la fonction d'interaction de paires dans la Section 3.7.

3.1 Hypothèses et présentation de l'estimateur

Pour plus de commodité, nous commencerons par rappeler les définitions de base du processus ponctuel à interaction par paires stationnaire. Le Processus ponctuel à interaction par paires est caractérisé par son intensité conditionnelle de Papangelou définie par

$$\lambda(u, \mathbf{x}) = g_0(u) \exp\left(-\sum_{v \in \mathbf{x} \setminus u} g_0(\{u, v\})\right),$$

processus ponctuel à interaction par paires est stationnaire, si $g_0(u) = \beta$ est une constante et $g_0(\{u, v\}) = g(v - u)$ est invariante par translation. Autrement dit, pour tout $u \in \mathbb{R}^d$, $\mathbf{x} \in \Omega$ (Ω l'espace des configurations localement finies) l'intensité conditionnelle de Papangelou du modèle considéré s'écrit

$$\lambda_{\beta^\star}(u, \mathbf{x}) = \beta^\star \exp\left(-\sum_{v \in \mathbf{x} \setminus u} g(v - u)\right) \tag{3.1.1}$$

avec $\beta^\star$ est la vraie valeur du paramètre d'intensité de Poisson, g représente le potentiel d'interaction défini sur $\mathbb{R}^d$ à valeur dans $\mathbb{R}$ et $G = \exp(-g)$ représente la fonction d'interaction de paires. Dans ce modèle semi-paramétrique (3.1.1), l'estimation de $\beta^\star$ a représenté la première étape de notre procédure dans le Chapitre 2. Nous souhaitons construire un estimateur non paramétrique de la fonction d'interaction de paires $G(t) = \exp(-g(t))$ sous des hypothèses aussi générales que possible.

Nos principales hypothèses sont les suivantes : Nous supposons que l'intensité conditionnelle de Papangelou satisfait la propriété suivante :

$$\lambda_{\beta^\star}(u, \mathbf{x}) = \lambda_{\beta^\star}(u, \mathbf{x}_{B(u,R)}). \tag{3.1.2}$$

Généralement pour les modèles de Gibbs utilisés dans la pratique, le potentiel g est supposé une fonction positive, ce qui implique la stabilité locale (et donc intégrabilité). Cela signifie que les points du processus se repoussent les uns les autres, de sorte que les réalisations du processus ont tendance à être plus régulières que pour un processus de Poisson.

Nous supposons que nous observons la réalisation du processus ponctuel à interaction par paires stationnaire $\mathbf{X}$ dans un domaine borné de $\mathbb{R}^d$ noté Λ_n où la suite $(\Lambda_n)_{n \geq 1}$ tend vers $\mathbb{R}^d$ quand n tend vers l'infini. Pour simplifier, supposons que $\Lambda_n = [-n, n]^d$ soit la suite des domaines d'observation. Nous allons introduire un estimateur de la fonction d'interaction de paires $G = \exp(-g)$ par la méthode de lissage,

cette méthode nous permet de donner un cadre général à l'estimation de la fonction d'interaction de paires. Cette technique et des exemples d'application ont été débattus par Baddeley et al. [8] et par Gaetan et Guyon [50] pour les champs des résidus lissés. A cet effet, nous avons choisi un noyau K une fonction mesurable définie sur $\mathbb{R}^d$ à valeur dans $\mathbb{R}$, associée à une suite de réels positifs $(b_n)_{n\geq 1}$ appelée paramètre de lissage. Nous supposons que le noyau K satisfait certain nombre de conditions qui sont considérées comme usuelles pour les noyaux et qui permettent d'étudier le biais, la variance et les résultats asymptotiques de l'estimateur qui en résulte.

Condition $\mathbf{K^0}$

(i) $\int_{\mathbb{R}^d} K(z)\mathrm{d}z = 1$,
(ii) $K(z) = 0$, pour tout $z \in \mathbb{R}^d$ avec $||z|| > R_1$ et pour un certain $R_1 \in \mathbb{R}_+$,
(iii) la fonction K est non négative et bornée,
(iv) pour $m \in \mathbb{N}^\star$

$$\int_{\mathbb{R}^d} z_1^{a_1} ... z_d^{a_d} K(z)\mathrm{d}z = 0$$

avec $a_1, ..., a_d \in \mathbb{N}^\star$ tels que $0 < \sum_{i=1}^d a_i < m$.

Nous trouvons dans la littérature des conditions sur la fonction du noyau analogues également nécessaires pour l'estimation de la densité de probabilité, voir Nadaraya [99] et Horváth [73], l'esimation de la fonction de corrélation de paires, voir Heinrich et Liebscher [72]. Nous trouverons des exemples de noyaux usuels dans l'article de Deheuvels [38] et David [35]. Nous avons besoin aussi d'imposer la condition suivante tout au long de ce chapitre sur la suite $(b_n)_{n\geq 1}$:

$$\lim_{n\to\infty} b_n = 0 \quad \text{et} \quad \lim_{n\to\infty} b_n^d |\Lambda_{n,R}| = \infty.$$

Rappelons la notion déjà utilisée au chapitre précédent, la fonction mesurable non négative h définie pour tout $u \in \mathbb{R}^d$, $\mathbf{x} \in \Omega$, par

$$h(u, \mathbf{x}) = \mathbb{1}(\mathbf{x} \cap B(u, R) = \emptyset). \tag{3.1.3}$$

Soit $T = \{t \in \mathbb{R}^d; g(t) > 0, \text{ pour } ||t|| < R\}$ le support de la fonction g. Il nous faut supposer certaines conditions de régularité :

Condition $\boldsymbol{\mathcal{P}(t)}$

$$\bar{F}(o, t) = \mathrm{E}[h(o, \mathbf{X})h(t, \mathbf{X})] = \mathbb{P}(\mathbf{X} \cap B(o, R) = \emptyset, \mathbf{X} \cap B(t, R) = \emptyset) \tag{3.1.4}$$

et $G(t) = \exp(-g(t))$ ou g sont des fonctions continues en $t \in T$.

Avec la convention $c/0 = 0$, pour tout réel c. Nous prenons en compte les problèmes d'effet de bord, par l'introduction du $2R$-intérieur de Λ_n défini par

$$\Lambda_{n,R} = \{u \in \Lambda_n : B(u, 2R) \subset \Lambda_n\}, \quad \text{avec} \quad |\Lambda_{n,R}| > 0.$$

Le choix de $\Lambda_{n,R}$ est induit par le fait que lorsque $v \in B(u, R)$ et $u \in \Lambda_{n,R}$, alors $h(u, \mathbf{X} \setminus \{u, v\})$ et $h(v, \mathbf{X} \setminus \{u, v\})$ sont calculabes à partir de $\mathbf{X}_{\Lambda_n}$. Dans ce chapitre et les suivants, nous supposons, pour simplifier la présentation que la porté R est connue. En pratique, estimer cette portée comme proposé dans le chapitre précédent est une piste raisonnable. Nous définissons l'estimateur de $\beta^\star G(t)$ par

$$\widehat{G}_n(t) = \frac{\widehat{H}_n(t)}{\widehat{L}_n(t)}, \tag{3.1.5}$$

où

(3.1.6)

$$\widehat{H}_n(t) = \frac{1}{b_n^d |\Lambda_{n,R}|} \sum_{\substack{u,v \in \mathbf{X} \\ v-u \in B(o,R)}}^{\neq} \mathbb{1}_{\Lambda_{n,R}}(u) h(u, \mathbf{X} \setminus \{u, v\}) h(v, \mathbf{X} \setminus \{u, v\}) K\left(\frac{v-u-t}{b_n}\right)$$

est un estimateur de type noyau estimant $\beta^{\star 2} \exp(-g(t)) \bar{F}(o, t)$ avec $\bar{F}(o, t) = \mathbb{P}(\mathbf{X} \cap B(o, \widetilde{R}) = \emptyset, \mathbf{X} \cap B(t, \widetilde{R}) = \emptyset)$.

$$\widehat{L}_n(t) = \frac{1}{|\Lambda_{n,R}|} \sum_{u \in \mathbf{X}_{\Lambda_{n,R}}} h(u, \mathbf{X} \setminus \{u\}) h(t-u, \mathbf{X} \setminus \{u\}) \tag{3.1.7}$$

étant quand à lui un estimateur de $\beta^\star \bar{F}(o, t)$.
Dans les sections qui suivent, nous allons développer certains résultats asymptotiques de statistique non paramétrique de notre modèle. Nous présentons quelques propriétés asymptotiques des estimateurs $\widehat{H}_n(t)$ et $\widehat{L}_n(t)$, que nous utilisons plus tard pour obtenir des résultats asymptotiques pour l'estimateur $\widehat{G}_n(t)$. Nous présentons d'abord les comportements asymptotiques de la moyenne et de la variance de l'estimateur $\widehat{H}_n(t)$, des résultats asymptotiques ponctuels puis des résultats asymptotiques globaux.

3.2 Comportements asymptotiques de la moyenne et de la variance de l'estimateur $\widehat{H}_n(t)$

Nous débutons l'étude de l'estimation non paramétrique de la fonction d'interaction de paires par les comportements asymptotiques de la moyenne et la variance de l'estimateur par lissage $\widehat{H}_n(t)$. Avant d'entreprendre cette étude, nous commençons par rappeler bièvement quelques définitions qui seront utilisées.

Soit $(Z_n)_{n\geq 1}$ une suite de variables aléatoires réelles et $(v_n)_{n\geq 1}$ une suite de nombres réels positifs. On dit que

$$Z_n = \mathcal{O}_{p.s.}(v_n)$$

s'il existe $b > 0$ tel que

$$\limsup_{n\to\infty} \frac{|Z_n|}{v_n} \leq b \quad \text{p.s.}$$

et on dit que

$$Z_n = o_{p.s}(v_n)$$

si

$$\lim_{n\to\infty} \frac{|Z_n|}{v_n} = 0 \quad \text{p.s.}$$

Proposition 3.2.1. *Soit g le potentiel de paires invariant par translation défini en (3.1.1). Sous la condition de portée d'interaction de paires finie (3.1.2) et si les trois premières conditions de $\mathbf{K}^0$ sont remplies. On a*

$$\lim_{n\to\infty} \mathrm{E}\,\widehat{H}_n(t) = \beta^{\star 2}\exp(-g(t))\bar{F}(o,t)$$

en tout point de continuité $t \in T$ des fonctions g et $\bar{F}$.
De plus si la quatrième condition de $\mathbf{K}^0$ est satisfaite et si la fonction $\exp(-g(t))\bar{F}(o,t)$ admet des dérivées partielles d'ordre m continues et bornées sur la boule ouverte $B^o(t,\delta)$, pour $t \in \mathring{T}$, pour un certain $\delta > 0$ et pour un certain $m \geq 1$. Alors, on a

$$\mathrm{E}\,\widehat{H}_n(t) = \beta^{\star 2}\exp(-g(t))\bar{F}(o,t) + \mathcal{O}(b_n^m) \quad \textit{quand} \quad n \to \infty. \tag{3.2.1}$$

Théorème 3.2.2. *Soit g le potentiel de paires invariant par translation défini en (3.1.1). Sous la condition de portée d'interaction de paires finie (3.1.2) et sous les trois premières conditions de $\mathbf{K}^0$, telles que $\int_{\mathbb{R}^d} K^2(z)\mathrm{d}z < \infty$. Alors, on a*

$$\lim_{n\to\infty} b_n^d|\Lambda_{n,R}|\,\mathrm{Var}\left(\widehat{H}_n(t)\right) = \beta^{\star 2}\exp(-g(t))\bar{F}(o,t)\int_{\mathbb{R}^d} K^2(z)\mathrm{d}z$$

en tout point de continuité $t \in T\backslash\{o\}$ des fonctions g et $\bar{F}$.

Théorème 3.2.3. *Sous les mêmes conditions du Théorème 3.2.2, soient $t_1 \neq t_2 \in T$. Alors, on a*

$$\lim_{n\to\infty} b_n^d|\Lambda_{n,R}|\,\mathrm{Cov}\left(\widehat{H}_n(t_1), \widehat{H}_n(t_2)\right) = 0.$$

La preuve du Théorème 3.2.3 est analogue à celle du Théorème 3.2.2 et est omise.
La Proposition 3.2.1 et le Théorème 3.2.2 ci-dessus fournissent une bande asymptotique de

$$\mathrm{EQM}\left(\widehat{H}_n(t)\right) = \mathrm{Var}\left(\widehat{H}_n(t)\right) + \left(Biais(\widehat{H}_n(t)\right)^2$$

à partir de laquelle une bande asymptotique optimale peut être déduite.

Corollaire 3.2.4. Sous les conditions de la Proposition 3.2.1 et du Théorème 3.2.2. Alors, on a

$$\mathrm{EQM}(\widehat{H}_n(t)) = o\big(\frac{1}{b_n^d|\Lambda_{n,R}|}\big) + \mathcal{O}\big(b_n^{2m}\big).$$

3.2.1 Preuve de la Proposition 3.2.1

Démonstration. Il est commode dans les preuves de ce chapitre de noter

$$H(u_1, ..., u_s, \mathbf{X}) = h(u_1, \mathbf{X})...h(u_s, \mathbf{X})$$

et

$$I(u_1, ..., u_s) = \mathbb{1}(u_1 \in B(o, R), ..., u_s \in B(o, R)).$$

Gardons à l'esprit la fonction $\bar{F}(o, t)$ définie en (3.1.4). Pour démontrer la première partie de la Proposition 3.2.1, soient alors les points $u, v \in \mathbb{R}^d$ et $\mathbf{x} \in \Omega$, nous définissons l'intensité conditionnelle de Papangelou du second ordre par

$$\lambda_{\beta^\star}(u, v, \mathbf{x}) = \lambda_{\beta^\star}(u, \mathbf{x})\lambda_{\beta^\star}(v, \mathbf{x} \cup \{u\}).$$

Il suffit d'utiliser la formule de GNZ (1.4.3) au second ordre, i.e.

$$\mathrm{E}\sum_{u,v,\in\mathbf{X}}^{\neq} h(u, v, \mathbf{X} \setminus \{u, v\}) = \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} \mathrm{E}\, h(u, v, \mathbf{X})\lambda(u, v, \mathbf{X})\mathrm{d}u\mathrm{d}v$$

pour

$$h(u, v, \mathbf{X}) = \mathbb{1}_{\Lambda_{n,R}}(u)I(v - u)H(u, v, \mathbf{X})K\left(\frac{v - u - t}{b_n}\right)$$

et la propriété de portée d'interaction finie (3.1.2) pour chaque fonction $\lambda_{\beta^\star}(u, \mathbf{x})$ et $\lambda_{\beta^\star}(v, \mathbf{x} \cup \{u\})$, il résulte que

$$\begin{aligned}\mathrm{E}\,\widehat{H}_n(t) &= \frac{1}{b_n^d|\Lambda_{n,R}|}\,\mathrm{E}\int_{\mathbb{R}^{2d}} \mathbb{1}_{\Lambda_{n,R}}(u)I(v - u)H(u, v, \mathbf{X})K\left(\frac{v - u - t}{b_n}\right)\lambda_{\beta^\star}(u, v, \mathbf{X})\mathrm{d}u\mathrm{d}v \\ &= \frac{\beta^{\star 2}}{b_n^d|\Lambda_{n,R}|}\,\mathrm{E}\int_{\mathbb{R}^{2d}} \mathbb{1}_{\Lambda_{n,R}}(u)I(v - u)H(u, v, \mathbf{X})K\left(\frac{v - u - t}{b_n}\right)G(v - u)\mathrm{d}u\mathrm{d}v.\end{aligned}$$

D'après la stationnarité de $\mathbf{X}$, nous déduisons

$$\begin{aligned}\mathrm{E}\,\widehat{H}_n(t) &= \frac{\beta^{\star 2}}{b_n^d}\int_{\mathbb{R}^d} I(s)\bar{F}(o, s)K\left(\frac{s - t}{b_n}\right)G(s)\mathrm{d}s \\ &= \beta^{\star 2}\int_{\mathbb{R}^d} I(b_n z + t)\bar{F}(o, b_n z + t)K(z)G(b_n z + t)\mathrm{d}z.\end{aligned}$$

Nous appliquons le théorème de convergence dominée et sous les trois premières conditions de $\mathbf{K}^0$, nous déduisons pour $t \in T$, le résultat recherché du premier énoncé de la Proposition 3.2.1.

Pour obtenir les résultats de la seconde partie de la Proposition 3.2.1, nous considérons $B^0(t,\delta)$ une boule ouverte de $\mathbb{R}^d$, $\exp(-g(t)\bar{F}(o,t))$ admet des dérivées partielles d'ordre m continues et bornées en chaque point de $B^0(t,\delta), t \in \mathring{T}$, alors pour tout point z de $\mathbb{R}^d$, il existe $\theta \in]0,1[$, tel que nous avons la formule de Taylor-Lagrange

$$G(t+b_n z) = G(t) + \sum_{k=1}^{m-1} \sum_{|\alpha|=k} \frac{(b_n z)^\alpha}{\alpha!} \frac{\partial^k G(t)}{\partial x^\alpha} + b_n^m R_m(z,t)$$

avec $R_m(z,t) = \sum_{|\alpha|=m} \frac{z^\alpha}{\alpha!} \frac{\partial^m G}{\partial x^\alpha}(t+\theta b_n z)$.

$$\bar{F}(o,t+b_n z) = \bar{F}(o,t) + \sum_{k=1}^{m-1} \sum_{|\alpha|=k} \frac{(b_n z)^\alpha}{\alpha!} \frac{\partial^k \bar{F}(o,t)}{\partial x^\alpha} + b_n^m R'_m(z,t)$$

avec $R'_m(z,t) = \sum_{|\alpha|=m} \frac{z^\alpha}{\alpha!} \frac{\partial^m \bar{F}}{\partial x^\alpha}(o,t+\theta b_n z)$.
Le produit admet un développement limité d'ordre m dont la partie régulière s'obtient en prenant dans le produit des parties régulières les termes de degré inférieure ou égale à m. Le produit $\bar{F}(o,b_n z+t)G(b_n z+t)$ se mettra donc sous la forme

$$\bar{F}(o,b_n z+t)G(b_n z+t) = \bar{F}(o,t)G(t) + \sum_{k=1}^{\alpha-1} T_n(t)(b_n z)^k + \mathcal{O}(b_n^m) \quad \text{quand } n \to \infty.$$

$$\begin{aligned} \mathrm{E}\,\widehat{H}_n(t) &= \beta^{\star 2} G(t)\bar{F}(o,t) \int_{\mathbb{R}^d} K(z)\mathrm{d}z \\ &+ \beta^{\star 2} \sum_{k=1}^{\alpha-1} T_n(t) b_n^k \int_{\mathbb{R}^d} z^k K(z)\mathrm{d}z \\ &+ \mathcal{O}(b_n^m) \quad \text{quand} \quad n \to \infty \end{aligned}$$

sous la condition $\mathbf{K}^0$, nous concluons le taux demandé de convergence. □

3.2.2 Preuve du Théorème 3.2.2

Démonstration. Nous déterminons le comportement asymptotique de la variance de l'estimateur $\widehat{H}_n(t)$. Pour cela nous nous servons du résultat suivant.

Lemme 3.2.5. *Soit* $\mathbf{X}$ *un processus ponctuel de Gibbs, son intensité conditionnelle de Papangelou* λ. *Pour toute fonction mesurable non négative* $f : \mathbb{R}^d \times \mathbb{R}^d \times \Omega \longrightarrow \mathbb{R}$,

alors

$$
\begin{aligned}
&\mathrm{Var}\Big(\sum_{u,v\in\mathbf{X}}^{\neq} f(u,v,\mathbf{X}\backslash\{u,v\})\Big)\\
&= \mathrm{E}\int_{\mathbb{R}^{2d}} f(u,v,\mathbf{X})\big[f(u,v,\mathbf{X})+f(v,u,\mathbf{X})\big]\lambda(u,v,\mathbf{X})\mathrm{d}u\mathrm{d}v\\
&+\mathrm{E}\int_{\mathbb{R}^{3d}} f(u,v,\mathbf{X})\big[f(v,w,\mathbf{X})+f(w,v,\mathbf{X})\big]\lambda(u,v,w,\mathbf{X})\mathrm{d}u\mathrm{d}v\mathrm{d}w\\
&+\mathrm{E}\int_{\mathbb{R}^{3d}} f(u,v,\mathbf{X})\big[f(u,w,\mathbf{X})+f(w,u,\mathbf{X})\big]\lambda(u,v,w,\mathbf{X})\mathrm{d}u\mathrm{d}v\mathrm{d}w\\
&+\mathrm{E}\int_{\mathbb{R}^{4d}} f(u,v,\mathbf{X})f(w,y,\mathbf{X})\lambda(u,v,w,y,\mathbf{X})\mathrm{d}u\mathrm{d}v\mathrm{d}w\mathrm{d}y\\
&-\int_{\mathbb{R}^{4d}} \mathrm{E}\left[f(u,v,\mathbf{X})\lambda(u,v,\mathbf{X})\right]\mathrm{E}\left[f(w,y,\mathbf{X})\lambda(w,y,\mathbf{X})\right]\mathrm{d}u\mathrm{d}v\mathrm{d}w\mathrm{d}y.
\end{aligned}
$$

Démonstration. Après un calcul simple, mais prolixe (voir Jolivet [80] et Heinrich [71]), nous obtenons la décomposition suivante :

$$
\begin{aligned}
\Big(\sum_{u,v\in\mathbf{X}}^{\neq} f(u,v,\mathbf{X}\backslash\{u,v\})\Big)^2 &= \sum_{u,v\in\mathbf{X}}^{\neq} f^2(u,v,\mathbf{X}\backslash\{u,v\})\\
&+\sum_{u,v\in\mathbf{X}}^{\neq} f(u,v,\mathbf{X}\backslash\{u,v\})f(v,u,\mathbf{X}\backslash\{u,v\})\\
&+\sum_{u,v,w\in\mathbf{X}}^{\neq} f(u,v,\mathbf{X}\backslash\{u,v,w\})f(v,w,\mathbf{X}\backslash\{u,v,w\})\\
&+\sum_{u,v,w\in\mathbf{X}}^{\neq} f(u,v,\mathbf{X}\backslash\{u,v,w\})f(w,v,\mathbf{X}\backslash\{u,v,w\})\\
&+\sum_{u,v,w\in\mathbf{X}}^{\neq} f(u,v,\mathbf{X}\backslash\{u,v,w\})f(u,w,\mathbf{X}\backslash\{u,v,w\})\\
&+\sum_{u,v,w\in\mathbf{X}}^{\neq} f(u,v,\mathbf{X}\backslash\{u,v,w\})f(w,u,\mathbf{X}\backslash\{u,v,w\})\\
&+\sum_{u,v,w,y\in\mathbf{X}}^{\neq} f(u,v,\mathbf{X}\backslash\{u,v,w,y\})f(w,y,\mathbf{X}\backslash\{u,v,w,y\}). \qquad (3.2.2)
\end{aligned}
$$

D'autre part nous avons
(3.2.3)
$$\mathrm{Var}\sum_{u,v\in\mathbf{X}}^{\neq} f(u,v,\mathbf{X}\backslash\{u,v\}) = \mathrm{E}\Big(\sum_{u,v\in\mathbf{X}}^{\neq} f(u,v,\mathbf{X}\backslash\{u,v\})\Big)^2 - \Big(\mathrm{E}\sum_{u,v\in\mathbf{X}}^{\neq} f(u,v,\mathbf{X}\backslash\{u,v\})\Big)^2.$$

Après réarrangement avec la formule de GNZ (1.4.3) et les formules (3.2.2) et (3.2.3), nous obtenons le résultat voulu. □

En combinant le Lemme 3.2.5 avec

$$f(u,v,\mathbf{x}) = \mathbb{1}_{\Lambda_{n,R}}(u) I(v-u) H(u,v,\mathbf{x}) K\Big(\frac{v-u-t}{b_n}\Big)$$

ceci nous amène à obtenir

$$\mathrm{Var}\,\widehat{H}_n(t) = T_1 + T_2 + T_3 + T_4 + T_5 + T_6 + T_7 + T_8,$$

avec

$$\begin{aligned}
T_1 =& \frac{1}{b_n^{2d}|\Lambda_{n,R}|^2}\,\mathrm{E}\int_{\mathbb{R}^{2d}} \mathbb{1}_{\Lambda_{n,R}}(u) I(v-u) H(u,v,\mathbf{X}) K^2\Big(\frac{v-u-t}{b_n}\Big)\lambda_{\beta^\star}(u,v,\mathbf{X})\mathrm{d}u\mathrm{d}v,\\
T_2 =& \frac{1}{b_n^{2d}|\Lambda_{n,R}|^2}\\
&\mathrm{E}\int_{\Lambda_{n,R}^2} I(v-u,u-v) H(u,v,\mathbf{X}) K\Big(\frac{v-u-t}{b_n}\Big) K\Big(\frac{u-v-t}{b_n}\Big)\lambda_{\beta^\star}(u,v,\mathbf{X})\mathrm{d}u\mathrm{d}v,\\
T_3 =& \frac{1}{b_n^{2d}|\Lambda_{n,R}|^2}\,\mathrm{E}\int_{\mathbb{R}^{2d}}\int_{\Lambda_{n,R}} I(v-u,w-v) H(u,v,w,\mathbf{X}) K\Big(\frac{v-u-t}{b_n}\Big) K\Big(\frac{w-v-t}{b_n}\Big)\\
&\times \lambda_{\beta^\star}(u,v,w,\mathbf{X})\mathrm{d}u\mathrm{d}v\mathrm{d}w,\\
T_4 =& \frac{1}{b_n^{2d}|\Lambda_{n,R}|^2}\,\mathrm{E}\int_{\mathbb{R}^{d}}\int_{\Lambda_{n,R}^2} I(v-u,v-w) H(u,v,w,\mathbf{X}) K\Big(\frac{v-u-t}{b_n}\Big) K\Big(\frac{v-w-t}{b_n}\Big)\\
&\times \lambda_{\beta^\star}(u,v,w,\mathbf{X})\mathrm{d}u\mathrm{d}v\mathrm{d}w,\\
T_5 =& \frac{1}{b_n^{2d}|\Lambda_{n,R}|^2}\,\mathrm{E}\int_{\mathbb{R}^{2d}}\int_{\Lambda_{n,R}} I(v-u,w-u) H(u,v,w,\mathbf{X}) K\Big(\frac{v-u-t}{b_n}\Big) K\Big(\frac{w-u-t}{b_n}\Big)\\
&\times \lambda_{\beta^\star}(u,v,w,\mathbf{X})\mathrm{d}u\mathrm{d}v\mathrm{d}w,\\
T_6 =& \frac{1}{b_n^{2d}|\Lambda_{n,R}|^2}\,\mathrm{E}\int_{\mathbb{R}^{d}}\int_{\Lambda_{n,R}^2} I(v-u,u-w) H(u,v,w,\mathbf{X}) K\Big(\frac{v-u-t}{b_n}\Big) K\Big(\frac{u-w-t}{b_n}\Big)\\
&\times \lambda_{\beta^\star}(u,v,w,\mathbf{X})\mathrm{d}u\mathrm{d}v\mathrm{d}w,
\end{aligned}$$

$$T_7 = \frac{1}{b_n^{2d}|\Lambda_{n,R}|^2} \mathrm{E} \int_{\mathbb{R}^{4d}} \mathbb{1}_{\Lambda_{n,R}}(u) I(v-u, y-w) H(u,v,w,y,\mathbf{X}) K\left(\frac{v-u-t}{b_n}\right) K\left(\frac{y-w-t}{b_n}\right) \times \lambda_{\beta^\star}(u,v,w,y,\mathbf{X}) \mathrm{d}u\mathrm{d}v\mathrm{d}w\mathrm{d}y,$$

$$T_8 = -\frac{1}{b_n^{2d}|\Lambda_{n,R}|^2} \int_{\mathbb{R}^{4d}} \mathbb{1}_{\Lambda_{n,R}}(u) \mathbb{1}_{\Lambda_{n,R}}(w) I(v-u, y-w) K\left(\frac{v-u-t}{b_n}\right) K\left(\frac{y-w-t}{b_n}\right) \times \mathrm{E}\left[H(u,v,\mathbf{X})\lambda_{\beta^\star}(u,v,\mathbf{X})\right] \mathrm{E}\left[H(w,y,\mathbf{X})\lambda_{\beta^\star}(w,y,\mathbf{X})\right] \mathrm{d}u\mathrm{d}v\mathrm{d}w\mathrm{d}y.$$

Nous étudions maintenant séparément le comportement asymptotique des termes $T_1, T_2, T_3, T_4, T_5, T_6, T_7$ et T_8.

Pour le comportement asymptotique du premier terme T_1, il nous suffit d'appliquer la propriété de portée d'interaction finie (3.1.2) et la stationnarité de $\mathbf{X}$, il résulte

$$\begin{aligned} T_1 &= \frac{1}{b_n^{2d}|\Lambda_{n,R}|^2} \mathrm{E} \int_{\mathbb{R}^{2d}} \mathbb{1}_{\Lambda_{n,R}}(u) I(v-u) H(u,v,\mathbf{X}) K^2\left(\frac{v-u-t}{b_n}\right) \lambda_{\beta^\star}(u,v,\mathbf{X}) \mathrm{d}u\mathrm{d}v \\ &= \frac{\beta^{\star 2}}{b_n^{2d}|\Lambda_{n,R}|^2} \int_{\mathbb{R}^{2d}} \mathbb{1}_{\Lambda_{n,R}}(u) I(v-u) \bar{F}(u,v) K^2\left(\frac{v-u-t}{b_n}\right) G(v-u) \mathrm{d}u\mathrm{d}v \\ &= \frac{\beta^{\star 2}}{b_n^{2d}|\Lambda_{n,R}|} \int_{\mathbb{R}^d} I(s) \bar{F}(o,s) K^2\left(\frac{s-t}{b_n}\right) G(s) \mathrm{d}s \\ &= \frac{\beta^{\star 2}}{b_n^{d}|\Lambda_{n,R}|} \int_{\mathbb{R}^d} I(b_n z + t) \bar{F}(o, b_n z + t) K^2(z) G(b_n z + t) \mathrm{d}z. \end{aligned}$$

Nous appliquons le théorème de convergence dominée, nous obtenons alors pour tout $t \in T$

$$\lim_{n\to\infty} b_n^d |\Lambda_{n,R}| T_1 = \beta^{\star 2} \bar{F}(o,t) G(t) \int_{\mathbb{R}^d} K^2(z) \mathrm{d}z.$$

Remarque 3.2.6. Nous appliquons dans la suite, le lemme suivant, établi par (David [35], Lemme A.2), une propriété de la séquence de fenêtres d'observation $(\Lambda_n)_{n\in\mathbb{N}}$ pour obtenir des résultats asymptotiques.

Lemme 3.2.7. *Soit $(\Lambda_n)_{n\in\mathbb{N}}$ une suite d'ensembles convexes dans $\mathbb{R}^d$ et soient $y, z \in \mathbb{R}^d$. Soit $(b_n)_{n\in\mathbb{N}}$ une suite de nombres réels positifs avec b_n tend vers zéro quand n tend vers l'infini. Alors on a*

$$\lim_{n\to\infty} \frac{\left|\Lambda_n \cap (\Lambda_n - b_n y - z)\right|}{|\Lambda_n|} = 1.$$

Nous allons maintenant montrer que tous les autres termes de la variance de l'estimateur $\widehat{H}_n(t)$ sont négligeables quand n tend vers l'infini. Avec des arguments similaires à ceux utilisés pour obtenir les résultats ci-dessus et le Lemme 3.2.7, nous pouvons également obtenir le résultat suivant

$$\begin{aligned}
T_2 &= \frac{1}{b_n^{2d}|\Lambda_{n,R}|^2} \\
&\mathrm{E}\int_{\mathbb{R}^{2d}} \mathbb{1}_{\Lambda_{n,R}}(u)\mathbb{1}_{\Lambda_{n,R}}(v) I(v-u,u-v)H(u,v,\mathbf{X})K\Big(\frac{v-u-t}{b_n}\Big)K\Big(\frac{u-v-t}{b_n}\Big)\lambda_{\beta^\star}(u,v,\mathbf{X})\mathrm{d}u\mathrm{d}v \\
&= \frac{\beta^{\star 2}}{b_n^{2d}|\Lambda_{n,R}|^2} \\
&\int_{\mathbb{R}^{2d}} \mathbb{1}_{\Lambda_{n,R}}(u)\mathbb{1}_{\Lambda_{n,R}}(v) I(v-u,u-v)\bar{F}(o,v-u)K\Big(\frac{v-u-t}{b_n}\Big)K\Big(\frac{u-v-t}{b_n}\Big)G(v-u)\mathrm{d}u\mathrm{d}v \\
&= \frac{\beta^{\star 2}}{b_n^{2d}|\Lambda_{n,R}|}\int_{\mathbb{R}^d}\frac{|\Lambda_{n,R}\cap(\Lambda_{n,R}-s)|}{|\Lambda_{n,R}|} I(s,-s)\bar{F}(o,s)K\Big(\frac{s-t}{b_n}\Big)K\Big(\frac{-s-t}{b_n}\Big)G(s)\mathrm{d}s \\
&= \frac{\beta^{\star 2}}{b_n^{d}|\Lambda_{n,R}|} \\
&\int_{\mathbb{R}^d}\frac{|\Lambda_{n,R}\cap(\Lambda_{n,R}-(b_nz+t))|}{|\Lambda_{n,R}|} I(b_nz+t,-b_nz-t)\bar{F}(o,b_nz+t)K(z)K(-z-\frac{2t}{b_n})G(b_nz+t)\mathrm{d}z.
\end{aligned}$$

Le schéma de la fin de la preuve est le même que celui la démonstration ci-dessus, en appliquant le théorème de convergence dominée, sous la condition que $t \neq o$ et comme le noyau K est à support compact, nous obtenons $b_n^d|\Lambda_{n,R}|T_2 \longrightarrow 0$ quand $n \to \infty$.

Pour le comportement asymptotique du troisième terme T_3, nous rappelons que l'intensité conditionnelle de Papangelou du troisième ordre, pour tous points $u, v, w \in \mathbb{R}^d$ et $\mathbf{x} \in \Omega$, par

$$\lambda_{\beta^\star}(u,v,w,\mathbf{x}) = \lambda_{\beta^\star}(u,\mathbf{x})\lambda_{\beta^\star}(v,\mathbf{x}\cup\{u\})\lambda_{\beta^\star}(w,\mathbf{x}\cup\{u,v\}).$$

Etant donné que $\mathbf{X}$ est un processus ponctuel à interaction par paires, les termes d'interaction dus aux triplets ou d'ordre supérieur sont égales à un, c'est à dire que nous avons donc le potentiel $g(\mathbf{y})$ est nul quand $n(\mathbf{y}) > 2$, pour $\emptyset \neq \mathbf{y} \subseteq \mathbf{x}$. En appliquant la propriété de portée d'interaction finie (3.1.2) aux fonctions $\lambda_{\beta^\star}(u,\mathbf{x})$, $\lambda_{\beta^\star}(v,\mathbf{x}\cup\{u\})$ et $\lambda_{\beta^\star}(w,\mathbf{x}\cup\{u,v\})$ et après un calcul élémentaire, nous avons

$$\lambda_{\beta^\star}(u,v,w,\emptyset) = \begin{cases} \beta^{\star 3}G(v-u)G(w-v) & \text{si} \quad u \notin B(w,R) \\ \beta^{\star 3}G(v-u) & \text{sinon.} \end{cases} \tag{3.2.4}$$

Ce qui assure que $\lambda_{\beta^\star}(u,v,w,\emptyset)$ est une fonction qui dépend uniquement uniquement des variables $v-u, w-v$, nous la notons par $G_3(v-u,w-v)$. D'après la stationnarité

de $\mathbf{X}$, il résulte que

$$\begin{aligned} T_3 &= \frac{1}{b_n^{2d}|\Lambda_{n,R}|}\int_{\mathbb{R}^{2d}} \frac{|\Lambda_{n,R}\cap(\Lambda_{n,R}-v)|}{|\Lambda_{n,R}|}\bar{G}_3(v,w-v)K\left(\frac{v-t}{b_n}\right)K\left(\frac{w-v-t}{b_n}\right)\mathrm{d}v\mathrm{d}w \\ &= \frac{1}{|\Lambda_{n,R}|}\int_{\mathbb{R}^{2d}} \frac{|\Lambda_{n,R}\cap(\Lambda_{n,R}-(b_nz+t))|}{|\Lambda_{n,R}|}\bar{G}_3(b_nz+t,b_nz'+t)K(z)K(z')\mathrm{d}z\mathrm{d}z', \end{aligned}$$

où $\bar{G}_3(v,w-v) = I(v,w-v)\bar{F}(-v,w-v)G_3(v,w-v)$.

En appliquant le théorème de convergence dominée et sous les conditions supposées dans le Théorème 3.2.2, nous obtenons $b_n^d|\Lambda_{n,R}|T_3 \longrightarrow 0$ quand $n\to\infty$. En ce qui concerne le comportement asymptotique des termes T_4, T_5 et T_6 se prouve de la manière analogue. Pour le comportement asymptotique du terme T_7, soient u,v,w,y des points quelconques de $\mathbb{R}^d$ et $\mathbf{x}\in\Omega$, nous définissons l'intensité conditionnelle de Papangelou du quatrième ordre par

$$\lambda_{\beta^\star}(u,v,w,y,\mathbf{x}) = \lambda_{\beta^\star}(u,\mathbf{x})\lambda_{\beta^\star}(v,\mathbf{x}\cup\{u\})\lambda_{\beta^\star}(w,\mathbf{x}\cup\{u,v\})\lambda_{\beta^\star}(y,\mathbf{x}\cup\{u,v,w\})$$

ensuite, nous introduisons l'hypothèse de portée d'interaction finie (3.1.2) et en raisonnant de façon analogue avec ce qui précède sur $\lambda_{\beta^\star}(u,v,w,\emptyset)$ définie en (3.2.4), ceci nous a conduit à considérer $\lambda_{\beta^\star}(u,v,w,y,\emptyset)$ comme étant une fonction qui dépend uniquement des variables $v-u, y-w, w-u, w-v$, nous la notons par $G_4(v-u, y-w, w-u, w-v)$. De plus par la stationnarité de $\mathbf{X}$, ceci implique que

$$\begin{aligned} T_7 &= \frac{1}{b_n^{2d}|\Lambda_{n,R}|} \\ &\int_{\mathbb{R}^{3d}} \frac{|\Lambda_{n,R}\cap(\Lambda_{n,R}-w)|}{|\Lambda_{n,R}|}\bar{G}_4(v,y-w,w,w-v)K\left(\frac{v-t}{b_n}\right)K\left(\frac{y-w-t}{b_n}\right)\mathrm{d}v\mathrm{d}w\mathrm{d}y \\ &= \frac{1}{|\Lambda_{n,R}|}\int_{\mathbb{R}^{3d}} \frac{|\Lambda_{n,R}\cap(\Lambda_{n,R}-w)|}{|\Lambda_{n,R}|}\bar{G}_4(b_nz+t,b_nz'+t,w,w-b_nz-t)K(z)K(z')\mathrm{d}z\mathrm{d}z'\mathrm{d}w. \end{aligned}$$

$\bar{G}_4(v,y-w,w,w-v) = I(v,y-w)\bar{F}(v,w,y)G_4(v,y-w,w,w-v)$.

Nous appliquons le théorème de convergence dominée, nous obtenons donc pour $t\in T$, $b_n^d|\Lambda_{n,R}|T_7 \longrightarrow 0$ quand $n\to\infty$.

Nous terminons maintenant la preuve par le comportement asymptotique du terme T_8. Sous les conditions supposées dans le Théorème 3.2.2, nous avons

$$\begin{aligned}
T_8 &= -\frac{1}{b_n^{2d}|\Lambda_{n,R}|^2}\int_{\mathbb{R}^{4d}} \mathbb{1}_{\Lambda_{n,R}}(u)\mathbb{1}_{\Lambda_{n,R}}(w) I(v-u,y-w)K\left(\frac{v-u-t}{b_n}\right)K\left(\frac{y-w-t}{b_n}\right)\\
&\times \mathrm{E}\left[H(u,v,\mathbf{X})\lambda_{\beta^\star}(u,v,\mathbf{X})\right]\mathrm{E}\left[H(w,y,\mathbf{X})\lambda_{\beta^\star}(w,y,\mathbf{X})\right]\mathrm{d}u\mathrm{d}v\mathrm{d}w\mathrm{d}y\\
&= -\frac{\beta^{\star 4}}{b_n^{2d}|\Lambda_{n,R}|^2}\\
&\int_{\mathbb{R}^{4d}} \mathbb{1}_{\Lambda_{n,R}}(u)\mathbb{1}_{\Lambda_{n,R}}(w)\bar{G}_5(u,v,w,y,v-u,y-w)K\left(\frac{v-u-t}{b_n}\right)K\left(\frac{y-w-t}{b_n}\right)\mathrm{d}u\mathrm{d}v\mathrm{d}w\mathrm{d}y\\
&= -\frac{\beta^{\star 4}}{b_n^{2d}|\Lambda_{n,R}|}\\
&\int_{\mathbb{R}^{3d}} \frac{|\Lambda_{n,R}\cap(\Lambda_{n,R}-w)|}{|\Lambda_{n,R}|}\bar{G}_5(o,v,w,y,v,y-w)K\left(\frac{v-t}{b_n}\right)K\left(\frac{y-w-t}{b_n}\right)\mathrm{d}v\mathrm{d}w\mathrm{d}y\\
&= -\frac{\beta^{\star 4}}{|\Lambda_{n,R}|}\\
&\int_{\mathbb{R}^{3d}} \frac{|\Lambda_{n,R}\cap(\Lambda_{n,R}-w)|}{|\Lambda_{n,R}|}\bar{G}_5(o,b_nz+t,w,b_nz'+t+w,b_nz+t,b_nz'+t)K(z)K(z')\mathrm{d}w\mathrm{d}z\mathrm{d}z'
\end{aligned}$$

où $\bar{G}_5(u,v,w,y,v-u,y-w) = I(v-u,y-w)\bar{F}(u,v)\bar{F}(w,y)G(v-u)G(y-w)$. En appliquant le théorème de convergence dominée, nous obtenons le résultat voulu. □

3.3 Consistance ponctuelle

Dans cette section, nous présentons la consistance ponctuelle de l'estimateur $\widehat{G}_n(t)$.

Lemme 1. *Soit* $\mathbf{X}$ *un processus ponctuel à interaction par paires stationnaire défini en* (3.1.1)*, vérifiant la condition de portée d'interaction de paires finie* (3.1.2)*. Alors,*

$$\widehat{L}_n(t) \xrightarrow[n\to\infty]{p.s} L(t) = \beta^\star \bar{F}(o,t)$$

en point de continuité $t \in T$ *de* $\bar{F}$*.*

Démonstration. Pour la preuve du Lemme 1, il suffit d'appliquer le théorème ergodique de Nguyen et Zessin (1979).

$$\widehat{L}_n(t) = \frac{1}{|\Lambda_{n,R}|}\sum_{u\in\mathbf{X}_{\Lambda_{n,R}}} h(u,\mathbf{X}\setminus\{u\})h(t-u,\mathbf{X}\setminus\{u\})$$

converge presque sûrement vers

$$L(t) = \frac{1}{|\Lambda_0|} \mathrm{E} \sum_{u \in \mathbf{X}_{\Lambda_0}} h(u, \mathbf{X} \setminus \{u\}) h(t-u, \mathbf{X} \setminus \{u\})$$

où Λ_0 est un domaine borné de $\mathbb{R}^d$. D'après la formule de GNZ (1.4.3), nous avons

$$L(t) = \frac{1}{|\Lambda_0|} \mathrm{E} \int_{\Lambda_0} h(u, \mathbf{X}) h(t-u, \mathbf{X}) \lambda_{\beta^\star}(u, \mathbf{X}) \mathrm{d}u.$$

Comme $\mathbf{X}$ est stationnaire avec une fonction d'interaction invariante par translation et sous la condition (3.1.2), il résulte que $L(t) = \beta^\star \bar{F}(o,t)$. □

La consistance forte de l'estimateur $\widehat{L}_n(t)$ implique ce qui suit :

Proposition 3.3.1. *Sous les conditions de la Proposition 3.2.1 et le Théorème 3.2.2. Alors*

$$\widehat{G}_n(t) \underset{n\to\infty}{\longrightarrow} \beta^\star G(t) \quad \textit{en probabilité} \quad \mathbb{P} \quad \textit{si}$$

$$\widehat{H}_n(t) \underset{n\to\infty}{\longrightarrow} \widetilde{H}(t) = \beta^{\star 2} \bar{F}(o,t) G(t) \quad \textit{en probabilité} \quad \mathbb{P}.$$

3.3.1 Preuve de la Proposition 3.3.1

Démonstration. Nous considérons l'erreur quadratique moyenne de $\widehat{H}_n(t)$ définie par

$$\mathrm{EQM}\left(\widehat{H}_n(t)\right) = \mathrm{Var}\left(\widehat{H}_n(t)\right) + \left(Biais(\widehat{H}_n(t)\right)^2.$$

L'estimateur $\widehat{H}_n(t)$ est asymptotiquement sans biais d'après la Proposition 3.2.1 et donc

$$\lim_{n\to\infty} (Biais(\widehat{H}_n(t))^2 = 0.$$

D'après le Théorème 3.2.2, nous avons

$$\lim_{n\to\infty} \mathrm{Var}(\widehat{H}_n(t)) = 0.$$

Par conséquent, $\widehat{H}_n(t)$ converge en moyenne quadratique et donc converge en probabilité vers $\widetilde{H}(t)$. En combinant la consistance de l'estimateur $\widehat{H}_n(t)$ et la consistance connue de l'estimateur $\widehat{L}_n(t)$ de $L(t)$, ainsi la consitance ponctuelle (convergence en probabilité) de $\widehat{G}_n(t)$ peut être déduite

$$\widehat{G}_n(t) = \frac{\widehat{H}_n(t)}{\widehat{L}_n(t)} \longrightarrow \frac{\widetilde{H}(t)}{L(t)} = \beta^\star G(t).$$

□

3.4 Inégalités de Kahane-Khintchine

Cette section est dédiée à la présentation de l'outil mathématique que nous utiliserons tout au long de notre travail, à savoir les espaces d'Orlicz associés aux fonctions de Young. Pour plus de détails sur la théorie des fonctions de Young et des espaces d'Orlicz, nous pourrons nous référer au livre de Krasnosel'skii et Rutickii [84], Kufner et al. [54].

Nous introduisons les notations suivantes, que nous garderons tout au long de notre travail. On munit le réseau $\mathbb{Z}^d$ de la relation d'ordre lexicographique comme suit : si $i = (i_1, ..., i_d)$ et $j = (j_1, ..., j_d)$ sont des éléments distincts de $\mathbb{Z}^d$, la notation $i <_{lex} j$ signifie que, ou bien $i_1 < j_1$ ou bien pour un certain p dans $\{2, 3, ..., d\}$, $i_p < j_p$ et $i_q = j_q$ pour $1 \leq q < p$. Soient les ensembles $\{V_i^k \,;\, i \in \mathbb{Z}^d \,,\, k \in \mathbb{N}^*\}$ définis de la façon suivante :

$$V_i^1 = \{j \in \mathbb{Z}^d \,;\, j <_{lex} i\}$$

et pour $k \geq 2$

$$V_i^k = V_i^1 \cap \{j \in \mathbb{Z}^d \,;\, |i - j| \geq k\} \quad \text{où} \quad |i - j| = \max_{1 \leq l \leq d} |i_l - j_l|.$$

Pour tout sous-ensemble Γ de $\mathbb{Z}^d$, on définit $\mathcal{F}_\Gamma = \sigma(X_i \,;\, i \in \Gamma)$ et on note

$$\mathrm{E}_{|k|}(X_i) = \mathrm{E}[X_i | \mathcal{F}_{V_i^{|k|}}], \quad k \in V_i^1.$$

Fonctions de Young et des espaces d'Orlicz

On dit qu'une fonction ψ est une fonction de Young si ψ est une fonction convexe, strictement croissante, définie sur $\mathbb{R}^+$ à valeurs dans $\mathbb{R}^+ \cup \{\infty\}$ telle que $\lim_{t \to \infty} \psi(t) = +\infty$ et $\psi(0) = 0$. L'espace d'Orlicz $\mathbb{L}_\psi$ associé à la fonction de Young ψ est l'espace des variables aléatoires Z définies sur l'espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$ telles que $\mathrm{E}[\psi(|Z|/\kappa)] < +\infty$, pour un certain $\kappa > 0$. On munit l'espace d'Orlicz $\mathbb{L}_\psi$ de la norme de Luxemburg définie pour toute variable aléatoire Z par

$$\|Z\|_\psi = \inf\{\, \kappa > 0 \,;\, \mathrm{E}[\psi(|Z|/\kappa)] \leq 1 \,\}.$$

De la définition de la norme de Luxemburg, nous déduisons qu'une condition nécessaire et suffisante pour que $Z \in \mathbb{L}_\psi$, est qu'il existe une constante $c > 0$ telle que

$$\mathrm{E}[c\psi(|Z|)] < +\infty.$$

Inégalités de Kahane-Khintchine

Soit $q > 0$. Nous notons par ψ_q la fonction de Young exponentielle définie pour tout $x \in \mathbb{R}^+$ par

$$\psi_q(x) = \exp((x+\xi_q)^q) - \exp(\xi_q^q) \quad \text{où} \quad \xi_q = ((1-q)/q)^{1/q}\mathbb{1}\{0 < q < 1\}.$$

Le théorème qui suit sera admis, mais le lecteur qui désire lire sa démonstration pourra la trouver dans (El Machkouri [88].Théorème 1). L'auteur présente un résultat sur l'inégalité de type Kahane-Khintchine dans des espaces d'Orlicz pour des champs de variables aléatoires dépendantes ayant des moments exponentiels finis. Sous l'hypothèse suivante :

$$\exists q \in]0,2[\quad \exists c > 0 \quad \mathrm{E}[\exp(c|X_0|^{\theta(q)})] < +\infty \tag{3.4.1}$$

pour tout réel $0 < q < 2$ avec $\theta(q) = 2q/(2-q)$ et que par convention $1/\theta(2) = 0$.

Théorème 3.4.1. *(El Machkouri, 2002) Soit $(X_i)_{i\in\mathbb{Z}^d}$ un champ de variables aléatoires réelles de moyennes nulles stationnaires satisfont l'hypothèse (3.4.1) pour un certain réel $0 < q < 2$. Alors il existe une constante universelle $M_1(q) > 0$ qui ne dépend que de q telle que pour toute famille $(a_i)_{i\in\mathbb{Z}^d}$ de nombres réels et toute partie finie Γ de $\mathbb{Z}^d$,*

$$\Big\| \sum_{i\in\Gamma} a_i X_i \Big\|_{\psi_q} \leq M_1(q) \Big(\sum_{i\in\Gamma} |a_i| b_{i,q}(X) \Big)^{1/2} \tag{3.4.2}$$

où

$$b_{i,q}(X) = |a_i| \big\|X_0\big\|^2_{\psi_{\theta(q)}} + \sum_{k\in V_0^1} |a_{k+i}| \left\| \sqrt{\big|X_k E_{|k|}(X_0)\big|} \right\|^2_{\psi_{\theta(q)}}.$$

Si $(X_i)_{i\in\mathbb{Z}^d}$ est borné alors pour tout réel $0 < q \leq 2$, il existe une constante universelle $M_2(q) > 0$ qui ne dépend que de q telle que pour toute famille $(a_i)_{i\in\mathbb{Z}^d}$ de nombres réels et toute partie finie Γ de $\mathbb{Z}^d$,

$$\Big\| \sum_{i\in\Gamma} a_i X_i \Big\|_{\psi_q} \leq M_2(q) \Big(\sum_{i\in\Gamma} |a_i| b_{i,\infty}(X) \Big)^{1/2} \tag{3.4.3}$$

où

$$b_{i,\infty}(X) = |a_i| \big\|X_0\big\|^2_{\infty} + \sum_{k\in V_0^1} |a_{k+i}| \big\|X_k E_{|k|}(X_0)\big\|_{\infty}.$$

Nous avons besoin aussi de la proposition suivante établie par Dedecker [37].

Proposition 3.4.2. *(Dedecker, 2001) Soit $(X_i)_{i\in\mathbb{Z}^d}$ un champ de variables aléatoires réelles centrées et soit Γ une partie finie de $\mathbb{Z}^d$. Pour tout réel $p > 2$, on a*

$$\left\| \sum_{i\in\Gamma} X_i \right\|_p \leq \left(2p \sum_{i\in\Gamma} c_i(X) \right)^{1/2} \tag{3.4.4}$$

où

$$c_i(X) = \|X_i\|_p^2 + \sum_{k\in V_i^1} \|X_k E_{|k-i|}(X_i)\|_{\frac{p}{2}}.$$

3.5 Consistance forte ponctuelle

En faisant le lien avec les résultats du paragraphe précédent, il est alors aisé d'établir les résultats de la consistance forte de l'estimateur $\widehat{G}_n(t)$. Pour d'autres résultats sur la consistance forte, en utilisant des hypothèses de mélange, nous l'étudierons dans le Chapitre 5.

Proposition 3.5.1. *Soit $\mathbf{X}$ un processus ponctuel d'interaction par paires stationnaire dans $\mathbb{R}^d$ d'intensité contitionelle de Papangelou définie en (3.1.1) satisfait la condition (3.1.2).*

$$\widehat{G}_n(t) \underset{n\to\infty}{\longrightarrow} \beta^\star G(t) \quad p.s. \quad si$$

$$\widehat{H}_n(t) \underset{n\to\infty}{\longrightarrow} \widetilde{H}(t) = \beta^{\star 2} \bar{F}(o,t) G(t) \quad p.s..$$

Nous donnons les conditions assurant ponctuellement la consistance forte de l'estimateur $\widehat{H}_n(t)$. Supposons que $\Lambda_{n,R}$ se décompose comme $\Lambda_{n,R} = \cup_{i\in\widetilde{I}_n} \Lambda_i$, où Λ_i sont des sous-cubes de $\Lambda_{n,R}$ centrés en i et $\widetilde{I}_n$ une suite de sous-ensembles de $\mathbb{Z}^d$, telle que le cardinal $|\widetilde{I}_n|$ de $\widetilde{I}_n$ tend vers l'infini, quand n tend vers l'infini. Notons pour tout $k \in \widetilde{I}_n$,

$$\bar{Z}_k = Z_k - \mathrm{E}\, Z_k, \quad \text{et} \quad S_n = \sum_{k\in\widetilde{I}_n} \bar{Z}_k$$

avec

$$Z_k = \sum_{\substack{u,v\in\mathbf{X} \\ v-u\in B(o,R)}}^{\neq} \mathbb{1}_{\Lambda_k}(u) h(u, \mathbf{X}\setminus\{u,v\}) h(v, \mathbf{X}\setminus\{u,v\}) K\left(\frac{v-u-t}{b_n}\right).$$

Proposition 3.5.2. *Soit $\mathbf{X}$ un processus ponctuel à interaction par paires stationnaire défini en (3.1.1). Sous les conditions $\mathbf{K}^0$ et $\mathcal{P}(t)$. De plus supposons que* $\exp(-g(t)\bar{F}(o,t))$

admet des dérivées partielles d'ordre m continues et bornées sur $B^o(t,\delta)$, pour $t \in \mathring{T}$, pour un certain $\delta > 0$ et pour tout $m \geq 1$. Alors, quand $n \to \infty$:
1) S'il existe un réel $0 < q < 2$ tel que la variable aléatoire $\bar{Z}_0$ appartient à $\mathbb{L}_{\psi_{\theta(q)}}$ et sous la condition

$$\sum_{k \in V_0^1} \left\| \sqrt{\left|\bar{Z}_k E_{|k|}(\bar{Z}_0)\right|} \right\|^2_{\psi_{\theta(q)}} < \infty. \tag{3.5.1}$$

Alors

$$\widehat{H}_n(t) - \widetilde{H}(t) = \mathcal{O}_{p.s.}\left(\frac{(\log n)^{1/q}}{(b_n\sqrt{n})^d}\right) + \mathcal{O}(b_n^m).$$

2) Si $\bar{Z}_0$ appartient à $\mathbb{L}^\infty$ et sous la condition

$$\sum_{k \in V_0^1} \left\|\bar{Z}_k E_{|k|}(\bar{Z}_0)\right\|_\infty < \infty. \tag{3.5.2}$$

Alors

$$\widehat{H}_n(t) - \widetilde{H}(t) = \mathcal{O}_{p.s.}\left(\frac{(\log n)^{1/2}}{(b_n\sqrt{n})^d}\right) + \mathcal{O}(b_n^m).$$

3) S'il existe un réel $p > 2$, tel que la variable aléatoire $\bar{Z}_0$ appartient à $\mathbb{L}^p$ et sous la condition

$$\sum_{k \in V_0^1} \left\|\bar{Z}_k E_{|k|}(\bar{Z}_0)\right\|_{\frac{p}{2}} < \infty. \tag{3.5.3}$$

Alors

$$\widehat{H}_n(t) - \widetilde{H}(t) = \mathcal{O}_{p.s.}\left(\frac{(\log n)^{1/p}}{(b_n\sqrt{n})^d}\right) + \mathcal{O}(b_n^m).$$

3.5.1 Preuve de la Proposition 3.5.2

Remarque 3.5.3. Faisons une remarque sur les conditions (3.5.1), (3.5.2), (3.5.3). D'après la propriéte de Markov, pour $i \neq 0$ qui ne sont pas des voisinages, alors $\mathrm{E}[\bar{Z}_0 | X_i; i \neq 0] = 0$. Comme $\sigma(Z_i, i \in V_0^k)$ est contenu dans $\sigma(X_i, i \neq 0)$ pour $k > l$, pour un certain l, il résulte que les conditions (3.5.1), (3.5.2), (3.5.3) sont satisfaites.

Démonstration. Dans ce qui suit C désignera une constante réelle strictement positive générique. Soit ε un réel positif, nous pouvons alors imaginer qu'il existe $(r_n)_{n>0}$ une suite de réels strictement positifs intermédiaire, tendant vers zéro. Une telle suite définit la vitesse de convergence de l'estimateur qui nous intéresse.

1) En utilisant l'inégalité de Markov , nous obtenous

$$\mathbb{P}(|\widehat{H}_n(t) - \mathrm{E}\,\widehat{H}_n(t)| > \varepsilon r_n) = \mathbb{P}\left(|S_n| > \varepsilon r_n (nb_n)^d\right)$$
$$\leq \exp\left[-\left(\frac{\varepsilon\, r_n (nb_n)^d}{||S_n||_{\psi_{\theta(q)}}} + \xi_q\right)^q\right] \mathrm{E}\exp\left[\left(\frac{|S_n|}{||S_n||_{\psi_{\theta(q)}}} + \xi_q\right)^q\right].$$

Par conséquent, nous supposons qu'il existe un réel $0 < q < 2$ fixé, tel que $\bar{Z}_0 \in \mathbb{L}_{\psi_{\theta(q)}}$, nous utilisons l'inégalité de Kahane-Khintchine (3.4.2) du Théorème 3.4.1, nous obtenons

$$\mathbb{P}(|\widehat{H}_n(t) - \mathrm{E}\,\widehat{H}_n(t)| > \varepsilon r_n) = \mathbb{P}\left(|S_n| > \varepsilon r_n (nb_n)^d\right)$$
$$\leq (1 + e^{\xi_q^q})\, \exp\left[-\left(\frac{\varepsilon\, r_n (nb_n)^d}{M_1(q)(\sum_{i\in\widetilde{I}_n} b_{i,q}(\bar{Z}))^{1/2}} + \xi_q\right)^q\right]$$

où

$$b_{i,q}(\bar{Z}) = \left\|\bar{Z}_0\right\|^2_{\psi_{\theta(q)}} + \sum_{k\in V_0^1} \left\|\sqrt{\left|\bar{Z}_k E_{|k|}(\bar{Z}_0)\right|}\right\|^2_{\psi_{\theta(q)}}.$$

Par définition de la norme $\left\|\bar{Z}_0\right\|_{\psi_{\theta(q)}}$ et sous la condition (3.5.1), il résulte donc qu'il existe une constante $C > 0$, telle que

$$\mathbb{P}(|\widehat{H}_n(t) - \mathrm{E}\,\widehat{H}_n(t)| > \varepsilon r_n) \leq (1 + e^{\xi_q^q})\, \exp\left[-\frac{\varepsilon^q r_n^q (b_n\sqrt{n})^{dq}}{C^q}\right].$$

Il suffit maintenant de choisir $r_n = (\log n)^{1/q}/(b_n\sqrt{n})^d$ et pour n suffisamment grand, alors

$$\mathbb{P}(|\widehat{H}_n(t) - \mathrm{E}\,\widehat{H}_n(t)| > \varepsilon r_n) \leq (1 + e^{\xi_q^q})\, \exp\left[-\frac{\varepsilon^q \log n}{C^q}\right].$$

2) En reformulant l'inégalité de Markov et l'inégalité (3.4.3) du Théorème 3.4.1, pour $\bar{Z}_0 \in \mathbb{L}^\infty$, entraîne que pour ε un réel positif,

$$\mathbb{P}(|\widehat{H}_n(t) - \mathrm{E}\,\widehat{H}_n(t)| > \varepsilon r_n) \leq 2\, \exp\left[-\left(\frac{\varepsilon\, r_n (nb_n)^d}{M_2(q)(\sum_{i\in\widetilde{I}_n} b_{i,\infty}(\bar{Z}))^{1/2}}\right)^2\right].$$

$$b_{i,\infty}(\bar{Z}) = \left\|\bar{Z}_0\right\|^2_\infty + \sum_{k\in V_0^1} \left\|\bar{Z}_k E_{|k|}(\bar{Z}_0)\right\|_\infty.$$

Ainsi si $r_n = (\log n)^{1/2}/(b_n\sqrt{n})^d$ pour n suffisamment et sous la condition (3.5.2), il résulte qu' il existe une constante $C > 0$ telle que

$$\mathbb{P}(|\widehat{H}_n(t) - \mathrm{E}\,\widehat{H}_n(t)| > \varepsilon r_n) \leq 2\, \exp\left[-\frac{\varepsilon^2 \log n}{C^2}\right].$$

D'aprés le Lemme de Borel-Cantelli et (3.2.1), nous concluons la preuve des deux premiers résultats de la Proposition 3.5.2.

3) Enfin, en toute rigueur, nous pouvons maintenant achever la démonstration de la Proposition 3.5.2. Soit $p > 2$ fixé, tel que $\bar{Z}_0$ appartient à l'espace $\mathbb{L}^p$. En utilisant l'inégalité de Markov, nous avons pour tout $\varepsilon > 0$,

$$\begin{aligned}\mathbb{P}(|\widehat{H}_n(t) - \mathrm{E}\,\widehat{H}_n(t)| > \varepsilon r_n) &= \mathbb{P}\left(|S_n| > \varepsilon r_n (nb_n)^d\right) \\ &\leq \frac{\varepsilon^{-p}\,\mathrm{E}\,|S_n|^p}{r_n^p (nb_n)^{pd}} \\ &\leq \frac{\varepsilon^{-p}}{r_n^p (nb_n)^{pd}} \left(2p \sum_{i \in \widetilde{I}_n} c_i(\bar{Z})\right)^{p/2}.\end{aligned}$$

La dernière inégalité résulte d'après la Proposition 3.4.2, où

$$c_i(Z) = \|\bar{Z}_i\|_p^2 + \sum_{k \in V_i^1} \|\bar{Z}_k E_{|k-i|}(\bar{Z}_i)\|_{\frac{p}{2}}.$$

Sous la condition (3.5.3), nous déduisons par conséquent qu'il existe $C > 0$ telle que

$$\mathbb{P}(|\widehat{H}_n(t) - \mathrm{E}\,\widehat{H}_n(t)| > \varepsilon r_n) \leq \frac{C\varepsilon^{-p}}{r_n^p (b_n\sqrt{n})^{pd}}$$

dès lors que $r_n = (\log n)^{1/p}/(b_n\sqrt{n})^d$, d'aprés le Lemme de Borel-Cantelli et (3.2.1), nous concluons la preuve de la Proposition 3.5.2. □

3.6 Consistance forte globale

Au meilleur de notre connaissance, en 1965, Nadaraja [98] a été le premier chercheur à formuler un résultat énonçant la convergence uniforme presque sûre de l'estimateur à noyau de la densité. Pour prouver la consistance forte globale de l'estimateur de la fonction d'interaction de paires, qui fera l'objet de notre étude dans cette section. Il est clair que l'obtention de résultats uniformes nécessite le renforcement de certaines hypothèses.

Condition $\mathcal{P}(T)$:

$\bar{F}$ et $G = \exp(-g)$ sont des fonctions continues sur T.

Un des intérêts de la condition présentée ci-dessus est qu'elle est nécessaire pour la preuve de la convergence de la partie biais. Nous avons besoin aussi de rajouter l' hypothèse suivante portant sur la classe des noyaux que nous utilisons.

Condition $\mathcal{Z}$:

La fonction noyau K est lipschitzienne, c'est à dire qu'il existe une constante $\eta > 0$ telle que

$$\left|K(u) - K(v)\right| \leq \eta||u - v|| \quad \text{pour tous} \quad u, v \in T.$$

La condition lipschitzienne est une hypothèse classique en estimation non paramétrique, qui est nécessaire pour la convergence de la partie stochastique, nous permet d'obtenir la vitesse de la convergence.
Nous établissons maintenant la consistance forte globale de l'estimateur $\widehat{G}_n(t)$ sur un compact T_0 de T.

Théorème 3.6.1. *Soit* $\mathbf{X}$ *un processus ponctuel d'interaction par paires stationnaire dans* $\mathbb{R}^d$ *d'intensité contitionelle de Papangelou définie en* (3.1.1) *satisfait la condition* (3.1.2).

$$\sup_{t\in T_o}|\widehat{G}_n(t)-\beta^\star G(t)|\longrightarrow 0 \quad p.s. \quad si$$

$$\sup_{t\in T_o}|\widehat{H}_n(t)-\widetilde{H}(t)|\longrightarrow 0 \quad p.s..$$

Nous donnons les conditions de la consistance glogale de l'estimateur $\widehat{H}_n(t)$.

Théorème 3.6.2. *Sous les conditions* $\mathbf{K}^0$, $\mathcal{P}(T)$ *et* $\mathcal{Z}$. *De plus, on suppose que* $\exp(-g(t))\bar{F}(o,t)$ *admet des dérivées partielles d'ordre* m *continues et bornées sur* T.
1) S'il existe un réel $0<q<2$ *tel que la variable aléatoire* $\bar{Z}_0$ *appartient à* $\mathbb{L}_{\psi_{\theta(q)}}$ *et sous la condition*

$$\sum_{k\in V_0^1}\left\|\sqrt{|\bar{Z}_k E_{|k|}(\bar{Z}_0)|}\right\|^2_{\psi_{\theta(q)}}<\infty.$$

Alors

$$\sup_{t\in T_o}\left|\widehat{H}_n(t)-\widetilde{H}(t)\right|=\mathcal{O}_{p.s.}\left(\frac{(\log n)^{1/q}}{(b_n\sqrt{n})^d}\right)+\mathcal{O}(b_n^m) \quad quand \quad n\to\infty.$$

2) Si la variable aléatoire $\bar{Z}_0$ *appartient à* $\mathbb{L}^\infty$ *et sous la condition*

$$\sum_{k\in V_0^1}\left\|\bar{Z}_k E_{|k|}(\bar{Z}_0)\right\|_\infty<\infty.$$

Alors

$$\sup_{t\in T_o}\left|\widehat{H}_n(t)-\widetilde{H}(t)\right|=\mathcal{O}_{p.s.}\left(\frac{(\log n)^{1/2}}{(b_n\sqrt{n})^d}\right)+\mathcal{O}(b_n^m) \quad quand \quad n\to\infty.$$

3) S'il existe un réel $p>2$ *tel que la variable aléatoire* $\bar{Z}_0$ *appartient à* $\mathbb{L}^p$ *et sous la condition*

$$\sum_{k\in V_0^1}\left\|\bar{Z}_k E_{|k|}(\bar{Z}_0)\right\|_{\frac{p}{2}}<\infty.$$

Soit $b_n=n^{-q_2}(\log n)^{q_1}$ *pour certaines constantes* $q_1,q_2>0$. *Soient* $a,b\geq 0$ *fixés, si*

$$a(p+d)-d^2/2-q_2 d>1 \quad et \quad b(p+d)+q_1 d>1.$$

Alors

$$\sup_{t\in T_o}\left|\widehat{H}_n(t)-\widetilde{H}(t)\right| = \mathcal{O}_{p.s.}\left(\frac{n^a(\log n)^b}{(b_n\sqrt{n})^d}\right)+\mathcal{O}(b_n^m) \quad \textit{quand} \quad n\to\infty.$$

3.6.1 Preuve du Théorème 3.6.2

Démonstration. Nous montrons la consistance forte globale de $\widehat{H}_n$ vers $\widetilde{H}$. Nous appliquons une décomposition en inégalité triangulaire permet d'avoir

$$\sup_{t\in T_0}\left|\widehat{H}_n(t)-\widetilde{H}(t)\right| \le \sup_{t\in T_0}\left|\widehat{H}_n(t)-\mathrm{E}[\widehat{H}_n(t)]\right| + \sup_{t\in T_0}\left|\mathrm{E}[\widehat{H}_n(t)]-\widetilde{H}(t)\right|. \tag{3.6.1}$$

Soit $(r_n)_{n\ge1}$ une suite de nombres positifs tend vers zéro. Suivant Carbon et al. [25], le compact T_0 peut être couvert par v_n cubes T_k ayant des côtés de longueur $l_n = r_n b_n^{d+1}$ de centre c_k, tels qu'il existe $c>0$, $v_n \le c/l_n^d$

$$A_1 = \max_{1\le k\le v_n}\sup_{t\in T_k}\left|\widehat{H}_n(t)-\widehat{H}_n(c_k)\right|$$

$$A_2 = \max_{1\le k\le v_n}\sup_{t\in T_k}\left|\mathrm{E}\,\widehat{H}_n(t)-\mathrm{E}\,\widehat{H}_n(c_k)\right|$$

$$A_3 = \max_{1\le k\le v_n}\left|\widehat{H}_n(c_k)-\mathrm{E}\,\widehat{H}_n(c_k)\right|$$

alors

$$\sup_{t\in T_o}\left|\widehat{H}_n(t)-\mathrm{E}\,\widehat{H}_n(t)\right| \le A_1+A_2+A_3. \tag{3.6.2}$$

Nous étudions les lemmes qui suivent qui améliorent sensiblement le résultat voulu.

Lemme 3.6.3. *Pour j=1,2, on a*

$$A_j = \mathcal{O}_{p.s.}(r_n).$$

Démonstration. Soit alors pour tous $t\in T_k$, sous la condition $\mathcal{Z}$, nous déduisons qu'il

existe une constante $\eta > 0$ telle que pour tout n suffisamment grand

$$\begin{aligned}
&\left|\widehat{H}_n(t) - \widehat{H}_n(c_k)\right| \\
&\leq \frac{1}{b_n^{d+1}}\eta\|t - c_k\|\left|\frac{1}{|\Lambda_{n,R}|}\sum_{\substack{u,v\in\mathbf{X}\\ v-u\in B(o,R)}}^{\neq} \mathbb{1}_{\Lambda_{n,R}}(u)h(u,\mathbf{X}\setminus\{u,v\})h(v,\mathbf{X}\setminus\{u,v\})\right| \\
&\leq \frac{1}{b_n^{d+1}}\eta r_n b_n^{d+1}\left|\frac{1}{|\Lambda_{n,R}|}\sum_{\substack{u,v\in\mathbf{X}\\ v-u\in B(o,R)}}^{\neq} \mathbb{1}_{\Lambda_{n,R}}(u)h(u,\mathbf{X}\setminus\{u,v\})h(v,\mathbf{X}\setminus\{u,v\})\right| \\
&\leq \eta r_n\left|\frac{1}{|\Lambda_{n,R}|}\sum_{\substack{u,v\in\mathbf{X}\\ v-u\in B(o,R)}}^{\neq} \mathbb{1}_{\Lambda_{n,R}}(u)h(u,\mathbf{X}\setminus\{u,v\})h(v,\mathbf{X}\setminus\{u,v\})\right|.
\end{aligned}$$

Nous appliquons aisément le théorème ergodique obtenu par Nguyen et Zessin [101] dans la dernière inégalité. □

Lemme 3.6.4. *Sous la condition* (3.5.1) *pour certain* $0 < q < 2$ *telle que* $\bar{Z}_0 \in \mathbb{L}_{\psi_{\theta(q)}}$ *et* $r_n = (\log n)^{1/q}/(b_n\sqrt{n})^d$ *ou sous la condition* (3.5.2) *telle que* $\bar{Z}_0 \in \mathbb{L}^\infty$ *et* $r_n = (\log n)^{1/2}/(b_n\sqrt{n})^d$. *Alors pour* n *suffisamment grand,*

$$A_3 = \mathcal{O}_{p.s.}(r_n).$$

Démonstration. Nous reprenons les mêmes étapes de la preuve de la Proposition 3.5.2, pour $\varepsilon > 0$, pour $0 < q < 2$ fixé et $r_n = (\log n)^{1/q}/(b_n\sqrt{n})^d$. Alors pour n assez grand nous obtenons,

$$\sup_{t\in T_0}\mathbb{P}(|\widehat{H}_n(t) - \mathrm{E}\,\widehat{H}_n(t)| > \varepsilon r_n) \leq (1 + e^{\xi_q^q})\,\exp\left[-\frac{\varepsilon^q\,\log n}{C^q}\right]. \tag{3.6.3}$$

Soit $0 < q \leq 2$ fixé, $r_n = (\log n)^{1/2}/(b_n\sqrt{n})^d$ nous obtenons ainsi,

$$\sup_{t\in T_0}\mathbb{P}(|\widehat{H}_n(t) - \mathrm{E}\,\widehat{H}_n(t)| > \varepsilon r_n) \leq 2\,\exp\left[-\frac{\varepsilon^2\,\log n}{C^2}\right]. \tag{3.6.4}$$

D'autre part, nous avons

$$\mathbb{P}\left(|A_3| > \varepsilon r_n\right) \leq v_n \sup_{t\in T_0}\mathbb{P}(|\widehat{H}_n(t) - \mathrm{E}\,\widehat{H}_n(t)| > \varepsilon r_n). \tag{3.6.5}$$

En combinant (3.6.5), (3.6.3) et (3.6.4), donc il en résulte grâce au Lemme de Borel-Cantelli, que

$$\mathbb{P}(\limsup_{n\to\infty}|A_3| > \varepsilon r_n) = 0.$$

□

Les deux premiers résultats du Théorème 3.6.2 découlent alors de l'expréssion (3.6.2) et en utilisant la continuité uniforme de $\widetilde{H}$, d'une part, la partie biais converge vers 0 d'après la Proposition 3.2.1, nous obtenons par un calcul classique

$$\sup_{t\in T_o} |\operatorname{E}\widehat{H}_n(t) - \widetilde{H}(t)| = \mathcal{O}(b_n^m). \tag{3.6.6}$$

Lemme 3.6.5. *S'il existe un réel $p > 2$ tel que la variable aléatoire $\bar{Z}_0$ appartient à $\mathbb{L}^p$ et sous la condition (3.5.3) avec $b_n = n^{-q_2}(\log n)^{q_1}$ pour certaines constantes $q_1, q_2 > 0$ et si*

$$a(p+d) - d^2/2 - q_2 d > 1 \quad \textit{et} \quad b(p+d) + q_1 d > 1,$$

pour $a, b \geq 0$ fixés et $r_n = n^a(\log n)^b/(b_n\sqrt{n})^d$. Alors

$$A_3 = \mathcal{O}_{p.s}(r_n).$$

Démonstration. Soit $p > 2$ fixé. Pour tout $\varepsilon > 0$, nous reprenons les mêmes étapes de la preuve de la Proposition 3.5.2, nous concluons

$$\begin{aligned}\mathbb{P}\left(|A_3| > \varepsilon r_n\right) &\leq v_n \sup_{t\in T_0} \mathbb{P}(|\widehat{H}_n(t) - \operatorname{E}\widehat{H}_n(t)| > \varepsilon r_n)\\ &\leq v_n \frac{\kappa\varepsilon^{-p}}{r_n^p (b_n\sqrt{n})^{pd}}.\end{aligned}$$

Comme $v_n \leq c/l_n^d$ et $l_n = r_n b_n^{1+d}$, donc pour $r_n = n^a(\log n)^b/(b_n\sqrt{n})^d$, il résulte pour n suffisamment grand,

$$\begin{aligned}\mathbb{P}\left(|A_3| > \varepsilon r_n\right) &\leq \frac{\kappa\varepsilon^{-p}}{n^{a(p+d)-d^2/2}(\log n)^{b(p+d)} b_n^d}\\ &\leq \frac{\kappa\varepsilon^{-p}}{n^{a(p+d)-d^2/2-q_2 d}(\log n)^{b(p+d)+q_1 d}}.\end{aligned}$$

Pour $a(p+d) - d^2/2 - q_2 d > 1$ et $b(p+d) + q_1 d > 1$, nous obtenons pour tout réel $\varepsilon > 0$

$$\sum_{n\geq 1} \mathbb{P}\left(|A_3| > \varepsilon r_n\right) < \infty.$$

□

Nous concluons la preuve du Théorème 3.6.2, grâce à l'inégalité (3.6.1) et le résultat (3.6.6). □

3.6.2 Preuve du Théorème 3.6.1

Démonstration. Nos preuves seront exprimées à l'aide des constantes α_1 et α_2, en utilisant le fait que en tout point $t \in T_0$,

$$\exists \alpha_1 > 0, \quad L(t) \geq \alpha_1$$

et

$$\exists \alpha_2 > 0, \quad \widetilde{H}(t) \leq \alpha_2,$$

ce qui entraîne,

$$|\widehat{G}_n(t) - \beta^\star G(t)| \leq \frac{|\widehat{H}_n(t) - \widetilde{H}(t)|}{\alpha_1 - |\widehat{L}_n(t) - L(t)|} + \alpha_2 \alpha_1^{-1} |\widehat{L}_n(t) - L(t)|. \tag{3.6.7}$$

Le Lemme 1 donne immédiatement, quand $n \to \infty$,

$$\left|\widehat{L}_n(t) - L(t)\right| \xrightarrow{p.s} 0. \tag{3.6.8}$$

En utilisant la monotonie des fonctions L et $\widehat{L}_n$, la relation d'ordre (partiel) sur $\mathbb{R}^d$ étant définie par

$$(u_i)_{1 \leq j \leq d} \leq (v_i)_{1 \leq j \leq d} \Leftrightarrow \forall \quad j = 1, ..., d \quad u_i \leq v_j$$

nous pouvons approcher les fonctions L et $\widehat{L}_n$ par leurs valeurs en un nombre fini de points. Réunissant cette remarque et le résultat (3.6.8), nous obtenons quand $n \to \infty$,

$$\sup_{t \in T_o} \left|\widehat{L}_n(t) - L(t)\right| \xrightarrow{p.s} 0. \tag{3.6.9}$$

Notre résultat découle de la consistance forte globale de $\widehat{H}_n$ vers $\widetilde{H}$ déjà étudiée dans le Théorème 3.6.2, de la relation (3.6.9) et de la relation (3.6.7). □

3.7 Normalité asymptotique

Dans cette section, nous établissons la normalité asymptotique de l'estimateur $\widehat{G}_n(t)$. Pour aboutir à nos résultats, rappelons un théorème central limite proposé par Wang et Woodroofe [132], pour les tableaux triangulaires stationnaires de champs aléatoires m_n-dépendant. Ce résultat améliore le théorème central limite établi par Heinrich [70].

La notation $i \preceq j$ signifie que $i_\tau \leq j_\tau$ pour tout $\tau = 1, ..., d$, pour $i, j \in \mathbb{Z}^d$ et $\mathbf{1}= (1, ..., 1) \in \mathbb{Z}^d$. Soit $\{Y_{n,i}; i \in \mathbb{N}^d\}$ un tableau triangulaire stationnaire de moyenne nulle et soit $(m_n)_{n\geq 1}$ une suite d'entiers positifs qui tend vers l'infini. $\{Y_{n,i}; i \in \mathbb{N}^d\}$ est m_n-dépendant dans le sens où $Y_{n,i}$ et $Y_{n,j}$ sont indépendants si $|i - j|_\infty \geq m$, où $|j|_\infty = \max\{|j_1|, ..., |j_d|\}$. En outre, les auteurs imposent la condition essentielle suivante :

$$\left\| \sum_{i\in\mathbb{N}^d, \mathbf{1}\preceq i\preceq j} Y_{n,i} \right\|_2 \leq C(j_1, ..., j_d)^{1/2} \quad \text{pour tout} \quad n \in \mathbb{N}, j \in \mathbb{N}^d. \tag{3.7.1}$$

L'inégalité (3.7.1) a été mis en place, sous diverses conditions de la dépendance des champs aléatoires stationnaires, par Dedecker [37], Wang et Woodroffe [131] et El Machkouri et al. [90].

Théorème 3.7.1. *(Wang et Woodroofe, 2011) On suppose qu'il existe une constante C telle que (3.7.1) est vérifiée. S'il existe une suite $\{l_n\}_{n\in\mathbb{N}} \subset \mathbb{N}$, $m_n/l_n \to 0$ et $l_n/n \to 0$ quand $n \to \infty$, telle que*

$$\lim_{n\to\infty} \frac{1}{l_n^d} \mathrm{E}\left[\left(\sum_{i\in[\![1,l_n]\!]^d} Y_{n,i}\right)^2\right] = \sigma^2, \tag{3.7.2}$$

et

$$\lim_{n\to\infty} \frac{1}{l_n^d} \mathrm{E}\left[\left(\sum_{i\in[\![1,l_n]\!]^d} Y_{n,i}\right)^2 \mathbb{1}\left\{\left|\sum_{i\in[\![1,l_n]\!]^d} Y_{n,i}\right| > n^{d/2}\epsilon\right\}\right] = 0, \tag{3.7.3}$$

pour tout $\epsilon > 0$. Alors

$$\frac{1}{n^{d/2}} \sum_{i\in[\![1,n]\!]^d} Y_{n,i} \xrightarrow{d} \mathcal{N}\big(0, \sigma^2\big) \quad \textit{quand} \quad n \to \infty.$$

où $a, b \in \mathbb{Z}$, $[\![a, b]\!] \equiv \{a, a + 1, ..., b\}$.

Afin de simplifier notre étude sur la normalité asymptotique de notre estimateur, nous commençons par décrire la structure mathématique nécessaire et introduire les notations correspondantes.

Nous supposons que $\Lambda_{n,R}$ se décompose comme $\Lambda_{n,R} = \cup_{i\in\widetilde{I}_n} \Lambda_i$, où Λ_i des sous-cubes de $\Lambda_{n,R}$ et nous supposerons que $\widetilde{I}_n$ croît vers $\mathbb{Z}^d$ et notons $X_i = X_{\Lambda_i}$, $i \in \mathbb{Z}^d$. Nous supposons un cadre où $\{X_i\}_{i\in\mathbb{Z}^d}$ $(d \in \mathbb{N})$ sont des champs aléatoires de la forme suivante :

$$X_i = F(\varepsilon_{i-k}, k \in \mathbb{Z}^d), i \in \mathbb{Z}^d \tag{3.7.4}$$

où les variables aléatoires $\{\varepsilon_i\}_{i\in\mathbb{Z}^d}$ sont des variables aléatoires indépendantes et identiquement distribuées et F une fonction mesurable. Soit $\{\varepsilon_i'\}_{i\in\mathbb{Z}^d}$ une suite de variables aléatoires i.i.d. copie de $\{\varepsilon_i\}_{i\in\mathbb{Z}^d}$. Nous considérons pour tout entier positif n la version associée $X_i^\star$ définie par $X_i^\star = F(\varepsilon_{i-k}^\star, k \in \mathbb{Z}^d)$, où

$$\varepsilon_j^\star = \begin{cases} \varepsilon_j & \text{si} \quad j \neq 0 \\ \varepsilon\prime_0 & \text{si} \quad j = 0. \end{cases}$$

Soient $i \in \mathbb{Z}^d$ et $p > 0$ fixé. Si $X_i \in \mathbb{L}^p$ ($\mathrm{E}\,|X_i|^p$ est fini), nous définissons $\delta_{i,p} = \|X_i - X_i^\star\|_p$, où $\|\cdot\|_p$ est une $\mathbb{L}^p$-norme. Une étude approfondie des propriétés théoriques sur d'autre forme, comme des champs aléatoires linéaires i.e.

$$X_i = \sum_{k\in\mathbb{Z}^d} a_k \varepsilon_{i-k}, i \in \mathbb{Z}^d \tag{3.7.5}$$

où a_k sont des coefficients et les variables aléatoires $\{\varepsilon_i\}_{i\in\mathbb{Z}^d}$ sont à valeurs réelles indépendantes et identiquement distribuées de moyenne nulle et de variance finie, par plusieurs auteurs comme Tjøstheim ([127] ; [128]). Ainsi, El Machkouri [89] a démontré la normalité asymptotique de l'estimateur à noyau de la densité, pour un champ aléatoire de la forme (3.7.4) sous une condition suffisante

$$\sum_{i\in\mathbb{Z}^d} |i|^{5d/2} \delta_{i,p} < \infty. \tag{3.7.6}$$

Hallin et al. [66] obtiennent le même résultat pour un champ aléatoire linéaire de la forme définie en (3.7.5) quand

$$|a_j| \leq C|i|^{-q} \tag{3.7.7}$$

où $q > \max(d+3, 2d+1/2)$ et $n^d b_n^{(2q-1+6d)/(2q-1-4d)}$ tend à l'infini. Ainsi, dans le cas particulier des champs aléatoires linéaires, l'hypothèse (3.7) est plus restrictive que la condition (3.7.7). Wang et Woodroffe [131] ont obtenu une condition suffisante sur les coefficients $(a_j)_{j\in\mathbb{N}^d}$ pour l'estimateur à noyau de la densite d'être asymptotiquement normal. Leur condition est moins restrictive que la condition (3.7). Dans le cas unidimensionnel a été examiné par Castellana et Leadbetter [26], Robinson [113], Bosq et al. [22], Wu et Mielniczuk [134].

Nous nous apparentons aux arguments de El Machkouri [89]. Dans la suite, nous considérons la séquence $(m_n)_{n\geq 1}$ définie par

$$m_n = \max\left\{ v_n, \left[\left(\frac{1}{b_n^3} \sum_{|i|>v_n} |i|^{\frac{5d}{2}} \delta_{i,2} \right)^{\frac{1}{3d}} \right] + 1 \right\}$$

où $v_n = [b_n^{-\frac{1}{2d}}]$ et $[\cdot]$ représente la partie entière. Nos résultats se démontrent essentiellement à travers le lemme suivant de El Machkouri [89] (Lemme 1.2) (est une version spatiale d'un résultat établi par Bosq et al.[22], pages 88-89 et cité dans le papier de El Machkouri [89]).

Lemme 3.7.2. *Si la condition (3.7) est stisfaite, alors*

$$m_n \to \infty, \quad m_n^d b_n^d \to 0 \quad et \quad \frac{1}{(m_n b_n)^{3d/2}} \sum_{|i|>m_n} |i|^{\frac{5d}{2}} \delta_{i,2} \to 0.$$

Nos résultats sont alors les suivants, sous la condition suivante.
Condition $\mathcal{W}$: Il existe une suite $(m_n)_{n\in\mathbb{N}}$ telle que $m_n \to \infty$ quand $n \to \infty$ et la limite suivante est satisfaite :

$$\frac{m_n^d \log^d n}{b_n^d n^d} \to 0.$$

Théorème 3.7.3. *Soit* $\mathbf{X}$ *un processus ponctuel à interaction par paires stationnaire défini en (3.1.1) et soient* $t_1, ..., t_s$, s *points distincts de* $T\backslash\{o\}$. *Sous les conditions* $\mathbf{K}^0$, $\mathcal{Z}$, $\mathcal{W}$, *(3.7.1) et* $b_n^{d+2m}|\Lambda_{n,R}| \to 0$ *quand* $n \to \infty$. *Alors*

$$\sqrt{b_n^d|\Lambda_{n,R}|}\left(\widehat{G}_n(t_i) - \beta^\star G(t_i)\right)_{i=1}^s \xrightarrow{d} \mathcal{N}(0, \sigma_G)$$

avec la matrice de covariance $\sigma_G = [\sigma_{i,j}]_{i,j=1,...,s}$ *où* $\sigma_{i,j}$ *a pour élément*

$$\sigma_{i,i} = \frac{G(t_i)}{\bar{F}(o,t_i)} \int K^2(z)\mathrm{d}z \quad pour\ i = 1, ..., s \quad et \quad \sigma_{i,j} = 0 \quad si \quad i \neq j.$$

La preuve du Théorème 3.7.3 est essentiellement basée sur l'analyse multivariée du théorème de la limite centrale (CLT) pour l'estimateur $\widehat{H}_n(t)$ de $\widetilde{H}(t)$.

Théorème 3.7.4. *Soit* $\mathbf{X}$ *un processus ponctuel à interaction par paires stationnaire défini en (3.1.1) et soient* $t_1, ..., t_s$, s *points distincts de* $T\backslash\{o\}$. *Sous les conditions* $\mathbf{K}^0$, $\mathcal{Z}$, $\mathcal{W}$ *et (3.7.1). Alors le vecteur ligne*

$$\sqrt{b_n^d|\Lambda_{n,R}|}\left(\widehat{H}_n(t_i) - \mathrm{E}\,\widehat{H}_n(t_i)\right)_{i=1}^s \xrightarrow{d} \mathcal{N}(0, \sigma_H)$$

avec la matrice de covariance $\sigma_H = [r_{i,j}]_{i,j=1,...,s}$ *où* $r_{i,j}$ *a pour élément*

$$r_{i,i} = \beta^{\star 2}\bar{F}(o,t_i)G(t_i) \int_{\mathbb{R}^d} K^2(z)\mathrm{d}z, \quad pour\ i = 1, ..., s \quad et \quad r_{i,j} = 0 \quad si \quad i \neq j.$$

Corollaire 3.7.5. Sous les conditions du Théorème 3.7.4 et soit $b_n^{d+2m}|\Lambda_{n,R}| \to 0$ quand $n \to +\infty$. De plus si $\widetilde{H}$ admet des dérivées partielles d'ordre m continues et bornées au voisinage de t_i pour $i = 1, ..., s$. Alors

$$\sqrt{b_n^d|\Lambda_{n,R}|}\Big(\widehat{H}_n(t_i) - \widetilde{H}(t_i)\Big)_{i=1}^{s} \xrightarrow{d} \mathcal{N}(0, \sigma_H)$$

avec la matrice de covariance $\sigma_H = [r_{i,j}]_{i,j=1,...,s}$ où $r_{i,j}$ a pour élément

$$r_{i,i} = \beta^{\star 2}\bar{F}(o,t_i)G(t_i)\int_{\mathbb{R}^d} K^2(z)\mathrm{d}z, \quad \text{pour } i = 1, ..., s \quad \text{et} \quad r_{i,j} = 0 \quad \text{si} \quad i \neq j.$$

3.7.1 Preuve du Théorème 3.7.4

Démonstration. Nos résultats de normalité asymptotique pour notre estimateur par fenêtres glissantes suivent directement par l'argument de la méthode de Cramér-Wold (1936), sans perte de généralité, nous considérons que le cas $s = 2$. Soient λ_1 et λ_2 deux constantes telles que $\lambda_1 + \lambda_2 \neq 0$ et soient $t_1, t_2 \in T\backslash\{o\}$ tels que $t_1 \neq t_2$ et $|\Lambda_{n,R}| \sim cn^d$. Notons

$$\lambda_1(b_n^d n^d)^{1/2}\Big(\widehat{H}_n(t_1) - \mathrm{E}[\widehat{H}_n(t_1)]\Big) + \lambda_2(b_n^d n^d)^{1/2}\Big(\widehat{H}_n(t_2) - \mathrm{E}[\widehat{H}_n(t_2)]\Big) = \frac{1}{n^{d/2}}\sum_{i\in\widetilde{I}_n}\bar{Y}_{n,i}$$

où

$$\bar{Y}_{n,i} = \lambda_1\Big(Y_i(t_1) - \mathrm{E}[Y_i(t_1)]\Big) + \lambda_2\Big(Y_i(t_2) - \mathrm{E}[Y_i(t_2)]\Big), i \in \mathbb{Z}^d \tag{3.7.8}$$

pour tout $j = 1, 2$, nous considérons

$$Y_i(t_j) = (b_n^d)^{-1/2}\sum_{\substack{u\in X_{\Lambda_i}, v\in\mathbf{X}\\ v-u\in B(o,R)}}^{\neq} h(u, \mathbf{X}\setminus\{u,v\})h(v, \mathbf{X}\setminus\{u,v\})K\Big(\frac{v-u-t_j}{b_n}\Big).$$

Nous allons adopter une approche m-approximation, pour démontrer le Théorème 3.7.4, nous construisons une suite de variables aléatoires réelles centrées $\{\widetilde{Y}_{n,i}; i \in \mathbb{Z}^d\}$ définie par

$$\widetilde{Y}_{n,i} = \lambda_1\Big(\widetilde{Y}_i(t_1) - \mathrm{E}[\widetilde{Y}_i(t_1)]\Big) + \lambda_2\Big(\widetilde{Y}_i(t_2) - \mathrm{E}[\widetilde{Y}_i(t_2)]\Big), \ i \in \mathbb{Z}^d$$

pour tout $j = 1, 2$, nous considérons

$$\widetilde{Y}_i(t_j) = (b_n^d)^{-1/2}\sum_{\substack{u\in X_{\Lambda_{i-k}}, v\in\mathbf{X}\\ v-u\in B(o,R)}}^{\neq} h(u, \mathbf{X}\setminus\{u,v\})h(v, \mathbf{X}\setminus\{u,v\})K\Big(\frac{v-u-t_j}{b_n}\Big).$$

Nous considérons, $\{\widetilde{Y}_{n,i}\}_{i\in\mathbb{Z}^d}$ sont m_n- dépendant. Nous allons utiliser $\{\widetilde{Y}_{n,i}; i\in\mathbb{Z}^d\}_{n\in\mathbb{N}}$ pour approximer $\{\bar{Y}_{n,i}; i\in\mathbb{Z}^d\}_{n\in\mathbb{N}}$. Nous décomposons donc

$$\frac{1}{n^{d/2}}\sum_{i\in\widetilde{I}_n}\bar{Y}_{n,i}=\frac{1}{n^{d/2}}\sum_{i\in\widetilde{I}_n}\widetilde{Y}_{n,i}+\frac{1}{n^{d/2}}\sum_{i\in\widetilde{I}_n}\big(\bar{Y}_{n,i}-\widetilde{Y}_{n,i}\big).$$

Alors pour aboutir à notre résultat, il suffit d'établir les deux résultats suivants.

Proposition 3.7.6. *Sous les conditions $\mathcal{W}$ et (3.7.1). Pour n suffisament grand, on a*

$$\frac{1}{n^{d/2}}\sum_{i\in\widetilde{I}_n}\widetilde{Y}_{n,i}\xrightarrow{d}\mathcal{N}\big(0,\sigma^2\big),$$

où

$$\sigma^2=\Big(\beta^{\star 2}\lambda_1^2\bar{F}(o,t_1)G(t_1)+\beta^{\star 2}\lambda_2^2\bar{F}(o,t_2)G(t_2)\Big)\int_{\mathbb{R}^d}K^2(z)\mathrm{d}z. \tag{3.7.9}$$

Proposition 3.7.7. *Sous les conditions du Lemme 3.7.2. On a pour n suffisament grand,*

$$\frac{1}{n^{d/2}}\sum_{i\in\widetilde{I}_n}\big(\bar{Y}_{n,i}-\widetilde{Y}_{n,i}\big)\xrightarrow{\mathbb{P}}0.$$

Preuve de la Proposition 3.7.6

Démonstration. Pour conclure à la convergence en loi de $\frac{1}{n^{d/2}}\sum_{i\in\widetilde{I}_n}\widetilde{Y}_{n,i}$ vers une variable gaussienne, nous devons vérifier que les deux conditions fondamentales (3.7.2) et (3.7.3) du Théorème 3.7.1, étant donné la suite $\{\widetilde{Y}_{n,i}\}_{i\in\mathbb{Z}^d}$ est m_n- dépendant.

Preuve de la Condition (3.7.2)

Soit $l_n=m_n\log n$ tel que $m_n/l_n\to 0$ et $l_n/n\to 0$ quand $n\to\infty$. Nous déduisons facilement par la définition de $\sum_{i\in[\![1,l_n]\!]^d}\widetilde{Y}_{n,i}$ i.e.

$$\frac{1}{l_n^{d/2}}\sum_{i\in[\![1,l_n]\!]^d}\widetilde{Y}_{n,i}=\lambda_1(b_n^dl_n^d)^{1/2}\Big(\widehat{H}_{l_n}(t_1)-\mathrm{E}[\widehat{H}_{l_n}(t_1)]\Big)+\lambda_2(b_n^dl_n^d)^{1/2}\Big(\widehat{H}_{l_n}(t_2)-\mathrm{E}[\widehat{H}_{l_n}(t_2)]\Big)$$

où pour $j=1,2$

$$\widehat{H}_{l_n}(t_j)=\frac{1}{b_n^dl_n^d}\sum_{\substack{u,v\in\mathbf{X}\\ v-u\in B(o,R)}}^{\neq}\mathbb{1}_{[\![1,l_n]\!]^d}(u)h(u,\mathbf{X}\setminus\{u,v\})h(v,\mathbf{X}\setminus\{u,v\})K\Big(\frac{v-u-t_j}{b_n}\Big)$$

que

$$\lim_{n\to\infty} \frac{1}{l_n^d} \operatorname{Var}\left(\sum_{i\in[\![1,l_n]\!]^d} \widetilde{Y}_{n,i} \right)$$
$$= \lim_{n\to\infty} b_n^d l_n^d \left(\lambda_1^2 \operatorname{Var}(\widehat{H}_{l_n}(t_1)) + \lambda_2^2 \operatorname{Var}(\widehat{H}_{l_n}(t_2)) + 2\lambda_1\lambda_2 \operatorname{Cov}(\widehat{H}_{l_n}(t_1), \widehat{H}_{l_n}(t_2)) \right).$$

Le Théorème 3.2.2 entraîne en remplaçant $\widehat{H}_n(t_j)$ par $\widehat{H}_{l_n}(t_j)$ que

$$\lim_{n\to\infty} b_n^d l_n^d \operatorname{Var}(\widehat{H}_{l_n}(t_j)) = \beta^{\star 2} G(t_j) \bar{F}(o, t_j) \int_{\mathbb{R}^d} K^2(u) \mathrm{d}u.$$

Ainsi, la preuve de

$$\lim_{n\to\infty} b_n^d l_n^d \operatorname{Cov}\left(\widehat{H}_{l_n}(t_1), \widehat{H}_{l_n}(t_2) \right) = 0$$

est justifié de la même façon en appliquant le Théorème 3.2.3. Ainsi, il vient alors que

$$\lim_{n\to\infty} \frac{1}{l_n^d} \operatorname{Var}\left(\sum_{i\in[\![1,l_n]\!]^d} \widetilde{Y}_{n,i} \right) = \sigma^2$$

où σ^2 est définie en (3.7.9).

Preuve de la Condition (3.7.3)

Notons $\xi_n = \sum_{i\in[\![1,l_n]\!]^d} \widetilde{Y}_{n,i}$, d'après l'inégalité de Cauchy-Schwarz et l'inégalité de Bienaymé-Tchebychev, nous avons

$$\begin{aligned} \mathrm{E}\left[\xi_n^2 \mathbb{1}\{|\xi_n| > \varepsilon n^{d/2}\}\right] &\leq ||\xi_n||_p^2 \big(\mathbb{P}(|\xi_n| > n^{d/2}\varepsilon)\big)^{(p-2)/p} \\ &\leq ||\xi_n||_p^2 \left(\frac{||\xi_n||_2^2}{n^d \varepsilon^2} \right)^{(p-2)/p}. \end{aligned}$$

D'après la condition (3.7.1), il résulte que

$$\|\xi_n\|_2 \leq C l_n^{d/2} \quad \text{pour tout} \quad n \in \mathbb{N}.$$

Pour $p > 2$ et puisque K est borné, nous avons

$$\|\xi_n\|_p \leq C \left(\frac{1}{b_n^{(p-2)/2}} \|\xi_n\|_2^2 \right)^{1/p} \leq C l_n^d b_n^{-(p-2)/p}.$$

Nous déduisons facilement à partir des majorations de ce qui précède,

$$\frac{1}{l_n^d} \mathrm{E}\left[\xi_n^2 \mathbb{1}\{|\xi_n| > \epsilon n^{d/2}\}\right] \leq C \left(\frac{l_n^d}{b_n^d n^d} \right)^{(p-2)/p}$$

sous la Condition $\mathcal{W}$ pour n assez grand, la condition (3.7.3) est satisfaite. □

Preuve de la Proposition 3.7.7

Démonstration. Soit $p > 1$ fixé. Nous suivons la preuve du Lemme 1.3 de El Machkouri [89]. Pour $i \in \mathbb{Z}^d$ et pour $t \in T$, notons $R_i = Y_i(t) - \widetilde{Y}_i(t)$. Comme il existe une fonction mesurable F telle que $R_i = F(\varepsilon_{i-s}; s \in \mathbb{Z}^d)$. Notons $\delta_{i,p}^{(n)} = \|R_i - R_i^\star\|_p$ où $R_i^\star = F(\varepsilon_{i-s}^\star; s \in \mathbb{Z}^d)$, avec

$$\varepsilon_j^\star = \begin{cases} \varepsilon_j & \text{si} \quad j \neq 0 \\ \varepsilon\prime_0 & \text{si} \quad j = 0. \end{cases}$$

Par application l'inégalité de Burkholder (c.f Hall Heyde [65] , page 23) et après un calcul techniques (voir El Machkouri [89]), nous obtenons :

$$\left\| \sum_{i \in \widetilde{I}_n} a_i R_i \right\|_p \leq \left(2p \sum_{i \in \widetilde{I}_n} a_i^2 \right)^{1/2} \sum_{i \in \mathbb{Z}^d} \delta_{i,p}^{(n)}.$$

Etant donné $(Y_i(t) - \widetilde{Y}_i(t))^\star = Y_i^\star(t) - \widetilde{Y}_i^\star(t)$, nous déduisons que $\delta_{i,p}^{(n)} \leq 2\|Y_i(t) - Y_i^\star(t)\|_p$. Comme le noyau K est une fonction lipschitzienne, nous obtenons

$$\delta_{i,p}^{(n)} \leq \frac{2c\delta_{i,p}}{b_n^d} \tag{3.7.10}$$

où $\delta_{i,p} = \|X_i - X_i^\star\|_p$. Par ailleurs, nous avons aussi

$$\delta_{i,p}^{(n)} \leq 2\|Y_0(t) - \widetilde{Y}_0(t)\|_p. \tag{3.7.11}$$

Combinons (3.7.11) et le Lemme 3.3 de El Machkouri [89], ou l'inégalité de Burkholder (c.f Hall Heyde [65] , page 23), pour obtenir

$$\delta_{i,p}^{(n)} \leq \frac{\sqrt{8p}}{b_n^d} \sum_{|j| > m_n} \delta_{i,p}. \tag{3.7.12}$$

Combinant (3.7.10) et (3.7.11), nous avons

$$\begin{aligned} \sum_{i \in \mathbb{Z}^d} \delta_{i,p}^{(n)} &\leq \frac{m_n^d \sqrt{8p}}{b_n^d} \sum_{|j| > m_n} \delta_{i,p} + \frac{2}{b_n^d} \sum_{|j| > m_n} \delta_{i,p} \\ &\leq \frac{2m_n^d \sqrt{8p}}{b_n^d} \sum_{|j| > m_n} \delta_{i,p}. \end{aligned} \tag{3.7.13}$$

Finalement, d'après (3.7.13) et sous les conditions du Lemme 3.7.2, nous obtenons

$$\frac{1}{n^{d/2}} \left\| \sum_{i \in \widetilde{I}_n} \left(\bar{Y}_{n,i} - \widetilde{Y}_{n,i} \right) \right\|_2 \leq c \frac{|\lambda_1| + |\lambda_2|}{(m_n b_n)^{3d/2}} \sum_{|i| > m_n} |i|^{\frac{5d}{2}} \delta_{i,2} = o(1).$$

□

Nous avons bien le résultat souhaité. □

3.7.2 Preuve du Corollaire 3.7.5

Démonstration. Par la continuité de la fonction $\widetilde{H}$ au voisinage des points $t_1, ..., t_s$ et par le fait que $\widetilde{H}(t_i)$ admet des dérivées partielles d'ordre m continues et bornées au voisinage de t_i pour $i = 1, ..., s$, par conséquent nous obtenons par l'équation (3.2.1)

$$\sqrt{b_n^d|\Lambda_{n,R}|}\big(\mathrm{E}\,\widehat{H}_n(t_i) - \widetilde{H}(t_i)\big) = \mathcal{O}\Big(\sqrt{b_n^{d+2m}|\Lambda_{n,R}|}\Big).$$

Sous la condition que $b_n^{d+2m}|\Lambda_{n,R}| \longrightarrow 0$, cela avec l'application Théorème 3.7.4, nous concluons donc le résultat annoncé. □

3.7.3 Preuve du Théorème 3.7.3

Démonstration. Pour établir la preuve du Théorème 3.7.3, nous fixons à cet effet $t \in \{t_1, ..., t_s\}$. Nous utilisons de manière cruciale la décomposition suivante

$$\begin{aligned}
\widehat{G}_n(t) - \beta^\star G(t) &= \frac{L(t)\widehat{H}_n(t) - \widetilde{H}(t)\widehat{L}_n(t)}{L(t)\widehat{L}_n(t)} \\
&= \frac{\widehat{H}_n(t) - \mathrm{E}\,\widehat{H}_n(t)}{\widehat{L}_n(t)} + \frac{\mathrm{E}\,\widehat{H}_n(t) - \beta^\star G(t)\widehat{L}_n(t)}{\widehat{L}_n(t)} \\
&= A_n^{(1)} + A_n^{(2)}
\end{aligned}$$

Nous appliquons le Théorème 3.7.4 avec la consistance forte de l'estimateur $\widehat{L}_n(t)$, ceux-ci impliquent que $\sqrt{b_n^d|\Lambda_{n,R}|}A_n^{(1)}$ tend vers une distribution gaussienne de moyenne nulle. L'application de la Proposition 3.2.1 conduit à $\sqrt{b_n^d|\Lambda_{n,R}|}\big(\mathrm{E}\,\widehat{H}_n(t) - \widetilde{H}(t)\big) \underset{n\to\infty}{\longrightarrow} 0$, sous la condition que $b_n^{d+2m}|\Lambda_{n,R}| \to 0$, avec la consistance forte de l'estimateur $\widehat{L}_n(t)$, ceux-ci impliquent que quand $n \to \infty$,

$$\sqrt{b_n^d|\Lambda_{n,R}|}A_n^{(2)} \xrightarrow{\mathbb{P}} 0.$$

Cela conclut la preuve □

Chapitre 4

Estimation non paramétrique de fonction d'interaction de paires d'un processus ponctuel à interaction par paires stationnaire et isotrope

Résumé

Nous adaptons un estimateur au modèle à interaction par paires stationnaire et isotrope, nous allons esquisser quelques preuves et souligner les principales différences apportées dans le cas ou le modèle à interaction par paires est uniquement stationnaire étudié dans le chapitre précédent. Nous présentons dans un premier temps le modèle que nous étudions, nous donnons des hypothèses techniques sur le modèle et nous définissons ensuite l'expression de l'estimateur non paramétrique que nous considérons dans la Section 4.1. Dans la Section 4.2, nous étudions les comportements asymptotiques de la moyenne et de la variance de l'estimateur non paramétrique par lissage qui nous intéresse. La consistance forte ponctuelle et la consistance forte globale des estimateurs obtenus avec des différentes vitesses de convergence sont étudiées dans les Sections 4.4 et 4.5. Nous évoquerons ensuite la normalité asymptotique de notre estimateur dans la Section 4.6.

4.1 Hypothèses et présentation de l'estimateur

Le processus ponctuel à interaction par paires caractérisé par son intensité conditionnelle de Papangelou définie par

$$\lambda(u, \mathbf{x}) = g_0(u) \exp\left(-\sum_{v \in \mathbf{x} \backslash u} g_0(\{u, v\})\right),$$

est appellé un processus ponctuel à interaction par paires stationnaire et isotrope, si $g_0(u) = \beta$ est une constante et $g_0(\{u, v\}) = \gamma(||u - v||)$ est invariante par translation et par rotation. L'interprétation du potentiel qui correspond uniquement à des interactions de paires est la suivante :

- $\gamma(r) > 0$ traduit une probabilité de trouver deux points à distance r, nous disons qu'il y a répulsion (ou inhibition) à distance r.
- $\gamma(r) < 0$ traduit une forte probabilité de trouver deux points à distance r, nous disons qu'il y a attraction à distance r.
- $\gamma(r) = 0$ traduit une répartition aléatoire.

Dans la suite de ce chapitre, nous considèrons le modèle d'interaction de paires stationnaire et isotrope caractérisé par son intensité conditionnelle de Papangelou définie par

$$\lambda_{\beta^\star}(u, \mathbf{x}) = \beta^\star \exp\left(-\sum_{v \in \mathbf{x} \backslash u} \gamma(||v - u||)\right), \ \forall u \in \mathbb{R}^d, \mathbf{x} \in \Omega \tag{4.1.1}$$

où $\beta^\star$ est la vraie valeur du paramètre d'interaction du premier ordre. γ est le potentiel d'interaction de paires, défini sur $\mathbb{R}^+$ à valeurs dans $\mathbb{R}$, nous supposons qu'il admet une portée d'interaction finie R, i.e.

$$\gamma(||v - u||) = 0 \quad \text{pour } ||v - u|| > R, \ \forall u \in \mathbb{R}^d, v \in \mathbf{x} \backslash u \in \Omega \tag{4.1.2}$$

et nous supposons que

$$\gamma(r) > 0, \quad \text{si} \quad r \in (0, R] \quad \text{pour} \quad R > 0. \tag{4.1.3}$$

Cette dernière condition implique que les réalisations typiques seront plus ou moins régulières par rapport à un dispositif complètement aléatoire. En particulier, des potentiels négatifs (ou partiellement négatifs) peuvent conduire à des distributions irrégulières ou instables (Kelly et Ripley [81] ; Gates et Westcott [55]).

Nous introduisons en outre une définition que nous utiliserons par la suite.

Définition 4.1.1. *On appelle* $K_1 : \mathbb{R} \longrightarrow \mathbb{R}$ *un noyau d'ordre* $\alpha \geq 1$ *si les fonctions* $u \mapsto u^j K_1(u), j = 0, 1, ..., \alpha$ *sont intégrables et si*

$$\int_{\mathbb{R}} K_1(u)\mathrm{d}u = 1, \int_{\mathbb{R}} u^j K_1(u)\mathrm{d}u = 0, j = 0, 1, ..., \alpha - 1, \quad \textit{pour } \alpha \geq 2.$$

Exemples 4.1.2. Les noyaux les plus couramment utilisés en pratique, voir Nadaraja [99], Silverman [120] sont :

- Le noyau d'Epanechnikov,

$$K_1(u) = \frac{d(d+2)\Gamma(d/2)}{4\pi^{d/2}R^d}\left(1 - \frac{||u||^2}{R^2}\right)\mathbb{1}_{B(o,R)}(u).$$

- Le noyau "conus",

$$K_1(u) = \frac{d(d+1)\Gamma(d/2)}{2\pi^{d/2}R^d}\left(1 - \frac{||u||}{R}\right)\mathbb{1}_{B(o,R)}(u).$$

Le noyau de Epanechnikov univariée (Fiksel [52], D. Stoyan et H. Stoyan [122]) est recommandé en dimension 1. Nous avons besoin aussi d'imposer les conditions suivantes tout au long de ce travail : Soit $(b_n)_{n\geq 1}$ une suite de réels strictement positifs telle que

$$b_n \underset{n\to\infty}{\longrightarrow} 0, \quad b_n|\Lambda_{n,R}| \underset{n\to\infty}{\longrightarrow} \infty.$$

L'évaluation de l'estimateur de la fonction d'interaction de paires, exige l'observation du processus ponctuel à interaction par paires stationnaire et isotrope noté $\mathbf{X}$ dans un domaine borné de $\mathbb{R}^d$ noté Λ_n. Supposons que la suite $(\Lambda_n)_{n\geq 1}$ tend vers $\mathbb{R}^d$ quand n tend vers l'infini et un noyau univarié K_1. Alors l'estimateur de $\beta^\star\Phi(r)$, où $\Phi(r) = \exp(-\gamma(r))$ corrigeant les effets de bord par l'introduction du $2R$-intérieur de Λ_n, i.e.

$$\Lambda_{n,R} = \{u \in \Lambda_n : B(u, 2R) \subset \Lambda_n\}, \quad \text{avec} \quad |\Lambda_{n,R}| > 0$$

est défini pour $r \in (0, R]$ par

$$\widehat{\Phi}_n(r) = \frac{\widehat{R}_n(r)}{\widehat{N}_n(r)} \tag{4.1.4}$$

où

(4.1.5)

$$\widehat{R}_n(r) = \frac{1}{b_n|\Lambda_{n,R}|\sigma_d} \sum_{\substack{u,v\in\mathbf{X} \\ ||v-u||\leq R}}^{\neq} \frac{\mathbb{1}_{\Lambda_{n,R}}(u)}{||v-u||^{d-1}} h(u, \mathbf{X}\backslash\{u,v\})h(v, \mathbf{X}\backslash\{u,v\})K_1\left(\frac{||v-u|| - r}{b_n}\right)$$

est un estimateur de $\beta^{\star 2}\Phi(r)J(r)$, où $J(r) = \int_{\mathbb{S}^{d-1}} \bar{F}(o, rv)\mathrm{d}v$.

$$\widehat{N}_n(r) = \frac{1}{|\Lambda_{n,R}|} \sum_{u \in \mathbf{X}} \mathbb{1}_{\Lambda_{n,R}}(u) h(u, \mathbf{X} \setminus \{u\}) h^\star(u, \mathbf{X} \setminus \{u\}) \tag{4.1.6}$$

estimateur de $\beta^\star J(r)$. Où h une fonction mesurable non négative définie pour tout $u \in \mathbb{R}^d$, $\mathbf{x} \in \Omega$ par

$$h(u, \mathbf{x}) = \mathbb{1}\left(\inf_{v \in \mathbf{x}} ||v - u|| > R\right) = \mathbb{1}\left(d(u, \mathbf{x}) > R\right)$$

$$h^\star(u, \mathbf{x}) = \int_{\mathbb{S}^{d-1}} h(rv - u, \mathbf{x})\mathrm{d}v$$

et $\sigma_d = \frac{2\pi^{d/2}}{\Gamma(d/2)}$ représente la mesure de la sphère unité de $\mathbb{R}^d$.

4.2 Comportements asymptotiques de la moyenne et de la variance de l'estimateur $\widehat{R}_n(r)$

Nous commençons avec les comportements asymptotiques de la moyenne et de la variance de l'estimateur $\widehat{R}_n(r)$.

Proposition 4.2.1. *Soit γ le potentiel d'interaction de paires défini en (4.1.1) vérifiant la condition de portée d'interaction de paires finie (4.1.2). Soit K_1 un noyau non négatif à support compact d'ordre $\alpha \geq 1$. Alors pour tout $r \in (0, R]$*

$$\lim_{n \to \infty} \mathrm{E}\, \widehat{R}_n(r) = \beta^{\star 2} J(r) \exp(-\gamma(r)).$$

De plus si la fonction $\exp(-\gamma(r))J(r)$ admet des dérivées partielles d'ordre α continues et bornées sur $(r - \delta, r + \delta)$ pour un certain $\delta > 0$ et pour tout $\alpha \geq 1$. Alors

$$\mathrm{E}\, \widehat{R}_n(r) = \beta^{\star 2} J(r) \exp(-\gamma(r)) + \mathcal{O}(b_n^\alpha) \quad \textit{quand} \;\; n \to \infty.$$

Théorème 4.2.2. *Soit γ le potentiel d'interaction de paires défini en (4.1.1) vérifiant la condition de portée d'interaction de paires finie (4.1.2). Soit K_1 un noyau non négatif à support compact d'ordre $\alpha \geq 1$, tel que $\int_{\mathbb{R}} K_1^2(\rho)\mathrm{d}\rho < \infty$. Alors pour tout $r \in (0, R]$,*

$$\lim_{n \to \infty} b_n |\Lambda_{n,R}| \operatorname{Var}(\widehat{R}_n(r)) = \frac{2\beta^{\star 2}}{\sigma_d r^{d-1}} J(r) \exp(-\gamma(r)) \int_{\mathbb{R}} K_1^2(\rho)\mathrm{d}\rho.$$

4.2.1 Preuve de la Proposition 4.2.1

Démonstration. Nous établissons le comportement asymptotique de la moyenne de l'estimateur $\widehat{R}_n(r)$. Tout d'abord, précisons quelques notations. Notons

$$H(u_1, ..., u_s, \mathbf{X}) = h(u_1, \mathbf{X})...h(u_s, \mathbf{X}), \quad I(||u_1||, ..., ||u_s||) = \mathbb{1}(||u_1|| \leq R, ..., ||u_s|| \leq R)$$

$$\bar{F}(u_1, ..., u_s) = \mathrm{E}[h(u_1, \mathbf{X}) \cdots h(u_s, \mathbf{X})].$$

En utilisant la formule de GNZ (1.4.3) du second ordre, la stationnarité du processus ponctuel $\mathbf{X}$ et en appliquant la propriété de la portée d'interaction finie (4.1.2), nous obtenons :

$$\begin{aligned}
\mathrm{E}\,\widehat{R}_n(r) &= \frac{1}{b_n|\Lambda_{n,R}|\sigma_d} \\
&\quad \mathrm{E}\int_{\mathbb{R}^{2d}} \frac{\mathbb{1}_{\Lambda_{n,R}}(u)}{||v-u||^{d-1}} I(||v-u||)H(u,v,\mathbf{X})K_1\left(\frac{||v-u||-r}{b_n}\right)\lambda_{\beta^\star}(u,v,\mathbf{X})\mathrm{d}u\mathrm{d}v \\
&= \frac{\beta^{\star 2}}{b_n|\Lambda_{n,R}|\sigma_d} \\
&\quad \mathrm{E}\int_{\mathbb{R}^{2d}} \frac{\mathbb{1}_{\Lambda_{n,R}}(u)}{||v-u||^{d-1}} I(||v-u||)H(u,v,\mathbf{X})K_1\left(\frac{||v-u||-r}{b_n}\right)\Phi(||v-u||)\mathrm{d}u\mathrm{d}v \\
&= \frac{\beta^{\star 2}}{b_n\sigma_d}\int_{\mathbb{R}^d} \frac{I(||s||)}{||s||^{d-1}}\,\mathrm{E}[H(o,s,\mathbf{X})]K_1\left(\frac{||s||-r}{b_n}\right)\Phi(||s||)\mathrm{d}s.
\end{aligned}$$

Rappelons une propriété de la théorie d'intégration (voir Briane et Pagès [24] ou Rudin [116]). Soit $\mathbb{S}^{d-1}$ la sphère unité de $\mathbb{R}^d$:

$$\mathbb{S}^{d-1} = \{u \in \mathbb{R}^d : ||u|| = 1\},$$

alors pour toute fonction borélienne $f : \mathbb{R}^d \longrightarrow \mathbb{R}_+$,

$$\int_{\mathbb{R}^d} f(u)\mathrm{d}u = \int_0^\infty \int_{\mathbb{S}^{d-1}} f(rz)r^{d-1}\sigma_d \mathrm{d}r\mathrm{d}z.$$

En combinant le résultat ci-dessus, nous obtenons donc :

$$\mathrm{E}\,\widehat{R}_n(r) = \frac{\beta^{\star 2}}{\sigma_d}\int_{-r/b_n}^\infty \int_{\mathbb{S}^{d-1}} I(b_n\varrho + r)\bar{F}(o,(b_n\varrho+r)v)K_1(\varrho)\Phi(b_n\varrho+r)\sigma_d\mathrm{d}\varrho\mathrm{d}v.$$

Par application du théorème de convergence dominé et sous les hypothèses sur le noyau K_1, d'où la première assertion de la Proposition 4.2.1.

Nous intéressons maintenant aux résultats de la seconde partie de la Proposition 4.2.1. Nous considérons pour un certain $\delta > 0$, $(r-\delta, r+\delta)$ un ouvert de $\mathbb{R}$, la fonction $\exp(-\gamma(r))J(r)$ admet des dérivées partielles d'ordre α continues et bornées en chaque point de $(r-\delta, r+\delta)$, alors pour tout $r \in (0, R]$, il existe $\theta \in]0,1[$, donc suivant la formule de Taylor-Lagrange, nous avons pour $n \to \infty$,

$$\Phi(b_n\varrho + r) = \Phi(r) + \sum_{k=1}^{\alpha-1} \frac{(b_n\varrho)^k}{k!}\frac{d\Phi}{dr}(r)dr + \frac{b_n^\alpha}{\alpha!}\frac{d^\alpha\Phi}{dr^\alpha}(r + b_n\varrho\theta).$$

$$\bar{F}(o,(b_n\varrho + r)v) = \bar{F}(o,rv) + \sum_{k=1}^{\alpha-1} \frac{(b_n\varrho)^k}{k!}\frac{d\bar{F}(o,rv)}{dr}dr + \frac{b_n^\alpha}{\alpha!}\frac{d^\alpha\bar{F}}{dr^\alpha}(o,(r + b_n\varrho\theta)v).$$

le produit admet un développement limité d'ordre α dont la partie régulière s'obtient en prenant dans le produit des parties régulières les termes de degré inférieure ou égale à α. Le produit $\bar{F}(o,(b_n\varrho + r)v)\Phi(b_n\varrho + r)$ se mettra donc sous la forme

$$\bar{F}(o,(b_n\varrho + r)v)\Phi(b_n\varrho + r) = \bar{F}(o,rv)\Phi(r) + \sum_{k=1}^{\alpha-1} T_n(rv,r)(b_n\varrho)^k + \mathcal{O}(b_n^\alpha) \quad \text{quand } n \to \infty.$$

En appliquant le théorème de convergence dominée, il résulte que pour $r \in (0, R]$,

$$\begin{aligned}\mathrm{E}\,\widehat{R}_n(r) &= \beta^{\star 2}\Phi(r)J(r) \\ &\quad + \beta^{\star 2}\int_{\mathbb{S}^{d-1}} \sum_{k=1}^{\alpha-1} T_n(rv,r)b_n^k \mathrm{d}v \int_{\mathbb{R}} \varrho^k K_1(\varrho)\mathrm{d}\varrho \\ &\quad + \mathcal{O}(b_n^\alpha) \quad \text{quand } n \to \infty.\end{aligned}$$

Les hypothèses sur le noyau K_1 terminent la preuve. □

4.2.2 Preuve du Théorème 4.2.2

Démonstration. Nous déterminons le comportement asymptotique de la variance de

$$\widehat{R}_n(r) = \frac{1}{b_n|\Lambda_{n,R}|\sigma_d} \sum_{\substack{u,v\in\mathbf{X} \\ ||v-u||\le R}}^{\neq} \frac{\mathbb{1}_{\Lambda_{n,R}}(u)}{||v-u||^{d-1}} h(u,\mathbf{X}\setminus\{u,v\})h(v,\mathbf{X}\setminus\{u,v\})K_1\left(\frac{||v-u|| - r}{b_n}\right).$$

La preuve de ce résultat est obtenue en utilisant le corollaire suivant.

Corollaire 4.2.3. Sous les conditions du Lemme 3.2.5 et si de plus la fonction f est symétrique. Alors

$$\begin{aligned}
&\operatorname{Var}\Big(\sum_{u,v\in\mathbf{X}}^{\neq} f(u,v,\mathbf{X}\backslash\{u,v\})\Big)\\
&= 2\,\mathrm{E}\int_{\mathbb{R}^{2d}} f^2(u,v,\mathbf{X})\lambda(u,v,\mathbf{X})\mathrm{d}u\mathrm{d}v\\
&+ 4\,\mathrm{E}\int_{\mathbb{R}^{3d}} f(u,v,\mathbf{X})f(v,w,\mathbf{X})\lambda(u,v,w,\mathbf{X})\mathrm{d}u\mathrm{d}v\mathrm{d}w\\
&+ \mathrm{E}\int_{\mathbb{R}^{4d}} f(u,v,\mathbf{X})f(w,y,\mathbf{X})\lambda(u,v,w,y,\mathbf{X})\mathrm{d}u\mathrm{d}v\mathrm{d}w\mathrm{d}y\\
&- \int_{\mathbb{R}^{4d}} \mathrm{E}[f(u,v,\mathbf{X})\lambda(u,v,\mathbf{X})]\,\mathrm{E}[f(w,y,\mathbf{X})\lambda(w,y,\mathbf{X})]\mathrm{d}u\mathrm{d}v\mathrm{d}w\mathrm{d}y.
\end{aligned}$$

Démonstration. En utilisant la formule (3.6.8) et en tenant compte du fait que f est symétrique, nous obtenons

$$\begin{aligned}
\Big(\sum_{u,v\in\mathbf{X}}^{\neq} f(u,v,\mathbf{X}\backslash\{u,v\})\Big)^2 &= 2\sum_{u,v\in\mathbf{X}}^{\neq} f^2(u,v,\mathbf{X}\backslash\{u,v\})\\
&+ 4\sum_{u,v,w\in\mathbf{X}}^{\neq} f(u,v,\mathbf{X}\backslash\{u,v,w\})f(v,w,\mathbf{X}\backslash\{u,v,w\})\\
&+ \sum_{u,v,w,y\in\mathbf{X}}^{\neq} f(u,v,\mathbf{X}\backslash\{u,v,w,y\})f(w,y,\mathbf{X}\backslash\{u,v,w,y\}). \qquad (4.2.1)
\end{aligned}$$

D'autre part, nous avons

(4.2.2)

$$\operatorname{Var}\sum_{u,v\in\mathbf{X}}^{\neq} f(u,v,\mathbf{X}\backslash\{u,v\}) = \mathrm{E}\big(\sum_{u,v\in\mathbf{X}}^{\neq} f(u,v,\mathbf{X}\backslash\{u,v\})\big)^2 - \big(\mathrm{E}\sum_{u,v\in\mathbf{X}}^{\neq} f(u,v,\mathbf{X}\backslash\{u,v\})\big)^2$$

et par conséquent, le résultat du Corollaire (4.2.3) découle, d'une application directe des versions itérées de la formule de GNZ (1.4.3) dans (4.2.1) combinée avec (4.2.2). □

Nous déduisons du corollaire précédent l'expression de la variance de l'estimateur $\widehat{R}_n(r)$ définie par

$$\operatorname{Var}\widehat{R}_n(r) = A_1 + A_2 + A_3 - A_4,$$

pour

$$f(u,v,\mathbf{X}) = \frac{\mathbb{1}_{\Lambda_{n,R}}(u)}{||v-u||^{d-1}} I(||v-u||)H(u,v,\mathbf{X})K_1\Big(\frac{||v-u||-r}{b_n}\Big),$$

où

$$
\begin{aligned}
A_1 &= \frac{2}{b_n^2|\Lambda_{n,R}|^2\sigma_d^2} \\
&\quad \mathrm{E}\int_{\mathbb{R}^{2d}} \frac{\mathbb{1}_{\Lambda_{n,R}}(u)}{||v-u||^{2(d-1)}} I(||v-u||)H(u,v,\mathbf{X})K_1^2\left(\frac{||v-u||-r}{b_n}\right)\lambda_{\beta^\star}(u,v,\mathbf{X})\mathrm{d}u\mathrm{d}v, \\
A_2 &= \frac{4}{b_n^2|\Lambda_{n,R}|^2\sigma_d^2}\,\mathrm{E}\int_{\mathbb{R}^{3d}} \frac{\mathbb{1}_{\Lambda_{n,R}}(u)\mathbb{1}_{\Lambda_{n,R}}(v)}{||v-u||^{d-1}||v-w||^{d-1}} I(||v-u||,||v-w||)H(u,v,w,\mathbf{X}) \\
&\quad K_1\left(\frac{||v-u||-r}{b_n}\right)K_1\left(\frac{||v-w||-r}{b_n}\right)\lambda_{\beta^\star}(u,v,w,\mathbf{X})\mathrm{d}u\mathrm{d}v\mathrm{d}w, \\
A_3 &= \frac{1}{b_n^2|\Lambda_{n,R}|^2\sigma_d^2}\,\mathrm{E}\int_{\mathbb{R}^{4d}} \frac{\mathbb{1}_{\Lambda_{n,R}}(u)\mathbb{1}_{\Lambda_{n,R}}(w)}{||v-u||^{d-1}||w-y||^{d-1}} I(||v-u||,||w-y||)H(u,v,w,y,\mathbf{X}) \\
&\quad \times K_1\left(\frac{||v-u||-r}{b_n}\right)K_1\left(\frac{||w-y||-r}{b_n}\right)\lambda_{\beta^\star}(u,v,w,y,\mathbf{X})\mathrm{d}u\mathrm{d}v\mathrm{d}w\mathrm{d}y,
\end{aligned}
$$

et

$$
\begin{aligned}
A_4 &= \frac{1}{b_n^2|\Lambda_{n,R}|^2\sigma_d^2} \\
&\quad \left(\mathrm{E}\int_{\mathbb{R}^d} \frac{\mathbb{1}_{\Lambda_{n,R}}(u)}{||v-u||^{d-1}} I(||v-u||)H(u,v,\mathbf{X})K_1\left(\frac{||v-u||-r}{b_n}\right)\lambda_{\beta^\star}(u,v,\mathbf{X})\mathrm{d}u\mathrm{d}v\right)^2.
\end{aligned}
$$

Nous pouvons maintenant facilement étudier le comportement asymptotique des termes A_1, A_2, A_3 et A_4.

Grâce à la stationnarité de $\mathbf{X}$ et en appliquant la propriété de portée d'interaction de paires finie (4.1.2), nous obtenons,

$$
\begin{aligned}
A_1 &= \frac{2}{b_n^2|\Lambda_{n,R}|^2\sigma_d^2} \\
&\quad \mathrm{E}\int_{\mathbb{R}^{2d}} \frac{\mathbb{1}_{\Lambda_{n,R}}(u)}{||v-u||^{2(d-1)}} I(||v-u||)H(u,v,\mathbf{X})K_1^2\left(\frac{||v-u||-r}{b_n}\right)\lambda_{\beta^\star}(u,v,\mathbf{X})\mathrm{d}u\mathrm{d}v \\
&= \frac{2\beta^{\star 2}}{b_n^2|\Lambda_{n,R}|^2\sigma_d^2} \\
&\quad \int_{\mathbb{R}^{2d}} \frac{\mathbb{1}_{\Lambda_{n,R}}(u)}{||v-u||^{2(d-1)}} I(||v-u||)\,\mathrm{E}[H(o,v-u,\mathbf{X})]K_1^2\left(\frac{||v-u||-r}{b_n}\right)\Phi(||v-u||)\mathrm{d}u\mathrm{d}v \\
&= \frac{2\beta^{\star 2}}{b_n|\Lambda_{n,R}|\sigma_d^2}\int_{-r/b_n}^{\infty}\int_{\mathbb{S}^{d-1}} \frac{I(b_n\varrho+r)}{(b_n\varrho+r)^{d-1}}\bar{F}(o,(b_n\varrho+r)w)K_1^2(\varrho)\Phi(b_n\varrho+r)\mathrm{d}\varrho\mathrm{d}\sigma(w).
\end{aligned}
$$

Nous appliquons le théorème de convergence dominée et sous les conditions suposées sur le noyau K_1, nous obtenons pour $r \in (0, R]$

$$\lim_{n\to\infty} b_n|\Lambda_{n,R}|A_1 = \frac{2\beta^{\star 2}}{\sigma_d r^{d-1}} J(r)\Phi(r) \int_{\mathbb{R}} K_1^2(\rho)\mathrm{d}\rho.$$

Nous allons maintenant montrer que tous les autres termes de la variance de $\widehat{R}_n(r)$ disparaissent lorsque n tend vers l'infini. Avec des arguments similaires à celles détaillés dans le chapitre précédent, $\lambda_{\beta^\star}(u, v, w, \emptyset)$ peut être notée $\Phi_1(||v-u||, ||w-v||)$, donc nous pouvons également obtenir le résultat suivant.

$$\begin{aligned}
A_2 &= \frac{4}{b_n^2|\Lambda_{n,R}|^2\sigma_d^2} \mathrm{E} \int_{\mathbb{R}^{3d}} \frac{\mathbb{1}_{\Lambda_{n,R}}(u)\mathbb{1}_{\Lambda_{n,R}}(v)}{||v-u||^{d-1}||v-w||^{d-1}} I(||v-u||, ||v-w||) H(u, v, w, \mathbf{X}) \\
&\times \Phi_1(||v-u||, ||w-v||) K_1\left(\frac{||v-u||-r}{b_n}\right) K_1\left(\frac{||v-w||-r}{b_n}\right) \mathrm{d}u\mathrm{d}v\mathrm{d}w \\
&= \frac{4}{|\Lambda_{n,R}|\sigma_d^2} \int_{-r/b_n}^{\infty} \int_{-r/b_n}^{\infty} \int_{\mathbb{S}^{d-1}} \int_{\mathbb{S}^{d-1}} \frac{|\Lambda_{n,R} \cap (\Lambda_{n,R} - (b_n\varrho + r)z)|}{|\Lambda_{n,R}|} \\
&\times \bar{F}(o, (b_n\varrho + r)z, (b_n\varrho' + r)z') \Phi_1(b_n\varrho + r, b_n\varrho' + r) K_1(\varrho) K_1(\varrho') \mathrm{d}\varrho \mathrm{d}\varrho' \mathrm{d}\sigma_d(z) \mathrm{d}\sigma_d(z').
\end{aligned}$$

En appliquant le théorème de convergence dominée, nous obtenons le résultat voulu $\lim_{n\to\infty} b_n|\Lambda_{n,R}|A_2 = 0$.
En procédant identiquement à celle détaillé dans le chapitre précédent, $\lambda_{\beta^\star}(u, v, w, y, \emptyset)$ peut être noté $\Phi_2(||v-u||, ||y-w||, ||w-u||, ||w-v||)$. Il résulte

$$\begin{aligned}
A_3 &= \frac{1}{b_n^2|\Lambda_{n,R}|^2\sigma_d^2} \mathrm{E} \int_{\mathbb{R}^4} \frac{\mathbb{1}_{\Lambda_{n,R}}(u)\mathbb{1}_{\Lambda_{n,R}}(w)}{||v-u||^{d-1}||w-y||^{d-1}} I(||v-u||, ||w-y||) H(u, v, w, y, \mathbf{X}) \\
&\times K_1\left(\frac{||v-u||-r}{b_n}\right) K_1\left(\frac{||w-y||-r}{b_n}\right) \lambda_{\beta^\star}(u, v, w, y, \mathbf{X}) \mathrm{d}u\mathrm{d}v\mathrm{d}w\mathrm{d}y \\
&= \frac{1}{|\Lambda_{n,R}|\sigma_d^2} \\
&\int_{\mathbb{R}^d} \int_{-r/b_n}^{\infty} \int_{-r/b_n}^{\infty} \int_{\mathbb{S}^{d-1}} \int_{\mathbb{S}^{d-1}} \frac{|\Lambda_{n,R} \cap (\Lambda_{n,R} - w)|}{|\Lambda_{n,R}|} \bar{F}(o, (b_n\varrho + r)z, (b_n\varrho' + r)z') \\
&\times \Phi_2^\star(b_n\varrho + r, b_n\varrho' + r, ||w||, ||(b_n\varrho + r)z - w||) K_1(\varrho) K_1(\varrho') \mathrm{d}w\mathrm{d}\varrho\mathrm{d}\varrho'\mathrm{d}\sigma_d(z)\mathrm{d}\sigma_d(z').
\end{aligned}$$

où $I(b_n\varrho+r, b_n\varrho'+r)\Phi_2(b_n\varrho+r, b_n\varrho'+r, ||w||, ||(b_n\varrho+r)z-w||)$ est noté par la fonction $\Phi_2^\star(b_n\varrho + r, b_n\varrho' + r, ||w||, ||(b_n\varrho + r)z - w||)$.
En appliquant le théorème de convergence dominée, nous obtenons $\lim_{n\to\infty} b_n|\Lambda_{n,R}|A_3 = 0$. Pour le comportement asymptotique de A_4, il suffit alors de reprendre les arguments

développés précédemment pour conclure le résultat suivant.

$$
\begin{aligned}
A_4 &= \frac{\beta^{\star 4}}{b_n^2|\Lambda_{n,R}|^2\sigma_d^2} \\
&\int_{\mathbb{R}^{4d}} \mathbb{1}_{\Lambda_{n,R}}(u)\mathbb{1}_{\Lambda_{n,R}}(w)\frac{I(||v-u||,||w-y||)}{||v-u||^{d-1}||w-y||^{d-1}}\,\mathrm{E}[H(u,v,\mathbf{X})]\,\mathrm{E}[H(w,y,\mathbf{X})] \\
&\times \Phi(||v-u||)\Phi(||y-w||)K_1\left(\frac{||v-u||-r}{b_n}\right)K_1\left(\frac{||w-y||-r}{b_n}\right)\mathrm{d}u\mathrm{d}v\mathrm{d}w\mathrm{d}y \\
&= \frac{\beta^4}{b_n^2|\Lambda_{n,R}|\sigma_d^2} \\
&\int_{\mathbb{R}^{3d}} \frac{|\Lambda_{n,R}\cap(\Lambda_{n,R}-w)|}{|\Lambda_{n,R}|||v-u||^{d-1}||w-y||^{d-1}}I(||v||,||w-y||)\,\mathrm{E}[H(o,v,\mathbf{X})]\,\mathrm{E}[H(w,y,\mathbf{X})] \\
&\times \Phi(||v||)\Phi(||y-w||)K_1\left(\frac{||v||-r}{b_n}\right)K_1\left(\frac{||w-y||-r}{b_n}\right)\mathrm{d}v\mathrm{d}w\mathrm{d}y \\
&= \frac{\beta^4}{|\Lambda_{n,R}|\sigma_d^2} \\
&\times \int_{\mathbb{R}^d}\int_{-r/b_n}^{\infty}\int_{-r/b_n}^{\infty}\int_{\mathbb{S}^{d-1}}\int_{\mathbb{S}^{d-1}} \frac{|\Lambda_{n,R}\cap(\Lambda_{n,R}-w)|}{|\Lambda_{n,R}|}I(b_n\varrho+r,b_n\varrho'+r)\bar{F}(o,(b_n\varrho+r)z) \\
&\times \bar{F}o,(b_n\varrho'+r)z')\Phi(b_n\varrho+r)\Phi(b_n\varrho'+r)K_1(\varrho)K_1(\varrho')\mathrm{d}w\mathrm{d}\varrho\mathrm{d}\varrho'\mathrm{d}\sigma_d(z)\mathrm{d}\sigma_d(z').
\end{aligned}
$$

En appliquant le théorème de convergence dominée, nous obtenons $\lim_{n\to\infty} b_n|\Lambda_{n,R}|A_4 = 0$. □

4.3 Consistance ponctuelle

Dans cette section, nous présentons la consistance ponctuelle de l'estimateur $\widehat{\Phi}_n(r)$. Par application directe du théorème ergodique classique pour les processus ponctuels spatiaux obtenus par Nguyen et Zessin [101], nous avons

$$\widehat{N}_n(r) \xrightarrow[n\to\infty]{p.s} \beta^\star J(r)$$

La consistance forte de l'estimateur $\widehat{N}_n(r)$ implique ce qui suit :

Proposition 4.3.1. *Sous les conditions de la Proposition 4.2.1 et le Théorème 4.2.2. Alors pour n assez grand, on a*

$$\widehat{\Phi}_n(r) \underset{n\to\infty}{\longrightarrow} \beta^\star\Phi(r) \quad \textit{en probabilité} \quad \mathbb{P} \quad \textit{si}$$

$$\widehat{R}_n(r) \underset{n\to\infty}{\longrightarrow} R(r) = \beta^{\star 2}J(r)\Phi(r) \quad \textit{en probabilité} \quad \mathbb{P}.$$

La preuve de ce résultat, se déduit du corollaire suivant.

Corollaire 4.3.2. Soit K_1 un noyau non négatif à support compact d'ordre $\alpha \geq 1$, tel que $\int_{\mathbb{R}} K_1^2(\rho)\mathrm{d}\rho < \infty$, on suppose de plus que $\exp(-\gamma(r))J(r)$ admet des dérivées partielles d'ordre α continues et bornées sur $(r-\delta, r+\delta)$, pour un certain $\delta > 0$. Alors pour $r \in (0, R]$

$$\mathrm{E}[\widehat{R}_n(r) - R(r)]^2 = \mathcal{O}(b_n^{2\alpha}) + o\big(\frac{1}{b_n n^d}\big), \quad \text{quand } n \to \infty.$$

Démonstration. Le schéma de la preuve du Corollaire 4.3.2 est le même que celui la démonstration du Corollaire 3.2.4. □

4.4 Consistance forte ponctuelle

Nous étudions dans ce paragraphe les propriétés ponctuelles de la consistance forte de l'estimateur $\widehat{\Phi}_n(r)$ défini en (5.3), cette convergence sera obtenue, à partir des propriétés asymptotiques de l'estimateur lissé $\widehat{R}_n(r)$ défini en (5.3) et de l'estimateur $\widehat{N}_n(r)$ défini en (5.3). La consistance forte des estimateurs $\widehat{N}_n(r)$ suit le schéma adapté à l'estimateur empirique étudié dans le chapitre précédent.

Supposons que $\Lambda_{n,R}$ se décompose comme $\Lambda_{n,R} = \cup_{i \in \widetilde{I}_n} \Lambda_i$, où Λ_i sont des sous-cubes de $\Lambda_{n,R}$ centrés en i et $\widetilde{I}_n$ une suite de sous-ensembles de $\mathbb{Z}^d$, telle que le cardinal $|\widetilde{I}_n|$ de $\widetilde{I}_n$ tend vers l'infini, quand n tend vers l'infini. Notons pour tout $k \in \mathbb{Z}^d$,

$$\bar{R}_k = R_k - \mathrm{E}\, R_k \quad \text{et} \quad S_n = \sum_{i \in \widetilde{I}_n} \bar{R}_i$$

avec

$$R_k = \sum_{\substack{u,v \in \mathbf{X} \\ ||v-u|| \leq R.}}^{\neq} \frac{\mathbb{1}_{\Lambda_k}(u)}{||v-u||^{d-1}} h(u, \mathbf{X} \setminus \{u,v\}) h(v, \mathbf{X} \setminus \{u,v\}) K_1\left(\frac{||v-u|| - r}{b_n}\right)$$

Proposition 4.4.1. *Soit K_1 un noyau non négatif à support compact d'ordre $\alpha \geq 1$, tel que $\int_{\mathbb{R}} K_1^2(\rho)\mathrm{d}\rho < \infty$, on suppose de plus que $\Phi(r)J(r)$ admet des dérivées partielles d'ordre α continues et bornées sur $(r-\delta, r+\delta)$, pour un certain $\delta > 0$. Alors pour $n \to \infty$,*

$$\widehat{\Phi}_n(r) \underset{n\to\infty}{\longrightarrow} \beta^\star \Phi(r) \quad p.s.$$

si l'une des propriétés suivantes est satisfaite :

1) S'il existe un réel $0 < q < 2$ tel que la variable aléatoire $\bar{R}_0$ appartient à $\mathbb{L}_{\psi_{\theta(q)}}$ et sous la condition

$$\sum_{k \in V_0^1} \left\| \sqrt{\left| \bar{R}_k E_{|k-i|}(\bar{R}_i) \right|} \right\|^2_{\psi_{\theta(q)}} < \infty. \tag{4.4.1}$$

Alors

$$\widehat{R}_n(r) - R(r) = \mathcal{O}_{p.s.}\left(\frac{(\log n)^{1/q}}{b_n n^{d/2}} \right) + \mathcal{O}(b_n^\alpha) \quad \textit{quand} \quad n \to \infty.$$

2) Si $\bar{R}_0$ appartient à $\mathbb{L}^\infty$ et sous la condition

$$\sum_{k \in V_i^1} \| \bar{R}_k E_{|k-i|}(\bar{R}_i) \|_\infty < \infty. \tag{4.4.2}$$

Alors

$$\widehat{R}_n(r) - R(r) = \mathcal{O}_{p.s.}\left(\frac{(\log n)^{1/2}}{b_n n^{d/2}} \right) + \mathcal{O}(b_n^\alpha) \quad \textit{quand} \quad n \to \infty.$$

3) S'il existe un réel $p > 2$ tel que la variable aléatoire $\bar{R}_0$ appartient à $\mathbb{L}^p$ et sous la condition

$$\sum_{k \in V_i^1} \| \bar{R}_k E_{|k-i|}(\bar{R}_i) \|_{\frac{p}{2}}. \tag{4.4.3}$$

Alors

$$\widehat{R}_n(r) - R(r) = \mathcal{O}_{p.s.}\left(\frac{(\log n)^{1/p}}{b_n n^{d/2}} \right) + \mathcal{O}(b_n^\alpha) \quad \textit{quand} \quad n \to \infty.$$

4.5 Consistance forte globale

Nous établissons maintenant la consistance forte globale de l'estimateur de la fonction d'interaction de paires dans l'intervalle $[r_1, r_2]$. Nous donnons tout d'abord les hypothèses portant sur le modèle permettant d'obtenir le résultat voulu.

Condition $\mathcal{P}(D)$:

J et $\Phi(\cdot) = \exp(-\gamma(\cdot))$ sont des fonctions continues sur le compact $[r_1, r_2]$.

S'ajoutent également l'hypothèse suivante sur le noyau.

Condition $\mathcal{Z}$:

Le noyau K_1 est une fonction lipschitzienne, c'est à dire qu'il existe une constante $\eta > 0$ telle que

$$\left| K_1(\rho) - K_1(\rho') \right| \leq \eta |\rho - \rho'| \quad \text{pour tous} \quad \rho, \rho' \in [r_1, r_2].$$

Théorème 4.5.1. *Sous les conditions $\mathcal{P}(D)$ et $\mathcal{Z}$. De plus si $\exp(-\gamma(r))J(r)$ admet des dérivées partielles d'ordre α continues et bornées sur $[r_1-\delta, r_2+\delta]$ pour un certain $\delta > 0$. Soit K_1 un noyau non négatif d'ordre $\alpha \geq 1$, tel que $\int_{\mathbb{R}} K_1^2(\rho)\mathrm{d}\rho < \infty$. Alors quand $n \to \infty$*

$$\sup_{r_1 \leq r \leq r_2} \left|\widehat{\Phi}_n(r) - \beta^\star \Phi(r)\right| \longrightarrow 0 \quad p.s.$$

si l'une des propriétés suivantes est satisfaite :
1) S'il existe un réel $0 < q < 2$ tel que la variable aléatoire $\bar{R}_0$ appartient à $\mathbb{L}_{\psi_{\theta(q)}}$ et sous la condition

$$\sum_{k \in V_0^1} \left\| \sqrt{\left|\bar{R}_k E_{|k|}(\bar{Z}_0)\right|} \right\|^2_{\psi_{\theta(q)}}. \tag{4.5.1}$$

Alors

$$\sup_{r_1 \leq r \leq r_2} \left|\widehat{R}_n(r) - R(r)\right| = \mathcal{O}_{p.s.}\left(\frac{(\log n)^{1/q}}{b_n n^{d/2}}\right) + \mathcal{O}(b_n^\alpha) \quad quand \quad n \to \infty.$$

2) Si $\bar{R}_0$ appartient à $\mathbb{L}^\infty$ et sous la condition

$$\sum_{k \in V_0^1} \left\|\bar{R}_k E_{|k|}(\bar{R}_0)\right\|_\infty. \tag{4.5.2}$$

Alors

$$\sup_{r_1 \leq r \leq r_2} \left|\widehat{R}_n(r) - R(r)\right| = \mathcal{O}_{p.s.}\left(\frac{(\log n)^{1/2}}{b_n n^{d/2}}\right) + \mathcal{O}(b_n^\alpha) \quad quand \quad n \to \infty.$$

3) S'il existe un réel $p > 2$ tel que la variable aléatoire $\bar{R}_0$ appartient à $\mathbb{L}^p$ et sous la condition

$$\sum_{k \in V_i^1} \|\bar{R}_k E_{|k-i|}(\bar{R}_i)\|_{\frac{p}{2}}. \tag{4.5.3}$$

Soit $b_n = n^{-q_2}(\log n)^{q_1}$ pour certaines constantes $q_1, q_2 > 0$. Soient $a, b \geq 0$ fixés, si

$$a(p+1) - d/2 - q_2 > 1 \quad et \quad b(p+1) + q_1 > 1.$$

Alors

$$\sup_{r_1 \leq r \leq r_2} \left|\widehat{R}_n(r) - R(r)\right| = \mathcal{O}_{p.s.}\left(\frac{n^a(\log n)^b}{b_n n^{d/2}}\right) + \mathcal{O}(b_n^\alpha) \quad quand \quad n \to \infty.$$

4.5.1 Preuve de la Proposition 4.4.1

Démonstration. 1) Soient ε un réel positif et $(r_n)_{n>0}$ une suite de réels strictement tendant vers zéro. En utilisant l'inégalité de Markov, nous obtenons

$$\begin{aligned}\mathbb{P}(|\widehat{R}_n(r) - \mathrm{E}\,\widehat{R}_n(r)| > \varepsilon r_n) &= \mathbb{P}\left(|S_n| > \varepsilon r_n b_n n^d\right)\\ &\leq \exp\left[-\left(\frac{\varepsilon\, r_n b_n n^d}{||S_n||_{\psi_{\theta(q)}}} + \xi_q\right)^q\right] \mathrm{E}\exp\left[\left(\frac{|S_n|}{||S_n||_{\psi_{\theta(q)}}} + \xi_q\right)^q\right].\end{aligned}$$

Par conséquent pour $0 < q < 2$ fixé tel que $R_0 \in \mathbb{L}_{\psi_{\theta(q)}}$, nous utilisons l'inégalité de Kahane-Khintchine (3.4.2) du Théorème 3.4.1, nous obtenons

$$\begin{aligned}\mathbb{P}(|\widehat{R}_n(r) - \mathrm{E}\,\widehat{R}_n(r)| > \varepsilon r_n) &= \mathbb{P}\left(|S_n| > \varepsilon r_n b_n n^d\right)\\ &\leq (1 + e^{\xi_q^q})\, \exp\left[-\left(\frac{\varepsilon\, r_n b_n n^d}{M_1(q)\|\bar{R}_0\|^2_{\psi_{\theta(q)}}(\sum_{i\in\widetilde{I}_n} b_{i,q}(\bar{R}))^{1/2}} + \xi_q\right)^q\right].\end{aligned}$$

où

$$b_{i,q}(\bar{R}) = \left\|\bar{R}_0\right\|^2_{\psi_{\theta(q)}} + \sum_{k\in V_0^1}\left\|\sqrt{\left|\bar{R}_k E_{|k|}(\bar{Z}_0)\right|}\right\|^2_{\psi_{\theta(q)}}.$$

Par définition de la norme $||\cdot||_{\psi_{\theta(q)}}$ et sous la condition (4.5.1), il existe donc une constante $C > 0$ telle que

$$\mathbb{P}(|\widehat{R}_n(r) - \mathrm{E}\,\widehat{R}_n(r)| > \varepsilon r_n) \leq (1 + e^{\xi_q^q})\, \exp\left[-\frac{\varepsilon^q r_n^q b_n^q (\sqrt{n})^{dq}}{C^q}\right].$$

Si $r_n = (\log n)^{1/q}/b_n(\sqrt{n})^d$ et pour n suffisamment grand, alors

$$\mathbb{P}(|\widehat{R}_n(r) - \mathrm{E}\,\widehat{R}_n(r)| > \varepsilon r_n) \leq (1 + e^{\xi_q^q})\, \exp\left[-\frac{\varepsilon^q\, \log n}{C^q}\right].$$

2) En reformulant l'inégalité de Markov et nous utilisons l'inégalité (3.4.3) du Théorème 3.4.1, nous obtenons, pour $Z_0 \in \mathbb{L}^\infty$ et pour ε un réel positif,

$$\begin{aligned}\mathbb{P}(|\widehat{R}_n(r) - \mathrm{E}\,\widehat{R}_n(r)| > \varepsilon r_n) &= \mathbb{P}\left(|S_n| > \varepsilon r_n b_n n^d\right)\\ &\leq (1 + e^{\xi_q^q})\, \exp\left[-\left(\frac{\varepsilon\, r_n b_n n^d}{M_2(q)\|\bar{R}_0\|_\infty(\sum_{i\in\widetilde{I}_n} b_{i,\infty}(\bar{R}))^{1/2}} + \xi_q\right)^q\right],\end{aligned}$$

où

$$b_{i,\infty}(R) = \left\|\bar{R}_0\right\|^2_\infty + \sum_{k\in V_0^1}\left\|\bar{R}_k E_{|k|}(\bar{R}_0)\right\|_\infty.$$

Si $r_n = (\log n)^{1/2}/b_n(\sqrt{n})^d$ et n suffisamment et sous la condition (4.5.2), il existe une constante $C > 0$ telle que

$$\mathbb{P}(|\widehat{R}_n(r) - \mathrm{E}\,\widehat{R}_n(r)| > \varepsilon r_n) \leq 2\,\exp\left[-\frac{\varepsilon^2\,\log n}{C^2}\right].$$

D'aprés le Lemme de Borel-Cantelli et la Proposition 4.2.1, d'où les deux premières assertions de la Proposition 4.4.1.

3) Pour $p > 0$ tel que $\bar{R}_k \in \mathbb{L}^p$, nous avons pour tout $\varepsilon > 0$,

$$\begin{aligned}\mathbb{P}(|\widehat{R}_n(r) - \mathrm{E}\,\widehat{R}_n(r)| > \varepsilon r_n) &= \mathbb{P}\left(|S_n| > \varepsilon r_n b_n n^d\right)\\ &\leq \frac{\varepsilon^{-p}\,\mathrm{E}\,|S_n|^p}{r_n^p b_n^p (\sqrt{n})^{pd}}\\ &\leq \frac{\varepsilon^{-p}}{r_n^p b_n^p (\sqrt{n})^{pd}}\left(2p\sum_{i\in\widetilde{I}_n} c_i(\bar{R})\right)^{p/2}.\end{aligned}$$

La dernière inégalité résulte d'après la Proposition 3.4.2, où

$$c_i(R) = \|\bar{R}_i\|_p^2 + \sum_{k\in V_i^1} \|\bar{R}_k E_{|k-i|}(\bar{R}_i)\|_{\frac{p}{2}}.$$

Nous déduisons par conséquent sous la condition (4.5.3), il existe $C > 0$ telle que

$$\mathbb{P}(|\widehat{R}_n(r) - \mathrm{E}\,\widehat{R}_n(r)| > \varepsilon r_n) \leq \frac{C\varepsilon^{-p}}{r_n^p b_n^p (\sqrt{n})^{pd}}.$$

Dès lors que $r_n = (\log n)^{1/p}/b_n(\sqrt{n})^d$, d'aprés le Lemme de Borel-Cantelli et la Proposition 4.2.1, nous concluons la preuve de la Proposition 4.4.1.

□

4.5.2 Preuve du Théorème 4.5.1

Démonstration. Nous procédons à un recouvrement de l'intervalle $[r_1, r_2]$ par des intervalles $C_i = [s_{i-1} - s_i]$, où $s_i = r_1 + i(r_2 - r_1)/N, i = 1, ..., N$. Choisissant N comme le plus grand entier satisfaisant $N \leq c/l_n$ et $l_n = r_n b_n^2$.
Sous la condition $\mathcal{Z}$, nous déduisons qu'il existe une constante $\eta > 0$ telle que pour tout n suffisamment grand

$$\begin{aligned}\sup_{s_{i-1}\leq r,\rho\leq s_i}\left|\widehat{R}_n(r) - \widehat{R}_n(\rho)\right| &\leq \frac{1}{b_n^2}\eta|r-\rho|\widetilde{R}_n\\ &\leq r_n\widetilde{R}_n\end{aligned}$$

où

$$\widetilde{R}_n = \frac{1}{\sigma_d |\Lambda_{n,R}|} \sum_{\substack{u,v\in\mathbf{X} \\ ||v-u||\leq R}}^{\neq} \frac{\mathbb{1}_{\Lambda_{n,R}}(u)}{||v-u||^{d-1}} h(u, \mathbf{X}\setminus\{u,v\}) h(v, \mathbf{X}\setminus\{u,v\}).$$

Nous appliquons aisément le théorème ergodique obtenu par Nguyen et Zessin [101] dans la dernière inégalité. De même

$$\sup_{s_{i-1}\leq r,\rho\leq s_i} \left| \mathrm{E}\,\widehat{R}_n(r) - \mathrm{E}\,\widehat{R}_n(\rho) \right| \leq r_n \,\mathrm{E}\,\widetilde{R}_n.$$

Lemme 4.5.2. *Pour tout $0 < q < 2$ tel que la variable aléatoire $\bar{R}_0$ appartient à $\mathbb{L}_{\psi_{\theta(q)}}$ on note $r_n = (\log n)^{1/q}/b_n(\sqrt{n})^d$ ou pour tout $0 < q \leq 2$ tel que la variable aléatoire $\bar{R}_0$ appartient à $\mathbb{L}_\infty$, on note $r_n = (\log n)^{1/2}/b_n(\sqrt{n})^d$. Alors pour n suffisamment grand,*

$$\sup_{s_{i-1}\leq\rho\leq s_i} \left| \widehat{R}_n(\rho) - \mathrm{E}\,\widehat{R}_n(\rho) \right| = \mathcal{O}_{p.s.}(r_n).$$

Démonstration. Pour $\varepsilon > 0$ fixé, nous avons

$$\mathbb{P}\left(\sup_{s_{i-1}\leq\rho\leq s_i} \left| \widehat{R}_n(\rho) - \mathrm{E}\,\widehat{R}_n(\rho) \right| > \varepsilon r_n \right) \leq N \sup_{r_1\leq r\leq r_2} \mathbb{P}(|\widehat{R}_n(r) - \mathrm{E}\,\widehat{R}_n(r)| > \varepsilon r_n).$$

Ensuite, nous reprenons les mêmes étapes de la preuve de la Proposition 4.4.1, pour $0 < q < 2$ et si $r_n = (\log n)^{1/q}/b_n(\sqrt{n})^d$, nous obtenons

$$\sup_{r_1\leq r\leq r_2} \mathbb{P}(|\widehat{R}_n(r) - \mathrm{E}\,\widehat{R}_n(r)| > \varepsilon r_n) \leq (1 + e^{\xi_q^q}) \exp\left[-\frac{\varepsilon^q \log n}{C^q} \right].$$

Pour $0 < q \leq 2$, $r_n = (\log n)^{1/2}/b_n(\sqrt{n})^d$, nous obtenons ainsi,

$$\sup_{r_1\leq r\leq r_2} \mathbb{P}(|\widehat{R}_n(r) - \mathrm{E}\,\widehat{R}_n(r)| > \varepsilon r_n) \leq 2 \exp\left[-\frac{\varepsilon^2 \log n}{C^2} \right].$$

Donc, il en résulte grâce au Lemme de Borel-Cantelli que

$$\mathbb{P}(\limsup_{n\to\infty} \sup_{s_{i-1}\leq\rho\leq s_i} \left| \widehat{R}_n(\rho) - \mathrm{E}\,\widehat{R}_n(\rho) \right| > \varepsilon r_n) = 0.$$

□

La pertinence des deux premiers résultats du Théorème 4.5.1 est étudiée par la continuité uniforme de $R(r)$. D'une part la partie biais converge vers 0 d'après la Proposition 4.2.1, nous obtenons par un calcul classique

$$\sup_{r_1\leq r\leq r_2} |\,\mathrm{E}\,\widehat{R}_n(r) - R(r)| = \mathcal{O}(b_n^\alpha) \quad \text{quand} \quad n \to \infty. \tag{4.5.4}$$

Lemme 4.5.3. *S'il existe un réel $p > 2$ tel que la variable aléatoire $\bar{R}_0$ appartient à $\mathbb{L}^p$ et $b_n = n^{-q_2}(\log n)^{q_1}$ pour certaines constantes $q_1, q_2 > 0$ et si*

$$a(p+1) - d/2 - q_2 > 1 \quad \textit{et} \quad b(p+1) + q_1 > 1,$$

pour $a, b \geq 0$ fixés. Alors

$$\sup_{s_{i-1} \leq \rho \leq s_i} \left| \widehat{R}_n(\rho) - \mathrm{E}\, \widehat{R}_n(\rho) \right| = \mathcal{O}_{p.s}(r_n) \quad \textit{quand} \quad n \to \infty.$$

Démonstration. Soit $p > 2$ fixé. Pour tout $\varepsilon > 0$, nous reprenons les mêmes étapes de la preuve de la Proposition 4.4.1, nous concluons

$$\begin{aligned}
\mathbb{P}\left(\sup_{s_{i-1} \leq \rho \leq s_i} \left| \widehat{R}_n(\rho) - \mathrm{E}\, \widehat{R}_n(\rho) \right| > \varepsilon r_n \right) &\leq N \sup_{r_1 \leq r \leq r_2} \mathbb{P}(|\widehat{R}_n(r) - \mathrm{E}\, \widehat{R}_n(r)| > \varepsilon r_n) \\
&\leq N \frac{\kappa \varepsilon^{-p}}{r_n^p b_n^p (\sqrt{n})^{pd}}.
\end{aligned}$$

Comme $N \leq c/l_n$ et $l_n = r_n b_n^2$, donc pour $r_n = n^a (\log n)^b / b_n (\sqrt{n})^d$, il résulte pour n suffisamment grand,

$$\begin{aligned}
\mathbb{P}\left(\sup_{s_{i-1} \leq \rho \leq s_i} \left| \widehat{R}_n(\rho) - \mathrm{E}\, \widehat{R}_n(\rho) \right| > \varepsilon r_n \right) &\leq \frac{\kappa \varepsilon^{-p}}{n^{a(p+1)-d/2} (\log n)^{b(p+1)} b_n} \\
&\leq \frac{\kappa \varepsilon^{-p}}{n^{a(p+1)-d/2-q_2} (\log n)^{b(p+1)+q_1}}.
\end{aligned}$$

Pour $a(p+1) - d/2 - q_2 > 1$ et $b(p+1) + q_1 > 1$, nous obtenons pour tout réel $\varepsilon > 0$

$$\sum_{n \geq 1} \mathbb{P}\left(\sup_{s_{i-1} \leq \rho \leq s_i} \left| \widehat{R}_n(\rho) - \mathrm{E}\, \widehat{R}_n(\rho) \right| > \varepsilon r_n \right) < \infty.$$

□

Compte tenu de ces arguments les preuves du Théorème 4.5.1 sont terminées, cela résulte d'une application directe du théoréme de Borel-Cantelli et du résultat (4.5.4). □

La démonstration de la consistance forte ponctuelle et globale consiste essentiellement à appliquer d'autres inégalités pour les processus ponctuels, nous renvoyons à Bosq [18], Collomb [32] et Peligrad [107].

4.6 Normalité asymptotique

Nos résultats sont alors les suivants, sous la condition suivante.
Condition $\mathcal{W}$: Pour n assez grand,

$$\frac{m_n^d \log^d n}{b_n n^d} \to 0.$$

Le théorème suivant établit la normalité asymptotique de l'estimateur $\widehat{\Phi}_n(r)$ de la fonction d'interaction de paires Φ.

Théorème 4.6.1. *Soit Φ la fonction d'interaction de paires définie en (4.1.1). Soit le q-uplet $(r_1, ..., r_q) \in (0, R]$, avec $r_i \neq r_j$ pour $i \neq j$ et r_i points de continuité de Φ pour $i = 1, ..., q$. K_1 un noyau non négatif à support compact d'ordre $\alpha \geq 1$, tel que $\int_{\mathbb{R}} K_1^2(\rho)\mathrm{d}\rho < \infty$. Sous les conditions $b_n^{1+2\alpha}|\Lambda_{n,R}| \to 0$ quand $n \to \infty$ et $\mathcal{W}$. Alors pour n assez grand,*

$$\sqrt{b_n|\Lambda_{n,R}|}\left(\widehat{\Phi}_n(r_i) - \beta^\star \Phi(r_i)\right)_{i=1}^q \xrightarrow{d} \mathcal{N}(0, \sigma_\Phi).$$

avec la matrice de covariance $\sigma_\Phi = [\sigma_{i,j}]_{i,j=1,...,q}$ où $\sigma_{i,j}$ a pour élément

$$\sigma_{i,i} = \frac{2\Phi(r_i)}{\sigma_d r_i^{d-1} J(r_i)} \int_{\mathbb{R}} K_1^2(\rho)\mathrm{d}\rho \quad \textit{pour } i = 1, ..., q \quad \textit{et} \quad \sigma_{i,j} = 0 \quad \textit{si} \quad i \neq j.$$

La preuve du Théorème 4.6.1 est basée sur les deux résultats suivants.

Théorème 4.6.2. *Soit Φ la fonction d'interaction de paires définie en (4.1.1) et soit le q-uplet $(r_1, ..., r_q) \in (0, R]$, avec $r_i \neq r_j$ pour $i \neq j$. Soit K_1 un noyau non négatif à support compact d'ordre $\alpha \geq 1$, tel que $\int_{\mathbb{R}} K_1^2(\rho)\mathrm{d}\rho < \infty$ et sous la condition $\mathcal{W}$. Alors le vecteur ligne*

$$\sqrt{b_n|\Lambda_{n,R}|}\left(\widehat{R}_n(r_i) - \mathrm{E}\,\widehat{R}_n(r_i)\right)_{i=1}^q \xrightarrow{d} \mathcal{N}(0, \sigma_\Phi).$$

avec la matrice de covariance $\sigma_\Phi = [r_{i,j}]_{i,j=1,...,q}$ où $r_{i,j}$ a pour élément

$$r_{i,i} = \frac{2\beta^{\star 2}}{\sigma_d r_i^{d-1}} J(r_i)\Phi(r_i) \int_{\mathbb{R}} K_1^2(\rho)\mathrm{d}\rho, \quad \textit{pour } i = 1, ..., q \quad \textit{et} \quad r_{i,j} = 0 \quad \textit{si} \quad i \neq j.$$

Corollaire 4.6.3. Sous les conditions du Théorème 4.6.2 et soit $b_n^{1+2\alpha}|\Lambda_{n,R}| \to 0$ quand $n \to +\infty$. De plus si $\exp(-\gamma(r))J(r)$ admet des dérivées partielles d'ordre α continues et bornées au voisinage de r_i pour $i = 1, ..., s$. Alors

$$\sqrt{b_n|\Lambda_{n,R}|}\left(\widehat{R}_n(r_i) - R(r_i)\right)_{i=1}^s \xrightarrow{d} \mathcal{N}(0, \sigma_\Phi).$$

avec la matrice de covariance $\sigma_{\Phi} = [r_{i,j}]_{i,j=1,\dots,s}$ où $r_{i,j}$ a pour élément

$$r_{i,i} = \frac{2\beta^{\star 2}}{\sigma_d r_i^{d-1}} J(r_i)\Phi(r_i) \int_{\mathbb{R}} K_1^2(\rho)\mathrm{d}\rho, \quad \text{pour } i = 1, ..., s \quad \text{et} \quad r_{i,j} = 0 \quad \text{si} \quad i \neq j.$$

Le schéma de la preuve de la normalité asymptotique de l'estimateur $\widehat{\Phi}_n(r)$ est le même que celui la démonstration dans le chapitre précédent.

Chapitre 5

Estimation non paramétrique de fonction d'interaction de paires d'un processus ponctuel à interaction par paires pour les champs mélangeants

Résumé

L'objectif dans ce chapitre est de reprendre les propriétés asymptotiques étudiées dans le Chapitre 3 et Chapitre 4 dans le cadre des champs mélangeants. Nous commençons par indiquer les hypothèses de dépendance usuelles dans la section 5.1. Nos résultats sur la consistance forte ponctuelle et globale pour le processus ponctuel à interaction par paires stationnaire mélangeant sont établis. Nous étudions également les vitesses de convergence de l'estimateur à base de noyau dans la Section 5.2, la normalité asymptotique est également examinée. Ensuite, nous établissons ces résultats pour le processus ponctuel à interaction par paires stationnaire et isotrope mélangeant dans la Section 5.3. Les preuves des résultats de ce chapitre sont présentées dans la Section 5.4. Nos preuves utilisent entre autres des inégalités exponentielles de type Kahane-Khintchine dans des espaces d'Orlicz induits par des fonctions de Young exponentielles établies par El Machkouri [88], fournies de nouveaux des critères pour les champs aléatoires mélangeants. La procédure fait intervenir des inégalités de mélange de Serfling (voir Serfling [118] ou McLeish [93]). Pour prouver nos résultats sur la normalité asymptotique, nous utiliserons le théorème central limite de Coeurjolly et Lavancier [29], ainsi qu'un contrôle des coefficients de mélange.

5.1 Coefficients de mélange

Il existe dans la littérature un nombre conséquent de travaux consacrés à l'étude des variables mélangeantes réelles dont les références de Bradley [23] Rosenblatt [115] et Ibragimov [75]. L'hypothèse de mélange a été introduite pour étendre le théorème central limite pour les chaînes de Markov satisfaisant des conditions Doeblin pour des classes plus générales de variables aléatoires dépendantes (voir Linnik et Ibragimov [76]). Il permet aussi de généraliser la convergence presque sûre et la normalité asymptotique obtenus avec des échantillons indépendants.

Pour commencer, nous rappelons quelques définitions utiles dans ce chapitre. Des conditions de dépendance sur le processus observé peuvent être considérées pour obtenir des estimateurs convergents. Les hypothèses de mélangeance sont les plus utilisées. Soit $(\Omega, \mathcal{F}, \mathbb{P})$ l'espace de probabilité. Donnons deux sous σ-algèbres $\mathcal{U}$ et $\mathcal{V}$ de $\mathcal{F}$. Nous sommes intéressés par le coefficient de mélange fort $\alpha(\mathcal{U}, \mathcal{V})$, défini par Ibragimov et Rozanov [74] (cf. aussi Doukhan [49]) par :

$$\alpha(\mathcal{U}, \mathcal{V}) = \sup\{|\mathbb{P}(U \cap V) - \mathbb{P}(U)\mathbb{P}(V)|,\, U \in \mathcal{U},\, V \in \mathcal{V}\}.$$

Notons $|\Gamma|$ le cardinal du tout sous-ensemble Γ de $\mathbb{Z}^d$. Nous utilisons également les coefficients α paramétrés par les tailles k et $l \in \mathbb{N}^\star \cup \{\infty\}$ de $\mathcal{U}$ et de $\mathcal{V}$; pour $n \in \mathbb{N}$:

$$\alpha_{k,l}(n) = \sup\{\alpha(\mathcal{F}_{\Gamma_1}, \mathcal{F}_{\Gamma_2}),\, |\Gamma_1| \leq k,\, |\Gamma_2| \leq l,\, \rho(\Gamma_1, \Gamma_2) \geq n\},$$

où la distance ρ est défini par $\rho(\Gamma_1, \Gamma_2) = \min\{|i - j|,\, i \in \Gamma_1,\, j \in \Gamma_2\}$. On dit que le champ aléatoire $(Z_k)_{k \in \mathbb{Z}^d}$ est α-mélangeant s'il existe un couple (k, l) dans $(\mathbb{N}^* \cup \{\infty\})^2$ tel que $\lim_{n\to\infty} \alpha_{k,l}(n) = 0$. Nous nous intéressons aussi à la propriété ϕ-mélangeant pour les processus ponctuels spatiaux, introduit en 1962 par Ibragimov [75],

$$\phi(\mathcal{U}, \mathcal{V}) = \sup\{\|\mathbb{P}(V|\mathcal{U}) - \mathbb{P}(V)\|_\infty\,,\, V \in \mathcal{V}\}.$$

On dit que le champ aléatoire $(X_k)_{k \in \mathbb{Z}^d}$ est ϕ-mélangeant s'il existe un couple (k, l) dans $(\mathbb{N}^* \cup \{\infty\})^2$ tel que $\lim_{n\to\infty} \phi_{k,l}(n) = 0$, où

$$\phi_{k,l}(n) = \sup\{\phi(\mathcal{F}_{\Gamma_1}, \mathcal{F}_{\Gamma_2}),\, |\Gamma_1| \leq k,\, |\Gamma_2| \leq l,\, \rho(\Gamma_1, \Gamma_2) \geq n\}.$$

$\mathbb{P}$ est également fortement mélangeant puisqu'on a toujours $2\alpha(\cdot) \leq \phi(\cdot)$ (voir Doukhan [49]).

Avec ces définitions, donnons un corollaire du Théorème 3.4.1 de El Machkouri [88] cité dans le Chapitre 3, ce résultat va nous permettre d'obtenir des résultats asymptotiques des estimateurs $\widehat{G}_n(t)$ défini en (5.2) et $\widehat{\Phi}_n(r)$ défini en (5.3) pour des champs aléatoires ϕ-mélangeants.

Corollaire 5.1.1. Soit $(X_i)_{i\in\mathbb{Z}^d}$ un champ de variables aléatoires réelles de moyennes nulles stationnaire satisfait l'hypothèse (3.4.1) pour un certain $0 < q < 2$. Alors il existe une constante universelle $M_1(q) > 0$ qui ne dépend que de q telle que pour toute famille $(a_i)_{i\in\mathbb{Z}^d}$ de nombres réels et toute partie finie Γ de $\mathbb{Z}^d$,

$$\left\| \sum_{i\in\Gamma} a_i X_i \right\|_{\psi_q} \leq M_1(q) \left\| X_0 \right\|_{\psi_{\theta(q)}} \left(\sum_{i\in\Gamma} |a_i| \widetilde{b}_{i,q}(X) \right)^{1/2} \tag{5.1.1}$$

où

$$\widetilde{b}_{i,q}(X) = |a_i| + C(q) \sum_{k\in V_0^1} |a_{k+i}| \sqrt{\phi_{\infty,1}(|k|)}.$$

$C(q)$ est une constante universelle strictement positive qui ne dépend que de q.
Si $(X_i)_{i\in\mathbb{Z}^d}$ est borné alors pour tout réel $0 < q \leq 2$, il existe une constante universelle $M_2(q) > 0$ qui ne dépend que de q telle que pour toute famille $(a_i)_{i\in\mathbb{Z}^d}$ de nombres réels et toute partie finie Γ de $\mathbb{Z}^d$,

$$\left\| \sum_{i\in\Gamma} a_i X_i \right\|_{\psi_q} \leq M_2(q) \left\| X_0 \right\|_{\infty} \left(\sum_{i\in\Gamma} |a_i| \widetilde{b}_{i,\infty}(X) \right)^{1/2} \tag{5.1.2}$$

où

$$\widetilde{b}_{i,\infty}(X) = |a_i| + 2 \sum_{k\in V_0^1} |a_{k+i}| \phi_{\infty,1}(|k|).$$

5.2 Résultats asymptotiques de l'estimateur $\widehat{G}_n(t)$

En introduisant des notations et des conditions similaires à celles présentées dans le Chapitre 3 et à partir des propriétés asymptotiques de l'estimateur $\widehat{L}_n(t)$ établis dans le Chapitre 3, nous établissons par la suite des résultats intéressants permettant d'établir la consistance forte ponctuelle et globale de l'estimateur $\widehat{G}_n(t)$, en donnant la vitesse de convergence de l'etimateur $\widehat{H}_n(t)$, sous la condition de l'unicité de Dobrushin (Dobrushin [47] ; Georgii [58] ; Guyon [62]).

L'estimateur non paramétrique de la fonction $\beta^\star G(t)$ tel qu'introduit dans le chapitre 3 défini par

$$\widehat{G}_n(t) = \frac{\widehat{H}_n(t)}{\widehat{L}_n(t)},$$

où

$$\widehat{H}_n(t) = \frac{1}{b_n^d|\Lambda_{n,R}|} \sum_{\substack{u,v\in\mathbf{X} \\ v-u\in B(o,R)}}^{\neq} 1\!\!1_{\Lambda_{n,R}}(u)h(u,\mathbf{X}\setminus\{u,v\})h(v,\mathbf{X}\setminus\{u,v\})K\left(\frac{v-u-t}{b_n}\right)$$

est un estimateur de type noyau estimant $\beta^{\star 2}\exp(-g(t))\bar{F}(o,t)$ avec $\bar{F}(o,t) = \mathbb{P}(\mathbf{X}\cap B(o,\widetilde{R}) = \emptyset, \mathbf{X}\cap B(t,\widetilde{R}) = \emptyset)$.

$$\widehat{L}_n(t) = \frac{1}{|\Lambda_{n,R}|} \sum_{u\in\mathbf{X}_{\Lambda_{n,R}}} h(u,\mathbf{X}\setminus\{u\})h(t-u,\mathbf{X}\setminus\{u\})$$

étant quand à lui un estimateur de $\beta^\star\bar{F}(o,t)$.

Nous reprenons les notations et les hypothèses introduites dans le Chapitre 3 qui nous seront utiles dans notre étude.

Condition $\mathbf{K^0}$

(i) $\int_{\mathbb{R}^d} K(z)\mathrm{d}z = 1$,

(ii) $K(z) = 0$, pour tout $z \in \mathbb{R}^d$ avec $||z|| > R_1$ et pour un certain $R_1 \in \mathbb{R}_+$,

(iii) la fonction K est non négative et bornée,

(iv) pour $m \in \mathbb{N}^\star$

$$\int_{\mathbb{R}^d} z_1^{a_1}...z_d^{a_d}K(z)\mathrm{d}z = 0$$

avec $a_1, ..., a_d \in \mathbb{N}^\star$ tels que $0 < \sum_{i=1}^d a_i < m$.

Condition $\boldsymbol{\mathcal{P}(T)}$:

$\bar{F}$ et $G = \exp(-g)$ sont des fonctions continues sur T.

5.2.1 Consistance forte ponctuelle

Proposition 5.2.1. *Soit* $\mathbf{X}$ *un processus ponctuel à interaction par paires stationnaire. Sous les conditions* $\mathbf{K}^0$ *et* $\mathcal{P}(t)$*. De plus si* $G(t)\bar{F}(o,t)$ *admet des dérivées partielles d'ordre* m *continues et bornées sur* $B^o(t,\delta)$*, pour* $t \in \mathring{T}$*, pour un certain* $\delta > 0$ *et pour tout* $m \geq 1$*. Alors,*

$$\widehat{G}_n(t) \xrightarrow[n\to\infty]{p.s} \beta^\star G(t)$$

si l'une des propriétés suivantes est satisfaite :

1) S'il existe un réel $0 < q < 2$ *tel que la variable aléatoire* $\bar{Z}_0$ *appartient à* $\mathbb{L}_{\psi_{\theta(q)}}$ *et sous la condition*

$$\sum_{k\in\mathbb{Z}^d} \sqrt{\phi_{\infty,1}(|k|)} < \infty. \tag{5.2.1}$$

Alors

$$\widehat{H}_n(t) - \widetilde{H}(t) = \mathcal{O}_{p.s.}\left(\frac{(\log n)^{1/q}}{(b_n\sqrt{n})^d}\right) + \mathcal{O}(b_n^m) \quad \textit{quand} \quad n \to \infty.$$

2) Si $\bar{Z}_0$ appartient à $\mathbb{L}^\infty$ et sous la condition

$$\sum_{k\in\mathbb{Z}^d} \phi_{\infty,1}(|k|) < \infty. \tag{5.2.2}$$

Alors

$$\widehat{H}_n(t) - \widetilde{H}(t) = \mathcal{O}_{p.s.}\left(\frac{(\log n)^{1/2}}{(b_n\sqrt{n})^d}\right) + \mathcal{O}(b_n^m) \quad \textit{quand} \quad n \to \infty.$$

3) S'il existe un réel $p > 2$, tel que la variable aléatoire $\bar{Z}_0$ appartient à $\mathbb{L}^p$ et sous la condition

$$\sum_{k\in\mathbb{Z}^d} \left(\int_0^{\alpha_{1,\infty}(|k|)} Q^p_{Z_0-\mathrm{E}\,Z_0}(u)\,du\right)^{2/p} < \infty \tag{5.2.3}$$

où $Q_{Z_0-\mathrm{E}\,Z_0}$ est l'inverse càdlàg de la fonction $t \to \mathbb{P}(|Z_0 - \mathrm{E}\,Z_0| > t)$ (i.e. pour tout $u \geq 0$, $Q_{Z_0-\mathrm{E}\,Z_0}(u) = \inf\{t > 0 \,|\, \mathbb{P}(|Z_0 - \mathrm{E}\,Z_0| > t) \leq u\}$). Alors

$$\widehat{H}_n(t) - \widetilde{H}(t) = \mathcal{O}_{p.s.}\left(\frac{(\log n)^{1/p}}{(b_n\sqrt{n})^d}\right) + \mathcal{O}(b_n^m) \quad \textit{quand} \quad n \to \infty.$$

5.2.2 Consistance forte globale

Nous donnons un résultat uniforme pour la convergence presque sûre de notre famille des estimateurs sur un compact T_0 de T. L'obtention de résultats uniformes nécessite la condition suivant :

Condition $\mathcal{Z}$:

La fonction noyau K est lipschitzienne, c'est à dire qu'il existe une constante $\eta > 0$ telle que

$$\big|K(u) - K(v)\big| \leq \eta||u - v|| \quad \text{pour tous} \quad u, v \in T.$$

Nous établissons maintenant la consistance forte globale de l'estimateur $\widehat{G}_n(t)$.

Théorème 5.2.2. *Sous les conditions $\mathbf{K}^0$, $\mathcal{P}(T)$ et $\mathcal{Z}$. De plus, on suppose que $G(t)\bar{F}(o,t)$ admet des dérivées partielles d'ordre m continues et bornées sur T. Alors,*

$$\sup_{t\in T_o} \big|\widehat{G}_n(t) - \beta^\star G(t)\big| \xrightarrow[n\to\infty]{p.s} 0$$

si l'une des propriétés suivantes est satisfaite :

1) S'il existe un réel $0 < q < 2$ tel que la variable aléatoire $\bar{Z}_0$ appartient à $\mathbb{L}_{\psi_{\theta(q)}}$ et sous la condition

$$\sum_{k\in\mathbb{Z}^d} \sqrt{\phi_{\infty,1}(|k|)} < \infty.$$

Alors

$$\sup_{t\in T_o} \left|\widehat{H}_n(t) - \widetilde{H}(t)\right| = \mathcal{O}_{p.s.}\left(\frac{(\log n)^{1/q}}{(b_n\sqrt{n})^d}\right) + \mathcal{O}(b_n^m) \quad \textit{quand} \quad n \to \infty.$$

2) Si $\bar{Z}_0$ appartient à $\mathbb{L}^\infty$ et sous la condition

$$\sum_{k\in\mathbb{Z}^d} \phi_{\infty,1}(|k|) < \infty.$$

Alors

$$\sup_{t\in T_o} \left|\widehat{H}_n(t) - \widetilde{H}(t)\right| = \mathcal{O}_{p.s.}\left(\frac{(\log n)^{1/2}}{(b_n\sqrt{n})^d}\right) + \mathcal{O}(b_n^m) \quad \textit{quand} \quad n \to \infty.$$

3) S'il existe un réel $p > 2$ tel que la variable aléatoire $\bar{Z}_0$ appartient à $\mathbb{L}^p$ et sous la condition

$$\sum_{k\in\mathbb{Z}^d} \left(\int_0^{\alpha_{1,\infty}(|k|)} Q^p_{Z_0-\mathrm{E}\,Z_0}(u)\,du\right)^{2/p} < \infty$$

où $Q_{Z_0-\mathrm{E}\,Z_0}$ est l'inverse càdlàg de la fonction $t \to \mathbb{P}(|Z_0 - \mathrm{E}\,Z_0| > t)$ (i.e. pour tout $u \geq 0$, $Q_{Z_0-\mathrm{E}\,Z_0}(u) = \inf\{t > 0 \,|\, \mathbb{P}(|Z_0 - \mathrm{E}\,Z_0| > t) \leq u\}$). Soit $b_n = n^{-q_2}(\log n)^{q_1}$ pour certaines constantes $q_1, q_2 > 0$. Soient $a, b \geq 0$ fixés si

$$a(p+d) - d^2/2 - q_2 d > 1 \quad \textit{et} \quad b(p+d) + q_1 d > 1.$$

Alors

$$\sup_{t\in T_o} \left|\widehat{H}_n(t) - \widetilde{H}(t)\right| = \mathcal{O}_{p.s.}\left(\frac{n^a(\log n)^b}{(b_n\sqrt{n})^d}\right) + \mathcal{O}(b_n^m) \quad \textit{quand} \quad n \to \infty.$$

5.2.3 Normalité asymptotique

Nous commençons par rappeler le théorème central limite qui a été établi par Coeurjolly et Lavancier [29]. Ce résultat va nous permettre d'obtenir la normalité asymptotique de l'estimateur $\widehat{G}_n(t)$.

Theorem 5.2.3. *Soit $X_{n,i}, n \in \mathbb{N}, i \in \mathbb{Z}^d$, une suite de champs dans un espace mesurable. Pour $n \in \mathbb{N}$, soit $K_n \subset \mathbb{Z}^d$ et pour tout $k \in K_n$, supposons $Z_{n,k} = f_{n,k}(X_{n,k+i}, i \in$*

I_0), où $I_0 = \{i \in \mathbb{Z}^d, |i| \leq 1\}$, $f_{n,k} : S^{I_0} \to \mathbb{R}^p$. Soit $S_n = \sum_{k \in K_n} Z_{n,k}$. Si
i) $sup_{n\in\mathbb{N}} sup_{k\in K_n} E|Z_{n,k}|^3 < \infty$,
ii) $\forall n \in \mathbb{N}, \forall k \in K_n, E\left[Z_{n,k}|X_{n,j}, j \neq k\right] = 0$.
iii) $|K_n| \to +\infty$ quand $n \to +\infty$.
iv) Il existe une matrice symétrique $\underline{\sum} \geq 0$ telle que

$$E\left[\left|\left||K_n|^{-1} \sum_{k\in K_n} \sum_{j\in B(k,1)\cap K_n} Z_{n,k} Z_{n,j}^T - \underline{\sum}\right|\right|\right] \longrightarrow 0,$$

alors $|K_n|^{\frac{-1}{2}} S_n \xrightarrow{d} \mathcal{N}(0, \underline{\sum})$.

Etudions maintenant la normalité asymptotique de notre famille des estimateurs.

Théorème 5.2.4. *Soit $\mathbf{X}$ un processus ponctuel à interaction par paires stationnaire défini en (3.1.1) et soient $t_1, ..., t_s$, s points distincts de $T\backslash\{o\}$. Sous les conditions $\mathbf{K}^0$, $b_n^{d+2m}|\Lambda_{n,R}| \to 0$ quand $n \to \infty$ et sous la condition*

$$\sum_{k\in\mathbb{Z}^d} \phi_{\infty,1}(|k|) < \infty. \tag{5.2.4}$$

Alors

$$\sqrt{b_n^d|\Lambda_{n,R}|}\left(\widehat{G}_n(t_i) - \beta^\star G(t_i)\right)_{i=1}^s \xrightarrow{d} \mathcal{N}(0, \sigma_G)$$

avec la matrice de covariance $\sigma_G = [\sigma_{i,j}]_{i,j=1,...,s}$ où $\sigma_{i,j}$ a pour élément

$$\sigma_{i,i} = \frac{G(t_i)}{\bar{F}(o,t_i)} \int_{\mathbb{R}^d} K^2(z)\mathrm{d}z \quad \textit{pour } i = 1, ..., s \quad \textit{et} \quad \sigma_{i,j} = 0 \quad \textit{si} \quad i \neq j.$$

5.3 Résultats asymptotiques de l'estimateur $\widehat{\Phi}_n(r)$

En introduisant des notations et des conditions similaires à celles présentées dans le Chapitre 4. L'estimateur de $\beta^\star\Phi(r)$, où $\Phi(r) = \exp(-\gamma(r))$ corrigeant les effets de bord par l'introduction du $2R$-intérieur de Λ_n, i.e.

$$\Lambda_{n,R} = \{u \in \Lambda_n : B(u, 2R) \subset \Lambda_n\}, \quad \text{avec} \quad |\Lambda_{n,R}| > 0$$

est défini pour $r \in (0, R]$ par

$$\widehat{\Phi}_n(r) = \frac{\widehat{R}_n(r)}{\widehat{N}_n(r)}$$

où

$$\widehat{R}_n(r) = \frac{1}{b_n|\Lambda_{n,R}|\sigma_d} \sum_{\substack{u,v\in\mathbf{X} \\ ||v-u||\leq R}}^{\neq} \frac{\mathbb{1}_{\Lambda_{n,R}}(u)}{||v-u||^{d-1}} h(u,\mathbf{X}\backslash\{u,v\})h(v,\mathbf{X}\backslash\{u,v\})K_1\left(\frac{||v-u||-r}{b_n}\right)$$

est un estimateur de $\beta^{\star 2}\Phi(r)J(r)$, où $J(r) = \int_{\mathbb{S}^{d-1}} \bar{F}(o,rv)\mathrm{d}v$.

$$\widehat{N}_n(r) = \frac{1}{|\Lambda_{n,R}|} \sum_{u\in\mathbf{X}} \mathbb{1}_{\Lambda_{n,R}}(u)h(u,\mathbf{X}\setminus\{u\})h^\star(u,\mathbf{X}\setminus\{u\})$$

estimateur de $\beta^\star J(r)$. Supposons que $\Lambda_{n,R}$ se décompose comme $\Lambda_{n,R} = \cup_{i\in\widetilde{I}_n}\Lambda_i$, où Λ_i sont des sous-cubes de $\Lambda_{n,R}$ centrés en i et $\widetilde{I}_n$ une suite de sous-ensembles de $\mathbb{Z}^d$, telle que le cardinal $|\widetilde{I}_n|$ de $\widetilde{I}_n$ tend vers l'infini, quand n tend vers l'infini. Notons pour tout $k \in \mathbb{Z}^d$,

$$\bar{R}_k = R_k - \mathrm{E}\, R_k \quad \text{et} \quad S_n = \sum_{i\in\widetilde{I}_n} \bar{R}_i$$

avec

$$R_k = \sum_{\substack{u,v\in\mathbf{X} \\ ||v-u||\leq R.}}^{\neq} \frac{\mathbb{1}_{\Lambda_k}(u)}{||v-u||^{d-1}} h(u,\mathbf{X}\setminus\{u,v\})h(v,\mathbf{X}\setminus\{u,v\})K_1\left(\frac{||v-u||-r}{b_n}\right)$$

5.3.1 Consistance forte ponctuelle

Nous étudions dans ce paragraphe les propriétés ponctuelles de la consistance forte de l'estimateur $\widehat{\Phi}_n(r)$, en tenant compte des résultats de la Section 5.1 et à partir des propriétés asymptotiques de l'estimateur lissé $\widehat{R}_n(r)$ et $\widehat{N}_n(r)$, ainsi sous la condition de l'unicité de Dobrushin (Dobrushin [47] ; Georgii [58] ; Guyon [62]), nous avons les résultats suivants.

Proposition 5.3.1. *Soit K_1 un noyau non négatif à support compact d'ordre $\alpha \geq 1$, tel que $\int_{\mathbb{R}} K_1^2(\rho)\mathrm{d}\rho < \infty$, on suppose de plus que $\exp(-\gamma(r))J(r)$ admet des dérivées partielles d'ordre α continues et bornées sur $(r-\delta, r+\delta)$, pour un certain $\delta > 0$. Alors,*

$$\widehat{\Phi}_n(r) \xrightarrow[n\to\infty]{p.s} \beta^\star\Phi(r)$$

si l'une des propriétés suivantes est satisfaite :
1) Si $\bar{R}_0$ appartient à $\mathbb{L}^\infty$ et sous la condition

$$\sum_{k\in\mathbb{Z}^d} \phi_{\infty,1}(|k|) < \infty.$$

Alors

$$\widehat{R}_n(r) - R(r) = \mathcal{O}_{p.s.}\left(\frac{(\log n)^{1/2}}{b_n n^{d/2}}\right) + \mathcal{O}(b_n^\alpha) \quad \textit{quand} \quad n \to \infty$$

2) S'il existe un réel $0 < q < 2$ tel que la variable aléatoire $\bar{R}_0$ appartient à $\mathbb{L}_{\psi_{\theta(q)}}$ et sous la condition

$$\sum_{k\in\mathbb{Z}^d} \sqrt{\phi_{\infty,1}(|k|)} < \infty.$$

Alors

$$\widehat{\Phi}_n(r) - \Phi(r) = \mathcal{O}_{p.s.}\left(\frac{(\log n)^{1/q}}{b_n n^{d/2}}\right) + \mathcal{O}(b_n^\alpha) \quad \textit{quand} \quad n \to \infty$$

3) S'il existe un réel $p > 2$ tel que la variable aléatoire $\bar{R}_0$ appartient à $\mathbb{L}^p$ et sous la condition

$$\sum_{k\in\mathbb{Z}^d} \left(\int_0^{\alpha_{1,\infty}(|k|)} Q^p_{R_0-\mathrm{E}\,R_0}(u)\,du\right)^{2/p} < \infty,$$

où $Q_{R_0-\mathrm{E}\,R_0}$ est l'inverse càdlàg de la fonction $t \to \mathbb{P}(|R_0 - \mathrm{E}\,R_0| > t)$. Soient $a, b \geq 0$ fixés et si

$$a(p+1) - d/2 - q_2 d > 1 \quad \textit{et} \quad b(p+1) + q_1 d > 1.$$

Alors

$$\widehat{R}_n(r) - R(r) = \mathcal{O}_{p.s.}\left(\frac{(\log n)^{1/p}}{b_n n^{d/2}}\right) + \mathcal{O}(b_n^\alpha) \quad \textit{quand} \quad n \to \infty.$$

5.3.2 Consistance forte globale

Nous établissons maintenant la consistance forte globale de l'estimateur $\widehat{\Phi}_n(r)$ de la fonction d'interaction de paires dans l'intervalle $[r_1, r_2]$.

Théorème 5.3.2. *Sous les conditions $\mathcal{P}(D)$ et $\mathcal{Z}$. De plus si $\exp(-\gamma(r))J(r)$ admet des dérivées partielles d'ordre α continues et bornées sur $[r_1-\delta, r_2+\delta]$ pour un certain $\delta > 0$. Soit K_1 un noyau non négatif d'ordre $\alpha \geq 1$, tel que $\int_{\mathbb{R}} K_1^2(\rho)\mathrm{d}\rho < \infty$. Alors*

$$\sup_{r_1\leq r\leq r_2} \left|\widehat{\Phi}_n(r) - \beta^\star\Phi(r)\right| \underset{n\to\infty}{\longrightarrow} \quad p.s.$$

si l'une des propriétés suivantes est satisfaite :
1) Si $\bar{R}_0$ appartient à $\mathbb{L}^\infty$ et sous la condition

$$\sum_{k\in\mathbb{Z}^d} \phi_{\infty,1}(|k|) < \infty.$$

Alors

$$\sup_{r_1 \le r \le r_2} \left|\widehat{R}_n(r) - R(r)\right| = \mathcal{O}_{p.s.}\left(\frac{(\log n)^{1/2}}{b_n n^{d/2}}\right) + \mathcal{O}(b_n^{\alpha}) \quad \textit{quand} \quad n \to \infty.$$

2) S'il existe un réel $0 < q < 2$ tel que la variable aléatoire $\bar{R}_0$ appartient à $\mathbb{L}_{\psi_{\theta(q)}}$ et sous la condition

$$\sum_{k \in \mathbb{Z}^d} \sqrt{\phi_{\infty,1}(|k|)} < \infty.$$

Alors

$$\sup_{r_1 \le r \le r_2} \left|\widehat{\Phi}_n(r) - \Phi(r)\right| = \mathcal{O}_{p.s.}\left(\frac{(\log n)^{1/q}}{b_n n^{d/2}}\right) + \mathcal{O}(b_n^{\alpha}) \quad \textit{quand} \quad n \to \infty.$$

3) S'il existe un réel $p > 2$ tel que la variable aléatoire $\bar{R}_0$ appartient à $\mathbb{L}^p$ et sous la condition

$$\sum_{k \in \mathbb{Z}^d} \left(\int_0^{\alpha_{1,\infty}(|k|)} Q^p_{Z_0 - \mathrm{E}\, Z_0}(u)\, du\right)^{2/p} < \infty$$

où Q_{X_0} est l'inverse càdlàg de la fonction $t \to \mathbb{P}(|X_0| > t)$ (i.e. pour tout $u \ge 0$, $Q_{Z_0 - \mathrm{E}\, Z_0}(u) = \inf\{t > 0 \,|\, \mathbb{P}(|Z_0 - \mathrm{E}\, Z_0| > t) \le u\}$). Soit $b_n = n^{-q_2}(\log n)^{q_1}$ pour certaines constantes $q_1, q_2 > 0$. Soient $a, b \ge 0$ fixés, si

$$a(p+1) - d/2 - q_2 d > 1 \quad \textit{et} \quad b(p+1) + q_1 d > 1.$$

Alors

$$\sup_{r_1 \le r \le r_2} \left|\widehat{R}_n(r) - R(r)\right| = \mathcal{O}_{p.s.}\left(\frac{n^a (\log n)^b}{b_n n^{d/2}}\right) + \mathcal{O}(b_n^{\alpha}) \quad \textit{quand} \quad n \to \infty.$$

5.3.3 Normalité asymptotique

Le théorème suivant établit la normalité asymptotique de l'estimateur $\widehat{\Phi}_n(r)$ de la fonction d'interaction de paires Φ.

Théorème 5.3.3. *Soit Φ la fonction d'interaction de paires définie dans (4.1.1). Soit le q-uplet $(r_1, ..., r_q) \in (0, R]$, avec $r_i \neq r_j$ pour $i \neq j$ tel que r_i points de continuité de Φ pour chaque $i = 1, ..., q$ et K_1 un noyau à support compact d'ordre $\alpha \ge 1$, tel que $\int_{\mathbb{R}} K_1^2(\rho)\mathrm{d}\rho < \infty$, $b_n^{1+2\alpha}|\Lambda_{n,R}| \to 0$ quand $n \to \infty$ et sous la condition*

$$\sum_{k \in \mathbb{Z}^d} \phi_{\infty,1}(|k|) < \infty.$$

Alors

$$\sqrt{b_n|\Lambda_{n,R}|}\Big(\widehat{\Phi}_n(r_i)-\beta^\star\Phi(r_i)\Big)_{i=1}^q \xrightarrow{d} \mathcal{N}(0,\sigma_\Phi).$$

avec la matrice de covariance $\sigma_\Phi=[\sigma_{i,j}]_{i,j=1,\dots,q}$ *où* $\sigma_{i,j}$ *a pour élément*

$$\sigma_{i,i}=\frac{2\Phi(r_i)}{\sigma_d r_i^{d-1}J(r_i)}\int_{\mathbb{R}}K_1^2(\rho)\mathrm{d}\rho \quad \textit{pour } i=1,\dots,q \quad \textit{et} \quad \sigma_{i,j}=0 \quad \textit{si} \quad i\neq j.$$

5.4 Preuves

5.4.1 Preuve de la Proposition 5.2.2

Démonstration. Nous prouvons maintenant la Proposition 5.2.2 en suivant le schéma de preuve que nous l'avait indiqué au Chapitre 3. Nous commence notamment par la décomposition T_0 par v_n cubes T_k ayant des côtés de longueur $l_n=r_nb_n^{d+1}$ de centre c_k, tels qu'il existe $c>0$, $v_n\leq c/l_n^d$ et soit $(r_n)_{n\geq 1}$ une suite de nombres positifs tendant vers zéro. Nous avons donc

$$\begin{aligned}\sup_{t\in T_o}|\widehat{H}_n(t)-\mathrm{E}\,\widehat{H}_n(t)| &\leq \max_{1\leq k\leq v_n}\sup_{t\in T_k}\big|\widehat{H}_n(t)-\widehat{H}_n(c_k)\big|\\ &+\max_{1\leq k\leq v_n}\sup_{t\in T_k}\big|\mathrm{E}\,\widehat{H}_n(t)-\mathrm{E}\,\widehat{H}_n(c_k)\big|\\ &+\max_{1\leq k\leq v_n}\big|\widehat{H}_n(c_k)-\mathrm{E}\,\widehat{H}_n(c_k)\big|\\ &=A_1+A_2+A_3 \qquad (5.4.1)\end{aligned}$$

Le premier terme A_1 et le deuxième terme A_2 ont été étudiés dans le Lemme 3.6.3, il résulte que

$$A_1+A_2=\mathcal{O}_{p.s.}(r_n).$$

Focalisons-nous sur le terme A_3.
1) Pour $\varepsilon>0$ fixé, $r_n=(\log n)^{1/q}/(b_n\sqrt{n})^d$, alors pour n assez grand

$$\mathbb{P}\left(|A_3|>\varepsilon r_n\right)\leq v_n\sup_{t\in T_0}\mathbb{P}(|\widehat{H}_n(t)-\mathrm{E}\,\widehat{H}_n(t)|>\varepsilon r_n).$$

Soint ε un réel positif et une suite $(r_n)_{n>0}$ de réels strictement positifs intermédiaire, tendant vers zéro. En utilisant l'inégalité de Markov, nous obtenous

$$\begin{aligned}\mathbb{P}(|\widehat{H}_n(t)-\mathrm{E}\,\widehat{H}_n(t)|>\varepsilon r_n) &\leq \mathbb{P}\left(|S_n|>\varepsilon r_n(b_nn)^d\right)\\ &\leq \exp\left[-\left(\frac{\varepsilon\, r_n(b_nn)^d}{||S_n||_{\psi_{\theta(q)}}}+\xi_q\right)^q\right]\mathrm{E}\exp\left[\left(\frac{|S_n|}{||S_n||_{\psi_{\theta(q)}}}+\xi_q\right)^q\right].\end{aligned}$$

Par conséquent pour $0 < q < 2$, nous utilisons l'inégalité (5.1.1) du Corollaire 5.1.1, nous obtenons

$$\mathbb{P}(|\widehat{H}_n(t) - \mathrm{E}\,\widehat{H}_n(t)| > \varepsilon r_n) = \mathbb{P}\left(|S_n| > \varepsilon r_n (b_n n)^d\right)$$
$$\leq (1 + e^{\xi_q^q}) \exp\left[- \left(\frac{\varepsilon\, r_n (b_n n)^d}{M_1(q) \|\bar{Z}_0\|^2_{\psi_{\theta(q)}} (\sum_{i \in \widetilde{I}_n} \widetilde{b}_{i,q}(\bar{Z}))^{1/2}} + \xi_q \right)^q \right].$$

où

$$\widetilde{b}_{i,q}(X) = 1 + C(q) \sum_{k \in V_0^1} \sqrt{\phi_{\infty,1}(|k|)}.$$

Ainsi sous la condition de l'unicité de Dobrushin (Dobrushin [47] ; Georgii [58] ; Guyon [62]) (voir le Théorème 2.3.1) les conditions (5.2.1) et (5.2.2) sont satisfaites, ce qui entraîne en plus la sationnarité de $\mathbf{X}$ et par définition de la norme $|| \cdot ||_{\psi_{\theta(q)}}$, qu' il existe une constante $C > 0$ telle que

$$\mathbb{P}(|\widehat{H}_n(t) - \mathrm{E}\,\widehat{H}_n(t)| > \varepsilon r_n) \leq (1 + e^{\xi_q^q}) \exp\left[- \frac{\varepsilon^q r_n^q (b_n \sqrt{n})^{dq}}{C^q} \right].$$

Si $r_n = (\log n)^{1/q}/(b_n \sqrt{n})^d$ et pour n suffisamment grand, alors

$$\mathbb{P}(|\widehat{H}_n(t) - \mathrm{E}\,\widehat{H}_n(t)| > \varepsilon r_n) \leq (1 + e^{\xi_q^q}) \exp\left[- \frac{\varepsilon^q \log n}{C^q} \right].$$

Nous obtenons

$$\sup_{t \in T_0} \mathbb{P}(|\widehat{H}_n(t) - \mathrm{E}\,\widehat{H}_n(t)| > \varepsilon r_n) \leq (1 + e^{\xi_q^q}) \exp\left[- \frac{\varepsilon^q \log n}{C^q} \right]. \tag{5.4.2}$$

2) En reformulant l'inégalité de Markov et nous utilisons l'inégalité (5.1.2) du Corollaire 5.1.1, nous obtenons, pour $Z_0 \in \mathbb{L}^\infty$ et pour ε un réel positif,

$$\mathbb{P}(|\widehat{H}_n(t) - \mathrm{E}\,\widehat{H}_n(t)| > \varepsilon r_n) = \mathbb{P}\left(|S_n| > \varepsilon r_n (b_n n)^d\right)$$
$$\leq (1 + e^{\xi_q^q}) \exp\left[- \left(\frac{\varepsilon\, r_n (b_n n)^d}{M_2(q) \|\bar{Z}_0\|_\infty (\sum_{i \in \widetilde{I}_n} \widetilde{b}_{i,\infty}(\bar{Z}))^{1/2}} + \xi_q \right)^q \right].$$

où

$$\widetilde{b}_{i,\infty}(\bar{Z}) = 1 + 2 \sum_{k \in V_0^1} \phi_{\infty,1}(|k|).$$

donc si $r_n = (\log n)^{1/2}/(b_n \sqrt{n})^d$ et n suffisamment, alors il existe une constante $C > 0$ telle que

$$\mathbb{P}(|\widehat{H}_n(t) - \mathrm{E}\,\widehat{H}_n(t)| > \varepsilon r_n) \leq 2 \exp\left[- \frac{\varepsilon^2 \log n}{C^2} \right].$$

Nous obtenons ainsi,

$$\sup_{t\in T_0} \mathbb{P}(|\widehat{H}_n(t) - \mathrm{E}\,\widehat{H}_n(t)| > \varepsilon r_n) \leq 2\,\exp\left[-\frac{\varepsilon^2\,\log n}{C^2}\right]. \tag{5.4.3}$$

En utilisant (5.4.2) et (5.4.3), pour ε suffisamment large, donc, il en résulte grâce au Lemme de Borel-Cantelli, que

$$\mathbb{P}(\limsup_{n\to\infty} |A_3| > \varepsilon r_n) = 0.$$

Les deux premiers résultats du Théorème 5.2.2 découlent alors de (5.4.1).

3) En utilisant l'inégalité de Markov, nous avons pour tout $\varepsilon > 0$,

$$\begin{aligned}\mathbb{P}(|\widehat{H}_n(t) - \mathrm{E}\,\widehat{H}_n(t)| > \varepsilon r_n) &= \mathbb{P}\left(|S_n| > \varepsilon r_n (b_n n)^d\right)\\ &\leq \frac{\varepsilon^{-p}\,\mathrm{E}\,|S_n|^p}{r_n^p (b_n n)^{pd}}\\ &\leq \frac{\varepsilon^{-p}}{r_n^p (b_n n)^{pd}}\left(2p\sum_{i\in\widetilde{I}_n} c_i(\bar{Z})\right)^{p/2}.\end{aligned}$$

La dernière inégalité résulte d'après la Proposition 3.4.2, où

$$c_i(Z) = \|\bar{Z}_i\|_p^2 + \sum_{k\in V_i^1} \|\bar{Z}_k E_{|k-i|}(\bar{Z}_i)\|_{\frac{p}{2}}.$$

En utilisant l'inégalité de Rio [110] (voir aussi Dedecker [37]), nous obtenons

$$\|\bar{Z}_k E_{|k|}(\bar{Z}_0)\|_{\frac{p}{2}} \leq 4\left(\int_0^{\alpha_{1,\infty}(|k|)} Q^p_{\bar{Z}_0}(u)\,du\right)^{2/p}.$$

Ainsi sous la condition de l'unicité de Dobrushin (Dobrushin [47] ; Georgii [58] ; Guyon [62]) et d'après le Théorème 2.13 de Guyon [62], la condition (5.3.2) est satisfaite. Nous déduisons par conséquent qu'il existe une constante $\kappa > 0$ telle que

$$\mathbb{P}(|\widehat{H}_n(t) - \mathrm{E}\,\widehat{H}_n(t)| > \varepsilon r_n) \leq \frac{\kappa\varepsilon^{-p}}{r_n^p b_n^p (\sqrt{n})^{pd}}.$$

Nous concluons

$$\begin{aligned}\mathbb{P}\left(|A_3| > \varepsilon r_n\right) &\leq v_n \sup_{t\in T_0} \mathbb{P}(|\widehat{H}_n(t) - \mathrm{E}\,\widehat{H}_n(t)| > \varepsilon r_n)\\ &\leq v_n \frac{\kappa\varepsilon^{-p}}{r_n^p (b_n\sqrt{n})^{pd}}.\end{aligned}$$

Comme $v_n \leq c/l_n^d$ et $l_n = r_n b_n^{1+d}$, donc pour $r_n = n^a(\log n)^b/(b_n\sqrt{n})^d$, il résulte pour n suffisamment grand,

$$\begin{aligned}\mathbb{P}\left(|A_3| > \varepsilon r_n\right) &\leq \frac{\kappa\varepsilon^{-p}}{n^{a(p+d)-d^2/2}(\log n)^{b(p+d)}b_n^d} \\ &\leq \frac{\kappa\varepsilon^{-p}}{n^{a(p+d)-d^2/2-q_2 d}(\log n)^{b(p+d)+q_1 d}}.\end{aligned}$$

Pour $a(p+d) - d^2/2 - q_2 d > 1$ et $b(p+d) + q_1 d > 1$, nous obtenons pour tout réel $\varepsilon > 0$

$$\sum_{n\geq 1} \mathbb{P}\left(|A_3| > \varepsilon r_n\right) < \infty.$$

Nous concluons la preuve du Théorème 5.2.2, grâce à l'inégalité (5.4.1) et la Proposition 3.2.1. et en utilisant la continuité uniforme de $\widetilde{H}$, d'une part la partie biais converge vers 0 d'après □

5.4.2 Preuve du Théorème 5.3.2

Démonstration. La preuve du Théorème 5.3.2 se fait avec les mêmes étapes que celles de la preuve du Théorème 5.2.2. □

5.4.3 Preuve du Théorème 5.2.4

Démonstration. Nous supposons que $\Lambda_n = [0, n)^d$, se décompose comme, $\Lambda_n = \cup_{i\in I_n}\Lambda_i$ et que le processus est observé dans $\Lambda_{n,R} = \cup_{i\in\widetilde{I}_n}\Lambda_i$, avec $\widetilde{I}_n = \{i \in I_n : |i-j| \leq 1, \quad \text{pour tout} \quad j \in I_n\}$, où $|j| = \max\{|j_1|, ..., |j_d|\}$. Nous supposerons que I_n croît vers $\mathbb{Z}^d$ et $\frac{|\partial I_n|}{|I_n|}$ tend vers 0 quand n tend vers l'infini, où $\partial I_n = \{i \in I_n \mid \exists j \notin I_n : |j-i| = 1\}$. Notons pour tout $l = 1, 2, ..., s$,

$$(b_n^d n^d)^{1/2}\left(\widehat{H}_n(t_l) - \mathrm{E}[\widehat{H}_n(t_l)]\right) = \frac{1}{n^{d/2}}\sum_{k\in\widetilde{I}_n}\bar{Z}_{n,k},$$

où

$$\bar{Z}_{n,k} = Z_{n,k}^{(l)} - \mathrm{E}\, Z_{n,k}^{(l)}, k \in \mathbb{Z}^d.$$

Nous considérons

$$Z_{n,k}^{(l)} = (b_n^d)^{-1/2}\sum_{\substack{u\in X_{\Lambda_k}, v\in\mathbf{X} \\ v-u\in B(o,R)}}^{\neq} h(u, \mathbf{X}\setminus\{u,v\})h(v, \mathbf{X}\setminus\{u,v\})K\left(\frac{v-u-t_l}{b_n}\right).$$

Pour commencer, il nous faut introduire quelques variables qui joueront un rôle clef dans nos résultats. Nous construisons deux suites de variables aléatoires réelles définies par

$$\bar{Y}_{n,k} = Z_{n,k}^{(l)} - \mathrm{E}\left[Z_{n,k}^{(l)} | X_{\Lambda_j}, j \neq k\right],\ k \in \mathbb{Z}^d \tag{5.4.4}$$

et

$$\widetilde{Y}_{n,k} = \mathrm{E}\left[Z_{n,k}^{(l)} | X_{\Lambda_j}, j \neq k\right] - \mathrm{E}\, Z_{n,k}^{(l)},\ k \in \mathbb{Z}^d.$$

Nous décomposons donc

$$\frac{1}{n^{d/2}} \sum_{k \in \widetilde{I}_n} \bar{Z}_{n,k} = \frac{1}{n^{d/2}} \sum_{k \in \widetilde{I}_n} \bar{Y}_{n,k} + \frac{1}{n^{d/2}} \sum_{k \in \widetilde{I}_n} \widetilde{Y}_{n,k}.$$

Alors pour aboutir à notre résultat, il suffit d'établir les deux résultats suivants.

Proposition 5.4.1. *Pour n suffisament grand, on a*

$$\frac{1}{n^{d/2}} \sum_{k \in \widetilde{I}_n} \bar{Y}_{n,k} \xrightarrow{d} \mathcal{N}\big(0, \sigma^2\big)$$

$$\sigma^2 = \beta^{\star 2} \bar{F}(o, t_l) G(t_l) \int_{\mathbb{R}^d} K^2(z) \mathrm{d}z, l = 1, 2, ..., s. \tag{5.4.5}$$

Proposition 5.4.2. *Sous la condition (5.2.4). Alors, pour n suffisament grand, on a*

$$\frac{1}{n^{d/2}} \sum_{k \in \widetilde{I}_n} \widetilde{Y}_{n,k} \xrightarrow{\mathbb{P}} 0.$$

Preuve de la Proposition 5.4.1

Démonstration. Soint $\{\bar{Y}_{n,k}; k \in \mathbb{N}^d\}$ un tableau triangulaire stationnaire asymptotiquement sans biais défini en (5.4.4). Nous vérifions les deux conditions fondamentales du Théorème 5.2.3 :

(i) $sup_{n \in \mathbb{N}} sup_{k \in \widetilde{I}_n} E|\bar{Y}_{n,k}|^3 < \infty$,

(ii) $\forall n \in \mathbb{N}, \forall k \in \widetilde{I}_n, E\left[\bar{Y}_{n,k} | X_j, j \neq k\right] = 0.$

La condition (ii) est évidente. Vérifions la condition (i). A cet égard, nous avons

$$\begin{aligned}\|\bar{Y}_{n,0}\|_3 &= \|Z_{n,0}^{(l)} - \mathrm{E}[Z_{n,0}^{(l)} \mid X_{\Lambda_0^c}]\|_3 \\ &\leq \|Z_{n,0}^{(l)}\|_3 + \|\,\mathrm{E}[Z_{n,0}^{(l)} \mid X_{\Lambda_0^c}\|_3 \\ &\leq 2\|Z_{n,0}^{(l)}\|_3,\end{aligned}$$

d'après Minkowski et les inégalités conditionnelles de Jensen. D'autre part, puisque le noyau K est borné

$$\begin{aligned}\|Z_{n,0}^{(l)}\|_3 &\leq \frac{c}{\sqrt{b_n^d}} \operatorname{Var}\left(\frac{1}{n^d}\sum_{k\in\widetilde{I}_n} Z_{n,k}^{(l)}\right) \\ &\leq \frac{c}{\sqrt{b_n^d n^d}} \operatorname{Var}\left(\sqrt{b_n^d n^d}\,\widehat{H}_n(t_l)\right) \\ &\leq \frac{c}{\sqrt{b_n^d n^d}}\left[\beta^{\star 2}\exp(-g(t_l))\bar{F}(o,t_l)\int_{\mathbb{R}^d} K^2(z)\mathrm{d}z + o(1)\right]\end{aligned}$$

Le Théorème 3.2.2 entraîne la dernière inégalité, ce qui achève notre preuve. □

Preuve de la Proposition 5.4.2

Démonstration. Soit $\mathcal{F}_{i,n} = \mathcal{F}_i = \sigma(X_{\Lambda_j, j\neq i})$ désigne la tribu engendré par les variables aléatoires X_{Λ_j}. Grâce à la propriété $\phi-$mélangeant, nous obtenons

$$\left|\left|\frac{1}{n^{d/2}}\sum_{k\in\widetilde{I}_n}\widetilde{Y}_{n,k}\right|\right|_2^2 \leq \frac{c}{b_n^d n^d}\sum_{k\in\mathbb{Z}^d}\phi_{\infty,1}(|k|) \tag{5.4.6}$$

Ainsi sous la condition (5.2.4) et comme $\frac{1}{b_n^d n^d} \to 0$, il vient alors que

$$\frac{1}{n^{d/2}}\sum_{k\in\widetilde{I}_n}\widetilde{Y}_{n,k} \quad \text{converge en quadratique.}$$

□

Ainsi, nous concluons le résultat recherché en vertu de la preuve du Théorème 5.2.4. □

La preuve du Théorème 5.3.3 suit pas à pas la preuve du Théorème 5.2.4.

Conclusion et Perspectives

Les processus de Gibbs forment une famille de modèles de processus ponctuels. Ils permettent de spécifier l'interaction précise entre points voisins. Cette dernière peut être attractive ou répulsive. Dans le cadre de ma thèse, nous nous sommes intéressés à l'inférence semi-paramétrique des modèles des processus ponctuels de Gibbs caractérisés par l'intensité conditionnelle de Papangelou. Toutes nos considérations dans ce travail de thèse, sont menées par la propriété de portée finie et que nous supposons ne disposer que d'une seule réalisation des processus ponctuels. Deux contextes sont étudiés. Dans le premier, nous avons introduit un nouvel estimateur du paramètre d'intensité de Poisson de l'intensité conditionnelle de Papangelou du processus ponctuels de Gibbs stationnaires. Dans le second, nous avons introduit des nouvelles familles des estimateurs de la fonction d'interaction de paires du processus ponctuel à interaction par paires.

Nous avons développé dans le Chapitre 2 une nouvelle méthode qui permet d'estimer le paramètre d'intensité de Poisson des modèles semi-paramétriques des processus ponctuels de Gibbs stationnaires sans avoir à estimer les termes d'interaction d'ordre supérieur de l'intensité conditionnelle de Papangelou. L'idée sur laquelle l'estimation est basée permet de négliger les termes du paramètre fonctionnel. Dans un premier temps, nous avons précisé une expression analytique de l'estimateur du paramètre d'intensité de Poisson. Sous l'hypothèse que la portée est finie non connue, la consistance forte et la normalité asymptotique sont prouvés. Nous avons illustré des résultats théoriques par des simulations numériques à l'aide des modèles de Gibbs classiquement définis par leur intensité conditionnelle de Papangelou montrant le bon comportement numérique de la méthode d'estimation proposée.

Contrairement à l'inférence statistique paramétrique pour les champs aléatoires de Gibbs, l'inférence non paramétrique parait être moins étudiée. Ainsi dans le Chapitre

3, nous présentons et analysons un estimateur non paramétrique à fonction d'interaction de paires du processus ponctuel à interaction par paires stationnaire, sous des hypothèses aussi générales que possible. En effet, l'estimation de cette fonction ou le potentiel joue un rôle très important en statistique, est appliquée à de nombreux domaines scientifiques tels que la biologie, de l'épidémiologie, de la géographie, de l'astrophysique, de la physique et de l'économie. Dans un premier temps, nous avons construit, pour ce modèle semi-paramétrique défini par son intensité conditionnelle de Papangelou, un nouvel estimateur non paramétrique de fonction d'interaction de paires défini comme un rapport d'un estimateur non paramétrique par lissage et d'un estimateur empirique. Le fait d'avoir explicité les termes dominants du biais et de la variance asymptotiques de l'estimateur par lissage, nous permet notamment d'illustrer la consistance de l'estimateur qui nous intéresse. Nous avons établi aussi la consistance forte ponctuelle et la consistance forte globale sur un compact de l'estimateur par lissage en donnant sa vitesse de convergence et en reposant sur la théorie des estimateurs empiriques, est également applicable pour établir la consistance forte ponctuelle et la consistance forte globale de l'estimateur empirique, ce qui nous permet de conclure la consistance, la consistance forte ponctuelle et la consistance forte globale de l'estimateur de la fonction d'interaction de paires. La preuve de ces résultats est obtenue en utilisant entre autres des inégalités exponentielles de type Kahane-Khintchine dans des espaces d'Orlicz induits par des fonctions de Young exponentielles et par application directe du théorème ergodique adapté aux processus ponctuels spatiaux. Nous avons terminé ce chapitre en illustrant la normalité asymptotique de l'estimateur qui nous intéresse, via le théorème central limite pour les tableaux triangulaires proposé par Wang et Woodroofe [132].

Dans la continuité de ce travail, il était judicieux de faire une étude sur l'estimation de la fonction d'interaction de paires lorsque le processus ponctuel à interaction par paires est stationnaire et isotrope, caractérisé par son intensité conditionnelle de Papangelou. Dans le Chapitre 4, nous proposons un estimateur non paramétrique de la fonction d'interaction de paires. Nous considérons principalement, un estimateur à base de noyau au dénominateur lié à une mesure de distance (distance euclidienne) entre points et un estimateur empirique au numérateur. Nous avons calculé l'erreur en moyenne quadratique de l'estimateur du dénominateur, critère qui s'insère dans le contexte d'estimation non paramétrique par lissage permettant de dériver la convergence en probabilité de l'estimateur. Les principaux résultats que nous avons également obtenus, la consistance forte ponctuelle et la consistance forte globale sur un compact de l'estimateur de la fonction d'interaction de paires. D'un point de vue théorique, la preuve de ces résultats s'appuient sur des inégalités de Kahane-Khintchine dans des espaces d'Orlicz. Enfin, il est possible de montrer la normalité asymptotique de l'estimateur proposé, via le théorème central limite proposé par Wang et

Woodroofe ([132], Théorème 2).

L'objectif dans Chapitre 5 est de reprendre les propriétés asymptotiques étudiées dans le Chapitre 3 et Chapitre 4 sous les conditions du coefficient de mélange. Nous appliquons de nouveaux types d'inégalités, qui s'avèrent aussi très utiles pour les champs aléatoires mélangeants. La procédure fait intervenir des inégalités de mélange de Serfling (voir Serfling [118] ou McLeish [93]). Nous nous sommes intéressés, dans un premier temps, à l'étude de la fonction d'interaction du modèle stationnaire et dans deuxième temps au modèle stationnaire et isotrope. Nous avons établi la consistance forte ponctuelle, la consistance forte globale et la normalité asymptotique des estimateurs de ces modèles.

Dans la continuité de ce travail, plusieurs perspectives sont envisageables. Notre objectif est d'étudier les résultats asymptotiques concernant l'erreur quadratique intégrée pour évaluer l'erreur entre la fonction d'interaction de paires des processus ponctuels à interaction par paires dans le cadre stationnaire et dans le cadre stationnaire et isotrope et ses estimateurs. En suivant cette direction, nous pourrions développer des théorèmes de limite centrale pour les erreurs quadratiques intégrées de nos estimateurs. Ces théorèmes peuvent être utilisés pour construire des tests d'ajustement. Ces travaux futurs seront illustrés par une étude par simulations pour justifier le bon comportement de nos estimateurs. Autres perspectives intéressantes en ce qui concerne les problèmes statistiques liées aux modèles spatiaux marqués. Ils connaissent ces derniers temps un grand intérêt en statistique. Cela s'avère assez utile en pratique lorsqu'on souhaite distinguer la nature des points : spectre d'émission des étoiles, prédateurs ou proies, cellules marquées par des marqueurs radioactifs, espèces d'arbres présents dans une forêt. Nous compterions étudier, toujours dans le cadre de la statistique non-paramétrique ces modèles.

Bibliographie

Bibliographie

[1] A. Baddeley, J. Møller and R.P. Waagepetersen. Non and semi-parametric estimation of interaction in inhomogeneous point partterns. *Statistica neerlandia*, 54 :329–350, 2000.

[2] M. P. Almeida and B. Gidas. A variational method for estimating the parameters of MRF from complete or incomplete data. *Annals of Applied Probability*, 3 :103–136, 1993.

[3] A. Baddeley, J.-F. Coeurjolly, E. Rubak, and R. Waagepetersen. Logistic regression for spatial point processes. *Biometrika*, 101(2) :377–392, 2014.

[4] A. Baddeley and A. Dereudre. Variational estimators for the parameters of Gibbs point process models. *Bernoulli*, 19(3) :905–930, 2013.

[5] A. Baddeley, J. Møller, and A.G. Pakes. Properties of residuals for spatial point processes. *Annals of the Institute of Statistical Mathematics*, 60(3) :627–649, 2008.

[6] A. Baddeley and R. Turner. Practical maximum pseudolikelihood for spatial point patterns (with discussion). *Australian and New Zealand Journal of Statistics*, 42 :283–322, 2000.

[7] A. Baddeley and R. Turner. Modelling spatial point patterns in `R`. *Journal of Statistical Software*, 12(6) :1–42, 2005.

[8] A. Baddeley, R. Turner, J. Møller, and M. Hazelton. Residual analysis for spatial point processes. *Journal of the Royal Statistical Society, Series B*, 67 :1–35, 2005.

[9] E. Bertin, J.-M. Billiot, and R. Drouilhet. Existence of "Nearest-Neighbour" Gibbs Point Models. *Annals of Applied Probability*, 31 :895–909, 1999.

[10] J. Besag. Spatial interaction and the statistical analysis of lattice system. *Journal of the Royal Statistical Society : Series B*, 26 :192–236, 1974.

[11] J. Besag, R. Milne, and S. Zachary. Point process limits of lattice processes. *Applied Probability*, 19 :210–216, 1982.

[12] G. Biau. Spatial kernel density estimation. *Mathematical Methods of Statistics*, 12 :371–390, 2003.

[13] P. Biber. Ein verfahren zum ausgleish von rabdeffekten bei der berechnung von konkurrenzindizes. *Deutscher verband forstlicher forschungsanstalten*, 189-202, 1999.

[14] J.M. Billiot. Estimation d'interactions et étude asymptotique pour les processus ponctuels de gibbs. *Thèse de doctorat en Mathématiques, Université de Montpellier*, 1994.

[15] J.M. Billiot, J.-F. Coeurjolly, and R. Drouilhet. Maximum pseudolikelihood estimator for exponential family models of marked Gibbs point processes. *Electronic Journal of Statistics*, 2 :234–264, 2008.

[16] J.M. Billiot and M. Goulard. An estimation Method of the Pair Potential Function for Gibbs Point Processes on Spheres. *Scandinavian Journal of Statistics*, 28 :185–203, 2001.

[17] E. Bolthausen. On the central limit theorem for stationary mixing random fields. *The Annals of Probability*, 10(4) :1047–1050, 1982.

[18] D. Bosq. Inégalité de Bernstein pour les processus stationnaires et mélangeants. *Applications. CRAS, A*, 281 :1095–98, 1975.

[19] D. Bosq. La statistique non paramétrique des processus. *Rend. Sem. Mat. Univers. Politcn. Torino*, 45(2) :1–24, 1987.

[20] D. Bosq. Nonparametric Statistics for Stochastic Processes-Estimation and Prediction-2nde Edition. *Lecture Notes in Statistics. Springer-Verlag, New York*, 1998.

[21] D. Bosq. *Linear Processes in Function Spaces : Theory and Applications.* 2000.

[22] D. Bosq, F.Merlevède, and M.Peligrad. Asymptotic normality for density kernel estimators in discrete and continuous time. *Journal of Multivariate Analysis*, 68(1) :78–95, 1999.

[23] R. Bradley. Introduction to Strong Mixing Conditions. *Kendrick Press*, 3 :1–3, 2007.

[24] M. Briane and G. Pagès. *Théorie de l'Intégration.* Vuibert, 2000.

[25] M. Carbon, M. Hallin, and B. Wu. Kernel density estimation for random fields : the l_1 theory. *Nonparametric Statist*, 6 :157–170, 1996.

[26] J. V. Castellana and M. R. Leadbettern. On smoothed probability density estimation for stationary processes. *Stochastic Processes and their Applications*, 21(2) :179–193, 1986.

[27] J.-F. Coeurjolly, D. Dereudre, R. Drouilhet, and F. Lavancier. Takacs-Fiksel method for stationary marked Gibbs point processes. *Scandinavian Journal of Statistics*, 49(3) :416–443, 2012.

[28] J.-F. Coeurjolly and R. Drouilhet. Asymptotic properties of the maximum pseudo-likelihood estimator for stationary Gibbs point processes including the lennard-jones model. *Electronic Journal of Statistics*, 4 :677–706, 2010.

[29] J.-F. Coeurjolly and F. Lavancier. Residuals and goodness-of-fit test for stationary marked Gibbs point processes. *Journal of the Royal Statistical Society, Series B*, 75(2) :247–276, 2013.

[30] J.-F. Coeurjolly and E. Rubak. Fast covariance estimation for innovations computed from a spatial Gibbs point process. *Scandinavian Journal of Statistics*, 40(4) :669–684, 2013.

[31] J.F. Coeurjolly and N. Morsli. Poisson intensity parameter estimation for stationary Gibbs point processes of finite interaction range. *Spatial Statistics*, 4 :45–56, 2013.

[32] G. Collomb. Proprietes de convergence presque complete du predicteur à noyau. *Zeitschift für W. V. G*, 66 :441–60, 1984.

[33] F. Comets and M. Janzura. A central limit theorem for conditionally centred random fields with an application to Markov fields. *Journal of Applied Probability*, 35(3) :608–621, 1998.

[34] D. Daley and D. Vere-Jones. *An introduction to the Theory of Point Processes*. Springer Verlag, New York, 1988.

[35] S.V. David. Central limit theorems for empirical product densities of stationary point processes. *Thèse de Doctorat d'Etat, Université Augsburg*, 2008.

[36] J. Dedecker. Central limit theorem for stationary random fields. *Probability Theory and Related Fields*, 110 :397–426, 1998.

[37] J. Dedecker. Exponential inequalities and functional central limit theorems for random fields. *ESAIM Probability and Statistics*, 5 :77–104, 2001.

[38] P. Deheuvels. Estimation non paramétrique de la densité par histogrammes généralisé. *Statistic. Appl*, 25(3) :5–42, 1977.

[39] M. Delecroix. *Thèse de Doctorat d'Etat, Université de Lille 1*, 1987.

[40] D. Dereudre. Gibbs Delaunay tessellations with geometric hardcore condition. *J. Stat. Phys.*, 131(1) :127–152, 2008.

[41] D. Dereudre, R. Drouilhet, and H.O. Georgii. Existence of Gibbsian point processes with geometry-dependent interactions. *Probability Theory and Related Fields*, 153 :647–670, 2012.

[42] D. Dereudre, K. Helisova, and F. Lavancier. Fitting all parameters of the quermass model by the takacs-fiksel method. *Scandinavian Journal of Statistics*, 2013.

[43] D. Dereudre and F. Lavancier. Campbell equilibrium equation and pseudo-likelihood estimation for non-hereditary Gibbs point processes. *Bernoulli*, 15(4) :1368–1396, 2009.

[44] P.J. Diggle. *Statistical analysis of spatial point patterns*. Academic Press, London, 2003.

[45] P.J. Diggle, D.J. Gates, and A. Stibbard. A nonparametric estimator for pairwise-interaction point processes. *Biometrika*, 74(4) :763–770, 1987.

[46] P.J. Diggle and R.J. Gratton. Monte Carlo methods of inference for implicit statistical models. *Journal of the Royal Statistical Society*, B 46 :193–227, 1984.

[47] R.L. Dobrushin. The description of a random field by means of conditional probabilities and condition of its regularity. *Theory of Probability and its Applications*, 13 :197–224, 1968.

[48] P. Doukhan and S. Louhichi. A new weak dependence condition and application to moment inequalities. *Stochastic Processes and their Applications*, 84 :313–342, 1999.

[49] R.l. Doukhan. *Mixing : Properties and Axemples, vol 85*. Lecture Notes in Statistics. Berlin, 1994.

[50] C. Gaetan et X. Guyon. *Modélisation et statistique spatiales*. Springer, 2000.

[51] T. Fiksel. Estimation of parameterized pair potentials of marked and non-marked Gibbsian point processes. *Elektronische Informationsverarbeitung und Kybernetik*, 20 :270–278, 1984.

[52] T. Fiksel. Edge-corrected density estimators for point processes. *Statistics*, 19 :67–75, 1988.

[53] T. Fiksel. Estimation of interaction potentials of Gibbsian point processes. *Statistics*, 19 :77–86, 1988.

[54] S. Fucik, O.John, and A. Kufner. Function Spaces. *Czechoslovak Mathematical Journal*, 29 :325, 1979.

[55] D.J. Gates and M. Westcott. Clustering estimes for spatial point distribution with unstable potentials. *Annals of the Institute of Statistical Mathematics*, 38 :123–135, 1986.

[56] S. Gema and D. Gema. Stochastic relaxation, Gibbs Distributions and the Bayesian Restoration of Images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6 :721–741, 1984.

[57] H.O. Georgii. Canonical and grand canonical Gibbs states for continuum systems. *Communications in Mathematical Physics*, 48(1) :31–51, 1976.

[58] H.O. Georgii. *Gibbs measures and Phase Transitions*. Walter de Gruyter, Berlin, 1988.

[59] C.J. Geyer. Likelihood inference for spatial point processes. In W.S. Kendall O.E. Barndorff-Nielsen and M.N.M. Van Lieshout (eds), editors, *Stochastic Geometry : Likelihood and Computation*, number 80, chapter 3, pages 79–140. 1999.

[60] L. Giraitis, H. L. Koul, and D. Surgailis. Asymptotic normality of regression estimators with long memory errors. *Statist. Probab. Lett*, 29(4) :317–335, 1996.

[61] X. Guyon. Méthode de pseudo-vraissemblance et codage pour les processus ponctuels de Gibbs. *Prépublication SAMOS no 6. Univ. Paris 1, Panthéon-Sorbonne, 19 p*, 1991.

[62] X. Guyon. Random Fields on a Network : Modeling, Statistics and Applications. *Springer, New York*, 1995.

[63] L. Györfi, W. Härdle, P. Sarda, and P. Vieu. Nonparametric curve estimation from time series. *Lecture Notes in Statistics 60 Springer-Verlag, Berlin*, page 153, 1989.

[64] P. Hall and J. Hart. Bootstrap test for dierene between means in nonparametric regression. *Journal of the American Statistical Association*, 85 :1039–1049, 1990.

[65] P. Hall and C.C. Heyde. Martingale limit theory and its application. *Academic Press, New York-London*, 285-297, 1980.

[66] M. Hallin, Z. Lu, and L.T. Tran. Density estimation for spatial linear processes. *Bernoulli*, 7(4) :657–668, 2001.

[67] M. Hallin, Z. Lu, and L.T. Tran. Local linear spatial regression. *Annals of Statistics*, 32(6) :2469–2500, 2004.

[68] K.H. Hanisch and D. Stoyan. Remarks on statistical inference and prediction for hard-core clustering model. *Math. Oper. statistics. Ser. Statistics*, 14 :559–567, 1986.

[69] L. Heinrich. Minimum contrast estimates for parameters of ergodic spatial point processes. *Mathematics Biometrika*, 57 :97–109, 1970.

[70] L. Heinrich. Asymptotic behaviour of an empirical nearest-neighbour distance function for stationary Poisson cluster processes. *Mathematische Nachrichten*, 136 :131–148, 1988.

[71] L. Heinrich. Asymptotic Gaussianity of some estimators for reduced factorial moment measures and product densities of stationary Poisson cluster processess. *Statistics*, 19 :87–106, 1988.

[72] L. Heinrich and E. Liebscher. Strong convergence of kernel estimators for product densities of absolutely regular point processes. *S. J. Nonparametric Statistics*, 1 :65–96, 1997.

[73] L. Horváth. On l_p-norms of multivariate density estimators. *Revue de Statistique Appliquée*, 19(4) :1933–1949, 1991.

[74] Y.A. Rozanov I.A. Ibragimov. *Processus aléatoires gaussiens.* Springer-Verlag, 1974.

[75] I.A.Ibragimov. Some limit theorems for stationary processes. *Theory of Probability and Applications*, 7(4) :349–382, 1962.

[76] I.A. Ibragimov and Yu.V. Linnik. Independent and Stationary Sequences of Random Variables. *Wolters-Nocrdhoff, Grcningen*, 1971.

[77] J.L. Jensen and H.R. Künsch. On asymptotic normality of pseudo likelihood estimates of pairwise interaction processes. *Annals of the Institute of Statistical Mathematics*, 46(3) :475–486, 1994.

[78] J.L. Jensen and J. Møller. Pseudolikelihood for exponential family models of spatial point processes. *Advances in Applied Probability*, 1 :445–461, 1991.

[79] B.W. Silverman J.O. Ramsay. *Functional data analysis.* 1997.

[80] E. Jolivet. Upper bounds of the speed of convergence of moment density estimators or stationary point processes. *Metrika*, 31 :349–360, 1984.

[81] F. Kelly and B. Ripley. A note on Strauss's model for clustering. *Biometrika*, 63 :357–360, 1976.

[82] F. Kelly and B. Ripley. Statistical Inference for Spatial Processes. *Cambridge University Press, Cambridge*, 1988.

[83] W.S. Kendell and J. Møller. Perfect simulation using dominating processes on ordered state spaces, with application to locally sable point processes. *Advances in Applied Probability*, 32 :844–865, 2000.

[84] M. A. Krasnosel'skii and Y. B. Rutickii. *Convex Functions and Orlicz Spaces, P.* Noordhoff LTD-Groningen-The Netherlands, 1961.

[85] F. Lavancier, J. Møller, and E. Rubak. Determinantal point process models and statistical inference. *Mathematics. Statistics Theory*, 2013.

[86] A.B. Lawson. *Statistical Methods in Spatial Epidemiology.* Wiley, New York, 2001.

[87] Van Lieshout. *Markov Point Processes and their Applications.* Imperial College Press, London, 2000.

[88] M. El Machkouri. Kahane-khintchine inequalities and functional central limit theorem for stationary real random fields. *Stochastic Processes and Their Applications*, 120 :285–299, 2002.

[89] M. El Machkouri. Kernel density estimation for stationary random fields. *ALEA, Lat. Am. J. Probab. Math. Stat*, 11(1) :259–279, 2014.

[90] M. El Machkouri, D.Volný, and W. B. Wu. A central limit theorem for stationary random fields. *Stochastic Processes and Their Applications*, 123 :1–14, 2013.

[91] S. Mase. Uniform LAN condition of planar Gibbsian point processes and optimality of maximum likelihood estimators of soft-core potential function. *Probability Theory and Related Fields*, 92 :51–67, 1992.

[92] B. Matérn. *Spatial Variation : Stochastic Models and their Applications to Some Problems in Forest Surveys and Other Sampling Investigations*. 2nd. Springer, Heidelberg, 1986.

[93] D.L. McLeish. A maximal inequality and dependent strong laws. *Annals of Probability*, 3(5) :829–839, 1975.

[94] J. Møller, A.R. Syversveen, and R.P. Waagepetersen. Log Gaussian Cox processes. *Scandinavian Journal of Statistics*, 25 :451–482, 1998.

[95] J. Møller and R. Waagepetersen. *Statistical Inference and Simulation for Spatial Point Processes*. Chapman and Hall/CRC, Boca Raton, 2004.

[96] J. Møller and R.P. Waagepetersen. Modern statistics for spatial point processes. *Scandinavian Journal of Statistics*, 34(4) :643–684, 2007.

[97] V.M.R. Muggeo. Segmented : an R package to fit regression models with broken-line relationships. *R news*, 8 :20–25, 2008.

[98] E. Nadaraja. On nonparametric estimates of density functions and regression curves. *Theory of Probability and Applications*, 10 :186–190, 1965.

[99] E.A. Nadaraja. Nonparametric estimation of probability densities and regression curves. *Kluwer Academic Publishers. Dordrecht (Translation from the Russian)*, 1989.

[100] J. Neyman and E.L. Scott. Statistical Approach to Problems of Cosmology. *Journal of the Royal Statistical Society, series B*, 20 :1–43, 1958.

[101] X.X. Nguyen and H. Zessin. Ergodic theorems for Spatial Process. *Z. Wahrscheinlichkeitstheorie verw. Gebiete*, 48 :133–158, 1979.

[102] X.X. Nguyen and H. Zessin. Integral and differential characterizations Gibbs processes. *Mathematische Nachrichten*, 88 :105–115, 1979.

[103] Y. Ogata and M. Tanemura. Estimation of interaction potentials of spatial point patterns through the maximum likelihood procedure. *Annals of the Institute of Statistical Mathematics*, 33 :315–318, 1981.

[104] Y. Ogata and M. Tanemura. Likelihood analysis of spatial point patterns. *Journal of the Royal Statistical Society, series B*, 46(3) :496–518, 1984.

[105] F. Papangelou. The conditional intensity of general point processes and application to line processes . *Zeitschrift für Wahscheinlichkeitstheorieund verwandte Gebiete*, 28 :207–227, 1974.

[106] F. Papangelou. The Armenian Connection : Reminiscences from a time of interactions . *Contemporary mathematical analysis*, 44(1) :14–19, 2009.

[107] M. Pelidrad. Properties of uniform consistency of the kernel estimators of density and of regression functions under dependence assumptions. *Reidel Publishing Company*, 33-43, 1987.

[108] A. Penttien. *Modelling interaction in spatial point partterns : parameter estimation by maximum likelihood method. Jyväskylä Studies in Computer Science.* Economics and Statistics 7, 1984.

[109] C.J. Preston. *Random fields.* Springer Verlag, 1976.

[110] E. Rio. Covariance for strongly mixing processes. *Annales l'IHP*, 29(4) :587–597, 1993.

[111] B. Ripley. Edge effects in spatial stochastic processes. *Ranneby B ed. In Statistics in Theory and Practice,Umea : Swedish University of Agricultural Sciences*, 242-62, 1982.

[112] B. Ripley and F. Kelly. Markov point processes. *Journal of the London Mathematical Society*, 15 :188–192, 1977.

[113] P. M. Robinson. Nonparametric estimators for time series. *Spectral Analysis Time Series*, 4(3) :185–207, 1983.

[114] P.M. Robinson. Large-sample inference for nonparametric regression with dependent errors. *Annals of Statistics*, 28 :2054–2083, 1997.

[115] M. Rosenblatt. A central limit theorem and a strong mixing condition. *Nat. Acad. Sci. USA*, 42(1) :43–47, 1956.

[116] W. Rudin. *Real and Complex Analysis.* McGraw Hill, 1974.

[117] D. Ruelle. *Statistical Mechanics.* Benjamin, New York-Amsterdam, 1969.

[118] R.J. Serfling. Contributions to central limit theory for dependent variables. *Annals of Statistics Mathematics*, 39(4) :1158–1175, 1968.

[119] J. Serra. *Image Analysis and Mathematical Morphology.* Academic Press, New York, 1982.

[120] B. W. Silverman. Density estimation for statistics and data analysis,. *Chapman and Hall, London*, 1986.

[121] D. Stoyan, W.S. Kendall, and J. Mecke. *Stochastic geometry and its applications.* John Wiley and Sons, Chichester, 1987.

[122] D. Stoyan and H. Stoyan. Estimating pair correlation functions of planar cluster processes. *Biometrical J*, 38(3) :259–271, 1996.

[123] D.J. Strauss. A model for clustering. *Biometrika*, 62 :467–475, 1975.

[124] R. Takacs. Estimator for the pair-potential of a Gibbsian point process. *Institutsbericht 238, Institut fur Mathematik, Johannes Kepler Universitat Linz, Austria*, 1983.

[125] R. Takacs. Estimator for the pair-potential of a Gibbsian point process. *Statistics*, 17 :429–433, 1986.

[126] H. Takahata. On the central limit theorem for weakly dependent random fields. *Yokohama Math.J*, 31 :67–77, 1983.

[127] D. Tjøstheim. Statistical spatial series modelling. *Advances in Applied Probability*, 10 :130–154, 1978.

[128] D. Tjøstheim. Statistical spatial series modelling ii : Further results on unilateral lattice processes. *Advances in Applied Probability*, 15 :562–584, 1983.

[129] L.T. Tran and S.Yakowitz. Nearest neighbor estimators for random fields. *Journal of Multivariate Analysis*, 44 :23–46, 1993.

[130] J.W. Tukey. *Spectral Analysis Time Series*. Wiley, New York, 1976.

[131] Y. Wang and M. Woodroofe. A new condition on invariance principles for stationary random fields. *Statistica Sinica*, 23 :1673–1696, 2013.

[132] Y. Wang and M. Woodroofe. On the asymptotic normality of kernel density estimators for linear random fields. *Journal of Multivariate Analysis*, 123 :201–213, 2014.

[133] W. B. Wu. Nonlinear system theory : another look at dependence. *Proc. Nat. Acad. Sci. USA*, 102(40) :14150–14154, 2005.

[134] W. B. Wu and J. Mielniczuk. Kernel density estimation for linear processes. *Annals of Statistics*, 30(5) :1441–1459, 2002.

[135] H. Zessin. Der papangelou prozess. *Contemporary mathematical analysis*, 44 :36–44, 2009.